

Session 2

Beyond yourself, but how?

Welfare in the large vs. small

Instructor: Guo Xu

Reminders

- Start discussing group project with your assigned group
- Assignment 2 (Dragonfly) posts today
- **Ungraded/optional**, recommended to build up your skills
- **Pattern recognition:** parallels in utilitarian analysis
- To be discussed during GSI section Friday after the break (4/5)

Today

- Start our normative trek
- Introduce utilitarianism
- Apply it to Ziqitza case

Previously, in Ethics

Principles of moral architecture

- Leaders get to design games
- Some games incentivize bad actions
- Game features interact with behavioral biases
 - Moral disengagement:
Diffusion of responsibility, slippery slope
- Need ethical values as antidote for bad incentives and behavioral biases – but *which* values?

Your non-negotiables

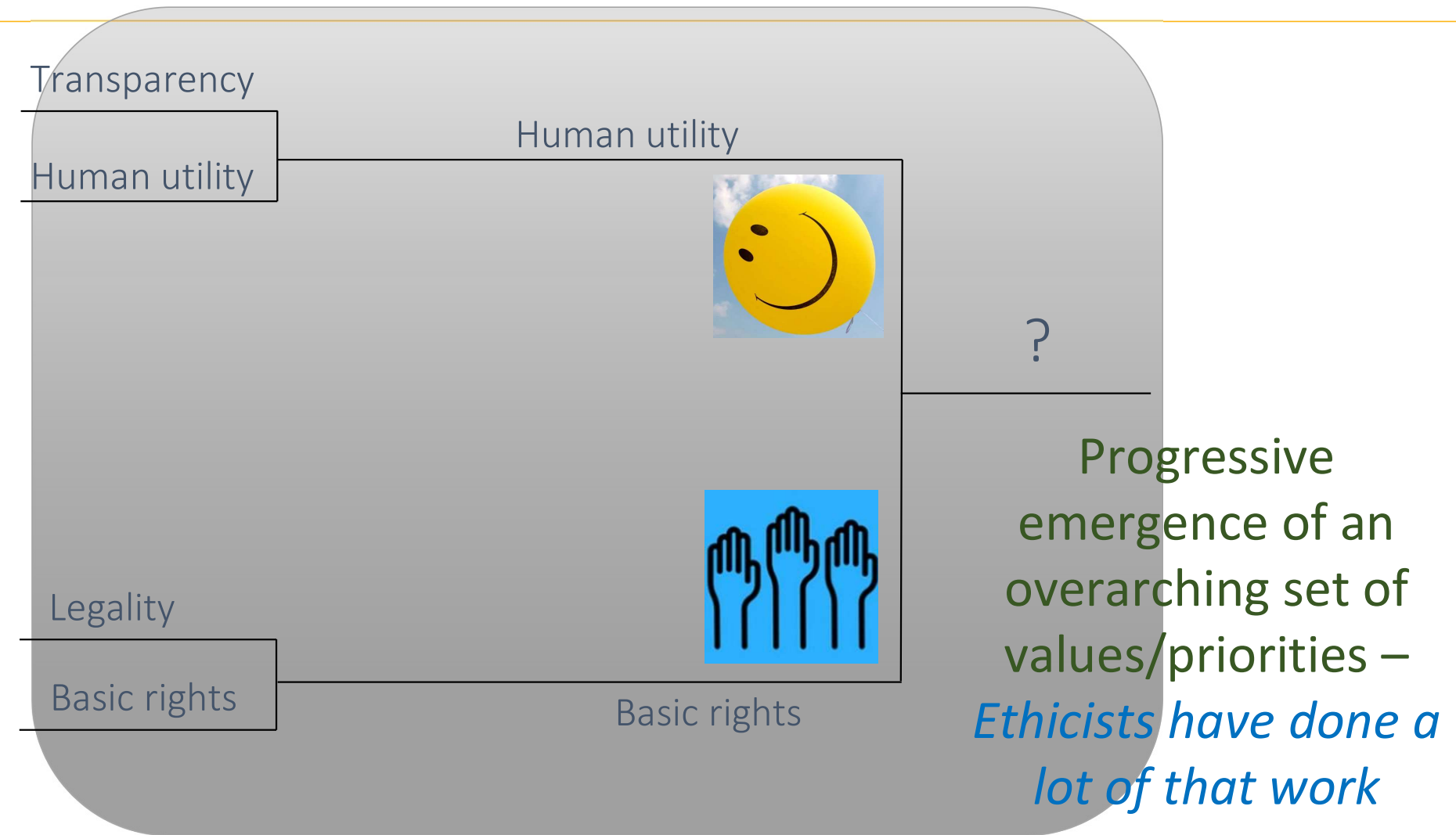


- What values would you fight the hardest not to compromise?

From description to decision

- Can think of Ethics as a “tournament” of moral ideas
 - Normative ethical frameworks as **decision tools**
 - **Proximate vs. fundamental** values
- **Ethics:** a systematic effort to identify values with *normative* power

Ethics: a “tournament” for moral values

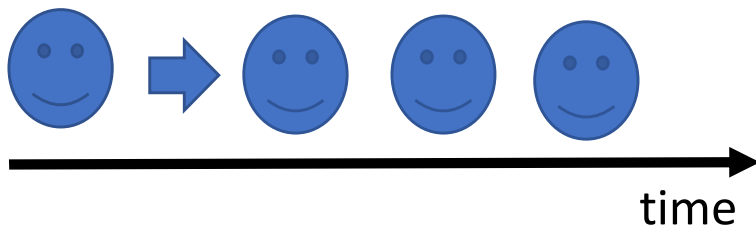




What makes values compelling?

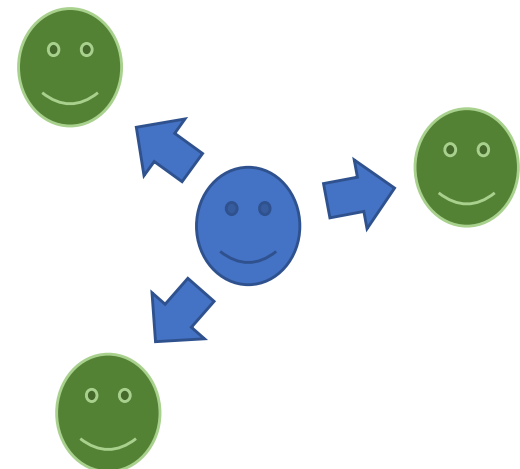
1. Interchangeability of perspectives - **impartiality**
2. A way to factor-in **others**
3. **Non-absolutist** – same decision criterion may yield different conclusions in different settings
 - Beyond yourself: morality is about internalizing other “selves”

My future selves – overcoming self-control



BerkeleyHaas

Other people's



Where we stand, where we go

Week	Topic	Utilitarianism	Duty ethics	Justice as fairness	Justice as entitlements	Virtue ethics
2	Corruption					
3	Workplace discrimination					
4	DEI & Sustainability					
5	CSR & Social media					
6	Gene editing					

Self-driving cars technical vs. ethical problems

The New York Times

Self-Driving Tesla Was Involved in Fatal Crash, U.S. Says



The New York Times

Self-Driving Uber Car Kills Pedestrian in Arizona, Where Robots Roam

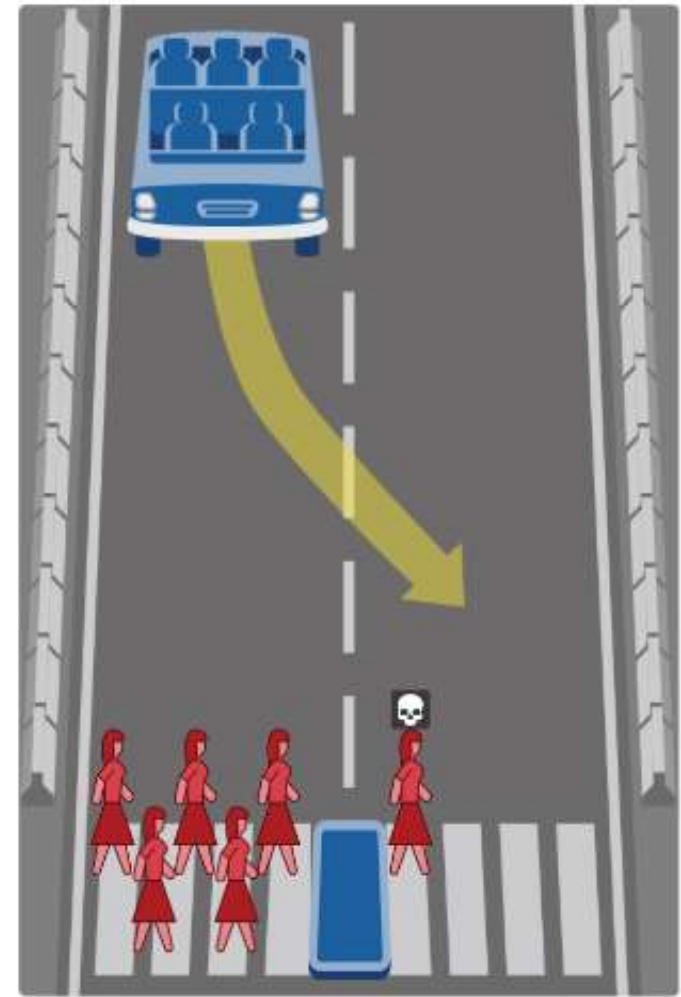
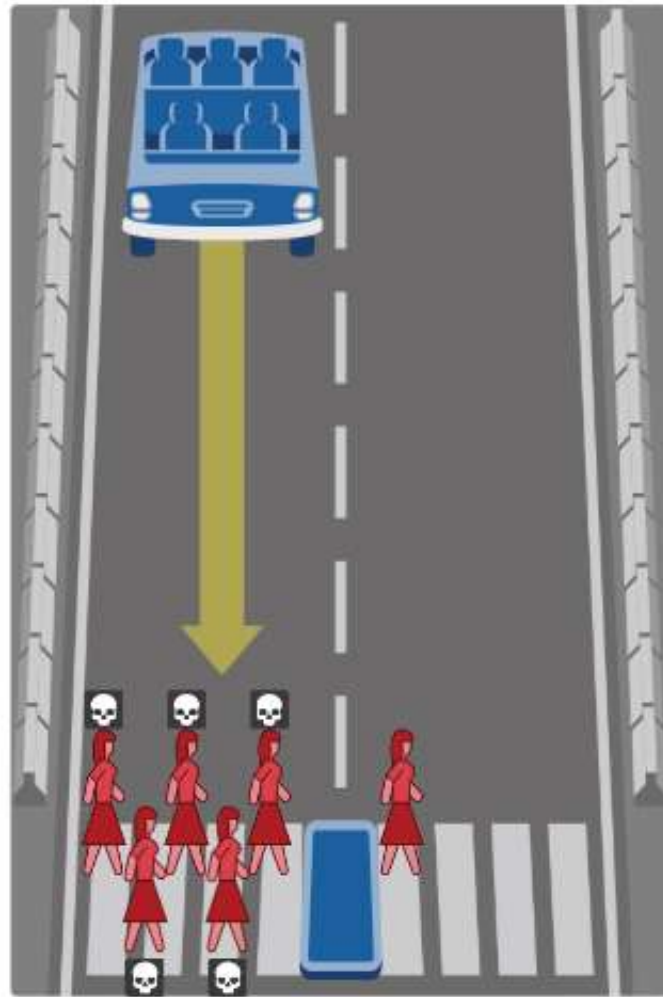


**Driver hits pedestrian, pushing her into path
of self-driving car in San Francisco**

Should the self-driving car go straight or swerve?

Five vs One

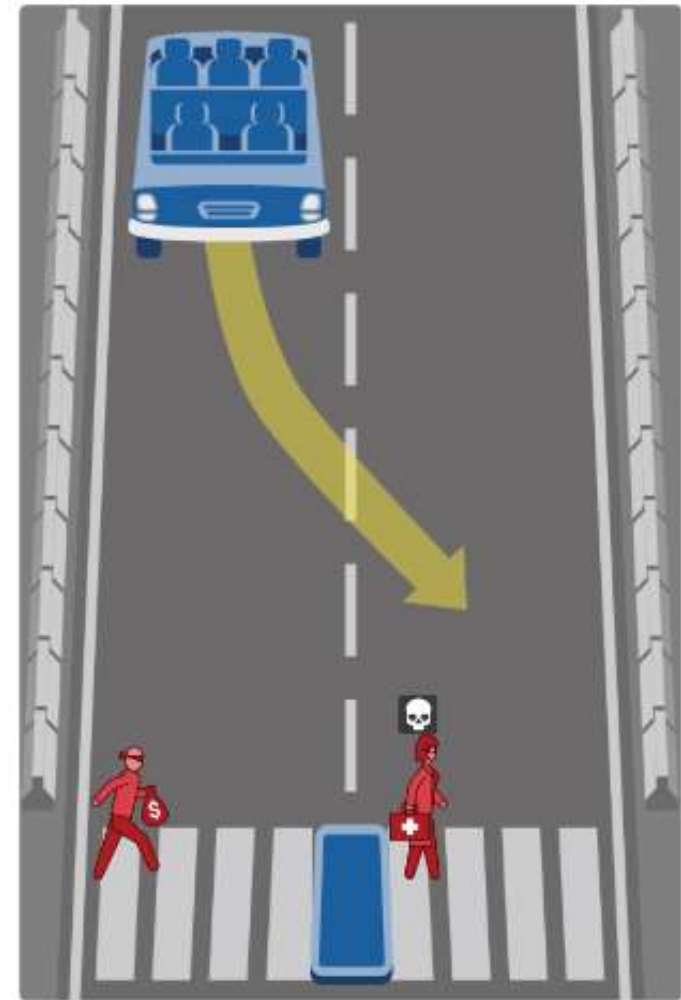
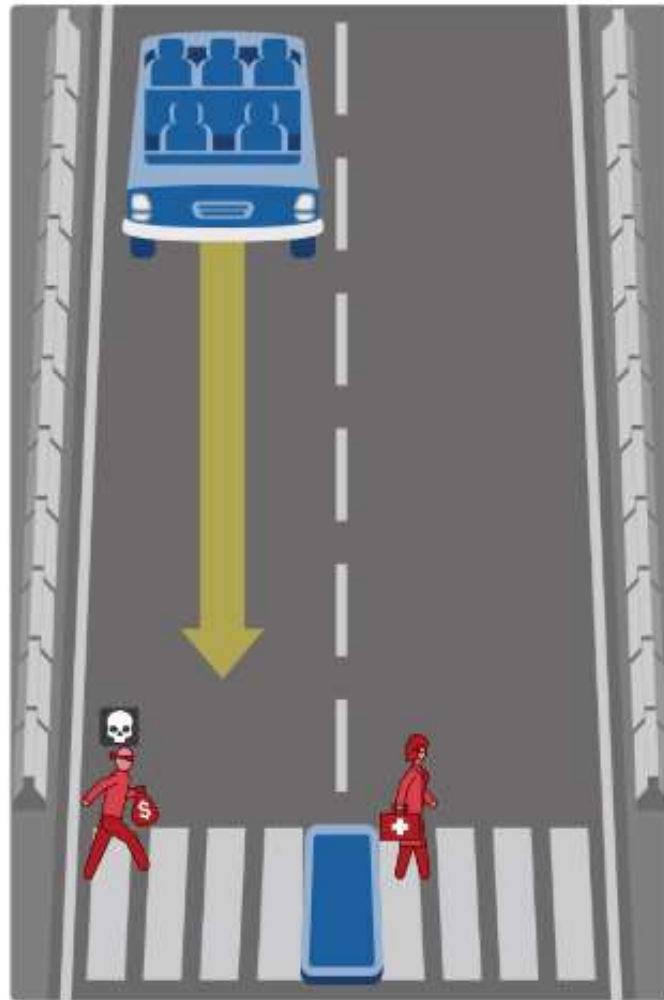
- A. Go straight
- B. Swerve



Should the self-driving car go straight or swerve?

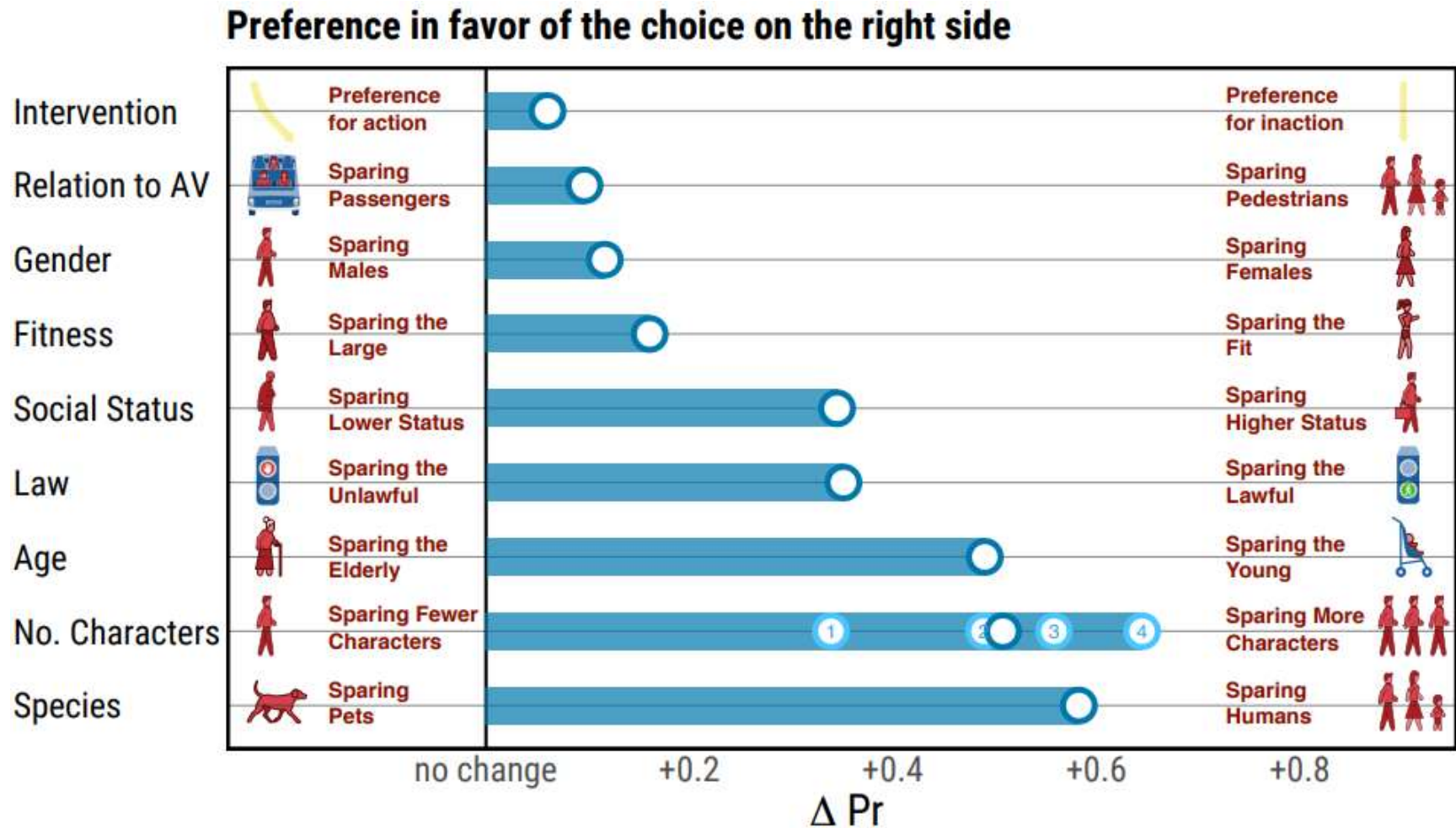
Criminal vs. Doctor

- A. Go straight
- B. Swerve



The moral machine experiment

Awad et al, Nature 2018



Value from human happiness



The right action is that which produces the greatest amount of pleasure for the greatest number of people.

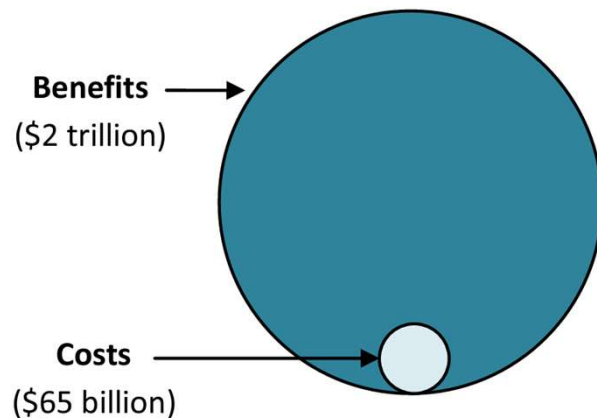
Basics of utilitarianism

- Objective: create the greatest possible well-being or *utility*
- *Consequentialist* view: an action is good if it has good *consequences* for well-being
- This does not imply selfish hedonism
 - The principle applies to aggregate or social well-being, **not one's own**
 - A social utility NPV, includes **both long and short-run**
- Historically linked with social reformers, today Effective Altruism
 - Widely used in public policy

$$\text{Max}_x \sum_{t=1}^{\infty} \sum_{i=1}^N d^t u_{it}(x)$$

Example: evaluating the Clean Air Act

“Our central benefits estimate exceeds costs by a factor of more than 30 to one, and the high benefits estimate exceeds costs by 90 times. Even the low benefits estimate exceeds costs by about three to one.”
— EPA 2023



The 1990 Clean Air Act Amendments prevent:

	Year 2010 (in cases)	Year 2020 (in cases)
Adult Mortality - particles	160,000	230,000
Infant Mortality - particles	230	280
Mortality - ozone	4300	7100
Chronic Bronchitis	54,000	75,000
Heart Disease - Acute Myocardial Infarction	130,000	200,000
Asthma Exacerbation	1,700,000	2,400,000
Emergency Room Visits	86,000	120,000
School Loss Days	3,200,000	5,400,000
Lost Work Days	13,000,000	17,000,000

Challenges

- Aggregating net effect – difficult to determine net effect when some gain and others lose
- Data-heavy – requires details about magnitudes of effects
 - PUB POL 259: Benefit-Cost Analysis and sustainability research at Haas
- Advantage: structured, objective approach – makes clear what the assumptions are

Pop quiz on utilitarianism

All Utilitarians care about welfare. But while Act-Utilitarians consider a world where everyone acts uniformly (“what if we all did this vs that”), a Rule-Utilitarian considers the impact of an isolated individual action taking as given the actions of others.

A: True

B: False

Two forms of utilitarianism

- Act-Utilitarianism
 - Choose the action that, given the action of others, will most increase aggregate social welfare
- Rule-Utilitarianism
 - Choose the action that, if taken by everybody when following a rule, would most increase aggregate social welfare
 - Linked to John Harsanyi, Haas

el Prize.



Ziqitza Healthcare



Facts?

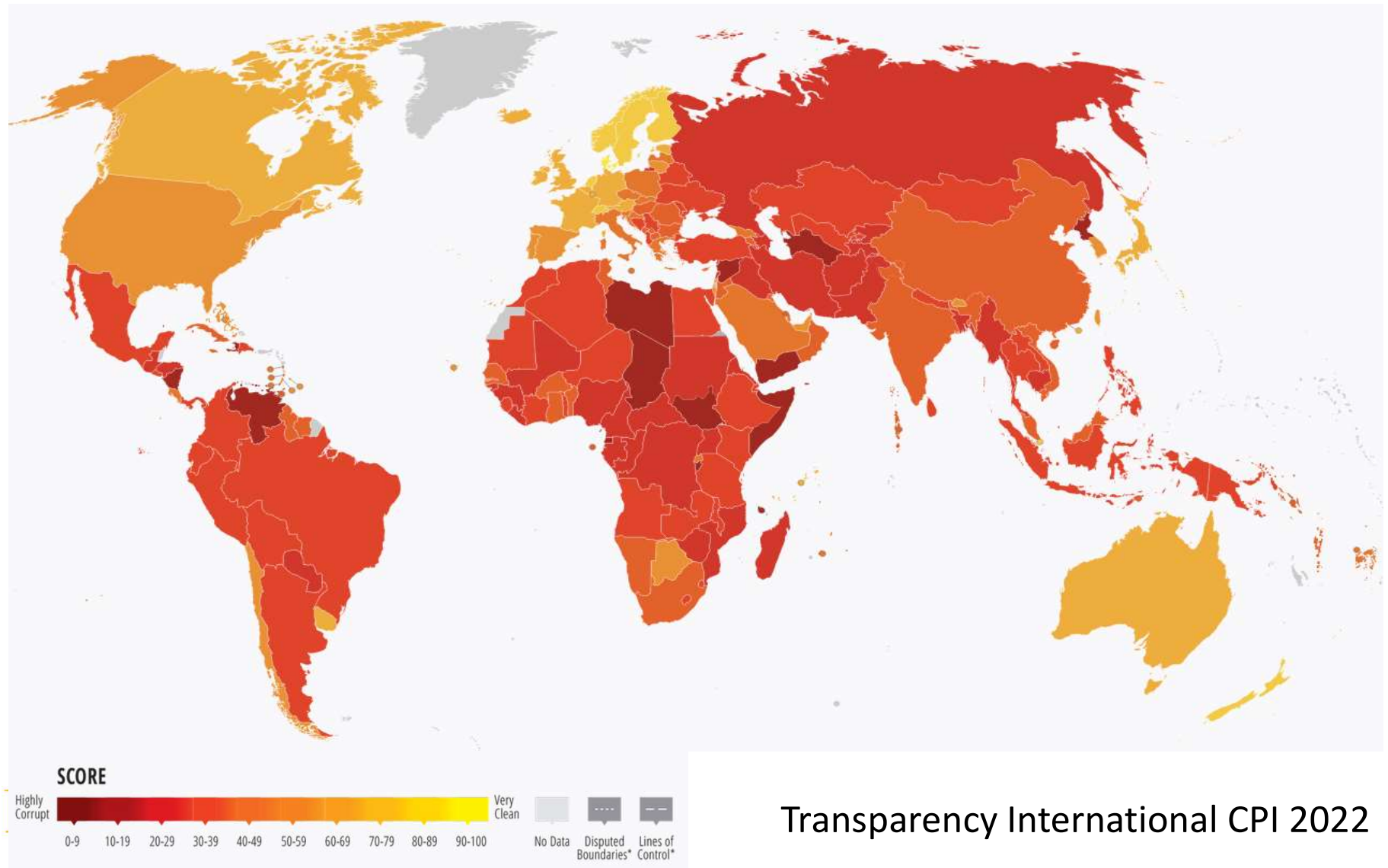
- Key players
- Goals of the organization
- Important contextual details
- Predicament

Ziqitza

- Founded in 2002 to provide high-quality ambulance service in India
- Mission-driven social enterprise
 - Providing service for disadvantaged populations
 - 20% patients have subsidized rates
 - Explicit pledge to refuse paying bribes
- Today, operating in 28 states of India, 3500+ ambulances, 12k+ employees, 47 million served
- High prevalence of [corruption](#) (TI 2020)
 - 39% report paying bribe for public service in past 12 months
 - 75% of population support “anti-corruption activities”



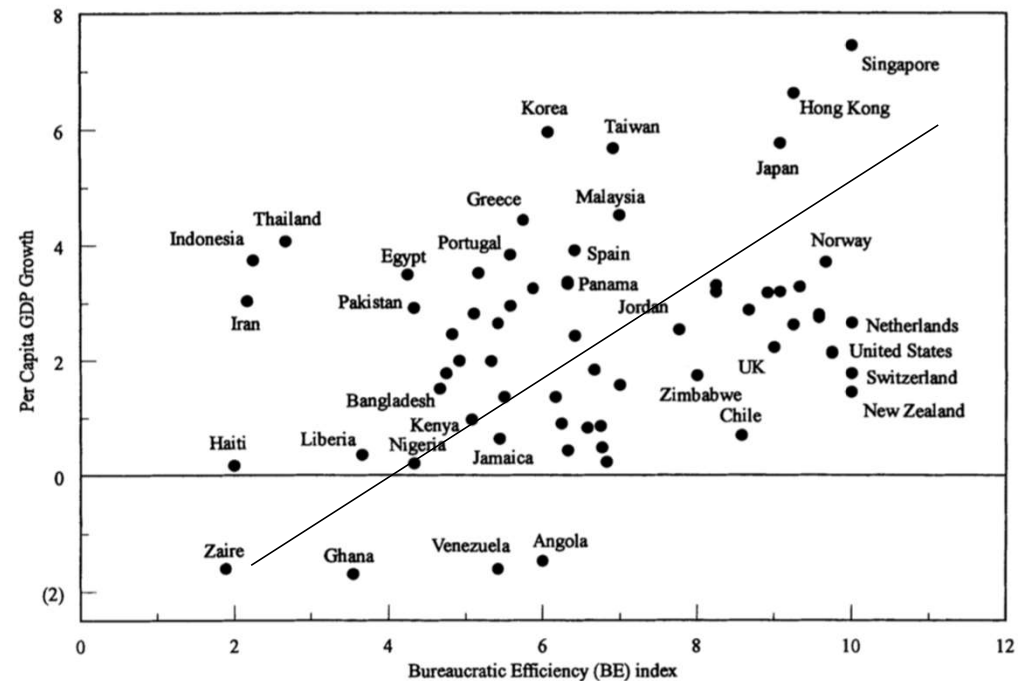
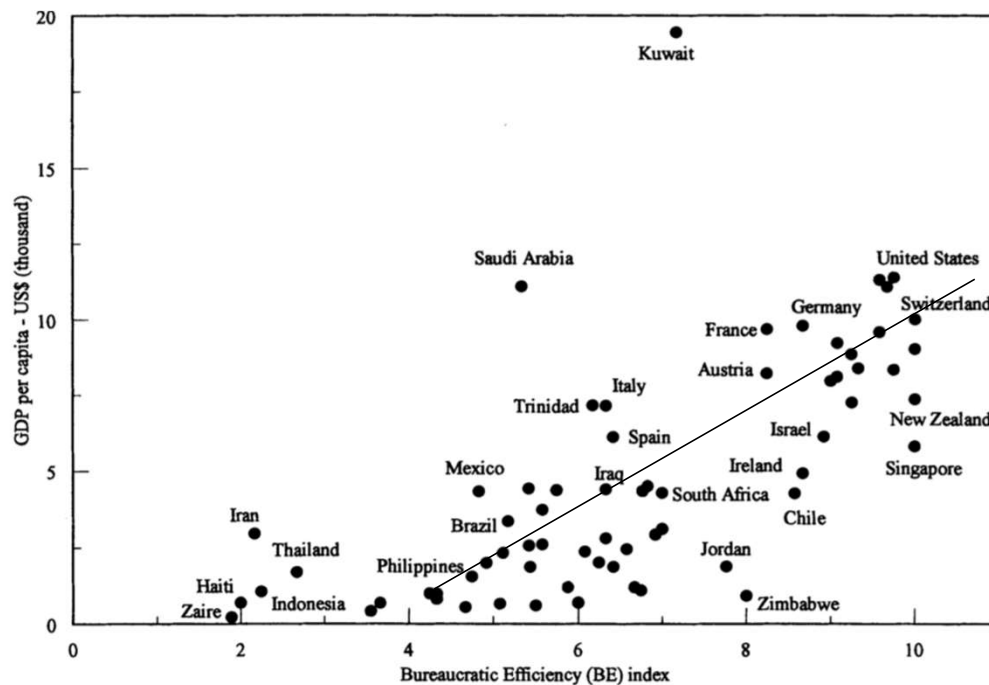
Corruption is common around the world



Transparency International CPI 2022

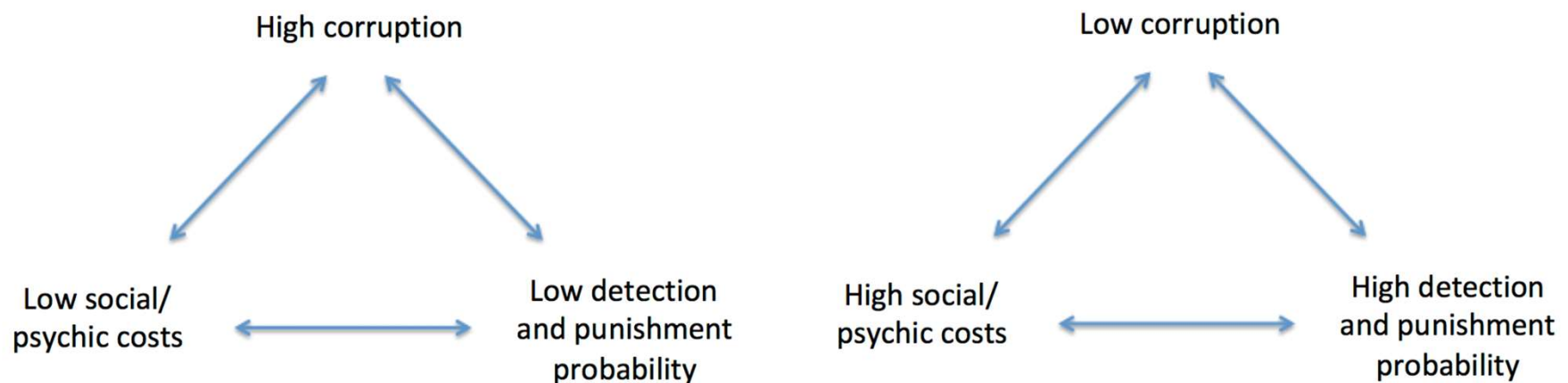
The view from Macro

- Corruption is correlated low incomes and slower growth



But paying a bribe is ‘rational’ in many settings: the “cost of doing business”

- **Multiple equilibria:** “Corruption happens as a result of interactions among individuals in which, given the choices others make, no one person can make herself better off by choosing any other course of action.” (Fisman and Golden, 2017, p. 4)



- NB: we are not *moralizing* corruption; our aim is to assess the net welfare consequences of paying versus not paying a bribe

The predicament

“The situation in one of the states where ZHL operates is getting critical. Unless the government pays what it owes us immediately, we will not be able to make payroll. We won’t be able to service our new ambulances, which will open us up to more accusations of negligence. Lives may be lost. This will devastate our morale and ruin our reputation. That bureaucrat will never stop.”

- Sanjay Rafati

Discussion: to pay or not?

- Pay bribe of 5%, or take out loan at 15% interest
- **According to utilitarianism, should Ziqitza pay or not?**
- Debate in groups
 - On my right: make the case for paying
 - On my left: make the case for not paying

The utilitarian decision framework

1. Specify the range of possible actions
 - **Here:** pay or not pay
2. Identify all who are affected by the action(s)
3. Assess the aggregate consequences of the actions on welfare
4. Two versions:
 - If **act-utilitarian**: Choose the action with the greatest welfare gain
 - If **rule-utilitarian**: Choose the action that - if taken by everyone when following a rule - yields the greatest welfare gain

Analogy archetype – The Ransom Game

- *Your relative gets kidnapped. You get asked for a ransom payment. You can afford it. It's the only way to save your relative.*
- Is it wrong to pay?
 - A: Yes
 - B: No

Everyone else is doing it...

Collective traps and rational irrationality

- If no one paid kidnappers, there would be no kidnapping – and yet, it is individually rational to pay
- Act utilitarianism is more contingent
 - Is actor “big” to have sufficiently large negative spillovers?
- Rule utilitarianism eliminates multiple equilibria
 - Calls out moral disengagement (diffusion of responsibility)

Welfare if small (no spillovers)

Actor	Utility from paying ransom
Family	+
Kidnapped person	+
Kidnappers	+
Rest of society	0

AU verdict: pay

Welfare if big (spillovers)

Actor	Utility from paying ransom
Family	+
Kidnapped person	+
Kidnappers	+
Rest of society	Σ^-

AU verdict: don't pay

The value scorecard

Make payment under assumption:	Act utilitarian	Rule utilitarian	Other principles we will study...
Ziqitza big , has spillover effects	not ok	not ok	?
Ziqitza small , no spillover effects	ok	not ok	?

Facilitation payment vs bribes for contract

- Bribe in case of Ziqitza was a **facilitation payment** - resolves the bureaucrat hold up
- **Consider another bribe:** Suppose municipal health authority opens tender for ambulance service
 - An ambulance provider decides to bribe health minister to get the contract, promising kickback once contract is awarded
- Is a **bribe for contract** different from a **facilitation payment**?

General conclusions

1. Human welfare as one of the cardinal values
 - Different utilitarian takes: Act and Rule
2. Moral rules are algorithms that can be used anywhere
 - Utilitarianism is one of them
3. Given different data/assumptions may get different answers
 - Can be both principled (in values) *and* flexible (in action)
 - **Value score card** helpful to assess the robustness of decisions – does it pass one ethical framework or many?

Case-specific conclusions

- Not all bribe payments have the same moral implications
- Tension for **Ziqitza**: immediate lives saved vs. incentivizing a high corruption environment
 - Act U: depends on **size** of negative spill-overs – whether it is setting a precedent
 - Rule U: “as if precedent-setting” – much weaker case for giving in to bribery request – guards against moral disengagement
- **Sweta Mangal** and **Ziqitza’s** “no bribery” stance resonates with being an agent of positive change (Act U-big), or Rule U

Coming up...

- Recommendation: tackle **Dragonfly case**
 - Look for parallels
 - Practice your Utilitarianism
- **Next class**
 - More frameworks – duty ethics
 - Cases on workplace discrimination and sexual harassment