

ІНСТИТУТ ДЕМОГРАФІЇ ТА СОЦІАЛЬНИХ ДОСЛІДЖЕНЬ ІМЕНІ М. В. ПТУХИ
НАЦІОНАЛЬНОЇ АКАДЕМІЇ НАУК УКРАЇНИ

ОЦІНЮВАННЯ СОЦІАЛЬНО- ЕКОНОМІЧНИХ ПОКАЗНИКІВ: ПРИКЛАДНІ АСПЕКТИ ЗАСТОСУВАННЯ НЕПРЯМИХ МЕТОДІВ

Монографія

Київ – 2012

УДК 330.101.52
ББК 65.051
С 20

*Рекомендовано до друку Вченою радою Інституту демографії та соціальних досліджень імені М. В. Птухи НАН України
(протокол № 17 від 10 грудня 2012 року)*

Рецензенти:

Єріна А. М., д.е.н., професор кафедри статистики, ДВНЗ «Київський національний економічний університет імені Вадима Гетьмана».

Пуначова М. В., д.е.н., Генеральний директор Науково-технічного комплексу статистичних досліджень Державної служби статистики України.

Серіогло В. Г.

С 20 Оцінювання соціально-економічних показників: прикладні аспекти застосування непрямих методів : [монографія] / Серіогло В. Г. – К. : Інститут демографії та соціальних досліджень імені М. В. Птухи НАН України, 2012. – 136 с.

ISBN 978-966-304-020-2

У монографії розглянуто особливості практичного застосування сучасних методологічних підходів до підвищення надійності оцінювання соціально-економічних показників за результатами вибірових обстежень домогосподарств (населення). На прикладі розв'язання проблеми недостатньої надійності показників бідності для регіонів України показано, що ефективним є статистичне об'єднання прямих та синтетичних оцінок цих показників на основі багатофакторної композиційної оціночної функції. При цьому можливість підвищення надійності показників обумовлена використанням інформації з додаткових, зовнішніх щодо вибірового обстеження, джерел. Наведено результати аналізу джерел зовнішньої інформації, яка потенційно може бути використана при оцінюванні бідності. Показано, що при розробці дизайну вибірки, зокрема при його оптимізації, доцільно враховувати застосування ускладнених методів оцінювання показників за результатами обстеження. Певну увагу приділено інструментарію автоматизації процедур оцінювання показників.

Матеріали монографії призначені для фахівців органів державного управління, насамперед органів державної статистики, науковців, які займаються питаннями підвищення якості інформаційного забезпечення соціальної та соціально-економічної політики.

**УДК 330.101.52
ББК 65.051**

ISBN 978-966-304-020-2

© Серіогло В. Г., 2012

ЗМІСТ

Передмова.....	5
Перелік скорочень.....	11
Визначення основних термінів.....	11
РОЗДІЛ 1. Методологічні підходи до оцінювання показників бідності на регіональному рівні на основі композиційних оціночних функцій.....	17
1.1. Проблеми забезпечення надійності оцінювання показників бідності.....	17
1.2. Загальна характеристика методології непрямого оцінювання показників бідності	32
1.3. Визначення надійності композиційних оцінок показників бідності.....	43
РОЗДІЛ 2. Особливості використання зовнішньої інформації при непрямому оцінюванні показників бідності.....	47
2.1. Джерела зовнішньої інформації для побудови синтетичних оціночних функцій.....	47
2.2. Визначення зовнішньої інформації для композиційного оцінювання показників бідності.....	62
2.3. Побудова синтетичних оцінок показників бідності на основі апроксимації розподілу доходів на мікрорівні.....	68
РОЗДІЛ 3. Прикладні аспекти побудови статистико-математичних моделей для непрямого оцінювання показників бідності.....	74
3.1. Аналіз результатів використання багатofакторних композиційних оціночних функцій для підвищення надійності оцінювання визначених показників бідності.....	74
3.2. Супровід використання композиційних оціночних функцій.....	80
3.3. Методологічні підходи до створення форматів подання користувачам результатів застосування методів непрямого оцінювання показників бідності.....	83

3.4. Програмне забезпечення для оцінювання визначених показників бідності на регіональному рівні з використанням композиційних оціночних функцій.....	87
РОЗДІЛ 4. Вплив методів непрямого оцінювання показників на результати вибірових обстежень.....	91
4.1. Узгодження дизайну вибірки та процедур непрямого оцінювання показників.....	91
4.2. Калібрація системи статистичних ваг для врахування результатів уточнення оцінок показників на мікрорівні.....	98
Заключні положення.....	103
Список використаних джерел.....	105
Додатки.....	109

ПЕРЕДМОВА

У теперішній час в Україні, як і в багатьох країнах світу, все більш актуальними стають питання підвищення ефективності реалізації соціальної, демографічної та соціально-економічної політики, насамперед програм соціально-економічного розвитку на національному, регіональному й місцевому рівнях. Значна увага приділяється оцінці фактичного впливу реалізованих та потенційного – запланованих заходів політики на якість життя населення. За таких умов суттєво зростають вимоги до інформаційного забезпечення діяльності органів державного управління різних рівнів, громадських організацій, наукових установ тощо. Все більш нагальною стає проблема недостатньої статистичної надійності визначених за результатами вибірових обстежень найважливіших соціально-економічних показників, таких як показники бідності, економічної активності населення, умов життя тощо, на регіональному та субрегіональному рівнях.

За наявності розвиненої системи державної (офіційної) статистики, сформованої відповідно до міжнародних стандартів, широкого кола актуальних досліджень, що здійснюються недержавними організаціями, а також сучасної технічної бази перспективним напрямом розв'язання проблем якості інформаційного забезпечення соціальної та економічної політики в Україні стає розробка та застосування підходів, які уможливають оцінювання найважливіших показників на основі комплексного використання інформації з різних джерел. Доцільно

зазначити, що застосування таких підходів передбачає суттєві зміни у механізмах отримання, надання користувачам різного рівня та подальшого використання відповідної інформації.

У запропонованій монографії розглянуто зміст та основні методологічні й методичні аспекти підходів до підвищення надійності оцінювання соціально-економічних показників з використанням даних вибірових обстежень і наявної додаткової (зовнішньої) інформації на прикладі показників бідності. Вибір саме цієї групи показників зумовлений тим, що реалізація ефективної соціальної політики в Україні потребує проведення постійного моніторингу бідності на загальнодержавному та регіональному рівнях, а це вимагає надійного вимірювання основних показників бідності. Достатній рівень їх надійності забезпечує адекватність аналізу бідності, у тому числі оцінки впливу на бідність заходів соціальної політики, та є базою для прийняття дієвих управлінських рішень у цій сфері.

В Україні показники бідності вимірюються за результатами державного вибіркового обстеження умов життя домогосподарств (далі – ОУЖД), методологія проведення якого повністю відповідає міжнародним стандартам для обстежень такого типу.

Слід зазначити, що проблема недостатньої надійності оцінювання статистичних показників є характерною для багатьох країн світу й особливо актуальною, якщо інформаційною базою для вимірювання показників слугують дані вибірових обстежень. У світовій практиці широко відомі дослідження, присвячені питанням забезпечення статистичної надійності показників бідності для різних рівнів агрегації даних. Водночас наразі проблема забезпечення достатньої статистичної надійності показників бідності залишається

невирішеною, що підсилює важливість наукових розвідок у цьому напрямі.

Як свідчать результати оцінки рівня надійності показників бідності, визначених як найбільш важливі для моніторингу бідності в Україні, пряме їх оцінювання за результатами обстеження умов життя домогосподарств не забезпечує їх достатньої надійності на рівні регіонів. Це призводить до нестабільності оцінок у динаміці та обмежує можливість ефективного моніторингу бідності.

У цілому проблема забезпечення надійності оцінок показників бідності в Україні може вирішуватись двома шляхами:

- підвищення статистичної ефективності дизайну вибірки, зокрема на основі збільшення обсягу вибіркової сукупності домогосподарств в обстеженні умов життя домогосподарств для регіонів України та / або оптимізації розподілу вибірки за стратами, кількості та розміру територіальних одиниць вибірки тощо;
- застосування спеціальних методологічних підходів до оцінювання показників бідності, у тому числі з урахуванням наявної додаткової (зовнішньої) інформації з інших джерел.

На практиці зазначені підходи, як правило, застосовуються одночасно, оскільки вони у загальному випадку є взаємозалежними. Суттєвим обмеженням досліджень варіантів дизайну вибірки є неможливість збільшення загального обсягу вибірки ОУЖД. Це пояснюється, головним чином, тим, що при побудові дизайну вибірки ОУЖД значною мірою враховані потреби мінімізації витрат на обстеження при дуже великому обсязі інформації, яка збирається. У зв'язку з обмеженістю ресурсів побудувати абсолютно оптимальну вибірку, яка б забезпечила бажану надійність оцінок різноманітних показників на регіональному рівні, як правило, неможливо. Таким чином, у

разі значної кількості регіонів, за якими надійність оцінювання певних показників є недостатньою, можливість застосування методів оптимізації дизайну вибірки є обмеженою.

Водночас досвід інших країн свідчить, що за рахунок удосконалення дизайну вибірки проблему недостатньої надійності оцінок окремих показників бідності за регіонами можна дещо пом'якшити. Певні кроки в цьому напрямі були зроблені Держстатом України при вдосконаленні дизайну вибірки ОУЖД для проведення обстежень у 2009–2013 роках. Однак розробка та впровадження ефективних підходів до оптимізації дизайну вибірки все ще залишається актуальним завданням. Крім того, слід урахувати, що зменшення варіативності рівня надійності оцінок показників бідності за регіонами може забезпечити більш адекватну інформаційну базу для застосування з метою забезпечення надійності й інших статистичних методів.

На основі результатів опрацювання існуючих підходів до підвищення рівня надійності результатів вибірових обстежень домогосподарств можна зробити висновок, що в теперішній час найбільш ефективною є методологія уточнення оцінок показників на базі так званих методів оцінювання для малих територій. Специфічна особливість цих методів полягає в комплексному використанні даних вибірового обстеження та релевантної інформації з додаткових (зовнішніх) джерел.

У першому розділі монографії розглядаються особливості визначення надійності статистичного оцінювання показників за результатами вибірових обстежень, зміст основних показників бідності, за якими здійснюється її моніторинг в Україні. Наводиться характеристика методу композиційного оцінювання – ефективного методологічного підходу до підви-

щення надійності оцінювання статистичних показників на основі використання додаткової інформації.

Другий розділ присвячено питанням визначення потенційних джерел додаткових даних, що можуть бути використані при побудові композиційних оціночних функцій для оцінювання показників бідності в Україні, та аналізу придатності інформації з цих джерел.

Прикладні аспекти застосування багатофакторних композиційних оціночних функцій для підвищення надійності показників бідності розглядаються у третьому розділі. Певна увага приділяється питанням методичного супроводу функціонування інструментарію композиційного оцінювання показників бідності та принципам надання інформації за результатами його застосування внутрішнім та зовнішнім користувачам. Подається загальна характеристика програмного забезпечення, розробленого для автоматизації розрахунків композиційних оцінок показників бідності у середовищі *SPSS (PASW Statistics)* з використанням пакета *R*.

У четвертому розділі висвітлено прикладні аспекти реалізації методів оптимізації дизайну вибірки з огляду на застосування композиційних оціночних функцій на етапі оцінювання показників. Також наводиться загальна характеристика підходу до урахування уточнених оцінок показників для калібрації систем статистичних ваг у масиві мікроданих вибіркового обстеження, що потенційно може забезпечити поліпшення надійності оцінювання й інших показників за результатами обстеження.

Доцільно зазначити, що методологічною основою досліджень, результати яких представлені у монографії, є, головним чином, розробки Н. Т. Лонгфорда, який працював як міжнародний консультант над питаннями створення підходів до

підвищення надійності оцінювання показників бідності в Україні упродовж коротких проміжків часу у 2009–2011 роках при реалізації проекту Світового банку "Удосконалення системи соціальної допомоги в Україні". Результати досліджень доведені до рівня, який дозволив розпочати їх упровадження в системі державної статистики України. Слід зазначити, що це було б неможливо без вимогливості та безпосередньої участі в апробації підходів і розробці методичних матеріалів начальника департаменту обстежень домогосподарств Держстату України І. І. Осипової та без підтримки досліджень з боку заступника Голови Держстату Н. С. Власенко. Також варто відмітити суттєвий вклад у проведення основних етапів досліджень та апробації їх результатів співробітників відділу соціально-демографічної статистики Інституту демографії та соціальних досліджень імені М. В. Птухи НАН України та фахівців департаменту обстежень домогосподарств Держстату.

Монографію підготовлено в рамках дослідження «Моделювання поточних та перспективних соціальних результатів реалізації бюджетних програм».

ПЕРЕЛІК СКОРОЧЕНЬ

- Держстат – Державна служба статистики України.
ДПА – Державна податкова адміністрація.
ДЦЗ – Державний центр зайнятості.
ОЕАН – обстеження населення (домогосподарств) з питань економічної активності.
ОУЖД – обстеження умов життя домогосподарств.

ВИЗНАЧЕННЯ ОСНОВНИХ ТЕРМІНІВ

(за матеріалами [1; 2])

- Гранична похибка вибірки* – гранично припустима розбіжність між оцінкою показника за результатами вибіркового обстеження та дійсним значенням показника (невідомим) для генеральної сукупності, визначена на основі стандартної похибки вибірки та встановленого рівня довірчої імовірності.
- Дизайн (план) вибірки* – організаційно-логічна модель структури вибіркової сукупності та принципи її формування. Якістю плану вибірки в значній мірі визначається репрезентативність вибіркової сукупності, отриманої за визначеними у плані правилами.
- Дизайн-ефект* – міра відносної статистичної ефективності певного плану вибірки порівняно з простою випадковою вибіркою (відношення дисперсії вибірових оцінок показника для плану вибірки, реалізованого в обстеженні, до дисперсії оцінки цього показника, отриманої за умови використання процедури простого випадкового відбору).

<i>Дисперсія вибіркових оцінок</i>	– середній квадрат відхилень оцінок показника за всіма можливими вибірками певного дизайну від середнього значення оцінки за всіма вибірками. – це інформація, що використовується при оцінюванні показників за результатами проведеного вибіркового обстеження, але отримана з іншого джерела даних (наприклад, це дані інших статистичних обстежень, дані цього ж обстеження, але по інших територіях або за інші періоди часу, дані статистичної звітності, адміністративні дані тощо).
<i>Додаткова (зовнішня) інформація</i>	
<i>Зміщення оцінки показника</i>	– різниця між очікуваною величиною оцінки та дійсним значенням показника для всієї сукупності. Особливість зміщення полягає в тому, що воно є постійною складовою частини похибки, яка при численній сукупності не зменшується зі збільшенням обсягу вибірки, тоді як випадкова похибка вибірки зі збільшенням обсягу вибірки в середньому зменшується. Визначають зміщення як різницю середнього значення оцінок у разі повторних спостережень і оцінюваної величини.
<i>Коваріата</i>	– змінна, яка статистично взаємопов'язана зі змінною, що оцінюється.
<i>Коефіцієнт варіації (відносна стандартна похибка)</i>	– відношення величини стандартної похибки вибірки до оцінки показника (частіше за все вимірюється у відсотках).

<i>Композиційна оціночна функція</i>	– оціночна функція, отримана на основі композиції (комбінації) різних оціночних функцій. У статистиці частіше за все використовуються композиція прямих та непрямих оціночних функцій.
<i>Критерій оптимальності</i>	– ознака, згідно якої визначається найкращий з можливих варіантів структури, стану об'єкта; в математичних моделях критерію оптимальності відповідає цільова функція, екстремальне значення якої і визначає оптимальну структуру, стан об'єкта моделювання.
<i>Межа бідності</i>	– рівень доходу, нижче від якого є неможливим задоволення основних потреб. Визначення критеріїв для здійснення оцінки бідності на національному та регіональному рівнях (національної межі бідності та регіональної межі бідності) регламентується чинними нормативними актами.
<i>Непряме оцінювання</i>	– оцінювання показника на основі непрямой оціночної функції, при побудові якої передбачається використання даних, отриманих не лише за результатами безпосереднього вимірювання (на основі конкретного обстеження), а і з використанням додаткової (зовнішньої) інформації (наприклад, даних аналогічних обстежень по інших територіях, за інші періоди часу, дані інших обстежень тощо).
<i>Обсяг вибірки</i>	– кількість одиниць у вибірковій сукупності. Необхідний для забезпечення репрезентативності та належної якості результатів

вибіркового обстеження обсяг вибірки розраховують залежно від способу відбору та типу вибірки. Величина обсягу вибірки залежить від варіації досліджуваних ознак та припустимої похибки вибірки.

*Оптимальний
план (дизайн)
вибірки*

– дизайн вибірки, який забезпечує найвищу статистичну ефективність (за певним критерієм оптимальності) оцінювання визначених показників за результатами обстеження.

Оцінка показника

– значення показника, яке розраховано на основі відповідної оціночної функції за результатами вибіркового обстеження.

Оціночна функція

– правило або співвідношення (формула) для оцінювання показника з використанням даних вибіркового обстеження. Оціночна функція відома ще до отримання результатів обстеження. Слід розрізняти поняття *оціночна функція* та *оцінка*, оскільки оціночна функція – це функція вибірки, тоді як оцінка є її значенням, знайденим на основі реалізації вибірки.

*Оцінювання для
малих територій*

– статистична методологія, яка розроблена з метою підвищення надійності оцінювання показників за результатами статистичних спостережень на основі прямих та непрямих оціночних функцій, комплексного використання інформації з різних джерел та сучасних методів статистико-математичного моделювання.

Похибка вибірки

– частина розбіжності між істинним значенням статистичного показника та його зна-

ченням, розрахованим за вибіркою, яка зумовлена застосуванням саме несущільного характеру обстеження.

Пряме оцінювання – оцінювання показника на основі прямої оціночної функції, при побудові якої передбачається, що оцінювання здійснюється на основі даних, отриманих за результатами безпосереднього вимірювання (на основі конкретного обстеження).

Пріоритет – показник, який використовується як параметр при розподілі вибірки за регіонами і відображає пріоритетність оцінювання показників для регіонального та національного рівнів у плані характеристик надійності оцінювання, що забезпечуються.

Прожитковий мінімум – вартісна величина достатнього для забезпечення нормального функціонування організму людини, збереження його здоров'я набору продуктів харчування, а також мінімального набору непродовольчих товарів та мінімального набору послуг, необхідних для задоволення основних соціальних і культурних потреб особистості.

Реплікація – незалежне повторення певного процесу (формування вибірки, оцінювання тощо).

Рівень бідності за національною межею – частка населення країни, рівень споживання (доходів) якого є нижчим від встановленої національної межі бідності.

Рівень бідності за регіональною межею – частка населення регіону, рівень споживання (доходів) якого є нижчим від встановленої регіональної межі бідності.

<i>Рівень бідності за прожитковим мінімумом</i>	– частка населення з загальними середньомісячними еквівалентними доходами, нижчими за прожитковий мінімум у середньому на одну особу.
<i>Середньоквадратична похибка</i>	– це загальна похибка оцінювання показника, що визначається як сума дисперсії оцінки показника та квадрату величини зміщення оцінки, яке є різницею між очікуваною величиною оцінки (наприклад, середнім із вибірових оцінок) та дійсним значенням показника для генеральної сукупності.
<i>Стандартна похибка</i>	– середньоквадратичне відхилення вибірових оцінок показника за всіма можливими вибірками певного дизайну від дійсного значення показника для генеральної сукупності.

РОЗДІЛ 1

МЕТОДОЛОГІЧНІ ПІДХОДИ ДО ОЦІНЮВАННЯ ПОКАЗНИКІВ БІДНОСТІ НА РЕГІОНАЛЬНОМУ РІВНІ НА ОСНОВІ КОМПОЗИЦІЙНИХ ОЦІНОЧНИХ ФУНКЦІЙ

1.1. Проблеми забезпечення надійності оцінювання показників бідності

Необхідність підвищення дієвості соціальної політики в Україні вимагає проведення ефективного моніторингу й аналізу бідності на загальнодержавному та регіональному рівнях на основі надійного й своєчасного вимірювання основних показників бідності. Достатній рівень надійності показників бідності забезпечує адекватність аналізу й оцінки бідності (під оцінкою бідності будемо розуміти встановлення основних факторів бідності та визначення їх впливу) і є, певною мірою, підґрунтям для прийняття дієвих управлінських рішень у різних сферах соціальної політики.

В Україні показники бідності вимірюються за результатами державного вибіркового обстеження умов життя домогосподарств (далі – ОУЖД), методологія проведення якого відповідає міжнародним стандартам для обстежень такого типу [3; 4]. Для ОУЖД характерними є: складний дизайн вибірки (стратифікація генеральної сукупності, застосування механізмів багатоступеневого відбору, у тому числі відбору з імовірностями, що залежать від розміру одиниць відбору, та ін.);

вимірювання значної кількості показників для різних типів одиниць спостереження (особи, домогосподарства), використання оцінок різних типів показників (оцінювання сумарних, середніх значень, часток) тощо. Це значною мірою ускладнює процедури аналізу надійності оцінювання показників за результатами ОУЖД, оскільки зумовлює необхідність використання доволі складних методів розрахунків відповідних похибок.

Зі статистичного погляду, оцінка певного показника за результатами вибіркового обстеження є надійною, якщо вибіркові оцінки, побудовані на основі всіх можливих вибірок однакового дизайну та обсягу, сконцентровані біля дійсного значення показника [5; 6]. Висока концентрація, тобто надійність, є бажаною тому, що в окремому процесі побудови вибірки та розрахунку значення показника існує лише певна імовірність наблизитися до дійсного його значення.

Близькість оцінки до реального значення показника, тобто її надійність, має два аспекти: 1) вибіркові оцінки мають малу дисперсію; 2) середнє значення оцінок показника близьке до дійсного значення, тобто зміщення оцінок відсутнє або має відносно незначну величину (рис. 1.1).

Реальна надійність будь-якого показника, оціненого за даними вибіркового обстеження, може бути визначена лише у винятковому випадку, оскільки на практиці процедура відбору реалізується лише один раз і дійсні значення показників невідомі. Таким чином, у більшості випадків надійність показників визначається лише з певним наближенням.

Однією з головних характеристик надійності оцінки показника, розрахованого на підставі даних вибіркового обстеження, є похибка вибірки, яка показує, наскільки оцінка може відрізнятися від дійсного значення при існуючих дизайні вибірки, обсязі вибіркової сукупності, дисперсії значень ознаки за одини-

цями вибіркової сукупності або наскільки оцінка показника, отримана за результатами вибіркового обстеження, може відрізнятися від значення показника, яке було б отримане за результатами ідеального (без помилок) суцільного обстеження одиниць генеральної сукупності.

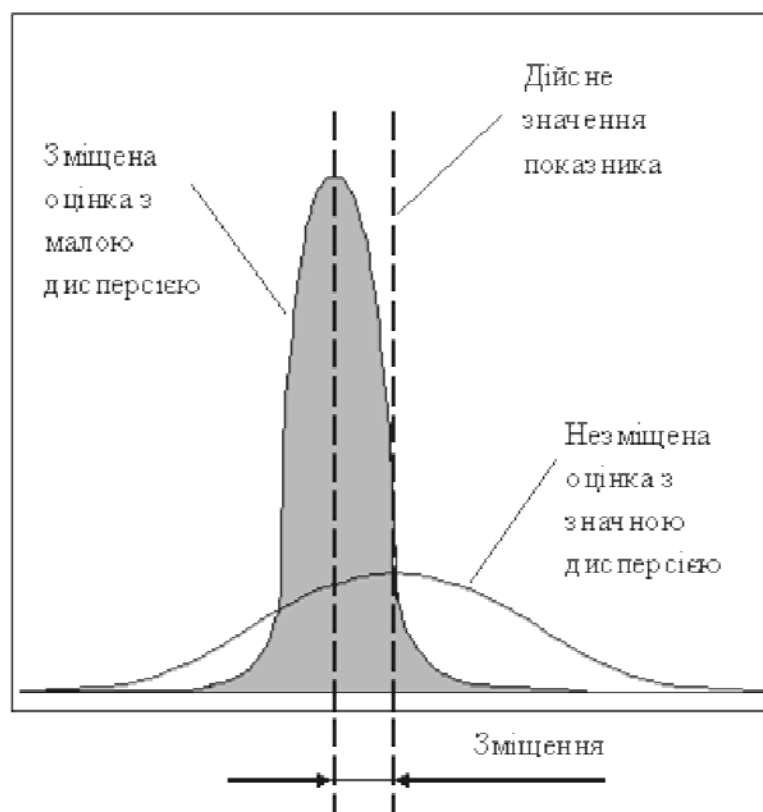


Рис. 1.1. Порівняння оцінок показника з великою дисперсією без зміщення та з малою дисперсією зі зміщенням

За сучасними міжнародними рекомендаціями, похибку вибірки, що може визначатись як абсолютна або відносна величина, необхідно публікувати для всіх основних показників, які вимірюються чи розраховуються за результатами обстеження. При аналізі похибки вибірки розраховують, як правило, стандартну похибку вибірки, граничну похибку вибірки та відносну стандартну похибку вибірки, яку частіше називають коефіцієнтом варіації.

Дисперсія вибірових оцінок $V(\hat{\theta})$ показника θ визначається за формулою [5]:

$$V(\hat{\theta}) = \frac{1}{s} \sum_{s=1}^S \left(\hat{\theta}_s - \bar{\hat{\theta}} \right)^2, \quad (1.1)$$

де $\hat{\theta}_s$ – оцінка показника по вибірці s , $s = 1, 2, \dots, S$; $\bar{\hat{\theta}}$ – середнє значення вибірових оцінок показника, $\bar{\hat{\theta}} = \frac{1}{s} \sum_{s=1}^S \hat{\theta}_s$; s – кількість усіх можливих вибірок певного дизайну та обсягу; θ – узагальнене позначення показника (тобто це може бути середнє або сумарне значення, частка, оцінка показника за регресією або за іншою статистичною моделлю, при цьому передбачається, що моделі побудовані за вибіровими даними).

Стандартна похибка вибірки $SE(\hat{\theta})$ при оцінюванні показника θ розраховується як корінь квадратний із дисперсії вибірових оцінок $V(\hat{\theta})$ [5; 6]:

$$SE(\hat{\theta}) = \sqrt{V(\hat{\theta})}. \quad (1.2)$$

На практиці розрахувати стандартну похибку з використанням формул (1.1) та (1.2) неможливо, оскільки у дослідника є лише одна вибірка з усіх можливих. Тому для визначення величини $SE(\hat{\theta})$ користуються спеціальними підходами.

Для вибірок із простим дизайном частіше за все застосовують зв'язок між величинами $V(\hat{\theta})$ і σ^2 – дисперсією значень ознаки, за якою оцінюється показник θ , по одиницях вибірки. Наприклад, для безповторної простої випадкової вибірки з малою часткою відбору f (зазвичай при $f < 0,05$) цей зв'язок має вигляд [6]:

$$V(\hat{\theta}) = \frac{\sigma^2}{n}. \quad (1.3)$$

де n – обсяг вибірки, $\hat{\theta}$ – оцінка показника θ за результатами вибірового обстеження.

Значення σ^2 обчислюється за формулою:

$$\sigma^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n - 1}, \quad (1.4)$$

де y – ознака (характеристика), за якою оцінюється показник θ ; y_i – значення ознаки для i -го домогосподарства (наприклад, якщо θ – грошовий дохід, то y_i – грошовий дохід i -го домогосподарства з вибірки); $\bar{y}$ – середнє значення ознаки по всіх одиницях вибірки.

Для вибірок зі складним дизайном величина $V(\hat{\theta})$ може бути розрахована на основі формули, аналогічної (1.3), але з урахуванням додаткового параметра $deff(\hat{\theta})$ [5]:

$$V(\hat{\theta}) = deff(\hat{\theta}) \cdot \frac{\sigma^2}{n}, \quad (1.5)$$

де $deff(\hat{\theta})$ – параметр, що відображає вплив дизайну вибірки на величину дисперсії вибіркових оцінок показника (дизайн-ефект). При використанні для побудови вибірки механізму простого випадкового відбору $deff(\hat{\theta}) = 1$.

Визначення величини $deff(\hat{\theta})$ для конкретного показника та конкретного дизайну вибірки потребує застосування спеціальних методів. Доцільно зазначити, що величина дизайн-ефекту при оцінюванні, наприклад, показників витрат та доходів на національному рівні за результатами ОУЖД може складати від 2,2 до 2,5.

Гранична похибка вибірки $ME(\hat{\theta})$ є максимально можливою похибкою вибірки для певної довірчої ймовірності p і визначається за формулою [6]:

$$ME(\hat{\theta}) = t \cdot SE(\hat{\theta}), \quad (1.6)$$

де t – довірче число (квантіль нормального розподілу або квантіль розподілу ймовірностей), яке визначає співвідношення граничної та стандартної похибки при певній імовірності p (p – імовірність того, що похибка вибірки для оцінки показника не перевищить величину граничної похибки $ME(\hat{\theta})$). Типові значення t наведені у табл. 1.1 (за даними [6]).

Гранична похибка вибірки використовується для побудови довірчих границь оцінок (границь довірчих інтервалів).

Таблиця 1.1

Взаємовідповідність між величинами довірчої імовірності p і довірчого числа t

p	0,50	0,80	0,90	0,95	0,99
t	0,67	1,28	1,64	1,96	2,58

Відносна стандартна похибка $RSE(\hat{\theta})$, або коефіцієнт варіації $CV(\hat{\theta})$, розраховується за формулою [5]:

$$RSE(\hat{\theta}) = CV(\hat{\theta}) = \frac{SE(\hat{\theta})}{\hat{\theta}} \cdot 100\%. \quad (1.7)$$

Використання відносних стандартних похибок при аналізі даних обстежень є особливо доцільним, коли ознаки, що розглядаються, є додатними величинами. У цьому випадку при порівнянні ступеня варіювання (і, як наслідок, точності оцінювання) різних показників забезпечується незалежність характеристик надійності від рівня показників та їх змісту.

У випадку отримання зміщених оцінок показників (наприклад, через наявність невибіркових систематичних похибок оцінювання або застосування зміщених оціночних функцій) при визначенні надійності замість величини дисперсії вибіркових

оцінок (1.1) використовують величину середньоквадратичної похибки оцінювання $MSE(\hat{\theta}; \theta)$ [5]:

$$MSE(\hat{\theta}; \theta) = V(\hat{\theta}) + B^2(\hat{\theta}; \theta), \quad (1.8)$$

де $B^2(\hat{\theta}; \theta)$ – зміщення оцінки показника $\hat{\theta}$ щодо дійсного значення цього показника θ ($B = \hat{\theta} - \theta$).

Усі інші характеристики надійності, зокрема граничні та відносні похибки, визначаються на основі величини $MSE(\hat{\theta}; \theta)$ за формулами, аналогічними (1.6), (1.7).

Слід зауважити, що визначення характеристик надійності, за результатами аналізу яких може бути зроблений висновок про достатній або недостатній рівень надійності показників бідності, у загальному випадку не є тривіальним завданням. Суть проблеми полягає у тому, що, як витікає з вищеведеного, при визначенні надійності оцінок показників можуть бути використані декілька характеристик (стандартна похибка, гранична похибка вибірки та коефіцієнт варіації). І хоча зазначені характеристики взаємопов'язані, вибір однієї з них як базової суттєво впливає на стратегію забезпечення надійності оцінювання показників. Це зумовлено тим, що, наприклад, стандартна похибка визначається лише дисперсією вибірових оцінок показника, а коефіцієнт варіації, який є відносною величиною, – стандартною похибкою, віднесеною до власне оцінки показника. Тобто за однакового рівня стандартної похибки нижчі значення показників бідності характеризуватимуться вищими значеннями коефіцієнтів варіації. Таким чином, за умови використання коефіцієнтів варіації виникає ситуація, коли при розробці заходів з підвищення надійності оцінювання показників на регіональному рівні більше уваги приділяється (і, відповідно, більше ресурсів виділяється) тим регіонам, де рівні бідності є суттєво нижчими.

Багатьма міжнародними статистичними організаціями розробляються спеціальні документи, що регламентують вимоги до надійності результатів вибірових обстежень. За сучасними рекомендаціями, для державних вибірових обстежень населення (домогосподарств) надійність оцінок показників для країни в цілому, яка характеризується коефіцієнтом варіації CV , повинна забезпечуватися на рівні $CV \leq 5\div 8\%$. Водночас для окремих територій (регіонів, районів) вимоги щодо надійності оцінок показників дещо послаблюються. Результати вважаються достатньо надійними, якщо коефіцієнт варіації не перевищує 10%, а для деяких обстежень – і 15%. Така надійність вважається допустимою, зокрема, і для загальнодержавних обстежень у США при оцінюванні показників для рівня штату [7].

Експертами з питань вимірювання бідності в Україні у 2011 році визначені основні показники бідності, надійність оцінювання яких має бути забезпечена на належному рівні: це рівень бідності за національною межею, рівень бідності за регіональною межею та рівень бідності за прожитковим мінімумом [1; 2; 8]. Розглянемо зміст цих показників та формули для їх розрахунку.

1. Рівень бідності за національною межею – це питома вага сімей (домогосподарств), у яких рівень споживання (доходів) на одну особу є нижчим від визначеної межі бідності. Цей показник позначається P_1 і розраховується за формулою [9]:

$$P_1 = \frac{Q}{N} \cdot 100\%, \quad (1.9)$$

де Q – чисельність населення, що визнається бідним відповідно до прийнятої національної межі бідності; N – загальна чисельність населення України.

Межа бідності визначається на підставі відносного критерію зарахування різних верств населення до категорії

бідних, який обчислюється за фіксованою часткою середньодушового доходу (витрат) – 75% медіанного рівня сукупних доходів (витрат) у розрахунку на умовного дорослого.

2. Рівень бідності за регіональною межею – це частка населення регіону, доходи яких є нижчими за розраховану межу бідності. Цей рівень бідності P_2 обчислюється за формулою:

$$P_{2,r} = \frac{Q_r}{N_r} \cdot 100\%, \quad (1.10)$$

де Q_r – чисельність населення регіону, яке визнається бідним відповідно до межі бідності для цього регіону; N_r – загальна чисельність населення регіону.

3. Рівень бідності за розміром прожиткового мінімуму – це частка населення з загальними середньомісячними еквівалентними доходами, нижчими за прожитковий мінімум у середньому на одну особу. Цей рівень бідності P_3 визначається за формулою:

$$P_{3,r} = \frac{Q_{r(m)}}{N_r} \cdot 100\%, \quad (1.11)$$

де $Q_{r(m)}$ – чисельність населення регіону з загальними середньомісячними еквівалентними доходами, нижчими за прожитковий мінімум; N_r – загальна чисельність населення регіону.

Базовою оціночною функцією при прямому оцінюванні показників бідності є функція Гурвіца – Томпсона, що передбачає обчислення показника за результатами вибіркового обстеження з урахуванням статистичних ваг обстежених одиниць сукупності, тобто дає можливість поширювати дані обстеження на генеральну сукупність [5; 10]. Якщо за результатами вибіркового обстеження необхідно оцінити, наприклад, сумарне значення $\hat{Y}$ за певною ознакою y_i , то розрахунок здійснюється на основі оціночної функції такого вигляду:

$$\hat{Y}_{HT} = \sum_{i=1}^n \left(w_i^{(d)} \cdot y_i \right), \quad (1.12)$$

де $\hat{Y}_{HT}$ – оціночна функція Гурвіца – Томпсона; $w_i^{(d)}$ – статистичні ваги дизайну, визначені як величини, обернені до імовірностей включення одиниць до вибірки π_i ; n – обсяг вибірки.

При розрахунку статистичної ваги i -ї обстеженої одиниці вибірки (наприклад, домогосподарства) враховується, що вона має загальну ймовірність потрапити до вибірки π_i , тобто репрезентує $w_i^{(d)} = 1/\pi_i$ одиниць генеральної сукупності та при визначенні оцінок показників за генеральною сукупністю її необхідно врахувати $w_i^{(d)}$ разів. Слід зазначити, що у реальних обстеженнях кінцева статистична вага одиниць вибірки визначається як імовірністю її відбору, так і ймовірністю її обстеження (отримання інформації від неї) [10]. Відповідно, прямі оціночні функції, що використовуються на практиці, лише за формою є функціями Гурвіца – Томпсона (1.12).

Розглянемо особливості розрахунку прямої оцінки рівня бідності за результатами вибіркового обстеження умов життя домогосподарств. Як вже зазначалося, рівень бідності – це відносний показник, що характеризує питому вагу населення, рівень споживання якого є нижчим за визначену межу. За даними ОУЖД він розраховується за такою узагальненою формулою:

$$P_U = \frac{Q_{HT}}{N_{HT}} \cdot 100\%, \quad (1.13)$$

де P_U – оцінка рівня бідності по генеральній сукупності (країні, регіону); Q_{HT} – оцінка чисельності бідного населення; N_{HT} – оцінка загальної чисельності неінституційного населення.

Оціночна функція для розрахунку чисельності бідного населення є аналогічною функції (1.12) і має вигляд:

$$Q_{HT} = \sum_{i=1}^n w_i \cdot h_i \cdot \alpha_i, \quad (1.14)$$

де w_i – кінцева статистична вага i -го домогосподарства; h_i – розмір i -го домогосподарства (кількість його членів); α_i – бінарна змінна, $\alpha_i = 1$ для i -го домогосподарства, якщо воно є бідним, і $\alpha_i = 0$, якщо ні.

Доцільно звернути увагу, що за такого підходу (який є дуже поширеним при оцінюванні бідності) поза увагою залишається питання розподілу доходів між членами домогосподарства, тобто всі його члени вважаються однаково бідними. Це припущення не дуже точно відповідає дійсності, але суттєво спрощує вимірювання бідності.

Слід зазначити, що при визначенні межі бідності, як правило, використовують показник середньодушових витрат (або доходів) у розрахунку на умовного дорослого. Для цього застосовується так звана шкала еквівалентності, що дозволяє врахувати економію від спільного користування товарами та послугами особами, які складають домогосподарство [9; 11]. Кожній окремій особі присвоюється певний коефіцієнт, який враховує її потреби, а сума коефіцієнтів для кожного члена домогосподарства і складає еквівалентний розмір останнього.

У міжнародній практиці для визначення еквівалентного розміру домогосподарства застосовують декілька шкал еквівалентності. Так, за Оксфордською шкалою приймається, що потреби першого дорослого члена сім'ї дорівнюють 1,0, кожного іншого дорослого – 0,7, кожної дитини – 0,5. В Україні в основу розрахунку еквівалентного розміру домогосподарства покладено шкалу, за якою потреби першого дорослого члена домогосподарства приймаються рівними 1,0,

а всіх інших – 0,7 [11]. При цьому формула для визначення еквівалентного розміру домогосподарства (h_i^e) має вигляд:

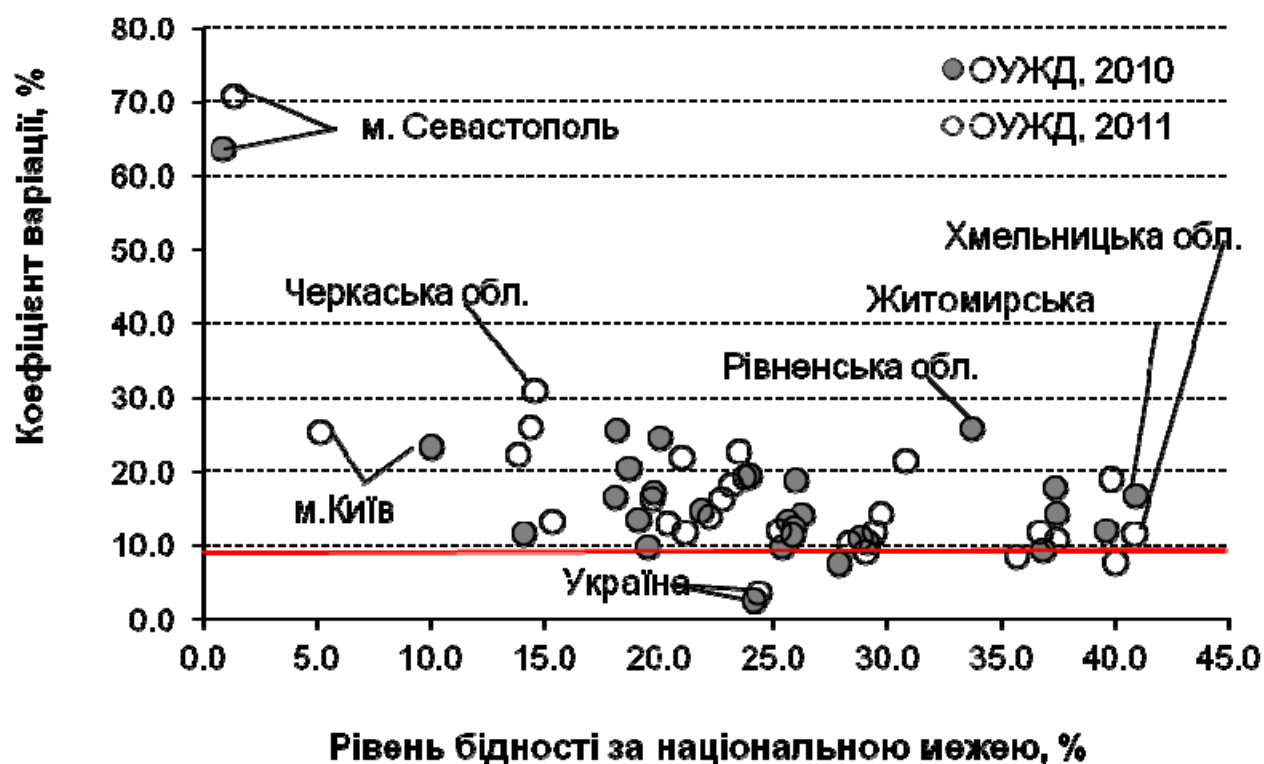
$$h_i^e = 1 + 0,7(h_i - 1). \quad (1.15)$$

Оціночна функція для розрахунку загальної чисельності неінституційного населення (знаменник у формулі (1.13)) має вигляд:

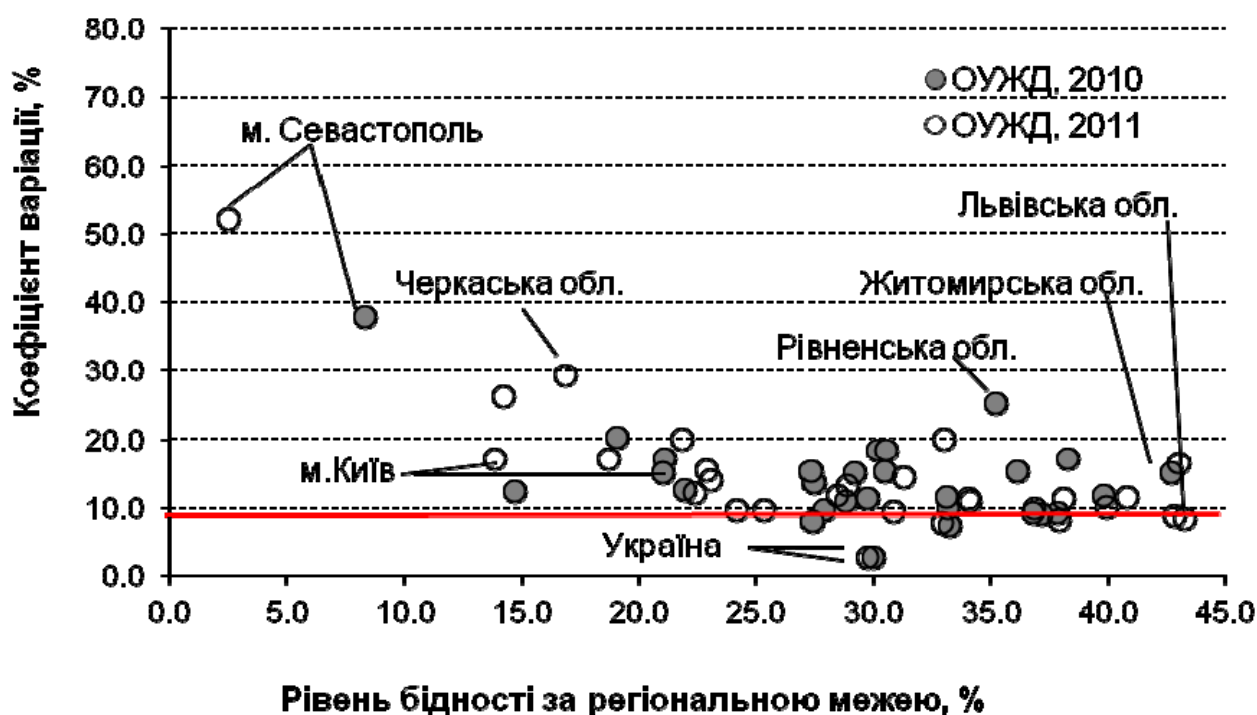
$$N_{HT} = \sum_{i=1}^n w_i \cdot h_i. \quad (1.16)$$

Результати виконаних досліджень свідчать, що в Україні при вимірюванні бідності вимоги щодо надійності оцінювання основних показників задовольняються не повною мірою. За існуючих системи показників бідності та інформаційної бази їх розрахунку статистична надійність прямих оцінок показників, визначених на рівні регіонів, є недостатньою, що проілюстровано даними, наведеними на рис. 1.2 та у додатку А.

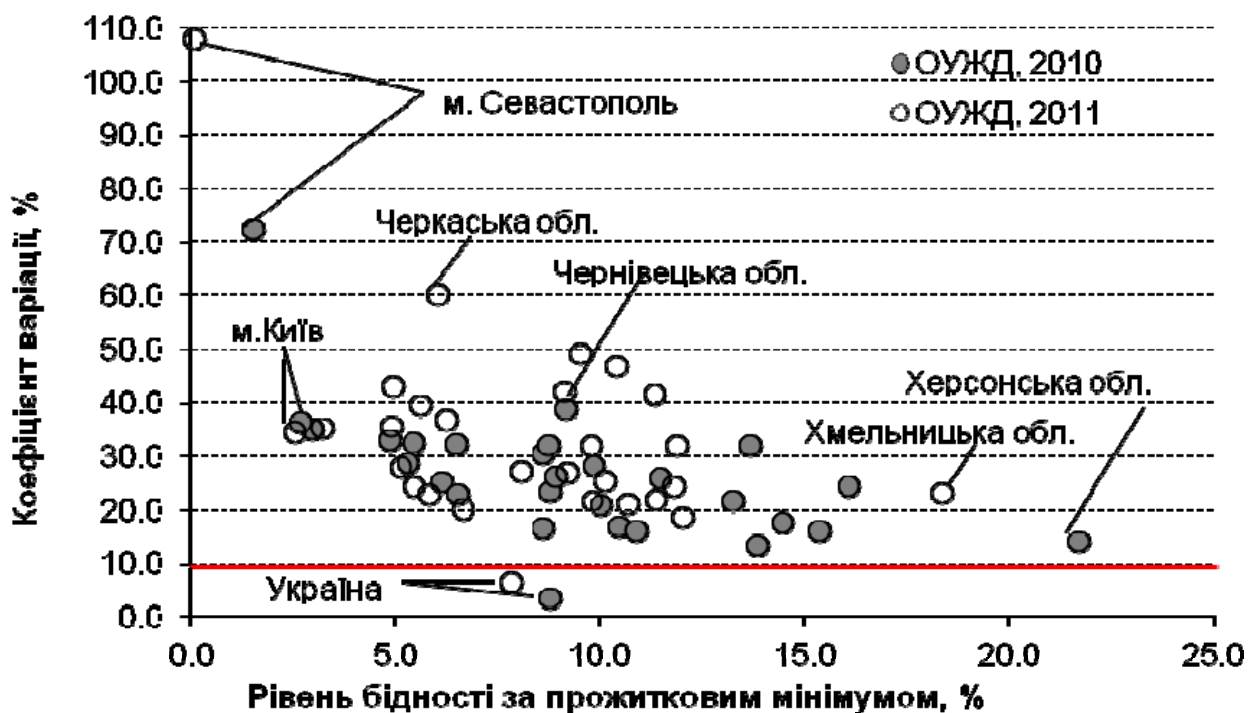
Обсяг і дизайн вибірки обстеження умов життя домогосподарств забезпечують достатню надійність річних оцінок показників бідності на загальнодержавному рівні. Так, рівень бідності у 2011 році за національною межею становив близько 24,4% з коефіцієнтом варіації 3,59%, за регіональною межею, відповідно, 29,8% і 2,68%, за прожитковим мінімумом – 7,8% і 6,47%. При цьому на регіональному рівні коефіцієнти варіації оцінки рівня бідності за національною межею досягають 25,35% (для м. Києва) та 30,96% (для Черкаської області). У цілому аналогічна картина спостерігалася і у попередні роки. По м. Севастополь показники бідності, враховуючи їх дуже низьку надійність, не публікуються.



а)



б)



в)

Рис. 1.2. Надійність прямих оцінок показників бідності для національного та регіонального рівня (ОУЖД, 2010–2011 рр.):

- а) рівень бідності за національною межею;
- б) рівень бідності за регіональною межею;
- в) рівень бідності за прожитковим мінімумом

Проблема недостатньої надійності прямого оцінювання показників для окремих територій або груп одиниць генеральної сукупності за результатами вибірових обстежень відома як проблема малих територій. Недостатній рівень надійності прямих оцінок основних показників бідності для регіонів України є прикладом зазначеної проблеми. Як було вказано вище, пряме оцінювання показників здійснюється за стандартними прямими оціночними функціями. За такого підходу оцінка показника розраховується на основі поточних даних вибіркового обстеження. Однак у багатьох випадках цього виявляється недостатньо для надійного оцінювання, що пояснюється відносно малими обсягами вибірки, на основі яких

здійснюється оцінювання показників для регіонів, порівняно з обсягом вибірки на національному рівні.

У цілому проблема забезпечення надійності оцінок показників бідності в Україні може розв'язуватися двома основними шляхами:

- оптимізація дизайну вибірки, зокрема збільшення обсягу вибіркової сукупності ОУЖД для регіонів з недостатнім рівнем надійності оцінювання показників;

- розробка та впровадження спеціальних методологічних підходів до оцінювання показників бідності на основі методів статистичного моделювання, у тому числі з урахуванням наявної додаткової (зовнішньої) інформації.

Слід зазначити, що ці підходи є взаємозалежними (у тому сенсі, що моделі враховують розподіл вибірки за регіонами, який, у свою чергу, має враховувати надійність оцінювання показників на основі моделей) і на практиці можуть застосовуватись одночасно. При цьому забезпечення надійності показників бідності на регіональному рівні шляхом збільшення обсягу вибіркової сукупності ОУЖД є проблематичним, оскільки потребує значних фінансових витрат (збільшення як кількості інтерв'юєрів для проведення обстеження, так і працівників, задіяних для обробки й аналізу даних, зростання витрат на плату домогосподарствам за участь в обстеженні та ін.). Крім того, вимірювання бідності не є основною метою ОУЖД, а лише однією з багатьох його цілей [3].

Досвід інших країн свідчить, що проблему підвищення надійності оцінок окремих показників можна дещо пом'якшити за рахунок перерозподілу обсягу вибірки між регіонами (оптимізації дизайну вибірки). Певні кроки в цьому напрямі зроблені Держстатом України при вдосконаленні дизайну вибірки ОУЖД для проведення обстежень у 2009 – 2013 роках

[4]. Однак розробка та впровадження ефективних підходів до оптимізації дизайну вибірки наразі залишається актуальним завданням.

Ефективним способом підвищення надійності показників бідності для різних рівнів агрегації даних є застосування спеціальних методологічних підходів, що базуються на методах непрямого оцінювання. Варто зауважити, що впровадження таких методів, як правило, вимагає певного ускладнення технології обробки даних обстеження, відповідного вдосконалення організаційних структур та підвищення рівня кваліфікації фахівців, які займаються розрахунками показників бідності.

1.2. Загальна характеристика методології непрямого оцінювання показників бідності

За результатами опрацювання існуючих підходів до підвищення рівня надійності оцінювання показників за даними вибірових обстежень (див., наприклад, [12–16]) можна зробити висновок, що в теперішній час найбільш ефективним підходом є застосування так званих методів «малих територій», які належать до непрямих методів оцінювання. Специфічна особливість цих методів полягає у комплексному використанні існуючої інформації, що характеризує досліджуване явище або процес, а також урахуванні статистичних властивостей сукупності даних вибірового обстеження. Головною ідеєю методів малих територій є те, що для підвищення надійності оцінювання показників, крім безпосередніх результатів певного обстеження, використовується інформація з додаткових джерел даних на мікрорівні, на рівні малої території або на рівні більшої території, до складу якої входить мала територія, що

розглядається. Наприклад, це можуть бути дані попередніх етапів обстеження, кореляти показників бідності, отримані з адміністративних джерел, за даними статистичної звітності, за результатами переписів тощо.

Низка методологічних підходів вже певний час застосовуються у міжнародній практиці та навіть отримали назву «традиційних методів малих територій». Зокрема, такі методи використовуються ще з вісімдесятих років минулого століття у США та Канаді для оцінювання показників доходів, витрат, бідності, зайнятості й безробіття, показників стану здоров'я тощо [13].

На сьогодні підходи до оцінювання офіційних статистичних показників на основі методології малих територій найбільш широко використовують у США. У Бюро переписів США розроблено та у 1997 році впроваджено методологію оцінювання показників доходів (медіанний дохід) та бідності (чисельність бідного населення за віковими групами) для малих територій – штатів та округів (Small Area Income and Poverty Estimates – SAIPE) [17]. За цією методологією здійснюються як поточні, так і ретроспективні розрахунки. Так, оцінки показників бідності для рівня штатів опубліковані за 1989, 1993 та за всі роки, починаючи з 1995 року. Оцінки показників бідності за цією методологією ґрунтуються на даних щодо складу домогосподарств та їх доходів, оцінених із обстежень та адміністративних джерел, а саме даних вибіркового Поточного обстеження населення (Current Population Survey – CPS), даних перепису населення, даних адміністративних реєстрів: «Продуктовий талон» (Food Stamp), додаткові програми безпеки доходів (Supplemental Security Income Programs) та реєстр повернення федеральних податків (Federal Tax Returns). При цьому використовуються дані CPS за три послідовні роки. З 2005 року

для оцінювання показників за методологією SAIPE застосовуються також дані вибіркового обстеження американських громад (American Community Survey), у якому також вимірюються показники доходів домогосподарств.

На рівні штатів і округів розраховують такі чотири показники:

- медіанний дохід домогосподарств;
- чисельність населення, рівень життя якого нижчий за межу бідності;
- чисельність дітей у віці до 18 років, рівень життя яких нижчий за межу бідності;
- чисельність дітей у віці 5–17 років у домогосподарствах, рівень життя яких нижчий за межу бідності.

При цьому на рівні штатів додатково визначається ще показник «Чисельність дітей у віці до 5 років, рівень життя яких нижчий за межу бідності».

Для оцінювання показників застосовуються методика, яка полягає у побудові спеціальних оціночних функцій на базі прямих оціночних функцій, що використовуються для оцінювання показників за результатами вибіркового обстеження, та синтетичних оціночних функцій, побудованих на основі регресійних моделей. Зазначимо, що з часу розробки перших оціночних функцій для уточнення оцінок бідності в США ці підходи пройшли декілька етапів удосконалення.

Дещо схожа за змістом стратегія оцінювання розроблена і впроваджена у Міністерстві праці США. Ця стратегія застосовується для оцінки показників безробіття за окремими адміністративно-територіальними одиницями і має назву «Статистика безробіття для локальних територій» (Local Area Unemployment Statistics).

У країнах ЄС до вирішення питань застосування методології малих територій в офіційній статистиці підійшли набагато пізніше, але більш фундаментально. У 2000 році був створений міжнародний консорціум, який за кошти спеціального проекту Євростату до 2004 року провів певні роботи з опрацювання існуючих методів малих територій і розробив рекомендації щодо використання цих методів в офіційній статистиці [16]. З 2008 до 2011 року було реалізовано новий аналогічний проект (Small Area Methods for Poverty and Living Condition Estimates – SAMPLE), двома основними завданнями якого стало дослідження з питань визначення нових показників бідності та диференціації населення за рівнем життя й розробка статистичних моделей малих територій для оцінювання таких показників і характеристик їх точності для рівнів територіального поділу країн, найбільш актуальних для користувачів. Це планується реалізувати через інтеграцію даних обстежень умов та рівня життя в ЄС (EU-SILC) з адміністративними даними країн – членів ЄС [18].

Серед традиційних методів малих територій, які пропонуються для застосовування в офіційній статистиці для вирішення проблеми недостатньої надійності оцінок показників, пріоритетне значення має метод композиційного оцінювання (composite estimation or shrinkage estimation) [12–16]. Історично метод композиційного оцінювання виник у результаті пошуку можливостей об'єднання (композиції) прямих і непрямих оціночних функцій. При цьому як непрямі можуть застосовуватися так звані синтетичні оціночні функції, побудовані на основі статистичних моделей з використанням зовнішньої (додаткової) інформації [13; 14]. У сучасній статистичній практиці цей метод став ефективним інструментом розв'язання

проблеми підвищення надійності оцінок показників на різних рівнях агрегації даних.

У спрощеному вигляді оціночна функція при композиційному оцінюванні певного цільового показника Y для регіону k може бути представлена у такому вигляді [13; 14]:

$$\hat{Y}_k^{(H)} = (1 - \gamma_k) \hat{Y}_k^{(D)} + \gamma_k \hat{Y}_k^{(S)}, \quad (1.17)$$

де $\hat{Y}_k^{(H)}$ – композиційна оціночна функція; $\hat{Y}_k^{(D)}$ – пряма оціночна функція; $\hat{Y}_k^{(S)}$ – синтетична оціночна функція; γ_k – ваговий коефіцієнт, який, як правило, визначається з умови мінімізації середньоквадратичної похибки оцінювання (цей коефіцієнт приймає значення від 0 до 1); $k = 1, 2, \dots, K$, K – загальна кількість регіонів.

Особливістю цього методу є те, що композиційна оціночна функція забезпечує можливість оцінювання показників навіть для територій, які не увійшли до територіальної вибірки і за якими, відповідно, неможливо здійснити пряме оцінювання. У таких випадках значення коефіцієнта γ_k у формулі (1.17) дорівнює 1.

Доцільно враховувати, що найбільший ефект щодо підвищення надійності оцінювання показників, зокрема і показників бідності, за результатами вибірових обстежень досягається при комплексному застосуванні різних підходів, наприклад при близьких до оптимального дизайнах вибірки, належній обробці даних, реалізації найбільш ефективних процедур оцінювання показників та ін. Але найбільш ефективним в умовах існуючих дизайну вибірки й технології обробки даних є застосування методології та процедур непрямого оцінювання показників. Цей підхід до підвищення рівня надійності оцінок показників бідності впроваджується як додаток до стандартних процедур –

збирання даних, їх обробки, обчислення прямих оцінок показників тощо, – які реалізуються на національному рівні (наприклад, у відповідних підрозділах Держстату України). Таким чином, загальна технологія обробки результатів обстежень не зазнає суттєвих змін. Новий етап обробки даних необхідно ввести лише на національному рівні, де створено найкращі умови для накопичення відповідного досвіду.

Існують і більш складні процедури оцінювання показників з використанням результатів вибірових обстежень та зовнішньої інформації, але їх застосування потребує поглиблених досліджень, які необхідно вести одночасно з апробацією та впровадженням. Для скорочення ризиків на цьому етапі варто зосередитися на відносно простих моделях, застосування яких забезпечує також краще розуміння аналітиками та користувачами процесів оцінювання, основних визначень і можливостей використання. Лише після накопичення необхідного досвіду щодо розробки, верифікації та впровадження моделей, детального дослідження властивостей отриманих оцінок і допоміжних джерел інформації слід у подальшому удосконалювати моделі для оцінювання.

При практичному впровадженні методу композиційного оцінювання слід урахувати, що його застосування доцільне для уточнення лише певних, найбільш важливих показників через складність визначення зовнішньої інформації та побудови непрямих оціночних функцій. Крім того, необхідно враховувати такі загальні вимоги до цього методу оцінювання показників бідності: розроблені статистико-математичні моделі мають забезпечувати оцінювання показників бідності для всіх регіонів України зі встановленим рівнем надійності на основі визначених джерел даних, у тому числі і джерел зовнішньої інформації; джерела даних, які використовуватимуться для уточнення показників бідності на основі моделей, повинні бути доступними

на регулярній основі; розроблені моделі та алгоритми повинні бути зручними для автоматизації програмними засобами, що використовуються в системі державної статистики України. Терміни обробки даних для моделювання та розрахунків уточнених показників мають бути мінімізовані наскільки це можливо.

Для адекватного визначення методологічних основ побудови композиційних оціночних функцій доцільно розглянути ряд базових показників та їх позначень.

Нехай за результатами ОУЖД за поточний рік розраховано оцінку j -го показника бідності для k -го регіону $P_k^{(j)}$ за прямою оціночною функцією (наприклад, це може бути оцінка рівня бідності за національною межею ($j=1$) для Автономної Республіки Крим ($k=1$)). Припустимо, що додаткова (зовнішня) інформація наявна для регіону k у вигляді деяких показників $x_{1k}^{(j)}, x_{2k}^{(j)}, \dots, x_{Mk}^{(j)}$, де M – загальна кількість показників із додаткових джерел інформації (коваріат).

Непряме оцінювання показників бідності на рівні регіонів України пропонується здійснювати на основі багатofакторної композиційної оціночної функції, що визначається за формулою [1; 2]:

$$\hat{P}_{comp,k}^{(j)} = \left(u^{(i)} - \hat{\gamma}_k^{(j)} \right)^T \hat{\theta}_k^{(j)} + \hat{\gamma}_k^{(j)T} \hat{\theta}^{(j)}, j = 1, 2, 3, \quad (1.18)$$

де $\hat{P}_{comp,k}^{(j)}$ – вектор оцінок j -го показника бідності для регіонів за композиційною оціночною функцією; $\hat{\theta}_k^{(j)}$ – вектор, який містить пряму оцінку j -го показника бідності та оцінки його коваріат (зовнішньої інформації) на регіональному рівні. Наприклад, якщо для j -го показника бідності визначено M коваріат, то

$$\hat{\theta}_k^{(j)} = \begin{pmatrix} \hat{p}_k^{(j)} \\ x_{1k}^{(j)} \\ \dots \\ x_{Mk}^{(j)} \end{pmatrix}, \text{ де } \hat{p}_k^{(j)} - \text{пряма оцінка } j\text{-го показника}$$

бідності для k -го регіону; $x_{1k}^{(j)}, \dots, x_{Mk}^{(j)}$ – коваріати j -го показника бідності для k -го регіону;

$\theta^{(j)}$ – вектор який містить пряму оцінку j -го показника бідності на національному рівні:

$$\theta^{(j)} = \begin{pmatrix} P^{(j)} \\ 0 \\ \dots \\ 0 \end{pmatrix}, \text{ де } P^{(j)} - \text{пряма оцінка } j\text{-го показника}$$

бідності для національного рівня;

$\hat{\gamma}_k^{(j)}$ – вектор оцінок вагових коефіцієнтів композиційної оціночної функції за регіонами;

$$u^{(j)} - \text{індексний вектор, } u^{(j)} = \begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \end{pmatrix};$$

індекс j відносить показники рівня бідності до відповідних критеріїв:

$j=1$ – «Рівень бідності за національною межею»; $j=2$ – «Рівень бідності за регіональною межею»; $j=3$ – «Рівень бідності за розміром прожиткового мінімуму»; індекс T позначає операцію транспонування вектора.

Доцільно зазначити, що у виразі (1.18) синтетична оцінка включена до вектору регіональних оцінок $\hat{\theta}_k^{(j)}$, а вагові коефіцієнти при коваріатах показників бідності на регіональному рівні за формою запису (але не за процедурою визначення та змістом) можна представити як добуток вагового коефіцієнта

однофакторної оціночної функції та коефіцієнтів регресії. За такого підходу не вимагається побудови у явному вигляді регресійної моделі для синтетичних оцінок, що спрощує процедуру оцінювання і не потребує врахування обмежень моделей регресії [14; 15].

Також заслуговує на увагу врахування у композиційній оціночній функції (1.18) прямої оцінки показника бідності для національного рівня $\theta^{(j)}$. Ця оцінка є класичною формою синтетичної оцінки для регіонального рівня. Її застосування у кінцевому варіанті запропонованих моделей для підвищення надійності оцінювання показників бідності в Україні зумовлено відсутністю зовнішніх даних (насамперед мова йде про адміністративну інформацію та дані суцільних спостережень), які б мали високий рівень придатності для використання при побудові моделей. Ефективність застосування оцінок національного рівня підтверджена результатами апробації запропонованих моделей.

При композиційному оцінюванні необхідно визначити такий вектор вагових коефіцієнтів $\gamma_k^{(j)}$, який забезпечує, за заданих умов, мінімальний рівень середньоквадратичної похибки оцінювання відповідних показників бідності для регіонів. Вектор $\gamma_k^{(j)}$ визначається за таким співвідношенням, записаним у матричній формі [14; 15]:

$$\gamma_k^{(j)} = \left[V_k^{(j)} + V^{(j)} - C_k^{(j)} - C_k^{(j)T} + \Sigma_b^{(j)} \right]^{-1} \left[V_k^{(j)} - C_k^{(j)} \right] u^{(j)}, \quad (1.19)$$

де $V_k^{(j)}$ – вектор дисперсій вибірових оцінок показників регіонального рівня;

$V^{(j)}$ – вектор дисперсій оцінок показників національного рівня;

$C_k^{(j)}$ – коваріаційна матриця регіональних та національних оцінок показників;

$\Sigma_b^{(j)}$ – матриця дисперсій регіональних оцінок показників щодо їх середнього значення за регіонами.

Для однофакторної композиційної оціночної функції співвідношення (1.19) має такий вигляд [15]:

$$\gamma_k^{(j)} = \frac{v_k^{(j)} - c_k^{(j)}}{v_k^{(j)} + v^{(j)} - 2c_k^{(j)} + \sigma_B^{(j)2}}, \quad (1.20)$$

де $v_k^{(j)}$ – дисперсія вибірових оцінок j -го показника бідності для k -го регіону $P_k^{(j)}$;

$v^{(j)}$ – дисперсія вибірових оцінок j -го показника бідності для національного рівня $P^{(j)}$;

$c_k^{(j)}$ – коваріація показників $P_k^{(j)}$ та $P^{(j)}$;

$\sigma_B^{(j)2}$ – дисперсія регіональних оцінок j -го показника бідності щодо їх середнього значення для k -го регіону:

$$\sigma_B^{(j)2} = \frac{1}{K} \sum_{k=1}^K \left(P_k^{(j)} - \bar{P}^{(j)} \right)^2, \quad (1.21)$$

де $\bar{P}^{(j)}$ – середнє значення j -го показника бідності за регіонами (у формулі (1.21) маються на увазі дійсні значення показників, а не їх оцінки).

Як показано у [15] дисперсія $\sigma_B^{(j)2}$ може бути апроксимована таким виразом:

$$\sigma_B^{(j)2} = \frac{1}{K} \sum_{k=1}^K \left(P_k^{(j)} - \bar{P}^{(j)} \right)^2 - \frac{1}{K} \sum_{k=1}^K v_k^{(j)} \left(1 - \frac{2N_k}{\frac{1}{K} \sum_{k=1}^K N_k} \right), \quad (1.22)$$

де N_k – середньорічна чисельність постійного населення k -го регіону.

Слід зазначити, що на практиці вагові коефіцієнти $\gamma_k^{(j)}$ лише оцінюють, оскільки більшість компонентів у виразах (1.19) та (1.20) також лише оцінюється, особливо для вибірок зі складним дизайном.

$$\hat{\gamma}_k^{(j)} = \frac{\hat{v}_k^{(j)} - \hat{c}_k^{(j)}}{\hat{v}_k^{(j)} + \hat{v}^{(j)} - 2\hat{c}_k^{(j)} + \hat{\sigma}_B^{(j)2}}. \quad (1.23)$$

Якщо взяти до уваги, що величини $\hat{v}^{(j)}$ і $\hat{c}_k^{(j)}$ набагато менші за величини $\hat{v}_k^{(j)}$, то вираз (1.23) можна спростити і представити у такому вигляді [15]:

$$\hat{\gamma}_k^{(j)} = \frac{\hat{v}_k^{(j)}}{\hat{v}_k^{(j)} + \hat{\sigma}_B^{(j)2}}. \quad (1.24)$$

Вираз (1.24) свідчить, що основними джерелами похибки при композиційному оцінюванні показників для регіонів є саме надійність оцінювання регіональних показників та варіація значень показника за регіонами.

За оцінки вагового коефіцієнта (1.24) наближена оцінка середньоквадратичної похибки при однофакторному композиційному оцінюванні може бути представлена у вигляді:

$$M\hat{S}E(P_{comp,k}^{(j)}) \approx \hat{v}_k^{(j)} - \frac{\hat{v}_k^{(j)2}}{\hat{v}_k^{(j)} + \hat{\sigma}_B^{(j)2}}. \quad (1.25)$$

У цілому побудова композиційної оціночної функції для уточнення оцінок визначених показників бідності за регіонами за певний рік складається з таких основних етапів:

- визначення прямих оцінок показників бідності за регіонами;
- оцінка дисперсій вибірових оцінок для прямих оцінок показників бідності за регіонами;

- вибір зовнішньої інформації;
- оцінка вагових коефіцієнтів композиційної оціночної функції $\hat{\gamma}_k$ та розрахунки композиційних оцінок показників бідності для регіонів і характеристик їх надійності.

Кінцевим критерієм ефективності оціночної функції є величина середньоквадратичної похибки $MSE(P_{comp,k}^{(j)})$. Тому оптимальний вид багатофакторної композиційної оціночної функції доцільно визначати саме за найменшою величиною середньоквадратичної похибки.

Доцільно враховувати, що композиційне оцінювання (уточнення) показників бідності за регіонами передбачає перерахунок відповідних показників для рівня України. При цьому для розрахунку показника на національному рівні застосовується така загальна формула:

$$P_{comp,U}^{(j)} = \frac{\sum_{k=1}^K N_k P_{comp,k}^{(j)}}{\sum_{k=1}^K N_k}, \quad (1.26)$$

де $P_{comp,U}^{(j)}$ – уточнена оцінка показника бідності на національному рівні;

$P_{comp,k}^{(j)}$ – композиційна оцінка показника бідності для k -го регіону.

1.3. Визначення надійності композиційних оцінок показників бідності

При оцінюванні показників бідності з використанням композиційних оціночних функцій за основний критерій надійності доцільно прийняти таку умову: значення абсолютної або відносної похибки оцінювання, які у загальному випадку є похідними від величини середньоквадратичної похибки $MSE(P_{comp,k}^{(j)})$, по-

винні бути меншими за встановлене граничне значення відповідної похибки.

Абсолютною похибкою оцінювання показників бідності з використанням композиційної оціночної функції є величина кореня квадратного від середньоквадратичної похибки композиційної оцінки:

$$RMSE(P_{comp,k}^{(j)}) = \sqrt{MSE(P_{comp,k}^{(j)})} . \quad (1.27)$$

Відносною похибкою є величина відношення кореня квадратного від середньоквадратичної похибки до композиційної оцінки показника бідності за регіонами:

$$\begin{aligned} RRMSE(P_{comp,k}^{(j)}) &= \frac{RMSE(P_{comp,k}^{(j)})}{P_{comp,k}^{(j)}} \cdot 100\% = \\ &= \frac{\sqrt{MSE(P_{comp,k}^{(j)})}}{P_{comp,k}^{(j)}} \cdot 100\% . \end{aligned} \quad (1.28)$$

Величина $RRMSE(P_{comp,k}^{(j)})$ є аналогом коефіцієнта варіації для прямих оцінок і відрізняється від нього тим, що враховує зміщення оцінок. Межі значень $RRMSE(P_{comp,k}^{(j)})$, що визначають рівень надійності показників, такі самі, як і для коефіцієнтів варіації. Композиційна оцінка показника бідності може вважатися достатньо надійною для регіонального рівня, якщо значення $RRMSE(P_{comp,k}^{(j)})$ не перевищує 10%, в окремих випадках – 15%. Якщо значення $RRMSE(P_{comp,k}^{(j)})$ знаходиться в інтервалі 15–25%, використання показника рекомендується лише для якісного аналізу. Якщо значення $RRMSE(P_{comp,k}^{(j)})$ перевищує 25%, показник взагалі не рекомендується використовувати для аналізу бідності.

Для визначення та аналізу надійності композиційного оцінювання показників також використовують величину граничної похибки вибірки $ME(P_{comp,k}^{(j)})$, яка для довірчої імовірності 95% визначається за формулою:

$$ME(P_{comp,k}^{(j)}) = 1,96 \cdot RMSE(P_{comp,k}^{(j)}). \quad (1.29)$$

Для оцінки результатів моделювання здійснюються такі основні перевірки:

а) перевіряються оцінки відносних похибок – відношень кореня квадратного від середньоквадратичної похибки до композиційної оцінки показника бідності за регіонами $RRMSE(P_{comp,k}^{(j)})$ – з метою підтвердження того, що їх значення відповідають встановленим вимогам щодо рівня надійності оцінювання показників бідності за регіонами;

б) порівнюються прямі $P_k^{(j)}$ та композиційні $P_{comp,k}^{(j)}$ оцінки показників з метою перевірки того, що останні потрапляють до довірчих інтервалів, побудованих для відповідних прямих оцінок – $P_k^{(j)} \pm ME(P_k^{(j)})$. Це дозволяє виявити потенційні випадки, коли композиційна оцінка для окремого регіону відхиляється від прямої на значну величину порівняно з довірчим інтервалом, що може свідчити про неадекватність моделі для цього регіону та вимагати проведення спеціального аналізу.

Середньоквадратичні похибки оцінок показників бідності на національному рівні $MSE(P_{comp,u}^{(j)})$ наближено можуть бути визначені на основі середньоквадратичних похибок оцінок показників для регіонів за формулою:

$$MSE(P_{comp,u}^{(j)}) = \frac{1}{N^2} \sum_{k=1}^K N_k^2 MSE(P_{comp,k}^{(j)}). \quad (1.30)$$

Похідні характеристики надійності оцінок показників бідності на національному рівні розраховуються на основі формул (1.27) – (1.29).

При застосуванні методів непрямого оцінювання на практиці серйозною проблемою є визначення саме величини $MSE(P_{comp,k}^{(j)})$. При оцінці цієї величини за наявності даних вибіркового обстеження та зовнішньої інформації можуть застосовуватись методи моментів, метод максимальної правдоподібності (*ML* – Maximum Likelihood) або максимальної правдоподібності з обмеженнями (*REML* – Restricted Maximum Likelihood) [13; 14]. Процедури оцінювання $MSE(P_{comp,k}^{(j)})$ детально розглянуті в існуючих публікаціях з питань оцінювання показників для малих територій і реалізовані у статистичному програмному забезпеченні, зокрема в *R*, яке працює як незалежно, так і у середовищі *SPSS*.

За досвідом застосування та практичністю при оцінюванні величини $MSE(P_{comp,k}^{(j)})$ перевагу іноді надають так званим реплікаційним методам. Ураховуючи, що в розпорядженні фахівців з непрямого оцінювання, як правило, є результати вибіркового обстеження на мікрорівні, величину дисперсії композиційної оцінки та її зміщення можна визначити, порівнявши композиційні оцінки за багаторазово побудованими реплікаціями (підвибірками з основної вибірки) з середнім значенням цих оцінок та середнім значенням прямої оцінки, яка вважається незміщеною. На кожній реплікації здійснюється повна процедура композиційного оцінювання. Кількість необхідних реплікацій залежить від властивостей композиційної оціночної функції та може досягати кількох тисяч. Процедура зупиняється за умови, що оцінка $MSE(P_{comp,k}^{(j)})$ на певній реплікації змінюється на величину, меншу за деяку наперед задану величину, наприклад менш ніж на 0,5%.

РОЗДІЛ 2

ОСОБЛИВОСТІ ВИКОРИСТАННЯ ЗОВНІШНЬОЇ ІНФОРМАЦІЇ ПРИ НЕПРЯМОМУ ОЦІНЮВАННІ ПОКАЗНИКІВ БІДНОСТІ

2.1. Джерела зовнішньої інформації для побудови синтетичних оціночних функцій

Джерела, структура та обсяги інформації, що може бути використана для цілей підвищення надійності оцінювання показників бідності, визначаються, головним чином, діючою методологією вимірювання бідності та наявністю й доступністю надійних статистичних і адміністративних даних (зовнішньої інформації) щодо рівня життя населення [1].

При визначенні джерел зовнішніх даних для оцінювання показників бідності з застосуванням моделей слід брати до уваги, що використання значної кількості джерел призводить, як правило, до зменшення величини зміщення оцінок за моделями і забезпечує повноту урахування всіх можливих взаємозв'язків показників бідності із зовнішніми даними, але майже завжди зумовлює зростання величини дисперсії оцінок показників, що пояснюється зростанням нестабільності оцінок параметрів моделей [14; 15]. Головною особливістю зовнішньої інформації, яка може бути використана при оцінюванні показників бідності із застосуванням моделей, є те, що вона має бути статистично (а бажано й економічно) тісно пов'язаною з відповідними показ-

никами і бути достатньо надійною. Це, певною мірою, ускладнює можливість отримання інформації, корисної для оцінювання показників, з інших джерел. Відповідно, перелік джерел зовнішньої інформації, які доцільно використовувати з метою підвищення надійності оцінюванні показників бідності частіше за все є достатньо обмеженим [1].

Для вибору незалежних зовнішніх змінних (корелят бідності) насамперед аналізують парні коефіцієнти кореляції Пірсона залежної змінної (прямих оцінок показників бідності або певних функцій від них) з кожною з L потенційних незалежних змінних $x_k^{(j)}, \dots, x_{Lk}^{(j)}$. Як правило, з сукупності потенційних незалежних змінних обираються лише ті змінні, коефіцієнти парної кореляції з якими перевищують 0,7, але в окремих випадках статистично ефективним виявляється включення до аналізу також змінних з коефіцієнтами кореляції близько 0,5 і вище. При цьому можуть застосовуватися також математичні перетворення змінних, наприклад операції логарифмування, що може покращити статистичний взаємозв'язок розподілів, які аналізуються. Далі зі сформованої сукупності зовнішніх змінних з використанням покрокової процедури включення змінних (або їх виключення) формується оптимальний набір зовнішніх змінних $x_k^{(j)}, \dots, x_{Mk}^{(j)}$. Доцільно враховувати, що кількість незалежних змінних, яку можна використати у моделі, є обмеженою. Наприклад, якщо мова йде про 27 регіонів України, то в моделі може бути використано, як правило, не більше трьох незалежних змінних, що пояснюється особливостями оцінювання її параметрів [19].

Інформаційною базою для підвищення надійності оцінювання показників бідності для регіонів України на основі непрямих методів можуть бути, насамперед, ті джерела, які дають можливість оцінити показники, що статистично тісно

пов'язані з показниками рівня життя населення. Такими джерелами інформації є, наприклад, державні вибіркові обстеження домогосподарств (населення) і, перш за все, безпосередньо дані ОУЖД [3]. Іншими потенційними джерелами є дані статистики праці, соціальної статистики, національних рахунків, переписів населення, адміністративні дані.

При проведенні досліджень зовнішньої інформації для підвищення надійності оцінювання показників бідності виконано аналіз низки показників, що розраховуються для регіонального рівня (табл. 2.1). При цьому окремі показники розглядалися лише на першому етапі досліджень, який було реалізовано у 2008–2009 роках, а показники, потенційно корисні для цілей дослідження, аналізувалися практично до кінця 2012 року.

Цінність даних ОУЖД обумовлена їх методологічною узгодженістю, що забезпечує порівнянність показників за різні періоди часу та по різних територіях.

Як свідчать результати виконаних досліджень, для уточнення показників бідності на цьому етапі розробки та впровадження статистико-математичних моделей для надійного оцінювання показників бідності, насамперед доцільно розглядати прямі оцінки цих показників за попередні періоди обстеження, а також прямі оцінки показників бідності за результатами поточного обстеження для територій, що включають регіони (наприклад, це можуть бути оцінки для національного рівня та рівня економічних районів).

Водночас слід зазначити, що використання лише показників ОУЖД для уточнення показників бідності характеризується їх окремими недоліками. Основний недолік полягає у тому, що ці показники визначаються з певною похибкою вибірки, а у випадку зміщеності оцінок показників ОУЖД (наприклад, у наслідок недостатнього охоплення заможних верств населення) прямі оцінки показників бідності будуть

також зміщеними і, відповідно, ступінь реального підвищення надійності їх оцінювання на основі непрямих методів може бути в цілому незначною.

Таблиця 2.1

Перелік потенційних показників з зовнішніх джерел, що можуть бути використані для непрямого оцінювання показників бідності

№ з/п	Показник	Джерело
1.	Рівень бідності за національною межею	Держстат, ОУЖД
2.	Рівень бідності за регіональною межею	Держстат, ОУЖД
3.	Питома вага населення з доходами, нижчими за розмір прожиткового мінімуму	Держстат, ОУЖД
4.	Середньодушові середньомісячні грошові доходи	Держстат, ОУЖД
5.	Середньодушові середньомісячні грошові витрати	Держстат, ОУЖД
6.	Середньодушові середньомісячні загальні доходи	Держстат, ОУЖД
7.	Середньодушові середньомісячні сукупні витрати	Держстат, ОУЖД
8.	Рівень безробіття	Держстат, ОЕАН
9.	Рівень зайнятості	Держстат, ОУЖД
10.	Рівень зареєстрованого безробіття	ДЦЗ
11.	Частка працівників, що отримують заробітну плату нижче мінімального рівня	Держстат
12.	Середні розміри призначених пенсій з цільовою грошовою допомогою	Держстат
13.	Середньомісячна заробітна плата фізичних осіб за місцем роботи	ДПА
14.	Середньомісячна заробітна плата фізичних осіб за місцем проживання	ДПА
15.	Середньомісячна заробітна плата штатних працівників за даними поточного обстеження підприємств з питань статистики праці	Держстат
16.	Середньомісячна заробітна плата осіб-членів домогосподарств	Держстат
17.	Наявний дохід у розрахунку на одну особу	Держстат
18.	Валовий регіональний продукт на душу населення	Держстат
19.	Кінцеві споживчі витрати домогосподарств	Держстат

Інший суттєвий недолік полягає у тому, що оцінки показників бідності та оцінки інших показників ОУЖД не є статистично незалежними, що зумовлено використанням фактично однієї і тієї самої інформаційної бази для розрахунків. Таким чином, доцільним є використання поряд із зазначеними даними також даних, отриманих за результатами вибірових обстежень, не пов'язаних з ОУЖД, даних статистичної звітності, даних адміністративних реєстрів тощо.

Як свідчать результати аналізу щільності взаємозв'язку показників державного вибірового обстеження економічної активності населення (далі – ОЕАН) [22] та показників бідності, вона є недостатньо високою. Так, для розподілів за регіонами визначених показників бідності та рівня безробіття за даними ОЕАН коефіцієнт кореляції не перевищує 0,5. Коефіцієнт кореляції зазначених показників бідності з рівнем зайнятості складає від $-0,3$ до $-0,4$. Виходячи з отриманих результатів, слід констатувати, що використання показників ОЕАН, які також характеризуються певною похибкою вибірки, як зовнішніх змінних при розробці моделей для уточнення показників бідності є недоцільним.

Коефіцієнт кореляції показників бідності для регіонів з рівнем зареєстрованого безробіття складає приблизно від 0,3 до 0,5, що в цілому не свідчить про наявність щільного взаємозв'язку. Оскільки рівень зареєстрованого безробіття є адміністративною інформацією, то дані вважаються абсолютно надійними в тому сенсі, що похибка вибірки оцінювання показників приймається рівною нулю. Систематичні похибки (можливі зміщення значень показників), пов'язані з методологією вимірювання показників зареєстрованого безробіття, похибками обробки даних тощо, можуть не братися до уваги (з певним наближенням), оскільки вони не впливають на коефіцієнти кореляції, що характеризують рівень взаємозв'язку між

розподілами показників. Це зауваження стосується всіх джерел адміністративної інформації або інформації, що отримана на основі статистичної звітності.

При проведенні досліджень було отримано інформацію Державної податкової адміністрації України (далі – ДПА) щодо сум доходів у вигляді заробітної плати, інших виплат та винагород, виплачених фізичним особам за місцем проживання та місцем роботи у 2009 році. Відповідно виникла можливість визначення придатності цієї інформації для використання при побудові моделей для непрямого оцінювання показників бідності. Для повноти аналізу одночасно з даними ДПА використовувалися дані щодо заробітної плати, отримані з обстеження підприємств та ОУЖД, що проводяться Держстатом. Таким чином, аналіз здійснювався за чотирма різними показниками середньомісячної заробітної плати:

- середньомісячні суми доходів у вигляді заробітної плати, інших виплат та винагород, виплачених фізичним особам за місцем роботи у 2009 році за даними ДПА (далі – середньомісячна заробітна плата фізичних осіб за місцем роботи);

- середньомісячні суми доходів у вигляді заробітної плати, інших виплат та винагород, виплачених фізичним особам за місцем проживання у 2009 році за даними ДПА (далі – середньомісячна заробітна плата фізичних осіб за місцем проживання);

- середньомісячна заробітна плата штатних працівників (поточне обстеження підприємств з питань статистики праці 2009 р.);

- середньомісячна заробітна плата осіб – членів домогосподарств (ОУЖД, 2009 р.).

Як свідчать результати аналізу, в цілому більш щільний взаємозв'язок за регіонами України спостерігається між

показниками бідності та заробітною платою фізичних осіб за місцем роботи (табл. 2.2).

Таблиця 2.2

Коефіцієнти кореляції між розподілами за регіонами показників бідності та показників середньомісячної заробітної плати, отриманими за даними ДПА, поточного обстеження підприємств з питань статистики праці та ОУЖД, 2009 р.

Показники заробітної плати	Рівень бідності за національною межею	Рівень бідності за регіональною межею	Рівень бідності за прожитковим мінімумом
Середньомісячна заробітна плата фізичних осіб за місцем роботи (ДПА)	-0,502	-0,436	-0,472
Середньомісячна заробітна плата фізичних осіб за місцем проживання (ДПА)	-0,497	-0,422	-0,453
Середньомісячна заробітна плата штатних працівників за даними поточного обстеження підприємств з питань статистики праці (Держкомстат)	-0,485	-0,409	-0,472
Середньомісячна заробітна плата осіб – членів домогосподарств (ОУЖД)	-0,553	-0,440	-0,457

Найбільшим значенням коефіцієнта кореляції Пірсона ρ характеризується взаємозв'язок між рівнем бідності за національною межею та середньомісячною заробітною платою фізичних осіб за місцем роботи ($\rho = -0,502$), для частки населення з доходами нижче прожиткового мінімуму $\rho = -0,472$, та для рівня бідності за регіональною межею він є найменшим: $\rho = -0,436$.

Щільність взаємозв'язку між показниками бідності та середньомісячною заробітною платою фізичних осіб за місцем проживання є дещо нижчою. Так, коефіцієнт кореляції цього показника з рівнем бідності за національною межею дорівнює $-0,497$, з рівнем бідності за регіональною межею – $(-0,422)$, для частки населення з доходами нижче прожиткового мінімуму – $(-0,453)$.

Щільність лінійного взаємозв'язку середньомісячної заробітної плати штатних працівників та показників бідності є також відносно невисокою (див. табл. 2.2). Так, для рівня бідності за національною межею та заробітної плати штатних працівників за даними 2009 року коефіцієнт кореляції складає близько $-0,485$. Для частки населення з доходами нижче прожиткового мінімуму він становить $(-0,472)$. Для рівня бідності за регіональною межею коефіцієнт кореляції є дещо меншим та дорівнює $(-0,409)$.

Доцільно навести також дані, що ілюструють щільність зв'язку між показниками бідності та середньомісячною заробітною платою осіб – членів домогосподарств, які взяли участь в ОУЖД. Цей показник розраховувався на основі річних даних ОУЖД як середньомісячна сума заробітної плати в грошовій та натуральній формах за основним місцем роботи та на інших роботах, дивідендів, компенсаційних виплат та ін. Як свідчать результати розрахунків, для рівня бідності за національною межею та заробітної плати осіб – членів домогосподарств за даними 2009 року коефіцієнт кореляції складає близько $-0,553$.

Для частки населення з доходами нижче прожиткового мінімуму він дорівнює $-0,457$. Для рівня бідності за регіональною межею коефіцієнт кореляції становить близько $-0,440$.

За отриманими результатами слід констатувати, що взаємозв'язок між визначеними показниками бідності та оцінками середньомісячної заробітної плати працівників є в

цілому недостатньо щільним – значення коефіцієнта кореляції Пірсона не перевищує величини 0,553 за модулем (і при цьому таке значення відповідає даним ОУЖД). Таким чином, ці показники не є кращими кандидатами для використання як зовнішньої інформації при оцінюванні показників бідності за моделями.

Необхідно зауважити, що характеристики розподілу за регіонами розглянутих показників середньомісячної заробітної плати, незалежно від джерела та методології отримання показників, є досить схожими. Це виражається у щільному лінійному взаємозв'язку показників заробітної плати (табл. 2.3 та рис. 2.1).

Таблиця 2.3

Коефіцієнти кореляції між розподілом середньомісячної заробітної плати осіб – членів домогосподарств за 2009 рік та середньомісячної заробітної плати, отриманої на основі даних статистичної звітності та з адміністративних джерел

Показники заробітної плати	Середньомісячна заробітна плата штатних працівників (Держкомстат)	Середньомісячна заробітна плата фізичних осіб за місцем роботи (ДПА)	Середньомісячна заробітна плата фізичних осіб за місцем проживання (ДПА)	Середньомісячна заробітна плата осіб – членів домогосподарств (ОУЖД)
Середньомісячна заробітна плата штатних працівників (Держкомстат)	1	0,996	0,992	0,891
Середньомісячна заробітна плата фізичних осіб за місцем роботи (ДПА)		1	0,988	0,884
Середньомісячна заробітна плата фізичних осіб за місцем проживання (ДПА)			1	0,906
Середньомісячна заробітна плата осіб – членів домогосподарств (ОУЖД)				1

Отже, як свідчать отримані дані, результати поточного обстеження підприємств з питань статистики праці є достатньо статистично близькими до даних ДПА і можуть використовуватись в аналізі як наближення останніх. За результатами аналізу взаємозв'язку результатів поточного обстеження підприємств та показників бідності за 2007, 2008, 2010, 2011 роки встановлено також, що щільність взаємозв'язку та співвідношення коефіцієнтів кореляції для різних показників бідності є приблизно такими самими, як і у 2009 році.

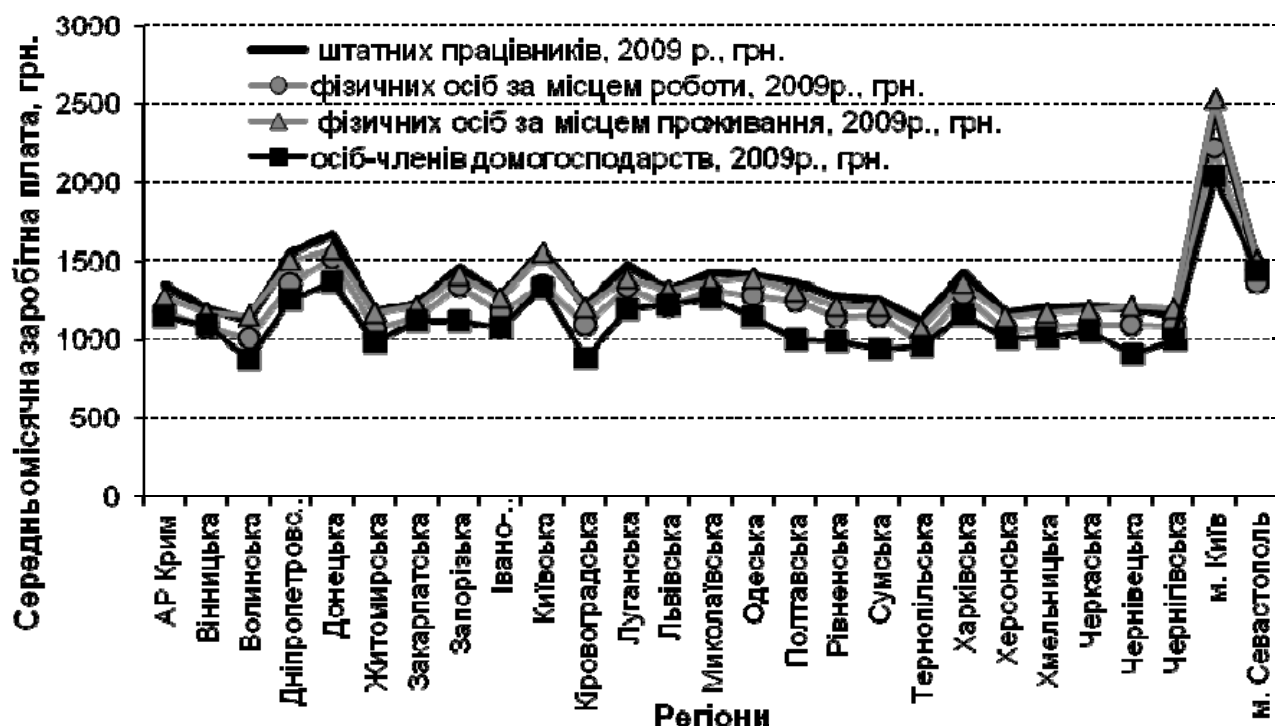


Рис. 2.1. Розподіл оцінок середньомісячної заробітної плати на основі даних з різних джерел: поточного обстеження підприємств з питань статистики праці щодо заробітної плати штатних працівників*; даних ДПА щодо доходів фізичних осіб за місцем роботи та місцем проживання*; даних ОУЖД щодо доходів осіб – членів домогосподарств

(*для цілей аналізу ці показники зменшено на 21%, оскільки, на відміну від даних ОУЖД, вони визначаються без урахування відповідних податків та зборів)

Як видно з наведених у табл. 2.4 даних, взаємозв'язок рівнів бідності за суміжні роки, наприклад 2008–2009 рр., 2009–2010 рр. і т.д., характеризується достатньо високими значен-

нями коефіцієнтів кореляції. При цьому найнижчий рівень кореляції спостерігається для рівня бідності за прожитковим мінімумом і складає для 2008–2009 рр. 0,607, найвищий – для рівня бідності за національною межею, який становить 0,808 для 2008–2009 рр.

Таблиця 2.4

Коефіцієнти кореляції між розподілами за регіонами показників бідності та показників, які потенційно можуть бути використані як зовнішня інформація при побудові моделей для непрямого оцінювання показників бідності

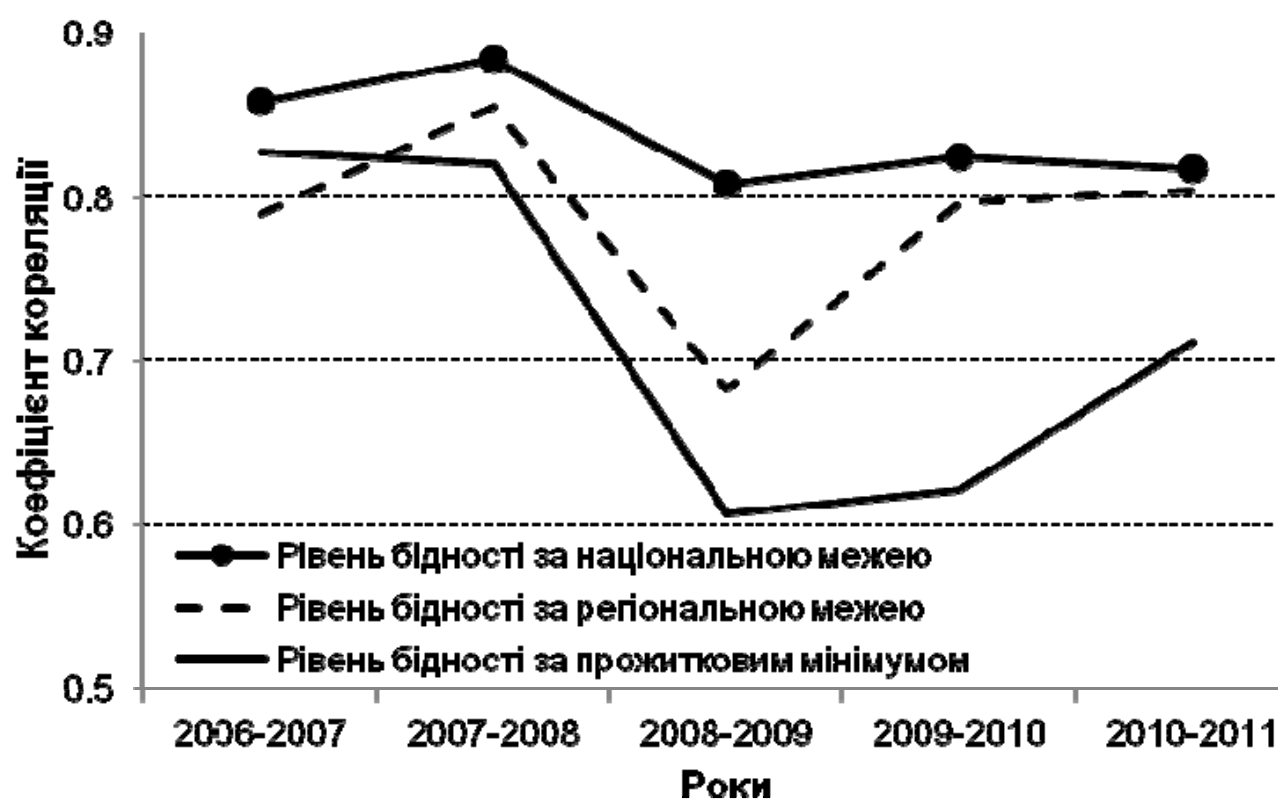
Показники	Рівень бідності за національною межею (2009 р.)	Рівень бідності за регіональною межею (2009 р.)	Рівень бідності за прожитковим мінімумом (2009 р.)
<i>Рівень бідності за національною межею (2008 р.)</i>	0,808	0,720	0,634
<i>Рівень бідності за регіональною межею (2008 р.)</i>	0,690	0,683	0,483
<i>Рівень бідності за прожитковим мінімумом (2008 р.)</i>	0,705	0,597	0,607
<i>Кінцеві споживчі витрати домогосподарств у розрахунку на одну особу (2009 р.)</i>	–0,470	–0,400	–0,436
<i>Середньомісячна заробітна плата фізичних осіб за місцем роботи (ДПА, 2009 р.)</i>	–0,502	–0,436	–0,472
<i>Наявний дохід у розрахунку на одну особу (2009 р.)</i>	–0,389	–0,361	–0,383
<i>Рівень бідності за національною межею (2010 р.)</i>	0,825	0,781	0,626
<i>Рівень бідності за регіональною межею (2010 р.)</i>	0,731	0,797	0,464
<i>Рівень бідності за прожитковим мінімумом (2010 р.)</i>	0,667	0,634	0,621

Таким чином, показники бідності за 2008 рік можуть бути обрані як зовнішня інформація при непрямому оцінюванні показників за 2009 рік, а показники за 2009 рік – при оцінюванні показників за 2010 рік. При цьому доцільно зауважити, що вибірки за суміжні роки є певною мірою незалежними оскільки, в ОУЖД здійснюється щорічна повна ротація домогосподарств (але не територіальних одиниць вибірки).

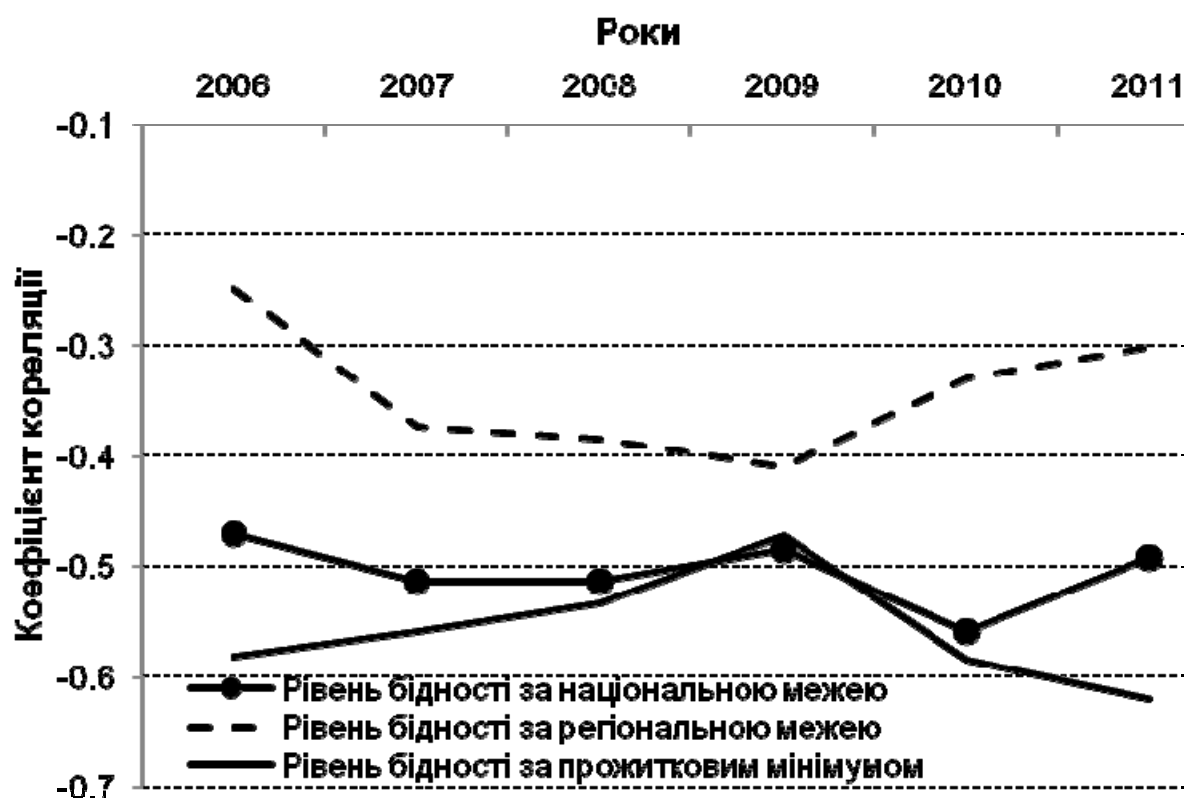
Дані, що характеризують зміну коефіцієнтів кореляції між розподілами показників бідності у часі, наведено на рис. 2.2. Як видно з представлених даних, кореляція між розподілами у 2008 та 2009 роках помітно погіршилася порівняно з попередніми роками. Це підкреслює необхідність постійного контролю рівня взаємозв'язку між зовнішніми змінними та показниками, що оцінюються на основі статистико-математичних моделей.

Усі інші розподіли з наведених у табл. 2.4, крім розподілу кінцевих споживчих витрат домогосподарств у розрахунку на одну особу (визначається за даними національних рахунків), у 2008, 2009 роках характеризуються близькими коефіцієнтами кореляції на рівні 0,4 – 0,5 за модулем. Тому при виборі потенційних змінних для моделі слід керуватися доступністю інформації, її наявністю за декілька послідовних років, рівнем її загальної якості тощо.

З урахуванням цих факторів за найбільш прийнятний варіант зовнішньої інформації на першому етапі впровадження методів непрямого оцінювання в офіційній статистиці України визначено прямі оцінки відповідних показників бідності за попередній рік та величину кінцевих споживчих витрат домогосподарств у розрахунку на одну особу за регіонами.



а)



б)

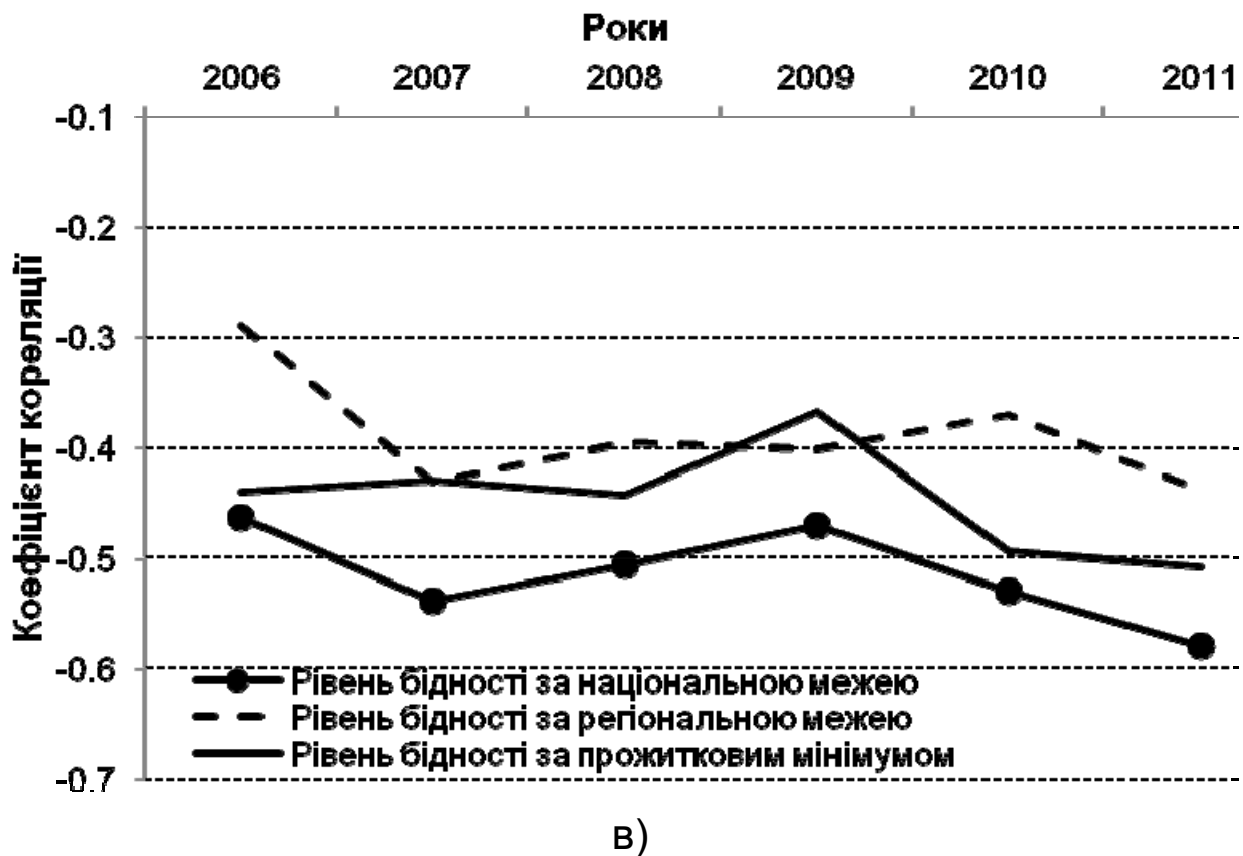


Рис. 2.2. *Зміна величини коефіцієнтів кореляції між розподілами визначених показників бідності та окремих зовнішніх показників за регіонами у часі:*

- а) між однойменними показниками бідності по суміжних роках;
- б) між показниками бідності та середньомісячною заробітною платою штатних працівників;
- в) між показниками бідності та кінцевими споживчими витратами домогосподарств у розрахунку на одну особу.

Для повноти аналізу представляється доцільним приділити певну увагу динаміці зміни коефіцієнтів кореляції між оцінками показників бідності зі збільшенням лагу (рис. 2.3). Як видно з даних, наведених на рисунку, у цілому спостерігається поступове зменшення коефіцієнтів кореляції зі збільшенням лагу. Проте у 2009 році оцінки всіх показників бідності змінилися дещо неузгоджено з загальним трендом. Це видно також із даних, наведених на рис. 2.2 а).

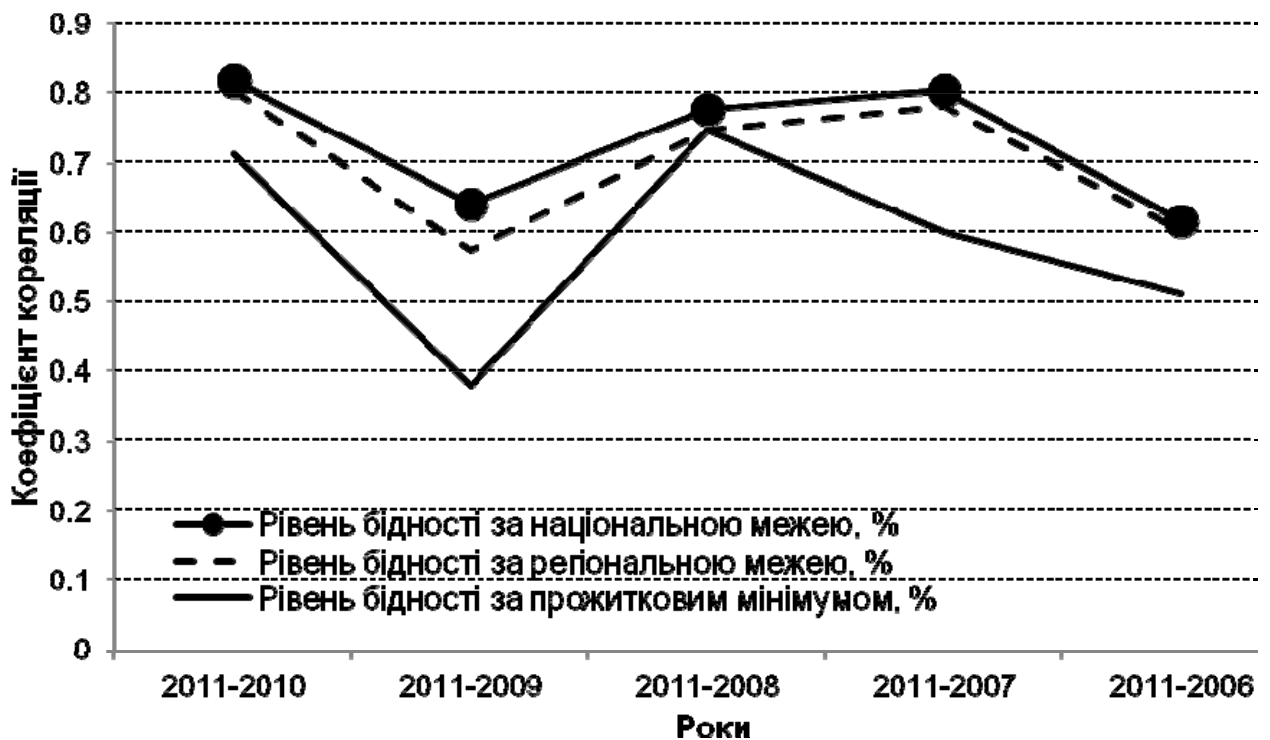


Рис. 2.3. Зміна коефіцієнтів кореляції між розподілами оцінок показників бідності за регіонами зі збільшенням лагу щодо 2011 року

Певними факторами такого «стрибка» кореляції могли бути зміна територіальної вибірки для проведення ОУЖД у 2009 – 2013 роках і недостатня надійність оцінювання показників бідності на регіональному рівні. Але тоді мала б бути відносно незначною і кореляція розподілу оцінок показників бідності за інші роки різних циклів ОУЖД (2004 – 2008рр. та 2009 – 2013 рр.), наприклад: 2011 р. з 2008 р., 2011 р. з 2007 р. і т. д. Водночас наведені на рис. 2.3 дані свідчать, що в цілому з'являється потенційна можливість використання з метою уточнення показників бідності відповідних даних не лише за суміжні роки, а і за більш відділені періоди.

2.2. Визначення зовнішньої інформації для композиційного оцінювання показників бідності

Як зазначалось у попередньому підрозділі, для вибору незалежних змінних, що можуть бути використані при непрямому оцінюванні показників бідності, насамперед аналізують парні коефіцієнти кореляції Пірсона прямих оцінок певного показника бідності за регіонами з кожною зі змінних, яка потенційно може бути використана як зовнішня інформація. За умови вибору декількох зовнішніх змінних доцільно проаналізувати відповідні регресійні моделі, що уможлиблює перевірку статистичної ефективності використання зовнішньої інформації. Як зовнішні змінні при оцінюванні, як правило, використовують ті, що відповідають найбільш адекватній регресійній моделі.

На основі сформованої сукупності зовнішніх змінних будується модель регресії й аналізується її адекватність. При цьому основними характеристиками адекватності регресійної моделі доцільно вважати коефіцієнт детермінації R^2 , який змінюється від 0 до 1 і показує частку змін результативної ознаки (дисперсії результативної ознаки), пояснюваної регресійною моделлю, і F -відношення, що відображає ступінь взаємозв'язку між результативною і факторними ознаками [19]. Чим ближчі значення R^2 до 1, тим краще зовнішні змінні здатні апроксимувати відповідні показники бідності. У загальному випадку значення R^2 , більші за 0,7, свідчать про те, що незалежні змінні достатньою мірою пояснюють варіацію відповідного показника бідності, а відповідна регресійна модель є адекватною. Варто також указати, що значення R^2 , менші за 0,7, не детермінують неможливість використання регресійної моделі.

Якщо розрахункове значення F -відношення перевищує порогове значення $R_{кр}$ (порогове значення визначається за спеціальними формулами або з довідників [19]), то це свідчить про наявність взаємозв'язку між розподілом показників бідності за регіонами та розподілами відповідних зовнішніх змінних. При побудові регресійних моделей необхідно також контролювати значення t -статистик для коефіцієнтів регресії, які мають перевищувати 2 за модулем [19].

Слід зазначити, що реалізований у цьому дослідженні методологічний підхід не передбачає використання регресійних моделей у явному вигляді, а це загалом послаблює вимоги щодо їх адекватності. Критерієм статистичної ефективності композиційної оціночної функції є, насамперед, зменшення середньоквадратичної похибки, що, у свою чергу, можливо лише за наявності статистичного взаємозв'язку між зовнішньою інформацією та показниками бідності.

На основі прямих оцінок показників бідності та зовнішніх даних доцільно обчислювати параметри відповідних регресійних моделей. Так, за умови визначення доцільності перевірки двофакторної композиційної оціночної функції (під двофакторністю тут мається на увазі використання двох зовнішніх змінних) варто виконати аналіз такої регресійної моделі (тут і далі розглядаються лише оцінки показників, тому приймається, наприклад, що $P_{regr,k}^{(j)} = \hat{P}_{regr,k}^{(j)}$):

$$P_{regr,k}^{(j)} = \beta_0^{(j)} + \beta_1^{(j)} x_{1k}^{(j)} + \beta_2^{(j)} x_{2k}^{(j)}, \quad (2.1)$$

де $\beta_0^{(j)}$, $\beta_1^{(j)}$, $\beta_2^{(j)}$ – коефіцієнти регресії, однакові для всіх регіонів України; $x_{1k}^{(j)}$ – перша зовнішня змінна (наприклад, це може бути рівень бідності у k -му регіоні у попередньому році); $x_{2k}^{(j)}$ – друга зовнішня змінна (наприклад, це може бути

величина кінцевих споживчих витрат домогосподарств у розрахунку на одну особу за даними національних рахунків для k -го регіону у поточному році).

Розглянемо приклади побудови регресійних моделей для оцінки придатності зовнішньої інформації для використання при непрямому оцінюванні показників бідності.

Оцінка рівня бідності за національною межею

При композиційному оцінюванні рівня бідності за національною межею у 2009 році регресійна модель для початкових даних, наведених у додатку А та табл. 2.5, має вигляд:

$$P_{regr,k}^{(1)} = 7,300 + 0,801 \cdot P_{k,2008}^{(1)} - 0,195 \cdot 10^{-3} \cdot x_{2k}^{(1)}, \quad (2.2)$$

де $P_{regr,k}^{(1)}$ – оцінка рівня бідності за національною межею у 2009 році за регресійною моделлю для k -го регіону; $P_{k,2008}^{(1)}$ – пряма оцінка рівня бідності за національною межею у 2008 році для k -го регіону; $x_{2k}^{(1)}$ – величина кінцевих споживчих витрат домогосподарств у розрахунку на одну особу за даними національних рахунків у 2009 році для k -го регіону.

Таблиця 2.5

Кінцеві споживчі витрати домогосподарств у розрахунку на одну особу

(грн.)

Регіон	Кінцеві споживчі витрати домогосподарств у розрахунку на одну особу	
	2008 р.	2009 р.
Україна	12592,0	12631,7
АР Крим	12253,4	12083,6
Вінницька	10451,4	10503,2
Волинська	10305,6	10377,3
Дніпропетровська	12790,9	12435,1
Донецька	12373,2	12372,43
Житомирська	9465,0	9696,93

Продовження таблиці 2.5

Закарпатська	8970,3	9328,03
Запорізька	12406,6	12776,8
Івано-Франківська	9291,3	9764,7
Київська	12327,9	12177,1
Кіровоградська	9617,9	9716,4
Луганська	9140,3	9629,2
Львівська	10515,8	10583,2
Миколаївська	11086,9	11500,8
Одеська	13609,2	14258,2
Полтавська	10805,9	11020,9
Рівненська	9088,3	9385,0
Сумська	9087,5	9211,5
Тернопільська	8169,9	8483,2
Харківська	14190,2	14563,2
Херсонська	10152,7	10261,8
Хмельницька	10155,8	10224,3
Черкаська	10650,7	10795,6
Чернівецька	9412,9	9531,1
Чернігівська	9377,0	9489,9
м. Київ	35044,7	33031,7
м. Севастополь	15738,9	15340,5

Характеристики адекватності регресійної моделі (2.2):

$$R^2 = 0,658; F = 24,00 (F_{кр} = 3,40).$$

Таким чином, регресійна модель може вважатися достатньо адекватною, оскільки $R^2 \approx 0,7$ і $F \geq F_{кр}$.

При композиційному оцінюванні рівня бідності за національною межею у 2008 році регресійна модель для аналогічної зовнішньої інформації має вигляд:

$$P_{regr,k}^{(1)} = 1,009 + 0,986 \cdot P_{k,2007}^{(1)} - 0,060 \cdot 10^{-3} \cdot x_{2k}^{(1)}, \quad (2.3)$$

де $P_{regr,k}^{(1)}$ – оцінка рівня бідності за національною межею у 2008 році за регресійною моделлю для k -го регіону; $P_{k,2007}^{(1)}$ – пряма оцінка рівня бідності за національною межею у 2007 році для k -го регіону; $x_{2k}^{(1)}$ – величина кінцевих споживчих витрат

домогосподарств у розрахунку на одну особу за даними національних рахунків у 2008 році для k -го регіону.

Характеристики адекватності регресійної моделі (2.3):

$$R^2 = 0,782; F = 44,92 (F_{кр} = 3,40).$$

Таким чином, регресійна модель може вважатися достатньо адекватною: $R^2 > 0,7$ і $F \geq F_{кр}$. Слід зазначити, що у 2009 році адекватність моделі дещо погіршилася порівняно з 2008 роком.

Якщо при композиційному оцінюванні рівня бідності за національною межею у 2009 році у початкових даних замість величини кінцевих споживчих витрат домогосподарств у розрахунку на одну особу за 2009 рік використати величину середньомісячної заробітної плати фізичних осіб за місцем роботи у цьому році (за даними ДПА), то модель матиме вигляд:

$$P_{regr,k}^{(1)} = 9,021 + 0,802 \cdot P_{k,2008}^{(1)} - 0,260 \cdot 10^{-3} \cdot x_{2k}^{(1)}, \quad (2.4)$$

де $P_{regr,k}^{(1)}$ – оцінка рівня бідності за національною межею у 2009 році за регресійною моделлю для k -го регіону; $P_{k,2008}^{(1)}$ – пряма оцінка рівня бідності за національною межею у 2008 році для k -го регіону; $x_{2k}^{(1)}$ – величина середньомісячної заробітної плати фізичних осіб за місцем роботи у 2009 році для k -го регіону.

Характеристики адекватності регресійної моделі (2.4):

$$R^2 = 0,656; F = 22,87 (F_{кр} = 3,40).$$

Таким чином, хоча коефіцієнт кореляції розподілу величини середньомісячної заробітної плати фізичних осіб за місцем роботи у 2009 році за регіонами та розподілу оцінок рівня бідності за національною межею є дещо вищим, ніж для розподілу величини кінцевих споживчих витрат домогосподарств у розрахунку на одну особу за 2009 рік (див. табл. 2.4),

регресійна модель (2.4) є певною мірою менш адекватною, ніж модель (2.2).

Оцінка рівня бідності за регіональною межею

При композиційному оцінюванні рівня бідності за регіональною межею у 2009 році регресійна модель має вигляд:

$$P_{regr,k}^{(2)} = 14,70 + 0,633 \cdot P_{k,2008}^{(2)} - 0,293 \cdot 10^{-3} \cdot x_{2k}^{(2)}, \quad (2.5)$$

де $P_{regr,k}^{(2)}$ – оцінка рівня бідності за регіональною межею у 2009 році за регресійною моделлю для k -го регіону; $P_{k,2008}^{(2)}$ – пряма оцінка рівня бідності за регіональною межею у 2008 році для k -го регіону; $x_{2k}^{(2)}$ – величина кінцевих споживчих витрат домогосподарств у розрахунку на одну особу за даними національних рахунків у 2009 році для k -го регіону.

Характеристики адекватності регресійної моделі (2.5):

$$R^2 = 0,487; \quad F = 11,85 \quad (F_{кр} = 3,40).$$

Адекватність цієї регресійної моделі є не дуже високою, оскільки $R^2 < 0,7$, але при цьому виконується умова $F \geq F_{кр}$.

Оцінка рівня бідності за прожитковим мінімумом

При композиційному оцінюванні рівня бідності за прожитковим мінімумом у 2009 році регресійна модель має вигляд:

$$P_{regr,k}^{(3)} = 3,729 + 0,457 \cdot P_{k,2008}^{(3)} - 0,095 \cdot 10^{-3} \cdot x_{2k}^{(3)}, \quad (2.6)$$

де $P_{regr,k}^{(3)}$ – оцінка рівня бідності за прожитковим мінімумом у 2009 році за регресійною моделлю для k -го регіону; $P_{k,2008}^{(3)}$ – пряма оцінка рівня бідності за прожитковим мінімумом у 2008 році для k -го регіону; $x_{2k}^{(3)}$ – величина кінцевих споживчих витрат домогосподарств у розрахунку на одну особу за даними національних рахунків у 2009 році для k -го регіону.

Характеристики адекватності регресійної моделі (2.6):

$$R^2 = 0,380; F = 7,67 (F_{кр} = 3,40).$$

Адекватність цієї регресійної моделі є також не дуже високою, оскільки $R^2 < 0,7$, але при цьому $F \geq F_{кр}$.

Слід зазначити, що модель (2.6) у цілому не забезпечує потрібну надійність оцінювання рівня бідності за прожитковим мінімумом. Це пов'язано, насамперед, з низьким рівнем бідності за цим критерієм і, відповідно, з високою відносною похибкою прямого оцінювання цього показника за результатами ОУЖД (див. додаток А).

2.3. Побудова синтетичних оцінок показників бідності на основі апроксимації розподілу доходів на мікрорівні

Одним з альтернативних підходів до побудови синтетичних оцінок показників бідності є їх визначення на основі апроксимації форми розподілу імовірностей для середньодушових еквівалентних сукупних витрат за місяць логнормальною функцією, який визначається формулою [20; 21]:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{\ln x - \overline{\ln x}}{\sigma}\right)^2}, \quad (2.7)$$

де x – середньодушові еквівалентні сукупні витрати за місяць;
 σ – середньоквадратичне відхилення логарифмів середньодушових еквівалентних сукупних витрат за місяць.

За такої апроксимації синтетична оцінка рівня бідності для регіону визначається як площа під кривою розподілу, що розташована зліва від межі бідності (рис. 2.4.). Уся площа під цією кривою дорівнює одиниці.

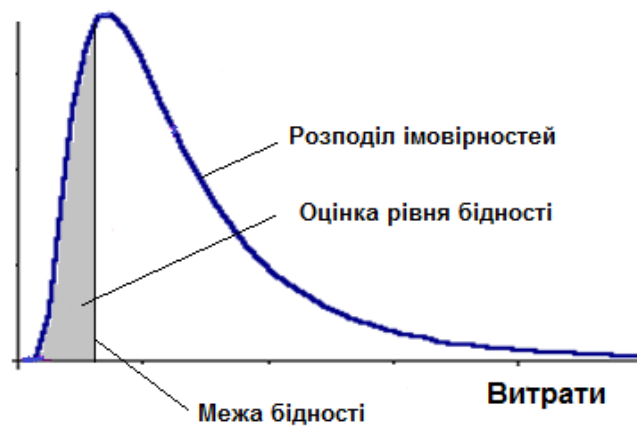


Рис. 2.4. Розподіл імовірностей для середньодушових еквівалентних сукупних витрат за місяць і межа бідності

Проблема полягає у тому, що, по-перше, емпіричний розподіл імовірностей (частот) за витратами домогосподарств, побудований за первинними даними ОУЖД, дещо відрізняється від теоретичного, і, по-друге, на регіональному рівні неможливо надійно оцінити параметри емпіричного розподілу (рис. 2.5–2.7). Останнє пояснюється, головним чином, незначним обсягом вибірки ОУЖД для регіонів.

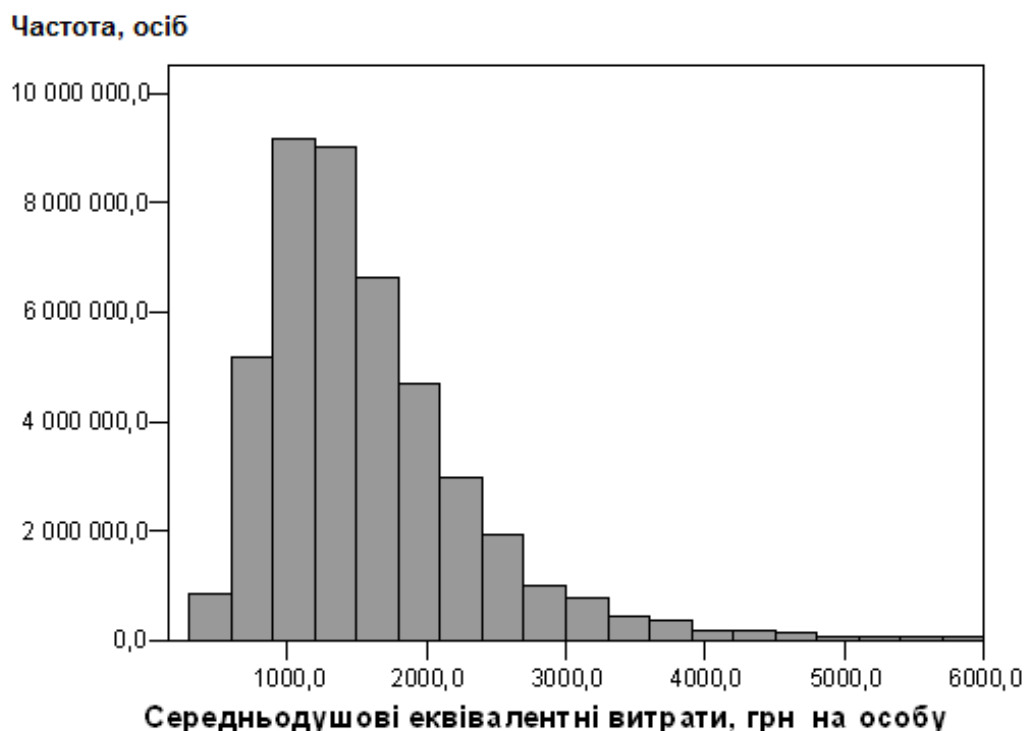


Рис. 2.5. Гістограма розподілу середньодушових еквівалентних сукупних витрат населення України (ОУЖД, 2011 р.)

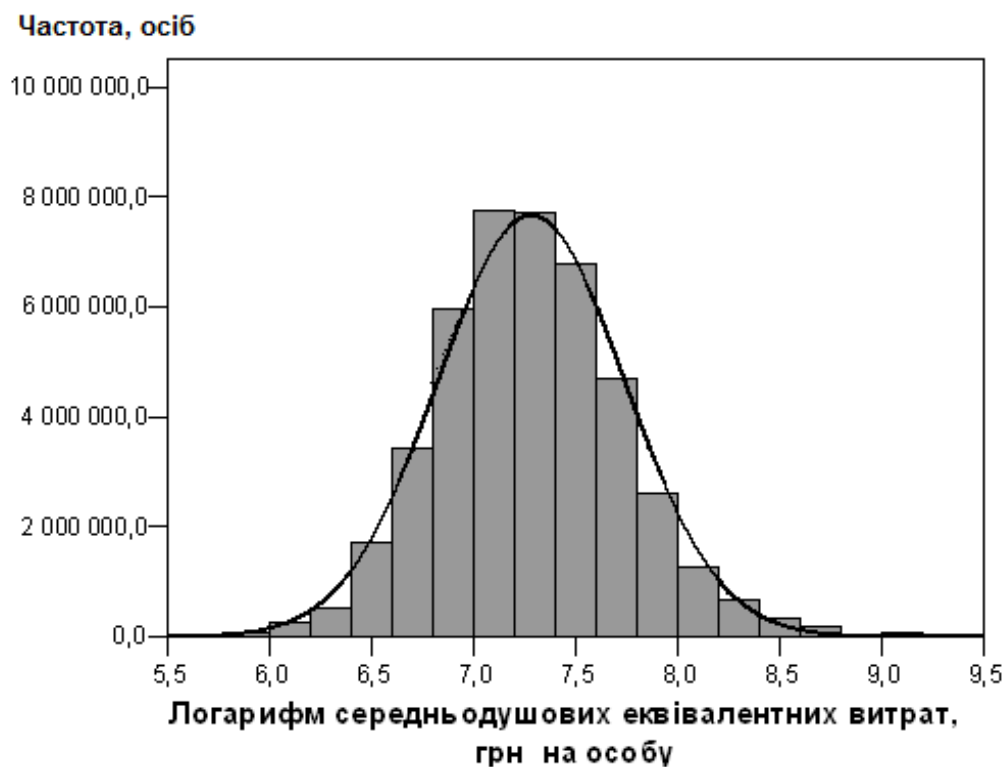


Рис. 2.6. Гістограма розподілу логарифмів середньодушових еквівалентних сукупних витрат населення України (ОУЖД, 2011 р.)

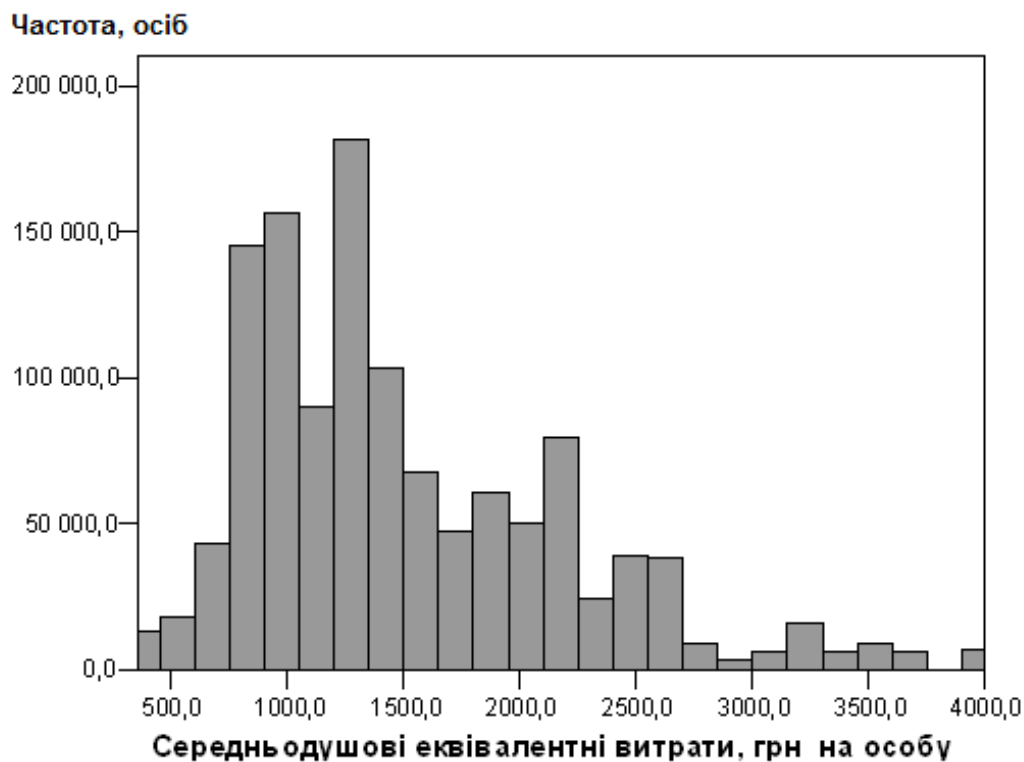


Рис. 2.7. Гістограма розподілу середньодушових еквівалентних сукупних витрат населення Закарпатської області (ОУЖД, 2011 р.)

Прийmemo, що оцінка рівня бідності для k -го регіону визначається за загальною формулою [20]:

$$\hat{P}_k^{(S)} = \frac{1}{\sigma_k \cdot \sqrt{2\pi}} \cdot \int_0^{\ln z_k} e^{-\frac{1}{2} \left(\frac{y - \overline{\ln x}}{\sigma_k} \right)^2} dy, \quad (2.8)$$

де $\overline{\ln x_k}$ – величина середнього логарифму середньодушових еквівалентних сукупних витрат за місяць по k -му регіону; $y = \ln x$, x_k – щільність імовірності для середньодушових еквівалентних сукупних витрат за місяць у k -му регіоні; z_k – межа бідності; σ_k – середньоквадратичне відхилення логарифмів середньо-душових еквівалентних сукупних витрат за місяць по k -му регіону.

Як свідчать результати досліджень [20; 22], оцінка бідності на регіональному рівні з використанням формули (2.8) може бути здійснена за такою схемою.

1) За масивом мікроданих ОУЖД поточного року розраховується змінна $y_i = \ln x_i$, де x_i – середньодушові еквівалентні сукупні витрати за місяць для i -го домогосподарства.

2) Для кожного регіону розраховується величина середнього логарифму середньодушових еквівалентних сукупних витрат за місяць $\overline{\ln x_k}$ за формулою:

$$\overline{\ln x_k} = \frac{\sum_{i=1}^{n_k} w_i \cdot m_i \cdot y_i}{\sum_{i=1}^{n_k} w_i \cdot m_i}, \quad (2.9)$$

де m_i – розмір i -го домогосподарства; w_i – статистична вага i -го домогосподарства; n_k – кількість домогосподарств, обстежених у регіоні.

3) По кожному економічному району q , до якого належить k -й регіон, за даними ОУЖД розраховується величина середньоквадратичного відхилення логарифмів середньодушових

еквівалентних сукупних витрат за місяць σ_q (з метою здійснення економічного та соціально-економічного аналізу Україна іноді поділяється на 8 економічних районів, що формуються з груп регіонів, відносно однорідних за певними характеристиками [23]).

4) По кожному регіону визначається відношення межі бідності z_k до середніх середньодушових еквівалентних сукупних витрат за місяць $\bar{x}_k$ для поточного року t за формулою:

$$k_{k(t)} = \frac{Z_{d(t)}}{\bar{x}_{d(t)}}. \quad (2.10)$$

5) По кожному регіону визначається середній за три останніх роки коефіцієнт відношення межі бідності до середніх середньодушових еквівалентних сукупних витрат за місяць за формулою:

$$\bar{k}_k = \frac{k_{k(t-2)} + k_{k(t-1)} + k_{k(t)}}{3}. \quad (2.11)$$

6) По кожному регіону визначається величина логарифму скоригованої межі бідності для поточного року $\ln(Z_{k(t)}^{\text{mod}})$ за формулою:

$$\ln(z_{k(t)}^{\text{mod}}) = \ln(\bar{k}_k \cdot \bar{x}_{k(t)}); \quad (2.12)$$

7) По кожному регіону визначається синтетична оцінка рівня бідності за формулою:

$$\hat{P}_k^{(S)} = \frac{1}{\sigma_k \cdot \sqrt{2\pi}} \cdot \int_0^{\ln(z_{k(t)}^{\text{mod}})} e^{-\frac{1}{2} \left(\frac{y_i - \overline{\ln x_k}}{\sigma_q} \right)^2} dy. \quad (2.13)$$

Похибка синтетичного оцінювання рівня бідності за співвідношенням (2.13) визначається наближено шляхом порівняння оцінок за реплікаціями, здійсненими за підвибірками територіальних одиниць на національному рівні. Тобто, наприклад,

нехай умовна національна вибірка ОУЖД складається з 800 територіальних одиниць, вибірка економічного району – в середньому з 240, а регіональна – з 80. За національною вибіркою будується логнормальний розподіл середньодушових еквівалентних сукупних витрат за місяць. Далі з національної вибірки формується достатньо велика кількість (100 і більше) вибірок (реплікацій) по 240 та по 80 територіальних одиниць, реалізується описана методика розрахунку синтетичних оцінок рівня бідності для умовного регіону й умовного економічного району і розраховується середньоквадратичне відхилення оцінок за кожною з реплікацій від оцінки за національною вибіркою, що приймається за точну. Зазначене відхилення приймається за середньоквадратичну похибку синтетичної оцінки.

Розглянутий підхід дає можливість помітно підвищити надійність оцінювання показників бідності за композиційною оціночною функцією (1.17) [20], але очевидними його недоліками є те, що уточнення здійснюється на основі лише даних ОУЖД за рахунок усереднення ефектів за певний період часу (декілька років) та використання аналітичного статистичного розподілу.

РОЗДІЛ 3

ПРИКЛАДНІ АСПЕКТИ ПОБУДОВИ СТАТИСТИКО-МАТЕМАТИЧНИХ МОДЕЛЕЙ ДЛЯ НЕПРЯМОГО ОЦІНЮВАННЯ ПОКАЗНИКІВ БІДНОСТІ

3.1. Аналіз результатів використання багатофакторних композиційних оціночних функцій для підвищення надійності оцінювання визначених показників бідності

Загальний вид багатофакторної оціночної функції, що визначена як найбільш ефективна статистична модель для підвищення надійності оцінювання показників бідності, представлено у розділі 1, формула (1.18). У ході виконаних досліджень апробовані однофакторні, двофакторні та трифакторні композиційні оціночні функції. Для розрахунків уточнених оцінок визначених показників бідності доцільно застосовувати трифакторні оціночні функції. Це зумовлено тим, що для таких функцій обсяги додаткових даних, які використовуються при оцінюванні, є найбільшими і, відповідно, слід очікувати зменшення величини зміщення оцінок. Водночас зазначимо, що за існуючих умов ступінь підвищення надійності оцінок за трифакторними функціями порівняно з двофакторними є загалом незначною. Цього слід було очікувати, з огляду на те, що розподіл за регіонами величини кінцевих споживчих витрат домогосподарств у розрахунку на одну особу за даними національних рахунків характеризується помітно слабшим

кореляційним взаємозв'язком з розподілом показників бідності, ніж їх розподіли за попередній рік.

Наведемо оціночні функції для кожного визначеного показника бідності.

Трифакторна композиційна оціночна функція для підвищення надійності оцінювання рівня бідності за національною межею

$$P_{comp,k}^{(1)} = (1 - \hat{\gamma}_{1k}^{(1)}) \cdot P_k^{(1)} - \hat{\gamma}_{2k}^{(1)} \cdot x_{1k}^{(1)} - \hat{\gamma}_{3k}^{(1)} \cdot x_{2k}^{(1)} + \hat{\gamma}_{1k}^{(1)} \cdot P^{(1)}, \quad (3.1)$$

де $P_{comp,k}^{(1)}$ – композиційна оцінка рівня бідності за національною межею: для k -го регіону у році, що розглядається; $P_k^{(1)}$ – пряма оцінка рівня бідності за національною межею для k -го регіону у році, що розглядається; $P^{(1)}$ – пряма оцінка рівня бідності за національною межею для національного рівня у році, що розглядається; $x_{1k}^{(1)}$ – пряма оцінка рівня бідності за національною межею для k -го регіону у попередньому році; $x_{2k}^{(1)}$ – величина кінцевих споживчих витрат домогосподарств у розрахунку на одну особу за даними національних рахунків у році, що розглядається.

Трифакторна композиційна оціночна функція для підвищення надійності оцінювання рівня бідності за регіональною межею

$$P_{comp,k}^{(2)} = (1 - \hat{\gamma}_{1k}^{(2)}) \cdot P_k^{(2)} - \hat{\gamma}_{2k}^{(2)} \cdot x_{1k}^{(2)} - \hat{\gamma}_{3k}^{(2)} \cdot x_{2k}^{(2)} + \hat{\gamma}_{1k}^{(2)} \cdot P^{(2)}, \quad (3.2)$$

де $P_{comp,k}^{(2)}$ – композиційна оцінка рівня бідності за регіональною межею для k -го регіону у році, що розглядається; $P_k^{(2)}$ – пряма

оцінка рівня бідності за регіональною межею для k -го регіону у році, що розглядається; $P^{(2)}$ – пряма оцінка рівня бідності за регіональною межею для національного рівня у році, що розглядається; $x_{1k}^{(2)}$ – пряма оцінка рівня бідності за регіональною межею для k -го регіону у попередньому році; $x_{2k}^{(j)}$ – величина кінцевих споживчих витрат домогосподарств у розрахунку на одну особу за даними національних рахунків у році, що розглядається.

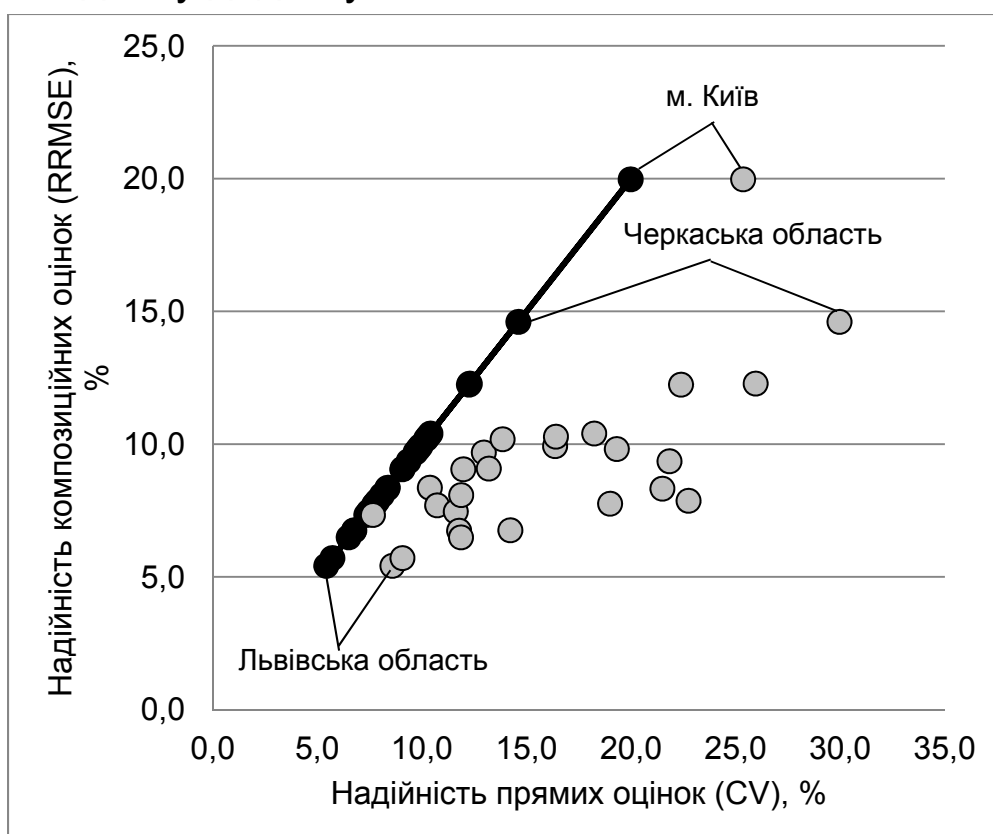
Трифакторна композиційна оціночна функція для підвищення надійності оцінювання рівня бідності за прожитковим мінімумом:

$$P_{comp,k}^{(3)} = (1 - \hat{\gamma}_{1k}^{(3)}) \cdot P_k^{(3)} - \hat{\gamma}_{2k}^{(3)} \cdot x_{1k}^{(3)} - \hat{\gamma}_{3k}^{(3)} \cdot x_{2k}^{(3)} + \hat{\gamma}_{1k}^{(3)} \cdot P^{(3)}, \quad (3.3)$$

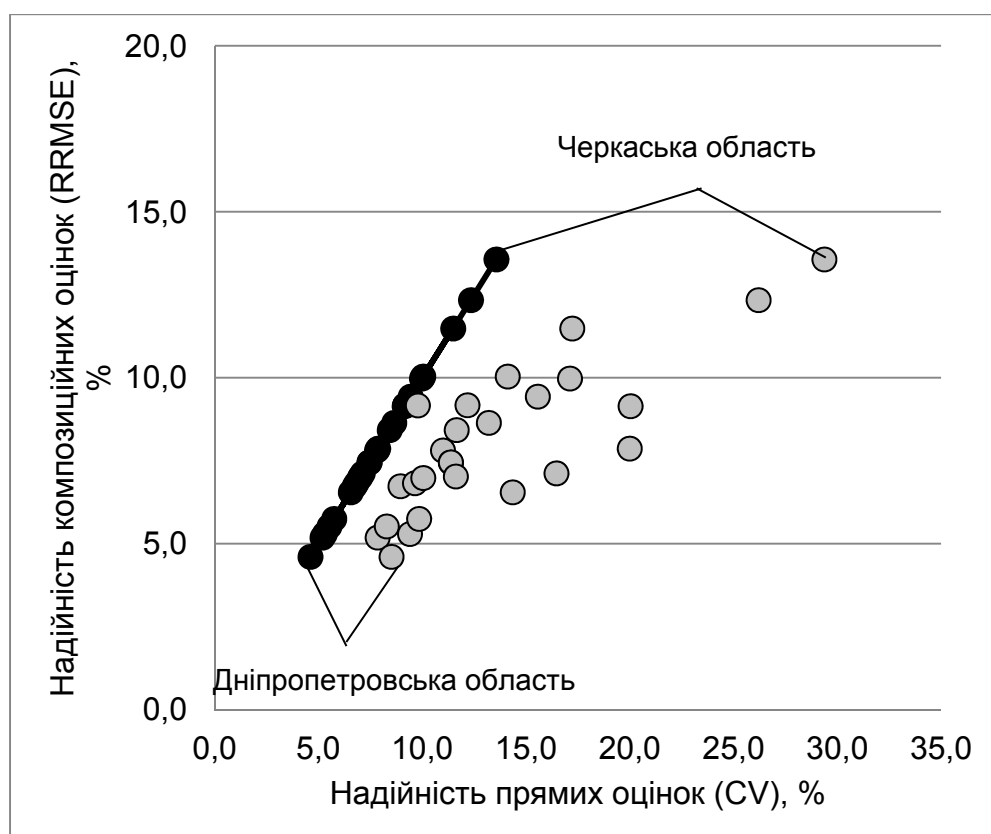
де $P_{comp,k}^{(3)}$ – композиційна оцінка рівня бідності за прожитковим мінімумом для k -го регіону у році, що розглядається; $P_k^{(3)}$ – пряма оцінка рівня бідності за прожитковим мінімумом для k -го регіону у році, що розглядається; $P^{(3)}$ – пряма оцінка рівня бідності за прожитковим мінімумом для національного рівня у році, що розглядається; $x_{1k}^{(3)}$ – пряма оцінка рівня бідності за прожитковим мінімумом для k -го регіону у попередньому році; $x_{2k}^{(3)}$ – величина кінцевих споживчих витрат домогосподарств у розрахунку на одну особу за даними національних рахунків у році, що розглядається.

Таким чином, моделі для трьох визначених показників бідності є однаковими за формою. Різними є лише інформація, що використовується для оцінювання, та параметри моделі – система вагових коефіцієнтів $\hat{\gamma}_{1k}^{(j)}$, $\hat{\gamma}_{2k}^{(j)}$, $\hat{\gamma}_{3k}^{(j)}$.

Основні результати розрахунків проілюстровані на рис. 3.1 та подані у додатку Б.



а)



б)

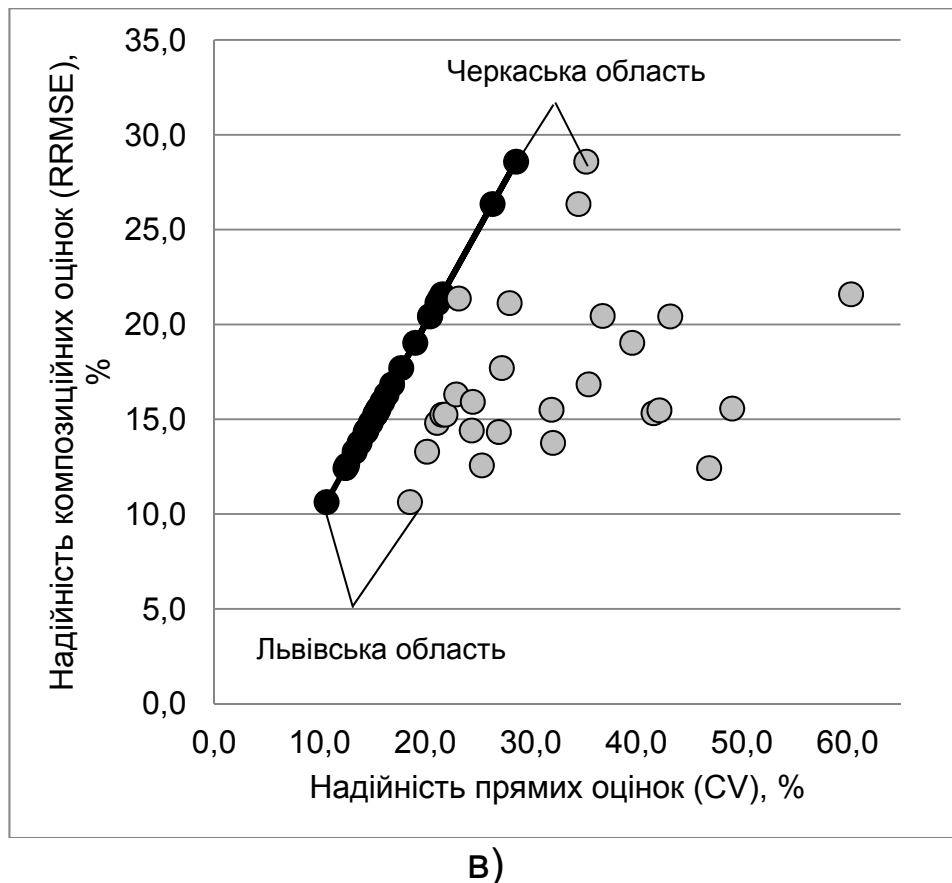


Рис. 3.1. Порівняння відносних похибок оцінювання за прямими та трифакторними композиційними оціночними функціями для визначених показників бідності:

- а) рівень бідності за національною межею;
- б) рівень бідності за регіональною межею;
- в) рівень бідності за прожитковим мінімумом

Як видно з наведених даних, в окремих випадках (для окремих регіонів) може спостерігатися незначне зменшення похибки оцінювання визначених показників бідності за композиційною оціночною функцією порівняно з прямими оцінками. Слід зазначити, що відносні похибки композиційного оцінювання показників можуть навіть збільшуватися порівняно з похибками прямого оцінювання у разі, якщо оцінка показника знижується більшою мірою, ніж абсолютна похибка його оцінювання.

Найбільші ефекти покращання надійності застосування композиційних оціночних функцій досягаються для регіонів, в

яких значення синтетичної оцінки показника і прямої оцінки є достатньо близькими.

Як уже зазначалося, параметри моделей для оцінювання показників бідності визначаються з умови мінімізації середньоквадратичної похибки композиційної оціночної функції на основі співвідношення (1.19). Оскільки це співвідношення містить об'єкти (варіаційні та коваріаційні матриці й вектори), оцінені за результатами вибіркового обстеження, для визначення середньоквадратичних похибок оцінювання застосовується реплікаційний підхід (див. п. 1.3).

Для реалізації реплікацій розроблено спеціальне програмне забезпечення (див. п. 3.4) на основі пакета програм *R*, який підтримується середовищем *SPSS*. Слід зазначити, що для непрямих оцінок показників бідності доцільно перевіряти, чи потрапляють вони до довірчих інтервалів відповідних прямих оцінок. Якщо за результатами такої перевірки встановлено, що кількість викидних значень перевищує 5 або відхилення композиційної оцінки перевищує 20% відповідної межі довірчого інтервалу, варто прийняти рішення щодо неприйнятності композиційних оцінок як уточнених оцінок показників бідності.

Якщо надійність композиційних оцінок певного показника для окремих регіонів виявиться нижчою, ніж надійність відповідних прямих оцінок, користувачам все одно необхідно надавати композиційні оцінки для забезпечення порівнянності даних. Якщо кількість регіонів, за якими спостерігатиметься погіршення надійності оцінювання певного показника, перевищить 50% від загальної кількості регіонів, слід відмовитися від використання композиційних оцінок.

Як свідчать отримані дані, застосування тривимірної композиційної оціночної функції приводить до помітного покра-

щання надійності оцінювання рівнів бідності. Так для показників за 2011 рік відносні похибки оцінювання рівнів бідності за національною та регіональною межею зменшуються в середньому для всіх регіонів у 1,61 та 1,58 рази відповідно. Для рівня бідності за прожитковим мінімумом ефективність моделей є навіть більш високою – відносна похибка зменшується в середньому у 1,74 рази. Але для цього показника не виконується умова надійності оцінювання за відотною похибкою ($RRMSE(P_{comp,k}^{(j)}) \leq 10\%$). І лише для восьми областей $RRMSE(P_{comp,k}^{(j)}) \leq 15\%$ (див. табл. Б.6 у додатку Б). Це можна пояснити значно нижчим рівнем бідності за прожитковим мінімумом, ніж за іншими критеріями – по Україні у 2011 році він становив 7,59% (рівень бідності за національною межею – 24,16%, за регіональною межею – 29,64%). Відповідно, задоволення вимог надійності за відотною похибкою суттєво ускладнюється через те, що у формулі для визначення останньої величина показника знаходиться у знаменнику.

3.2. Супровід використання композиційних оціночних функцій

Представляється доцільним упорядкувати основні положення, що визначають порядок застосування на практиці процедур підвищення надійності оцінювання показників бідності з використанням композиційних оціночних функцій. Основні з цих положень розглянуті у попередніх підрозділах монографії.

Для забезпечення функціонування інструментарію підвищення надійності оцінювання показників бідності необхідно передбачити формування та аналіз інформаційної бази по мірі надходження інформації з визначених додаткових джерел.

На щорічній основі здійснюються такі заходи:

1. Аналізуються прямі оцінки кожного визначеного показника бідності за регіонами.

2. Аналізуються похибки оцінювання (стандартні похибки, граничні похибки та коефіцієнти варіації) прямих оцінок кожного визначеного показника бідності за регіонами.

3. Визначаються коефіцієнти кореляції розподілів за регіонами прямих оцінок кожного показника бідності за поточний та попередній рік. За умови зменшення значення коефіцієнта кореляції нижче прийнятного рівня приймається рішення щодо доцільності використання розподілу показників бідності за попередній рік як зовнішньої інформації у композиційній оціночній функції.

4. Визначаються коефіцієнти кореляції розподілів за регіонами прямих оцінок кожного показника бідності за поточний рік з величиною кінцевих споживчих витрат домогосподарств у розрахунку на одну особу за даними національних рахунків за попередній рік (або з будь-якою іншою змінною, визначеною як потенційна зовнішня інформація для побудови композиційної оціночної функції). За умови зменшення значення коефіцієнта кореляції нижче припустимого рівня приймається рішення щодо доцільності використання розподілу величини кінцевих споживчих витрат домогосподарств у розрахунку на одну особу за даними національних рахунків за попередній рік як зовнішньої інформації у композиційній оціночній функції.

5. На основі розподілів за регіонами прямих оцінок кожного показника бідності за поточний рік та зовнішніх змінних, придатність яких для використання у композиційних оціночних функціях підтверджено за результатами аналізу, будуються й аналізуються регресійні моделі вигляду (2.1).

6. Для оцінок показників бідності за регресійними моделями перевіряється, чи потрапляють вони до довірчих інтервалів відповідних прямих оцінок. Якщо кількість викидних значень (значень, які не потрапляють до довірчих інтервалів) перевищує 5 (приблизно 20% випадків) або відхилення оцінки за регресією перевищує 20% відповідної межі довірчого інтервалу, може бути прийняте рішення щодо неадекватності отриманої регресійної моделі та необхідності пошуку іншої моделі.

За умови підтвердження адекватності регресійної моделі розраховуються композиційні оціночні функції з використанням запропонованого або подібного інструментарію.

7. Для оцінок показників бідності за композиційними оціночними функціями перевіряється, чи потрапляють вони до довірчих інтервалів відповідних прямих оцінок. Якщо кількість викидних значень перевищує 5 (приблизно 20% випадків) або відхилення композиційної оцінки перевищує 20% відповідної межі довірчого інтервалу, може бути прийняте рішення щодо неприйнятності композиційних оцінок як уточнених оцінок показників бідності.

8. У разі, якщо надійність композиційних оцінок певного показника для окремих регіонів буде нижчою, ніж надійність прямих оцінок, користувачам слід все одно надавати композиційні оцінки для забезпечення порівнянності даних. Якщо кількість регіонів, за якими спостерігатиметься погіршення надійності оцінювання певного показника, перевищить 50% від загальної кількості регіонів (це може відбутися, зокрема, внаслідок “старіння” моделі через зміну з часом статистичних взаємозв’язків показників бідності та зовнішньої інформації), необхідно здійснити пошук більш ефективної моделі або відмовитися від використання композиційних оцінок, передбачених цією методикою.

9. Якщо виявиться, що окремі зовнішні змінні не можуть бути використані як зовнішня інформація, розроблений інструментарій уможлиблює використання двомірних (або навіть одновірних) композиційних оціночних функцій.

На постійній основі доцільно також проводити роботи щодо пошуку та аналізу зовнішньої інформації, що може бути використана при побудові моделей з метою підвищення надійності оцінювання показників бідності за регіонами.

3.3. Методологічні підходи до створення форматів подання користувачам результатів застосування методів непрямого оцінювання показників бідності

Вихідна (результуюча) інформація за результатами оцінювання визначених показників бідності для регіонів України на основі композиційних оціночних функцій призначена як для внутрішнього використання з метою перевірки поточних результатів та ефективності роботи інструментарію, так і для застосування при підготовці інформації для надання зовнішнім користувачам [24].

У статистичній практиці інших країн оприлюднення результатів застосування моделей для підвищення надійності оцінювання показників здійснюється, як правило, один раз на рік із суттєвим запізненням даних, яке може складати один-два роки щодо звітного. Це пояснюється, з одного боку, складністю розрахунків показників за моделями та оцінювання характеристик якості цих показників. З іншого боку, можливість використання моделей вимагає наявності інформації з різних джерел, яка надходить неодноразово і до якої також застосовуються спеціальні процедури перетворення, узгодження тощо. Важливим є

також той факт, що зі зміною показників та зовнішньої інформації параметри моделей можуть потребувати певного коригування.

Результати моделювання, як правило, мають надаватися користувачам у вигляді окремої публікації, до складу якої доцільно включити методологічні пояснення. Нижче наведено стислий опис змісту методологічних пояснень щодо оцінювання показників бідності з використанням методів статистичного моделювання, розроблений на основі аналізу існуючих закордонних публікацій [1].

Загальний опис проблеми

Пояснюється необхідність підвищення рівня надійності оцінювання визначених показників бідності за результатами вибіркового обстеження.

Система показників

Зазначається, які саме показники бідності розраховано, викладається методика (наводиться формула) розрахунку кожного показника.

Рівень оцінювання

Указується, для якого саме територіального рівня (за якими розрізами) проводяться розрахунки показників бідності.

Методи оцінювання показників

Вказується, на основі якого типу моделей отримано оцінки показників бідності для кожного територіального рівня. Надається характеристика зовнішньої інформації, що використовується. Обчислюються та наводяться характеристики адекватності моделей.

Підходи до оцінювання надійності показників бідності

Подаються характеристики та критерії надійності, формули розрахунку основних характеристик надійності, наведених у публікації.

Технічні зауваження

Зазначається програмне забезпечення, на основі якого отримано оцінки показників бідності.

Користування результатами

Окреслюється коло користувачів, на яке розраховані наведені в публікації результати оцінки показників бідності.

Додаткова інформація

Наводиться контактна інформація (поштова адреса, електронна адреса, веб-сторінка), за якою можна звернутися для отримання додаткової інформації щодо методології оцінювання показників бідності, результати застосування якої наведені в публікації.

Основні терміни та поняття

Надається визначення основних понять і термінів, які використовуються в публікації.

Вихідну інформацію за результатами моделювання доцільно розділити на два типи: для внутрішнього використання (у підрозділі, де застосування інструментарію) та для надання зовнішнім користувачам.

Вихідна інформація для внутрішнього використання

За результатами застосування композиційних оціночних функцій щодо кожного з показників бідності наводиться така інформація:

- прямі оцінки показника за регіонами;
- оцінки дисперсії прямих оцінок, граничних похибок та коефіцієнтів варіації за регіонами;
- синтетичні оцінки показника за регіонами;
- вектор-стовпець оцінки коефіцієнтів регресії;
- коефіцієнт детермінації;
- F -відношення;
- оцінка дисперсії синтетичної оціночної функції;
- оцінки вагових коефіцієнтів композиційної оціночної функції за регіонами;

- композиційні оцінки показника за регіонами;
- оцінки величин середньоквадратичних похибок композиційних оціночних функцій за регіонами;
- величини відношення кореня квадратного від середньоквадратичної похибки до композиційної оцінки показника бідності за регіонами;
- оцінка величини граничної похибки композиційного оцінювання за регіонами.

Інформація для внутрішнього використання необхідна для визначення ефективності оцінювання показників для регіонів на основі прийнятої композиційної оціночної функції.

Для оцінки результатів композиційного оцінювання здійснюються перевірки, викладені у попередньому підрозділі.

Вихідна інформація для надання зовнішнім користувачам

Зовнішнім користувачам інформація надається фактично у тому самому обсязі, що і за результатами прямого оцінювання. При цьому замість прямих оцінок визначених показників бідності за регіонами надаються уточнені (композиційні) оцінки, замість коефіцієнтів варіації – відповідні відносні похибки (величини відношення кореня квадратного від середньоквадратичної похибки до композиційної оцінки у відсотках). Розрахунки обов'язково супроводжуються методологічними поясненнями щодо оцінювання показників бідності з використанням методів статистико-математичного моделювання.

При підготовці інформації слід урахувати та звертати увагу користувачів, що композиційне оцінювання показників бідності за регіонами вимагає перерахунку показників рівня України та рівня економічних районів.

Результати оцінювання показників бідності на основі моделей, як і результати прямого оцінювання, доцільно представляти у вигляді таблиць, діаграм та графіків.

3.4. Програмне забезпечення для оцінювання визначених показників бідності на регіональному рівні з використанням композиційних оціночних функцій

Програмне забезпечення, в якому реалізовано розрахунки за композиційними оціночними функціями, розроблено міжнародним консультантом Н. Т. Лонгфордом (з використанням пакета програм *R*) та національним фахівцем з питань використання пакета *PASW Statistics (SPSS)* Н. Г. Шапран [25]. Розроблені також інструктивні матеріали для користувачів програмним забезпеченням.

Схеми розміщення файлів та директорій для коректної роботи програмного комплексу наведені на рис. 3.2.

Послідовність дій при композиційному оцінюванні показників бідності в автоматизованому режимі включає такі етапи:

- налаштування середовища *PASW Statistics (SPSS)* для інтеграції з мовою програмування *R*;
- налаштування середовища *PASW Statistics* для оптимальної роботи в режимі «виробничого завдання»;
- встановлення на комп'ютер робочих папок програмного модуля з оцінювання показників бідності згідно зі схемою на рис. 3.2;
- формування вхідних даних;
- запуск на виконання програмного модуля оцінювання визначених показників бідності у режимі «виробничого завдання»;
- відкриття файла вихідних даних, під час якого в автоматизованому режимі здійснюється формування стандартних вихідних таблиць (додаток В).

Файл вхідних даних *C:\SAE\data_model\input_data.xls* містить електронні сторінки з даними за роками, при цьому кожен файл налічує 10 вхідних змінних (див. табл. 3.1, додаток В).

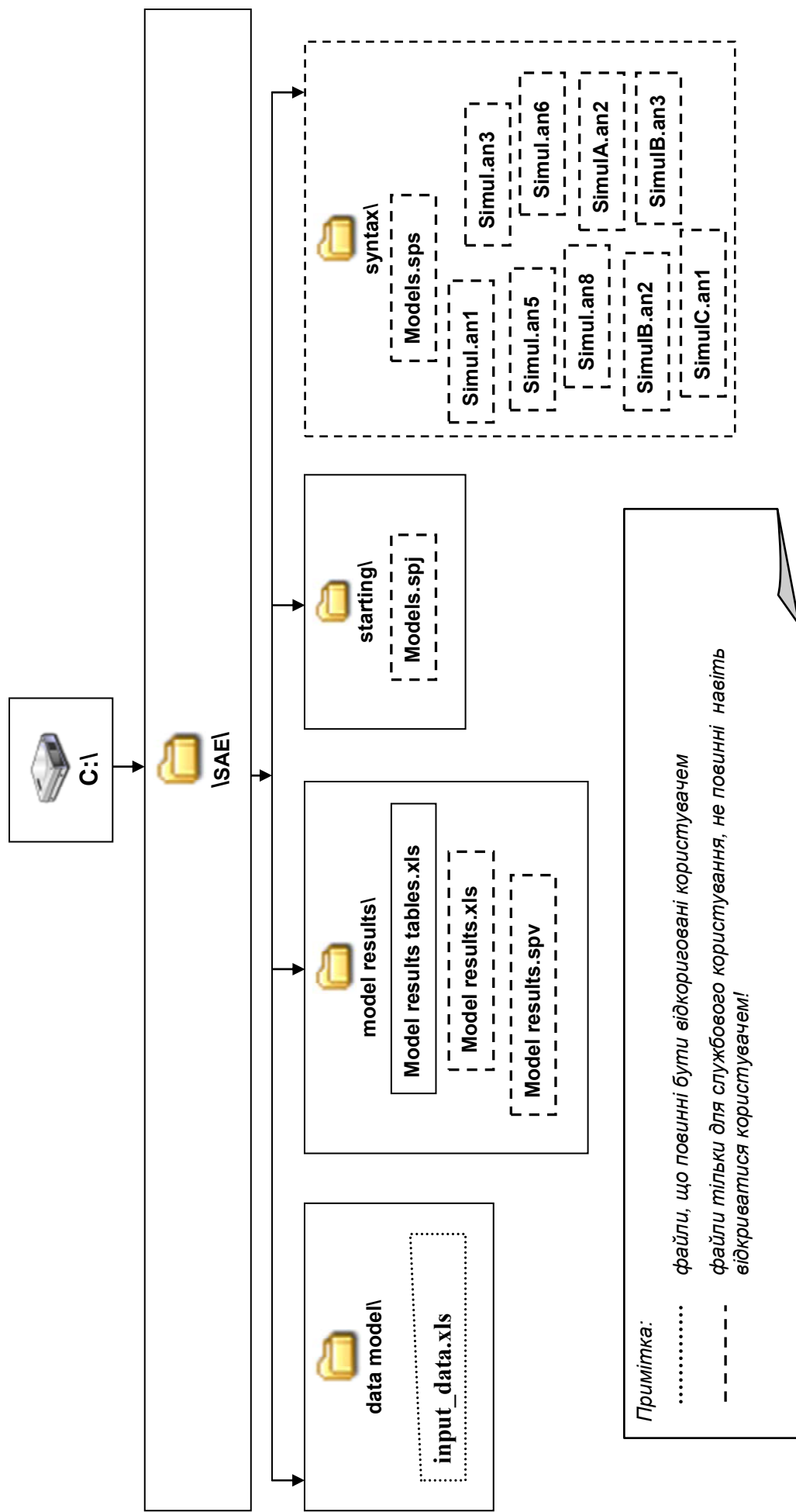


Рис. 3.2. Розміщення файлів та директорій, що забезпечує коректний запуск програмного модуля з композитного оцінювання визначених показників бідності в автоматизованому режимі

Таблиця 3.1

Змінні файла вхідних даних

№ з/п	Назва змінної	Опис змінної
1	Регіон	Назви регіонів
2	Чисельність постійного населення, осіб	—
3	Обсяг вибірки	Кількість домогосподарств, що відбираються для обстеження умов життя домогосподарств за кожним регіоном
4	Рівень бідності за національною межею $P1$, %	Відповідні значення по Україні та регіонах
5	Коефіцієнт варіації $CV(P1)$, %	
6	Рівень бідності за регіональною межею $P2$, %	
7	Коефіцієнт варіації $CV(P2)$, %	
8	Рівень бідності за прожитковим мінімумом $P3$, %	
9	Коефіцієнт варіації $CV(P3)$, %	Зовнішня змінна (у прикладах, що розглядаються, – це величина кінцевих споживчих витрат домогосподарств), необхідна для оцінювання рівня бідності за регіонами
10	Зовнішня інформація	

У режимі «виробничого завдання» запускається файл *C:\SAE\starting\Models.spj*. При цьому на екрані з'являється запит щодо характеристик моделі, а саме: рік, для якого розраховуються композиційні оцінки показника, номер показника та кількість факторів у моделі (рис. 3.3).

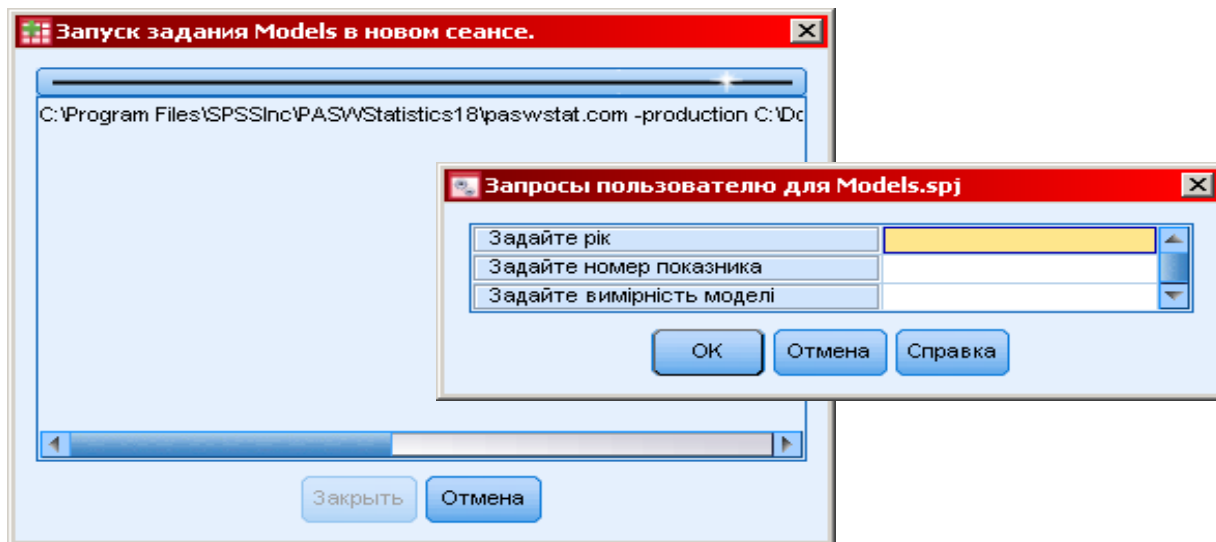


Рис. 3.3. Запит щодо характеристик моделі

Наприклад, для двофакторної моделі для показника *P2* «Рівень бідності за регіональною межею» у 2009 році характеристики моделі матимуть вигляд:

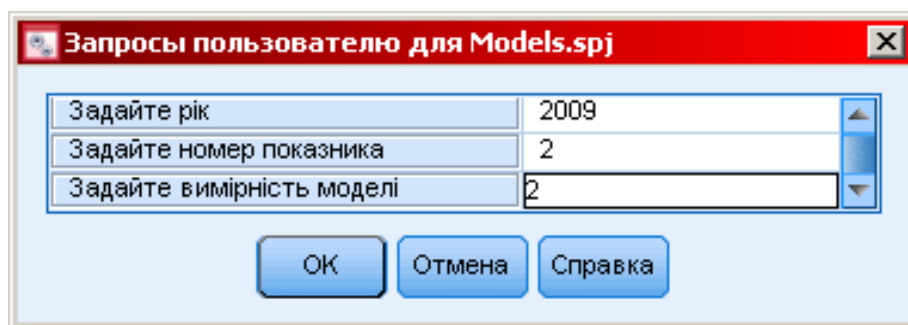


Рис. 3.4. Запит на розрахунки композиційних оцінок за двофакторною моделлю для показника *P2* «Рівень бідності за регіональною межею» у 2009 році

Залежно від характеристик комп'ютера та характеристик моделі процедура виконання завдання може тривати 30 хв. і більше.

Файл вихідних даних *C:\SAE\model results\ model results tables.xls* містить вихідні дані для внутрішніх та зовнішніх користувачів. Приклади таблиць зі вхідною інформацією та вихідних таблиць, сформованих за результатами роботи програми розрахунку композиційних оцінок показників бідності, наведені у додатку В.

РОЗДІЛ 4

ВПЛИВ МЕТОДІВ НЕПРЯМОГО ОЦІНЮВАННЯ ПОКАЗНИКІВ НА РЕЗУЛЬТАТИ ВИБІРКОВИХ ОБСТЕЖЕНЬ

4.1. Узгодження дизайну вибірки та процедур непрямого оцінювання показників

Застосування методів оцінювання показників на основі моделей, як правило, призводить до зростання актуальності оптимізації дизайну вибірки обстеження, за результатами якого вони визначаються. Це пояснюється тим, що статистична ефективність оцінювання показників може бути помітно підвищена за умови узгодження ефектів від оптимізації дизайну вибірки і застосування моделей. У цьому підрозділі розглядається методологічний підхід до оптимізації дизайну вибірки з урахуванням особливостей оцінювання показників на основі моделей. Під оптимізацією дизайну вибірки тут розуміється вибір такого дизайну серед можливих його варіантів з урахуванням наявних ресурсів, який забезпечує оцінювання показників бідності за результатами ОУЖД та з використанням зовнішньої інформації з найвищою можливою надійністю [8; 26].

Метою оптимізації дизайну вибірки ОУЖД є отримання такого розподілу фіксованого загального обсягу вибірки між регіонами України, за якого похибка оцінювання визначених показників бідності буде мінімальною. З огляду на те, що оцінювання показників здійснюється з використанням статис-

тико-математичних моделей, а саме багатофакторної композиційної оціночної функції, за критерій ефективності оціночної функції приймається величина середньоквадратичної похибки, абсолютної ($MSE(P_{comp,k}^{(j)})$) або відносної ($RMSE(P_{comp,k}^{(j)}) = \frac{MSE(P_{comp,k}^{(j)})}{(P_{comp,k}^{(j)})^2}$).

Ураховуючи зазначене обмеження, оптимізація дизайну вибірки ОУЖД в умовах застосування статистико-математичних моделей для оцінювання визначених показників бідності за регіонами забезпечується мінімізацією зваженої загальної суми MSE регіональних оцінок та дисперсії вибірових оцінок національної оцінки рівня бідності [26; 27]:

$$\sum_{j=1}^3 \sum_{k=1}^K (R_k \cdot MSE(P_{comp,k}^{(j)}) + G \cdot R_+ \cdot V(P^{(j)})) \rightarrow \min \quad (4.1)$$

з обмеженням

$$\sum_{k=1}^K n_k = n, \quad (4.2)$$

де R_k – ваговий коефіцієнт, пріоритет для k -го регіону; $MSE(P_{comp,k}^{(j)})$ – середньоквадратична похибка оцінювання показника $P_k^{(j)}$ для k -го регіону на основі моделі (як і раніше, індекс j визначає рівень бідності, а саме: рівень бідності за національною межею ($j=1$), рівень бідності за регіональною межею ($j=2$), рівень бідності за розміром прожиткового мінімуму ($j=3$)); n – загальний обсяг вибірки; k – підрядковий індекс, що відносить величину до регіону, $k = \overline{1, K}$, $K=27$; G – відносний пріоритет оцінки рівня бідності на національному рівні; $R_+ = R_1 + \dots + R_D$ – сума пріоритетів регіонів; $V(P^{(j)})$ – дисперсія оцінки рівня бідності на національному рівні.

Пріоритети R_k виражаються як [26–28]:

$$R_k = N_k^q, \quad (4.3)$$

де N_k – чисельність постійного населення k -го регіону;
 $q \in (0,2)$ – параметр пріоритету, який необхідно визначити.
 Екстремальні значення параметра $q=0$ і $q=2$ відповідають розподілу обсягів вибірки, що не враховують чисельності населення регіонів ($q=0$), та близькому до оптимального розподілу для оцінки національного рівня бідності ($q=2$).
 Значення параметра q може також визначатися ступенем, до якого припустимо зменшити надійність оцінки показника на національному рівні.

За необхідності оптимізації дизайну вибірки за відносними характеристиками надійності – відносною середньою квадратичною похибкою та відносною дисперсією – постановка задачі матиме такий вигляд:

$$\sum_{j=1}^3 \sum_{k=1}^K (R_k \cdot \frac{MSE(P_{comp,k}^{(j)})}{P_{comp,k}^{(j)2}} + G \cdot R_+ \cdot \frac{V(P^{(j)})}{P^{(j)2}}) \rightarrow \min \quad (4.4)$$

з обмеженням:

$$\sum_{k=1}^K n_k = n. \quad (4.5)$$

Розв'язання задач (4.1), (4.2) та (4.4), (4.5) здійснюється з використанням методу множників Лагранжа.

Розв'язання задачі (4.1), (4.2) може бути представлено у вигляді такої системи рівнянь [26]:

$$\begin{cases} \sum_{j=1}^3 (R_k + \frac{P_{comp,k}^{(j)}(1-P_{comp,k}^{(j)}) \cdot \sigma_{B_j}^{(j)4}}{\{P_{comp,k}^{(j)}(1-P_{comp,k}^{(j)}) + n_k \sigma_{B_j}^{(j)2}\}^2}) \\ + G \cdot R_+ \cdot \frac{P_{comp,k}^{(j)}(1-P_{comp,k}^{(j)}) N_k^2}{n_k N^2} = const; \\ n_1 + n_2 + \dots + n_K = n, \end{cases} \quad (4.6)$$

де вираз $\frac{P_{comp,k}^{(j)}(1-P_{comp,k}^{(j)}) \cdot \sigma_{B_j}^{(j)4}}{\left\{P_{comp,k}^{(j)}(1-P_{comp,k}^{(j)}) + n_k \sigma_{B_j}^{(j)2}\right\}^2}$ є наближеною оцінкою

середньоквадратичної похибки оцінювання $MSE(P_{comp,k}^{(j)})$ на основі композиційної оціночної функції показника $P_{comp,k}^{(j)}$;

$\sigma_{B_j}^{(j)2} = \frac{1}{K} \sum_{k=1}^K (P_k^{(j)} - \bar{P}^{(j)})^2$ – дисперсія регіональних оцінок j -го показника бідності;

$\bar{P}^{(j)} = \frac{1}{K} \sum_{k=1}^K P_k^{(j)}$ – середнє значення j -го показника бідності за регіонами.

Розв'язання задачі (4.4), (4.5) має аналогічний вигляд:

$$\left\{ \begin{aligned} & \sum_{j=1}^3 \sum_{k=1}^K (R_k \frac{P_{comp,k}^{(j)}(1-P_{comp,k}^{(j)}) \cdot \sigma_{B_j}^{(j)4}}{(P_{comp,k}^{(j)})^2 \cdot \left\{P_{comp,k}^{(j)}(1-P_{comp,k}^{(j)}) + n_k \sigma_{B_j}^{(j)2}\right\}^2} + \\ & + G \cdot R_+ \cdot \frac{P_{comp,k}^{(j)}(1-P_{comp,k}^{(j)}) \frac{N_k^2}{N^2}}{(P_{comp,k}^{(j)})^2 \cdot n_k} = const; \\ & n_1 + n_2 + \dots + n_K = n. \end{aligned} \right. \quad (4.7)$$

Системи рівнянь (4.6) та (4.7) розв'язуються щодо n_K з використанням ітераційних методів.

Інформаційна база для реалізації процедур розрахунку обсягу вибірки ОУЖД за регіонами України складається з таких даних:

- оцінки на основі математико-статистичних моделей визначених показників бідності за регіонами України та на національному рівні;
- величини середньоквадратичних похибок оцінок показників бідності на основі моделей;
- оцінки чисельності постійного населення за регіонами, N_K ;

- загальний обсяг вибірки ОУЖД, n ;
- величини параметрів пріоритетів q та G .

За умови прийняття за критерій надійності величини відносної похибки оцінювання – відносної середньоквадратичної похибки, обсяги вибірки за регіонами розраховуються з такої системи рівнянь:

$$\left\{ \begin{array}{l} \sum_{j=1}^3 (R_1 \frac{P_{comp,1}^{(j)}(1-P_{comp,1}^{(j)}) \cdot \sigma_{B_j}^{(j)4}}{(P_{comp,1}^{(j)})^2 \cdot \{P_{comp,1}^{(j)}(1-P_{comp,1}^{(j)}) + n_1 \sigma_{B_j}^{(j)2}\}^2} + \\ + G \cdot R_+ \cdot \frac{P_{comp,1}^{(j)}(1-P_{comp,1}^{(j)}) N_2^2}{P_{comp,1}^{(j)2} n_1^2} \frac{N_2^2}{N^2}) - \lambda = 0; \\ \sum_{j=1}^3 (R_2 \frac{P_{comp,2}^{(j)}(1-P_{comp,2}^{(j)}) \cdot \sigma_{B_j}^{(j)4}}{(P_{comp,2}^{(j)})^2 \cdot \{P_{comp,2}^{(j)}(1-P_{comp,2}^{(j)}) + n_2 \sigma_{B_j}^{(j)2}\}^2} + \\ G \cdot R_+ \cdot \frac{P_{comp,2}^{(j)}(1-P_{comp,2}^{(j)}) N_2^2}{P_{comp,2}^{(j)2} n_2^2} \frac{N_2^2}{N^2}) - \lambda = 0; \\ \dots \\ \sum_{j=1}^3 (R_K \frac{P_{comp,K}^{(j)}(1-P_{comp,K}^{(j)}) \cdot \sigma_{B_j}^{(j)4}}{(P_{comp,K}^{(j)})^2 \cdot \{P_{comp,K}^{(j)}(1-P_{comp,K}^{(j)}) + n_K \sigma_{B_j}^{(j)2}\}^2} + \\ G \cdot R_+ \cdot \frac{P_{comp,K}^{(j)}(1-P_{comp,K}^{(j)}) N_K^2}{P_{comp,K}^{(j)2} n_K^2} \frac{N_K^2}{N^2}) - \lambda = 0; \\ n_1 + n_2 + \dots + n_K = n. \end{array} \right. \quad (4.8)$$

За результатами розв'язання системи рівнянь (4.8), що складається з 28-ми рівнянь, визначаються величини $n_1, n_2, \dots, n_K$ та їх сума n . Розв'язання можна здійснити на основі ітераційних методів.

Алгоритм розрахунку може бути, зокрема, такий [26].

1. Задається певне значення обсягу вибірки для першого регіону $n_1 - \hat{n}_1$ (у реалізованому алгоритмі $\hat{n}_1$ – це існуючий обсяг вибірки ОУЖД);
2. З першого рівняння на основі значення $\hat{n}_1$ визначається значення константи $\hat{\lambda}$;
3. Із рівнянь 2 – 27 визначаються величини $\hat{n}_2, \dots, \hat{n}_K$;
4. Із двадцять восьмого рівняння визначається величина загального обсягу вибірки $\hat{n}_1 + \hat{n}_2 + \dots + \hat{n}_K = \hat{n}$;
5. Визначається різниця обсягів вибірки $\Delta = n - \hat{n}$;
6. Величина $\hat{n}_1$ коригується на основі співвідношення $\hat{n}'_1 = \hat{n}_1 + \Delta \cdot \frac{N_1}{N}$;
7. Повторюються етапи 2 – 6 доти, доки величина $\Delta \cdot \frac{N_1}{N}$ не стане менше певної наперед заданої величини, наприклад 0,5.

За умов, коли оптимальний розподіл вибірки, визначений на узгоджених засадах застосування удосконаленого дизайну вибірки обстеження умов життя домогосподарств та методів статистико-математичного моделювання, необхідно узгодити з іншим варіантом розподілу, наприклад розподілу, отриманого відповідно до стандартної методики формування вибірки ОУЖД [29], доцільно реалізувати такі етапи розрахунків.

Якщо показники ОУЖД мають вищий пріоритет, ніж показники бідності, за методикою визначаються обсяги вибірки за регіонами, які забезпечують мінімальні вимоги щодо якості оцінювання основних показників ОУЖД (наприклад, такі обсяги, що забезпечують $CV = 10\%$ для найбільш важливого показника для всіх регіонів). У випадку, якщо таким умовам задовольнятиме певний мінімальний загальний обсяг вибірки, який складає менше 80% обсягу вибірки, що фактично засто-

совується в ОУЖД, за запропонованим алгоритмом оптимального розподілу обсягу вибірки з урахуванням моделей розподіляється лише обсяг вибірки, що є різницею між фактичним та зазначеним мінімальним обсягом вибірки. Наприклад, при побудові вибірки для проведення ОУЖД у 2009–2013 роках за основний показник було визначено середню за 2004–2006 роки частку домогосподарств із середньодушовими сукупними витратами на місяць, нижчими за прожитковий мінімум. За фактичного обсягу вибірки ОУЖД у 2006 році у 10499 домогосподарств, виявилось, що для забезпечення умови $CV = 10\%$ щодо всіх регіонів достатньо було використати 4355 домогосподарств (або 41,5% від загального обсягу вибірки). Тобто для перерозподілу за запропонованим алгоритмом оптимізації дизайну вибірки залишається 6144 домогосподарства.

Якщо зазначена різниця між фактичним та мінімально необхідним обсягом вибірки складатиме менше 20% від фактичного обсягу вибірки ОУЖД, то застосування алгоритму оптимального розподілу обсягу вибірки за регіонами з урахуванням використання моделей для підвищення надійності оцінювання показників бідності не має сенсу, оскільки ефект буде дуже незначним.

За умови, що визначені показники бідності мають вищий пріоритет, ніж показники ОУЖД, обсяг вибірки розподіляється тільки за запропонованим алгоритмом.

Після розрахунків оптимальних обсягів вибірки здійснюється оцінка отриманих результатів. Це пов'язано з тим, що оптимізація є математичною процедурою і її результати можуть не задовольняти деяким критеріям, які неможливо врахувати при формалізації задач оптимізації. Основною метою оцінки й контролю результатів оптимізації дизайну вибірки є визначення

ступеня прийнятності отриманого дизайну вибірки та його погодження зі всіма фахівцями, залученими до питань організації і проведення обстеження.

Ураховуючи, що результати обстеження можуть використовуватися для оцінки низки важливих соціально-економічних показників, крім визначених показників бідності, доцільно обмежити мінімальний обсяг вибірки. Наприклад, в ОУЖД для регіонального рівня мінімально припустимий обсяг вибірки складає 300 домогосподарств [29].

Існує окрема проблема визначення обсягів вибірки для регіонів з низькими значеннями показників бідності (насамперед це стосується м. Севастополь та м. Київ), оскільки процедура оптимізації за відносною похибкою може генерувати для цих регіонів неприйнятно високі обсяги вибірки. При розв'язанні задачі оптимізації вказані регіони доцільно виключити, а обсяги вибірки для них визначати на основі експертних оцінок або виходячи зі значень, які забезпечують вимоги до оцінювання основних показників ОУЖД. Приклади застосування алгоритму оптимізації дизайну вибірки наведені у додатку Г.

У разі виявлення незадовільних результатів оптимізації внаслідок значного перерозподілу вибірки, замалого обсягу однієї зі страт тощо, слід переглянути вимоги до оптимізації і повторити зазначену послідовність етапів робіт.

4.2. Калібрація системи статистичних ваг для врахування результатів уточнення оцінок показників на мікрорівні

Представляється доцільним приділити увагу окремим підходам до вирішення проблеми узгодження оцінок показників за результатами вибіркового обстеження домогосподарств для

різних рівнів агрегації даних при використанні на певних рівнях процедур уточнення оцінок показників, зокрема на основі методів малих територій. Необхідність узгодження оцінок показників при використанні різних методів оцінювання для різних груп даних комплексних вибіркового обстежень вже відзначалася окремими науковцями, які займаються проблемами малих територій (див., наприклад, [24]). Водночас універсальних рекомендацій для цілей офіційної статистики щодо цього питання не існує. Проблема полягає в тому, що при застосуванні методів підвищення надійності оцінювання показників виникає неоднозначність в оцінках показників – вони можуть бути розраховані безпосередньо за даними обстеження (прямі оцінки) і за деякою оціночною функцією, побудованою на основі статистичних моделей спеціально для рівня малих територій.

Розглянемо окреслену проблему в цілому. Нехай на підставі вибіркового обстеження необхідно оцінити деякий показник, наприклад грошові витрати домогосподарств, для двох різних рівнів, наприклад, державного і регіонального. Позначимо ці оцінки відповідно $\hat{Y}_{(I)}$ і $\hat{Y}_{(II)k}$, де $k = 1, 2, \dots, K$; K – кількість регіонів. При цьому територіально регіони є складовими держави.

За даними вибіркового обстеження для кожної територіальної одиниці на кожному з розглянутих рівнів може бути побудована пряма оцінка $\hat{Y}^{(D)}$ і непряма (наприклад, композиційна) оцінка $\hat{Y}^{(H)}$.

Якщо для уточнення оцінок по територіальній одиниці II рівня була використана композиційна оціночна функція, то оцінка на загальнодержавному рівні, отримана з урахуванням

цього уточнення, буде відрізнятися від тієї, що відповідає прямій оціночній функції.

Один із методологічних підходів, який може бути застосований для розв'язання цієї проблеми, полягає в тому, що композиційні оцінки слід використовувати для калібрації системи статистичних ваг і вже потім одержувати оцінки для будь-яких рівнів на підставі прямих оціночних функцій.

Процедура калібрації статистичних ваг є багатоетапною [25]. На першому етапі за результатами вибіркового обстеження розраховуються прямі оцінки показників для загальнодержавного рівня. Потім перевіряється надійність цих оцінок, і якщо вона задовільна, то реалізується другий етап процедури. Слід зазначити, що на загальнодержавному рівні, як правило, вже здійснена калібрація ваг для забезпечення відповідності результатів вибіркового обстеження наявним зовнішнім даним, наприклад даним щодо демографічних характеристик населення.

На другому етапі будуються прямі оцінки показників для всіх регіонів. Якщо для регіонального рівня надійність оцінювання незадовільна, оцінки уточнюються на підставі спеціально побудованих моделей, наприклад з використанням композиційних оціночних функцій. Уточнені оцінки використовуються для калібрації статистичних ваг по цих регіонах і побудови оцінок на основі прямої оціночної функції. При цьому значення інших показників, визначені на попередніх етапах (наприклад, чисельність населення, кількість домогосподарств), можуть використовуватись також як умова (обмеження) при калібрації ваг на другому етапі.

Очевидно, що розглянутий підхід може бути застосований і для системи територіальних одиниць, що складається з більшого числа рівнів. Однак на деякому етапі може виявитися, що

процедура калібрації ваг не може бути використана через значне викривлення ваг, які відповідають дизайну вибірки.

Нижче приведено приклад застосування розглянутого підходу. Нехай на першому рівні отримана оцінка показника “Грошові доходи” для Центрального економічного району України (м. Київ, Київська та Черкаська області) $\hat{Y}_{(I)}^{(D)}$. Коефіцієнт варіації цієї оцінки є прийнятним і складає близько 5%. При цьому на основі калібрації системи статистичних ваг забезпечується узгодження оцінки кількості домогосподарств економічного району Hh за результатами вибіркового обстеження з наявними зовнішніми даними.

На другому рівні розглядаються три регіони, що складають Центральний економічний район. Оцінки кількості домогосподарств для цих регіонів становлять, відповідно, $Hh_{(II)1}^{(D)}$, $Hh_{(II)2}^{(D)}$, $Hh_{(II)3}^{(D)}$, при цьому $Hh_{(II)1}^{(D)} + Hh_{(II)2}^{(D)} + Hh_{(II)3}^{(D)} = Hh_{(I)}^{(D)} = Hh$. Оцінки показника “Грошові доходи” для них, відповідно, $\hat{Y}_{(II)1}^{(D)}$, $\hat{Y}_{(II)2}^{(D)}$, $\hat{Y}_{(II)3}^{(D)}$ і надійність оцінювання є, у цілому, незадовільною (коефіцієнт варіації перевищує 10%). Нехай оцінки $\hat{Y}_{(II)1}^{(D)}$, $\hat{Y}_{(II)2}^{(D)}$, $\hat{Y}_{(II)3}^{(D)}$ уточнюються на підставі методу композиційного оцінювання, що забезпечує надійні композиційні оцінки показників $\hat{Y}_{(II)1}^{(H)}$, $\hat{Y}_{(II)2}^{(H)}$, $\hat{Y}_{(II)3}^{(H)}$.

Далі здійснюється калібрація системи статистичних ваг у такий спосіб, щоб забезпечити оцінки $D_{(II)1}^{(D)}$, $D_{(II)2}^{(D)}$, $D_{(II)3}^{(D)}$ і $\hat{Y}_{(II)1}^{(H)}$, $\hat{Y}_{(II)2}^{(H)}$, $\hat{Y}_{(II)3}^{(H)}$. Для цього розв’язується відповідна задача мінімізації з обмеженнями:

$$\left\{ \begin{array}{l} \sum_{i=1}^n \frac{1}{2} \cdot \frac{(w_i^{(c)} - w_i^{(d)})^2}{w_i^{(d)}} \rightarrow \min \\ \sum_{i=1}^{n_1} w_i^{(c)} = Hh_{(II)1}^{(D)}, \\ \sum_{i=1}^{n_2} w_i^{(c)} = Hh_{(II)2}^{(D)}, \\ \sum_{i=1}^{n_3} w_i^{(c)} = Hh_{(II)3}^{(D)}, \\ \frac{1}{Hh_{(II)1}^{(D)}} \sum_{i=1}^{n_1} c_i \cdot w_i^{(c)} = \hat{Y}_{(II)1}^{(H)}, \\ \frac{1}{Hh_{(II)2}^{(D)}} \sum_{i=1}^{n_2} c_i \cdot w_i^{(c)} = \hat{Y}_{(II)2}^{(H)}, \\ \frac{1}{Hh_{(II)3}^{(D)}} \sum_{i=1}^{n_3} c_i \cdot w_i^{(c)} = \hat{Y}_{(II)3}^{(H)} \end{array} \right., \quad (4.9)$$

де n – обсяг вибіркової сукупності по Центральному економічному району; n_1, n_2, n_3 – обсяги вибірових сукупностей по регіонах, що складають Центральний економічний район ($n_1 + n_2 + n_3 = n$); $w_i^{(d)}$ – статистичні ваги домогосподарств до калібрації; $w_i^{(c)}$ – статистичні ваги домогосподарств після калібрації; c_i – середньомісячні грошові доходи i -го домогосподарства.

Розв'язання задачі (4.9) може бути здійснено за допомогою методу множників Лагранжа.

За результатами застосування процедури калібрації потенційно можуть бути покращені й інші оцінки показників, оскільки з'являється можливість оцінювати показники за файлом мікроданих з поліпшеною системою статистичних ваг (розрахованою з використанням більшого обсягу інформації).

ЗАКЛЮЧНІ ПОЛОЖЕННЯ

Результати досліджень, спрямованих на розробку теоретико-методологічних засад та прикладного інструментарію підвищення надійності оцінювання соціально-економічних показників за результатами вибірових обстежень, зокрема і представлені у цій публікації, свідчать про ефективність методів непрямого оцінювання. Одним із таких методів, достатньо широко апробованим у статистичній практиці багатьох країн, є багатофакторне композиційне оцінювання. Цей метод забезпечує підвищення надійності оцінювання показників на основі ефективного (оптимального в плані надійності композиційної оцінки показника) об'єднання прямої та синтетичної оціночних функцій. Синтетична оціночна функція формується на основі додаткової (зовнішньої щодо поточного вибірового обстеження та сукупності одиниць спостереження, за якими оцінюється показник) інформації – даних вибірового обстеження за попередні періоди та / або за іншими кластерами, даних інших суцільних або вибірових обстежень та / або адміністративних даних, що щільно статистично пов'язані (є корелятами) з показниками, надійність яких необхідно підвищити.

На прикладі побудови композиційних оціночних функцій для підвищення надійності оцінювання показників бідності на регіональному рівні в Україні показано, що як зовнішню інформацію доцільно використовувати оцінки цих показників на національному рівні та за попередні цикли обстеження умов життя домогосподарств, а також дані макроекономічної статистики щодо величини кінцевих споживчих витрат домогосподарств. Підхід апробовано при оцінюванні рівнів бідності за національною межею, за регіональною межею та за прожитковим мінімумом за 2008–2011 роки. Встановлено, що

застосований інструментарій композиційного оцінювання в цілому дає можливість вирішити проблему надійності на регіональному рівні при оцінюванні лише перших двох із зазначених показників. Це пояснюється як суттєво нижчою надійністю прямого оцінювання рівня бідності за прожитковим мінімумом, так і застосуванням як критерію надійності відносних величин при незначних абсолютних значеннях цього показника.

Реалізація процедур підвищення надійності оцінювання показників накладає певні обмеження на процедури оптимізації дизайну вибірки. Розглянуто результати апробації такої процедури при оцінюванні показників бідності на основі вибірки обстеження умов життя домогосподарств.

Показано, що процедури композиційного оцінювання можуть бути ефективно автоматизовані на основі популярних і поширених статистичних пакетів *SPSS (PASW Statistics)* та *R*. Наведено приклади розрахунків.

Матеріали монографії можуть використовуватись фахівцями органів державного управління, насамперед органів державної статистики, науковцями, викладачами які займаються питаннями підвищення якості інформаційного забезпечення соціальної та соціально-економічної політики.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Методологічні положення для оцінки показників бідності на регіональному рівні із застосуванням методів статистико-математичного моделювання) : затверджені Наказом Держкомстату від 01.10.2010 р. № 408 [Електронний ресурс]. – Режим доступу :

http://ukrstat.org/uk/metod_polog/metod_doc/sod/sod_01.pdf

2. Методика побудови статистико-математичної моделі для підвищення надійності оцінок показників бідності на регіональному рівні) : затверджена Наказом Держкомстату від 23.12.2011 р. № 390 [Електронний ресурс]. – Режим доступу :

http://ukrstat.org/uk/metaopus/2013/0134001_2013.htm –

Назва з титул. екрана.

3. Витрати і ресурси домогосподарств України у 2002 році (за даними вибіркового обстеження умов життя домогосподарств України) : [стат. зб.] / відп. за вип. І. І. Осипова ; Держкомстат України. – К. : ІВЦ Держкомстату України, 2003. – 344 с.

4. Методика формування вибірових сукупностей для проведення у 2009–2013 роках вибірових обстежень населення (домогосподарств): умов життя домогосподарств, економічної активності населення та сільськогосподарської діяльності населення у сільській місцевості ; затверджена Наказом Держкомстату від 14.08.2009 р. № 308 [Електронний ресурс]. – Режим доступу :

http://ukrstat.org/uk/metod_polog/metod_doc/2009/308/metod.htm

5. Kish L. Survey sampling / Kish L. – New York : John Wiley & Sons, 1995. – 643 p.

6. Кокрен У. Методы выборочного исследования / Кокрен У. – М. : Статистика, 1976. – 440 с.

7. Kasprzyk D. Reporting Sources of Error in U.S. Federal Government Surveys / D. Kasprzyk, L. Giesbrecht // Journal of Official Statistics. – 2003. – Vol. 19, № 4. – P. 343–363.
8. Методологічні розробки з питань підвищення надійності оцінювання визначених показників бідності на регіональному рівні. – К. : Міністерство соціальної політики України, 2011. – 111 с.
9. Методика комплексної оцінки бідності ; затверджена Наказом Міністерства праці та соціальної політики України, Міністерства фінансів України, Міністерства економіки та з питань європейської інтеграції України, Державного комітету статистики України, Національної академії наук України від 05.04.2002 р. № 171/238/100/149/2нд [Електронний ресурс]. – Режим доступу :
<http://zakon1.rada.gov.ua/laws/show/z0401-02>
10. Сариогло В. Г. Проблеми статистичного зважування вибірових даних : [моногр.] / Сариогло В. Г. – К. : ІВЦ Держкомстату України, 2005. – 264 с.
11. 8. Обратить реформы на благо всех и каждого. Бедность и неравенство в странах Европы и Центральной Азии. – Вашингтон : Всемирный Банк, 2001. – 495 с.
12. Ghosh M. Small Area Estimation: An Appraisal / M. Ghosh, J. N. K. Rao // Statistical Science. – 1994. – Vol. 9, № 1. – P. 55–93.
13. Rao J. N. K. Small Area Estimation / Rao J. N. K. – New York : John Wiley&Sons, 2003. – 314 p.
14. Longford N. T. Missing Data and Small-Area Estimation. Modern Analytical Equipment for the Survey Statistician (Statistics for Social Science and Behavioral Sciences) / Longford N. T. – New York : Springer-Verlag, 2005. – 357 p.

15. Longford N. T. Estimation of the poverty rates in the regions of Ukraine / N. T. Longford // The report prepared for the Social Assistance System Modernization Project, Ukraine. – Barcelona : Universitat Pompeu Fabra, 2009. – 18 p.
16. Enhancing Small Area Estimation Techniques to meet European Needs (EURAREA) [Electronic resource] // Project IST-2000-26290. – Access mode :
http://www.ine.es/en/docutrab/eurarea/eurarea_05_en.pdf
17. Use of ACS Data to Produce SAIPE Model-Based Estimates of Poverty for Counties // Report. – Washington : U.S. Census Bureau, 2007. – 117 p.
18. Tzavidis N. Small Area Methods for Poverty and Living Condition Estimates [Electronic resource] / N. Tzavidis // CCSR Cathie Marsh Centre for Census and Survey Research. – Access mode :
<http://www.ccsr.ac.uk/ssrg/documents/NKPosterBAEUProjectLaunchDay.pdf>
19. Афифи А. Статистический анализ: Подход с использованием ЭВМ / Афифи А., Эйзен С. – М. : Мир, 1982. – 488 с.
20. Саріогло В. Г. Методологічні підходи до підвищення надійності оцінок рівня бідності для регіонів України / В. Г. Саріогло, М. Ю. Огай // Статистика України. – 2003. – № 4. – С. 32–35.
21. Хастингс Н. Справочник по статистическим распределениям / Хастингс Н., Пикок Дж. – М. : Статистика, 1980. – 86 с.
22. Огай М. Ю. Статистичне оцінювання бідності для регіонів України: автореф. дис. на здобуття наук. ступеня канд. екон. наук : спец. 08.03.01 «Статистика» / М. Ю. Огай. – К., 2004. – 22 с.

23. Методологічні засади комплексного розвитку і розміщення продуктивних сил регіонів. – К. : РВПС України НАН України, 1998. – 161 с.
24. Методологічні положення з підготовки стандартних звітів щодо якості результатів державних вибірових обстежень населення (домогосподарств) ; затверджені Наказом Держкомстату від 05.06.2008 р. № 178 [Електронний ресурс]. – Режим доступу :
http://ukrstat.org/uk/metod_polog/metod_doc/2008/178/metod.htm
25. Описи програмного забезпечення для реалізації процедур підвищення надійності оцінювання показників бідності на регіональному рівні та ін. – К. : Міністерство соціальної політики України, 2011. – 94 с.
26. Longford N. T. Sample size calculation for small-area estimation of the poverty rates in the oblasts of Ukraine / Longford N. T. // The report prepared for the Social Assistance System Modernization Project, Ukraine. – Barcelona : Universitat Pompeu Fabra, 2010.
27. Longford N. T. Sample Size Calculation for Small Area Estimation // Survey Methodology. – 2006. – Vol. 32, No. 1. – P. 87–96.
28. Гладун О. М. Вибіркові обстеження населення: методологія, методика, практика / Гладун О. М. – Ніжин : ТОВ «Видавництво «Аспект – Поліграф», 2008. – 348 с.
29. Методичні рекомендації щодо оптимізації плану вибірки обстеження умов життя домогосподарств з метою підвищення надійності оцінки показників бідності на регіональному рівні. – К. : Держкомстат України, 2010. – 41 с.

ДОДАТКИ

Додаток А

Надійність прямого оцінювання показників бідності у 2007 – 2011 роках

Таблиця А.1

Надійність прямих оцінок визначених показників бідності для національного та регіонального рівнів (ОУЖД, 2007 р.)

(%)

Регіон	Рівень бідності за національ- ною межею, $\hat{P}_{1k}$	Коефіцієнт варіації, $CV(\hat{P}_{1k})$	Рівень бідності за регіональ- ною межею, $\hat{P}_{2k}$	Коефіцієнт варіації, $CV(\hat{P}_{2k})$	Рівень бідності за прожит- ковим мінімумом, $\hat{P}_{3k}$	Коефіцієнт варіації, $CV(\hat{P}_{3k})$
Україна	27,27	2,47	32,30	2,04	12,86	4,72
АР Крим	29,20	13,04	37,33	10,82	20,58	23,05
Вінницька область	23,03	14,43	24,63	14,88	17,29	18,00
Волинська область	39,87	13,18	38,31	13,75	25,66	16,23
Дніпропетровська область	27,85	11,26	32,94	9,43	12,35	15,32
Донецька область	22,43	10,07	30,61	8,37	6,97	15,76
Житомирська область	35,61	9,26	40,12	8,68	16,01	18,42
Закарпатська область	24,28	9,26	27,83	11,79	10,47	32,51
Запорізька область	21,84	12,91	24,32	12,36	12,28	17,15
Івано-Франківська область	16,88	22,48	21,64	21,42	9,25	28,13
Київська область	29,98	14,12	42,24	9,72	9,47	20,04
Кіровоградська область	45,57	10,22	46,96	11,42	25,53	21,09
Луганська область	39,34	7,24	42,59	6,90	12,56	13,00
Львівська область	29,79	8,82	34,02	8,44	13,73	15,21
Миколаївська область	19,63	14,93	20,42	14,07	14,10	18,84
Одеська область	34,61	11,83	42,58	9,55	13,68	15,43
Полтавська область	27,47	12,41	31,84	11,72	11,03	19,11
Рівненська область	34,78	10,81	35,76	10,92	14,24	24,28
Сумська область	30,87	11,84	34,60	10,52	12,94	20,71
Тернопільська область	33,30	10,74	36,97	6,94	19,89	21,90
Харківська область	28,18	12,00	32,77	10,49	8,74	24,55
Херсонська область	27,44	10,52	26,42	10,24	17,30	26,60
Хмельницька область	34,69	7,39	35,07	7,35	15,24	11,76
Черкаська область	16,29	12,98	24,28	10,69	7,44	22,03
Чернівецька область	28,10	11,66	38,96	9,68	19,62	20,78
Чернігівська область	27,26	15,22	30,33	11,62	17,21	21,12
м. Київ	8,00	25,36	16,75	17,90	5,46	39,83
м. Севастополь	7,91	47,17	18,32	33,21	0,28	115,75

Таблиця А.2

**Надійність прямих оцінок визначених показників бідності
для національного та регіонального рівнів (ОУЖД, 2008 р.)**
(%)

Регіон	Рівень бідності за національною межею, $\hat{P}_{1k}$	Коефіцієнт варіації, $CV(\hat{P}_{1k})$	Рівень бідності за регіональною межею, $\hat{P}_{2k}$	Коефіцієнт варіації, $CV(\hat{P}_{2k})$	Рівень бідності за життєвим мінімумом, $\hat{P}_{3k}$	Коефіцієнт варіації, $CV(\hat{P}_{3k})$
Україна	26,97	2,41	32,62	2,15	7,10	5,40
АР Крим	24,89	9,21	34,36	6,96	7,80	26,30
Вінницька область	24,30	13,86	27,14	12,71	6,90	21,70
Волинська область	47,30	10,21	47,74	9,93	15,60	22,40
Дніпропетровська область	27,22	13,01	32,25	11,48	5,60	22,60
Донецька область	25,58	9,66	32,16	7,95	4,10	24,40
Житомирська область	34,22	14,72	39,14	11,29	9,60	22,30
Закарпатська область	20,30	22,07	24,00	19,93	8,30	39,00
Запорізька область	24,39	15,47	30,35	12,91	6,80	33,70
Івано-Франківська область	20,09	18,93	24,44	14,81	3,50	43,10
Київська область	21,17	23,63	35,13	16,69	5,00	51,90
Кіровоградська область	47,09	9,14	47,09	9,14	21,30	14,90
Луганська область	29,20	8,74	34,32	9,77	6,40	21,80
Львівська область	29,25	8,49	38,66	8,12	7,10	19,10
Миколаївська область	19,81	16,35	19,81	16,35	6,70	18,60
Одеська область	38,48	12,04	43,99	10,24	10,70	21,90
Полтавська область	27,47	12,36	32,48	10,60	6,90	28,40
Рівненська область	39,28	11,51	39,61	11,93	12,90	19,00
Сумська область	29,07	12,16	30,25	11,85	4,90	25,10
Тернопільська область	42,77	10,93	46,50	10,10	10,10	39,30
Харківська область	20,59	9,90	31,23	9,00	2,80	28,20
Херсонська область	32,96	12,80	32,96	12,80	12,30	18,90
Хмельницька область	29,79	7,79	32,79	8,71	11,30	22,40
Черкаська область	18,37	20,64	21,44	22,08	6,20	24,90
Чернівецька область	29,60	15,88	35,78	12,91	7,00	32,10
Чернігівська область	28,62	11,88	31,13	10,88	10,00	17,50
м. Київ	9,63	17,11	19,36	11,09	1,10	72,90
м. Севастополь	2,70	73,05	14,59	56,86	0,80	62,30

Таблиця А.3

Надійність прямих оцінок визначених показників бідності для національного та регіонального рівнів (ОУЖД, 2009 р.)

(%)

Регіон	Рівень бідності за національною межею, $\hat{P}_{1k}$	Коефіцієнт варіації, $CV(\hat{P}_{1k})$	Рівень бідності за регіональною межею, $\hat{P}_{2k}$	Коефіцієнт варіації, $CV(\hat{P}_{2k})$	Рівень бідності за прожитковим мінімумом, $\hat{P}_{3k}$	Коефіцієнт варіації, $CV(\hat{P}_{3k})$
Україна	26,37	2,63	31,73	1,99	5,77	6,66
АР Крим	25,84	15,20	34,35	11,32	2,91	29,58
Вінницька область	22,45	14,23	26,81	12,84	8,13	21,11
Волинська область	42,09	8,18	42,09	8,18	10,80	22,39
Дніпропетровська область	26,35	9,68	32,26	8,13	3,19	24,07
Донецька область	23,62	11,11	30,65	10,71	5,39	21,01
Житомирська область	35,73	14,95	41,78	14,72	7,98	27,19
Закарпатська область	23,03	23,08	33,01	20,44	4,94	38,21
Запорізька область	19,72	14,02	21,40	12,88	2,12	28,52
Івано-Франківська область	23,10	16,38	29,55	13,88	5,03	32,49
Київська область	20,81	11,60	32,17	8,90	2,06	16,68
Кіровоградська область	42,01	10,01	42,48	10,10	11,63	17,24
Луганська область	28,15	10,36	30,75	9,12	5,74	19,33
Львівська область	31,39	7,87	37,00	5,36	9,02	24,00
Миколаївська область	20,89	16,30	21,72	15,79	3,62	30,80
Одеська область	34,17	12,56	40,79	10,44	10,88	16,91
Полтавська область	28,40	14,72	30,41	13,56	5,99	28,24
Рівненська область	36,85	15,31	38,58	13,15	5,48	28,77
Сумська область	45,81	8,70	46,11	8,36	12,52	28,18
Тернопільська область	29,44	15,99	32,31	15,46	7,72	27,79
Харківська область	20,74	17,18	31,58	9,08	1,73	43,26
Херсонська область	50,28	8,60	50,28	8,60	15,58	19,14
Хмельницька область	31,02	10,59	30,59	10,72	5,70	35,60
Черкаська область	13,45	18,43	19,65	13,68	6,04	34,08
Чернівецька область	18,48	13,72	23,66	9,75	3,11	30,20
Чернігівська область	22,44	18,64	25,77	21,71	5,09	25,28
м. Київ	9,51	17,83	18,50	11,93	1,75	43,72
м. Севастополь	5,00	58,47	15,79	32,33	2,90	70,87

Таблиця А.4

**Надійність прямих оцінок визначених показників бідності
для національного та регіонального рівнів (ОУЖД, 2010 р.)**
(%)

Регіон	Рівень бідності за національною межею, $\hat{P}_{1k}$	Коефіцієнт варіації, $CV(\hat{P}_{1k})$	Рівень бідності за регіональною межею, $\hat{P}_{2k}$	Коефіцієнт варіації, $CV(\hat{P}_{2k})$	Рівень бідності за прожитковим мінімумом, $\hat{P}_{3k}$	Коефіцієнт варіації, $CV(\hat{P}_{3k})$
Україна	24,14	2,49	29,93	2,62	8,78	3,51
АР Крим	25,40	9,73	37,82	9,34	13,67	32,07
Вінницька область	26,23	14,03	27,45	13,56	10,01	20,93
Волинська область	37,41	14,25	36,12	15,25	10,47	16,85
Дніпропетровська область	27,86	7,53	33,21	7,48	6,13	25,09
Донецька область	19,51	9,80	27,89	9,57	6,51	22,90
Житомирська область	40,85	16,81	42,68	15,09	14,45	17,62
Закарпатська область	25,66	13,15	33,18	9,94	15,37	16,13
Запорізька область	18,06	16,56	21,95	12,57	8,78	23,47
Івано-Франківська область	21,84	14,74	29,15	15,06	8,94	26,21
Київська область	18,15	25,57	30,17	18,32	5,47	32,44
Кіровоградська область	39,59	11,92	39,82	11,72	13,25	21,75
Луганська область	19,99	24,65	30,47	15,33	5,33	28,76
Львівська область	29,12	10,40	37,16	9,00	10,88	16,00
Миколаївська область	18,71	20,54	19,03	20,35	6,47	32,35
Одеська область	25,90	12,48	33,12	11,45	9,84	28,27
Полтавська область	25,98	18,77	30,52	18,46	8,61	30,56
Рівненська область	33,71	25,81	35,17	25,26	11,48	25,97
Сумська область	23,96	19,42	27,30	15,21	8,74	31,89
Тернопільська область	37,34	17,70	38,25	17,06	16,07	24,48
Харківська область	19,09	13,37	27,38	7,96	2,96	34,93
Херсонська область	36,82	9,32	36,82	9,32	21,70	14,17
Хмельницька область	28,81	10,92	28,81	10,92	13,86	13,26
Черкаська область	14,06	11,68	14,71	12,31	8,63	16,53
Чернівецька область	25,85	11,37	29,70	11,25	9,16	38,81
Чернігівська область	19,72	17,15	21,08	17,19	4,86	32,99
м. Київ	10,04	23,33	21,03	15,13	2,70	36,23
м. Севастополь	0,87	63,55	8,26	37,80	1,51	72,47

Таблиця А.5

Надійність прямих оцінок визначених показників бідності для національного та регіонального рівнів (ОУЖД, 2011 р.)

(%)

Регіон	Рівень бідності за національною межею, $\hat{P}_{1k}$	Коефіцієнт варіації, $CV(\hat{P}_{1k})$	Рівень бідності за регіональною межею, $\hat{P}_{2k}$	Коефіцієнт варіації, $CV(\hat{P}_{2k})$	Рівень бідності за прожитковим мінімумом, $\hat{P}_{3k}$	Коефіцієнт варіації, $CV(\hat{P}_{3k})$
Україна	24,35	3,59	29,75	2,68	7,81	6,47
АР Крим	23,47	22,74	31,26	14,35	10,40	46,87
Вінницька область	20,34	12,95	22,36	12,17	9,20	26,95
Волинська область	28,37	10,37	25,35	9,78	10,67	21,10
Дніпропетровська область	29,05	9,06	32,93	7,84	6,64	20,15
Донецька область	21,14	11,77	30,85	9,40	5,44	24,38
Житомирська область	37,44	10,72	39,92	10,04	11,36	21,89
Закарпатська область	30,76	21,49	32,99	19,99	11,33	41,60
Запорізька область	13,78	22,38	18,67	17,11	4,94	43,18
Івано-Франківська область	25,25	11,97	28,42	11,65	9,13	42,16
Київська область	22,71	16,36	34,05	11,37	3,23	35,23
Кіровоградська область	36,71	11,87	38,08	11,37	8,05	27,23
Луганська область	29,45	11,86	37,92	8,28	5,83	22,91
Львівська область	35,66	8,58	43,27	8,51	12,04	18,53
Миколаївська область	14,31	25,95	14,21	26,20	5,63	39,58
Одеська область	29,71	14,22	36,87	9,84	10,12	25,34
Полтавська область	19,72	16,40	22,88	15,56	5,13	27,97
Рівненська область	40,00	7,65	42,81	8,94	11,89	31,94
Сумська область	23,16	18,23	28,84	13,20	6,24	36,78
Тернопільська область	39,77	18,99	43,03	16,46	11,81	24,48
Харківська область	15,26	13,19	24,18	9,64	4,90	35,45
Херсонська область	20,97	21,83	21,84	20,03	9,77	32,08
Хмельницька область	40,80	11,61	40,80	11,61	18,37	23,15
Черкаська область	14,53	30,96	16,83	29,37	6,03	60,31
Чернівецька область	23,77	19,31	34,10	10,99	9,51	49,05
Чернігівська область	22,17	13,85	23,06	14,11	9,81	21,59
м. Київ	5,12	25,35	13,85	17,22	2,57	34,49
м. Севастополь	1,31	70,90	2,46	52,22	0,09	107,82

Додаток Б

Результати композиційного оцінювання показників бідності у 2011 році

Таблиця Б.1

**Вагові коефіцієнти однофакторної композиційної
оціночної функції для оцінювання рівня бідності за
національною межею у 2011 році**

Регіон	$\hat{\gamma}_{1k}^{(1)}$
АР Крим	0,221715
Вінницька область	0,064887
Волинська область	0,079666
Дніпропетровська область	0,064790
Донецька область	0,058307
Житомирська область	0,138752
Закарпатська область	0,304112
Запорізька область	0,086857
Івано-Франківська область	0,083713
Київська область	0,121307
Кіровоградська область	0,159591
Луганська область	0,108741
Львівська область	0,085609
Миколаївська область	0,121198
Одеська область	0,151468
Полтавська область	0,094699
Рівненська область	0,085628
Сумська область	0,151304
Тернопільська область	0,363236
Харківська область	0,038940
Херсонська область	0,173268
Хмельницька область	0,183277
Черкаська область	0,168321
Чернівецька область	0,174034
Чернігівська область	0,086168
м. Київ	0,016569
м. Севастополь	0,008554

Таблиця Б.2

**Результати композиційного оцінювання рівня бідності
за національною межею у 2011 році на основі
однофакторної композиційної оціночної функції
(%)**

Регіон	Оцінка, $P_{comp,k}^{(1)}$	Відносна похибка оцінювання, $RRMSE(P_{comp,k}^{(1)})$	Гранична похибка, $ME(P_{comp,k}^{(1)})$
Україна	23,21	2,73	1,24
АР Крим	23,67	18,94	8,78
Вінницька область	20,54	12,21	4,92
Волинська область	27,73	9,99	5,43
Дніпропетровська область	29,01	8,64	4,91
Донецька область	21,60	11,03	4,67
Житомирська область	35,18	10,26	7,07
Закарпатська область	32,79	15,73	10,11
Запорізька область	15,25	18,94	5,66
Івано-Франківська область	24,29	11,68	5,56
Київська область	23,02	14,72	6,64
Кіровоградська область	34,48	11,18	7,55
Луганська область	30,24	10,64	6,30
Львівська область	35,13	8,17	5,62
Миколаївська область	16,90	20,04	6,64
Одеська область	27,38	13,74	7,37
Полтавська область	20,67	14,57	5,90
Рівненська область	38,26	7,50	5,62
Сумська область	25,71	14,62	7,37
Тернопільська область	33,74	16,51	10,92
Харківська область	16,21	12,06	3,83
Херсонська область	19,98	20,04	7,85
Хмельницька область	37,17	11,05	8,05
Черкаська область	18,95	20,84	7,74
Чернівецька область	22,16	18,10	7,86
Чернігівська область	22,31	12,90	5,64
м. Київ	5,40	23,73	2,51
м. Севастополь	1,34	68,75	1,81

Таблиця Б.3

Вагові коефіцієнти трифакторної композиційної оціночної функції для оцінювання рівня бідності за національною межею у 2011 році

Регіон	$\hat{\gamma}_{1k}^{(1)}$	$\hat{\gamma}_{2k}^{(1)}$	$\hat{\gamma}_{3k}^{(1)}$
АР Крим	0,097027	-0,081328	0,000115
Вінницька область	0,105645	-0,087315	0,000129
Волинська область	0,149933	-0,119878	0,000196
Дніпропетровська область	0,076322	-0,064979	0,000087
Донецька область	0,057774	-0,049776	0,000064
Житомирська область	0,176522	-0,139362	0,000236
Закарпатська область	0,164056	-0,133237	0,000208
Запорізька область	0,076235	-0,064182	0,000089
Івано-Франківська область	0,144283	-0,118679	0,000178
Київська область	0,136466	-0,112987	0,000166
Кіровоградська область	0,188409	-0,148377	0,000253
Луганська область	0,100246	-0,084523	0,000117
Львівська область	0,096528	-0,081414	0,000113
Миколаївська область	0,082283	-0,068883	0,000098
Одеська область	0,105475	-0,087751	0,000127
Полтавська область	0,110685	-0,090337	0,000139
Рівненська область	0,177058	-0,141885	0,000230
Сумська область	0,149174	-0,121098	0,000189
Тернопільська область	0,190151	-0,149551	0,000256
Харківська область	0,055353	-0,047264	0,000063
Херсонська область	0,104654	-0,085303	0,000132
Хмельницька область	0,182543	-0,146939	0,000235
Черкаська область	0,120735	-0,100392	0,000145
Чернівецька область	0,148062	-0,119775	0,000189
Чернігівська область	0,137593	-0,113168	0,000170
м. Київ	0,029000	-0,024928	0,000032
м. Севастополь	0,059201	-0,050858	0,000066

Таблиця Б.4

**Результати композиційного оцінювання рівня бідності
за національною межею у 2011 році на основі
трифакторної композиційної оціночної функції
(%)**

Регіон	Оцінка, $P_{comp,k}^{(1)}$	Відносна похибка оцінювання, $RRMSE(P_{comp,k}^{(1)})$	Гранична похибка, $ME(P_{comp,k}^{(1)})$
Україна	24,16	1,66	0,79
АР Крим	23,53	7,98	3,68
Вінницька область	20,44	9,81	3,93
Волинська область	28,13	8,72	4,81
Дніпропетровська область	28,86	5,65	3,19
Донецька область	21,15	6,99	2,90
Житомирська область	36,53	7,53	5,39
Закарпатська область	30,42	8,38	5,00
Запорізька область	14,07	12,27	3,39
Івано-Франківська область	25,16	9,19	4,53
Київська область	22,79	9,97	4,45
Кіровоградська область	35,68	8,12	5,68
Луганська область	29,17	6,50	3,72
Львівська область	35,32	5,38	3,73
Миколаївська область	14,58	12,07	3,45
Одеська область	29,49	6,60	3,82
Полтавська область	19,92	10,44	4,08
Рівненська область	38,87	7,36	5,61
Сумська область	23,18	10,17	4,62
Тернопільська область	38,59	7,82	5,91
Харківська область	15,48	9,13	2,77
Херсонська область	21,15	9,44	3,91
Хмельницька область	39,64	7,36	5,72
Черкаська область	14,96	14,68	4,30
Чернівецька область	23,78	9,83	4,58
Чернігівська область	22,30	9,98	4,36
м. Київ	5,28	20,11	2,08
м. Севастополь	1,75	63,96	2,19

Таблиця Б.5

Вагові коефіцієнти трифакторної композиційної оціночної функції для оцінювання рівня бідності за регіональною межею у 2011 році

Регіон	$\hat{\gamma}_{1k}^{(1)}$	$\hat{\gamma}_{2k}^{(1)}$	$\hat{\gamma}_{3k}^{(1)}$
АР Крим	0,097027	-0,081328	0,000115
Вінницька область	0,105645	-0,087315	0,000129
Волинська область	0,149933	-0,119878	0,000196
Дніпропетровська область	0,076322	-0,064979	0,000087
Донецька область	0,057774	-0,049776	0,000064
Житомирська область	0,176522	-0,139362	0,000236
Закарпатська область	0,164056	-0,133237	0,000208
Запорізька область	0,076235	-0,064182	0,000089
Івано-Франківська область	0,144283	-0,118679	0,000178
Київська область	0,136466	-0,112987	0,000166
Кіровоградська область	0,188409	-0,148377	0,000253
Луганська область	0,100246	-0,084523	0,000117
Львівська область	0,096528	-0,081414	0,000113
Миколаївська область	0,082283	-0,068883	0,000098
Одеська область	0,105475	-0,087751	0,000127
Полтавська область	0,110685	-0,090337	0,000139
Рівненська область	0,177058	-0,141885	0,000230
Сумська область	0,149174	-0,121098	0,000189
Тернопільська область	0,190151	-0,149551	0,000256
Харківська область	0,055353	-0,047264	0,000063
Херсонська область	0,104654	-0,085303	0,000132
Хмельницька область	0,182543	-0,146939	0,000235
Черкаська область	0,120735	-0,100392	0,000145
Чернівецька область	0,148062	-0,119775	0,000189
Чернігівська область	0,137593	-0,113168	0,000170
м. Київ	0,029000	-0,024928	0,000032
м. Севастополь	0,059201	-0,050858	0,000066

Таблиця Б.6

**Результати композиційного оцінювання рівня бідності
за регіональною межею у 2011 році на основі трифакторної
композиційної оціночної функції**

(%)

Регіон	Оцінка, $P_{comp,k}^{(2)}$	Відносна похибка оцінювання, $RRMSE(P_{comp,k}^{(2)})$	Гранична похибка, $ME(P_{comp,k}^{(2)})$
Україна	29,64	1,46	0,85
АР Крим	31,21	6,56	4,02
Вінницька область	22,61	9,13	4,05
Волинська область	25,57	9,18	4,60
Дніпропетровська область	32,82	5,31	3,41
Донецька область	30,79	5,27	3,18
Житомирська область	39,23	7,03	5,41
Закарпатська область	32,72	7,82	5,02
Запорізька область	19,05	9,95	3,71
Івано-Франківська область	28,50	8,38	4,68
Київська область	33,75	7,52	4,98
Кіровоградська область	37,48	7,45	5,47
Луганська область	37,57	5,54	4,08
Львівська область	42,80	4,67	3,91
Миколаївська область	14,65	12,15	3,49
Одеська область	36,56	5,69	4,08
Полтавська область	23,18	9,30	4,22
Рівненська область	41,91	6,64	5,45
Сумська область	28,85	8,72	4,93
Тернопільська область	42,03	7,07	5,83
Харківська область	24,32	6,82	3,25
Херсонська область	22,16	9,28	4,03
Хмельницька область	39,97	6,98	5,47
Черкаська область	17,44	13,36	4,57
Чернівецька область	33,79	7,75	5,13
Чернігівська область	23,41	9,82	4,50
м. Київ	14,21	11,42	3,18
м. Севастополь	3,03	49,65	2,95

Таблиця Б.7

**Вагові коефіцієнти трифакторної
композиційної оціночної функції для оцінювання рівня
бідності за прожитковим мінімумом у 2011 році**

Регіон	$\hat{\gamma}_{1k}^{(3)}$	$\hat{\gamma}_{2k}^{(3)}$	$\hat{\gamma}_{3k}^{(3)}$
АР Крим	0,220238	-0,112688	0,000099
Вінницька область	0,237808	-0,121868	0,000107
Волинська область	0,302241	-0,148937	0,000145
Дніпропетровська область	0,118325	-0,065621	0,000046
Донецька область	0,091947	-0,051341	0,000035
Житомирська область	0,318951	-0,148701	0,000166
Закарпатська область	0,312114	-0,146499	0,000161
Запорізька область	0,143801	-0,075334	0,000062
Івано-Франківська область	0,274188	-0,138865	0,000126
Київська область	0,138508	-0,073293	0,000059
Кіровоградська область	0,274968	-0,129267	0,000141
Луганська область	0,135587	-0,073947	0,000055
Львівська область	0,202747	-0,107772	0,000086
Миколаївська область	0,169190	-0,089876	0,000072
Одеська область	0,210103	-0,109069	0,000092
Полтавська область	0,168967	-0,086163	0,000077
Рівненська область	0,325477	-0,158469	0,000159
Сумська область	0,229264	-0,114982	0,000108
Тернопільська область	0,336331	-0,152222	0,000182
Харківська область	0,103417	-0,058196	0,000039
Херсонська область	0,232901	-0,109200	0,000120
Хмельницька область	0,408271	-0,192286	0,000209
Черкаська область	0,239789	-0,120256	0,000112
Чернівецька область	0,296494	-0,147356	0,000141
Чернігівська область	0,302096	-0,159549	0,000130
м. Київ	0,088441	-0,049358	0,000034
м. Севастополь	0,190587	-0,099721	0,000083

Таблиця Б.8

**Результати композиційного оцінювання рівня бідності
за прожитковим мінімумом у 2011 році на основі
трифакторної композиційної оціночної функції**
(%)

Регіон	Оцінка, $P_{comp,k}^{(3)}$	Відносна похибка оцінювання, $RRMSE(P_{comp,k}^{(3)})$	Гранична похибка, $ME(P_{comp,k}^{(3)})$
Україна	7,59	3,29	0,49
АР Крим	10,07	12,42	2,45
Вінницька область	8,98	14,33	2,52
Волинська область	10,12	14,81	2,94
Дніпропетровська область	6,67	13,29	1,74
Донецька область	5,54	14,41	1,56
Житомирська область	10,63	15,24	3,18
Закарпатська область	10,60	15,32	3,18
Запорізька область	5,12	20,42	2,05
Івано-Франківська область	8,90	15,47	2,70
Київська область	3,50	28,59	1,96
Кіровоградська область	7,93	17,70	2,75
Луганська область	5,97	16,31	1,91
Львівська область	11,59	10,63	2,41
Миколаївська область	5,80	19,03	2,16
Одеська область	9,84	12,57	2,42
Полтавська область	5,31	21,13	2,20
Рівненська область	11,01	15,50	3,35
Сумська область	6,41	20,45	2,57
Тернопільська область	10,93	15,92	3,41
Харківська область	5,03	16,84	1,66
Херсонська область	9,47	13,75	2,55
Хмельницька область	15,56	21,37	6,52
Черкаська область	6,23	21,59	2,64
Чернівецька область	9,16	15,57	2,80
Чернігівська область	9,45	15,24	2,82
м. Київ	2,76	26,35	1,42
м. Севастополь	0,85	97,41	1,63

Додаток В

Приклади вхідної та вихідної інформації при застосуванні інструментарію композиційного оцінювання

Таблиця В.1

Приклад вхідної інформації для роботи програмного забезпечення з оцінювання рівня
бідності з використанням одно-, дво- та трифакторних композиційних оціночних функцій

Регіон, область	Чисельність постійного населення, осіб	Обсяг вибірки, од.	Рівень бідності за національною межею, P1, %	Коефіцієнт варіації, CV(P1), %	Рівень бідності за регіональною межею, P2, %	Коефіцієнт варіації, CV(P2), %	Рівень бідності за прожитковим мінімумом, P3, %	Коефіцієнт варіації, CV(P3), %	Зовнішня інформація
Україна	45525731	10641	24,35	3,59	29,75	2,68	29,75	6,47	18944,08
АР Крим	1954506	480	23,47	22,74	31,26	14,35	31,26	46,87	18481,64
Вінницька	1630609	389	20,34	12,95	22,36	12,17	22,36	26,95	15250,05
Волинська	1035145	313	28,37	10,37	25,35	9,78	25,35	21,10	15183,54
Дніпропетровська	3325092	714	29,05	9,06	32,93	7,84	32,93	20,15	19668,61
Донецька	4405209	774	21,14	11,77	30,85	9,40	30,85	24,38	18704,87
Житомирська	1276872	290	37,44	10,72	39,92	10,04	39,92	21,89	15128,91
Закарпатська	1246221	300	30,76	21,49	32,99	19,99	32,99	41,60	13265,01
Запорізька	1795702	419	13,78	22,38	18,67	17,11	18,67	43,18	20554,41
Івано-Франківська	1377221	316	25,25	11,97	28,42	11,65	28,42	42,16	14831,88
Київська	1712824	315	22,71	16,36	34,05	11,37	34,05	35,23	18745,49
Кіровоградська	999789	266	36,71	11,87	38,08	11,37	38,08	27,23	14558,74
Луганська	2277376	538	29,45	11,86	37,92	8,28	37,92	22,91	15520,16
Львівська	2524473	617	35,66	8,58	43,27	8,51	43,27	18,53	16168,40
Миколаївська	1180038	401	14,31	25,95	14,21	26,20	14,21	39,58	16994,41

Продовження таблиці В.1

Одеська	2377424	499	29,71	14,22	36,87	9,84	36,87	25,34	22192,59
Полтавська	1474635	354	19,72	16,40	22,88	15,56	22,88	27,97	15847,55
Рівненська	1152331	303	40,00	7,65	42,81	8,94	42,81	31,94	14567,37
Сумська	1154747	282	23,16	18,23	28,84	13,20	28,84	36,78	14511,19
Тернопільська	1079028	270	39,77	18,99	43,03	16,46	43,03	24,48	13607,13
Харківська	2733008	646	15,26	13,19	24,18	9,64	24,18	35,45	22589,68
Херсонська	1084370	395	20,97	21,83	21,84	20,03	21,84	32,08	15616,14
Хмельницька	1320386	295	40,80	11,61	40,80	11,61	40,80	23,15	15106,91
Черкаська	1277715	259	14,53	30,96	16,83	29,37	16,83	60,31	16542,57
Чернівецька	901706	290	23,77	19,31	34,10	10,99	34,10	49,05	14571,18
Чернігівська	1084817	307	22,17	13,85	23,06	14,11	23,06	21,59	15588,99

**Приклад вихідної інформації за регіонами для внутрішнього використання за
результатами композиційного оцінювання рівня бідності за національною межею
у 2011 році на основі трифакторної композиційної оціночної функції**

Регіон, область	Пряма оцінка показника	Дисперсія прямої оцінки	Коефіцієнт варіації прямої оцінки	Довірчий інтервал для прямої оцінки		Синтетична оцінка	Ваговий коефіцієнт			Композиційна оцінка	Середньоквадратична похибка композиційної оцінки	Гранична похибка композиційної оцінки	Відносна похибка композиційної оцінки
				нижня межа	верхня межа		$\hat{\gamma}_{1k}$	$\hat{\gamma}_{2k}$	$\hat{\gamma}_{3k}$				
АР Крим	23,5	48,31	29,58	9,88	37,12	24,86	0,097026	-0,081328	0,000114	23,5	3,52	3,68	7,98
Вінницька	20,3	18,36	21,11	11,90	28,70	26,51	0,105645	-0,087315	0,000129	20,4	4,02	3,93	9,81
Волинська область	28,4	40,43	22,39	15,94	40,86	35,18	0,149933	-0,119877	0,000195	28,1	6,02	4,81	8,72
Дніпропетровська	29,0	48,74	24,07	15,32	42,68	26,57	0,076322	-0,064979	0,000087	28,9	2,66	3,19	5,65
Донецька	21,1	19,66	21,01	12,41	29,79	20,18	0,057773	-0,049776	0,000064	21,1	2,18	2,90	6,99
Житомирська	37,4	103,43	27,19	17,47	57,33	38,25	0,176521	-0,139362	0,000236	36,5	7,56	5,39	7,53
Закарпатська	30,8	138,49	38,21	7,73	53,87	26,55	0,164055	-0,133237	0,000207	30,4	6,50	5,00	8,38
Запорізька	13,8	15,49	28,52	6,09	21,51	18,68	0,076235	-0,064181	0,000089	14,1	2,98	3,39	12,27
Івано-Франківська	25,2	67,05	32,49	9,15	41,25	23,22	0,144282	-0,118678	0,000178	25,2	5,35	4,53	9,19
Київська область	22,7	14,34	16,68	15,28	30,12	19,17	0,136466	-0,112986	0,000165	22,8	5,16	4,45	9,97
Кіровоградська	36,7	40,04	17,24	24,30	49,10	37,22	0,188409	-0,148377	0,000253	35,7	8,39	5,68	8,12
Луганська	29,4	32,31	19,33	18,26	40,54	21,74	0,100246	-0,084523	0,000117	29,2	3,60	3,72	6,50
Львівська	35,7	73,40	24,00	18,91	52,49	28,52	0,096527	-0,081414	0,000112	35,3	3,61	3,73	5,38
Миколаївська	14,3	19,40	30,80	5,67	22,93	20,13	0,082283	-0,068882	0,000097	14,6	3,10	3,45	12,07
Одеська	29,7	25,22	16,91	19,86	39,54	24,12	0,105474	-0,087750	0,000126	29,5	3,79	3,82	6,60
Полтавська	19,7	30,96	28,24	8,79	30,61	26,11	0,110684	-0,090337	0,000138	19,9	4,33	4,08	10,44
Рівненська	40,0	132,42	28,77	17,45	62,55	32,73	0,177057	-0,141884	0,000230	38,9	8,19	5,61	7,36
Сумська	23,2	42,74	28,18	10,39	36,01	25,10	0,149174	-0,121098	0,000189	23,2	5,55	4,62	10,17
Тернопільська	39,8	122,29	27,79	18,13	61,47	35,94	0,190150	-0,149551	0,000256	38,6	9,10	5,91	7,82
Харківська	15,3	43,81	43,26	2,33	28,27	18,69	0,055353	-0,047264	0,000062	15,5	2,00	2,77	9,13

Продовження таблиці В.2.1

Херсонська	21,0	16,15	19,14	13,12	28,88	34,78	0,104654	-0,085303	0,000131	21,1	3,98	3,91	9,44
Хмельницька	40,8	211,02	35,60	12,33	69,27	28,61	0,182542	-0,146939	0,000235	39,6	8,51	5,72	7,36
Черкаська	14,5	24,42	34,08	4,81	24,19	16,71	0,120735	-0,100392	0,000145	15,0	4,82	4,30	14,68
Чернівецька	23,8	51,67	30,20	9,71	37,89	26,50	0,148062	-0,119774	0,000189	23,8	5,46	4,58	9,83
Чернігівська	22,2	31,49	25,28	11,20	33,20	21,54	0,137592	-0,113167	0,000169	22,3	4,96	4,36	9,98
м. Київ	5,1	4,97	43,72	0,73	9,47	3,69	0,028999	-0,024928	0,000031	5,3	1,13	2,08	20,11
м. Севастополь	1,3	0,85	70,87	0,00	3,11	4,07	0,059201	-0,050858	0,000066	1,7	1,25	2,19	63,96

Таблиця В.2.2

Приклад вихідної інформації для внутрішнього використання за результатами композиційного оцінювання рівня бідності за національною межею у 2011 році на основі трифакторної композиційної оціночної функції

Назва показника	Оцінка показника
Коефіцієнт регресії (вільний член), $\hat{\beta}_0$	10,7737
Коефіцієнт регресії, $\hat{\beta}_1$	0,7833
Коефіцієнт регресії, $\hat{\beta}_2$	-0,0004
Коефіцієнт детермінації, R^2	0,6983
F-відношення	28,9381
Дисперсії синтетичної оціночної функції, $\hat{\sigma}_v^2$	69,7887

Таблиця В.2.3

**Приклад вихідної інформації для зовнішнього
використання за результатами композиційного
оцінювання рівня бідності за національною межею за
регіонами у 2011 році на основі трифакторної
композиційної оціночної функції**

(%)

Region	Уточнена оцінка рівня бідності	Відносна похибка оцінювання, <i>RRMSE</i>	Довірчий інтервал	
			нижня межа	верхня межа
Україна	24,16	1,66	23,37	24,95
АР Крим	23,53	7,98	19,85	27,21
Вінницька область	20,44	9,81	16,51	24,37
Волинська область	28,13	8,72	23,32	32,94
Дніпропетровська область	28,86	5,65	25,67	32,05
Донецька область	21,15	6,99	18,25	24,04
Житомирська область	36,53	7,53	31,14	41,92
Закарпатська область	30,42	8,38	25,43	35,42
Запорізька область	14,07	12,27	10,69	17,46
Івано-Франківська область	25,16	9,19	20,62	29,69
Київська область	22,79	9,97	18,34	27,24
Кіровоградська область	35,68	8,12	30,00	41,35
Луганська область	29,17	6,50	25,45	32,89
Львівська область	35,32	5,38	31,59	39,04
Миколаївська область	14,58	12,07	11,13	18,03
Одеська область	29,49	6,60	25,67	33,31
Полтавська область	19,92	10,44	15,85	24,00
Рівненська область	38,87	7,36	33,26	44,48
Сумська область	23,18	10,17	18,56	27,80
Тернопільська область	38,59	7,82	32,68	44,50
Харківська область	15,48	9,13	12,71	18,25
Херсонська область	21,15	9,44	17,24	25,06
Хмельницька область	39,64	7,36	33,92	45,35
Черкаська область	14,96	14,68	10,66	19,27
Чернівецька область	23,78	9,83	19,20	28,36
Чернігівська область	22,30	9,98	17,94	26,66
м. Київ	5,28	20,11	3,20	7,36
м. Севастополь	1,75	63,96	0,00	3,94

Таблиця В.2.4

Приклад вихідної інформації для зовнішнього використання за результатами композиційного оцінювання рівня бідності за національною межею за економічними районами у 2011 році на основі трифакторної композиційної оціночної функції

(%)

Економічний район	Уточнена оцінка рівня бідності	Відносна похибка оцінювання, <i>RRMSE</i>	Довірчий інтервал	
			нижня межа	верхня межа
Східний	18,36	5,72	16,30	20,42
Донецький	23,88	4,89	21,59	26,17
Придніпровський	23,68	5,15	21,29	26,07
Причорноморський	22,49	4,23	20,63	24,36
Подільський	31,59	4,72	28,67	34,51
Центральний	16,05	5,84	14,21	17,88
Карпатський	30,28	3,77	28,04	32,51
Поліський	31,82	4,12	29,25	34,38

Додаток Г

Приклади оптимізації розподілу обсягу вибірки обстеження умов життя домогосподарств за регіонами з урахуванням композиційного оцінювання показників бідності

Таблиця Г.1

**Оптимальні обсяги вибірки при мінімізації відносної
середньоквадратичної похибки оцінювання визначених
показників бідності**
(загальний обсяг вибірки відповідає кількості обстежених
домогосподарств в ОУЖД у 2009 році, $n = 10459$)

Регіон	Розподіл кількості обстежених домогосподарств в ОУЖД у 2009 році	Оптимальний розподіл вибірки для $q = 1,25$ та $G=0,1$	Оптимальний розподіл вибірки для $q = 1,25$ та $G=0,1$ (за умови визначення мінімаль- ного обсягу вибірки для регіону у 300 д/г)
Україна	10459	10459	10459
АР Крим	474	545	471
Вінницька область	419	302	300
Волинська область	307	147	300
Дніпропетровська область	707	742	642
Донецька область	732	706	609
Житомирська область	278	223	300
Закарпатська область	327	310	300
Запорізька область	419	627	543
Івано-Франківська область	319	333	300
Київська область	333	603	522
Кіровоградська область	289	136	300
Луганська область	532	433	372
Львівська область	594	354	301
Миколаївська область	392	367	316
Одеська область	358	297	300
Полтавська область	341	314	300
Рівненська область	305	261	300
Сумська область	280	140	300
Тернопільська область	267	213	300
Харківська область	586	889	771
Херсонська область	387	105	300
Хмельницька область	293	294	300

Продовження таблиці Г.1

Черкаська область	293	331	300
Чернівецька область	293	330	300
Чернігівська область	309	289	300
м. Київ	467	934	812
м. Севастополь	158	234	300

Таблиця Г.2

Оптимальні обсяги вибірки при мінімізації відносної середньоквадратичної похибки оцінювання визначених показників бідності

Регіон	Обсяг вибірки, який забезпечує надійність умовного показника ОУЖД на рівні $CV = 10\%$ для всіх регіонів	Оптимальний розподіл вибірки для $q = 1,25$ та $G=0,1^*$	Сумарний обсяг вибірки	Сумарний обсяг вибірки, який враховує обмеження мінімального обсягу вибірки
Україна	8367	2092	10459	10459
АР Крим	163	124	287	300
Вінницька область	605	51	656	636
Волинська область	202	10	212	300
Дніпропетровська область	599	173	772	698
Донецька область	257	156	413	342
Житомирська область	573	26	599	588
Закарпатська область	157	56	213	300
Запорізька область	202	149	351	300
Івано-Франківська область	457	62	519	465
Київська область	294	143	437	397
Кіровоградська область	398	10	408	382
Луганська область	237	84	321	300
Львівська область	233	58	291	300
Миколаївська область	191	76	267	300
Одеська область	488	43	531	509
Полтавська область	170	53	223	300
Рівненська область	184	40	224	300
Сумська область	276	10	286	300
Тернопільська область	392	26	418	407
Харківська область	285	219	504	446
Херсонська область	211	10	221	300
Хмельницька область	274	49	323	300
Черкаська область	222	65	287	300
Чернівецька область	223	68	291	300
Чернігівська область	201	51	252	300
м. Київ	658	230	888	789
м. Севастополь	215	50	265	300

* Обсяг вибірки для оптимального розподілу визначено як різницю між обсягом вибірки ОУЖД у 2009 році $n = 10459$ та обсягом, мінімально необхідним для ОУЖД, $n = 8367$.

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.

[illegible]

[illegible]

[illegible]

Наукове видання

Саріогло Володимир Георгійович

**ОЦІНЮВАННЯ СОЦІАЛЬНО-ЕКОНОМІЧНИХ
ПОКАЗНИКІВ: ПРИКЛАДНІ АСПЕКТИ
ЗАСТОСУВАННЯ НЕПРЯМИХ МЕТОДІВ**

Монографія

Редагування

В. Г. Чебанова

Комп'ютерне верстання

Д. В. Греков

Підписано до друку 18.12.2012. Формат 60x84/16
Папір офсет. Ум. арк. 7,68
Тираж 300. Зам. № 223

Видавець та виготівник “Сочінський”
20300, м. Умань, вул. Тищика, 18/19
тел. (04744) 4-64-88, 4-67-77
e-mail: vizavi08@mail.ru

Свідоцтво суб'єкта видавничої справи
ДК № 2521 від 08.06.2006.