

ЗАТВЕРДЖЕНО
Наказ Державного комітету
статистики України
23.12.2011 № 390

**МЕТОДИКА ПОБУДОВИ СТАТИСТИКО-МАТЕМАТИЧНОЇ МОДЕЛІ
ДЛЯ ПІДВИЩЕННЯ НАДІЙНОСТІ ОЦІНОК ПОКАЗНИКІВ БІДНОСТІ НА
РЕГІОНАЛЬНОМУ РІВНІ**

Київ - 2011

АНОТАЦІЯ

Методика побудови статистико-математичної моделі для підвищення надійності оцінок показників бідності на регіональному рівні (далі – Методика) визначає порядок побудови статистико-математичних моделей для оцінювання на регіональному рівні основних показників бідності - рівнів бідності за національною та регіональною межею, а також питомої ваги населення з доходами, нижчими за розмір прожиткового мінімуму.

Методика підготовлена експертом проекту Світового банку "Удосконалення системи соціальної допомоги в Україні" Саріогло В.Г. з урахуванням рекомендацій міжнародного експерта Лонгфорда Н.Т. за участю фахівців департаменту обстежень домогосподарств Держкомстату:

Осипової І.І. – директора департаменту обстежень домогосподарств;

Спесивої-Ухової С.В. – заступника директора цього департаменту;

Пліско К.Ю. – заступника директора департаменту – начальника відділу методологічного та організаційного забезпечення обстеження умов життя домогосподарств.

Методика призначена для використання працівниками органів державної статистики, міністерств, інших органів виконавчої влади на національному рівні, науковцями, які здійснюють дослідження у сфері вибіркового обстеження домогосподарств та питань бідності, іншими користувачами статистичної інформації.

Зміст

Передмова.....	4
1. Основні терміни та визначення.....	5
2. Принципи побудови статистико-математичних моделей для підвищення надійності оцінювання показників бідності.....	8
3. Статистико-математичні моделі для непрямого оцінювання визначених показників бідності	13
4. Вибір зовнішньої інформації для композиційного оцінювання показників бідності	16
5. Супровід використання композиційних оціночних функцій	19
Додатки	21
Перелік використаних джерел	44

Передмова

Методика побудови статистико-математичної моделі для підвищення надійності оцінок показників бідності на регіональному рівні (далі – Методика) підготовлена на виконання заходів Стратегії розвитку державної статистики на період до 2012 р., схваленої розпорядженням Кабінету Міністрів України від 05.11.2008 р. № 413–р.

Метою створення Методики є розробка методичного забезпечення для побудови статистико-математичних моделей, призначених підвищити якість оцінювання на регіональному рівні основних показників бідності – рівнів бідності за національною та регіональною межею, а також питомої ваги населення з доходами, нижчими за розмір прожиткового мінімуму.

Методика містить визначення основних понять і термінів, опис порядку побудови статистико-математичних моделей для оцінювання основних показників бідності на регіональному рівні.

У документі розглянуто інформаційне забезпечення для побудови моделей та розрахунку показників бідності на їх основі, надано конкретні моделі, які пропонуються для впровадження у практику роботи органів державної статистики України

Методика складається з передмови, п'яти розділів, додатків та переліку використаних джерел.

Методика базується на сучасних науково обґрунтованих принципах і міжнародному досвіді побудови статистико-математичних моделей для оцінювання показників.

1. ОСНОВНІ ТЕРМІНИ І ВИЗНАЧЕННЯ

<i>Дисперсія вибірових оцінок</i>	– середній квадрат відхилень оцінок показника за всіма можливими вибірками даного дизайну (реплікаціями) від середнього значення оцінки за всіма вибірками
<i>Додаткова (зовнішня) інформація</i>	– у контексті цієї методики це інформація, яка 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>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, яка оцінюється
<i>Коефіцієнт варіації (відносна стандартна похибка)</i>	– відношення величини стандартної похибки вибірки до оцінки показника (найчастіше вимірюється у відсотках)
<i>Композиційна оціночна функція</i>	– оціночна функція, отримана на основі композиції (комбінації) різних оціночних функцій. У статистиці найчастіше використовується композиція прямих і непрямих оціночних функцій
<i>Критерій оптимальності</i>	– ознака, згідно з якою визначається найкращий з можливих варіантів структури, стану об'єкта; у 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 непрямої 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>Оцінка показника</i>	– значення показника, яке розраховано на основі відповідної оціночної функції за результатами вибіркового обстеження
<i>Оціночна функція</i>	– правило або співвідношення (формула) для оцінювання показника з використанням даних вибіркового обстеження. Оціночна функція відома ще до отримання результатів обстеження. Слід розрізняти поняття <i>оціночна функція</i> та <i>оцінка</i> , оскільки оціночна функція – це функція вибірки, тоді як оцінка є її значенням, знайденим на основі реалізації вибірки
<i>Оцінювання для малих територій</i>	– статистична методологія, яка розроблена з метою підвищення надійності оцінювання показників за результатами статистичних спостережень на основі прямих і непрямих оціночних функцій, комплексного використання інформації з різних джерел і сучасних методів статистико-математичного моделювання
<i>Пряме оцінювання</i>	– 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)
<i>Прожитковий мінімум</i>	– 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
<i>Рівень бідності за національною межею</i>	– частка населення країни, рівень споживання (доходів) якого є нижчим від установленої національної межі бідності
<i>Рівень бідності за регіональною межею</i>	– частка населення регіону, рівень споживання (доходів) якого є нижчим від установленої регіональної межі бідності

<i>Рівень бідності за прожитковим мінімумом</i>	– частка населення із загальними середньомісячними еквівалентними доходами, нижчими за прожитковий мінімум у середньому на одну особу
<i>Середньоквадратична похибка</i>	– загальна похибка оцінювання показника, яка визначається як сума дисперсії вибірових оцінок показника та квадрату величини зміщення оцінки, яке є різницею між очікуваною величиною оцінки (тобто середнім із вибірових оцінок) та дійсним значенням показника для генеральної сукупності
<i>Стандартна похибка</i>	– середньоквадратичне відхилення вибірових оцінок показника за всіма можливими вибірками даного дизайну від дійсного значення показника для генеральної сукупності
<i>Статистико-математичне моделювання</i>	– методи математичної статистики, що забезпечують вибір серед можливих теоретико-ймовірнісних моделей такої, яка щонайкраще (у певному визначеному сенсі) відповідає наявним статистичним даним, що характеризують досліджуване явище або процес

2. Принципи побудови статистико-математичних моделей для підвищення надійності оцінювання показників бідності

В Україні показники бідності вимірюються за даними вибіркового обстеження умов життя домогосподарств (далі – ОУЖД), яке проводиться органами державної статистики. За міжнародними рекомендаціями для державних вибірових обстежень населення (домогосподарств) надійність оцінок показників для країни в цілому (характеризується коефіцієнтом варіації CV) повинна забезпечуватися на рівні $CV \leq 5\% \div 8\%$. Для окремих територій (регіонів, районів) оцінки показників уважаються достатньо надійними, якщо коефіцієнт варіації не перевищує 10%, а для деяких обстежень і 15%. В Україні для вимірювання бідності зазначені вимоги щодо надійності оцінювання показників бідності задовольняються не повною мірою. За існуючих обсягів вибірки та наявної інформаційної бази статистична надійність показників бідності, визначених на рівні регіонів, є недостатньою.

Як свідчить досвід інших країн та результати проведених в Україні досліджень, підвищення надійності оцінювання визначених показників бідності (рівня бідності за національною межею, рівня бідності за регіональною межею та рівня бідності за розміром прожиткового мінімуму) за регіонами України може бути досягнуто на основі застосування спеціальної методології оцінювання – так званих методів малих територій. Основним принципом, на якому будуються методи оцінювання для малих територій, є те, що для підвищення надійності оцінювання показників, крім безпосередніх результатів вибіркового обстеження, за даними якого здійснюється пряма оцінка цих показників, за поточний період використовується додаткова інформація на рівні регіону або на рівні більшої території, до складу якої входить регіон. Наприклад, це можуть бути дані попередніх етапів обстеження, кореляти показників бідності, отримані з адміністративних джерел або за даними статистичної звітності тощо.

Непряме оцінювання показників бідності на рівні регіонів України здійснюється на основі багатовимірної композиційної оціночної функції, яка визначається за формулою:

$$P_{comp,k}^{(j)} = (u^{(j)} - \gamma_k^{(j)})^T \cdot \theta_k^{(j)} + \gamma_k^{(j)T} \cdot \theta^{(j)}, \quad j = 1, 2, 3. \quad (1)$$

де $P_{comp,k}^{(j)}$ - вектор оцінок j -го показника бідності для регіонів за композиційною оціночною функцією;

$\theta_k^{(j)}$ - 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 $\theta_k^{(j)} = \begin{Bmatrix} P_k^{(j)} \\ x_{1k}^{(j)} \\ \dots \\ x_{Mk}^{(j)} \end{Bmatrix}$, де $P_k^{(j)}$ -

пряма оцінка j -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

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

національного рівня;

$\gamma_k^{(j)}$ - вектор вагових коефіцієнтів композиційної оціночної функції за регіонами;

$$u^{(j)} - \text{індексний вектор, } u^{(j)} = \begin{Bmatrix} 1 \\ 0 \\ \dots \\ 0 \end{Bmatrix} \quad (\text{індексний вектор - це вектор, елементами}$$

якого є визначені цілі числа – у цьому випадку елементи вектора можуть приймати лише значення 1 або 0);

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 вектору.

Доцільно зазначити, що наведений вигляд композиційної оціночної функції цілком відповідає методології формування оціночних функцій, затвердженій у відповідних методологічних положеннях і наведений у [3, 4 формула (1)]. Запропоновані зміни є насамперед уточнюючими. Так, дещо змінено спосіб запису вагових коефіцієнтів $\gamma_k^{(j)}$ в оціночній функції (1). В оціночній функції, представлений у [3, 4], коефіцієнт $\gamma_k^{(j)}$ є вагою прямої оцінки j -го показника бідності для k -го регіону, а у формулі (1) – вагою оцінки показника національного рівня. При цьому інші коефіцієнти $-\gamma_k^{(j)}$ – вагами коваріат j -го показника бідності для k -го регіону. Для прямої оцінки показника ваговий коефіцієнт дорівнює $(1-\gamma_k^{(j)})$. Така зміна доцільна, оскільки до композиційної

оціночної функції (1) включено оцінку національного рівня [2]. Таким чином, найбільш суттєвою зміною у формі представлення композиційної оціночної функції (1) є урахування прямої оцінки показників бідності для національного рівня $\theta^{(j)}$. Водночас слід зазначити, що такий підхід цілком відповідає методології композиційного оцінювання [2, 5]. Пряма оцінка показника бідності національного рівня є класичною формою зовнішньої інформації при оцінюванні показників для регіонального рівня. Її застосування в кінцевому варіанті запропонованих моделей для підвищення надійності оцінювання показників бідності в Україні зумовлено відсутністю зовнішніх даних (насамперед йдеться 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.

Таким чином, реалізована у цій Методиці методологія підвищення надійності оцінювання показників бідності відповідає методологічним положенням, викладеним у [3, 4]. Зміни форми представлення композиційної оціночної функції (1) порівняно з формулою (1) методологічних положень [4] пов'язані з доцільністю спрощення процедур розрахунку оцінок показників та доцільністю використання при непрямому оцінюванні прямих оцінок відповідних показників для національного рівня.

При композиційному оцінюванні необхідно визначити такі вагові коефіцієнти $\gamma_k^{(j)}$, які забезпечують за заданих умов (визначеного набору зовнішніх змінних і статистичних властивостей їх розподілів за регіонами та взаємозв'язків з показниками бідності, що оцінюються) мінімальний рівень похибки оцінювання відповідних показників бідності для регіонів. Значення коефіцієнтів $\gamma_k^{(j)}$ визначається за узагальненим методом найменших квадратів таким співвідношенням, записаним у матричній формі (див. [1, 2]):

$$\gamma_k^{(j)} = [V_k^{(j)} + V^{(j)} - C_k^{(j)} - C_k^{(j)T} + \Sigma_b^{(j)}]^{-1} \cdot [V_k^{(j)} - C_k^{(j)}] \cdot u^{(j)}, \quad (2)$$

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

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

C_k - коваріаційна матриця регіональних і національних оцінок показників;

Σ_b^j - матриця дисперсій регіональних оцінок показників відносно їх середнього значення за регіонами.

У цілому побудова композиційної оціночної функції для уточнення оцінок визначених показників бідності за регіонами за певний рік складається з таких основних етапів [3, 4]:

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

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

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

$$P_{comp,U}^{(j)} = \frac{\sum_{k=1}^K N_k P_{comp,k}^{(j)}}{\sum_{k=1}^K N_k}, \quad (3)$$

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

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

N_k – середньорічна чисельність постійного населення k -го регіону.

Для економічних районів уточнені оцінки показників бідності розраховуються за аналогічними формулами з урахуванням даних лише по тих регіонах, які входять до складу відповідних економічних районів.

При оцінюванні показників бідності з використанням композиційних оціночних функцій за основний критерій надійності доцільно прийняти умову, за якої значення абсолютної або відносної похибки оцінювання, які у загальному випадку є похідними від величини середньоквадратичної похибки $MSE(P_{comp,k}^{(j)})$, повинне бути меншим за встановлене граничне значення відповідної похибки.

Величина середньоквадратичної похибки оцінки показників $MSE(P_{comp,k}^{(j)})$ визначається за формулою:

$$MSE(P_{comp,k}^{(j)}) = V(P_{comp,k}^{(j)}) + B^2(P_{comp,k}^{(j)}), \quad (4)$$

де $V(P_{comp,k}^{(j)})$ - дисперсія вибірових оцінок показника $P_{comp,k}^{(j)}$;

$B(P_{comp,k}^{(j)})$ - зміщення оцінки показника $P_{comp,k}^{(j)}$.

На практиці, враховуючи неможливість визначення величин дисперсії вибірових оцінок показника та її зміщення на основі аналітичних залежностей для випадків складного дизайну вибірки та застосування математико-статистичних моделей для оцінювання показників, середньоквадратична похибка оцінюється на основі наближених методів. У методиці величина $MSE(P_{comp,k}^{(j)})$ визначається на основі методу реплікацій, реалізованому в розробленому пакеті програм (див. п.2 методики).

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

$$RMSE(P_{comp,k}^{(j)}) = \sqrt{MSE(P_{comp,k}^{(j)})} \quad (5)$$

Величина $RMSE(P_{comp,k}^{(j)})$ є аналогом стандартної похибки для прямих оцінок, але вона враховує також зміщення оцінок.

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

$$RRMSE(P_{comp,k}^{(j)}) = \frac{\sqrt{MSE(P_{comp,k}^{(j)})}}{P_{comp,k}^{(j)}} \cdot 100\% . \quad (6)$$

Величина $RRMSE(P_{comp,k}^{(j)})$ є аналогом коефіцієнта варіації для прямих оцінок і відрізняється від нього тим, що враховує зміщення оцінок за моделлю. Межі значень $RRMSE(P_{comp,k}^{(j)})$ визначають рівень надійності показників такі ж, 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 і 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 не рекомендується використовувати для аналізу бідності.

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

$$ME(P_{comp,k}^{(j)}) = 1.96 \cdot RMSE(P_{comp,k}^{(j)}), \quad (7)$$

Для оцінки результатів моделювання здійснюються такі основні перевірки:

- перевіряються значення оцінок вагових коефіцієнтів композиційної оціночної функції з метою підтвердження того, що значення потрапляють до визначеного інтервалу: $0 \leq |\hat{\gamma}_k| \leq 1$;
- 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 установленим вимогам щодо рівня надійності оцінювання показників бідності за регіонами;
- порівнюються прямі $P_k^{(j)}$ та композиційні $P_{comp,k}^{(j)}$ оцінки показників з метою перевірки того, що композиційні оцінки 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 і, можливо, побудови іншої моделі.

Середньоквадратичні похибки оцінок показників бідності на національному рівні $MSE(P_{comp,U}^{(j)})$ визначаються на основі середньоквадратичних похибок оцінок показників для регіонів за формулою:

$$MSE(P_{comp,U}^{(j)}) = \frac{1}{N^2} \sum_{k=1}^K N_k^2 MSE(P_{comp,k}^{(j)}). \quad (8)$$

Похідні характеристики надійності оцінок показників бідності на національному рівні розраховуються на основі формул (5)–(7).

3. Статистико-математичні моделі для непрямого оцінювання визначених показників бідності

Для непрямого оцінювання визначених показників бідності для регіонів України у певному році як основна застосовується математико-статистична модель (1), яка описується тривимірною композиційною оціночною функцією і в розгорнутому вигляді для k -го регіону визначається за формулою:

$$P_{comp,k}^{(j)} = (1 - \hat{\gamma}_{1k}^{(j)}) \cdot P_k^{(j)} - \hat{\gamma}_{2k}^{(j)} \cdot x_{1k}^{(j)} - \hat{\gamma}_{3k}^{(j)} \cdot x_{2k}^{(j)} + \hat{\gamma}_{1k}^{(j)} \cdot P^{(j)}, \quad j = 1, 2, 3. \quad (9)$$

де $P_{comp,k}^{(j)}$ - оцінка j -го показника бідності для k -го регіону за композиційною оціночною функцією в році, що розглядається;

$P_k^{(j)}$ - пряма оцінка j -го показника бідності для k -го регіону в році, що розглядається;

$P^{(j)}$ - пряма оцінка j -го показника бідності для національного рівня;

$x_{1,k}^{(j)}$ - пряма оцінка j -го показника бідності для k -го регіону в попередньому році;

$x_{2,k}^{(j)}$ - величина кінцевих споживчих витрат домогосподарств у розрахунку на одну особу за даними національних рахунків у році, що розглядається;

$\hat{\gamma}_{1k}^{(j)}, \hat{\gamma}_{2k}^{(j)}, \hat{\gamma}_{3k}^{(j)}$ - вагові коефіцієнти композиційної оціночної функції для k -го регіону при оцінюванні j -го показника бідності;

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».

Таким чином, у композиційній оцінчій функції як зовнішня інформація застосовуються такі дані: пряма оцінка відповідного показника бідності для національного рівня; розподіл прямих оцінок відповідного показника бідності за

регіонами у попередньому році; розподіл за регіонами кінцевих споживчих витрат домогосподарств у розрахунку на одну особу за даними національних рахунків у році, що розглядається.

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

Параметри моделей для показників оцінюються за умови мінімізації середньоквадратичної похибки композиційної оціночної функції на основі співвідношення (2). Оскільки це співвідношення містить об'єкти (варіаційні та коваріаційні матриці і вектори), оцінені за результатами вибіркового обстеження, для визначення параметрів моделей, композиційних оцінок і середньоквадратичних похибок оцінювання застосовується реплікаційний підхід. За цим підходом генерується значна кількість вибірок дизайну, аналогічного дизайну вибірки в ОУЖД, для кожної реплікації (кожної зі згенерованих вибірок) оцінюються параметри композиційної оціночної функції та власне рівні бідності [1]. На основі всіх реплікацій визначаються статистичні властивості розподілу реплікаційних оцінок – середні значення параметрів моделі та композиційних оцінок показників, зміщення композиційних оцінок і середньоквадратичні похибки.

Для реалізації реплікацій розроблено спеціальне програмне забезпечення на основі пакета програм R, який підтримується середовищем SPSS. Кількість реплікацій, передбачена програмним забезпеченням, становить 5000. Відповідна процедура є складовою розробленого програмного забезпечення для оцінювання показників бідності на основі математико-статистичних моделей.

При застосуванні моделей для підвищення надійності оцінювання визначених показників бідності доцільно враховувати, що на практиці композиційна оціночна функція для кожного показника будується окремо. Доцільно навести моделі для кожного визначеного показника бідності.

У додатку А1 як приклад наведено початкову інформацію для оцінювання визначених показників бідності у 2009 році на основі тривимірної композиційної оціночної функції.

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

$$P_{comp,k}^{(1)} = (1 - \hat{\gamma}_{1k}^{(1)}) \cdot P_k^{(1)} - \hat{\gamma}_{2k}^{(1)} \cdot x_{1k}^{(1)} - \hat{\gamma}_{3k}^{(1)} \cdot x_{2k}^{(1)} + \hat{\gamma}_{1k}^{(1)} \cdot P^{(1)}, \quad (10)$$

де $P_{comp,k}^{(1)}$ - композиційна оцінка рівня бідності за національною межею для k -го регіону в році, що розглядається;

$P_k^{(1)}$ - пряма оцінка рівня бідності за національною межею для k -го регіону в році, що розглядається;

$P^{(1)}$ - пряма оцінка рівня бідності за національною межею для національного рівня в році, що розглядається;

$x_{1,k}^{(1)}$ - пряма оцінка рівня бідності за національною межею для k -го регіону в попередньому році;

$x_{2,k}^{(1)}$ - величина кінцевих споживчих витрат домогосподарств у розрахунку на одну особу за даними національних рахунків у році, що розглядається.

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

$$P_{comp,k}^{(2)} = (1 - \hat{\gamma}_{1k}^{(2)}) \cdot P_k^{(2)} - \hat{\gamma}_{2k}^{(2)} \cdot x_{1k}^{(2)} - \hat{\gamma}_{3k}^{(2)} \cdot x_{2k}^{(2)} + \hat{\gamma}_{1k}^{(2)} \cdot P^{(2)}, \quad (11)$$

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

$P_k^{(2)}$ - пряма оцінка рівня бідності за регіональною межею для k -го регіону в році, що розглядається;

$P^{(2)}$ - пряма оцінка рівня бідності за регіональною межею для національного рівня в році, що розглядається;

$x_{1,k}^{(2)}$ - пряма оцінка рівня бідності за регіональною межею для k -го регіону в попередньому році;

$x_{2,k}^{(j)}$ - величина кінцевих споживчих витрат домогосподарств у розрахунку на одну особу за даними національних рахунків у році, що розглядається.

Тривимірна композиційна оціночна функція для підвищення надійності оцінювання рівня бідності за прожитковим мінімумом.

$$P_{comp,k}^{(3)} = (1 - \hat{\gamma}_{1k}^{(3)}) \cdot P_k^{(3)} - \hat{\gamma}_{2k}^{(3)} \cdot x_{1k}^{(3)} - \hat{\gamma}_{3k}^{(3)} \cdot x_{2k}^{(3)} + \hat{\gamma}_{1k}^{(3)} \cdot P^{(3)}, \quad (12)$$

де $P_{comp,k}^{(3)}$ - оцінка рівня бідності за прожитковим мінімумом для k -го регіону за тривимірною композиційною оціночною функцією в році, що розглядається;

$P_k^{(3)}$ - пряма оцінка рівня бідності за прожитковим мінімумом для k -го регіону в році, що розглядається;

$P^{(3)}$ - пряма оцінка рівня бідності за прожитковим мінімумом для національного рівня в році, що розглядається;

$x_{1,k}^{(3)}$ - пряма оцінка рівня бідності за прожитковим мінімумом для k -го регіону в попередньому році;

$x_{2,k}^{(3)}$ - величина кінцевих споживчих витрат домогосподарств у розрахунку на одну особу за даними національних рахунків у році, що розглядається.

У додатку А2 наведено результати оцінки параметрів тривимірної композиційної оціночної функції та показників бідності у 2009 році.

4. Вибір зовнішньої інформації для композиційного оцінювання показників бідності

Слід зазначити, що при відборі зовнішньої інформації для використання у композиційній оціночній функції було проаналізовано статистичний взаємозв'язок із показниками бідності низки змінних. Крім оцінок показників бідності за регіонами за попередні роки і величини кінцевих споживчих витрат домогосподарств за даними національних рахунків, аналізувалися розподіли за регіонами середньомісячних сум доходів у вигляді заробітної плати, інших виплат і винагород, виплачених фізичним особам за місцем роботи та за місцем проживання у 2009 році за даними ДПА, середньомісячної заробітної плати штатних працівників (за даними поточного обстеження підприємств із питань статистики праці), наявних доходів домогосподарств у розрахунку на одну особу за даними національних рахунків, рівня зайнятості та безробіття за даними обстеження економічної активності населення, частки працівників, що отримують заробітну плату нижче мінімального рівня, середніх розмірів призначених пенсій з цільовою грошовою допомогою, валового регіонального продукту на душу населення та ін. Окремі результати такого аналізу представлені у матеріалах [6, 7]. На основі виконаного аналізу на цьому етапі впровадження моделей було встановлено доцільність використання саме зовнішньої інформації, визначеної для моделі (9).

Доцільно зауважити, що хоча в моделях виду (1) та (9) явно не використовуються регресійні моделі, їх застосування залишається необхідним на етапі підбору зовнішньої інформації.

Для вибору незалежних змінних при побудові композиційної оціночної функції для кожного показника бідності насамперед аналізуються парні коефіцієнти кореляції Пірсона прямих оцінок показників бідності з кожною зі змінних, яка використовується або потенційно може бути використана як зовнішня інформація [3]. Як правило, із сукупності змінних обираються лише ті змінні, коефіцієнти парної кореляції з якими перевищують 0,7 але в окремих випадках доцільно послабити цю вимогу і включити до аналізу також змінні з коефіцієнтами кореляції, які складають близько 0,5. Наприклад, у табл. 1. представлені коефіцієнти парної кореляції для окремих змінних, які випробувалися як зовнішня інформація для оцінювання визначених показників бідності.

Як видно з даних, наведених у табл. 1, взаємозв'язок рівнів бідності за 2009 і 2008 роки характеризується достатньо високими коефіцієнтами кореляції. Таким чином, показники за 2008 рік можуть бути обрані як зовнішня інформація. Але при побудові моделей слід урахувувати, що ці показники визначаються за даними вибіркового обстеження (ОУЖД) і, відповідно, характеризуються певною похибкою вибірки. При визначенні зовнішньої інформації за близьких значень коефіцієнтів кореляції у різних показників перевагу слід надавати показникам, які

визначаються за результатами суцільних спостережень або на основі адміністративної інформації, а не за результатами вибірових спостережень.

Усі інші розподіли з наведених, крім розподілу наявного доходу в розрахунку на одну особу, характеризуються близькими коефіцієнтами кореляції – на рівні 0,4–0,5. 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 найбільш прийнятним варіантом зовнішньої інформації на цьому етапі впровадження моделей є саме величина кінцевих споживчих витрат домогосподарств у розрахунку на одну особу за регіонами.

Показник наявного доходу у розрахунку на одну особу недоцільно використовувати як зовнішню інформацію з метою підвищення надійності оцінювання рівня бідності за національною межею, оскільки його взаємозв'язок з останнім значно слабший, ніж у інших показників.

Далі на основі сформованої сукупності зовнішніх змінних будується рівняння регресії й аналізується його адекватність. Основними характеристиками адекватності регресійної моделі є коефіцієнт детермінації R^2 , який змінюється від 0 до 1 і показує частку змін результативної ознаки (дисперсії результативної ознаки), яка пояснюється регресійною моделлю, і F - відношення, яке відображає ступінь взаємозв'язку між результативною і факторними ознаками. Чим ближчі значення R^2 до 1, тим краще зовнішні змінні здатні апроксимувати відповідні показники бідності. У загальному випадку, якщо значення R^2 є більшими за 0,7, то це свідчить про те, що незалежні змінні у достатній мірі пояснюють варіацію відповідного показника бідності і відповідна регресійна модель є адекватною. Слід зазначити, що значення R^2 , менші за 0,7, не свідчать про неможливість використання регресійної моделі.

Таблиця 1 – Коефіцієнти кореляції Пірсона між розподілами за регіонами показників бідності за національною межею та показників, які потенційно можуть бути використані як зовнішня інформація при побудові статистико-математичних моделей для підвищення надійності оцінювання показників бідності

	Рівень бідності за національною межею (2009), %	Рівень бідності за регіональною межею (2009), %	Рівень бідності за прожитковим мінімумом (2009), %
Рівень бідності за національною межею (2008)	0,808	0,720	0,634
Рівень бідності за регіональною межею (2008), %	0,690	0,683	0,483
Рівень бідності за прожитковим мінімумом (2008), %	0,705	0,597	0,607
Кінцеві споживчі витрати домогосподарств у розрахунку на одну особу за даними національних рахунків (2009)	-0,470	-0,400	-0,436

Середньомісячна заробітна плата фізичних осіб за місцем роботи (ДПА, 2009)	-0,502	-0,436	-0,472
Наявний дохід у розрахунку на одну особу (2009)	-0,389	-0,361	-0,383

Якщо розрахункове значення F - відношення перевищує порогове значення $F_{кр}$ (порогове значення визначається за спеціальними формулами або з довідників [3]), то це свідчить про наявність взаємозв'язку між розподілом показників бідності за регіонами та розподілами відповідних зовнішніх змінних. Крім того, як уже було зазначено, реалізований підхід (оціночна функція (1)) не передбачає використання регресійних моделей у явному вигляді, що в цілому послаблює вимоги щодо їх адекватності. Критерієм статистичної ефективності композиційної оціночної функції є, насамперед, зменшення середньоквадратичної похибки, що, у свою чергу, можливо за наявності статистичного взаємозв'язку між зовнішньою інформацією та показниками бідності.

На основі прямих оцінок показників бідності та зовнішніх даних доцільно визначати параметри відповідних регресійних моделей. Наприклад, для тривимірної композиційної оціночної функції (9) доцільно виконати аналіз таких регресійних моделей:

$$P_{reg,k}^{(j)} = \beta_0^{(j)} + \beta_1^{(j)} x_{1k}^{(j)} + \beta_2^{(j)} x_{2k}^{(j)}, \quad (13)$$

де $P_{reg,k}^{(j)}$ - оцінка j -го показника бідності за регресійною моделлю для k -го регіону;

$\beta_0^{(j)}, \beta_1^{(j)}, \beta_2^{(j)}$ - коефіцієнти регресії, однакові для всіх регіонів України;

$x_{1k}^{(j)}$ - рівень бідності у k -му регіоні в попередньому році;

$x_{2k}^{(j)}$ - кінцеві споживчі витрати домогосподарств у розрахунку на одну особу за даними національних рахунків для k -го регіону в попередньому році.

Приклад побудови моделей регресії для оцінок рівнів бідності за різними критеріями наведено в додатку Б.

Аналіз регресійних моделей забезпечує можливість перевірки ефективності використання зовнішньої інформації за наявності декількох зовнішніх змінних.

Доцільно зауважити, що, враховуючи відносно незначну кількість регіонів в Україні, для побудови математико-статистичних моделей недоцільно використовувати більше трьох зовнішніх змінних (уключаючи оцінку показника бідності на національному рівні). Таким чином, для побудови регресійних моделей доцільно використовувати не більше двох змінних.

Як зовнішні змінні, як правило, використовуються змінні, які відповідають найбільш адекватній регресійній моделі.

5. Супровід використання композиційних оціночних функцій

На цьому етапі впровадження моделей для підвищення надійності оцінювання показників бідності доцільним є їх застосовування для уточнення річних оцінок визначених показників бідності за регіонами України.

Для забезпечення функціонування моделей необхідно формувати та аналізувати інформаційну базу (див. п.2) в міру надходження інформації з визначених 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 парної кореляції розподілів за регіонами прямих оцінок кожного показника бідності за поточний і попередній рік.

За умови зниження значення коефіцієнта кореляції нижче критичного рівня (див. п.3 та [3]) приймається рішення щодо доцільності використання розподілу показників бідності за попередній рік як зовнішньої інформації у композиційній оціночній функції.

4. Визначаються коефіцієнти парної кореляції розподілів за регіонами прямих оцінок кожного показника бідності за поточний рік із 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 зниження значення коефіцієнта кореляції нижче критичного 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 і зовнішніх змінних, придатність яких для використання у композиційних оціночних функціях підтверджено за результатами аналізу, будуються й аналізуються регресійні моделі виду (13).
6. Для оцінок показників бідності за регресійними моделями перевіряється, чи потрапляють вони до довірчих інтервалів відповідних прямих оцінок [4]. 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 прийнято рішення щодо неадекватності регресійної моделі та здійснено пошук іншої регресійної моделі. За умови підтвердження адекватності регресійної моделі розраховуються композиційні оцінки визначених показників бідності за запропонованою методикою з використанням розробленого інструментарію.

7. Перевіряються значення оцінок вагових коефіцієнтів композиційної оціночної функції з метою підтвердження того, що значення потрапляють до визначеного інтервалу: $0 \leq |\hat{\gamma}_k| \leq 1$.
8. 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 прийнято рішення щодо неприйнятності композиційних оцінок як уточнених оцінок показників бідності.
9. Якщо надійність композиційних оцінок певного показника для окремих регіонів виявиться нижчою, ніж надійність прямих оцінок, користувачам слід надавати композиційні оцінки для забезпечення порівнянності даних. Якщо кількість регіонів, по яких 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), слід здійснити пошук більш ефективної моделі або відмовитися від використання композиційних оцінок, передбачених цією методикою.
10. Якщо виявиться, що окремі зовнішні змінні не можуть бути використані як зовнішня інформація, розроблений інструментарій дає можливість використання двомірних композиційних оціночних функцій або навіть одномірних композиційних оціночних функцій.

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

Результати моделювання мають надаватися користувачам у вигляді окремої публікації, до складу якої доцільно включити методологічні пояснення. Проект методологічних пояснень наведено в додатку В.

ДОДАТКИ

Додаток А1

Таблиця А1.1 – Початкова інформація, необхідна для розрахунків оцінок рівня бідності за національною межею на основі тривимірної композиційної оціночної функції

Регіон	Пряма оцінка рівня бідності за національною межею, 2009р.,%	Коефіцієнт варіації прямої оцінки рівня бідності за національною межею, 2009р., %	Пряма оцінка рівня бідності за національною межею, 2008р., %	Коефіцієнт варіації прямої оцінки рівня бідності за національною межею, 2008р., %	Кінцеві споживчі витрати домогосподарств у розрахунку на одну особу, 2009р., грн.
Україна	26,37	2,63	26,97	2,41	12631,73
АР Крим	25,84	15,20	24,89	9,21	12083,61
Вінницька	22,45	14,23	24,3	13,86	10503,23
Волинська	42,09	8,18	47,3	10,21	10377,27
Дніпропетровська	26,35	9,68	27,22	13,01	12435,14
Донецька	23,62	11,11	25,58	9,66	12372,42
Житомирська	35,73	14,95	34,22	14,72	9696,90
Закарпатська	23,03	23,08	20,3	22,07	9328,03
Запорізька	19,72	14,02	24,39	15,47	12776,77
Івано-Франківська	23,10	16,38	20,09	18,93	9764,65
Київська	20,81	11,60	21,17	23,63	12177,06
Кіровоградська	42,01	10,01	47,09	9,14	9716,35
Луганська	28,15	10,36	29,2	8,74	9629,15
Львівська	31,39	7,87	29,25	8,49	10583,23
Миколаївська	20,89	16,30	19,81	16,35	11500,80
Одеська	34,17	12,56	38,48	12,04	14258,24
Полтавська	28,40	14,72	27,47	12,36	11020,92
Рівненська	36,85	15,31	39,28	11,51	9385,04
Сумська	45,81	8,70	29,07	12,16	9211,51
Тернопільська	29,44	15,99	42,77	10,93	8483,18
Харківська	20,74	17,18	20,59	9,9	14563,17
Херсонська	50,28	8,60	32,96	12,8	10261,79
Хмельницька	31,02	10,59	29,79	7,79	10224,27
Черкаська	13,45	18,43	18,37	20,64	10795,57
Чернівецька	18,48	13,72	29,6	15,88	9531,08
Чернігівська	22,44	18,64	28,62	11,88	9489,91
м. Київ	9,51	17,83	9,63	17,11	33031,74
м. Севастополь	5,00	58,47	2,7	73,05	15340,52

Таблиця А1.2 - Початкова інформація, необхідна для розрахунків оцінок рівня бідності за регіональною межею на основі тривимірної композиційної оціночної функції

Регіон	Пряма оцінка рівня бідності за регіональною межею, 2009р.,%	Коефіцієнт варіації прямої оцінки рівня бідності за регіональною межею, 2009р., %	Пряма оцінка рівня бідності за регіональною межею, 2008р., %	Коефіцієнт варіації прямої оцінки рівня бідності за регіональною межею, 2008р., %	Кінцеві споживчі витрати домогосподарств у розрахунку на одну особу, 2009р., грн.
Україна	31,73	1,99	32,62	2,15	12631,73
АР Крим	34,35	11,32	34,36	6,96	12083,61
Вінницька	26,81	12,84	27,14	12,71	10503,23
Волинська	42,09	8,18	47,74	9,93	10377,27
Дніпропетровська	32,26	8,13	32,25	11,48	12435,14
Донецька	30,65	10,71	32,16	7,95	12372,42
Житомирська	41,78	14,72	39,14	11,29	9696,90
Закарпатська	33,01	20,44	24,00	19,93	9328,03
Запорізька	21,40	12,88	30,35	12,91	12776,77
Івано-Франківська	29,55	13,88	24,44	14,81	9764,65
Київська	32,17	8,90	35,13	16,69	12177,06
Кіровоградська	42,48	10,10	47,09	9,14	9716,35
Луганська	30,75	9,12	34,32	9,77	9629,15
Львівська	37,00	5,36	38,66	8,12	10583,23
Миколаївська	21,72	15,79	19,81	16,35	11500,80
Одеська	40,79	10,44	43,99	10,24	14258,24
Полтавська	30,41	13,56	32,48	10,60	11020,92
Рівненська	38,58	13,15	39,61	11,93	9385,04
Сумська	46,11	8,36	30,25	11,85	9211,51
Тернопільська	32,31	15,46	46,50	10,10	8483,18
Харківська	31,58	9,08	31,23	9,00	14563,17
Херсонська	50,28	8,60	32,96	12,80	10261,79
Хмельницька	30,59	10,72	32,79	8,71	10224,27
Черкаська	19,65	13,68	21,44	22,08	10795,57
Чернівецька	23,66	9,75	35,78	12,91	9531,08
Чернігівська	25,77	21,71	31,13	10,88	9489,91
м. Київ	18,50	11,93	19,36	11,09	33031,74
м. Севастополь	15,79	32,33	14,59	56,86	15340,52

**Таблиця А1.3 - Початкова інформація, необхідна для розрахунків оцінок рівня бідності
за прожитковим мінімумом на основі
тривимірної композиційної оціночної функції**

Регіон	Пряма оцінка рівня бідності за прожитковим мінімумом, 2009р.,%	Коефіцієнт варіації прямої оцінки рівня бідності за прожитковим мінімумом, 2009р., %	Пряма оцінка рівня бідності за прожитковим мінімумом, 2008р.,%	Коефіцієнт варіації прямої оцінки рівня бідності за прожитковим мінімумом, 2008р., %	Кінцеві споживчі витрати домогоспо- дарств у розрахунку на одну особу, 2009р., грн.
Україна	5,77	6,66	7,1	5,4	12631,73
АР Крим	2,91	29,58	7,8	26,3	12083,61
Вінницька	8,13	21,11	6,9	21,7	10503,23
Волинська	10,80	22,39	15,6	22,4	10377,27
Дніпропетровська	3,19	24,07	5,6	22,6	12435,14
Донецька	5,39	21,01	4,1	24,4	12372,42
Житомирська	7,98	27,19	9,6	22,3	9696,90
Закарпатська	4,94	38,21	8,3	39	9328,03
Запорізька	2,12	28,52	6,8	33,7	12776,77
Івано-Франківська	5,03	32,49	3,5	43,1	9764,65
Київська	2,06	16,68	5	51,9	12177,06
Кіровоградська	11,63	17,24	21,3	14,9	9716,35
Луганська	5,74	19,33	6,4	21,8	9629,15
Львівська	9,02	24,00	7,1	19,1	10583,23
Миколаївська	3,62	30,80	6,7	18,6	11500,80
Одеська	10,88	16,91	10,7	21,9	14258,24
Полтавська	5,99	28,24	6,9	28,4	11020,92
Рівненська	5,48	28,77	12,9	19	9385,04
Сумська	12,52	28,18	4,9	25,1	9211,51
Тернопільська	7,72	27,79	10,1	39,3	8483,18
Харківська	1,73	43,26	2,8	28,2	14563,17
Херсонська	15,58	19,14	12,3	18,9	10261,79
Хмельницька	5,70	35,60	11,3	22,4	10224,27
Черкаська	6,04	34,08	6,2	24,9	10795,57
Чернівецька	3,11	30,20	7	32,1	9531,08
Чернігівська	5,09	25,28	10	17,5	9489,91
м. Київ	1,75	43,72	1,1	72,9	33031,74
м. Севастополь	2,90	70,87	0,8	62,3	15340,52

Додаток А2

**Таблиця А2.1 – Вагові коефіцієнти тривимірної
композиційної оціночної функції для оцінювання рівня бідності
за національною межею у 2009 році**

Регіон	$\hat{\gamma}_{1k}^{(1)}$	$\hat{\gamma}_{2k}^{(1)}$	$\hat{\gamma}_{3k}^{(1)}$
АР Крим	0,09300561	-0,07523291	0,00006267
Вінницька	0,09581386	-0,07723444	0,00006574
Волинська	0,15933207	-0,12264451	0,00013235
Дніпропетровська	0,06521414	-0,05366681	0,00004026
Донецька	0,05770342	-0,04755946	0,00003531
Житомирська	0,16748508	-0,13140362	0,00012945
Закарпатська	0,11934080	-0,09517026	0,00008613
Запорізька	0,08747046	-0,07065258	0,00005931
Івано-Франківська	0,12162910	-0,09688601	0,00008818
Київська	0,11055270	-0,08870614	0,00007746
Кіровоградська	0,16781983	-0,12933918	0,00013882
Луганська	0,08808632	-0,07160941	0,00005801
Львівська	0,08401698	-0,06829452	0,00005530
Миколаївська	0,09621033	-0,07682552	0,00006875
Одеська	0,13328773	-0,10587054	0,00009751
Полтавська	0,13141513	-0,10555942	0,00009156
Рівненська	0,15472720	-0,11964522	0,00012649
Сумська	0,17912902	-0,14088772	0,00013692
Тернопільська	0,15508771	-0,11821101	0,00013379
Харківська	0,06689039	-0,05487685	0,00004189
Херсонська	0,13675452	-0,10712791	0,00010612
Хмельницька	0,15253579	-0,12028666	0,00011543
Черкаська	0,09329502	-0,07546882	0,00006282
Чернівецька	0,10965102	-0,08429534	0,00009178
Чернігівська	0,12400203	-0,09856659	0,00009108
м. Київ	0,04555342	-0,03774696	0,00002606
м. Севастополь	0,07338284	-0,05921414	0,00004897

**Таблиця А2.2 – Результати композиційного оцінювання рівня бідності
за національною межею у 2009 році на основі тривимірної
композиційної оціночної функції**

Регіон	Пряма оцінка, % $P_k^{(1)}$	Коефіцієнт варіації прямої оцінки, % $CV(P_k^{(1)})$	Композиційна оцінка, % $P_{comp,k}^{(1)}$	Гранична похибка композиційної оцінки, % $ME(P_{comp,k}^{(1)})$	Відносна похибка композиційної оцінки, % $RRMSE(P_{comp,k}^{(1)})$
Україна	26,37	2,63	26,23	1,27	1,61
АР Крим	25,84	15,20	25,79	3,84	7,60
Вінницька	22,45	14,23	22,67	3,94	8,87
Волинська	42,09	8,18	41,04	5,62	6,99
Дніпропетровська	26,35	9,68	26,42	3,19	6,15
Донецька	23,62	11,11	23,68	3,05	6,56
Житомирська	35,73	14,95	35,09	5,51	8,01
Закарпатська	23,03	23,08	23,13	4,45	9,81
Запорізька	19,72	14,02	19,90	3,70	9,48
Івано-Франківська	23,10	16,38	23,13	4,41	9,73
Київська	20,81	11,60	21,05	4,31	10,44
Кіровоградська	42,01	10,01	40,91	5,79	7,22
Луганська	28,15	10,36	28,11	3,65	6,63
Львівська	31,39	7,87	31,27	3,55	5,79
Миколаївська	20,89	16,30	21,08	3,95	9,56
Одеська	34,17	12,56	33,77	4,76	7,19
Полтавська	28,40	14,72	28,26	4,53	8,17
Рівненська	36,85	15,31	36,24	5,20	7,32
Сумська	45,81	8,70	44,36	6,15	7,07
Тернопільська	29,44	15,99	29,21	5,22	9,11
Харківська	20,74	17,18	20,82	3,24	7,94
Херсонська	50,28	8,60	49,01	5,47	5,69
Хмельницька	31,02	10,59	30,70	4,99	8,29
Черкаська	13,45	18,43	13,93	4,01	14,69
Чернівецька	18,48	13,72	18,89	4,41	11,92
Чернігівська	22,44	18,64	22,57	4,47	10,10
м. Київ	9,51	17,83	9,73	2,71	14,20
м. Севастополь	5,00	58,47	5,55	3,80	35,00

**Таблиця А2.3 – Вагові коефіцієнти тривимірної
композиційної оціночної функції для оцінювання рівня бідності
за регіональною межею у 2009 році**

Регіон	$\hat{\gamma}_{1k}^{(2)}$	$\hat{\gamma}_{2k}^{(2)}$	$\hat{\gamma}_{3k}^{(2)}$
АР Крим	0,10691829	-0,06842331	0,00010997
Вінницька	0,10648193	-0,06826821	0,00010924
Волинська	0,16347234	-0,09898944	0,00018293
Дніпропетровська	0,07185254	-0,04740609	0,00007015
Донецька	0,06637654	-0,04383271	0,00006466
Житомирська	0,17848355	-0,11020998	0,00019418
Закарпатська	0,14573208	-0,09159924	0,00015440
Запорізька	0,09224256	-0,05905811	0,00009468
Івано-Франківська	0,14074499	-0,08824521	0,00014963
Київська	0,14033393	-0,08774472	0,00014970
Кіровоградська	0,17240564	-0,10456298	0,00019253
Луганська	0,09192599	-0,05960788	0,00009253
Львівська	0,08974179	-0,05804417	0,00009069
Миколаївська	0,09953705	-0,06327125	0,00010348
Одеська	0,14281542	-0,08986085	0,00015079
Полтавська	0,13568007	-0,08641340	0,00014075
Рівненська	0,16085258	-0,09796388	0,00017860
Сумська	0,18154552	-0,11288089	0,00019550
Тернопільська	0,16629530	-0,09917517	0,00019033
Харківська	0,08492098	-0,05534694	0,00008464
Херсонська	0,13916717	-0,08624578	0,00015049
Хмельницька	0,15318850	-0,09535475	0,00016475
Черкаська	0,12169905	-0,07823335	0,00012424
Чернівецька	0,13147766	-0,07856930	0,00015017
Чернігівська	0,13555975	-0,08529790	0,00014342
м. Київ	0,07560404	-0,04959339	0,00007319
м. Севастополь	0,16738141	-0,09476159	0,00020269

**Таблиця А2.4 – Результати композиційного оцінювання рівня бідності
за регіональною межею у 2009 році на основі тривимірної
композиційної оціночної функції**

Регіон	Пряма оцінка, % $P_k^{(2)}$	Коефіцієнт варіації прямої оцінки, % $CV(P_k^{(2)})$	Композиційна оцінка, % $P_{comp,k}^{(2)}$	Гранична похибка композиційної оцінки, % $ME(P_{comp,k}^{(2)})$	Відносна похибка композиційної оцінки, % $RRMSE(P_{comp,k}^{(2)})$
Україна	31,73	1,99	31,64	1,16	1,41
АР Крим	34,35	11,32	34,15	4,09	6,11
Вінницька	26,81	12,84	27,12	4,07	7,65
Волинська	42,09	8,18	41,08	5,54	6,88
Дніпропетровська	32,26	8,13	32,22	3,28	5,20
Донецька	30,65	10,71	30,67	3,25	5,40
Житомирська	41,78	14,72	40,73	5,57	6,98
Закарпатська	33,01	20,44	32,90	4,73	7,33
Запорізька	21,40	12,88	21,90	3,92	9,13
Івано-Франківська	29,55	13,88	29,69	4,59	7,88
Київська	32,17	8,90	32,14	4,61	7,31
Кіровоградська	42,48	10,10	41,31	5,63	6,95
Луганська	30,75	9,12	30,80	3,72	6,17
Львівська	37,00	5,36	36,68	3,80	5,29
Миколаївська	21,72	15,79	22,28	4,15	9,50
Одеська	40,79	10,44	40,03	4,90	6,25
Полтавська	30,41	13,56	30,49	4,54	7,60
Рівненська	38,58	13,15	37,98	5,06	6,80
Сумська	46,11	8,36	44,60	6,14	7,03
Тернопільська	32,31	15,46	32,26	5,08	8,04
Харківська	31,58	9,08	31,59	3,61	5,83
Херсонська	50,28	8,60	48,74	5,61	5,87
Хмельницька	30,59	10,72	30,71	4,81	8,00
Черкаська	19,65	13,68	20,46	4,69	11,71
Чернівецька	23,66	9,75	24,25	4,78	10,06
Чернігівська	25,77	21,71	26,19	4,56	8,88
м. Київ	18,50	11,93	19,05	3,68	9,86
м. Севастополь	15,79	32,33	17,38	6,62	19,34

**Таблиця А2.5 – Вагові коефіцієнти тривимірної
композиційної оціночної функції для оцінювання рівня бідності
за прожитковим мінімумом у 2009 році**

Регіон	$\hat{\gamma}_{1k}^{(3)}$	$\hat{\gamma}_{2k}^{(3)}$	$\hat{\gamma}_{3k}^{(3)}$
АР Крим	0,07313332	-0,03375318	0,00002404
Вінницька	0,18080702	-0,08329046	0,00005981
Волинська	0,26640605	-0,10618640	0,00011391
Дніпропетровська	0,05312848	-0,02561195	0,00001574
Донецька	0,07821298	-0,03810786	0,00002266
Житомирська	0,24260775	-0,10547352	0,00009076
Закарпатська	0,14738851	-0,06477325	0,00005393
Запорізька	0,08259129	-0,03825973	0,00002685
Івано-Франківська	0,15932580	-0,07454500	0,00005161
Київська	0,09957752	-0,04631673	0,00003227
Кіровоградська	0,28819785	-0,10984653	0,00013146
Луганська	0,11250015	-0,05315395	0,00003546
Львівська	0,14575145	-0,06855882	0,00004621
Миколаївська	0,10050709	-0,04513240	0,00003500
Одеська	0,24445181	-0,10825496	0,00008669
Полтавська	0,16949400	-0,07810442	0,00005613
Рівненська	0,16803163	-0,06921826	0,00006867
Сумська	0,32568266	-0,14927288	0,00011006
Тернопільська	0,23815559	-0,09895207	0,00009662
Харківська	0,06033676	-0,02941561	0,00001743
Херсонська	0,28695511	-0,12029117	0,00011359
Хмельницька	0,17972705	-0,07689583	0,00006888
Черкаська	0,19440615	-0,08954030	0,00006470
Чернівецька	0,11810844	-0,05098283	0,00004456
Чернігівська	0,16113336	-0,07094352	0,00005882
м. Київ	0,07668574	-0,03718533	0,00002269
м. Севастополь	0,21352450	-0,09222268	0,00007872

**Таблиця А2.6 – Результати композиційного оцінювання рівня бідності
за прожитковим мінімумом у 2009 році на основі тривимірної
композиційної оціночної функції**

Регіон	Пряма оцінка, % $P_k^{(3)}$	Коефіцієнт варіації прямої оцінки, % $CV(P_k^{(3)})$	Композиційна оцінка, % $P_{comp,k}^{(3)}$	Гранична похибка композиційної оцінки, % $ME(P_{comp,k}^{(3)})$	Відносна похибка композиційно ї оцінки, % $RRMSE(P_{comp,k}^{(3)})$
Україна	5,77	6,66	5,57	0,71	3,97
АР Крим	2,91	29,58	3,04	1,50	25,13
Вінницька	8,13	21,11	7,83	2,33	15,16
Волинська	10,80	22,39	9,79	3,34	17,42
Дніпропетровська	3,19	24,07	3,26	1,29	20,14
Донецька	5,39	21,01	5,41	1,58	14,91
Житомирська	7,98	27,19	7,57	2,65	17,85
Закарпатська	4,94	38,21	5,00	2,15	21,90
Запорізька	2,12	28,52	2,29	1,35	29,94
Івано-Франківська	5,03	32,49	5,06	2,18	21,96
Київська	2,06	16,68	2,30	1,52	33,72
Кіровоградська	11,63	17,24	10,33	3,77	18,60
Луганська	5,74	19,33	5,72	1,86	16,62
Львівська	9,02	24,00	8,69	2,17	12,74
Миколаївська	3,62	30,80	3,74	1,78	24,20
Одеська	10,88	16,91	10,00	3,11	15,89
Полтавська	5,99	28,24	5,93	2,26	19,41
Рівненська	5,48	28,77	5,46	2,23	20,86
Сумська	12,52	28,18	10,94	4,25	19,84
Тернопільська	7,72	27,79	7,34	2,66	18,51
Харківська	1,73	43,26	1,88	1,05	28,57
Херсонська	15,58	19,14	13,52	4,98	18,80
Хмельницька	5,70	35,60	5,68	2,34	21,03
Черкаська	6,04	34,08	5,95	2,37	20,33
Чернівецька	3,11	30,20	3,28	1,92	29,79
Чернігівська	5,09	25,28	5,11	2,21	22,11
м. Київ	1,75	43,72	1,94	1,18	31,00
м. Севастополь	2,90	70,87	3,27	2,40	37,42

Додаток Б

Приклад побудови регресійних моделей для оцінки придатності зовнішньої інформації для використання при побудові статистико-математичних моделей

Оцінка рівня бідності за національною межею

При композиційному оцінюванні рівня бідності за національною межею у 2009 році регресійна модель для початкових даних, наведених у табл. А1.1, має вид:

$$P_{regr,k}^{(1)} = 7.300 + 0.801 \cdot P_{k,2008}^{(1)} - 0.195 \cdot 10^{-3} \cdot x_{2k}^{(1)}, \quad (Б.1)$$

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.

Характеристики адекватності регресійної моделі (Б.1):

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

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

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

$$P_{regr,k}^{(1)} = 1.009 + 0.986 \cdot P_{k,2007}^{(1)} - 0.060 \cdot 10^{-3} \cdot x_{2k}^{(1)}, \quad (Б.2)$$

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 (Б.2):

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

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 (табл. А1.1) замість величини кінцевих споживчих витрат домогосподарств у розрахунку на одну особу за 2009 рік використати величину середньомісячної заробітної плати фізичних осіб за місцем роботи (ДПА, 2009), то модель матиме вид:

$$P_{reg,k}^{(1)} = 9.021 + 0.802 \cdot P_{k,2008}^{(1)} - 0.260 \cdot 10^{-3} \cdot x_{2k}^{(1)}, \quad (Б.3)$$

де $P_{reg,k}^{(1)}$ – 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 (Б.5):

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

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, регресійна модель (Б.3) є певною мірою менш адекватною, ніж модель (Б.1), для якої вищими є і $R^2 = 0,658$, і $F = 24,00$.

Оцінка рівня бідності за регіональною межею

При композиційному оцінюванні рівня бідності за регіональною межею у 2009 році регресійна модель для початкових даних, наведених у табл. А1.2, має вид:

$$P_{reg,k}^{(2)} = 14.70 + 0.633 \cdot P_{k,2008}^{(2)} - 0.293 \cdot 10^{-3} \cdot x_{2k}^{(2)}, \quad (Б.4)$$

де $P_{reg,k}^{(2)}$ – 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 (Б.4):

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

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

Оцінка рівня бідності за прожитковим мінімумом

При композиційному оцінюванні рівня бідності за прожитковим мінімумом у 2009 році регресійна модель для початкових даних, наведених у табл. А1.3, має вид:

$$P_{reg,k}^{(3)} = 3.729 + 0.457 \cdot P_{k,2008}^{(3)} - 0.095 \cdot 10^{-3} \cdot x_{2k}^{(3)}, \quad (Б.5)$$

де $P_{reg,k}^{(3)}$ – 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 (Б.5):

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

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

Додаток В

Проект методологічних пояснень щодо оцінювання показників бідності з використанням методів математико-статистичного моделювання

У статистичному збірнику представлені показники бідності для національного та регіонального рівнів, отримані на основі застосування спеціального підходу до оцінювання показників на основі методів статистичного моделювання. Метою застосування цих методів є підвищення надійності оцінювання показників бідності на регіональному рівні.

Загальний опис проблеми

Показники бідності в Україні вимірюються на основі даних щоквартального державного вибіркового обстеження умов життя домогосподарств (далі – ОУЖД) у відповідності із чинними методичними та нормативними документами. Це обстеження проводиться щороку на постійній основі, починаючи з 1999р., і відповідає міжнародним стандартам.

Статистичний аналіз характеристик надійності показників бідності, визначених за даними ОУЖД, засвідчив, що прямі оцінки показників на рівні окремих регіонів вимірюються з недостатньою точністю. Це призводить до нестабільності оцінок у динаміці, ускладнює порівняння показників і в цілому обмежує можливість ефективного моніторингу бідності в Україні.

Ефективним способом підвищення надійності показників бідності для різних рівнів агрегації даних є застосування спеціальних методологічних підходів до їх оцінювання, що базуються на методах непрямого оцінювання. Застосування таких методів, як правило, здійснюється для оцінювання лише найбільш важливих показників, оскільки вимагає реалізації додаткових складних технологічних процедур [1,2].

Система показників

У публікації представлені результати непрямого оцінювання таких показників бідності:

1) Рівень бідності за національною межею $P_k^{(1)}$ – питома вага сімей (домогосподарств), у яких рівень споживання (доходів) на одну особу є нижчим від визначеної межі бідності. Рівень бідності розраховується за формулою:

$$P_k^{(1)} = \frac{Q_k^{(1)}}{N_k} \cdot 100\%, \quad (1)$$

де $Q_k^{(1)}$ - чисельність населення k -го регіону, що визнається бідним відповідно до національної межі бідності;

N_k - загальна чисельність населення k -го регіону.

Межа бідності визначається на підставі відносного критерію зарахування різних верств населення до категорії бідних, який розраховується за фіксованою часткою середньодушового доходу (витрат) – 75% медіанного рівня сукупних доходів (витрат) у розрахунку на умовного дорослого.

2) Рівень бідності за регіональною межею $P_k^{(2)}$ – частка населення регіону, доходи яких є нижчими за розраховану регіональну межу бідності. Цей рівень бідності розраховується за формулою:

$$P_k^{(2)} = \frac{Q_k^{(2)}}{N_k} \cdot 100\%, \quad (2)$$

де $Q_k^{(2)}$ - чисельність населення k -го регіону, що визнається бідним відповідно до регіональної межі бідності;

N_k - загальна чисельність населення k -го регіону.

3) Рівень бідності за розміром прожиткового мінімуму $P_k^{(3)}$ – частка населення з загальними середньомісячними еквівалентними доходами, нижчими за прожитковий мінімум у середньому на одну особу. Цей рівень бідності визначається за формулою:

$$P_k^{(3)} = \frac{Q_k^{(3)}}{N_k} \cdot 100\%, \quad (3)$$

де $Q_k^{(3)}$ - чисельність населення k -го регіону з загальними середньомісячними еквівалентними доходами, нижчими за прожитковий мінімум;

N_k - загальна чисельність населення k -го регіону.

Наведений перелік показників бідності, для оцінки яких доцільно застосувати методи непрямого оцінювання, визначено спеціально створеною робочою групою, яка включала провідних експертів з питань моніторингу бідності.

Рівень оцінювання

Підвищення рівня надійності визначених показників бідності забезпечується для регіонів України. Водночас на основі уточнених оцінок показників для регіонів розраховуються уточнені оцінки показників для рівня економічних районів та України в цілому. Це пояснюється доцільністю забезпечення узгодженості оцінок показників на різних рівнях агрегації даних.

Методи оцінювання показників

Для непрямого оцінювання показників бідності на рівні регіонів України застосовано композиційну оціночну функцію, яка визначається за формулою:

$$P_{comp,k}^{(j)} = (1 - \hat{\gamma}_{1k}^{(j)}) \cdot P_k^{(j)} - \hat{\gamma}_{2k}^{(j)} \cdot x_{1k}^{(j)} - \hat{\gamma}_{3k}^{(j)} \cdot x_{2k}^{(j)} + \hat{\gamma}_{1k}^{(j)} \cdot P^{(j)}, \quad j = 1, 2, 3. \quad (4)$$

де $P_{comp,k}^{(j)}$ - оцінка j -го показника бідності для k -го регіону за композиційною оціночною функцією у році, що розглядається;

$P_k^{(j)}$ - пряма оцінка j -го показника бідності для k -го регіону у році, що розглядається;

$P^{(j)}$ - пряма оцінка j -го показника бідності для національного рівня;

$x_{1k}^{(j)}$ - пряма оцінка j -го показника бідності для k -го регіону в попередньому році;

$x_{2,k}^{(j)}$ - величина кінцевих споживчих витрат домогосподарств у розрахунку на одну особу за даними національних рахунків у році, що розглядається;

$\hat{\gamma}_{1k}^{(j)}, \hat{\gamma}_{2k}^{(j)}, \hat{\gamma}_{3k}^{(j)}$ - вагові коефіцієнти композиційної оціночної функції для k -го регіону при оцінюванні j -го показника бідності;

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».

Для забезпечення узгодженості даних на різних рівнях їх агрегації на основі уточнених оцінок показників регіонального рівня розраховуються уточнені оцінки показників для рівня економічних районів та для національного рівня. Слід зауважити, що ці оцінки можуть дещо відрізнятися від прямих оцінок, розрахованих безпосередньо за даними обстеження умов життя домогосподарств.

При визначенні уточненого рівня бідності $P_{comp,U}^{(j)}$ за будь-яким з визначених критеріїв на національному рівні застосовується така формула:

$$P_{comp,U}^{(j)} = \frac{\sum_{k=1}^K N_k P_{comp,k}^{(j)}}{\sum_{k=1}^K N_k}, \quad (5)$$

N_k - чисельність населення k -го регіону.

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

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

N_k – середньорічна чисельність постійного населення k -го регіону.

Для економічних районів уточнені оцінки показників бідності розраховуються за аналогічними формулами, але лише з урахуванням інформації по тих регіонах, які входять до складу відповідних економічних районів.

Оцінка надійності показників бідності

Основною характеристикою надійності непрямих оцінок показників бідності є величина середньоквадратичної похибки $MSE(P_{comp,k}^{(j)})$, яка визначається з використанням спеціальних статистичних процедур [2]. На основі оцінок величини $MSE(P_{comp,k}^{(j)})$ розраховуються величини відношення кореня квадратного від середньоквадратичної похибки до композиційної оцінки показника бідності за регіонами:

$$RRMSE(P_{comp,k}^{(j)}) = \frac{\sqrt{MSE(P_{comp,k}^{(j)})}}{P_{comp,k}^{(j)}} \cdot 100\%. \quad (6)$$

Величина $RRMSE(P_{comp,k}^{(j)})$ є аналогом коефіцієнтів варіації для прямих оцінок і може бути прийнята за основний критерій надійності результатів оцінювання показників бідності для регіонів України. Непряма оцінка показника бідності може вважатися достатньо надійною для регіонального рівня, якщо значення $RRMSE(P_{comp,k}^{(j)})$ не перевищує 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 не рекомендується використовувати для аналізу бідності.

Важливою характеристикою надійності оцінювання показників бідності є величина граничної похибки вибірки $ME(P_{comp,k}^{(j)})$ за регіонами, яка для довірчої імовірності 95% визначається за формулою:

$$ME(P_{comp,k}^{(j)}) = 1.96 \cdot RMSE(P_{comp,k}^{(j)}), \quad (7)$$

де $RMSE(P_{comp,k}^{(j)}) = \sqrt{MSE(P_{comp,k}^{(j)})}$ і є аналогом стандартної похибки. Гранична похибка $ME(P_{comp,k}^{(j)})$ визначає межі довірчих інтервалів композиційних оцінок: $P_{comp,k}^{(j)} \pm ME(P_{comp,k}^{(j)})$.

Технічні зауваження

Усі розрахунки за наведеними моделями здійснюються за затвердженою методологією з використанням спеціального програмного забезпечення, розробленого у середовищі SPSS.

Користування результатами

Статистичні дані та методологічні матеріали, наведені у публікації, призначені для використання органами державного управління, які здійснюють моніторинг бідності в Україні, фахівцями науково-дослідних установ та вищих освітніх навчальних закладів, які здійснюють дослідження рівня життя населення України.

Додаткова інформація

Контакти для отримання додаткової інформації щодо результатів обстеження, відповідного методико-методологічного забезпечення, а також довідок щодо умов поширення результатів ОУЖД такі:

адреса: 01601, м. Київ, МСП, вул. Шота Руставелі, 3

телефони: 234-01-34, 235-31-22, 234-32-13, 287-12-11

факс: 235-37-39

електронна пошта: households@ukrstat.gov.ua;

I.Osipova@ukrstat.gov.ua

Основні терміни та поняття

Додаткова (зовнішня) інформація – це інформація, яка 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>Композиційна оціночна функція</i>	– оціночна функція, отримана на основі композиції (комбінації) різних оціночних функцій. У статистиці найчастіше використовується композиція прямих і непрямих оціночних функцій
<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>Оцінка показника</i>	– значення показника, яке розраховано на основі відповідної оціночної функції за результатами вибіркового обстеження
<i>Оціночна функція</i>	– правило або співвідношення (формула) для оцінювання показника з використанням даних вибіркового обстеження. Оціночна функція відома ще до отримання результатів обстеження. Слід розрізняти поняття <i>оціночна функція</i> та <i>оцінка</i> , оскільки оціночна функція – це функція вибірки, тоді як оцінка є її значенням, знайденим на основі реалізації вибірки
<i>Пряме оцінювання</i>	– 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)
<i>Середньоквадратична похибка</i>	– загальна похибка оцінювання показника, яка визначається як сума дисперсії вибірових оцінок показника та квадрату величини зміщення оцінки, яке є різницею між очікуваною величиною оцінки (тобто середнім із вибірових оцінок) та дійсним значенням показника для генеральної сукупності

Перелік посилань

1. Rao J.N.K. Small Area Estimation / John Wiley&Sons, 2003. – 314 p.
2. Enhancing Small Area Estimation Techniques to meet European Needs (EURAREA) // Project IST-2000-26290. The Final Report. 2004.

Додаток Г

Шаблони вихідної інформації для внутрішніх та зовнішніх користувачів за результатами застосування статистико-математичної моделі для підвищення надійності оцінок показників бідності на регіональному рівні

Вихідна інформація для внутрішнього використання

За результатами застосування оціночних функцій (9) – (12) для оцінювання показників бідності по кожному з показників, надійність яких має бути підвищена, виводиться така інформація (див. табл. Г.1–Г.2).

- прямі оцінки показника бідності у році, що розглядається, $P_k^{(j)}$;
- дисперсія прямих оцінок за регіонами, $V_k(P_k^{(j)})$ та коефіцієнти варіації, $CV_k(P_k^{(j)}) = (\sqrt{V_k(P_k^{(j)})} / P_k^{(j)}) \cdot 100\%$;
- оцінки показника за регресійною моделлю за регіонами, $P_{regr,k}^{(j)}$;
- вектор-стовпець оцінки коефіцієнтів регресії, $\hat{\beta}$;
- коефіцієнт детермінації, регресійної моделі, R^2 ;
- F - відношення (F - критерій) для регресійної моделі;
- оцінки вагових коефіцієнтів композиційної оціночної функції за регіонами, $\gamma_k^{(j)}$;
- композиційні оцінки показника за регіонами, $P_{comp,k}^{(j)}$;
- оцінки величини середньоквадратичних похибок композиційних оціночних функцій за регіонами $MSE(P_{comp,k}^{(j)})$;
- величини відношення кореня квадратного від середньоквадратичної похибки до композиційної оцінки показника бідності за регіонами $RRMSE(P_{comp,k}^{(j)})$;
- оцінка величини граничної похибки вибірки $ME(P_{comp,k}^{(j)})$ за регіонами.

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

Слід зазначити, що для оцінки результатів моделювання здійснюються ряд перевірок (див. розділ 5 Методики). При цьому найважливішими перевітками є: порівняння прямих оцінок показників $P_k^{(j)}$ та їх оцінок за регресійною моделлю $P_{regr,k}^{(j)}$ (п. 6 розділ 5); порівняння прямих $P_k^{(j)}$ та композиційних $P_{comp,k}^{(j)}$ оцінок показників (п. 8 розділ 5); перевірка оцінок вагових коефіцієнтів композиційної оціночної функції (п. 7 розділ 5). Остаточне рішення щодо ефективності композиційної оціночної функції приймається за величинами відношення кореня квадратного від середньоквадратичної похибки до композиційної оцінки показника бідності за регіонами $RRMSE(P_{comp,k}^{(j)})$ (п. 9 розділ 5).

Макет вихідної інформації за регіонами за результатами застосування статистико-математичних моделей для оцінювання показників бідності для внутрішнього використання

[illegible]

Таблиця Г.2

Макет вихідної інформації за результатами застосування регресійних моделей, що використовуються при оцінювання показників бідності для внутрішнього використання

Назва показника	Оцінка показника
Коефіцієнт регресії (вільний член), $\hat{\beta}_0$	
Коефіцієнт регресії, $\hat{\beta}_1$	
Коефіцієнт регресії*, $\hat{\beta}_2$	
Коефіцієнт детермінації, R^2	
F - відношення	
Дисперсія оцінки за регресією, $\hat{\sigma}_v^2$	

*Для визначеності тут прийнято, що оцінка за регресією побудована на основі двох зовнішніх змінних:

$$P_{regr,k} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2.$$

Інші параметри моделей, що надаються внутрішнім користувачам, мають значення допоміжної інформації, яка може використовуватись для поглибленого аналізу результатів моделювання та контролю адекватності моделей. Наприклад, порівняння вихідної інформації поточного року з даними попереднього року може дозволити пояснити зміни окремих параметрів моделі (моделей) або рівнів надійності оцінювання показників. Про адекватність моделі регресії наближено можна зробити висновок якщо $R^2 \geq 0,7$ та $F > F_{кр}$. Якщо ці умови (або одна з умов) не виконується це може свідчити про неадекватність регресійної моделі і необхідність побудови іншої моделі. При виконанні зазначених умов, але суттєвих змінах коефіцієнтів регресії, доцільно проаналізувати кореляційні зв'язки залежних та незалежної змінних та з'ясувати причини змін. Величину $\hat{\sigma}_v^2$ доцільно порівняти з дисперсіями прямих оцінок $V_k(P_k)$ - вона повинна бути близької до середнього значення цих дисперсій або меншою за нього.

Вихідна інформація для надання зовнішнім користувачам

За результатами застосування статистико-математичних моделей зовнішнім користувачам інформація надається фактично у тому ж 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 (див. додаток В).

При підготовці інформації слід враховувати і 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 математико-статистичних моделей доцільно представляти у вигляді таблиць, діаграм та графіків. Нижче наведені рекомендовані макети таблиць (табл. Г.3 – Г.5) для представлення інформації щодо показника рівня бідності, визначеного за критерієм національної межі.

Таблиця Г.3

Уточнені оцінки рівнів бідності, визначених за національною межею, та характеристики їх надійності за регіонами України у 20 році

	Уточнена оцінка рівня бідності за попередній рік (довідково), %	Уточнена оцінка рівня бідності, %	Відносна похибка оцінювання, %	Довірчий інтервал, %
Україна				
АР Крим				
Вінницька область				
Волинська область				
Дніпропетровська область				
Донецька область				
Житомирська область				
Закарпатська область				
Запорізька область				
Івано-Франківська область				
Київська область				
Кіровоградська область				
Луганська область				
Львівська область				
Миколаївська область				
Одеська область				
Полтавська область				
Рівненська область				
Сумська область				
Тернопільська область				
Харківська область				
Херсонська область				
Хмельницька область				
Черкаська область				
Чернівецька область				
Чернігівська область				
м. Київ				
м. Севастополь				

Таблиця Г.4

Уточнені оцінки рівня бідності за національною межею та характеристики їх надійності за економічними районами у 20__ році

	Уточнена оцінка рівня бідності за попередній рік (довідково), %	Уточнена оцінка рівня бідності, %	Відносна похибка оцінювання, %	Довірчий інтервал, %
Східний				
Донецький				
Придніпровський				
Причорноморський				
Подільський				
Центральний				
Карпатський				
Поліський				

Для візуалізації результатів непрямого оцінювання показників бідності доцільно представити користувачам карти бідності для регіонів та економічних районів України (рис. Г.1, Г.2).

Аналогічні таблиці та карти наводяться по кожному показнику бідності, для якого визначаються уточнені оцінки, зокрема, для рівня бідності, визначеного за регіональною межею, та для рівня бідності, визначеного за розміром прожиткового мінімуму.



Рис. Г.1. Карта бідності за показником рівня бідності за регіональною межею за регіонами України у 20__ році.



Рис. Г.2. Карта бідності за показником рівня бідності за національною межею за економічними районами України у 20__ році.

Перелік використаних джерел

1. Longford N.T. Simulation of small-area estimators of the poverty rates in the oblasts of Ukraine. - SNTL and UPF, Barcelona, Spain. The report prepared for the Social Assistance System Modernization Project, Ukraine, Kyiv, 2010.
2. Longford N. T. Missing Data and Small-Area Estimation. Modern Analytical Equipment for the Survey Statistician. Springer-Verlag, New York, 2005.
3. Додаток 5 до сьомого звіту за договором №4807-56 від 22 жовтня 2008 року щодо надання консультативних послуг з моніторингу та оцінки соціальної політики «Пропозиції щодо побудови прототипів статистико-математичних моделей для оцінювання бідності, які забезпечуватимуть належні рівні надійності визначених показників бідності на регіональному рівні та здійснити апробацію прототипів моделей».
4. Методологічні положення для оцінки показників бідності на регіональному рівні із застосуванням методів статистико-математичного моделювання, затверджені наказом Держкомстату від 01.10.10 № 408.
5. Rao J.N.K. Small Area Estimation / John Wiley&Sons, 2003. – 314 p.
6. Додаток 3 до одинадцятого звіту за договором №4807-56 від 22 жовтня 2008 року щодо надання консультативних послуг з моніторингу та оцінки соціальної політики «Аналітичні матеріали щодо результатів апробації процедур надійного оцінювання визначених показників бідності на регіональному рівні на узгоджених засадах функціонування моделей та вдосконалення дизайну вибірки ОУЖД».
7. Додаток 2 до другого звіту за договором №4807-56 від 22 жовтня 2008 року щодо надання консультативних послуг з моніторингу та оцінки соціальної політики «Аналітичні матеріали щодо інформаційних джерел, які можуть бути використані при розробці моделей для оцінювання бідності, та їх подальшого практичного застосування».