



**Пілотні Проекти Соціальної Допомоги
Семінар**

**“Представлення та обговорення проміжних результатів
впровадження та оцінки пілотних проектів соціальної допомоги”**

**Опрацювання моделі непрямой оцінки
доходів (НОД)**

Семко Роман
CASE Україна

Березень, 2010

1. Вступ до моделювання
2. Аналіз даних
3. Методи для оцінки моделі
4. Симулювання
5. Дохід від використання активів (у сільському господарстві)
6. Результати експерименту (“подвійний сліпий метод”)
7. Порівняння моделей та висновки

- Світовий Банк розробив методологію для оцінки доходів, що ґрунтується на регресійному аналізі – НЕПРЯМА ОЦІНКА ДОХОДІВ (НОД)
- На основі методу НОД оцінюється право на отримання СД з використанням моделювання доходу домогосподарства
- Загальний дохід розділяється на дві частини: задекларований (напр., пенсії, стипендія) і незадекларований (напр., дивіденди, зарплата у конверті)
- Кінцевою метою є оцінка частки незадекларованого доходу на основі показників, що можуть бути точно виміряні і власне відображають незадекларований дохід
- Незадекларований дохід розділяється на дохід, що генерується не довгостроковими активами (оцінюється регресійною моделлю) і дохід від активів (оцінюється формулами)

Головною ціллю моделі є прогнозування сукупного доходу сім'ї з максимально можливою точністю



Методи

Дані та
знання

Модель

Рівняння, що оцінює дохід заявника на основі доступної інформації:

$$Y = \beta_1 * X_1 + \beta_2 * X_2 + \beta_3 * X_3 + \dots$$

Y	X ₁ , X ₂ , X ₃ , ...
➤ сукупний дохід ➤ незадекларований дохід	➤ структура сім'ї ➤ вид і сектор зайнятості ➤ освіта ➤ регіон ➤ інші

Критерії

1. Теоретичне підґрунтя
2. Простота
3. Максимізація коефіцієнта детермінації
4. Значимість незалежних змінних

Застосування та Симуляція

Ефективна модель повинна використовувати всю наявну релевантну інформацію для прогнозування доходів



ОУЖД 2008

10622 спостережень домогосподарств із загальним доходом



Необхідно/можливо використовувати більше інформації, щоб гарантувати прийнятний рівень точності

Дані пілотів

- > 3000 спостережень сімей із задекларованим доходом
- Не може використовуватися окремо для оцінки моделі, оскільки відсутній сукупний дохід

Задекларований дохід (ЗД) є важливим показником при оцінці сукупного доходу (СД)

СПІВСТАВЛЕННЯ

$$\begin{array}{c} \text{СД} \quad \text{Характеристики} \quad + \quad \text{Характеристики} \quad \text{ЗД} \\ \hline \text{СД} \quad \text{Характеристики} \quad \text{ЗД} \end{array}$$

Спостереження співставляються так, щоб гарантувати максимальну схожість між ними



ОУЖД 2008

Спостереж. 1 Спостереж. ... Спостереж. K_1	Група 1
Спостереж. 1 Спостереж. ... Спостереж. $K_{...}$	Група ...
Спостереж. 1 Спостереж. ... Спостереж. K_N	Група N

Процедура

1. **Формування груп** на основі наступних змінних: місце проживання, вид СД, розмір домогосподарства, кількість дітей, зайнятих, пенсіонерів, стать голови
2. **Співставляємо кожне спостереження з ОУЖД зі спостереженням з бази даних пілотів з однакових груп на основі схожих характеристик: вік голови, освіта тощо використовуючи функцію відстані Евкліда**
3. **Кожне спостереження з бази даних пілотів використовується для співставлення не більше 2 разів**
4. **Агрегування груп, якщо немає відповідних спостережень для ОУЖД у відповідній групі з бази даних пілотів, і співставляємо знову**

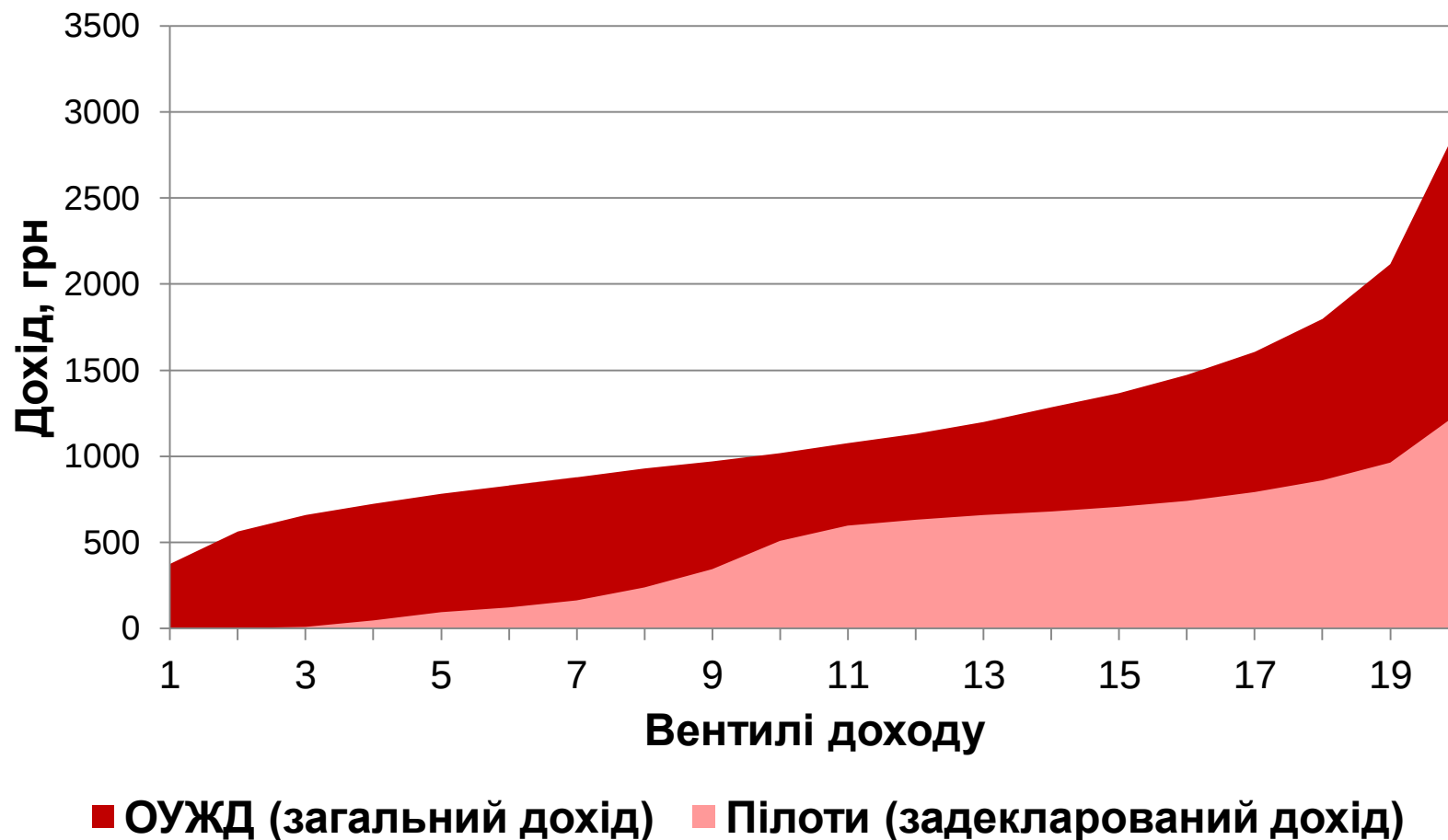
Дані пілотів

Група 1	Спостереж. 1 Спостереж. ... Спостереж. L_1
Група ...	Спостереж. 1 Спостереж. ... Спостереж. $L_{...}$
Група N	Спостереж. 1 Спостереж. ... Спостереж. L_N

Порівняння даних: ключова різниця між ОУЖД і даними пілотів, полягає доходи, в той час як інші характеристики схожі



Загальний і задекларований дохід (без СД)



У деяких областях середній дохід на основі ОУЖД значно відрізняється від індивідуального використовуваного доходу (ІВД)



ІВД, Держкомстат



ОУЖД



Різниця у доході без СД на душу населення у порівнянні з Чернівецькою областю, грн



Економетрика Баєса дозволяє поєднувати дані з агрегованими показниками



Калібрація

Коефіцієнт(и) моделі визначаються дослідником самостійно, напр., якщо регіональні макродані показують, що дохід в Києві на 1108 грн більший ніж в АР Крим, то прогнозований дохід заявника в Києві також на 1108 грн більший ніж в АР Крим за інших рівних умов

Економетрика Баєса

Комбінація обидвох підходів. Розраховані коефіцієнти лежать між каліброваними коефіцієнтами та коефіцієнтами, що розраховані на основі стандартної оцінки

Стандартна оцінка

Коефіцієнти визначаються на основі зібраних спостережень використовуючи стандартні методи оцінки (класична економетрика)



Не спричиняє значних змін у прогнозованому доході в межах області, але по областях України зміни значні: середній прогнозований дохід для областей змінився

Лінійна модель є найпростішою



Лінійна

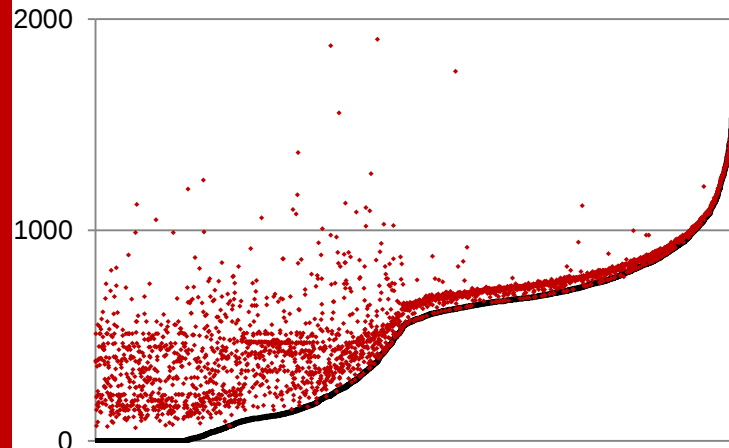
Опис

- Лінійний зв'язок між доходом і характеристиками сім'ї
- Залежна змінна у формі логарифма (лог-лінійна)
- Незалежні змінні (НЗ) включають легкооцінюваний дохід
- Інші НЗ: кількість дітей, зайнятих, людей похилого віку, тип і галузь зайнятості голів сім'ї, рівень освіти

R^2

58 % (великі міста – 65%, малі міста – 63%, села – 48%)

Прогнози



Концепція: чим більший задекларований дохід, тим менший дооцінений дохід, – схоже із “грою з нульовою сумою”

■ – задекларований дохід ■ – прогнозований дохід

Нелінійна модель працює ефективно, якщо різниця у доходах сімей значна – тобто на основі всієї бази ОУЖД



Нелінійна

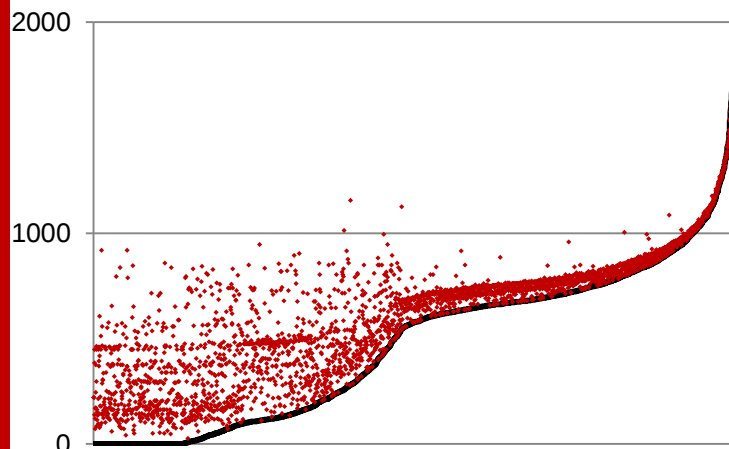
Опис

- Нелінійний зв'язок з доходом та характеристиками сім'ї. Форма зв'язку: кубічна чи квадратична – оскільки загальний дохід відсортований у зростаючому порядку збільшується як поліном 2 чи 3 порядку
- Залежна змінна у формі логарифма (лог-лінійна)
- Незалежні змінні такі, як і в Лінійній моделі

R^2

Коефіцієнт детермінації не лежить в $[0\%, 100\%]$ інтервалі

Прогнози



Концепція: як і в лінійній моделі

■ – задекларований дохід ■ – прогнозований дохід

Двокрокова модель найбільш ефективна, коли є велика кількість сімей як з нульовим, так і з ненульовим важкооцінюваним доходом



Двокрокова

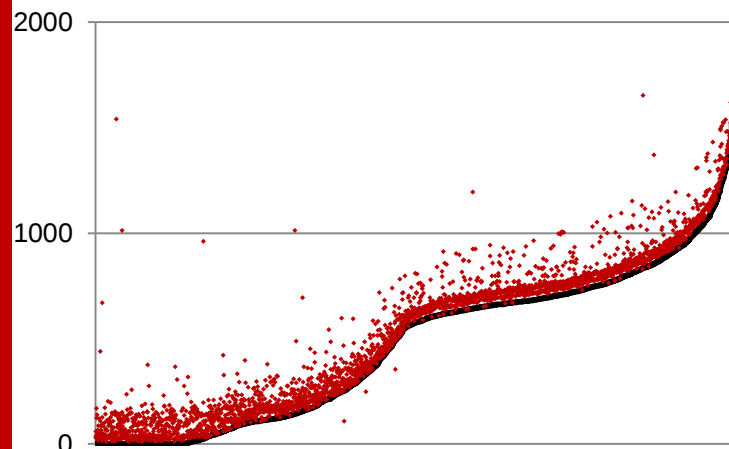
Опис

- На першому кроці оцінюється ймовірність того, що сім'я генерує важкооцінюваний дохід, а потім використовується для оцінки лінійна залежність між доходом та характеристиками сім'ї і ризиком наявності важкооцінюваного доходу
- Залежна змінна у формі логарифма і не включає зарплату

R^2

47 % (без поділу по типу населених пунктів)

Прогнози



Концепція: Стабільний дооцінений дохід додається до задекларованого – “гра із сталим виграшем”.

■ – задекларований дохід ■ – прогнозований дохід

Кожну модель необхідно відкоригувати для кінцевого використання



Корекція

Опис

1. ЗАЛЕЖНА ЗМІННА

Такі елементи незадекларованого доходу як зарплата у конверті повинні бути включені в залежну змінну

2. НЕЗАЛЕЖНІ ЗМІННІ (НЗ)

Деякі НЗ, які використовуються для прогнозування, важко точно виміряти. Наприклад, кількість мобільних телефонів

3. ЧАСОВИЙ АНАЛІЗ

Щоб порівнювати доходи отримані в різні проміжки часу, застосовуються середній темп росту ІВД та його елементів

4. ГОЛОВА СІМ'Ї

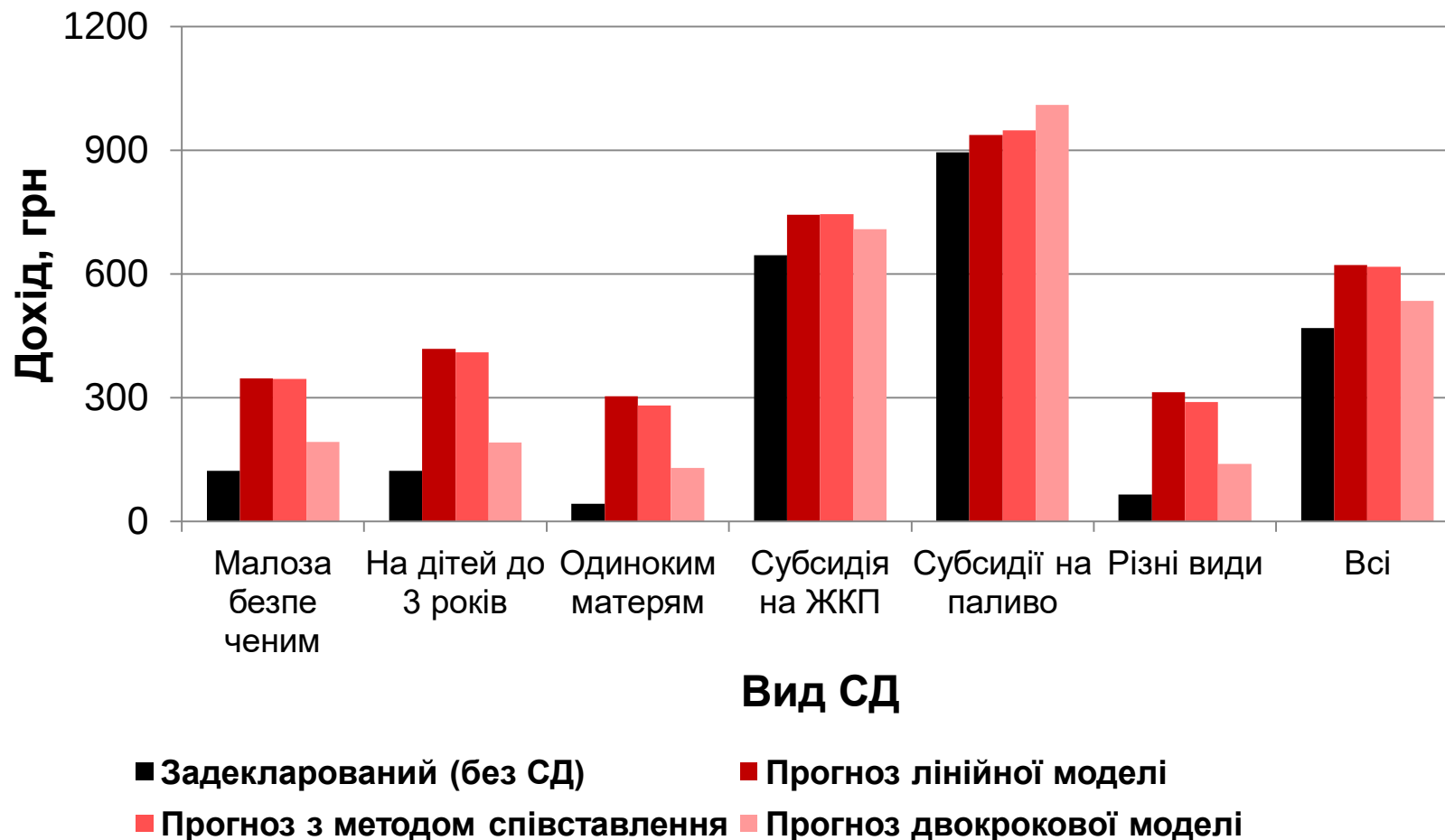
Визначення голови сім'ї стандартизовано: чоловік співголова та жінка співголова замість довільних визначень

Прогнози не змінюються значно при впровадженні цих змін, крім першого випадку. Якщо залежна змінна класифікована по іншому, то дооцінений дохід є більш стабільним і спадає при зростанні задекларованого доходу з меншою швидкістю

Середній прогнозований дохід перевищує задекларований на 26%



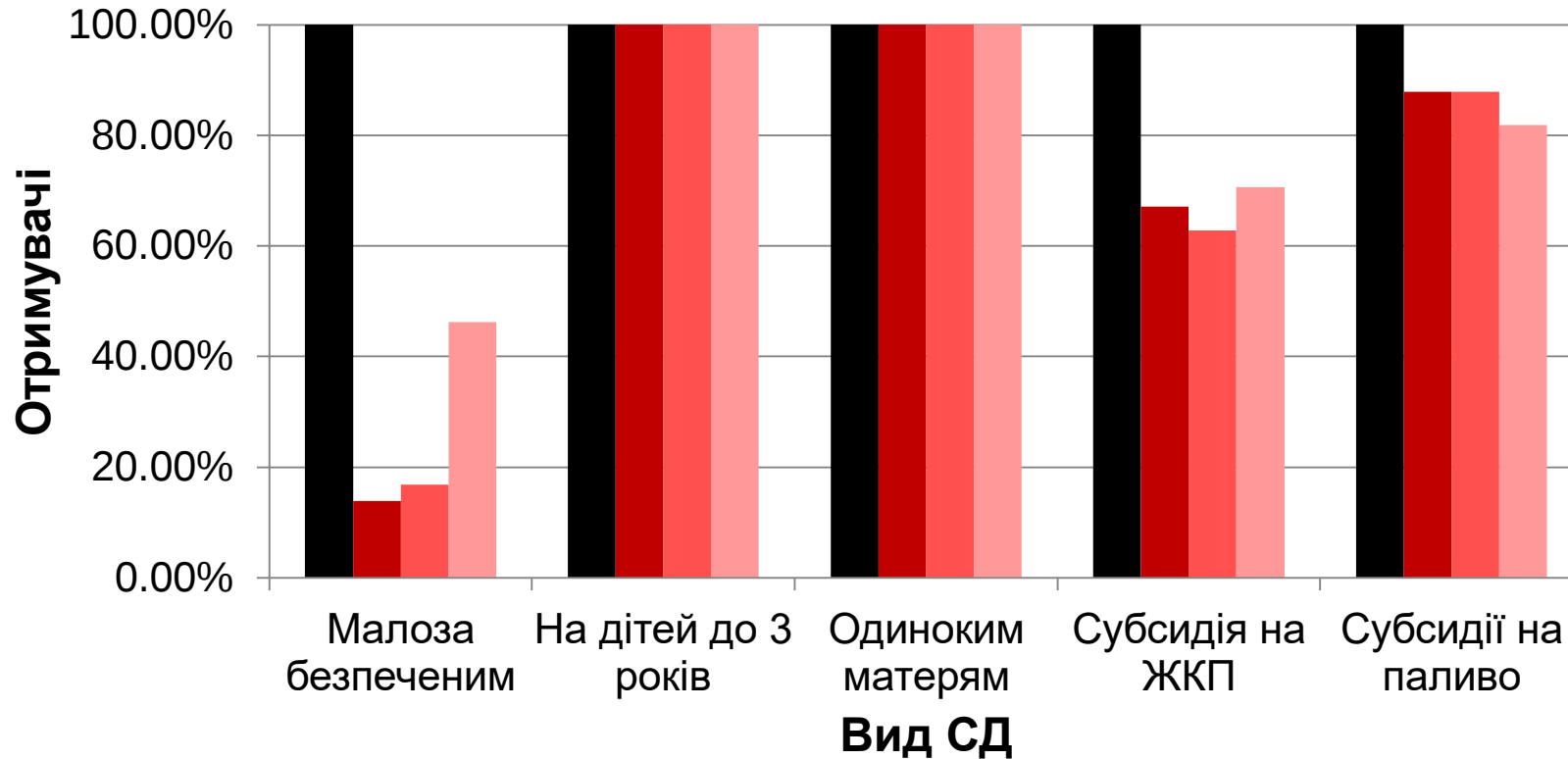
Задекларований і прогнозований дохід (по моделях)



27% сімей будуть виключені із програми СД



Кількість отримувачів (гіпотетичний сценарій)



■ Статус-кво

■ Лінійна модель

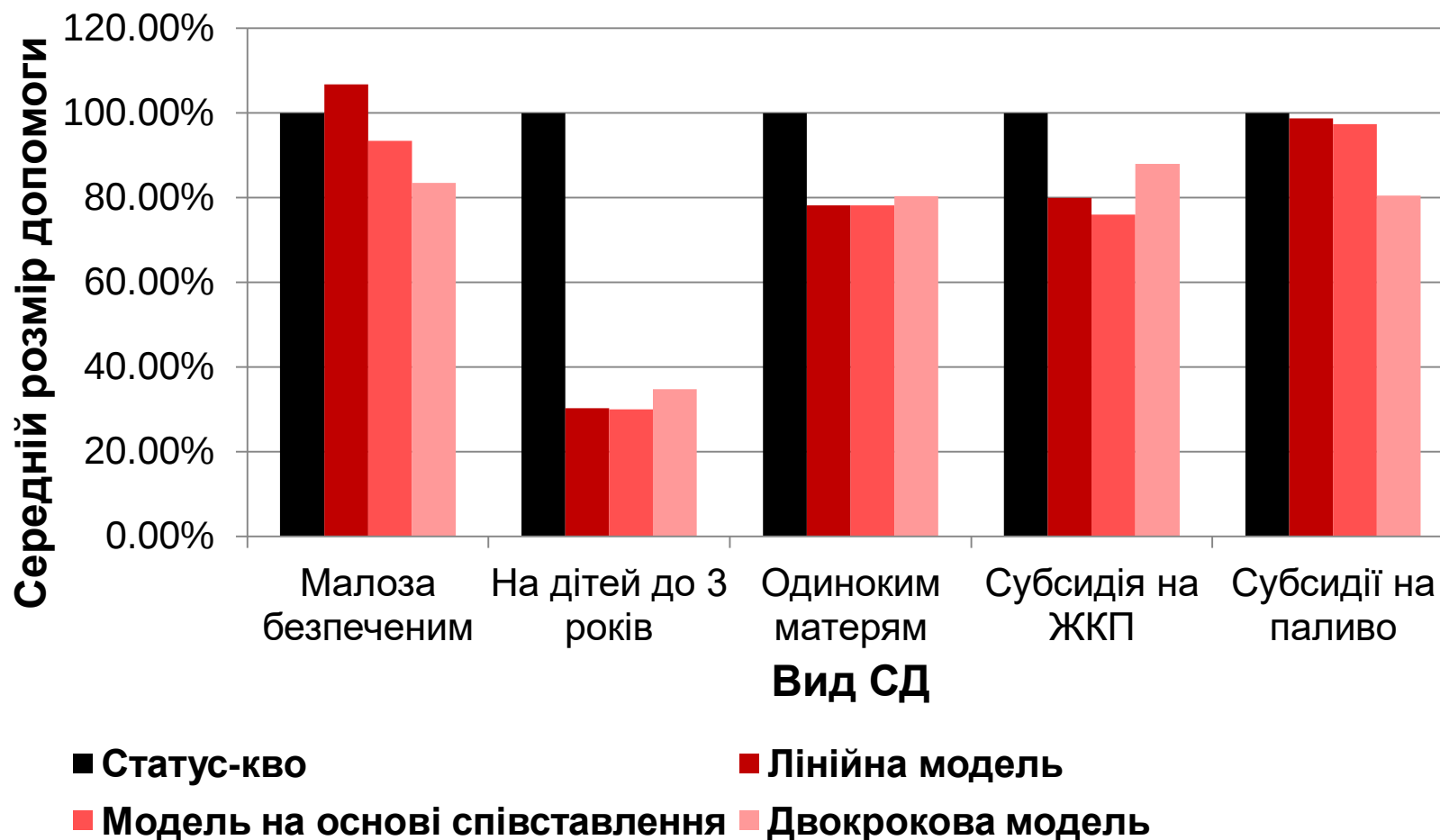
■ Модель на основі співставлення

■ Двоступенчата модель

Середній розмір допомоги значно зменшиться для всіх видів СД, крім допомоги малозабезпеченим і субсидій на паливо



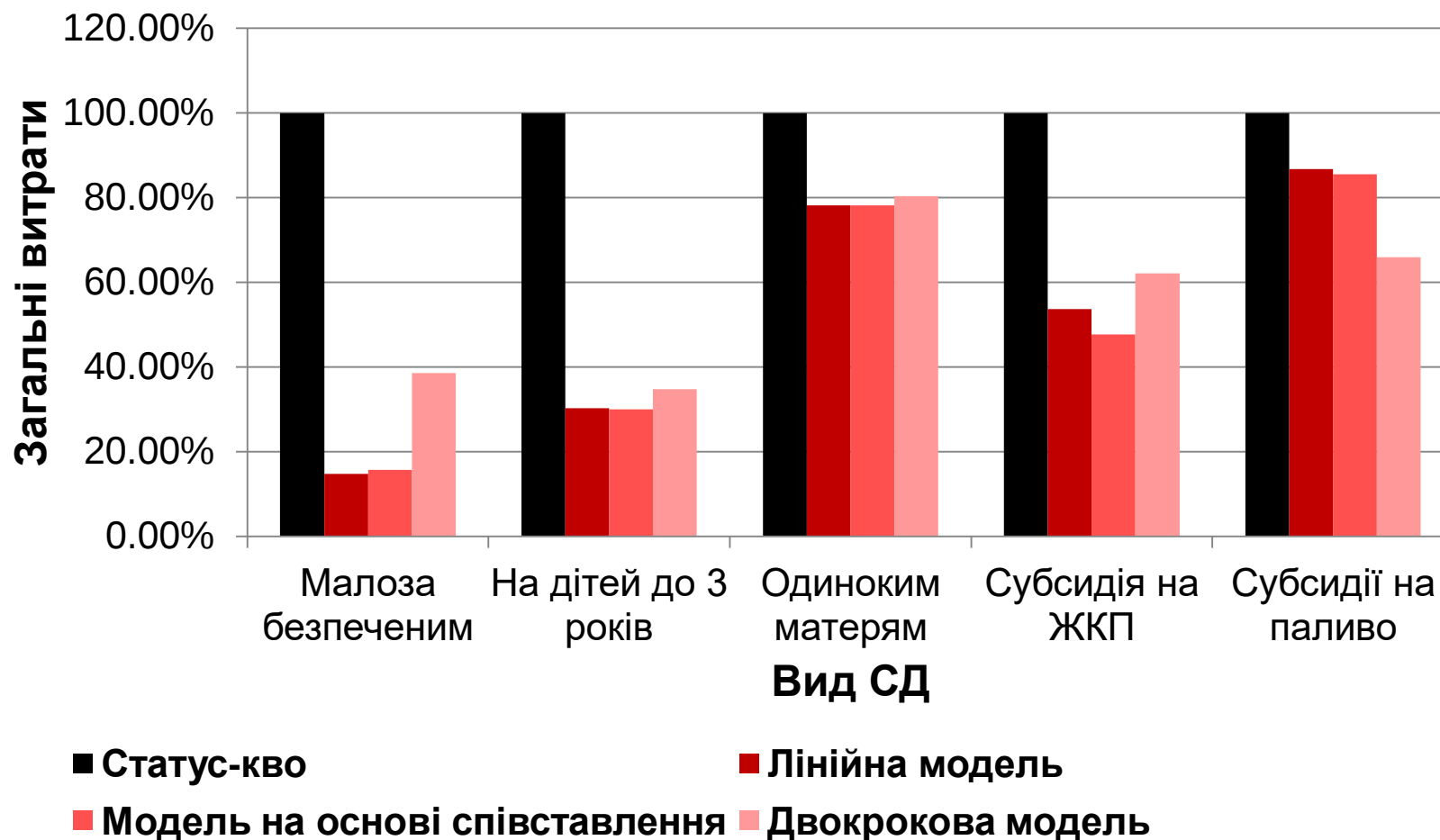
Середній розмір СД (гіпотетичний сценарій)



Загальний бюджет на виплату СД зменшиться на 27%



Загальні витрати на СД (гіпотетичний сценарій)



Дохід від сільського господарства розраховується на основі розроблених нормативів



Існуюча ситуація

- Агродохід розраховується як дохід на гектар
- Нормативи не уніфіковані по областях

Новий підхід

- Агродохід розраховується як дохід на гектар та одиницю худоби
- Нормативи розрізняються для сіл і міст
- Нормативи уніфіковані і ґрунтуються на однаковій методології і даних

Процедура розрахунку

Інформація, підтверджена довідками із сільської / міської ради

Не застосовується для сімей, де є інваліди чи люди літнього віку (>70)

Якщо заявник живе ближче ніж 10 км від міста – нормативи для міст

Дохід від землі є добутком площі та нормативу

Дохід від паїв розраховується окремо

Дохід від худоби є добутком кількості тварин та нормативів

Середній прогнозований дохід перевищує задекларований на 28%

Приклад розрахунку доходу від сільського господарства

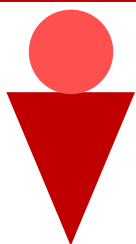


	РОСЛИННИЦТВО			+	ТВАРИННИЦТВО			=	АГРОДОХІД
	ПЛОЩА	×	НОРМАТИВ		ТВАРИНИ	×	НОРМАТИВ		
Існуюча ситуація	Тільки ділянка площею 0,56 га у селі (Донецька область)	×	127,62 грн на га в місяць	+	Є одна корова і 10 курей	×	Тільки через сіножаті і пасовища	=	+ 63,81 грн в місяць
Новий підхід	Тільки ділянка площею 0,56 га у селі (Донецька область)	×	412,44 грн на га в місяць	+	Є одна корова і 10 курей	×	270,83 для корови і 4,48 для однієї курки	=	+ 521,85 грн в місяць

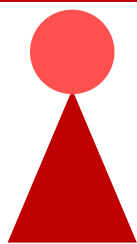
Експеримент (“подвійний сліпий метод”): приклад



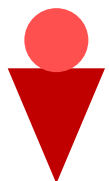
Характеристика сім'ї



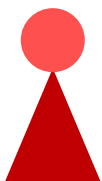
Батько: безробітний і
не зареєстрований в
центрі зайнятості



Мати:
домогосподарка



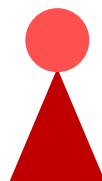
Вік: <18



Вік: <3



Вік: <3



Вік: <3

Комісія та інспекція на дому

ВІДМОВА

Результати моделі

Задекларований дохід = 211 грн

МЕНШЕ ніж

Мінімальний рівень = 255 грн



НАДАВАТИ?

В ТОЙ ЧАС, ЯК

Прогноз моделі = 308 грн

рішення – сім'я в зоні ризику,
необхідна інспекція на дому

БІЛЬШЕ ніж

Мінімальний рівень = 255 грн



ВІДМОВА

Відмова у наданні СД комісією на основі обстеження на дому

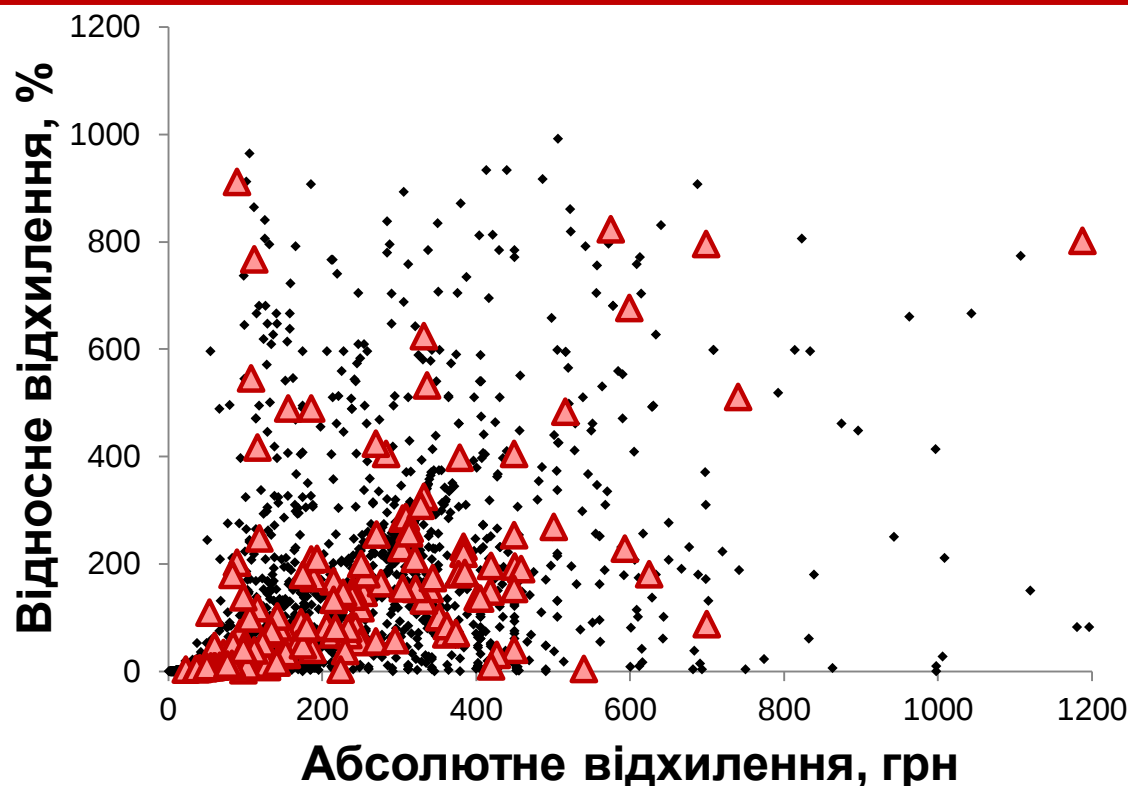


Вид СД	Задекларований дохід	Прогнозований дохід	Абсолютне відхилення	Відносне відхилення
	УАН	УАН	УАН	%
Малозабезпечені	9	133	124	1320
Малозабезпечені	211	308	97	46
Малозабезпечені	264	344	80	30
На дітей до 3 років	10	133	123	1238
На дітей до 3 років	12	127	114	931
На дітей до 3 років	209	286	77	37
На дітей до 3 років	250	327	77	31
На дітей до 3 років	260	365	105	40
На дітей до 3 років	273	330	56	21
Одиноким матерям	0	68	68	-
Одиноким матерям	222	291	69	31

Прогнозований дохід допомагає вибрати сім'ї для інспекції на дому



Сім'ї вибрані для інспекції



• Не вибрані **▲** Вибрані

Коментарі

- Кожна сім'я має можливість бути вибраною для інспекції
- Ймовірність вибору зростає, якщо прогнозований дохід значно відрізняється від задекларованого у абсолютному і відносному вимірі

Порівняння моделей



Характеристики	МОДЕЛЬ				
	Лінійна (без співставлення)	Нелінійна	Двокрокова	Лінійна (співставлення)	Лінійна (підхід Баєса +/- співстав.)
Теоретична база	Слабка, бо лінійні залежності малопоширені	Сильніша, бо враховує нелінійну поведінку	Сильна, якщо очікується зміщення відбору	Як і для лінійної (без співставлення)	Як і для лінійної
Простота	Дуже проста	Проста	Дещо складна	Проста в оцінці, але складна при співставленні	Дуже складна оцінка
R ²	Високий	-	Високий	Високий	-
Інформація для розрахунків	Тільки ОУЖД	Тільки ОУЖД	Тільки ОУЖД	ОУЖД і дані пілотів	Дуже ефективно використання
Вплив на заявників	Великий	Великий	Середній	Великий	Великий

- 1** Оцінений різними моделями дохід значно відрізняється від задекларованого заявниками доходу
- 2** Модель потребує більше емпіричних тестів
- 3** Початкові результати моделі повинні використовуватись як порада, а не критерій для надання СД
- 4** Моделі можуть використовуватись як інструмент для відбору заявників для інспекції на дому