

Measuring Growth and Development

Jonathan Colmer

University of Virginia

The Importance of Data

- We don't need data to realize that there are large differences in living standards across/within countries.
- But we do need data if we want to comprehensively and systematically understand differences in living standards.
- The minimum requirement is data that allows us to compare real living standards across (and within) countries and over time ... easier said than done.

What should we care about?

Roadmap

1. GDP's Origin Story
2. GDP in Practice
3. Beyond GDP
4. GDP and Welfare

A Short History of GDP

- The most common indicator of development is per capita GDP and the rate of its growth.
- It was invented in the 1940s — “One of the Great Inventions of the 20th Century” ([U.S. Bureau of Economic Analysis, 2000](#)).
- A measure of production, not welfare.

From National Income to National Accounts

- The Great Depression increased demand for national statistics.
 - Simon Kuznets, developed methods introduced by Colin Clark in the UK, and applied them to the U.S. economy.
 - He was a meticulous collector and assembler of data.
 - Paid very careful attention to the circumstances in which stats were gathered and what their flaws might be.
 - First report to Congress (Jan, 1934) showed that America's national income had halved between 1929 and 1932.
 - National income estimates massively increased the scope for policy (Recovery Program, Roosevelt's 1938 supplemental budget).
 - Hoover had to make do with industrial stats like the share price indexes and freight car loadings, which were much less compelling as a call to action.

Kuznets' Vision

Kuznets saw his task as how to measure national economic welfare, rather than just output.

"It would be of great value to have national income estimates that would remove from the total the elements which, from the standpoint of a more enlightened social philosophy than that of an acquisitive society represent dis-service rather than service.

Such estimates would subtract from the present national income totals all expenses on armament, most of the outlays on advertising, a great many of the expenses involved in financial and speculative activities, and what is perhaps most important, the outlays that have been made necessary in order to overcome difficulties that are, properly speaking, costs implicit in our economic civilization.

All the gigantic outlays in our urban civilization, subways, expensive housing, etc., which in our usual estimates we include at the value of the net product they yield on the market, do not really represent net services to the individuals comprising the nation but are, from their viewpoint, an evil necessary in order to be able to make a living."

- This didn't go down too well. . . The government didn't want a measure that showed the economy shrinking even if government expenditures were expanding output.

From National Income to GDP

- Led to heated debates on the meaning of economic growth and why statisticians were measuring it.
 - Department of Commerce's view was that the aim was to have a measurement that was useful to the government in running its fiscal policy.

"By including all government purchases as part of national products, the GNP statistics established the role of national government in the economy as that of an ultimate consumer, that is as a purchase of goods and services for final use."

- Kuznets argued this method tautologically ensured that fiscal spending would increase measured economic growth regardless of whether it actually benefited individuals' economic welfare.
- Kuznets lost.

An Increasing Role of Government

- For two centuries “the economy” had been the private sector.
- Government had played a small role in economic life and mainly featured to raise taxes to pay for wars.
- Its role expanded steadily over the centuries, due to the provision of public goods + historic roles of defense and justice.
- GDP came into existence because of the increasing conception that government added to national income rather than subtracted from it.
- Led to Keynesian macroeconomic theory becoming the basis for how governments ran their economies in the postwar era.

What is GDP?

- There is no entity called GDP, it is an empirical construct.
- The system for measuring GDP and its components is very complicated and has become more complicated over time.
 - The first guide, which all countries were supposed to follow, was published in 1953 and had fewer than fifty pages.
 - The 2025 SNA document has 1,260 pages.
 - Very few people exist who really understand how GDP statistics are actually constructed. . .
- GDP can be measured in three ways:
 - Add up all the output: value added
 - Add up all the expenditures: $C + I + G + (X - M)$
 - Add up all the income: compensation + rent + profits + interest + (taxes - subsidies) + depreciation

How would you measure these in practice?

The Penn World Tables

- The creation of the Penn World Tables (PWT) was revolutionary.
- Provided the first real opportunity to systematically compare real living standards across countries and over time with broad coverage.

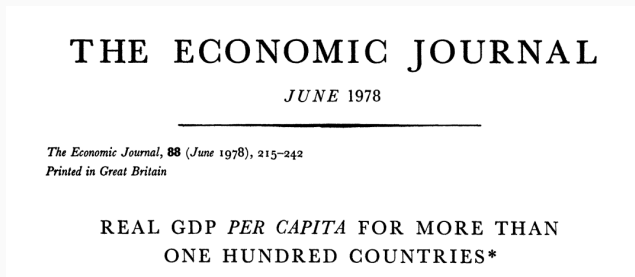


Figure 1: *

Kravis, Heston, and Summers (1978)

Why might it be hard to measure GDP in a country over time?

Why might it be hard to measure GDP in a country over time?

Why might it be hard to measure GDP across countries at a point in time?

The “Real” Challenge

- Measuring nominal GDP in a given country at a given time is hard enough as discussed.
- Measures of *real* GDP, accounting for differences in the cost of living, are better measures of differences in living standards but much harder to produce.
- Producing measures of real GDP across countries at a given point in time is especially difficult.
- Traditional strategy: use exchange rates to convert nominal GDPs into a common currency and use this as a measure of relative living standards.

What about Purchasing Power?

- Exchange rates deal with the units problem, but what about non-tradables?
- The purchasing power of \$1 varies substantially across countries.
- Measuring purchasing power across countries is incredibly difficult.
 - Standard approach is to use a “common basket”. Does one exist?
 - “Advanced economy” items can be very expensive in low income countries, leading one to conclude that purchasing power is very low.
 - The International Comparison Program (ICP) update purchasing power comparisons every few years. Can have a large effect on quantitative assessments of differences in living standards.
 - See [Deaton \(2010\)](#) and [Feenstra, Inklaar, and Timmer \(2015\)](#) for theory and practice

Almas (2012): Measuring PPP Bias using Engel Curves

- PPP bias likely understates cross-country inequality. Is it a big problem?
- **Basic idea:** if two households in different countries have the same PPP adjusted income and the same demographic characteristics (the same age and number of children and adults), any difference in the budget share for food can be attributed to PPP bias.
- Why? Engel curve for food is log-linear and stable, both over time and across societies.

Almas (2012): Measuring PPP Bias using Engel Curves

- P_j , the composite price of consumption in country j , can be expressed as

$$P_j = P'_j \times E_j$$

where P'_j is the biased macro price of consumption given in the PWT and E_j is the PPP bias for this country.

- The Engel curve of the Almost Ideal Demand System can be expressed as

$$m_{h,j} = a + b(\ln y_{h,j} - \ln P'_j) + \gamma(\ln P_{r,j}^f - \ln P_{r,j}^n) + \theta X_{h,j} + \sum_{j=1}^N d_j D_j + \epsilon_{h,j}$$

where D_j is the country dummy. The country dummy coefficient, d_j , is a function of the PPP bias, E_j , and the coefficient for the logarithm of household income, b :

$$d_j = -b \ln E_j, \text{ therefore, } E_j = \exp\left(-\frac{d_j}{b}\right)$$

Almas (2012): Measuring PPP Bias using Engel Curves

TABLE 3—THREE DIFFERENT INCOME MEASURES

	Y^{PWT}	Y^{EC}	Y^{EX}
UK	15,088	15,088	15,088
Spain	11,935	11,507	10,162
Hungary	5,651	3,363	2,780
Brazil	4,818	3,899	3,235
Bulgaria	3,027	1,073	1,106
Peru	2,839	818	1,575
Azerbaijan	1,739	939	303
Côte D'Ivoire	1,471	491	634
Nepal	829	211	151
Tanzania	372	97	111

Note: The table shows the income measured by PWT, EC incomes, and EX incomes for the ten base countries.

$$Y_j^{EX} = Y_j; Y_j^{PWT} = \frac{Y_j}{P_j^*}; Y_j^{EC} = \frac{Y_j}{P_j^* E_j}$$

- PPP bias leads to substantial underestimation of international inequality

Argente, Hsieh, and Lee (2023)

- Measurement is hard.
- PPP bias may exist for many reasons: Issues of imputation, sampling, quality, and variety.
- [Argente, Hsieh, and Lee \(2023\)](#) quantify potential biases in the ICP using a new decomposition framework to get at these components.
- They do this using Nielsen scanner data for U.S. and Mexico.
 - ICP overstates Mexican prices due to imputation, sampling, and variety bias
 - ICP understates Mexican prices due to quality bias

Open Questions

- Are poor countries cheap for everyone?
 - See [Handbury and Weinstein \(2014\)](#) and [Handbury \(2021\)](#) for cities in the US.
 - Distributional CPI for the U.S. by [Jaravel \(2024\)](#)
- What are the causes and consequences of price differences across countries?
 - High price of cement in Africa comes from higher marginal costs and higher markups ([Leone et al., 2024](#))
- How does informality and household production shape our understanding of cost of living and GDP?

When would $\text{GDP} = \text{Welfare}$?

Beyond GDP?

- Despite its popularity GDP is limited as a measure of welfare or well-being:
 - Excludes non-market goods & services outside the “production boundary”, e.g., unpaid housework/care, environmental goods and services, non-use values, etc.
 - It counts many “bads” as positive, e.g., defensive expenditures and disaster recovery.
 - Doesn’t account for risk and uncertainty, i.e., security.
 - Rights and freedoms are not reflected in GDP.
 - Distributional considerations are not reflected, e.g., poverty and inequality.
- What are the alternatives? How bad a proxy for well-being is it?

Subjective Well-Being

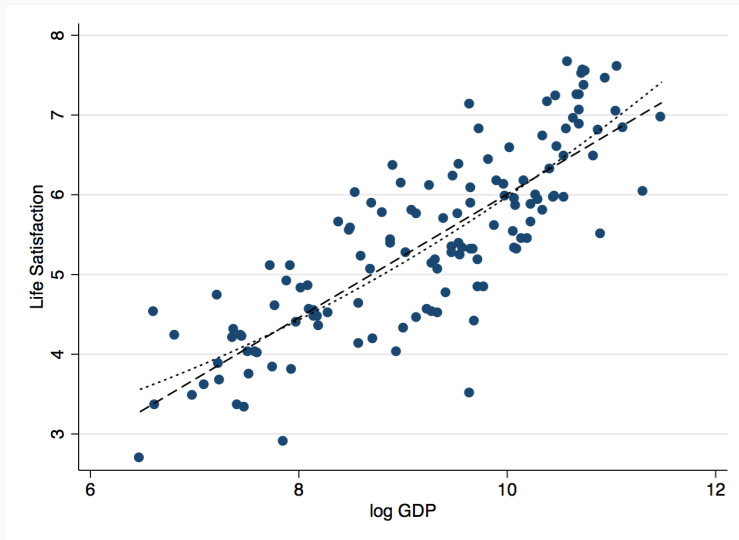
- What about indicators like life satisfaction and happiness?



Subjective Well-Being: What Are We Measuring?

- Different approaches:
 - **Evaluation (Life Satisfaction):**
 - Captures an overall judgment of one's life as a whole.
 - **Experience (Affect/Mood):**
 - Flow utility from time use, health, social contact, environment
 - **Eudaemonia (Purpose/meaning):**
 - Captures value alignment/capabilities; not reducible to momentary affect.
- Measures can be general or domain-specific, e.g., in the context of health, work, finances, relationships, etc.
- Beyond SWB, survey methods can be very effective for eliciting attitudes, preferences, and beliefs ([Stancheva, 2022](#); [Haaland et al., 2025](#))

Life Satisfaction vs. GDP per Capita



The Easterlin Paradox

- Within the cross-section, at any given point in time life satisfaction is strongly correlated with income.
- However, over the long-run in time series data, there seems to be diminishing marginal “happiness” returns to income above a threshold.
 - Could be an artifact of trying to correlate a stationary time series (life satisfaction) with a non-stationary time series (GDP per capita).
 - Correlation with GDP growth (a stationary time series) remains positive over time (Stevenson and Wolfers, 2013; Deaton, 2008), although this is debated (Easterlin and O’Conner, 2022).

Issues with Subjective Well-Being

- No objective yardstick
- Depends entirely on individual perception
- Highly sensitive to minor events.
- Hedonic adaptation leads people to evaluate in relation to a reference point or norm.
- Scale norming (re-anchoring) leads to folks using qualitatively different scales to reply to the same questions across survey waves.
- Happiness is not a substitute for Poverty eradication.

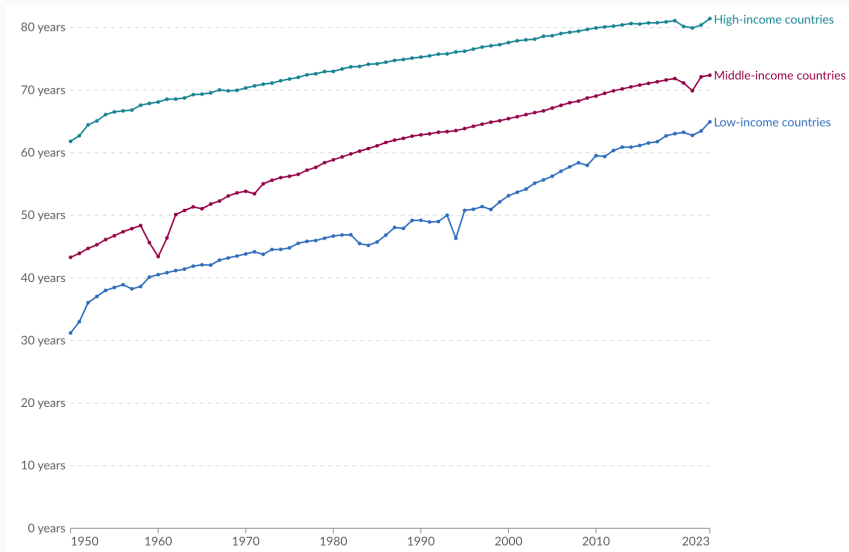
Capabilities

- Amartya Sen's capability theory – development should expand people's "budget sets", broadly defined as things they are able to do.
- The goal of development becomes the gradual enhancement of an individual's capability:
 - An individual's well-being cannot be determined by policies or cardinally measured
 - It's hard to evaluate what people do with their budget and whether it makes them happy.
 - However, enhancing someone's capability will enable them to realize their goals.

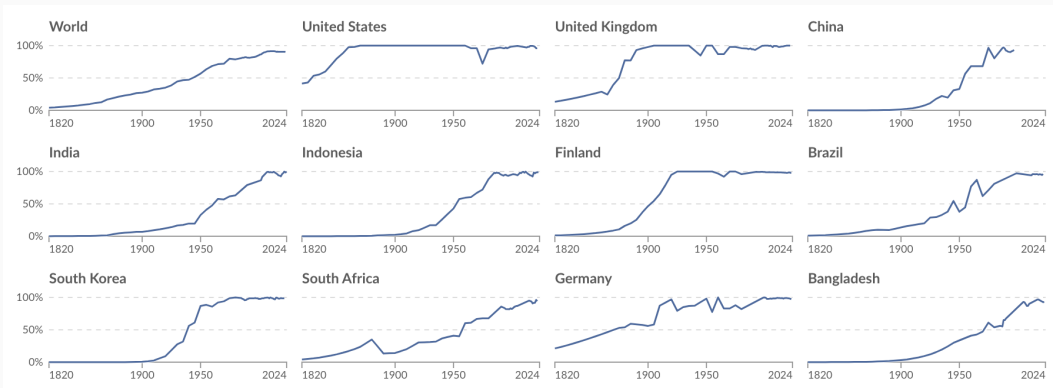
The Human Development Index

- Formulated by Mahbub ul Haq and Amartya Sen and included in the UN Development Report since 1990.
- It is an average of three different indices – per capita income, the mortality rate, and an index of education.
- Education and health do not merely contribute to the rise of national income, they are also important indicators of quality of life.

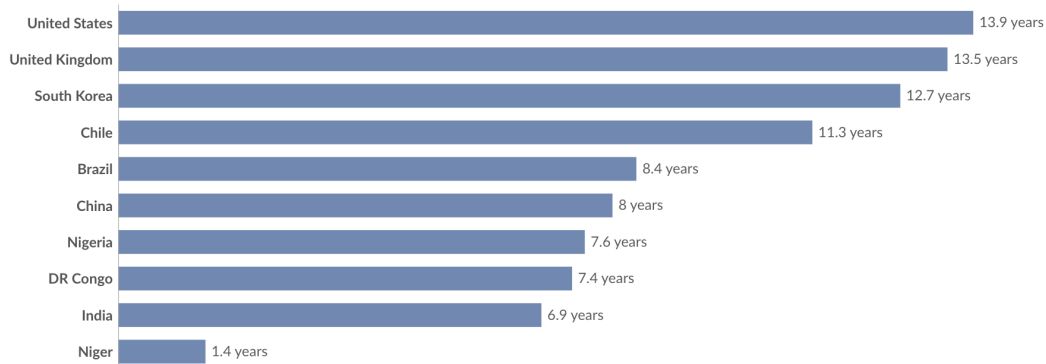
Life Expectancy



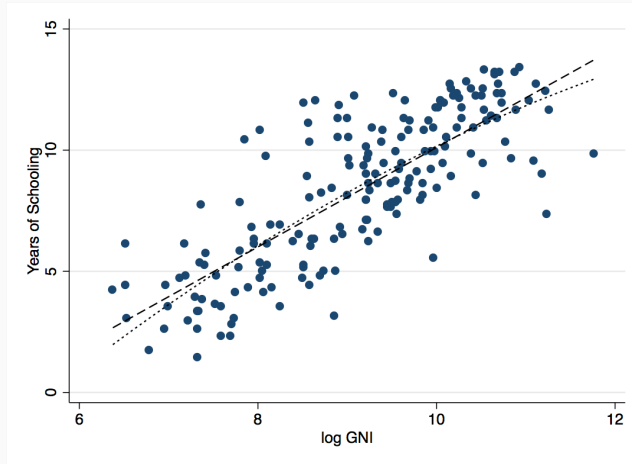
Primary School Enrollment



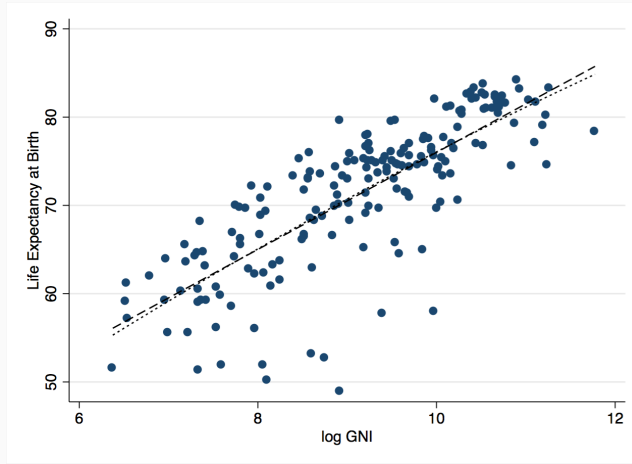
Years of Schooling (2023)



The Human Development Index



The Human Development Index



Beyond Health and Education

- Capability cannot be determined by education and health alone
- Capability depends considerably on the rights that individuals are permitted.
- Violations may include:
 - curbing freedom of expression
 - discrimination

Inequality

- Per capita income fails to capture inequality among sections of the population.
- Economists tend to be more concerned about inequality of opportunity.
 - Unequal societies can limit individuals from reaching their full potential.
 - Too much equality can limit incentives.
- There are reasons to be concerned about inequality as an end in itself.

Inequality

- Generally, inequality is a relative indicator while poverty is an absolute one.
 - Absolute poverty is objectionable by most welfare criteria.
 - The importance of inequality depends on the social welfare function.
- What if there is no absolute poverty. Is inequality still a bad thing?

From GDP to Welfare

- Utility depends on consumption, leisure, life expectancy, inequality, etc.
- But GDP per capita “only” measures income. . .
- The U.S. has higher private consumption, but France has:
 - More leisure
 - Less inequality
 - More public consumption
 - Longer life expectancy
- Which country has higher welfare?

The Veil of Ignorance Experiment

- We are “behind the veil” waiting to be born:
 - One of us will be born “unproductive” and have low consumption.
 - The rest will be productive and have high consumption.
 - We don’t know who will be who.

Veil of Ignorance Experiment

- Which would you choose:

- 1) Social contract: we commit behind the veil to help the person that turns out to be unproductive ex post.
- 2) Autarky: we let the uncertainty realize and bad luck to whoever turns out to be unproductive.

Veil of Ignorance Experiment

- **Result:** As long as people are risk-averse, they are strictly better off under the social contract.
- Thus, Rawls argues that we should hold people to this contract *ex post*.
- That is, we should enact policies that redistribute wealth to the neediest.

- Not arguing for redistributive policies.
- Instead asking that a representative individual for each country go behind the veil of ignorance.
- “Rawls” will live entire life as a random individual in some country/year.

Expected Lifetime Utility

$$U = E \left[\sum_{a=1}^{100} \beta^a u(C_a, \ell_a) S(a) \right]$$

- C_a : Consumption at age a
- ℓ_a : Leisure at age a
- $S(a)$: Probability of survival to age a
- β^a : Discount factor (less than one)

Consumption-Equivalent Welfare

- **Choice one:** live in country j , born as one randomly selected person in j
 - Face j 's survival rates; age of death is random
 - Draw randomly from j 's consumption and leisure distribution
- **Choice two:** Move to the U.S., live as one randomly selected person, get fraction λ times consumption.
- Consumption-equivalent welfare of j is defined as the λ that makes “Rawls” indifferent between **Choice one** and **Choice two**.

Consumption-Equivalent Welfare

If,

$$E \left[\sum_{a=1}^{100} \beta^a u(C_a, \ell_a) S(a) \right] = E \left[\sum_{a=1}^{100} \beta^a u(\lambda C_a^{US}, \ell_a^{US}) S^{US}(a) \right]$$

For some number λ , then λ is the welfare of country j .

Operationalizing the Welfare Metric

- Need distribution of consumption and leisure
 - Use household surveys for 13 countries
- Need survival rates by age,
 - Take from UN statistics
- Need functional forms for utility function
- Need to calibrate utility function

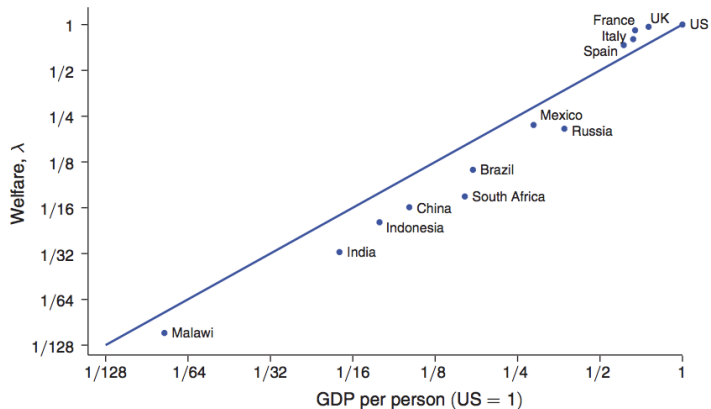
Utility Function

$$u(C_a^j, \ell_a^j) = u^* + \log(C_a^j) + \log(\ell_a^j)$$

- Key parameter is u^*
- u^* is the fixed utility value just for being alive.
- What is the value of staying alive one extra year?
- “Value of Statistical Life” (VSL): In the U.S., the expected value of remaining life for a 40 year old is estimated to be around \$6 million dollars ([Hall and Jones, 2007](#))

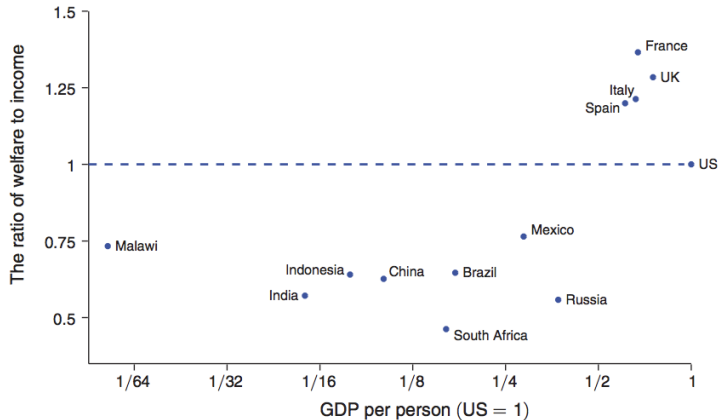
Welfare and Income are highly correlated ($\hat{\beta}=0.95$)

Panel A. Welfare and income are highly correlated at 0.98



But welfare does meaningfully differ from Income: by 35% on average

Panel B. But this masks substantial variation in the ratio of λ to GDP per capita



U.S. vs. France in 2005

				<i>Decomposition</i>				
	λ	Income	Log Ratio	Life Exp.	C/Y	Leis.	Cons. Ineq.	Leis. Ineq
U.S.	100.0	100.0	.000	.000	.000	.000	.000	.000
				77.4	.854	834	.539	1082
France	91.1	67.2	.305	.149	-.115	.083	.064	.124
				80.1	.768	527	.417	743

- Higher taxes and generous social safety net benefits may reduce work effort and GDP
- But these programs have benefits that are not measured by GDP...

U.S. vs. Western Europe

	λ	Income	Log Ratio	Life Exp.	<i>Decomposition</i>			
					<i>C/Y</i>	<i>Leis.</i>	<i>Cons. Ineq.</i>	<i>Leis. Ineq</i>
U.S.	100.0	100.0	.000	.000	.000	.000	.000	.000
				77.4	.854	834	.539	1082
U.K.	96.6	75.2	.250	.083	-.055	.073	.052	.097
				78.7	.815	570	.449	824
France	91.1	67.2	.305	.149	-.115	.083	.064	.124
				80.1	.768	527	.417	743
Italy	79.6	66.1	.185	.175	-.203	.078	.060	.075
				80.7	.697	567	.415	899
Spain	72.8	61.1	.175	.128	-.096	.070	-.000	.073
				79.1	.759	609	.528	898

U.S. vs. Lower-Middle Income Countries

				<i>Decomposition</i>				
	λ	Income	Log Ratio	Life Exp.	C/Y	Leis.	Cons. Ineq.	Leis. Ineq
U.S.	100.0	100.0	.000	.000	.000	.000	.000	.000
				77.4	.854	834	.539	1082
Mexico	22.0	28.6	-.262	-.149	-.011	-.010	-.088	-.005
				74.2	.844	862	.622	1092
Russia	21.1	37.0	-.563	-.480	-.130	.035	-.021	.032
				67.1	.743	736	.498	1023
Brazil	11.2	17.2	-.425	-.229	-.002	.005	-.204	.006
				71.2	.835	807	.713	1039
S. Africa	7.7	16.0	-.738	-.521	.036	.054	-.302	-.006
				60.9	.852	623	.850	1079

U.S. vs. Lower Income Countries

	λ	Income	Log Ratio	<i>Decomposition</i>				
				Life Exp.	C/Y	Leis.	Cons. Ineq.	Leis. Ineq
U.S.	100.0	100.0	.000	.000	.000	.000	.000	.000
				77.4	.854	834	.539	1082
China	6.4	10.1	-.458	-.163	-.261	-.016	-.004	-.014
				71.7	.647	873	.503	1093
Indo.	5.1	7.8	-.425	-.318	-.098	-.001	.032	-.041
				67.2	.774	836	.443	1170
India	3.3	5.6	-.528	-.407	-.120	-.019	.046	-.028
				62.8	.764	876	.428	1132
Malawi	1.0	1.3	-.249	-.326	.092	-.020	-.024	.028
				50.4	.920	867	.534	991

Welfare Growth

- Average Growth rates, 1980s–2000s
 - Welfare growth: 3.1%
 - Per capita GDP growth: 2.1%
- Welfare growth is 50% faster because of declining mortality.
 - At 3% welfare doubles every 24 years (20-fold over a century)
 - At 2% income doubles every 35 years (7-fold over a century)

Overview

- GDP per capita and welfare are highly correlated in both levels and growth rates.
- But there are important differences between income and welfare
 - Western Europe looks much closer to U.S. living standards
 - Most other countries are further behind, primarily due to lower life expectancy
 - Growth is 50% faster than we thought, largely because of significant declines in mortality.
- **We're probably still not fully capturing welfare. What else matters?**

Population and Welfare: The Greatest Good for the Greatest Number

- Much of what we've discussed has considered things in per capita terms
 - Puts zero weight on the number of people.
- Consider, two countries with the same growth path. One has constant N but rising c , the other has rising N but constant c
 - Example: Japan is $6\times$ richer p.c. than in 1960 and has $1.3\times$ more people. Mexico is $3\times$ richer but has $3\times$ more people.
- **Key question:** How much has population growth contributed to aggregate welfare growth? (Adhami, Bils, Jones, and Klenow, 2025)

Philosophy and Social Welfare

- Longstanding debate: both “average” and “total” views considered valuable.
 - Critique of “average” utilitarian approach: sadistic conclusion
 - Critique of “total” utilitarian approach: repugnant conclusion ([Parfit, 1984](#))
 - [Zuber et al. \(2020\)](#) (29 philosophers and economists) argue that this is no decisive.
- Focus of [Adhami et al. \(2025\)](#) is to measure how the “average” and “total” views differ in practice.

The Total Utilitarian Welfare Criterion

- Aggregate flow utility:

$$W(N_t, c_t) = N_t \cdot u(c_t)$$

- Existence $\Rightarrow u(c)$, non-existence $\Rightarrow 0$
- Growth in Consumption-Equivalent Aggregate Welfare:

$$\frac{u(c_t)}{u'(c_t)c_t} \cdot \frac{dW_t}{W_t} = \underbrace{\frac{u(c_t)}{u'(c_t)c_t}}_{\equiv \nu(c)} \cdot \frac{dN_t}{N_t} + \frac{dc_t}{c_t}$$

- $\nu(c)$ = value of having one more person live for a year – expressed relative to one year of per capita consumption.
- 1 pp of population growth is worth $\nu(c)$ pp of consumption growth.
- [Adhami et al. \(2025\)](#) calibrate $\nu(c) \approx 4.87$ based on estimates of the VSL.

Evaluating the Contribution of Population Growth

$$g_\lambda = \nu(c)g_N + g_c$$

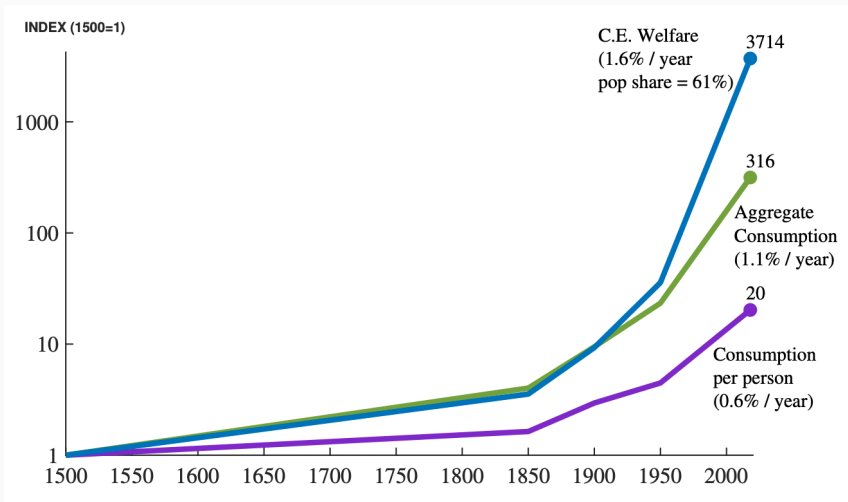
- $\nu(c) = 0$ is the per capita case, $g(\lambda) = g_c$
- $\nu(c) = 1 \Rightarrow$ CE-welfare growth is aggregate consumption growth
- $\nu(c) = 3$ or $5 \Rightarrow$ puts more weight on population growth

Decomposing Welfare Growth in Select Countries, 1960–2019

	g_λ	g_c	g_N	$v(c)$	$v(c) \cdot g_N$	Pop Share
Mexico	8.6	1.8	2.1	3.4	6.8	79%
Brazil	7.9	3.1	1.8	2.8	4.8	61%
South Africa	7.8	1.4	2.1	3.1	6.4	82%
United States	6.5	2.2	1.0	4.4	4.3	66%
China	5.8	3.8	1.3	1.8	2.0	34%
India	5.4	2.6	1.9	1.6	2.8	52%
Japan	4.9	3.2	0.5	3.8	1.7	34%
Ethiopia	4.4	2.5	2.7	0.7	1.9	44%
Germany	3.7	2.9	0.2	4.0	0.8	22%

Pop share of CE-Welfare Growth $\geq 50\%$ in 77 out of 101 countries.

World Cumulative Growth, 1500–2018



Why might this matter?

- Per capita vs. total approaches lead to different assessments of adverse events
 - Black Death, HIV/AIDS (“Gift of the Dying”, [Young \(2005\)](#)), Covid-19.
 - Per capita focus emphasizes outcomes for survivors
 - A total utilitarian perspective accounts for direct loss of human life.
- Interesting policy implications:
 - Immigration
 - Fertility
 - Inequality
 - Discounting (larger future populations \Rightarrow bigger weight on future welfare)