Lecture 1 – Do Markets Shape Our Morals?

Cowen, T., & Tabarrok, A. (2015). *Modern principles of economics*, Chapter 24. Macmillan International Higher Education.

Markets work best when traders know exactly what is being traded. When this is the case, markets will attract both buyers and sellers because each side expects trade to be mutually profitable. **Asymmetric information** is defined as when one party in an exchange has more or better information than the other party.

The **principal-agent problem** occurs when one party has different intentions than the other. How can a principal (employer) incentivize an agent (employee) to work in the principal's interest even when the agent has information that the principal does not?

Moral hazard is when an agent tries to exploit an information advantage in a dishonest or undesirable way.

There are two solutions to asymmetric information and moral hazard problems:

1. Provide more information, thereby reducing the asymmetry.
2. Reduce the incentive for the knowledgeable party to exploit their information advantage.

Overcoming Moral Hazard by Providing More Information

The internet has played a big role in helping overcome asymmetric information by making it easier for buyers to pool their knowledge. The internet has increased the value of having a good **reputation**. **Reviews** have two advantages. They make it easier to avoid deceitful sellers and they raise the cost to sellers of exploiting their information advantage. Nevertheless, consumer reviews can be easily faked to make one's own store look better, or a rival store look inferior. The general problem is that the more important a rating becomes, the greater the incentive to fake that rating or manipulate it. Rating companies, such as *Consumer Reports*, might be useful to overcome this problem, because they test and rate a wide variety of products themselves and do not accept advertisements or payments from the companies whose products are being rated. Rating companies would have better incentives if they were paid directly by buyers, but again, the buyers will not pay because of **free rider** problems (an individual who consumes but does not pay). In general, information problems are difficult to solve completely because dumping down one problem often inflates another.

Overcoming Moral Hazard by Creating Better Incentives

A second solution to asymmetric information is to give sellers less of an incentive to exploit their information advantage, or to align the buyer's and seller's incentives. For instance, lawyers are incentivized to work hard on their case with a contingent fee – the lawyer is paid, typically a third of any judgment – only if the case is won. Moral hazard is everywhere, as self-interest is everywhere.

Adverse selection occurs when an offer conveys negative information about the product being offered. E.g., sick people are most likely to want health insurance. The offer (health insurance) conveys the negative information that the buyer is likely to have above-average health costs. A market failure may arise when buyers and sellers have asymmetric information about the quality of the product traded. An example is the market for used cars. Even if there are high- and low-quality used cars, all sell for an average price, because buyers have little or wrong information about the quality of the product.

Lecture 1 – Do Markets Shape Our Morals?

Cowen, T., & Tabarrok, A. (2015). *Modern principles of economics*, Chapter 24. Macmillan International Higher Education.

Markets work best when traders know exactly what is being traded. When this is the case, markets will attract both buyers and sellers because each side expects trade to be mutually profitable. **Asymmetric information** is defined as when one party in an exchange has more or better information than the other party.

The **principal-agent problem** occurs when one party has different intentions than the other. How can a principal (employer) incentivize an agent (employee) to work in the principal's interest even when the agent has information that the principal does not?

Moral hazard is when an agent tries to exploit an information advantage in a dishonest or undesirable way.
There are two solutions to asymmetric information and moral hazard problems:

1. Provide more information, thereby reducing the asymmetry.
2. Reduce the incentive for the knowledgeable party to exploit their information advantage.

Overcoming Moral Hazard by Providing More Information
The internet has played a big role in helping overcome asymmetric information by making it easier for buyers to pool their knowledge. The internet has increased the value of having a good **reputation**. **Reviews** have two advantages. They make it easier to avoid deceitful sellers and they raise the cost to sellers of exploiting their information advantage. Nevertheless, consumer reviews can be easily faked to make one's own store look better, or a rival store look inferior. The general problem is that the more important a rating becomes, the greater the incentive to fake that rating or manipulate it. Rating companies, such as *Consumer Reports*, might be useful to overcome this problem, because they test and rate a wide variety of products themselves and do not accept advertisements or payments from the companies whose products are being rated. Rating companies would have better incentives if they were paid directly by buyers, but again, the buyers will not pay because of **free rider** problems (an individual who consumes but does not pay). In general, information problems are difficult to solve completely because dumping down one problem often inflates another.

Overcoming Moral Hazard by Creating Better Incentives
A second solution to asymmetric information is to give sellers less of an incentive to exploit their information advantage, or to align the buyer's and seller's incentives. For instance, lawyers are incentivized to work hard on their case with a contingent fee – the lawyer is paid, typically a third of any judgment – only if the case is won. Moral hazard is everywhere, as self-interest is everywhere.

Adverse selection occurs when an offer conveys negative information about the product being offered. E.g., sick people are most likely to want health insurance. The offer (health insurance) conveys the negative information that the buyer is likely to have above-average health costs. A market failure may arise when buyers and sellers have asymmetric information about the quality of the product traded. An example is the market for used cars. Even if there are high- and low-quality used cars, all sell for an average price, because buyers have little or wrong information about the quality of the product.

The outcome is that high-quality cars owners will be reluctant to sell because of the low market price, so high-quality cars will get crowded out from the market until only low-quality cars will be up for sale. There are ways, however, to overcome adverse selection. In the case of used cars, warranties, expert checks, and reports of the car's repair history are all ways to increase trade and avoid adverse selection.

Adverse Selection in Health Insurance

Health insurance might startle an "adverse selection death spiral". The sick are more likely to want health insurance than the healthy. However, if too many sick people versus healthy people buy insurances, then the company has to raise its rates. In turn, the higher rates make health insurance even more undesirable for healthy individuals, thus pushing them out of the market. Some market features help limit this adverse selection problem:

- **Inspections** or checkups. An insurance company can adjust rates based on periodic health checkups of its customers.
- **Group plans**. An insurance company can emphasize sales to groups, such as through someone's workplace, to increase the chances of signing up both healthy and unhealthy customers.
- **Conscientiousness**. Some individuals value health insurance more than others, independently from their health status. Insurance companies will target individuals who value and are careful of their health. This will lead not to adverse selection, but to positive selection which moderates the adverse selection death spiral.
- The **Affordable Care Act (ACA)** subsidizes the purchase of health insurance for the poorer people, which leads to avoidance of the adverse selection death spiral.

Another way to avoid the adverse selection spiral is to compel all citizens to purchase health care insurance by force of law. On the one hand, this reduces the likelihood of having a majority of sick buyers. On the other hand, however, it creates a moral hazard problem: When people are insured, they tend to behave less prudently.

Signaling as a Response to Asymmetric Information

A **signal** is an expensive action that is taken to reveal information. E.g., Hyundai in 1998 wanted to start producing more luxurious cars but they knew consumers would not buy if Hyundai could not shake its low-quality reputation. Therefore, they introduced "America's Best Warranty", a very favorable insurance for buyers that worked as a signal of Hyundai's new quality.

Signaling in the Job Market

If you bid for a lower yearly salary in order to get the job, the employer might get the feeling that you are willing to be paid less because you are less qualified for the job. This signals the wrong intention, as you just want that job very badly. How can you signal that you are, in fact, of high quality? Degrees might fix this problem. A degree signals positivity: quality, knowledge, hard work. Still, there are some inefficiencies involved, as getting a degree then essentially comes down to earning a higher wage, rather than actually learning something.

Bowles, S. (2008). Policies designed for self-interested citizens may undermine "the moral sentiments": Evidence from economic experiments. *Science, 320*, 1605-1609.

This article stresses the idea that incentives that appeal to self-interest may fail when they undermine the moral values that lead people to act altruistically or out of ethical drive. The author presents some behavioral experiments that support his theory.

Introduction

David Hume, philosopher, and economist claimed that constitutions and public policies should be designed for "knaves" motivated only by their "private interest". He believed that people were selfish, making choices out of egoism without knowing what is actually good for them. In this case, policies designed to align self-interest with the social interest may be counterproductive. E.g., One behavioral study focused on a daycare center that introduced a fine when children were picked up late. The idea was that many parents do not want a fine, so they would pick up their children earlier. Instead, the number of late pick-ups increased dramatically. The ethical commitment disappeared, and parents saw the pick-up as something that was up for sale.

Incentives and Market Failures

Market failures arise when individuals do not take into account the effects of their actions on others (**external effects** or **spillovers**). **Complete contracts** eliminate the spillovers, internalizing the external effects of an action, so that self-interested individuals implement outcomes that are socially efficient. However, contracts are rarely complete, mainly because of asymmetric information and the impossibility of writing an enforceable contract (e.g., a contract between labor markets and credit markets that specifies that the employee will work hard and well).

The mechanism designer's task is to find a way to assign to each actor the entire benefits and costs (to themselves and to others) of his actions, thereby providing a surrogate for complete contracts. Behavioral economics has recently revealed that the joint effect of intrinsic motivation (**ethical motives**) and extrinsic motivation (**explicit incentive**) is not simply the sum of their effects considered separately (**non-separability principle**). The **assumption of separability** shows that policies pointed towards self-interest do not affect ethical and other social preferences. The effects that moral sentiments and material interests have on our behavior are not interactive, but additive. In most cases, incentives undermine ethical motives. Four reasons have been suggested for the failure of the separability assumption.

1. **Framing**. Incentives may frame a decision problem and thereby suggest self-interest as the appropriate behavior.
2. **Endogenous preferences**. Incentives may affect the long-term development of preferences.
3. **Overdetermination**. Incentives may compromise the individual's sense of autonomy.
4. **Information content of incentives**. Incentives may convey information affecting behavior.

Experiments illustrating these four explanations follow below.

Framing

The first experiment shows how the frame-shifting effects of incentives may occur in cases of government-imposed incentives. The subjects of the experiment were asked to choose how much to withdraw from a mutually beneficial common pool, namely "the forest".

The experiment revealed that in the absence of "governmental regulations", the subjects acted partially ethically and withdrew amounts that were about midway between the self-interest and the group-beneficial levels.

However, when a "fine" was introduced to reduce the amount of resource withdrawal, subjects started acting almost entirely out of self-interest. The government regulation crowded out ethical motives and induced individuals to maximize their personal payoff.

Endogenous Preferences

Preferences are endogenous if one's experiences result in durable changes in motivations, and therefore a change in behavior in given situations. Incentives may change preferences because they affect key aspects of how we acquire our motivations. An experiment found that in anonymous experimental settings, individuals from the more market-integrated societies were also more fair-minded, in that they made more generous offers to their experimental partners and more often chose to receive nothing rather than accept an unfair offer. This kind of fair-mindedness is essential to the exchange process in market-oriented societies, where individuals engaging in mutually beneficial exchanges with strangers represent models of successful behavior who are then copied by others.

Overdetermination

Incentives may reduce the individual's sense of autonomy, startle control aversion, and induce pessimism. An experiment showed how incentives may tackle self-determination. Experimental agents (employees) had to choose a level of production. Before the agent's decision, the principal (the employer) could choose whether to leave the choice of the level of production completely to the agent's discretion or impose a lower bound on the agent's production. Because of control aversion and desire for self-determination, agents chose a lower level of production when the principal imposed the bound versus when they did not. The limit showed distrust in the agents.

Incentives Convey Information

Incentives reveal information about the principal's preferences, the nature of the task, and his beliefs concerning the agent, and can crowd out ethical behavior. An experiment showed that the use of a fine reduced "back-transfer" by the "trustee" to the "investor". A strong intrinsic positive response was instead recorded in the case where "punishments" were not used.

Why are Counterproductive Incentives Common?

Counterproductive incentives are common because even if they reduce the total gains associated with a project, their use may give the principal a sufficiently larger slice of the smaller pie to motivate the principal to use them. Small differences in institutional design could have large impacts on the outcomes.

Why Moral Sentiments and Material Interests Are Not Separable

Behavior is both acquisitive and constitutive. Individuals act not only to obtain something, but also to be or become a good person. The reasons for non-separability may reflect the fact that appeals to self-interest and to morality activate different cognitive and neurological processes, with appeals to the former competing with the latter by a process of cognitive overload.

Discussion
Good policies and constitutions are those that support socially valued ends not only by harnessing selfish preferences to public ends but also by evoking, cultivating, and empowering public-spirited motives.

Falk, A., & Szech, N. (2013). Morals and markets. *Science*, 340(6133), 707-711.

This article presents controlled experimental evidence on how market interaction changes how individuals value harm and damage done to third parties.

The Mouse Paradigm
In the experiment, subjects had to decide between either saving the life of a mouse (and not receiving money) or receiving money (and killing the mouse). Three different conditions were analyzed:

1. Individual treatment (one individual makes the decision).
2. Bilateral trading market (a buyer and a seller bargain to set prices).
3. Multilateral trading market (multiple buyers and sellers trade with each other).

In the individual treatment, subjects faced a simple binary choice. The subject could decide whether to save the life of the mouse and receive nothing, or to kill the mouse and get 10 euros. In the bilateral double auction market, one seller and one buyer bargained over killing a mouse for a total gain of 20 euros that the two parties could split up. The multilateral double auction market treatment was exactly like the bilateral market treatment, except that in this condition there were several buyers and sellers. The hypothesis of the authors was that markets would display a tendency to erode moral standards, because of three essential features of market interaction:

1. In markets, it takes two people who agree on trading to complete a trade, implying that **responsibility** and **feelings of guilt** may be shared and thus diminished.
2. Market interaction reveals **social information** about prevailing norms. Observing others trading may induce individuals to ignore moral standards and pursue self-interest instead, by engaging in trade.
3. Markets provide a **strong framing** and focus on materialistic aspects such as bargaining, negotiation, and competition, and may divert attention from possible adverse consequences and moral implications of trading.

The experiment revealed that in the individual decision treatment, 45.9% of subjects were willing to kill their mouse for 10 euros. In contrast, 72.2% of sellers in the bilateral market were willing to trade for prices below or equal to 10 euros. Finally, in the multilateral market, the willingness to kill for less than or equal to 10 euros was 75.9%. Subsequently, a price-list treatment was used to show that subjects in the individual condition required way greater amounts of money in order to kill the mouse compared to in the bilateral and multilateral markets.

Moral Versus Morally Neutral Values
The hypothesis of this final step is that for moral values the decay is more pronounced than for private consumption values, where trading involves opportunity costs of consumption rather than costs to third parties.

This assumption was tested through an experiment analogous to the mouse paradigm. Subjects were endowed with a coupon. In case they accepted a monetary amount (in the price-list condition) or decided to trade (in the market condition), they had to return their coupon, which was then invalidated. For coupons, the effect of markets was revealed to be much smaller and insignificant than that for mice. In conclusion, when identical procedures, parameters, and market institutions are used, moral values decline significantly more than values that are morally neutral.

Robustness and Discussion

The experiment's finding can raise three main concerns:

1. One could argue that the main treatment effect partly depended on the fact that total surplus was bigger in markets than in individual condition (20 versus 10 euros).
2. Subjects may have perceived killing the mouse as a side-effect of the act of trading in the market treatments, whereas in the individual treatment subjects may have perceived killing the mouse as a direct means to earn money.
3. The authors decided to use the minimum trading price as their main dependent variable to assess a seller's willingness to kill a mouse in markets because they most likely represent an upper bound for the sellers' reservation values.

Additional experiments were run to confute the first two points. In conclusion, the authors agree that we as a society have to think about where markets are appropriate – and where they are not.

Wight, Jonathan B. (2014). Economics within a pluralist ethical tradition. *Review of Social Economy*, 72:4, 417-435

Introduction

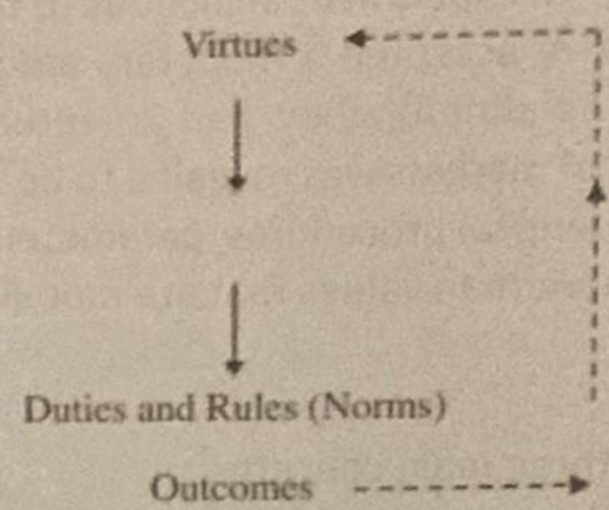
Ethical pluralism is the recognition that multiple ethical frameworks operate in social settings to solve problems of moral hazard. The thesis of this article is that economic actors often approach positive and normative questions from a pluralist ethical position. In fact, norms are often driven by sentiments or feelings in addition to rational calculations (e.g., duties and virtues). The key point is that a single ethical approach is inadequate for resolving coordination and cooperation problems; ethical pluralism allows for synergies, complementarities, and necessary back-ups.

Pluralism Defined

Pluralism in ethics is the degree of acceptance of two or more mutually inconsistent theoretical frameworks, which pertain to the same or overlapping domains of reality (Fullbrook, 2012).

For the purpose of this article, pluralism has two meanings:

1. The divergence of theories across ethical domains (**vertical pluralism**). It focuses mainly on a person's character and virtue, the action itself and the evaluation of outcomes. E.g., when you find $50 while being at work. Do you keep it, or do you give it to the company you work for, considering you found the money whilst being at work?
2. The divergence of theories within ethical domains (**horizontal pluralism**). It focuses on the different ideas within the same kind of ethics, such as virtue ethics (honesty).



Note. From Wight, Jonathan B. (2014). Economics within a pluralist ethical tradition. *Review of Social Economy*, 72:4, 417-435

The figure above illustrates vertical ethical pluralism. There are three main ethical frameworks, which focus on different aspects of human behavior and choice:

1. **Virtue ethics** relate to the person's character and virtue.
2. **Duty ethics** relate to the action itself and explore the rules and duties for action that could be derived from God or rational thought.
3. **Outcome ethics** (utilitarian) relate to the evaluation of outcomes.

As the figure above shows, a virtuous economic agent will have the self-control to obey rules and duties, and thereby obtain desirable outcomes. Nevertheless, such positive outcomes are not the motive for virtuous behavior and the calculation of outcomes is not the adopted method of ethical discernment. The upward dotted arrow posits a positive feedback loop between outcomes and virtues, which suggests that being part of a well-developed market makes people more inclined toward positive, other-regarding behaviors. Thus, trade can be defended on pluralist grounds of both good outcomes and promotion of virtue ethics. In fact, moral sentiments about fairness and trust – rather than simply a calculation of personal gain or loss – undergird many exchange relationships. Vertical pluralism offers researchers the insight that multiple ethical frameworks are at work in market processes and public policies.

Virtue ethics: Adam Smith ↔ Aristotle ↔ Confucius ↔ Buddha

Duty ethics: Divine Command ↔ Kantian ↔ Natural Rights

Outcome ethics: Utilitarian ↔ Neoclassical Economic ↔ Saving Lives ↔ National Security ↔ Freedom ↔ Fairness ↔ Dynamic Efficiency

Note. From Wight, Jonathan B. (2014). Economics within a pluralist ethical tradition. *Review of Social Economy*, 72:4, 417-435

The figure on this page illustrates horizontal pluralism. Within each different ethical framework there are multiple possible modes of analysis. This gives rise to horizontal pluralism. For example, honesty could be explained in virtue ethics from the perspective of Adam Smith, Aristotle, Confucius, Buddha, or others.

AthenaSummary

Efficiency in a Pluralist System

Economists often rely on pluralist ethics. The author agrees that critical thinking requires that economists in each case should explore alternative goals (horizontal pluralism) and ultimately make a qualitative ethical judgment about the value of efficiency compared with other worthy goals. Standard welfare economics does not make interpersonal utility comparisons.

Two examples of **goal-switching** that rely on comparison of utilities are:

- **Static efficiency**, which satisfies the preferences of the existing consumers.
- **Dynamic efficiency**, which promotes the preferences of future consumers.

Switching goals involves interpersonal comparisons of the gains and losses from such a switch. When is it permissible to switch social goals, and on what ethical basis should this be done? Economic policies generally produce multiple consequences that are amenable to trade-offs. Complex choice situations rarely entail a solution maximization, and more often require choosing the **least-worst** option, after considering the limitations imposed by rules, duties, and virtues.

Pluralism in Business Ethics

In a competitive market, maximizing profits produces desirable outcomes for consumers by producing the kinds and qualities of products desired at the lowest possible prices. Thus, one could argue that the only moral obligation of a firm is to maximize profits (Friedman, 1970). Nevertheless, the principal-agent problem tells us that managers have different interests from the owners of the corporation. Therefore, as Friedman points out, corporations need agents who abide by deontological ethics, following their **duty**. To produce individuals who obey their duties requires the cultivation of **virtue**. Finally, cultivation of virtue builds upon **feelings** to develop self-control. Vertical pluralism unites virtues, duties, and outcomes as a holistic account of the complexity of human organization and cooperation.

Ethical Pluralism in Science

The science market provides considerable incentives (money and status) to fudge, falsify, or even plagiarize outputs. In the standard economic view, people calculate their self-interest and tell the truth when the benefits of doing so exceed the costs. When this holds, researchers are honest because of incentives that align self-interest with social interests. However, the author believes that this narrow foundation on incentives only is not able to sustain the health of science in the long run, because it neglects two other powerful foundations, namely duty and virtue. Moral commitments should thus be explicitly studied and promoted as part of a pluralist understanding of what makes economic interaction work.

Conclusion

The author points out that some economic actors do rely on ethical pluralism for reducing transaction costs in exchange and for addressing problems of asymmetric information and moral hazard. Consequences, duties, and virtues all play a role in promoting the search for truth within the economic research community.

Lecture 2 – Do Markets Have Moral Limits?

Satz, D. (2010). *Why some things should not be for sale,* Chapter 4: Noxious markets. *The moral limits of markets*. Oxford University Press.

Abstract Markets Versus Noxious Markets

What is wrong about the nature of some markets? How should our social policies respond to undesirable trades? When and why should we seek to regulate or block a market?

The author begins by stating two basic assumptions of her account:

- First, she aims to discuss some specific markets that she classifies as "**noxious**" – immoral and toxic to individuals and/or society as a whole. E.g., trade in human body parts, child labor, toxic waste, sex, and life-saving medicines.
- Second, she is not taking in consideration cases of extreme scarcity, where "tragic choices" are to be made. The author argues that the problem of noxious markets is related to the standing of the parties involved in the trade. In short, noxious markets undermine the conditions for equality within society.

The approach to markets she defends recognizes market heterogeneity and considers other values besides efficiency and distributional equality narrowly conceived.

Noxious Markets: The Basic Parameters

According to Satz, there are four parameters that can be used to judge the nature of the "noxiousness" of a particular market.

Two parameters which characterize the consequences of particular markets:

1. Some markets produce extremely **harmful outcomes** for individuals. Such outcomes leave our set of basic interests unsatisfied and can be classified as *welfare harms* – a significant setback to a person's overall good – or *agency harms* – a harm to someone's ability to participate in deciding matters that bear on their good.
2. Some markets can be extremely **harmful for society**. Such markets jeopardize equality within society by undermining the conditions needed for people to interact as equals, or worse: they threaten equal democratic citizenship.

Two parameters that characterize the sources of particular markets:

3. Some markets are characterized by very **weak or highly asymmetric knowledge and agency** on the part of market participants. Information failures – poor information or highly asymmetric information – are relevant to judge the outcomes of a particular market. For instance, if one or both of the parties to a contract are mistaken about the information or the future consequences of their contract, we cannot assume that the exchange is a Pareto improvement (a Pareto improvement occurs when a change in allocation harms no one and helps at least one person, given an initial allocation of goods for a set of people). Within this category fall markets that involve young children, markets that fail to present alternative opinions about a political matter, and markets based on fraud and deception of the parties involved, even when there are no harmful consequences. Agency issues also occur when one of the affected parties does not directly trade but depends on other parties to transact for them (e.g., child labor).
4. Some markets reflect the underlying extreme **vulnerabilities** of one of the transacting parties. Rousseau wrote that no citizen should "be wealthy enough to buy another, and none poor enough to be forced to sell himself." Satz argues that markets can either reflect or exacerbate such vulnerabilities. Reflect means that some markets exploit the poor's misery, while exacerbate means that some markets make the poor more vulnerable.

TABLE 1. What Makes a Market Noxious?

Source: Weak Agency	*Source: Vulnerability*
Inadequate information about the nature of and/or consequences of a market; others enter the market on one's behalf	Markets in a desperately needed good with limited suppliers; markets with origins in poverty and destitution; markets whose participants have very unequal needs for goods being exchanged
Outcome: Extreme Harms for Individual	*Outcome: Extreme Harms for Society*
Produces destitution; produces harm to the basic welfare and/or agency interests of the individual	Promotes servility and dependence; undermines democratic governance; undermines other regarding motivations

Note. From '*Why some things should not be for sale,* Chapter 4: Noxious markets, p. 91-105. *The moral limits of markets*. Oxford University Press.

Equal Status in a Democracy

An equal status implies that all individuals have an equal moral worth. For two individuals to have equal status they need to acknowledge each other's existence and understand that each of them has independent claims they can press without needing the other's permission to do so. This requires that each have rights and liberties of certain kinds as well as very specific resources, such as a level of education. According to the British social theorist T. H. Marshall, the status of equal citizenship requires that all have:

1. Equal basic **political rights and freedoms**, including rights to speech and participation in the political process.
2. Equal **rights and freedoms within civil society**, including rights to own property.
3. Equal **rights to a threshold of economic welfare** and to "share to the full in the social heritage and to live the life of a civilized being according to the standards prevailing in the society".

Markets are usually able to supply these basic rights. However, these goods should be intrinsic in a society, and not dependent on individual preferences or income. If we want to prevent inequality from arising in a society, then we should seek to enforce or guarantee a certain level of goods, such as education or health care, even if some citizens would rather trade these goods. In some cases, markets can promote equal citizenship. However, some other markets ought to be blocked altogether since they undermine the conditions for equality. In these cases, the transfer of wealth (money) will not always be sufficient to maintain the conditions for citizen's equality. Giving someone wealth does not make them a functioning part of society. Distribution of money throughout a society does not make for equal citizenship, whereas supply of education and healthcare does.

Regulating Markets, Blocking Markets

In this section, Satz discusses how to empirically apply the theoretical analysis she presented above to answer the question: How should we decide what approach to take to a noxious market? Obviously which policies it makes sense to adopt depends on the source of the market's noxiousness. By using the four parameters, we can determine which factors contribute to the noxiousness of the market, and decide accordingly (e.g., weak agency, underlying vulnerability, threat to equality). For instance, if the problem behind a market is asymmetric information – a weak agency issue – then regulation of the market might be the best solution.

Differently, if the market entails harmful consequences that cannot be avoided – such as child labor – then we should seek to block the market.

Cases: The Titanic and Toxic Waste

The author illustrates her theory on noxious markets by applying it to the Titanic case and toxic waste. Beginning with the Titanic case, she argues that the ticket market was noxious because of weak agency and harmful consequences. People were allowed to buy tickets either with or without lifeboat access. There was weak agency due to misinformation (the ship's claimed unsinkability), and obviously extremely harmful consequences because people who did not have access to a lifeboat drowned. Therefore, the distribution of safety according to ticket price was a noxious market. Some could argue that it is paternalistic to enforce the purchase of lifeboat access. However, as a commitment to equal citizenship, there are some rights that should not be able to be contracted away. In conclusion, people who can and cannot afford lifeboats, should not be traveling on the same boat, meaning that we are to ensure the same safety to everyone within the same community. If our basic rights were tradeable, some people would be subject to servitude and subordination. There are two basic reasons why democratic societies might want to provide all individuals with some fundamental rights:

1. Inequalities in some goods, such as education or political influence, go against the idea that we are each other's equals.
2. Markets enable people to opt out of relationships with particular producers and to take up new relationships so as to find new ways of satisfying their preferences. However, this "exit" option might not be optimal for the common interest (e.g., voting).

For instance, when agents can buy private solutions for themselves in education, healthcare, or other goods, this may have problematic outcomes for the public provision of these goods. Thus, the best way to ensure that the public's interests are taken into account is to align self-interest with the social interest, by giving both the public and the decision maker the same interests. Then, Satz argues that the toxic waste market was noxious because of weak agency, vulnerability, and harmful consequences for individuals. Purely monetarily speaking, trade in toxic waste would benefit both poor and wealthy countries, by making them both better off. However, the market was noxious for three reasons. The first two are concerned with the sources, while the last one regards the consequences of such market:

1. There was vulnerability in the trading position between rich and poor countries. The less developed countries were exploited; had they not been so poor, they would not have consented to the transfer of toxic waste to their lands or would have required better terms.
2. There was weak agency, as the corrupt governments did not act for the well-being of the individuals.
3. There were extremely harmful outcomes for the individuals. The shipping and storing of toxic waste leads people to suffer or even die, and by some forms even future generations will be affected. Also, because these countries are poor, they have less capacity to monitor and regulate the pollution, which may lead to even more pollution.

The Limits of My Approach

Finally, Satz discusses some limits of her approach:

1. We cannot immediately conclude from the fact that a market is noxious that we legally ought to ban it. Banning a market may indeed have worse consequences.
2. People may have different views of the appropriate trade-offs between the different parameters Satz appealed to, as well as between these parameters and other values.
3. In her theory, she has not settled on exactly how to operationalize values, namely she did not give a quantitative way to assess noxious markets.

4. Her interpretation is sensitive to changing circumstances, so that markets that are currently noxious may emerge under other conditions as perfectly acceptable (or the reverse).
5. It must be admitted that other accounts of noxious markets are possible.

Nevertheless, the author concludes by saying that her theory provides a useful starting point for thinking about moral matters involved in markets.

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., & Rahwan, I. (2018). The moral machine experiment. *Nature, 563*(7729), 59-64

Abstract

The article is based on concern about developments of AI and societal expectations about its moral decisions. **The Moral Machine experiment** aimed to address the challenge of determining societal expectations for the ethical principles that should guide machine behavior, particularly in autonomous vehicles.

Introduction

Today machines we are using are not only used to promote wellbeing but also distribute harm that cannot be eliminated. Unavoidable harm scenario is a car crash scenario. Here autonomous (driverless) vehicles would need to decide between different stakeholders on the road. Regardless, of how rare these cases are, ethical principles should be established.

The challenge arises and cannot be solved by one party only. First challenge is high dimensionality of the problem: thousands of combinations of preferences must be analyzed which challenges conventional data collection methods. Secondly, to make universal ethical principles for AI we need to understand local preferences, which again requires substantial sample size.

Hence, in response, Awad et al. (2018) designed the **Moral Machine experiment** – multilingual online 'serious game' for collecting large-scale data on how people want autonomous vehicles to solve moral dilemmas in the context of unavoidable accidents.

Experiment focuses on **9 factors**: sparing humans (versus pets), staying on course (versus swerving), sparing passengers (versus pedestrians), sparing more lives (versus fewer lives), sparing men (versus women), sparing the young (versus the elderly), sparing pedestrians who cross legally (versus jaywalking), sparing the fit (versus less fit), and sparing those with higher social status (versus lower social status).

Global Preferences

The strongest preference globally observed **sparing humans over animals, sparing more lives, sparing younger lives.** Hence, these preferences can become fundamental building blocks for machine ethics. Although shown globally, these preferences still spark controversy, and hence to proceed further with AI ethics should be agreed upon among all stakeholders.

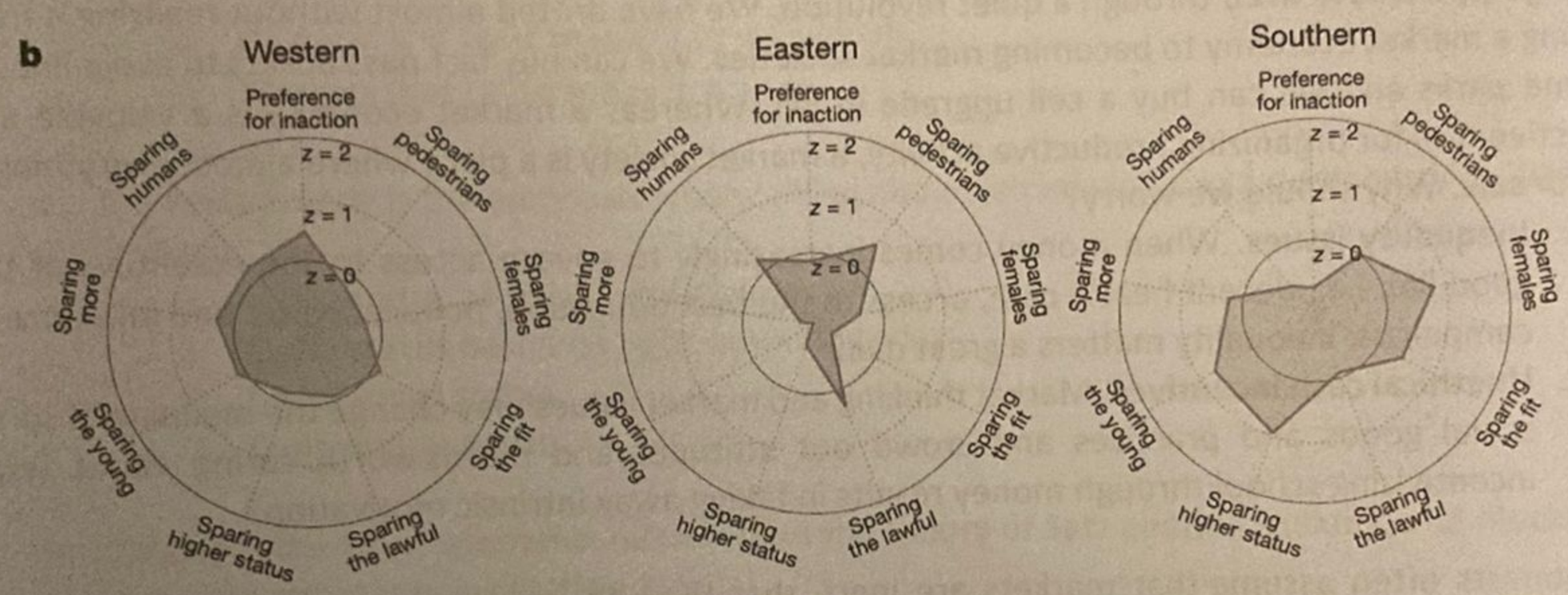
For example, German Ethical Rule (on Automated and Connected Driving) number 9 also states that any dis- tinction based on personal features, such as age, should be prohibited. This rule directly clashes with global preference to 'save' younger lives versus the elderly.

Individual variations

Individual variations have no sizable impact on any 9 variations, although theoretically may be important.

Cultural clusters

Geolocation allowed to locate clusters of countries with homogeneous vectors of moral preference, i.e. Western, Eastern, and Southern clusters. This clustering patterns suggest that groups of territories can find common preferences for machine ethics. However, this also means that between-cluster differences will impose challenges. Whereas the ethical preferences of the public should not necessarily be the primary arbiter of ethical policy, the people's willingness to buy autonomous vehicles and tolerate them on the roads will depend on the palatability of the ethical rules that are adopted.



Note. From '*The Moral Machine Experiment*' (2018)

Country-level predictors

Preferences revealed by the experiment are highly correlated to cultural and economic variation between countries.

Key 4 cultural and economic **predictors** of Moral machine preference:

1. Individualistic vs collectivistic cultures. Participants form individualistic countries prefer to spare the greater number of individuals. Hence, these differences may prove to be an obstacle.
2. Following vs violating law pedestrians. Should those who violate the law be protected the same? Higher prosperity and quality of laws and institutions correlate with greater preference against jaywalking.
3. Higher country-level economic inequality corresponds to how characters of different social status are treated (less equally)
4. Different treatment of male and female corresponds to country-level gender gap.

Discussion

We need to have a global conversation to express our preferences to the companies that will design moral algorithms, and to the policymakers that will regulate them.

The Moral Machine project was atypical in its objectives and ambitions: no research has previously attempted to measure moral preferences using a nine-dimensional experimental design in more than 200 countries.

Even the strongest preferences expressed through the Moral Machine showed substantial cultural variations, and our project builds on a long tradition of investigating cultural variations in ethical judgments, however, they might not be fatal.

Video: Michael Sandel: Why We Shouldn't Trust Markets with Our Civic Life

Michael Sandel discusses why we should not trust markets with our civic life. He asks the following question: What should be the role of money and markets in our societies? Over the past three decades, he argues, we have lived through a quiet revolution. We have drifted almost without realizing it from having a market economy to becoming **market societies**. We can buy fast pass tickets to avoid lines in theme parks and we can buy a cell upgrade in jail. Whereas a market economy is a valuable and effective tool for organizing productive activity, a market society is a place where almost everything is up for sale. Why should we worry?

1. **Inequality issues**. When money comes increasingly to govern access to the essentials of the good life, i.e., decent health care, access to the best education, political voice, and influence in campaigns, inequality matters a great deal.
2. **Unethical cash incentives**. Market thinking and market values may change the meaning of some social goods and practices and crowd out attitudes and norms worth caring about (e.g., incentivizing school through money results in taking away intrinsic motivation).

Economists often assume that markets are inert, that they do not touch or taint the goods they exchange. Market exchange, they assume, does not change the meaning or value of the goods being exchanged. This may be true enough if we are talking about material goods, but not for non-material goods and social practices. In fact, in these cases, markets may undermine or crowd out relevant non-market values. They change the goods. Once we see that markets change the character of goods, we have to debate amongst ourselves these bigger questions and decide how to value goods.

In the end, the question is actually about how we want to live together. Do we want a society where everything is up for sale, or are there certain moral and civic goods that markets do not honor, and money cannot buy?

Lecture 3 – A Fair Playing Field?

Cowen, T., & Tabarrok, A. (2015). *Modern principles of economics*, Chapter 21. Macmillan International Higher Education.

The main goal of this chapter is to raise questions about the moral limitations of economics. The chapter considers normative claims made by critics of economics and how they intersect with positive economics.

- **Positive economics** entails describing, explaining, or predicting economic events.
- **Normative economics** entails recommendations or arguments about what economic policy should be.

The Case for Exporting Pollution and Importing Kidneys

The case for exporting pollution and importing kidneys is based on a simple economic idea: trade makes people better off. One person values a kidney more than money, and the other values money more than their kidney, so trade makes both better off. Similarly, the rich value their health more than money, and the poor value money more than health, so rich countries exporting pollution to poorer countries makes both better off. Some of the objections to standard economic reasoning are the following:

1. Exploitation;
2. Meddlesome preferences;
3. Fair and equal treatment;
4. Cultural goods and paternalism;
5. Poverty, inequality, and the distribution of income;
6. Who counts? Should some count for more?

Exploitation

Is the seller of a kidney being exploited? It is difficult to see the line that separates exploitation from exchange. Why is selling a kidney different from any other source of remunerated job or exchange? One response is that for a poor person the money is exploitative because the circumstances of poor people give them little choice but to sell things they would rather keep. Consider which of the following cases is most exploitative:

- Case 1: Someone asks you to donate a kidney but offers you nothing in return.
- Case 2: Someone offers you $5,000 to donate a kidney.
- Case 3: Someone offers you $500,000 to donate a kidney.

Few people would consider case 1 to be the most exploitative. It follows that if case 2 is exploitative, then case 3 must be even more exploitative because it offers more money. However, generally sellers are exploited when they are offered too little. When this holds, case 2 is more exploitative than case 3, therefore case 1 is the most exploitative of all. Hence, determining which of the three cases is the most exploitative is a difficult task, which concerns controversial values and opinions.

Meddlesome Preferences

Many people have a gut feeling that trading kidneys for money is just wrong. **Meddlesome preferences** are preferences about what other people do, even when those other people do not interfere in any direct way with what you do. Meddlesome preferences may involve homosexuality, interracial dating, religious practices, or even meat eating. How much should these meddlesome preferences count when thinking about justice?

"Live and let live" might be the key to solve these controversies, however resolving meddlesome preferences with other important values such as freedom, religious liberty, and rights is often difficult.

Fair and Equal Treatment

The concept of fair and equal treatment may go against the value of trade and efficiency. When thinking about trade-offs, we need to be aware of the resulting tensions and subtleties. For instance, when NYC had to decide whether to equip all buses to be capable of accepting passengers in wheelchairs or buy each disabled person free taxi trips for life, a trade-off between efficiency and morality arose. Reimbursing each wheelchair user for their taxi trips would have been a lot cheaper, however the principle of "equal treatment" was considered to be more important than economic efficiency. Therefore, the city decided to invest in the improvement of the buses.

Cultural Goods and Paternalism

Should governments support some goods even when the public is not willing to pay for them? Most governments invest in subsidizing culture, education, and related "higher values". Supporters of this kind of policy assume that the government knows best what is beneficial for the citizens. Contrarily, the pragmatic criticism of this kind of policy is that it tends to be counterproductive and wasteful. The more philosophical criticism is that people should be able to decide how they want to spend their money (anti-paternalistic).

Poverty, Inequality, and the Distribution of Income

What is a just distribution of income? How much is owed to the poor? How much is owed to the rich? Questions like these are at the heart of many debates about foreign aid, trade, taxation, health care, and immigration. Three normative policy judgements have been particularly influential:

1. John Rawl's maximin principle.
2. Utilitarianism.
3. Robert Nozick's entitlement theory of justice.

John Rawl's Maximin Principle

Justice requires maximizing the benefits going to society's most disadvantaged group. The worst-off group is prioritized in this approach. Rawl's argument is that if no one knew what position they held in a society, then they would want a rule that maximized the position of the worst-off just in case they turned out to be the worst-off. In economic terms, Rawls believed that people were risk averse. The average income can be neglected.

Utilitarianism

The idea that the best society maximizes the sum of utility. Utilitarian advocates redistribute income up to the point where the marginal change in utility created from the redistribution is positive. Differently from Rawls, they try to maximize the total sum of utility, not the utility of the worst-off person. Therefore, the average income is much more important. The utilitarian is likely to suggest that some amount of money be redistributed from rich people toward poor people. However, this redistribution might be limited by incentives. In fact, taking money away from the rich might decrease their incentive to earn, while giving money to the poor might incentivize them to remain poor.

Robert Nozick's Entitlement Theory of Justice

The distribution of income in a society is considered just if property is justly acquired and voluntarily exchanged. In the words of Nozick, all capitalist acts between consenting adults should be allowed.

Nozick criticizes the view that equality of outcome is important. According to him, what matters is the justness of the process that leads to differences in wealth.

For instance, if a person gets rich by legally selling a good or performing a service, there is no reason why the difference of wealth between them and another person should be unfair.

Who Counts? Immigration

When a public policy like trade or immigration is evaluated, one can count the benefits and costs to all individuals equally, focus on the welfare of their citizens, or prioritize that of foreigners. Immigration can be considered to be a small benefit or a small cost to rich countries. However, it is undeniable that it represents a huge benefit for immigrants. Thus, when designing policies, the worth given to foreigners and citizens matters a great deal.

If foreigners are considered equal to or more important than citizens, then immigration is seen as a very positive affair. However, if the welfare of the citizens is prioritized, then immigrants may not be as beneficial after all.

Economic Ethics

Even though the predictions of economics are independent of any ethical theory, there are ethical ideas behind normative economic reasoning. Economists tend to respect other people's choices, and respect for preferences and choices naturally leads to respect for trade. Economists tend to treat all market demand equally, no matter which person they come from. When it comes to ethics, it is always difficult to judge what is right or wrong. Therefore, economists are not always right in their ethical consideration.

Shah, A. K., Mullainathan, S., & Shafir, E. (2012). Some consequences of having too little. *Science, 338*(6107), 682- 685.

Background

The poor often behave in ways that reinforce poverty: 1) often play lottery, 2) fail to enroll in assistance programs, 3) save too little 4) and borrow too much. These behaviors have 2 key explanations.

1. Circumstances of poverty (e.g. political representation, education etc). The poor live in socio-economic and political environments that promote such behaviors.
2. Second view is about personality traits: resource scarcity creates its own mindset, changing how people look at problems and make decisions.

Theory of the article

For this article core mechanism is number 2: **having less elicits greater focus**. This view is not bound to specific circumstances or dispositions of poor. So first, authors assume that fundamental feature of poverty is having less. Hence, **poor people focus on problems where scarcity is most salient.**

Second assumption follows from the first: **because scarcity elitists greater engagement in the problem, it leads to neglect of others**. From this idea follows explanation of extensive borrowing. Scarcity, of any kind, will create a tendency to borrow, with insufficient attention to whether the benefits outweigh the costs.

Experiments

5 experiments were conducted to test the theory. In all experiments participants were allocated to be either 'poor' or 'rich', with 'poor' having smaller budgets and different borrowing conditions. Experiment 1 and 2 show how scarcity creates focus. Experiments 2 to 5 how scarcity makes people borrow more.

Experiment 1: 60 participants played Wheel of Fortune. Budgeting was different in terms of chances to guess letters in a word puzzle, where 'rich' had more guesses than 'poor'. In this experiments poor performed worse. Scarcity increased the engagement in the problem, which made poor more fatigued (tired) and depleted.

Experiment 2: Shows how scarcity changes engagement. Participants played 'Angry Birds' with different budgets in terms of number of shots. Poor were aging more engaged which led to higher performance per shot. Hence, if scarcity can engage us too much, abundance might engage us too little. Moreover, poor participants borrowed more which was counterproductive, hence, suggesting that poor overborrowed. So, scarcity' focusing effects: The more focused the poor were on the current round, the more they neglected (and borrowed away from) future rounds.

Experiment 3: Here another form of scarcity was considered: having too little time in the game Family Feud. Some 'poor' participants also could borrow with different conditions. Regardless of interest rate, poor participants borrowed a greater proportion of their budget. The poor performed best when they could not borrow, less when could borrow without interest and the worst when borrowed with interest.

Experiment 4: Family Feud was modified, so that borrowing would create 'debt' in subsequent rounds. Poor participants borrowed greater proportion of their budgets, moreover, they did not adjust borrowing strategies as they accumulated debt.

Experiment 5: Offers more support for notion that scarcity creates attentional neglect. Participants again were playing Family Feud but now some participants could see previews oof subsequent rounds questions. In the experiment, previews helped rich to save more. Although poor could benefit with previews by borrowing less, there focus was on current round, hence, they overborrowed again.

Discussion

Taken together, these studies provide compelling support for the notion that scarcity elicits greater engagement and that a focus on some problems leads to neglect of others (manifesting in behaviors such as overborrowing). Poor and rich in experiments did not approach tasks in the same way: poor were more engaged. Another explanation is that scarcity creates a cognitive load, thereby, diminishing performance. This consumes attentional resources and leaves less for elsewhere.

This gives clear policy implications: Interventions that draw people's attention to specific future needs should be particularly effective at increasing savings. A scarcity mindset leads to the most convenient response to pressing demand, leading to constant financial juggling. this approach can help us to understand circumstances even broader than poverty, because scarcity underlies problems as dire as hunger and as mundane as busyness.

Robeyns, I. (2005). The Capability Approach: a theoretical survey, *Journal of Human Development*, 6:1, 93- 105

The goal of this article is to present a theoretical survey of the capability approach.

What is the Capability Approach?

The **capability approach** is an economic theory conceived in the 1980s by Amartya Sen and Martha Nussbaum – as an alternative approach to welfare economics – that focuses on what people are able to do and be (their capabilities). It is a normative framework for the assessment of individual and societal well-being, the design of policies, and suggestions about changes in society. It can be used to evaluate several aspects of people's and society's well-being, such as inequality, poverty, and social policies. However, the capability approach is not a theory that can explain poverty, inequality, or well-being; instead, it provides a tool and a framework within which to conceptualize and evaluate these phenomena. The approach focuses on the plural aspects of well-being and highlights the difference between means and ends, and between substantive freedoms (capabilities) and outcomes (achieved functionings). **Functionings** are effective opportunities that make life valuable and include working, resting, being literate, being healthy, being part of a community, being respected, and so forth. What is ultimately important is that people have the freedoms or valuable opportunities (capabilities) to lead the kind of lives they want to lead, to do what they want to do, and to be the person they want to be.

An Alternative Framework for Well-Being and Justice

The capability approach is a framework of thought about normative issues. It is a paradigm that can be used to evaluate several aspects of human well-being. This approach entails a critique of other evaluative approaches, mainly the welfarist approach, and on utilitarian and income-based or resources-based theories.

Welfarist theories rely exclusively on utility and thus exclude non-utility information from our moral judgements, whereas the capability approach is also concerned with the information that is excluded from a normative evaluation. San, indeed, rejects normative theories that rely exclusively on **mental states**. San has also criticized Rawl's use of primary goods – instead of capabilities – for interpersonal comparisons because they fail to reflect human diversity. Moreover, he has critiqued other resource-based normative theories, which rely on equality of resources for welfare comparisons.

Means Versus Functionings

A crucial distinction in the capability approach is the distinction between the means, such as goods and services, on the one hand, and functionings and capabilities on the other hand. A good cannot be judged in itself, but it can have certain characteristics that enable a functioning. For instance, a bike is only an object composed of certain materials, with a certain shape and color, but it enables the functioning of mobility.

The relation between a good and the functionings can be influenced by three groups of **conversion factors**:

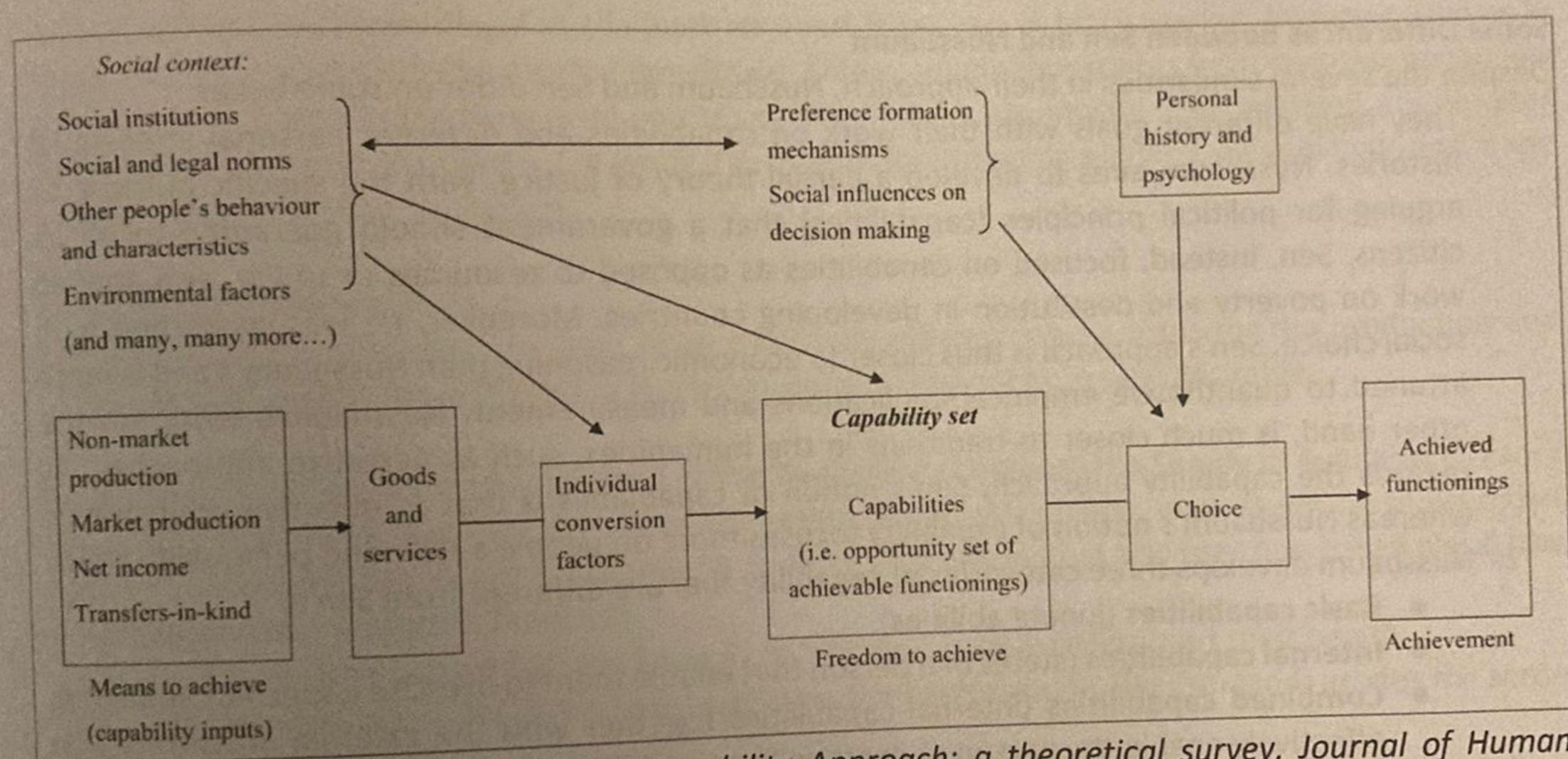
1. **Personal conversion factors** (e.g., physical condition, sex, personal skills). For instance, a disabled person would find a bike of limited help to enable the functioning of mobility.
2. **Social conversion factors** (e.g., policies, social norms, religious beliefs). For example, if there are heavy restrictions to bike transit, a bike can hardly be used to enable the functioning of mobility.
3. **Environmental conversion factors** (e.g., climate, geographical location). For instance, if the area is mountainous, then a bus could not enable the functioning of mobility.

The capability approach thus considers human diversity in many ways:

1. By focusing on the plurality of functionings and capabilities.
2. By enhancing personal and socio-environmental conversion factors of commodities into functionings.
3. By focusing on the whole social and institutional context that affects the conversion factors and also the capability set directly.

The capability approach not only presents an evaluation of people's capability sets, but also stresses the need to analyze the economic and social context to judge whether the circumstances in which people choose from their opportunity sets are enabling and just.

The figure, taken from the article, shows a stylized non-dynamic representation of a person's capability set and her social and personal context.



Note. From: Robeyns, I. (2005). The Capability Approach: a theoretical survey, Journal of Human Development, 6:1, 93- 114.

Achieved Functionings Versus Capabilities

The author makes some remarks about functionings and capabilities. The first two are terminological; the other two focus on the use of functionings versus capabilities in evaluative exercise and policy design.

1. The term **capability**. In Sen's earliest work, each capability referred to one person. In this terminology, a capability is synonymous with a capability set, which consists of a combination of potential functionings. Functionings can be potential or achieved. Later on, Nussbaum and many other scholars have defined a capability as a potential functioning. In this terminology, the capability set consists of a number of capabilities.
2. The term **basic capabilities**. In Sen's work, basic capabilities are a subset of all capabilities. Basic capabilities refer to the freedom to do some basic things that are necessary for survival and to avoid poverty.
3. There are cases where it's more sensical to investigate peoples achieved functionings directly, instead of assessing their capabilities. For example, situations of extreme material and bodily deprivation in very poor societies require a closer focus on functionings rather than capabilities.

4. As a liberal philosophical framework, the capability approach respects people's different ideas of the good life, which is why capability, and not achieved functioning, is the appropriate political goal. However, it is important to question to what extent people genuinely have access to all the capabilities in their capability set, and whether or not they are influenced by their family, tribal, religious, community, or cultural ties and background.

Distinguish Well-Being from Agency

Sen has also distinguished well-being from agency. The standard of living is "personal well-being related to one's own life". If we add the outcomes resulting from sympathies (i.e., feeling better from helping another person), we measure well-being. If well-being is supplemented with commitments (i.e., an action that is not beneficial to the agent herself), then we have overall agency. Moreover, all of these concepts can either be achievements (actual achieved outcomes) or freedoms (the freedom to achieve such outcomes).

Some Differences Between Sen and Nussbaum

Despite the several similarities in their approach, Nussbaum and Sen differ on some issues.

1. They have different goals with their work on capabilities and different personal intellectual histories. Nussbaum aims to develop a partial theory of justice, with the specific purpose of arguing for political principles (capabilities) that a government should guarantee to all its citizens. Sen, instead, focused on capabilities as opposed to resources or utility, and applied work on poverty and destitution in developing countries. Moreover, he was concerned with social choice. Sen's approach is thus closer to economic reasoning than Nussbaum's and is more attuned to quantitative empirical applications and measurement. Nussbaum's work, on the other hand, is much closer to traditions in the humanities, such as narrative approaches. As regards the capability approach, Sen's notion of capabilities is that of effective opportunity, whereas Nussbaum's notion of capability focuses more on people's skills and personality traits.
2. Nussbaum develops three categories of capability that are different from Sen's:
   - **Basic capabilities** (innate abilities).
   - **Internal capabilities** (states of a person that enable them to exercise a specific capability).
   - **Combined capabilities** (internal capabilities together with the external provisions that effectively enable the person to exercise the capability).

   However, Nussbaum and Sen agree that politics should focus on combined capabilities.
3. Nussbaum proposes a concrete list of capabilities, whereas Sen has always refused to endorse a specific list.
4. Nussbaum believes that the government has the duty to formulate constitutional principles that respect and reflect citizens' capabilities, whereas Sen's scope is broader and does not directly address governments.
5. Nussbaum does not endorse the agency-well-being distinction that Sen advocates and argues that "all the important distinctions can be captured as aspects of the capability/functioning distinction".

Satz, Debra (2008). 'The Moral Limits of Markets: The Case of Human Kidneys', *Proceedings of the Aristotelian Society*, 108:1(3), 269–288

This paper examines the morality of kidney markets through the lens of choice, inequality, and weak agency. The author sides against a market for human kidneys and stresses the idea that the choice for some to sell their kidneys can have external effects on those who do not want to do so. The author's account is a philosophical one. Her main purpose is to discuss the values at stake in a market for kidneys, and how such a market differs from markets in "normal goods".

Libertarian perspective views an organ ban as an illegitimate infringement with personal liberty.
Non-libertarian perspective views an organ ban is morally dubious, as lives would be saved by the increased supply.

Brief Background: The Status Quo Systems of Kidney Procurement

Kidney selling is currently illegal in all countries apart from Iran. A black market, however, thrives in many countries, especially the most undeveloped. Kidney supply thus comes from donations only. The organs can be transplanted from dead corpses or alive individuals, mostly close relatives, or intimates of the patient in need of a kidney. Currently, the queues for obtaining a kidney are very long, and many people die waiting for a transplant.

Anti-Market Considerations

The author agrees that markets tend to be extremely efficient mechanisms for the production and distribution of goods. Nevertheless, a market in kidneys may not be totally desirable.

A. An Important Side Issue: Does a Market Ban Necessarily Decrease the Supply of Available Organs?

The author reports Richard Titmuss' argument against a market for blood. In his study, he showed that a system of donated blood (the British one) is superior to a system that also uses purchased blood (the American one) for two main reasons:

1. A remunerated system for blood supply incentivizes people to lie about their health.
2. The introduction of blood markets "represses the expression of altruism [and] erodes the sense of community [...]".

Titmuss' conjecture received mixed support from some recent work in experimental economics. On the one hand, Bruno Frey and others have found that **intrinsic motivation** is partially crowded out when price incentives are introduced. On the other hand, different studies found that the introduction of price incentives "crowded in" civic action (e.g., carpooling in San Diego). Nevertheless, it remains unclear whether a kidney market would increase the supply of kidneys, since this depends on how many people actually donate out of altruism versus how many would sell their kidneys for extrinsic motivations.

B. Are Kidney Markets Really Examples of Free Exchange? A Distinction: Individual Choices Versus Choice Sets

Many people would side against a market for kidneys even if it would increase supply because it is seen as a "desperate exchange", an undesirable trade that nobody would ever engage into unless forced by external situations. The author emphasizes that even though the choice of selling organs is voluntary, it might be coercive.

Satz especially wants to point out that a restriction on kidney sales can be autonomy-inhibiting from the viewpoint of a particular individual, but it may be autonomy-enhancing from the point of view of society. This is because allowing such markets as a widespread practice, backed up by laws, may have effects on the nature of the choices that are available to people. The introduction of kidney markets can change environments, individual preferences, capacities, and values. Even in the case of a fair trade and an uncoerced choice, as Ronald Dworkin argues, body parts should not be part of our social resources at all.

The objection is based on the idea that people should not have to pay a cost for refusing to sell their body parts. A person's relationship to their body is so important that it should enjoy a special protection against the effects of market forces. Nevertheless, especially in the third world, governments have been struggling with a thriving black market. If the poor are as likely to participate in a black market for kidneys as in a legal market, then banning the exchange does not seem a good way of honoring their autonomy.

C. Objectionable Inequality and Asymmetric Vulnerability

A kidney market would almost certainly entail that poor people would disproportionately be the organ sellers of the world, while rich people the recipients. By contrast, a procurement system that relies on donations is much more likely to have suppliers that come from all classes of people. Some critics raise the legitimate concern that the introduction of kidney markets will actually worsen existing inequality. A kidney market would seem to mean that kidneys go to the highest bidders.

But should kidneys not be allocated to people on the basis of need, length of time waiting, and medical suitability, and not on the basis of ability to pay? Theoretically, however, governments might engage in regularization of such markets, in order to provide funding for the poor, or even create a monopoly in which it is the only legal buyer of organs. Nevertheless, it remains difficult to make kidney allocation completely independent of the wealth of the donor.

D. Agency Failures

Some markets do not involve fully and equally informed participants. Sellers of kidneys, especially in developing countries, are often not fully conscious of the consequences of an organ transplant. Moreover, health risks are likely to be greater in places where people have little access to clean water or adequate nutrition and are often engaged in difficult manual labor. In our non-ideal world, namely the third world countries, information problems and possibility of harm (weak agency) make it undesirable to establish a legal kidney market. Nevertheless, the argument from weak agency (inadequate information) does not single out what is problematic about the kidney market and might indeed discourage altruistic organ donation as well.

Concluding Thoughts

The author agrees that under ideal circumstances, market regulation may solve problems along the dimensions of weak information and inequality. Moreover, harmful health complications from a kidney transplant might be mitigated through follow-up care. However, the choice issue might not be so easy to solve through regulation. The author believes that the strongest argument against allowing kidney markets in ideal circumstances is that they make people incur a price that they can reasonably think that they should not have to pay. A person's relationship with their body is so important that it is not something that we should allow markets to influence. At the same time, this market would probably not be an ideal one. Instead, it is more likely to entail several market failures (e.g., weak agency, monopoly power, inequality). To conclude, the author considers four recent proposals for kidney markets and analyses them in terms of weak agency, inequality, and choice, as can be seen in the table below.

The four markets he considers are:

1. Markets governing both the supply and demand for kidneys.
2. Markets governing the supply of kidneys only while distributing on the basis of need, not the ability to pay.
3. Futures markets where a person's organs are only given up upon death.
4. In-kind exchanges in which a patient with a willing donor who has an incompatible blood type can exchange with another such incompatible patient donor pair.

A plus indicates a high (worrying) score along a dimension, while a minus indicates a relatively low score.

	Weak Agency	Inequality	Restricts Choice Sets
Markets in supply and demand	+*	+	+
Markets in supply only	+*	−	+
Futures Markets	+	− ?	?−
Matching markets	+	−	−

* Since harvesting kidneys from living donors involves risks, in these cases the lack of information is coupled with the potential for harmful outcomes.[27]

Note. From: Satz, Debra (2008). 'The Moral Limits of Markets: The Case of Human Kidneys', *Proceedings of the Aristotelian Society*, 108:1(3), 269–288.

Lecture 4 – How Do We, As Economists and Business Administrators, Retain Mortality in Markets?

Cowen, T., & Tabarrok, A. (2015). *Modern principles of economics*, Chapter 22. Macmillan International Higher Education.

Lesson One: You Get What You Pay For

When designing an incentive scheme, policymakers have to remember that you get what you pay for. A problem arises when what you pay for is not exactly what you want. In this case, careful design of an incentive scheme can narrow the gap between **what you want** and **what you pay for**. For example, Chicago public school students take a standardized test every year. Schools with low scores are closed, teachers are reassigned, and principals are fired. The idea is to give educators stronger incentives to work harder and better. However, teachers also have a great incentive to cheat (4-5% of teachers cheated in 1996). Therefore, policymakers had to institute new procedures to reduce cheating.

Prisons for Profit?

Should the management of prisons be contracted out to the private sector? Many economists say no. The owners of a private firm have a strong incentive to cut costs and improve productivity because they get the resulting profits. We care about costs, but also about prisoner rehabilitation, civil rights, and low levels of inmate and guard violence. The principle is that a strong incentive scheme that incentivizes the wrong thing can be worse than a weak incentive scheme. Cutting quality is the way to cut costs. However, when quality and cost cutting go together, a private firm has a strong incentive to increase quality.

Price Rates vs. Hourly Wages

Piece rate is any payment system that pays workers directly for their output. When should workers be paid by the hour and when by the piece? Piece rates increase the incentive to work hard and well when output is easy to measure, so "what you pay for" is close to "what you want". Piece rates do not work well when quality is important, but quality control is expensive.

The advantage of piece rates is that they can greatly increase productivity. However, although both firms and workers can benefit from piece rates, they are sometimes not implemented because of mistrust. E.g., workers will not work harder if they expect that higher productivity will be punished with lower piece rates.

Lesson Two: Tie Pay to Performance to Reduce Risk

An auto dealer who wants to motivate her sales staff can implement strong incentives (piece rates) to increase sales. In fact, paying salespeople by the number of cars sold can incentivize them to work harder and better. However, auto sales also depend on factors that the staff have no control over, such as price and quality of the car, the price of gas, and the state of the economy. If incentives tie earnings directly to sales, the sales staff are going to do great when the economy is booming, but poorly when the economy is in a recession. Thus, the sales staff and the owner have to agree on who is to bear the risk of a recession and benefit from an economic boom. Is the best incentive a fixed salary or a piece rate?

Tournament Theory
When sales depend heavily on outside factors such as the state of the economy, tying bonuses directly to sales will reward or penalize agents for outcomes that are often beyond their control, thus shifting risk to the agent but giving them little incentive to exert effort.

To solve this, a manager can implement a tournament to tie rewards more closely to actions that a sales agent does control. A **tournament** is a compensation scheme in which payment is based on relative performance (compared to other agents). It limits the amount of risk from the external environment.

Improving Executive Compensation with Pay for Relative Performance
How can the idea of pay for relative performance be used to tie executive pay more closely to actions that executive's control? Most executives are paid based on how well their stocks perform, which in turn depends on factors that are out of the executive's control (e.g., economic booms and busts, price of related products). Paying executives for relative performance sounds like a good idea but it has not been widely adopted. Many issues, including incentives, are involved in the paying of managers. For example, a bank manager has a big incentive to take too many risks, since if the risk pays off, the manager gets a big bonus, whereas if the risk fails there are no big consequences or punishments for them.

Environment Risk and Ability Risk
Tournaments isolate the rewards from **environment risk**, or risk due to outside factors that are common to all players, but creates **ability risk**, or risk due to the ability of another agent.

A manager who wants a lot of effort will structure tournaments so that rewards are closely tied to effort. In fact, a person who just started a new job should not compete with people who have been working in the same position for several years. For example, a manager might create junior and senior sales positions with tournaments played within each class.

Tournaments and grades
Some professors grade on a curve, while others use an absolute scale. Grading on a curve creates better incentives to study when the big risk is that the professor will be bad (an environment risk), because in such a case everyone is likely to perform poorly. However, grading on a curve reduces the incentive to study when students are of very different abilities (ability risk) and hinders cooperation. Differently, grading on an absolute scale creates better incentives to study when students are of very different abilities (ability risk) but reduces the incentive to study when the big risk is that the professor will be bad (environment risk).

Lesson Three: Money Is Not Everything
Not all incentives are for money. Intrinsic motivation is doing something simply for feelings of enjoyment and pride.

Corporate culture is the shared collection of values and norms that govern how people interact in an organization or firm. A good manager will encourage workers to identify with the corporation and its goals in the same way that sports fans identify with their team. Monetary rewards work best when they are supported by intrinsic motivation and measures of social status.

If a friend asks you to drive him/her to the airport, you would be willing to do so gladly because of friendship. When he/she offers you $20, you might be offended, as it is now no longer because of friendship but because of a monetary reward.

Lesson Four: Nudges Can Work

Behavior can also be influenced by how a choice is presented or "framed". Planning where, when, and how choices are made is called **choice architecture**.

Cass Sunstein and Richard Thaler wrote a book called Nudge, which considers a variety of ways that governments and other organizations can use small changes in the choice architecture (nudges) to change or improve decisions.

DellaVigna, S., & Linos, E. (2022). RCTs to scale: Comprehensive evidence from two nudge units. *Econometrica, 90*(1), 81-116

Introduction

The text introduces nudges, defined by Thaler and Sunstein in 2008 as **behavioral interventions altering behavior predictably without restricting options or changing economic incentives.** These interventions, like simplification and social-norm comparison, gained popularity in research and later expanded into government applications globally. With over 200 government teams using behavioral science, the text emphasizes the unique opportunity to compare the impact of these interventions when implemented by researchers versus larger-scale government initiatives. The lack of comprehensive data on Nudge Unit interventions has previously hindered comparisons about consistency of nudge impacts.

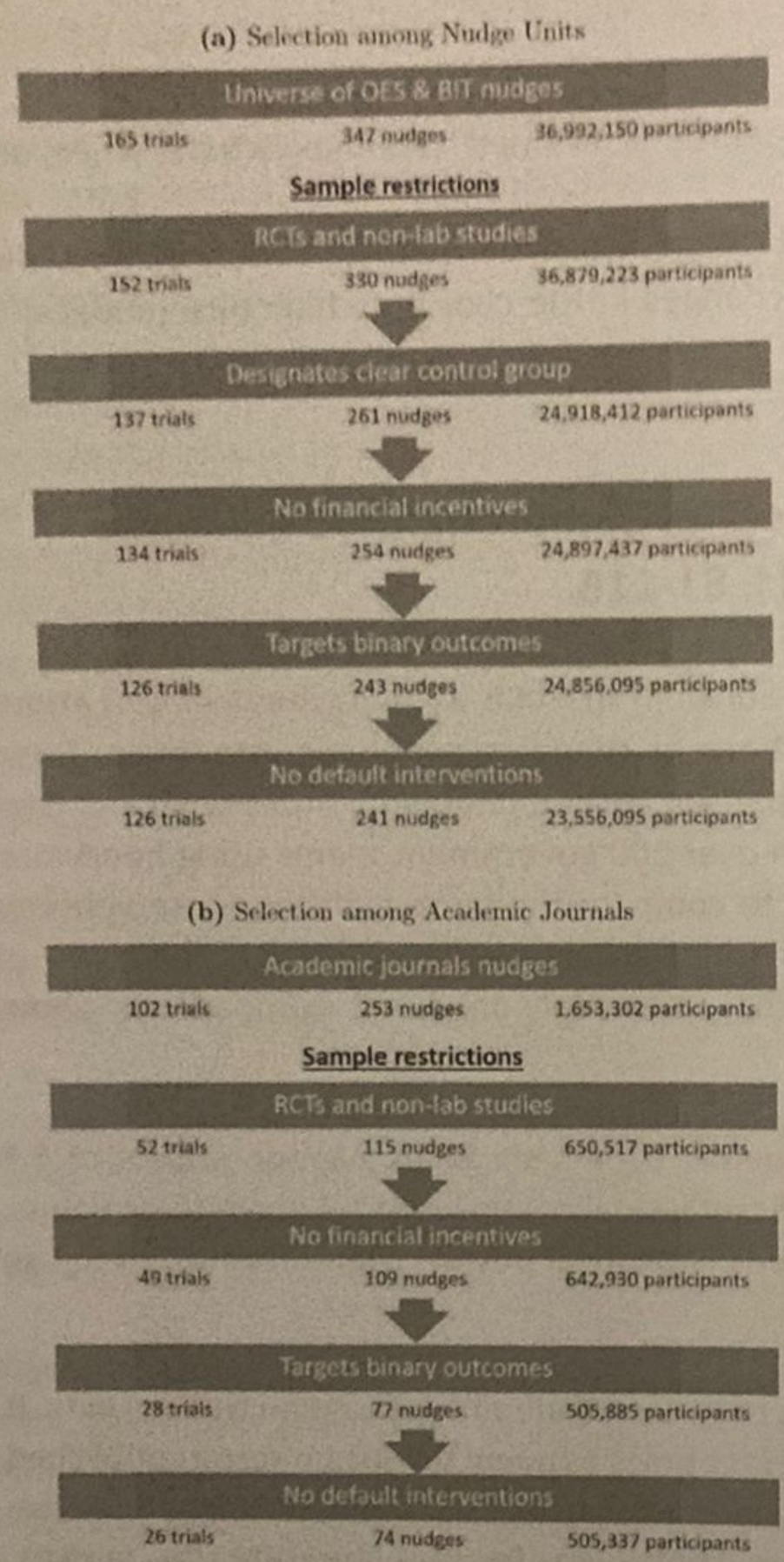
Comparing results with literature-based nudges, academic papers show an average impact of 8.7 percentage points, while Nudge Unit trials have a significantly smaller impact of 1.4 percentage points. The study attributes this difference to three dimensions: statistical power, characteristics of nudge interventions, and selective publication.

The paper identifies differences in nudge interventions between academic studies and Nudge Units. It highlights the impact of selective publication, with over four times as many significant results published in academic journals. The evaluation emphasizes the need to consider biases introduced by selective publication and low statistical power. The findings contribute to the literature on nudges' effectiveness, acknowledging potential institutional constraints faced by Nudge Units and the influence of trial design on subsequent interventions.

How can I use this article to paper?

Setting and Data

Nudge units



The paper examines trials by two major US Nudge Units: The Office of Evaluation Sciences (OES) and the Behavioral Insights Team's North America office (BIT-NA). Both units conduct Randomized Controlled Trials (RCTs) using behavioral interventions to improve government services. They prioritize good trial design, research transparency, and documentation of protocols. The two units have shared goals: to use behavioral science to improve the delivery of government services through rigorous RCTs, and to build the capacity of government agencies to use RCTs. Final sample consists of 126 randomized trials that include 241 nudges and involve 23.5 million participants (see. Figure)

Trials in Academic Journals

In a recent meta-analysis, Hummel and Maedche (2019) selected 100 papers screened out of over 2000 initial papers identified as having "nudge" or "nudging" in the title, abstract, or keyword. A second meta-analysis that covers several areas is Benartzi et al. (2017), which does a cost-benefit comparison of a few behavioral interventions to traditional incentive-based interventions. We thus combine the behavioral trials in Hummel and Maedche (2019) and in Benartzi et al. (2017), for a total of 102 trials (see figure).

Nudge Categories

A typical "**revenue & debt**" trial involves nudging people to pay fines after being delinquent on a utility payment, while an example of a "**benefits & programs**" trial encourages individuals to enroll in a government program, such as pre- and post-natal care for Medicaid-eligible mothers. The "**workforce and education**" category includes prompting jobseekers to plan their job search strategy. A "**health**" intervention urges people to get vaccinated or sign up for a doctor's appointment. A "**registration**" nudge asks business owners to register their business online as opposed to in-person, and a "**community engagement**" intervention motivates community members to attend a local town hall meeting. Compared to the Nudge Units sample, the **Academic Journals** sample has a **larger share of** trials about health outcomes and environmental choices and **fewer** about revenue and debt, benefits, and workforce and education.

Additionally, authors identified a few differences in trial features between the Academic Journals sample and the Nudge Units sample. The Nudge Unit interventions are less likely to be led by academics, tend to face higher institutional constraints, and involve fewer personnel. These features (or at least the first two) seem typical of an "at scale" intervention. Nudge Unit trials may have a different objective function, where more emphasis is placed on moving the policy outcome and less on disentangling the exact mechanism.

Impact of Nudges

The study reveals unweighted impact comparisons of nudges in two samples: Academic Journals and Nudge Units.

Academic Journals: The average treatment effect for 74 nudges across 26 trials is 8.68 pp. (s.e. = 2.47), representing a significant increase compared to the average control group take-up of 25.97 percent. It shows substantial heterogeneity in impact.

Nudge Units: The unweighted average impact of 241 nudge treatments in 126 trials is 1.39 pp. (s.e. = 0.30), relative to an average control take-up of 17.33 pp. While this estimated treatment effect remains significant and precise, it is **one-sixth the size of the effect observed in the Academic Journal papers.**

Two key differences between published academic papers and Nudge Unit interventions:

A. Difference in estimated treatment effects, which are generally larger, and more dispersed in the Academic Journals sample.
B. Statistical precision of the estimates: the confidence intervals are much tighter for the Nudge Unit studies, which are typically run with a much larger sample.

Nudge Units versus academic journal nudges

Authors assume that researchers and nudge units alike design experiments aiming to provide evidence on a particular treatment, in our case a nudge. They propose that the observed differences between two samples can be largely explained by differences in those three components, **d, β, and γ**:

1. **Different research design**. Both academic researchers and Nudge Units design an experiment to detect an effect size **d with 0.80 statistical power**.
2. There is a **true effect size of the nudge intervention β distributed with a random effect.**
3. **Results that are not statistically significant are published by academic researchers with some probability γ < 1**, while **results that are statistically significant are published with probability 1.**

Results:

For the Nudge Unit trials, the meta-analytic estimate is consistent with the unweighted estimate of 1.4 pp. For the Academic Journal trials, selective publication accounts for about 70 percent of the larger effect size relative to the Nudge Unit trials. If key nudge characteristics are included as predictors, these characteristics explain most of the remaining difference, since the observable characteristics of nudges in the Academic Journals are more often associated with larger effect sizes.

Discussion and Conclusion

An ongoing question in both policy circles and in academia asks: what would it look like if governments began using the "gold standard of evaluation"—RCTs—more consistently to inform policy decisions?

By studying the universe of trials run across two large Nudge Units in the United States, covering over 23 million people, and comparing our results to published meta-analyses, authors make progress on this question. They find, on average, nudge interventions have a meaningful and statistically significant impact on the outcome of 1.4 pp. This estimated effect is significantly smaller than in academic journal articles (8.7 pp.). Using a meta- analysis model, this difference is decomposed. **The largest source of the discrepancy** is selective publication in the Academic Journals sample, exacerbated by low statistical power in that sample. If we include key nudge characteristics as predictors of the effect size, these characteristics explain most of the remaining difference, since the characteristics of nudges in the Academic Journals are more often associated with larger effect sizes.

Jensen, Michael C. (2002). Value maximization, stakeholder theory, and the corporate objective function. *Business Ethics Quarterly*, 12:2, 235-256

Introduction

In this article, the author wants to clarify the concept of enlightened **value maximization**, the proper relation between value maximization and stakeholder theory. "This house believes that in implementing organizational change, managers must have a criterion for deciding what is better, and better should be measured by the increase in long-term market value of the firm."

- **Value maximization** states that managers should make all decisions so as to increase the total long-run market value of the firm.
- **Stakeholder theory** says that managers should make decisions so as to take account of the interests of all the stakeholders in a firm. Those include all individuals who can substantially affect the welfare of the firm, such as financial claimants, employees, customers, and communities.

The Logical Structure of the Problem

In discussing whether firms should maximize value or not, the author separates two distinct issues to what he gives the following answers:

1. Should the firm have a single-valued objective? Purposeful behavior requires the existence of a single-valued objective function.
2. Should that objective be value maximization or something else? Total firm value maximization makes society better off.

Purposeful Behavior Requires the Existence of a Single-Valued Objective Function

It is logically impossible to maximize in more than one dimension at the same time unless the dimensions are monotone transformations of one another. For example, if an increase in current-year profits can come only at the expense of market share (and vice versa), then a manager is faced with a tradeoff. Maximize market share or profits? A firm can solve this issue by setting an objective function that shows the tradeoffs among the various dimensions and seek to maximize it.

Total Firm Value Maximization Makes Society Better Off

Given that a firm must have a single objective that tells us what is better and what is worse, we then must face the issue of what the definition of better is. Social welfare is maximized when all firms in an economy maximize total firm value, which is the sum of the market values of equity, debt, and any other contingent claims outstanding on the firm.

Value Maximizing and Social Welfare

Profit Maximization

The issue to be addressed is what firm behavior will get the most out of society's limited resources. Profit maximization leads to an efficient social outcome because a firm should increase output as long as an additional dollar of resources taken out of the economy is valued by consumers more than a dollar. Therefore, when the firm acquires an additional unit of any input to produce an additional unit of any output, it increases social welfare at least by the amount of its profit – the difference between the value of the output (revenue) and the cost of the input required in producing it (cost).

Value and Tradeoffs Through Time

Interest rates tell us the cost of giving up a unit of a good today for receipt at some time in the future, namely the value of moving a marginal dollar of resources (inputs or consumption goods) forward or backward in time. The corporate objective function that maximizes social welfare thus becomes "maximize total market value of a firm".

Stakeholder Theory

The Specification of Trade-Offs and the Incompleteness of Stakeholder Theory

Any theory of action must tell the managers how to choose among multiple competing and inconsistent constituent interests. Value maximization provides the following answer to the tradeoff question: Spend an additional dollar on any constituency to the extent that the long-term value added to the firm from such expenditure is one dollar more.

Conflicts Between Family and Markets and Their Role in Stakeholder Theory

Stakeholder theory holds a conflict between family and markets. Markets have implemented a system for voluntary and decentralized coordination of human action that has brought a huge increase in the welfare of humans and in their freedom of action. This, however, is opposed to the idea of commitment to the family and tribe.

混合する

Enlightened Value Maximization

The way out of the conflict between value maximization and stakeholder theory lies in melding together enlightened value maximization and enlightened stakeholder theory. Enlightened value maximization recognizes that communication with, and motivation of, an organization's managers, employees, and partners is extremely difficult. People need enough structure to understand what maximizing value means so that they can be guided by it, and therefore have a chance at actually achieving it. They must be turned in by the vision or the strategy in the sense that it taps into some desire deep in the passions of human beings (e.g., a desire to build the best automobile).

Enlightened Stakeholder Theory

Enlightened stakeholder theory takes stakeholder theory and adds to it the simple specification that the objective function of the firm is to maximize total long-term firm market value. The firm has to adapt to the varying conditions of the market and create value in the long run.

Measurability and Imperfect Knowledge
The world is governed by complex dynamic systems that are difficult to optimize in the usual sense. However, we must not confuse optimization with value creation or value seeking. To create value, we need not know exactly where and what maximum value is, but only how to seek it, that is how to institute changes and strategies that cause value to rise. That should be our scorecard, which gives us the notion of "better".

The Balanced Scorecard
The balanced scorecard, designed by Kaplan and Norton, is the managerial equivalent of stakeholder theory. The author explains how this is flawed because it presents managers with a scorecard that gives no score – that is, no single valued measure of how they performed. In fact, the balanced scorecard complements financial measures of past performance with measures of the drivers of future performance. However, the author points out that it misses on performance measurement and how rewards and punishments are linked to the performance measure for each person within an organization.

Managers, therefore, cannot make a purposeful decision since they do not know the tradeoffs between the multiple measures. The solution is to define a true (single dimensional) score for measuring performance for the organization or division, and as long as their score is defined properly, this will enhance their contribution to the firm.

In a **decentralized** solution, headquarters determine the appropriate performance measure for the unit. In the **centralized** solution, headquarters determine the performance measure by giving a functional form which lists all the drivers and describes the importance of each driver in the determination of performance measure.

Sandel, Michael J. (2011). Justice: What's the right thing to do? *Boston University Law Review*, 91:4, 1303-1310

This lecture is derived from the author's namesake book, which lays out three approaches to **justice**:

1. The utilitarian idea of maximizing welfare or happiness.
2. The idea that justice means respecting freedom and human dignity.
3. The idea that justice has to do with honoring and recognizing virtues, and the goods implicit in social practices.

The author defends the third idea about justice and refuses the first two, seeing them as inadequate. He believes we cannot detach questions of justice and rights from debates about the nature of the good being distributed. In short, justice is judgmental. This idea is indebted to Aristotle in two respects:

1. Justice is theological (we must find the purpose, **telos**, of the good being distributed).
2. Justice is honorific (we must reason about what virtue the social practice should honor).

The author illustrates his point by discussing the just distribution of four goods: musical instruments, university admission, golf carts, and marriage. His point is that justice is not simply a matter of maximizing utility or of securing freedom of choice and non-discrimination. The logical reaction is that the best flutes go to the best players. The purpose of having music performances is not only to make the audience happy, but to honor the musicians as well.

Another example is that of Casey Martin, a golf player with a bad leg. He wanted to have a golf cart, but the official association refused. The reason for it was that walking to the next course is an essential part of the sport. This argument was about the purpose, honor, and recognition of the sport.

Justice is not simply a matter of maximizing utility or securing freedom of choice and non-discrimination. To achieve a just society, we have to reason about the meaning of the good life, and by this, we cannot avoid entering into substantive moral disagreements about the purpose of the social practice, and the virtues and goods the social practice honors.