Model of Income and Measurement error: Simplification

Setup

Assume that E_r and U_r are the earnings and Ubenefits of individuals, based on registered data. And for simplicity, we assume they are measured without error.

In this setup, we would observe 4 cases for the earnings/ubenefits of individuals.

Table 1: Earnings Setup

Case	Description	Prob
I	$E_r = 0 \& U_r = 0$	NA
II	$E_r > 0 \& U_r = 0$	$p1$
III	$E_r = 0 \& U_r > 0$	$p2$
IV	$E_r > 0 \& U_r > 0$	$1 - p1 - p2$

For now, lets assume CASE I is uninteresting.

From the Survey perspective we could have multiple cases:

- S1: Data is reported without error
- S2: Earnings are reported with error but Ubenefits are reported without error
- S3: Earnings are witout error but Ubenefits are reported with error
- S4: Both earnings and Ubenefits are reported with error

S1: $E_s = E_r$ and $U_s = U_r$

S2: $E_s = E_r + _1 $ and $U_s = U_r$

S3: $E_s = E_r $ and $U_s = U_r + \epsilon_2$

S4: $E_s = E_r + \epsilon_3$ and $U_s = U_r + \epsilon_4$

Im using different ϵ_k for now.

S1: Combined with Cases of R

S1 & CII CIII CIV: $E_s = E_r$ and $U_s = U_r$ with p_s

S2: Combined with Cases of R

S2 & CII : - $ln(E_s) = ln(E_r) + \epsilon_2$ and and $U_s = 0$ with probability $p_s s2$
- $E_s = 0$ and and $U_s = 0$ with probability $1 - p_s s2$

S2 & CIII : - $ln(E_s) = \epsilon_1$ and and $ln(U_s) = ln(U_r)$

S2 & CIV : - $ln(E_s) = ln(E_r) + \epsilon_2$ and and $ln(U_s) = ln(U_r)$ with probability $p_s s2$ - $E_s = 0$ and and $U_s = 0$ with probability $1 - p_s s2$

ϵ_1 is similar to Mismatch, because people declare some earnings, even tho they dont have any ($E_r = 0$).

ϵ_2 is just measurement error.

$p_s 2$ is the probability of people declaring earnings, and $1 - ps2$ is the probability of people not declaring earnings.

S3: Combined with Cases of R

S3 & CII :
- $E_s = E_r$ and and $ln(U_s) = \lambda_1$

S3 & CIII : - $E_s = 0$ and and $ln(U_s) = ln(U_r) + \lambda_2$ with probability $p_s s3$ - $E_s = 0$ and and $U_s = 0$ with probability $1 - p_s s3$

S3 & CIV : - $E_s = E_r$ and and $ln(U_s) = ln(U_r) + \lambda_2$ with probability $p_s s3$ - $E_s = E_r$ and and $U_s = 0$ with probability $1 - p_s s3$

λ_1 is similar to Mismatch, because people declare some Ubenefits even tho they dont have any ($U_r = 0$). The caveat, we cannot identify cases here Registered UB data is positive, but Survey UB data is 0. But we can just add a Probability of this happening.

S4: Combined with Cases of R

This is the only one with very interesting cases.

S4 & CII : - $ln(E_s) = ln(E_r) + \epsilon_2$ and $ln(U_s) = \lambda_1$ - $E_s = 0$ and $ln(U_s) = \lambda_1$

S4 & CIII : - $ln(E_s) = \epsilon_1$ and $ln(U_s) = ln(U_r) + \lambda_2$ - $ln(E_s) = \epsilon_1$ and $U_s = 0$

S4 & CIV : - $ln(E_s) = ln(E_r) + \epsilon_2 + \gamma$ and $ln(U_s) = ln(U_r) + \lambda_2 - \gamma$ - $ln(E_s) = ln(E_r) + \epsilon_2 + \gamma$ and $U_s = 0$ - $E_s = 0$ and $ln(U_s) = ln(U_r) + \lambda_2 - \gamma$ - $E_s = 0$ and $U_s = 0$

II | $E_r > 0 \& U_r = 0$ | $p1$ |
III | $E_r = 0 \& U_r > 0$ | $p2$ |
IV | $E_r > 0 \& U_r > 0$ | $1 - p1 - p2$ |

S4: $E_s = E_r + \epsilon_3$ and $U_s = U_r + \epsilon_4$