Discussion of:

Following the Money: An Econometric Investigation into Health Valuations

by: Tinna Laufey Ásgeirsdóttir and Viktor Andri Kárason

Nicholas W. Papageorge

AWEW — 2025

Quality Adjusted Life Years (QALYs)

- $\rightarrow$ This paper is about valuing QALYs.
- $\rightarrow$ A nice way to close this conference.
- $\rightarrow$ It raises a problem.
- $\rightarrow$ QALYs are measures widely-used in policy...
- $\rightarrow$ ...that are computed problematically.

Quality Adjusted Life Years:
Bringing Bad 'Metrics and Bad Theory Together

The **quality-adjusted life year** (**QALY**) is a generic measure of disease burden, including both the quality and the quantity of life lived.[1][2] It is used in economic evaluation to assess the value of medical interventions.[1] One QALY equates to one year in perfect health.[2] QALY scores range from 1 (perfect health) to 0 (dead).[3] QALYs can be used to inform health insurance coverage determinations, treatment decisions, to evaluate programs, and to set priorities for future programs.[3]

Critics argue that the QALY oversimplifies how actual patients would assess risks and outcomes, and that its use may restrict patients with disabilities from accessing treatment.[4] Proponents of the measure acknowledge that the QALY has some shortcomings, but that its ability to quantify tradeoffs and opportunity costs from the patient, and societal perspective make it a critical tool for equitably allocating resources.[4]

Valuation of Life Years or Quality Adjusted Life Years.
  → OK, but then how do we value QALYs?
      → Ask people questions (stated preferences / utility).
      → Calculate values or losses for disabilities.
      → Hypotheticals (i.e., ask people more complicated questions).

Valuation of Life Years or Quality Adjusted Life Years.
- → Types of questions:
    - → *Imagine you have 10 years left to live with condition X. How many years would you be willing to give up to be in full health for the remaining time?*
    - → *You have condition X. Would you choose a treatment that offers a chance of returning to full health or a chance of dying?*
    - → *On a scale of 0 to 100, where 0 is dead and 100 is perfect health, what number would you give to your current health state*

Valuation of Life Years or Quality Adjusted Life Years.

$\rightarrow$ Other methods to value QALYs?

$\rightarrow$ Look at wage differentials / risk premia.

$\rightarrow$ Look at actual payments for treatment.

Valuation of Life Years or Quality Adjusted Life Years.

- $\rightarrow$ Another approach: use more economic theory.
- $\rightarrow$ Essentially: calculate a compensating variation.
- $\rightarrow$ Yet, how—precisely—this is done matters.
- $\rightarrow$ Different modeling assumptions lead to vastly different conclusions.

This Paper.

1. Reconciles two QALY calculations that are vastly different *and that use the same date, setting, measures (mostly) and approach (mostly).*

2. Asks a broader question: *Might we consider standardization?*

3. Contributes to an even broader question: *Might we think harder if two studies purporting to do the same thing get absolutely different answers?*

Rest of Discussion.

1. A brief overview of the paper and what they found.
2. Some time on the question of QALYs and standardization of valuation.

Method:
$$LS = f(H, Y, X)$$

where:

- $\rightarrow$ LS: Life satisfaction
- $\rightarrow$ H: Health
- $\rightarrow$ Y: Income
- $\rightarrow$ X: Additional Variables.

Method:

Translate to an estimation equation:

$$LS = \alpha + \beta_1 H + \beta_2 Y + \beta_3 X + \epsilon$$

Next:

- → Estimate this equation.
- → Calculate COMPENSATING INCOME VARIATION.
- → With continuous income, that is $-\frac{\beta_1}{\beta_2}$.
- → (Keeping LS constant how much does Y go up if H goes down).

Method:

What went wrong?

$$LS = \alpha + \beta_1 H + \beta_2 Y + \beta_3 X + \epsilon$$

$\rightarrow$ Two studies get vastly different answers.

$\rightarrow$ Study 1: A\$42-83K for a gain of 1 QALY

$\rightarrow$ Study 2: About A\$225K for a loss of 0.1 QALY

$\rightarrow$ Not even in the same ballpark: S2 is 30-40 times S1.

$\rightarrow$ Both use HILDA data, similar variables.

This Paper: Detective Work.

- $\rightarrow$ Follow each study and see if there are errors.
- $\rightarrow$ Interrogate each difference.
- $\rightarrow$ See which one(s) matter(s).

- $\rightarrow$ There are no important errors.
- $\rightarrow$ Most assumptions (e.g., looking at 1 versus 2 years) don't matter a lot.
- $\rightarrow$ Bottom line: Income modeling drives differences.

This Paper: Detective Work.

$\rightarrow$ "[T]he differences in the monetary valuations of health are first and foremost driven by whether the income change is perceived as a windfall or a setback."

$\rightarrow$ Setbacks have relatively big effects on income.

$\rightarrow$ Use losses (e.g., as an IV for income changes) and LS goes down a lot.

$\rightarrow$ Then, β_2 is very large.

$\rightarrow$ Thus, $-\frac{\beta_1}{\beta_2}$ is relatively small.

$\rightarrow$ Perhaps due to curvature of utility? Loss aversion? Because setbacks are signs of more devastating changes (e.g., illness or job loss) than are improvements (e.g., a raise).

Thoughts on the Paper and the Topic of Valuing QALYs:

→ Wow, we use this for policy?

→ I agree with the point made by the authors: interrogate differences.

→ The step-by-step replication is useful.

→ I would add more on: *OK, but what should we do?*

Rant: Somewhat Frustrating.

$\rightarrow$ There are hosts of questions raised in conferences like this one.

$\rightarrow$ Here, particularly, we think about measurement and theory.

   $\rightarrow$ How should we think about errors of ratios?

   $\rightarrow$ What variables are informative or exogenous (or not)?

   $\rightarrow$ What measures contain additional information that cast doubt on their use?

   $\rightarrow$ How do we extract information that we might care about?

   $\rightarrow$ What values can structural models offer that RF/RCT approaches don't?

   $\rightarrow$ What values can RF/RCT approaches offer to improve structural models?

$\rightarrow$ This paper links to these issues with a widely used method.

Some Musings:

- → How ought we use income? Are some measures worse than others?
- → Do people prefer lost income and better health or more income but worse health? Are there assymmetries?
- → Why are WTP and WTAP different? What model might rationalize this?
    - → These kinds of differences need not require strange utility functions.
    - → Low MUC can make any health improvement seem infinitely valuable.
- → Do income changes capturing factors we do not want to include?

Some (More) Musings:

→ The CIV uses a model of income and health, but is linear and includes a few X variables.

→ These are really strong assumptions!

→ It is probably a reduced form of a very particular model that does not distinguish between different measurements of income, health, and utility.

Valuing QALYs:

→ The standard approach is a counterfactual policy analysis...

→ ...with a model with very strong simplifying assumptions.

→ We probably need more flexibility and more structure.

Valuing QALYs: Structural to the Rescue?

$\rightarrow$ Common practice ought not to be a model so sensitive to reasonable modeling choices.

$\rightarrow$ We still probably want to value QALYs, but well.

$\rightarrow$ A structural lifecycle model might be a place to start.

$\rightarrow$ Key ingredients:

1. Lifetime utility maximization.
2. A rich health process.
3. Health affecting utility, labor outcomes, leisure, cons. utility.
4. Consumption utility function needs curvature (!!)
5. Income gains and losses that capture accompanying shocks.
6. Standard data needs to permit replication.
7. Calculations of WTP and WTAP to get a range of values.
8. Additional factors, e.g., expectations, human capital, etc.

Valuing QALYs:

→ Bottom line: this paper suggests the weakness of the common method.

→ Perhaps better to introduce a far more complicated, but agreed upon model.

→ Consider, e.g., how epidemiological and economic models can be unified.