Market Efficiency and Market Failures

Jón Steinsson*
UC Berkeley

February 19, 2022

We live in an age of capitalism. Private property and voluntary exchange between private individuals play a central role in the organization of economic activity. We frequently describe economic activity as occurring on free "markets." In actuality, however, few products are traded on organized markets. Rather, we have come to use of the term "markets" as a short-hand metaphor to describe voluntary exchange of goods and services by individuals (and groups formed by individuals). Markets, conceived of in this broad sense, are a central institution in the modern world.

This heavy reliance on markets is a relatively recent phenomenon. Until a few hundred years ago, most economic activity consisted of small groups with strong kin ties, such as extended families, villages, or tribes, engaged in subsistence farming. Specialization and trade was limited due to high levels of violence, pervasive lack of freedom (e.g., serfdom, slavery, class and caste restrictions, etc.), absence of individual property rights (e.g., property rights over land), poor infrastructure, piracy, tolls, and various other obstacles to trade.

The rise of markets over the past few hundred years has coincided with an enormous increase in economic well-being for a large and growing fraction of the world's human population. Yet the growing importance of markets has been extremely controversial, and still is. Critics lament alienation and exploitation of the weak by the strong. Karl Marx, the most influential among these critics, argued that with the growth of capitalism "grows the mass of misery, oppression, slavery,

*I would like to thank Kenneth Ayotte, Gregory Mankiw, Emi Nakamura, Jessica Pham, Dmitriy Taubinsky, and Jonathan Vogel for helpful conversations. I would like to thank Vivek Adury, Emma Berman, Zhuowen Cheng, Adrian Fan, Jesse Flowers, Lawrence Garafola, Jacob Goldstein, Arnold Kim, Benjamin Nhan, Paul Underwood, Akhalya Valluvan, John Vincent and Kevin Wang for finding errors and typos. First posted in September 2021.

degradation, exploitation" (Marx, 1867/1887, ch. 32). Critics of free market capitalism often argue that other ways of organizing society would yield more overall economic well-being or a more equal distribution of economic well-being. Few topics outside of sports and religion generate as much heated debate as does free market capitalism.

Are these critics correct? Have we come to rely too heavily on markets for our own good? What is the appropriate role of markets in society? These are some of the most central questions that the field of economics seeks to shed light on. Proponents of free market capitalism put forth two main intellectual justifications for giving markets a prominent role in society. One justification is that freedom is good in and of itself. Free market capitalism is an economic system that delivers a large amount of a particular type of freedom. This feature arguably makes it a good system simply because freedom is a desirable end. Prominent scholars that have argued for free market capitalism on these grounds include Frederick A. Hayek, Milton Friedman, and Robert Nozick. Sandel (2009) provides an excellent recent "textbook" treatment of this idea (chapter 3 of that book) as well as several other competing notions of justice.

A second intellectual justification for free market capitalism is the notion that allowing individuals to engage in free exchange will yield higher levels of material well-being than other systems for organizing society. Proponents of free market capitalism such as Hayek, Friedman, and Nozick often make use of both of these arguments and sometimes don't distinguish between them clearly. But these two arguments are fundamentally different one from the other. Whether a person values freedom as an end—rather than a means to some other end—is a matter of preference, or perhaps a matter of morality. In sharp contrast, whether free market capitalism yields a higher level of material well-being than other systems of government is a scientific question that can in principle be settled by empirical evidence. Amassing evidence that bears on this question is one of the central goals of economics.

The degree of material well-being generated by free market capitalism relative to other systems for organizing society is of course a very big and complex question for which complete and conclusive evidence does not yet exist (and perhaps never will). Throughout this book, I will discuss pieces of evidence that bear on this question. However, in this chapter, we will focus on discussing under what conditions this argument for free market capitalism is valid. This theoretical question is useful because it can help guide our search for relevant empirical evidence. If condition X is crucial for free markets to generate more material well-being than some other way

of organizing society, then it is important for empirical economists to figure out to what extent condition X is satisfied in reality. Understanding the conditions needed for free market capitalism to generate good outcomes regarding material well-being will also help us organize our discussion of potential drawbacks of free market capitalism. In the last part of this chapter, we will discuss some of the main potential limitations of free market capitalism as a vehicle for generating material well-being.

1 Maximizing Material Well-Being: Two Challenges

Suppose we set ourselves the task of designing from scratch a system of government — i.e., laws, regulations, and institutions — that attains a maximal level of material well-being for the citizens of the society. What should such a system look like?

Several potential challenges are important in designing such a system. One challenge is the *information challenge*. This is the challenge that much of the information needed to allocate resources efficiently in a society is dispersed among the individuals that make up that society. Several types of information are relevant. First, peoples' preferences are important. Some people like apples, while others like oranges. Some people like scary movies, while others like romantic comedies. Some people like living in a bustling urban area, while others like living in a quiet suburb. The complexity of people's preferences is staggering. A trip to the local supermarket should make this clear. How many different varieties of peanut butter are available? (creamy vs. crunchy, natural vs. not, salt vs. no salt, extra sugar vs. no extra sugar, etc.) Or consider the huge number of products available to someone that is shopping for a living room sofa, bathroom tiles, or shoes. The variety of offerings is endless. One reason for this is presumably that people have very heterogeneous preferences regarding these goods. To produce a large amount of material well-being, a society must produce the "right" amount of each of these numerous goods and services and allocate them to the "right" people.

The second type of information that is important is peoples' abilities. Some people are dexterous, while others are thoughtful. Some people are good at physics and math, while others are good at writing and languages, and so on. To produce a large amount of material well-being, a society must allocate workers to the jobs that they perform well in. But people's preferences also matter for allocating workers. A person may be a very good accountant but enjoy working as a realtor so much more that it makes sense for them to do that instead. To complicate matters even

further, many tasks in society are best handled by teams of workers. In this case, it is important to know which sets of workers work well together and can therefore form a good team.

A third type of knowledge that is important is technical knowledge about how things are made and knowledge about physical resources. Efficient production and distribution of goods and services depends a great deal on huge amounts of scientific and engineering knowledge having to do with production processes. Furthermore, as Adam Smith emphasized, efficient production typically features an enormous amount of division of labor—i.e., the combination of many highly specialized of tasks in a complex supply chain. For example, building an automobile is a stupendously complicated task involving thousands of specialized parts. Building automobiles by the tens of thousands at an average cost of less than $20,000 a piece (at 2020 prices) is a truly amazing achievement. Detailed knowledge of how to make these parts, from where to source them or the inputs used to make them, how to finance this activity, and how to assemble them into an final product are crucial to efficient production.

Ever since the rise of capitalism a fierce debate has raged regarding the relative merits of *central planning* versus free-market capitalism. Both systems potentially face a severe information challenge. In the words of Hayek (1945):

> Which of these systems is likely to be more efficient depends mainly on the question under which of them we can expect that fuller use will be made of the existing knowledge. And this, in turn, depends on whether we are more likely to succeed in putting at the disposal of a single central authority all the knowledge which ought to be used but which is initially dispersed among many different individuals, or in conveying to the individuals such additional knowledge as they need in order to enable them to fit their plans in with those of others. (page 521)

While the individuals in society have a great deal of crucial information about themselves, their circumstances, and their surroundings, it is not clear that they have enough information to make good choices. For example, each individual knows something about their own abilities. Someone may think that they are a good plumber. But what does it mean to be a good plumber? Perhaps it means that the person is a better plumber than most other people. To reach such a judgment, the person must have knowledge of the distribution of other people's abilities as plumbers. Suppose for the sake of the argument, that the person does know this.

Does this mean that they should work as a plumber? Actually, it does not. One of the most basic insights of economics is that production is maximized if people do the things that they have a *comparative advantage* doing, not the things they have an absolute advantage doing. The master civil engineer may be better at taking measurements than her assistants. But this does not mean she should take all measurements herself. A person that has relatively low ability at all tasks should not sit idle. To maximize the economic output of society, one must compare the relative ability of a person at one task with their relative ability at other tasks and assign them to the task for which their relative ability is highest among all tasks, i.e., the task for which they have a comparative advantage.

This implies that basic economic choices such as what job a particular person should perform depend not only on that person's abilities (which he or she has unique insight into) but also on the abilities of everyone else in society across all jobs. Since each person knows only their abilities, but not necessarily those of others, it may seem like a recipe for disaster to let people choose themselves what they should do for a living without any direction from a central authority.

The enormous complexity and information burden associated with the efficient allocation of resources has lead many thinkers to conclude that a rational and scientific allocation of resources can only be attained by a great deal of careful planning by a central authority. Others, however, reach the opposite conclusion: How could a central authority possibly gather the relevant information? Just asking probably won't work too well. In many cases, people have an incentive to be less than truthful. In particular, they have an incentive to overstate their needs and understate their abilities which makes organizing society according to the famous motto "from each according to his abilities and to each according to his needs" difficult for a central planner.

A second challenge faced by those seeking to create an economic system that maximizes material well-being is the *incentive challenge*: Even if we know what needs to be done to bring about maximal material well-being, how do we provide people with the necessary incentives to do these things? The severity of this challenge and how it is best overcome depends a great deal on human nature. If humans by their nature have a strong concern for the common good, it is not necessary to provide them with incentives. In this case, they will want to do whatever is required of them to serve the common good.

Social thinkers on the left of the political spectrum tend to be optimistic about human nature in this regard. A rich and highly influential tradition of social thought

on the left proposes systems for organizing society that rely heavily on the idea that human beings are inherently good: if society is organized in a just way, people will on the whole behave altruistically and cooperatively to attain the common good. Nineteenth century socialist and communist thinking strongly has this characteristic. John Lennon's famous song Imagine promotes this viewpoint in a particularly eloquent manner:

> Imagine no possessions
> I wonder if you can
> No need for greed or hunger
> A brotherhood of man
>
> Imagine all the people
> Sharing all the world
>
> You may say that I'm a dreamer
> But I'm not the only one
> I hope someday you'll join us
> And the world will live as one

Social thinkers on the right of the political spectrum tend to be less optimistic about human nature. They tend to view people as being largely motivated by self-interest and therefore emphasize the need to provide people with incentives that align their self-interest with the common good. Within economics, skepticism about the wisdom of relying heavily on altruistic behavior by individuals when choosing how to run the economy is widespread. Josef Stiglitz—hardly a right-wing zealot—captured well the conventional wisdom among economists when he wrote: "Self-interest is a more persistent characteristic of human nature than concern to do good" (Stiglitz, 2000, p. 56).

Economists have produced a great deal of empirical evidence supporting the notion that people pursue their self-interest. For example, there is pervasive evidence that people work harder when their effort is rewarded and also when their work is monitored. Conversely, workers that face weak incentives (either because their effort is not rewarded or because their work is not monitored) will often shirk and cut corners to make their work easier. Readers that have managed a home renovation will likely have strong first-hand experience with this. I as a teacher certainly have strong first-hand knowledge that students respond to incentives. All too many students only seem to care about what will be on the exam.

Not all selfish behavior is in conflict with the common good. For example, people may choose not to enter a particular profession because the pay is bad. This is a selfish motive. But if the pay is bad because the products produced by this profession have little value, it is efficient for few people to enter. In many cases, however, the conflict between selfish behavior and the common good is obvious. The example of a carpenter that doesn't properly seal a vent because he knows the owner won't notice until problems with water damage appear 15 years down the line is a simple example. The enormous efforts of US tobacco companies in the 20th century to undermine mounting scientific evidence that the products they sell cause cancer (and a variety of other health problems) is a more serious example. Sadly, many similar examples exist.

Karl Marx argued that capitalists are selfish (exploit the proletariat) and before them aristocratic landowners were selfish (exploited rural peasants). But he seems to have believed that the proletariat (the working class) are fundamentally different. He advocated a revolution aimed at reversing power relations and establishing a "dictatorship of the proletariat." In Marx' view, this would result in a just society without exploitation. Social thinkers on the right have always been skeptical of this notion, believing that power corrupts whether one is a capitalist or worker-turned-revolutionary-leader.

The history of communist revolutions in the 20th century certainly does not support Marx' rosy view of dictatorships of the proletariat, to say the least. Yet despite this, the viewpoint seems to persist on the left of the political spectrum that corporations and capitalists are evil, while ordinary people are virtuous. This viewpoint is at odds with a great deal of experimental evidence showing that a large fraction of people will act dishonestly (cheat others) if they are given "plausible deniability" or think they can get away with it. For example, Kleven et al. (2011) show that 45% of people that report self-employment income in Denmark underreport their self-employment income, i.e., cheat on their taxes. Cheating is not confined to a few bad apples. A large fraction of people cheat if they have the option to. Furthermore, people seem to be moral hypocrites, i.e., perceive themselves as acting more fairly than they actually do (Batson et al., 1999).

Darwin's theory of evolution by natural selection provides a theoretical basis for the idea that we humans are selfish. Our genes are selfish, as evolutionary biologist Richard Dawkins memorably put it (Dawkins, 1976). This means that we have evolved to promote our self-interest to the extent that this promotes our survival. The theory of evolution does not, however, rule out all altruism. To the contrary,

it helps make sense of the forms human altruistic behavior takes. Certain forms of altruistic and pro-social behavior can help promote the survival of our genes. It makes evolutionary sense that we would evolve mental mechanisms that promote these types of altruistic and social behavior.

The type of altruistic behavior that most obviously makes evolutionary sense is the love and care we show towards our offsprings and relatives. After all, our offsprings and relatives carry a substantial fraction of the same genes as we do. Their survival therefore promotes the survival of our (selfish) genes. Altruism towards non-kin is harder to explain from an evolutionary point of view. However, in a seminal paper, evolutionary biologist Robert Trivers argued that reciprocal altruism imparts a survival advantage if cooperation is important for survival (Trivers, 1971). Specifically, Trivers argued that humans have evolved moral emotions that lead us to play a "tit-for-tat" strategy in our interactions with others: We cooperate initially, but then only continue to cooperate with those that have reciprocated our cooperation, while we shun or punish those that have taken advantage of us in the past.

Playing tit-for-tat has been shown to perform extremely well in formal settings where cooperation is valuable. It is easy to demonstrate that tit-for-tat yields higher returns to the individual than both universal altruism ("help anyone in need") which invites exploitation, or completely opportunistic behavior ("always grab everything") which works well in the short run but not in the long run. What is more impressive is that playing tit-for-tat seems to yield high returns when it is played against a wide variety of other strategies. In a famous experiment, economist Robert Axelrod invited experts in game theory to submit strategies for a computer tournament in which the different strategies played an iterated prisoner's dilemma game against each other. Despite being the simplest strategy submitted, tit-for-tat won the tournament. More remarkably, tit-for-tat won a follow up tournament in which participants knew that tit-for-tat had won the first tournament (Axelrod, 1984).

In his excellent book *The Righteous Mind*, moral psychologist Jonathan Haidt extends this argument further: he argues that we humans have evolved to be tribal or groupish as well as selfish (Haidt, 2012, part III). While our selfish behavior promotes survival of the individual in competition with other individuals, groupish behavior promotes the survival of our group in competition with other groups. Haidt is evoking the idea of *group selection*, which is controversial within evolutionary biology. Darwin originally put forth the idea of group selection in a famous passage in *The Descent of Man*, where he argued that a tribe that "included a great number

of courageous, sympathetic and faithful members, who were always ready to warn each other of danger, to aid and defend each other" would be able to conquer other tribes made up of less morally upright individuals.

However, Darwin also realized the challenge that such group level adaptations face in that each individual member of the tribe has an incentive to cheat. Darwin said: "He who was ready to sacrifice his life, as many a savage has been, rather than betray his comrades, would often leave no offspring to inherit his noble nature." In other words, "noble" behavior may be good for the group, but each individual is more likely to survive if he or she free rides on everyone else's noble behavior. This *free rider problem* is the principle challenge group level adaptations must overcome. For groupish traits to survive, their benefit to the group in competition with other groups must outweigh their cost to the individual in competition within the group (more precisely, their cost to the individual's genes).

There are several ways this can work. First, if all individuals in the group are siblings—as is the case with bees in a hive—the cost to the individual's genes of the individual making a sacrifice for the good of the group is limited: the individual's genes live on in his or her siblings for which the sacrifice was made. A second way group selection can work is if the group evolves traits that suppress competition within the group: traits such as honor, loyalty, shame, and sanctity. If those members of the group that cheat face ostracism because they have broken a sacred social code (acted without honor and loyalty) and this means that they are more likely to be killed or shunned by potential mates, this will limit the benefits of cheating and thereby shift the balance of evolutionary forces in favor of traits that are good for the group (Haidt, 2012, ch. 9).

The ability of a species to evolve traits that allow large tribes or colonies to function as a cohesive unit can impart enormous survival benefits. However, making this evolutionary transition seems difficult judging by how rare such superorganisms are. Colonial insects are the most obvious examples. They make up only about 2% of insect species but account for the majority of insects on Earth by weight. Haidt argues that we humans are unique among mammals in the extent to which we have managed to evolve groupish traits. Crucial in this evolution was the emergence of human morality, parts of which Haidt argues are a group-related evolutionary adaptation aimed both at making us act under certain circumstances for the good of the group and also to make free riding within the group costly.

Among left-leaning, secular, western, educated elites there is a strong consensus that morality is about the well-being of individuals: it is wrong to harm other in-

dividuals or treat them unfairly. Haidt's empirical research (building on work by anthropologists Richard Shweder) has shown that left-leaning, secular, western, educated elites are actually outliers among humans in how narrowly they define issues of right and wrong. For most people, morality is more sociocentric: it is not only the individual that matters from a moral point of view but also the family, tribe, nation, caste, race, team, company, etc. Most people's moral matrix involves a strong sense of loyalty towards the groups that they belong to and it views actions of individuals that harm or betray these groups as morally wrong. Haidt argues that we are wired to view loyalty towards groups and betrayal of groups in moral terms and that this is an adaptation that evolved to help humans form cohesive coalitions—i.e., Darwin's noble tribes. He furthermore argues that "religion is (probably) an evolutionary adaptation for binding groups together and helping them to create communities with a shared morality."

If Haidt is correct on these points, it has bearing on how we can best solve the incentive challenge discussed above. While appealing solely to our selfish nature with simple "stick and carrot" type incentives may work well in certain circumstances, it may not be universally the best way to motivate people. In some circumstances, motivating people by appealing to our tribal nature or a common goal of a group is likely to be a powerful complement to individual rewards and punishment. One example is war. It is arguably foolish to try to fight a war without appealing to people's patriotism. But this applies much more broadly. Leaders of companies, sports teams, and all manner of other groups routinely motivate members of the group by creating team spirit and rallying team members around their common goals. Purely transactional leadership is arguably much less effective. Economists rarely incorporate the tribal nature of humans into their analysis. Keeping this in mind is important when one seeks to apply economic analysis to reality. In some circumstances, the emphasis of economic analysis on transactional incentives risks blinding us to more effective ways of motivating people.

Some of our evolved moral behavior has no doubt rubbed off from its original context to become generalized altruism. After all, most of us are influenced in many setting be the general notion of doing the right thing. However, the insights of moral psychology and evolutionary biology, suggest that generalized altruism is weak and fragile (only a byproduct of the forms of specific altruism discussed above). These insights therefore suggest that the amount of generalized altruism needed to support the successful functioning of a socialistic society (at least beyond a small commune or kibbutz) does not exist. Marx tried to appeal to worker' mem-

bership in the proletariat class (a group). This may have some promise, but is very ambitious. In Marx' view and that of many socialists, the proletariat class is a world-wide group (recall the International) that therefore cuts across other groups such as nations and religions that people care deeply about. Also, once the proletariat is in charge and the capitalist class has been vanquished, the proletariat would include virtually all of humanity. Research in anthropology and moral psychology suggests that people's altruism does not extend this far on a reliable basis.

Furthermore, an economic system with very weak individual incentives (such as socialism) is likely to run foul of people's strong desire for fairness and aversion to situations where cheaters are not punished. Anyone who has worked along side a co-worker who took advantage of lax oversight and low-powered incentives by performing little work knows how poisonous such behavior is for the morale of the rest of the workplace. Even if most workers are motivated to perform diligently the work they have been asked to do, it can be hard to maintain such a motivation when others shirk their duties. It is even more infuriating when the fact that others shirk leaves more work for those that are diligent. If those that work diligently see that the slackers are not punished, they will in all likelihood get disillusioned and reduce their own effort levels. This can lead to a vicious cycle of diminishing effort by all. Arguably, this vicious cycle is an important part of the reason why societies with weak individual incentives usually do not function well.

For the reasons given in the last two paragraphs, many economists believe that people on the far left of the political spectrum often implicitly base their ideas about how to organize society on too rosy a view of human nature. Incentives actually matter! On the other hand, as I have emphasized above, economists often ignore the forms of altruism that psychological research has found humans to be endowed with. Furthermore, people on the far right of the political spectrum arguably base their ideas about how to organize society on too rosy a view of how markets work. We will explore this idea in more detail later in the chapter.

2 A Surprising Idea

Simple "economies"—such as single households—often rely heavily on kin-based altruism and reciprocal altruism. Some families are good examples of socialism in action, as they manage to function well based primarily on kin-based altruism ('from each according to their ability, to each according to their need'). Other fam-

ilies rely more on tit-for-tat behavior. Tribes and villages also rely heavily on kin-based altruism and reciprocal altruism. But since these forces are weaker in larger settings, tribes and villages have traditionally relied on humans' innate group loyalty (tribalism) as well as another group level adaptation emphasized by Haidt (2012): respect for authority and respect for the existing hierarchical order.

Building on the work of anthropologist Alan Fiske, Haidt argues that humans are wired to respect authority and existing hierarchies because this trait helped create needed order within tribes. The survival benefit that such order carries with it is perhaps best illustrated by the fact that all modern military organizations rely extremely heavily on respect for authority and existing hierarchies. Respect for authority and existing hierarchies is pervasive in most societies today, and was even more so in the past. Examples include class and caste systems, respect for parents, husbands, teachers, religious authorities, and superiors in the workplace. Haidt's research demonstrates clearly that most humans—though not contemporary, left-leaning, secular, western, educated elites—view acts that disrespect existing hierarchies that they view as legitimate as morally wrong.

As we move to even larger and more complex economies, kin-based altruism, reciprocal altruism, and group loyalty become still weaker mechanisms to base the organization of economic activity on. One approach to managing large economies is to rely more and more heavily on respect for the existing hierarchical order as well as raw coercion. This is arguably how most large societies were organized throughout most of history (e.g., kingdoms and empires, both religious and secular). Some would argue that reciprocal altruism and loyalty to the large group may have played an important role in some cases even at these large levels (e.g., the idea that the feudal relationship between the king and his subjects may have involved significant reciprocity). But coercion and respect for the existing hierarchical order were surely crucial in most cases.

Organizing a large economy along these lines potentially has serious drawbacks. First, relying heavily on respect for the existing hierarchical order can make the society rigid, which may be very inefficient. At the macro level, changes in the environment or technology will call for changes to how the economy is best organized. At the micro level, people's skills may differ from their parents. If the hierarchical order dictates that people's place in society be the same as their parents' (as was common), much talent will then be misallocated. Another potentially serious drawback of relying heavily on respect for hierarchies is that this way of organizing society has no natural way to solve the information problem discussed above. Those at the top of

the hierarchy can impose their will on those below. But how do they know what to order their subordinates to do? Finally, how do such societies incentivize those at the top of the hierarchy to act for the common good as oppose to acting in their self interest?

Given these drawbacks of traditional social organization, it is natural to ask: Is there some other better way to organize a large economy? Perhaps the most deeply surprising and interesting theoretical idea of all of economics is that society can in fact attain a high level of material well-being without relying at all on altruism, group loyalty, respect for existing hierarchies, planning, or centralized control. This truly radical idea was first developed by a group of enlightenment thinkers—the foremost being Adam Smith. In *The Wealth of Nations*, Smith argues that an economic system where people are simply granted the freedom to pursue their self-interest—i.e., choose freely what to produce and consume—and allowed to engage in free exchange with each other to realize their choices will lead to higher levels of material well-being than economic management based on altruism, group loyalty, respect for authority, planning, or centralized control. As Smith famously put it: individuals intending only their own gain will in many cases be led by an "invisible hand" to promote the interests of society more effectively than when they intended to promote it (Smith, 1776/2000).

At first blush, this idea seems outlandish. How can everyone pursuing their self-interest possibly promote the common good? An important element of the answer lies in the fact that large numbers of people pursuing their self-interest create competition and competition restrains exploitation. This is (arguably) the most important insight of all of economics. When sellers face little or no competition, they can exploit their customers' lack of alternative options by charging high prices for the goods and services they sell, potentially much higher prices than their cost of producing the goods and services. Similarly, when buyers face little or no competition, they can exploit those trying to sell goods and services by offering only very low prices. An important example of the latter is when employers are able to hire workers at very low wages because the workers have no other options for employment. A central theme of economics is that competition limits the scope for such exploitation. Competition among sellers drives prices down towards production cost, thereby limiting the scope of each seller to exploit buyers. Competition among buyers drives prices (and wages) up towards the marginal value of the good or service being purchased, thereby limiting the scope of each buyer to exploit sellers.

Socialists and others on the left rightly emphasize the importance for social wel-

fare of limiting exploitation of the poor. The methods they advocate to achieve this aim vary from collective action (e.g., union power) and central planning to relying on people's inherent good nature. Although this is often not appreciated by those on the left, mainstream economics in the tradition of Adam Smith also has as its central aim to limit exploitation. The difference is just that economists in this tradition emphasize the power of competition to achieve this end.

It is instructive to consider how free market capitalism solves the information challenge and the incentive challenge that we discussed in the last section. First, by relying primarily on people pursuing their self-interest, free market capitalism overcomes the incentive challenge: people do not need to be incentivized to pursue their self-interest. Furthermore, as Hayek (1945) forcefully and eloquently argued, a critical advantage of well-functioning free market capitalism is that it produces a powerful mechanism for solving the information problem: the price system.

Prices aggregate an enormous amount of information in a market economy. We discussed above the challenge of allocating workers to tasks. Efficient allocation dictates that workers be allocated to the tasks for which they have a comparative advantage. But this depends on their skills relative to the skills of all others at all tasks. How could they ever know this? In a well-functioning market economy, this turns out to be quite simple: the task at which the person earns the most is the task for which they have a comparative advantage. Rather than knowing everyone's skill at all tasks, the person only needs to know the market wage for all tasks and their own ability at all tasks. The prices and wages in the economy encode the abilities of others (as well as demand for the different tasks). Magical, isn't it? Any economic system that does not make heavy use of prices to solve the information challenge is fighting an uphill battle.

Smith's arguments and those of his contemporaries were relatively informal and imprecise. These scholars did not spell out just how much could be attained by free market capitalism. In the 1950's, however, the economists Kenneth Arrow and Gerard Debreu were able to show that under certain conditions, free exchange will yield an economic outcome that is maximally efficient (i.e., as efficient as possible). The conditions needed are principally that people are able to choose what is best for them among the choices they face, that there exist competitive markets for all goods, and that property rights are well defined over all goods and perfectly and costlessly enforced. Given these conditions (which are not at all trivial), simply allowing the individuals of society to engage in free exchange brings about an outcome that cannot be improved upon — in a certain sense — even by an omniscient, benevolent

and all powerful central planner. Far from bringing about chaos, economic freedom accompanied by markets and property rights can bring about maximal economic well-being!

This is a startlingly strong result. And it is also one of the principle intellectual justifications used by proponents of free market capitalism. To gain a better grasp for this idea, it is helpful to write down a formal model of the economy. By doing this we can see what exactly I meant by "under certain conditions" and "in a certain sense" in the preceding paragraph. Once we have done this, we will be in a better position to think carefully about the limitations of these results and also be in a better position to compare them with other ideas about how to design a good society.

3 A Model Economy

We start by describing only the physical environment of our model economy. We make no assumptions about institutions that might exist such as markets, laws, courts, property rights, or firms. Instead, we first derive what outcome in this environment would be the best outcome in terms of maximizing efficiency. We then consider the question of what kinds of institutions can bring about such an ideal outcome.

Consider a very simple "desert island" economy where Robinson Crusoe and Friday have washed ashore. They start off with no other resources than their own time — which they can use for either labor or leisure — and the land of the island. In addition to these physical resources, however, Robinson Crusoe and Friday have access to knowledge of two technologies, one for harvesting coconuts and another for building shelter. This simple economy thus has two people (Robinson Crusoe and Friday), two inputs to production (land and labor), and three final goods (coconuts, shelter, and leisure). The points we will make in this chapter, however, hold true in a much more general setting. Those that are interested in seeing a more general exposition of these results can consult, e.g., Mas-Colell et al. (1995, chapters 15 and 16), Malinvaud (1985, chapters 4 and 5), or Kreps (2013, chapters 14 through 16).

Let's denote the endowments of land and time by ω_N and ω_T, respectively. The symbol ω is the Greek letter omega. We are using the letter N to represent land (the letter L will be used for labor) and the letter T to represent time. The endowment of land ω_N may for example be 100 acres and the endowment of time ω_T may be 16 hours per day (if we assume that Robinson Crusoe and Friday must sleep for 8

hours a day).

The technologies for producing coconuts and shelter can be represented by production functions. The production function for coconuts is

$$Y_C = f_C(L_C, N_C). \tag{1}$$

Here, Y_C denotes the quantity of coconuts produced, L_C denotes the amount of labor used to produce coconuts, and N_C denotes the amount of land used to produce coconuts. Analogously, the production function for shelter is

$$Y_S = f_S(L_S, N_S). \tag{2}$$

We assume that both of these production functions are increasing and concave in each input holding the other input constant. This implies that if more land, say, is devoted to the production of coconuts (holding the labor devoted to coconut production constant), this will increase the amount of coconuts produced, but at a diminishing rate. In addition, it is natural to assume that the production functions are homogeneous of degree one, i.e., that they are "constant returns to scale." This implies that if we double the amount of both labor and land devoted to the production of coconuts (say), the production of coconuts will double. Finally, we assume for simplicity that the labor of Robinson Crusoe and Friday can be viewed as interchangeable in production (i.e., they are equally productive and provide identical labor). This allows us to simply lump together the labor provided by the two of them in each activity (e.g., L_C is the sum of the hours worked by Robinson Crusoe and Friday in coconut production).

The preferences of Robinson Crusoe and Friday can be represented by utility functions. The utility function of Robinson Crusoe is

$$U_R(X_{C,R}, X_{S,R}, X_{Z,R}), \tag{3}$$

where $X_{C,R}$ denotes the consumption of coconuts by Robinson Crusoe, $X_{S,R}$ denotes the consumption of shelter by Robinson Crusoe and $X_{Z,R}$ denotes the "consumption" of leisure by Robinson Crusoe. Analogously, Friday's utility function is

$$U_F(X_{C,F}, X_{S,F}, X_{Z,F}). \tag{4}$$

We assume that both utility functions are increasing and concave in all three arguments. In other words, Robinson Crusoe and Friday like consuming each of the three goods but at a diminishing rate.

Notice that these utility functions do not allow for altruism, tribalism, respect for authority, or norms. Robinson Crusoe and Friday are purely selfish in our model. I make this assumption not because I think it is realistic, but rather because I want to make the point that "perfect" markets can achieve an efficient allocation of resources even in the most "pathological" case of a world in which everyone is purely selfish. However, it is also true, for better or worse, that most economic analysis makes this extreme assumption.

4 Pareto Efficiency

Next we need to introduce a few new concepts to be able to make precise what we mean when we say that an economic outcome is efficient. The first is the notion of an *allocation*. An allocation is a particular outcome for all the endogenous quantities in the economy. In the desert island economy we are considering, the endogenous quantities are: 1) the inputs used (L_C, L_S, N_C, N_S), 2) the outputs produced (Y_C, Y_S), and 3) the amounts consumed of each good by each of Robinson Crusoe and Friday ($X_{C,R}$, $X_{C,F}$, $X_{S,R}$, $X_{S,F}$, $X_{Z,R}$, $X_{Z,F}$). An allocation is a list of particular values for each of these variables.

Notice that the notion of an allocation does not make any reference to prices since we have not yet assumed that markets exist. The notion of an allocation simply states how each of the resources that is available in the economy is allocated, how much is produced of each good, and how much each of the agents in the economy gets to consume of each of these goods.

Next, we define a *feasible allocation* to be an allocation that satisfies all the constraints imposed by the physical environment. The first of these constraints is that the amount of land used for the production of coconuts and shelter cannot sum to more than the total amount of land available: $N_C + N_S \leq \omega_N$. In principle, some land could be left idle (i.e., the inequality could be strict). But in practice as long as either Robinson Crusoe or Friday is not satiated in both coconuts and shelter, they will want to use all the land available to them so as to increase the production of coconuts or shelter. (Make sure you understand why this is the case.) In practice, therefore, all land will be used and we will have $N_C + N_S = \omega_N$. The same argument can be made for all the other constraints discussed below. We will therefore present these constraints as equality constrains for simplicity.

The second constraint is that the amount of time allocated to production of co-

conuts, shelter, and to leisure by Robinson Crusoe and Friday cannot sum to more than the total amount of time available: $L_C + L_S + X_{Z,R} + X_{Z,F} = \omega_T$. The third constraint is that the consumption of coconuts by Robinson Crusoe and Friday cannot sum up to more than the production of coconuts: $X_{C,R} + X_{C,F} = Y_C$. Fourth, the consumption of shelter by Robinson Crusoe and Friday cannot sum up to more than the production of shelter: $X_{S,R} + X_{S,F} = Y_S$. Finally, the production functions $Y_C = f_C(L_C, N_C)$ and $Y_S = f_S(L_S, N_S)$ are constraints: output cannot exceed what the production function indicates for each good given the inputs supplied to producing that good. Allocations that satisfy these constraints are said to be feasible allocations.

Our ultimate goal is to figure out which of the feasible allocations are *efficient*. So, we need a definition of efficiency. How should we define what is efficient? In economics, as in lay English, the term "efficiency" refers to the absence of waste. If we use an unnecessarily large amount of inputs to produce some output, this is inefficient since effectively we are wasting some of the inputs. Another way in which we can be wasteful is if we devote too many resources to the production of one good (e.g., coconuts) relative to another good (e.g., shelter). Suppose the economy is already producing so many coconuts that the last few coconuts (the ones that result from the last hour of labor devoted to coconut production) have very low value to Robinson Crusoe and Friday relative to the additional shelter that could be produced with one extra hour of labor. In this case, it is wasteful to allocate so many resources to coconut production. Robinson Crusoe and Friday could be made better off by producing less coconuts and more shelter.

Since what we ultimately care about is the utility of the people in the economy, we will define efficiency in terms of whether we are squandering opportunities to make some person better off in terms of utility without making anyone else worse off in terms of utility. This way of defining efficiency goes back to the Italian economist Vilfredo Pareto. In homage to Pareto we say:

> An allocation is *Pareto efficient* if no one can be made better off without making someone else worse off.

More formally, we say that a feasible allocation A is Pareto efficient if no other feasible allocation B exists such that $U_{i,B} \geq U_{i,A}$ for all i and $U_{i,B} > U_{i,A}$ for at least one i, where i is an index for the different people in the economy — e.g., Robinson Crusoe and Friday in our desert island economy — and $U_{i,A}$ denotes the utility of person i under allocation A.

An allocation A is said to be a *Pareto improvement* relative to allocation B if someone is better off under allocation A relative to allocation B but no one is worse off. In this case, allocation B is said to be inefficient. Intuitively, allocation B is inefficient because it is a pure waste not to move to allocation A: by moving from B to A you get something for nothing (make someone better off without making anyone worse off). Certain economists are fond of saying that "there is no such thing as a free lunch." Well, moving from an inefficient allocation to an efficient allocation is tantamount to a free lunch. By any reasonable metric it should be done.

Let's consider an example. Suppose the government builds a bridge over a river and charges a toll for people to cross the bridge. Suppose that the toll revenue is sufficient to pay for the bridge. This is *potentially* a Pareto improvement. The people that cross the bridge are presumably choosing to use the bridge and pay the toll because this option is better than their best alternative option (the best option they had access to before the bridge was built). So, they are better off. But is anyone worse off? Well, it may be that the bridge diverts traffic from other roads and owners of restaurants, gas stations, etc. on those roads are made worse off. It may also be that the shadow from the bridge or the noise from it lowers property values in the surrounding area making the owners of those properties worse off. If this is the case, the construction of the bridge is not a Pareto improvement. However, if the revenue from the toll is high enough, it may *in principle* be possible to compensate all of those that would be made worse off. If this is done, the construction of the bridge is a Pareto improvement. In practice, such compensation schemes are difficult to administer (partly because it is difficult to estimate the damages the bridge causes) and are rarely undertaken.

When using the concept of Pareto efficiency, it is important to understand that it is a rather "weak" concept in terms of being able to rank allocations. In particular, there are typically many allocations that are Pareto efficient. For example, the allocation in which one of the people in the economy gets everything is Pareto efficient since all other allocations make him/her worse off. (Make sure you understand this point.)

A much "stronger" concept in terms of ranking allocations from best to worst is the notion of a *utilitarian social welfare function*. Consider the following weighted sum of Robinson Crusoe's utility and Friday's utility

$$\theta_R U_R(X_{C,R}, X_{S,R}, X_{Z,R}) + \theta_F U_F(X_{C,F}, X_{S,F}, X_{Z,F}). \tag{5}$$

This type of function is called a utilitarian social welfare function and θ_R and θ_F are

called the social welfare weights put on the utility of Robinson Crusoe and Friday. (θ is the Greek letter theta.) One (perhaps natural) example would be to set θ_R and θ_F equal to $1/2$. In this case, the social welfare function would put equal weight on Robinson Crusoe's utility and Friday's utility. This social welfare function can be used to rank all allocations by calculating the value that the social welfare function would take for each allocation. The higher the social welfare function, the better the allocation from this utilitarian social welfare perspective. The optimal allocation from this utilitarian social welfare perspective is then the allocation that delivers the largest value of the social welfare function, i.e., the allocation that maximizes the social welfare function.

A fundamental difference between the utilitarian social welfare function and Pareto efficiency as metrics to rank allocations is that taking resources from the rich and giving to the poor can in some cases raise social welfare according to a utilitarian social welfare function, while such a transfer is not a Pareto improvement. Why can taking money from the rich and giving to the poor raise social welfare? This is because the utility function of each person is concave. The rich therefore have lower marginal utility than the poor.

It is easiest to see why this implies that redistribution can raise social welfare by considering an example. Suppose Robinson Crusoe is much richer than Friday. This means that Robinson Crusoe has many more coconuts and much more shelter than Friday. In this case, it is likely that the last coconut Robinson Crusoe eats is not very valuable to him (doesn't yield much utility). He has already eaten a huge number of coconuts and is getting pretty full when he eats the last one. Friday, however, has many fewer coconuts and, therefore, values having one extra coconut more than Robinson Crusoe values the last coconut he is eating. In this case, taking one coconut away from Robinson Crusoe and giving it to Friday will raise the social welfare function since the loss to Robinson Crusoe will be smaller than the gain to Friday.

Pareto efficiency is a "weak" concept because it is silent as to whether it is "good" to perform such redistribution. Pareto efficiency focuses on making sure there is no pure waste, i.e., no unambiguously good way to make the economic pie bigger. An influential school of thought within economics holds that economists should focus on attaining Pareto efficiency and should avoid questions relating to whether redistribution is good or bad. According to this school of thought, redistribution raises thorny ethical issues that are outside of the scope of economists expertise. Economists that ascribe to this point of view are not necessarily saying that redistri-

bution is an unimportant issue. They are simply saying that pondering the welfare consequences of redistribution should be left to experts on that topic (ethicists).

Not all economists ascribe to this point of view. Plenty of economists analyze economic policies from the perspective of the utilitarian social welfare function (and other social welfare functions) and think carefully about the consequences of different policies for the level of inequality in society. Some are even harshly critical of those that think redistribution should be outside of the scope of economics. One reason for this is that in many cases there is a trade-off between equality and Pareto efficiency (since redistribution typically weakens economic incentives). A sharp focus on Pareto efficiency will, therefore, often yield conclusions that imply a great deal of inequality. Some economists believe that an exclusive focus on efficiency biases the policy recommendations of economists towards policies that yield a great deal of inequality. If both equality and efficiency are important social objectives, it is problematic if totally different groups of researchers are experts on these two objectives. In that case there will be no one who can think carefully about how to trade off these two objectives.

5 Conditions for Efficiency

Recall that our goal is to derive the best economic outcome possible in terms of efficiency. Armed with the new language we introduced in section 4, we can rephrase our goal as deriving the Pareto efficient allocation. In the desert island economy described above, there are two types of conditions that need to be met for an allocation to be Pareto efficient:

1. **Exchange Efficiency**: There cannot be any scope for mutually advantageous trade in final goods.

2. **Production Efficiency**: There cannot be any scope for Pareto improving reallocation of factors of production between their different uses.

The basic idea of a Pareto efficient allocation is that all resources are put to their best use. So, the thing we have to check is that there isn't some way to reallocate a resource in such a way as to make someone better off without making anyone else worse off. Exchange efficiency is about ruling out reallocation of final goods (coconuts and shelter) between Robinson Crusoe and Friday. Production efficiency

is about reallocation of factors of production (labor and land) between their different uses (production of coconuts, shelter, or leisure).

5.1 Exchange Efficiency

Let's start by deriving conditions for exchange efficiency. We do this using a variational argument—sometimes also called a perturbation argument. The basic idea behind this type of argument is the following thought process: Suppose we have an allocation that we think might be Pareto efficient. Let's call this candidate efficient allocation P (for Pareto efficient) and use a superscript $*$ to denote the particular value of various variables in this allocation. For example, let $X^*_{C,R}$ denote the consumption of coconuts by Robinson Crusoe in allocation P. Since we aren't sure that allocation P is Pareto efficient, we decide to compare allocation P to some other "nearby" allocation. This nearby allocation is called a "variation" relative to allocation P. If allocation P is in fact Pareto efficient, then moving to this nearby allocation cannot result in a Pareto improvement. If it did, this would contradict the notion that allocation P was Pareto efficient in the first place.

This logic is a very powerful way to derive conditions that must be true about the efficient allocation P. The trick is to find the right "nearby" allocation to compare allocation P with. Not all nearby allocations are going to yield interesting information. For example, consider taking 5 coconuts from Friday and give them to Robinson Crusoe. Can this be a Pareto improvement? No! Friday is clearly made worse off by this reallocation. The same is true of any other reallocation which takes a resource from either Friday or Robinson Crusoe and gives it to the other without some sort of "compensation." To learn something interesting, we need to contemplate slightly more complicated variations.

Let's consider a variation that takes ϵ (the Greek letter epsilon) units of shelter from Friday and gives these to Robinson Crusoe and "in exchange" takes $\eta\epsilon$ coconuts from Robinson Crusoe and gives them to Friday (η is the Greek letter eta). Here, η is the number of coconuts per unit of shelter in the "exchange," and ϵ is the size of the exchange. Can this type of variation yield a Pareto improvement? To answer this question, we need to figure out whether—for some ϵ and η—someone is made better of by this type of variation without anyone else being made worse off.

Before we proceed, we need to introduce one more piece of notation. Let's denote the marginal utility Robinson Crusoe gets from an extra coconut when he starts off

at the allocation P as

$$U^*_{R,C} \equiv \frac{\partial U_R(X^*_{C,R}, X^*_{S,R}, X^*_{Z,R})}{\partial X_{R,C}}. \tag{6}$$

It is important to understand what the different subscripts and superscripts in $U^*_{R,C}$ denote. The subscript R denotes that we are considering Robinson Crusoe's utility. The subscript C, however, denotes that we are taking a partial derivative of Robinson Crusoe's utility function with respect to the consumption of coconuts. Finally, the superscript $*$ denotes that we are taking this partial derivative at the point $(X^*_{C,R}, X^*_{S,R}, X^*_{Z,R})$, i.e., at the candidate efficient allocation P. We use analogous notation for other partial derivaties of Robinson Crusoe's and Friday's utility function. So, for example, $U^*_{F,S}$ denotes Friday's marginal utility of shelter at allocation P.

Since we are assuming that Robinson Crusoe's utility function is concave, his marginal utility of coconuts depends on how many coconuts he already has (and potentially also how much shelter and leisure). For this reason, it is important to keep in mind that we are interested in the marginal utility of coconuts to Robinson Crusoe at the allocation P. For example, if Robinson Crusoe has lots of coconuts under allocation P, his marginal utility from coconuts will be low. This same argument applies to all the other marginal utilities that we consider.

Given this notation, let's consider how the variation we describe above affects Robinson Crusoe's utility. His utility after the variation is

$$U_R(X^*_{C,R} - \eta\epsilon, X^*_{S,R} + \epsilon, X^*_{Z,R}).$$

Let's differentiate this expression with respect to ϵ (the "size" of the variation) and evaluate the resulting derivative at $\epsilon = 0$. Doing this, we get

$$\frac{dU_R}{d\epsilon} = -U^*_{R,C}\eta + U^*_{R,S}. \tag{7}$$

This allows us to see how Robinson Crusoe's utility changes as the size of the variation changes at the margin (i.e., for very small ϵ). For a variation with a positive ϵ, Robinson Crusoe loses coconuts, which lowers his utility. This is the first term of the derivative above (which has a negative sign). On the other hand, he gets shelter, which raises his utility. This is the second term of the derivative above (which has a positive sign). Which term is larger depends on the size of the two partial derivatives and the size of η.

Friday's utility after the variation is

$$U_R(X^*_{C,F} + \eta\epsilon, X^*_{S,F} - \epsilon, X^*_{Z,F}),$$

Differentiating this with respect to ϵ and evaluating the derivative at $\epsilon = 0$ yields

$$\frac{dU_F}{d\epsilon} = U^*_{F,C}\eta - U^*_{F,S}. \tag{8}$$

Looking at expressions (7) and (8), it is clear that if η is very large, Robinson Crusoe will be made worse off on the margin by the variation (i.e., by moving from zero to a small positive ϵ). He will lose a lot of coconuts but get only very few units of shelter. Similarly, if η is very small, Friday will be made worse off. To have any hope of finding a variation that yields a Pareto improvement, we must therefore choose η carefully.

One way to choose η that will turn out to be useful is to choose it in such a way that the variation leaves Robinson Crusoe's overall utility unchanged. This is the case if

$$-U^*_{R,C}\eta + U^*_{R,S} = 0.$$

Simple manipulation of this equation yields

$$\eta = \frac{U^*_{R,S}}{U^*_{R,C}}. \tag{9}$$

This equation describes how many coconuts Robinson Crusoe needs to get in exchange for a unit of shelter at the margin so as to keep his utility unchanged when he is at the candidate efficient allocation P.

Let's next calculate the change in Friday's utility for the variation that leaves Robinson Crusoe's utility unchanged. We do this by plugging the expression for η in equation (9) into expression (8). This yields

$$\frac{dU_F}{d\epsilon} = U^*_{F,C}\frac{U^*_{R,S}}{U^*_{R,C}} - U^*_{F,S} \tag{10}$$

Since the variation is chosen to keep Robinson Crusoe's utility unchanged, it yields a Pareto improvement if (and only if) it makes Friday better off. Equivalently, for the original allocation P to be Pareto efficient, this variation can't make Friday better off. This implies that the derivative in equation (10) must be zero. If the derivative takes any non-zero value, Friday is made better off either by a small positive variation or a small negative variation. So we have that

$$U^*_{F,C}\frac{U^*_{R,S}}{U^*_{R,C}} - U^*_{F,S} = 0.$$

which implies

$$\frac{U^*_{F,S}}{U^*_{F,C}} = \frac{U^*_{R,S}}{U^*_{R,C}}. \tag{11}$$

The reasoning we have just gone through implies that equation (11) is a necessary condition for allocation P to be a Pareto efficient allocation. This is the condition that guarantees exchange efficiency in our desert island economy.

Let's now try to understand the economics that is embedded in equation (11). To do this, let's first understand how to interpret the ratio on the left-hand-side of the equation. Recall that $U^*_{F,S}$ denotes Friday's marginal utility of shelter at the allocation P, while $U^*_{F,C}$ denotes his marginal utility of coconuts at this allocation. Since an extra coconut yields $U^*_{F,C}$ "utils" (i.e., units of utility) at the margin, $1/U^*_{F,C}$ coconuts yield one util at the margin and $U^*_{F,S}/U^*_{F,C}$ coconuts yield $U^*_{F,S}$ utils. This shows that $U^*_{F,S}/U^*_{F,C}$ is the number of coconuts needed to exactly compensate Friday for the loss of one unit of shelter at the margin. Or in other words, it is the value Friday places on an extra unit of shelter measured in coconuts. This ratio is often referred to as Friday's marginal rate of substitution between coconuts and shelter (but we will try to avoid such jargon). By an analogous argument, the ratio of the right-hand-side of equation (11) is the value Robinson Crusoe places on an extra unit of shelter measured in units of coconuts.

As we have discussed above, Friday's marginal utility of coconuts and shelter will depend on the number of coconuts and shelter that he has. If he has lots of coconuts and not so much shelter, $U^*_{F,C}$ will be small and $U^*_{F,S}$ large, implying that the ratio $U^*_{F,S}/U^*_{F,C}$ will be large. Moving from such an allocation to an allocation where Friday has less coconuts and more shelter will lead the ratio $U^*_{F,S}/U^*_{F,C}$ to fall since the marginal value of coconuts will rise and the marginal value of shelter will fall.

Equation (11) states that for P to be a Pareto efficient allocation, Robinson Crusoe and Friday must value an extra unit of shelter (measured in coconuts) equally. Suppose to the contrary that Friday values shelter more than Robinson Crusoe. For concreteness, suppose that Friday is willing to give up 10 coconuts to get one extra unit of shelter—i.e., $U^*_{F,S}/U^*_{F,C} = 10$—while Robinson Crusoe is only willing to give up five coconuts to get one extra unit of shelter—i.e., $U^*_{R,S}/U^*_{R,C} = 5$. In this case, a trade in which Friday gives Robinson Crusoe seven coconuts (any number between 5 and 10 would work) in exchange for one unit of shelter makes both Robinson Crusoe and Friday better off. Friday is better off since he values the extra unit of shelter he is getting more highly than the seven coconuts he is giving up, while Robinson

Crusoe is better off since he values the seven coconuts that he is getting more highly than the unit of shelter he is giving up. The fact that it is possible to construct a mutually advantageous trade in this example implies that the original allocation is not Pareto efficient. (The trade is a Pareto improvement.)

When equation (11) holds, Robinson Crusoe and Friday value shelter equally (measured in coconuts). It is only in this case that it is not possible to construct a mutually advantageous trade in shelter and coconuts between Robinson Crusoe and Friday. For this reason, equation (11) is a necessary condition for allocation P to be a Pareto efficient allocation.[1]

The discussion in the last few paragraphs has been a little bit imprecise in that I have talked about variations of finite size (e.g., 7 coconuts for a unit of shelter). In fact, Equation (11) holds at point P. But as we start to move away from point P by a finite amount, the derivatives in equation (11) start to change. I have ignored this detail for ease of exposition above. The idea is that 7 coconuts and a unit of shelter are small amounts, small enough to not materially change the derivative in (11). I will continue to use this rhetorical sleight of hand below.

5.2 Production Efficiency

Let's next consider production efficiency. There are several conditions that need to hold for production efficiency to be attained. We will derive a set of these conditions that if combined in various ways imply all these conditions. Let's start by considering how much Robinson Crusoe works gathering coconuts. The candidate efficient allocation P involves Robinson Crusoe working some number of hours gathering coconuts. What we are interested in figuring out is whether it is possible to engineer a Pareto improvement by making Robinson Crusoe work either more or less hours gathering coconuts. To this end, let's contemplate the idea that Robinson Crusoe works ϵ more hours gathering coconuts and that he gets to keep the extra coconuts he gathers.

What is the benefit Robinson Crusoe derives from this variation? This depends on how many coconuts Robinson Crusoe can gather per hour at the margin, i.e., the marginal product of labor in coconut production. Let's denote the marginal product of labor in coconut production when we start off at the allocation P by

$$F^*_{C,L} = \frac{\partial F_C(L^*_C, N^*_C)}{\partial L_C}. \tag{12}$$

As before, it is very important to understand what the subscripts and superscripts in

this expression denote. The subscript C denotes that we are considering the production function for coconuts. The subscript L, however, denotes that we are taking a partial derivative of F_C with respect to labor. Finally, the superscript $*$ denotes that we are evaluating this partial derivate at the point (L_C^*, N_C^*), i.e., at the candidate efficient allocation P. As we discussed above, we are also implicitly assuming for simplicity that Robinson Crusoe and Friday are equally productive at gathering coconuts (otherwise the partial derivative would depend on which of them is putting in the extra effort).

The marginal product of labor in coconut production will depend on the amount of labor that is already allocated to coconut production as well as the amount of land allocated to coconut production. If Robinson Crusoe and Friday are between themselves spending a large amount of time gathering coconuts on a small amount of land, it is likely that additional time spent gathering coconuts will not yield many coconuts. In this case, $F_{C,L}^*$ will be small. On the other hand, if they are spending very little time gathering coconuts from a large field, an extra hour spent gathering is likely to yield a large number of coconuts. In this case, $F_{C,L}^*$ will be large.

We can now derive an expression for the marginal benefit Robinson Crusoe derives from working more gathering coconuts. Measured in terms of extra coconuts, this is equal to the number of coconuts he gathers per extra hour ($F_{C,L}^*$). To measure the marginal benefit in terms of utility, we must multiply by the marginal utility of coconuts ($U_{R,C}^*$). This yields a marginal benefit of $U_{R,C}^* F_{C,L}^*$ measured in utils per hour.

How about the marginal cost of working extra hours? The cost to Robinson Crusoe is that he must give up time that otherwise could be used for leisure. Robinson Crusoe's marginal utility of leisure is given by $U_{R,Z}^*$. His marginal cost of working extra hours is therefore equal to $U_{R,Z}^*$ utils per hour.

For the initial allocation P to be Pareto efficient, working more (or less) can't make Robinson Crusoe better off. This is the case when the marginal benefit of working more is equal to the marginal cost of working more:

$$U_{R,C}^* F_{C,L}^* = U_{R,Z}^*. \tag{13}$$

Suppose to the contrary that the marginal benefit of working more was larger than the marginal cost. If this were true, it would be a Pareto improvement for Robinson Crusoe to work more. He would be better off and no one would be worse off. Similarly, if the marginal benefit to Robinson Crusoe of working more was smaller than

the marginal cost, it would be a Pareto improvement for Robinson Crusoe to work less.

Simple manipulation of equation (13) yields

$$\frac{U^*_{R,Z}}{U^*_{R,C}} = F^*_{C,L}. \tag{14}$$

Written this way, the condition states that the marginal cost to Robinson Crusoe of working more measured in units of coconuts per hour (left-hand side) must equal the marginal benefit of working more measured in units of coconuts per hour (right-hand side). The only difference between equations (13) and (14) is that the former measures marginal benefit and marginal cost in units of utils per hour, while the latter measures them in units of coconuts per hour. These conditions are necessary conditions for production efficiency.

The same line of argument implies several additional necessary conditions for production efficiency. First, the marginal benefit of Friday working more on gathering coconuts must be equal to Friday's marginal cost of doing this:

$$\frac{U^*_{F,Z}}{U^*_{F,C}} = F^*_{C,L}. \tag{15}$$

Second, the marginal benefit of Robinson Crusoe working more on building shelter must be equal to Robinson Crusoe's marginal cost of doing this:

$$\frac{U^*_{R,Z}}{U^*_{R,S}} = F^*_{S,L}. \tag{16}$$

Third, the marginal benefit of Friday working more on building shelter must be equal to Friday's marginal cost of doing this:

$$\frac{U^*_{F,Z}}{U^*_{F,S}} = F^*_{S,L}. \tag{17}$$

Equations (14)-(17) can furthermore be combined to yield various other implications of production efficiency such as $U^*_{R,C}F^*_{C,L} = U^*_{R,S}F^*_{S,L}$, which says that extra time gathering coconuts should yield the same amount of utility as extra time building shelter at the margin. (Otherwise time should be reallocated to the one of these activities that yields more utility at the margin.)

But production efficiency doesn't only imply that labor should be used efficiently. It also implies that land should be used efficiently. For simplicity, we have

assumed that Robinson Crusoe and Friday do not derive utility directly from unused land (they don't enjoy lying on the beach, etc.). This implies that it is efficient to make use of all available land for coconut and shelter production. The only question that remains regarding the efficient use of land is how to divide up the land between coconut and shelter production.

The candidate efficient allocation P implies a particular division of the land between coconut and shelter production. Let's consider a variation where Robinson Crusoe transfers ϵ acres of land from coconut production to shelter production. The marginal benefit of this variation is the marginal product of land in shelter production ($F^*_{S,N}$) times the marginal utility of shelter for Robinson Crusoe ($U^*_{R,S}$), i.e., $U^*_{R,S}F^*_{S,N}$ utils per acre. The marginal cost of this variation is the marginal product of land in coconut production ($F^*_{C,N}$) times the marginal utility of coconuts for Robinson Crusoe ($U^*_{R,C}$), i.e., $U^*_{R,S}F^*_{S,N}$ utils per acre. Efficiency requires that the marginal benefit from this type of variation must equal the marginal cost:

$$U^*_{R,S}F^*_{S,N} = U^*_{R,C}F^*_{C,N}. \tag{18}$$

Manipulating this equation yields

$$\frac{U^*_{R,S}}{U^*_{R,C}} = \frac{F^*_{C,N}}{F^*_{S,N}}. \tag{19}$$

We have now derived all the conditions that need to hold for allocation P to be Pareto efficient. As we mentioned above, the conditions we have derived imply many additional conditions, i.e., you can combine the conditions we have derived in various ways to get other conditions. But any allocation that satisfies all the conditions we have derived above for our desert island economy will be Pareto efficient. In fact, given the assumptions we have made about the production functions and utility functions, there is a unique allocation that satisfies these conditions. This is shown formally, for example in Mas-Colell et al. (1995, chapters 15 and 16).

6 Bringing About Efficiency

Now that we have described what must be true for an allocation to be Pareto efficient, the next natural question is: How can we bring about this outcome? One theoretically possible way to do this—although perhaps not particularly plausible—is through central planning. Suppose there exists a benevolent, omniscient, and omnipotent central authority in the economy. Since the central authority is omnipotent,

it can bring about any outcome it wants. Since it is omniscient, it can figure out which allocations are Pareto efficient and which are not. And since it is benevolent, it will want to bring about a Pareto efficient allocation.

It is worth stressing how unrealistic it is to suppose that central planning will in fact yield anything close to a Pareto efficient allocation in reality. First, for this to occur, the central authority must know "everything." It must know everyone's preferences in full detail (who likes peanut butter, who likes to watch sports on TV, who likes to lie on the beach, etc.). It must also have full knowledge of the production process of all goods (What is the marginal product of labor in coconut production when coconuts are being grown on 40 acres with 70 hours of labor?, etc.). Second, it must be arbitrarily powerful since it needs to be able to get everyone to do what it says they should. In particular, it must be able to motivate people to work as much and as hard as it wants them to. And it must do this without wasting resources on monitoring them at work. Clearly, this is a tall order. Third, it must be benevolent. Even a casual reading of history strongly suggests that central authorities that are both knowledgeable and powerful are often far from fully benevolent.

So, is there an alternative? Is there some other set of institutions that might bring about a Pareto efficient allocation? Surprisingly, there actually is! As we discussed earlier in this chapter, a crown jewel of economic theory is the idea that well-defined property rights that are perfectly enforced coupled with a complete set of *competitive* markets will, in fact, bring about a Pareto efficient allocation.

By a *competitive* market, we mean a market where no person trading in the market believes that they can affect the price in that market. Intuitively, if each person trading in a market is small relative to the overall volume in the market, no single person can affect the price by buying or selling more or less. Each person is a "drop in the bucket" and has a negligible effect on the price. If all people trading in the market believe that they are small in this sense, they will all take the price as given (i.e., assume that they can't affect it). If this is the case, the market is said to be competitive.

To show that a complete set of competitive markets can bring about a Pareto efficient allocation, let's start by supposing that there exists a competitive *goods market* in our desert island economy where coconuts can be exchanged for shelter. This means that there is a price p for shelter in terms of coconuts (i.e., one unit of shelter costs p coconuts). Both Robinson Crusoe and Friday are able to buy or sell shelter in this market and both take the price in the market as given and unaffected by their own trading.

What will the existence of this market imply about the amount of coconuts and shelter Robinson Crusoe and Friday will end up with (after trading)? Let's start by thinking about what Robinson Crusoe will do in this market. Recall that $U_{R,S}/U_{R,C}$ is the marginal value of shelter to Robinson Crusoe measured in coconuts. The marginal cost of acquiring additional shelter on the market is simply the price p. Suppose the marginal value of shelter to Robinson Crusoe is higher than the marginal cost of shelter on the market—i.e., $U_{R,S}/U_{R,C} > p$. In this case, Robinson Crusoe can increase his utility by purchasing shelter on the market in exchange for coconuts. Robinson Crusoe will, therefore, purchase shelter on the market (in exchange for coconuts) until the marginal value of shelter to him has fallen enough that it equals the price of shelter in the market. On the other hand, if the marginal value of shelter to Robinson Crusoe is lower than the marginal cost of shelter on the market—i.e., if $U_{R,S}/U_{R,C} < p$—he will sell shelter on the market in exchange for coconuts. He will keep selling shelter until the marginal value of shelter has risen enough that it equals the price of shelter in the market.

Another way of describing what Robinson Crusoe does in the market is that he buys or sells shelter so as to equalize the marginal value and marginal cost of shelter. As usual, setting marginal value equal to marginal cost maximizes utility. The upshot of this is that irrespective of how much shelter and how many coconuts Robinson Crusoe starts off with, he will trade coconuts for shelter until

$$\frac{U_{R,S}}{U_{R,C}} = p. \tag{20}$$

The exact same logic holds for Friday. He will also trade on the market so as to equalize the marginal value and marginal cost of shelter in terms of coconuts. This implies that irrespective of how much shelter and how many coconuts Friday starts off with, he will trade coconuts for shelter until

$$\frac{U_{F,S}}{U_{F,C}} = p. \tag{21}$$

Now notice that equations (20) and (21) can be combined to yield

$$\frac{U_{R,S}}{U_{R,C}} = p = \frac{U_{F,S}}{U_{F,C}} \tag{22}$$

or simply

$$\frac{U_{R,S}}{U_{R,C}} = \frac{U_{F,S}}{U_{F,C}}. \tag{23}$$

Recall that this is the condition for exchange efficiency—see equation (11). We have therefore shown that the existence of a competitive goods market implies that the resulting allocation will satisfy exchange efficiency.[2]

How is the price in the goods market determined? Robinson Crusoe and Friday can trade to equalize their marginal value of shelter measured in coconuts to any value. However, if the price of shelter is "too high," they will both want to sell shelter. But then who would buy? In this case, we say that the market does not clear. The same thing will happen for a very low price: both Robinson Crusoe and Friday will want to buy shelter at a sufficiently low price. But who would sell? Again, the market does not clear. Under the assumptions that we have made earlier in this chapter, there is actually a unique price that clears the market, i.e., implies that demand for and supply of shelter on the market are equal. But how does the market find this price? This is a bit mysterious and is seldom modeled explicitly. The basic idea is that if demand is larger than supply the price in the market will be bid up, while if supply is larger than demand the price in the market will be bid down. This process will continue until the market clears. The idea that this type of process does in fact clear the market continuously is part of what we mean when we say a market is perfectly competitive.

Next, suppose there exists a competitive *labor market* in our desert island economy, where workers are hired and paid in coconuts. Let's denote the wage rate by w (i.e., producers pay workers w coconuts for each hour of labor). What will Robinson Crusoe do in this market? He will supply labor up until the point where the marginal cost of supplying more labor is equal to the marginal benefit. The marginal cost of supplying labor is foregone leisure. Measured in coconuts, the marginal cost of supplying labor for Robinson Crusoe is, therefore, $U_{R,Z}/U_{R,C}$. The marginal value of supplying labor is simply the wage w. Robinson Crusoe will, therefore, supply labor up until the point where

$$\frac{U_{R,Z}}{U_{R,C}} = w. \tag{24}$$

Again, the exact same logic holds for Friday. He will also set the marginal cost of supplying labor equal to the marginal benefit. In his case, this implies that he will supply labor up until the point where

$$\frac{U_{F,Z}}{U_{F,C}} = w. \tag{25}$$

The "buyers" in the labor market—i.e., those demanding labor—will be the coconut and shelter producers. They will demand labor up until the point where the

marginal product of labor is equal to its marginal cost. Consider first the coconut producers. The marginal product of labor in coconut production is $F_{C,L}$ coconuts per hour of labor. The marginal cost is the wage w. Coconut producers will, therefore, demand labor up until the point where

$$F_{C,L} = w. \tag{26}$$

The marginal product of labor in shelter production is $F_{S,L}$ units of shelter per hour of labor. Since we are quoting the wage in terms of coconuts, we need to convert the marginal product of labor in shelter production into its value in terms of coconuts. We do this by asking how many coconuts we would get if we sold the shelter produced by the extra hour of labor in the goods market. We would get p coconuts per unit of shelter, implying that the marginal product of labor in shelter production is $pF_{S,L}$ coconuts per hour. The marginal cost is again the wage w. Producers of shelter will, therefore, produce shelter up to the point where the marginal product of labor in shelter production (measured in terms of coconuts) is equal to the wage:

$$pF_{S,L} = w. \tag{27}$$

Now notice that combining equations (24) and (26) yields

$$\frac{U_{R,Z}}{U_{R,C}} = w = F_{C,L}, \tag{28}$$

which is one of our conditions for productions efficiency—equation (14)—if we eliminate the w in the middle. Notice also that combining equations (25) and (26) yields

$$\frac{U_{F,Z}}{U_{F,C}} = F_{C,L}, \tag{29}$$

which is another one of our conditions for productions efficiency—equation (15).

Next, combine equations (24) and (27) to get

$$\frac{U_{R,Z}}{U_{R,C}} = pF_{S,L}. \tag{30}$$

But we also know from our analysis of the goods market that the price p is equal to the marginal value of shelter in terms of coconuts—equation (20). Using this fact to eliminate the price p from the above equation yields

$$\frac{U_{R,Z}}{U_{R,C}} = \frac{U_{R,S}}{U_{R,C}} F_{S,L}. \tag{31}$$

This equation can be simplified to

$$\frac{U_{R,Z}}{U_{R,S}} = F_{S,L}, \tag{32}$$

which is yet another one of our conditions for production efficiency—equation (16). Finally, combining equations (21), (25), and (27) yields

$$\frac{U_{F,Z}}{U_{F,S}} = F_{S,L}, \tag{33}$$

which is our fourth and final condition for production efficiency that relates to the allocation of labor—equation (17). We have therefore shown that the existence of a competitive labor and goods markets implies that labor will be efficiently allocated.

Next, suppose that there is a competitive *rental market for land*, where an acre of land can be rented for r coconuts (i.e., r denotes the rental price of land). Consider first what the producers of coconuts will do in this market. They will rent land until the marginal product of land in coconut production $F_{C,N}$ is equal to the marginal cost of land, which is the rental price r. Demand for land by the producers of coconuts will therefore imply that

$$F_{C,N} = r. \tag{34}$$

Analogously, the producers of shelter will rent land until the marginal product of land in shelter production is equal to its marginal cost. The only tricky thing here is that we must calculate the marginal product and the marginal cost in the same units (e.g., coconuts). The marginal product of land in shelter production is $F_{S,N}$ units of shelter. We can convert this into coconuts by multiplying by p (the price of shelter in terms of coconuts). The marginal product of land in shelter production measured in coconuts is therefore $pF_{S,N}$. Shelter producers will therefore demand land up to the point where

$$pF_{S,N} = r. \tag{35}$$

Combining these last two equations yields

$$p = \frac{F_{C,N}}{F_{S,N}}. \tag{36}$$

Using the fact that $p = U_{R,S}/U_{R,C}$—equation (20)—yields

$$\frac{U_{R,S}}{U_{R,C}} = \frac{F_{C,N}}{F_{S,N}}. \tag{37}$$

Notice that this is the final condition for production efficiency we derived in section 5.

Together the derivations we have done in this section show that the existence of a perfectly competitive goods market, labor market, and rental market for land coupled with well-defined property rights that are perfectly and costlessly enforced imply that the equilibrium allocation will be Pareto efficient.

7 The Welfare Theorems

The result we derived in section 6 is a specific example of a much more general theorem called the *First Welfare Theorem*. Let me now state this theorem more generally:

First Welfare Theorem. *Suppose the following three conditions hold:*

1. ***Rationality**: Everyone in the economy is able to choose what is best for them from among the set of options they face.*
2. ***Competitive Markets**: There exist competitive markets for all goods and services in the economy (both inputs and outputs).*
3. ***Private Property**: Property rights over all goods and services are well defined and costlessly enforceable.*

Then the equilibrium outcome in the economy will be Pareto efficient.

In the simple desert island economy we considered in sections 3 through 6, competitive markets for all goods and services means a competitive goods market (shelter for coconuts), a competitive labor market (labor for coconuts), and a competitive land market (land for coconuts). Furthermore, property rights over all goods and services means that it is clearly defined (and costless to enforce) who owns each parcel of land, Robinson Crusoe's time, Friday's time, the coconuts and shelter that are produced, and any organization created to produce coconuts or shelter (e.g., any firms). Finally, rationality means that Robinson Crusoe and Friday are able to pick what is best for them from among the choices they face.

In the desert island economy, there are only two output goods and two factors of production. But the First Welfare Theorem holds in a situation with any number of goods and services and any number of factors of production as long as there are competitive markets for each and every one of these. In the desert island economy, we did not explicitly talk about how time unfolds. In fact, The First Welfare Theorem holds in an economy that unfolds over time (i.e., an economy with multiple

time periods), but with the important caveat that either all people must exist at all times or all people must belong to families in which parents are perfectly altruistic towards their children (i.e., treat their children's welfare in the future as though it was their own). Finally, in our desert island economy there was no uncertainty or risk. But the First Welfare Theorem holds in a world with uncertainty and risk. Again, there is an important caveat: information must be "symmetric," i.e., things that are known to anyone must be known to all (people can't have private information). Furthermore, markets must exist for all goods and services in each future "state of the world" (e.g., a market for an umbrella for delivery tomorrow *if it rains*). We will discuss the various situations in which the First Welfare Theorem does not hold in greater detail in section 9.

As I have said before, the First Welfare Theorem is one of the crown jewels of theoretical economics. It is a formalization of a remarkable insight that goes back to Adam Smith. In the words of Milton Friedman and Rose Friedman (1980):

> Adam Smith's flash of genius was his recognition that the prices that emerge from voluntary transactions between buyers and sellers—for short, in a free market—could coordinate the activity of millions of people, each seeking his [or her] own interest, in such a way as to make everyone better off. It was a startling idea then, and it remains one today, that economic order can emerge as the unintended consequence of the actions of many people, each seeking his [or her] own interest. (pages 13-14)

The First Welfare Theorem tells us that competitive markets yield a Pareto efficient outcome. However, earlier in the chapter we discussed the fact that there are many Pareto efficient allocations. For example, the allocation where one person gets everything is Pareto efficient. If market forces can only guarantee that the outcome will be Pareto efficient, perhaps that is not enough to guarantee that the outcome is "good." After all, we may feel that some Pareto efficient allocations are quite unappealing. It is therefore natural to ask: Is it possible to steer the market equilibrium towards a particular Pareto efficient allocation that we find desirable?

The answer to this question is called the *Second Welfare Theorem*. Let me first state the theorem and then explain what it means:

Second Welfare Theorem. *Suppose the conditions for the First Welfare Theorem hold. Suppose furthermore that a few (relatively weak) additional conditions hold.*[3] *Then any Pareto efficient allocation can be attained for some initial distribution of wealth.*

To explain the power of this theorem, let's suppose that a government exists that can alter the initial distribution of wealth. In other words, the government can redistribute wealth once at the economy's "starting point." The Second Welfare Theorem states that if the conditions of the First Welfare Theorem hold, such a government can achieve *any* Pareto efficient allocation.

Notice that no matter which (utilitarian) social welfare function you prefer, the best outcome from the perspective of this social welfare function will be a Pareto efficient outcome. This follows from the simple fact that it never makes sense from a utilitarian point of view to waste resources. The Second Welfare Theorem therefore implies that the best outcome for any utilitarian social welfare function can be obtained with a government that only does the following things: enforces property rights and contracts, maintains any institutions needed for markets to be competitive, and performs a single initial redistribution of wealth. That is it. No other government intervention is needed. In this sense, the First and Second Welfare Theorems provide an intellectual justification for *laissez faire* government policy.

Several remarks are in order, lest readers get too excited about this result. First, the difficulty of creating a government that can carry out even the "minimal functions" entailed by laissez faire policy—namely low cost enforcement of property rights and contracts and the creation and maintenance of institutions needed for markets to be competitive—should not be underestimated. Arguably, creating conditions for widespread competitive markets involves building very strong institutions of various kinds, something most governments don't even come close to being able to do. The idea that the government can just recede into the background and "let the market take care of things" is a gross misconception of what it takes to create a competitive economy. We will discuss this in more detail in the next section.

Second, even a cursory glance at the world reveals that the conditions needed for the First Welfare Theorem to hold are nowhere near being met in practice. Most markets are far from being perfectly competitive; markets for many goods and services simply don't exist; and informational asymmetries are widespread; just to name a few important "market failures." Given this, it is reasonable to ask whether the First and Second Welfare Theorems are of any practical relevance. Why do we care that *in principle* markets can yield a Pareto efficient allocation?

Views on this question vary widely. Let me describe a few perspectives. First, there are those that think the First and Second Welfare Theorems are totally uninteresting from a practical point of view. Since the conditions they rely on are obviously not satisfied, we cannot learn anything from them about whether markets

improve welfare or laissez faire policy is good policy. Some go so far as to argue that these theorems and the allure of their mathematical elegance and rigor has led the economics profession to put mathematical rigor ahead of real-world relevance. Furthermore, it is sometimes argued that the attention paid to these theorems has created an unconscious bias among economists in the direction of believing that the theorems actually hold in practice: economists think so much about models in which these theorems hold that they come to believe that the theorems actually describe the world to a greater extent than is justified. For example, Paul Krugman argued in the wake of the financial crisis of 2008 that "the economics profession went astray because economists, as a group, mistook beauty, clad in impressive-looking mathematics, for truth" (Krugman, 2009).

A different perspective is that the First and Second Welfare Theorems demonstrate that there is nothing inherently superior about some other way of organizing society, such as socialism. In other words, free market capitalism cannot be handicapped on purely theoretical grounds. Given this, what is really relevant is whether markets in practice work better than other ways of organizing society. Do the theorems help us think about this? Some economists argue that they help us organize our thoughts about "market failures," i.e., the types of circumstances that lead markets to function poorly. In this view, the theorems are a helpful benchmark and a helpful goal. A common way of thinking about a practical situation for an economist is: What is the deviation from the First Welfare Theorem in this situation? Is it monopoly power, an externality, a missing market, moral hazard, or something else? Once that has been identified clearly, this can help focus our thinking about how to improve the situation.

A common way for economists to think about the role of government in the economy is that its role is to fix market failures. According to this view, when one is contemplating whether the government should intervene somewhere in the economy one should ask two questions: First, what is the market failure—deviation from the First Welfare Theorem—that needs to be fixed? Second, does the government have an ability to fix this market failure? This can be a very powerful framework for thinking about the role of government in the economy. It is important to remember, though, that this approach is only useful in thinking about ways in which the government may improve efficiency. It may also be desirable for the government to intervene in the economy on grounds of equity.

Arguments about whether the government should do some policy X between an enthusiast of free markets and someone skeptical of free markets often take the fol-

lowing form. The free market enthusiast will say: 'let the market take care of this,' or 'let people solve this problem using voluntary agreements.' This can be a powerful rhetorical position. The skeptic will however often argue that 'markets don't work that well,' and the enthusiast will respond 'yes they do.' This will often lead to an impasse with neither side changing their view. (And perhaps one or both sides loosing their temper.) Arguably, a more productive way to debate this question is for the skeptic to try to identify the market failure that leads markets not to function well in the situation at hand and to argue that policy X will solve this market failure and that the government can actually carry out policy X. This will focus the debate. The market enthusiast might propose a market based solution to the market failure in question or argue that the government is actually not able to carry out policy X in a way that improves the situation. The market enthusiast and market skeptic may not reach an agreement, but at least they may be able to identify more specifically what they disagree about and perhaps try to find empirical evidence that informs those issues.

Yet another (related) perspective on the usefulness of the First Welfare Theorem is that it is useful because it helps identify inefficiencies. Inefficiencies arise when one or more of the conditions needed for the First Welfare Theorem do not hold. Let's consider one of the labor market conditions in section 6, the one for coconut production. This condition says that the marginal product of labor in the production of coconuts ($F_{C,L}$) should equal the marginal value of leisure measured in coconuts ($U_{R,Z}/U_{R,C}$). In economics, a difference between these two (if they are not equal) is often referred to as a "wedge." Suppose for example that the marginal value of leisure measured in coconuts is lower than the marginal product of labor in coconut production. Then there is some $\tau_C > 1$ which yields

$$\tau_C \frac{U_{F,Z}}{U_{F,C}} = F_{C,L}. \tag{38}$$

Here, τ_C is the wedge and the degree to which τ_C differs from one is a measure of how large the deviation from efficiency is in the labor market. If we can measure the size of the wedges, this may help us figure out in which markets inefficiencies are greatest and where to direct our attention in making the economy more efficient.

One way to think about the goal of economic policy is that it is to reduce wedges. Intuitively, the idea is that reducing wedges brings us closer to the "first best." If all the wedges could be eliminated, the First Welfare Theorem would hold and the economy would be efficient (the first best). There is however an important caveat to

the "eliminate wedges" perspective on economic policy. It turns out that if there are wedges (i.e., inefficiencies) in more than one market, it is not necessarily the case that eliminating a particular wedge will make the economy more efficient overall. This result is called the General Theory of Second Best and was established by Lipsey and Lancaster (1956).

The classic example of a situation where eliminating a wedge makes the overall economy less efficient is the case of a customs union. Consider three countries that all levy tariffs on imports of a particular product. These tariffs result in wedges. Suppose now that two of the countries form a customs union and eliminate tariffs on imports between each other (but not the third country). This will shift demand away from the third country and increase trade within the customs union. But if the third country is the lowest cost producer of the good, this will actually lead to less efficient production of the good.

While the General Theory of the Second Best is certainly something to be aware of when thinking about economic policy through the "reduce wedges" lens. It is not clear how often this logic will lead to such perverse results in practice. A few famous examples aside, this logic does not generate perverse results very often in applied theoretical work to my knowledge. And practically speaking, even in the case of regional trade deals, views vary considerably on its importance.

Finally, it is important to realize that government policies that reduce wedges will typically not result in a Pareto improvement. Take for example the case of a market that is monopolized by a single firm. This firm will set the price above marginal cost implying that there will be a wedge in the market for that good. Suppose now that the government manages to improve competition in this market. This will lead the price in the market to fall toward marginal cost. As a result, the wedge in this market will fall. This policy is not a Pareto improvement: the owners of the firm are made worse off. Before the policy they were exploiting their monopoly power. The policy reduced their ability to do this. The policy therefore made consumers better off partly at the expense of the prior monopoly firm. The policy improved welfare in the sense that the sum of consumer and producer surplus rose. This means that it is in principle possible to compensate the owners of the firm for their losses. But real world policy rarely involves such compensation schemes and it is not clear that compensating a monopoly for lost profits makes sense from an ethical point of view.

This example highlights the fact that wedges often result from exploitation. In the example above the firm was using its market power to exploit consumers. But other market failures can also be described as exploitation. For example, a polluter

is exploiting those that are harmed by the pollution; an agent exploits his or her principle by not putting in much efforts; etc. Viewing things this way, reducing wedges involves reducing exploitation in the economy. This is clearly rarely Pareto improving (the exploiters lose), but nevertheless makes sense as a goal for economic policy.

8 Coase and the Economics of Transaction Costs

We have emphasized how competitive markets can direct resources to their most efficient uses. This result is often used to argue that economic planning is not necessary or desirable: the price system is a powerful coordinating mechanism that results in economic order without any need for centralized control.

But is this really so? In a seminal article published in 1937, Ronald H. Coase drew attention to the role of firms in real world capitalist economies (Coase, 1937). Most production in capitalist economies takes place within firms and resource allocation within these firms is for the most part not determined by a price system. Rather, it is determined by managerial control, a form of economic planning. Take for example the allocation of workers to tasks within a firm. This is not determined by workers choosing freely what to do based on a system of piece rates. Rather, workers are directed by superiors within the firm to perform certain tasks and not others.

Capitalist economies thus clearly have "islands of conscious power" within an "ocean of unconscious co-operation" as D.H. Robertson once put it (Robertson, 1928). And some of these islands of conscious power are quite large. For example, the largest company in the world by employment as of the time of this writing (i.e., 2020) is Walmart Inc. with over two million employees of which about one and a half million are in the United States. Quite a few companies have more than one hundred thousand employees. More generally, there are firms of all sizes and these firms are ubiquitous in capitalist economies. This led Coase to ask: "in view of the fact that it is usually argued that co-ordination will be done by the price mechanism, why is such organization necessary? Why are there these 'islands of conscious power"' where the price mechanism is suspended?

Understanding this question was particularly pressing in the 1930s. The contrast between mass unemployment in Western economies during the Great Depression and rapid economic growth in the Soviet Union associated with its five year plans for industrialization was quite stark. This contrast led many to question the wisdom

of laissez faire capitalism and its heavy reliance on the price system and instead argue that central planning was necessary for sustained prosperity. Lenin had argued that the Soviet economy could be run like a large factory. Many western economists ridiculed this notion. But why? After all, enormous firms played an integral part in western economies. Why couldn't the whole economy be run like one large firm?

8.1 The Importance of Transaction Costs

In his 1937 article, Coase proposed an answer to these questions. The reason he proposed for why not all economic coordination occurs through markets is that it is costly to engage in "market" transactions. In the stylized models we usually write down, all markets are perfectly competitive. Furthermore, such markets exist for every conceivable good and service we may want, and they can be used without any effort or cost. This is obviously far from an accurate description of the real world. Take for example the "market" for the services of a plumber. In any large city there are many plumbers offering their services. So, one can (loosely speaking) say that there is a market for plumbers in the city. But this market is far from perfectly competitive and it involves significant effort to transact in this market. The first time one seeks the services of a plumber in a city one must find a "good" plumber. Even with the aid of modern technology (the internet, etc.) this can be quite time consuming. A major issue is assessing quality. Once one has settled on a plumber, one must contact them, discuss the job, discuss terms, and discuss the plumber's availability. If the plumbing job is more than something very minor, one will likely want to get a quote before agreeing to hire the plumber. For a substantial job, one may want to get several quotes.

My wife and I got more than 10 quotes for our recent seismic retrofit renovation. This involved waiting at home 10 mornings for contractors to come and then having a 30 minute conversation with each contractor about the job. Why did we think this made sense? First, we needed to educate ourselves about how to evaluate contractors. Talking to contractors proved useful for this (one could cross-check what one contractor said with the next). Second, not all the contractors turned out to be high quality (as perceived imperfectly by us). Third, the bids we got from them were highly heterogeneous. The price difference between the highest and lowest bid was more than a factor of two! Clearly, the market for seismic retrofits—even in Northern California where everyone is waiting for "the big one"—is far from perfectly competitive.

But getting a good quote from a high quality contractor is not the end of the story. After the contract has been signed and the work begun, one must monitor the work. Often there are unforeseen circumstances that call for "change orders." These can involve significant bargaining. Finally, in the case of specialized contracting work it is important to get the work inspected.

This example illustrate that purchasing the services of a contractor often involves substantial *transaction costs*. The same is true of other market transactions to a larger or smaller degree. Coase argued that the existence of transaction costs will lead buyers and sellers to seek ways to reduce the transaction costs and this quest is what gives rise to much of the institutional structure of modern economies. In particular, he argued that the firm is a device that exists to reduce transaction costs. In the place of myriad market transactions with associated contracts and transaction costs is substituted a single contract: the employment contract. This contract stipulates not all the things that employee should do, but rather simply that the employee should obey the directions of the employer within certain limits. It also avoids the employee having to contract with others involved in the work in question. In this way, the firm cuts down on the complexity and cost of writing a potentially complicated web of contracts. Economizing on transactions costs in this way may be crucial to making the transaction worthwhile at all.

Coase emphasized that organized markets such as stock exchanges and commodities exchanges are also institutions that exist to facilitate trade, i.e., to lower transactions costs. Interestingly, all such exchanges regulate trade in great detail: "What can be traded, when it can be traded, the terms of settlement and so on are all laid down by the authorities of the exchange. There is, in effect, a private law. Without such rules and regulations, the speedy conclusion of trades would not be possible" (Coase, 1992, p. 718). This highlights the importance of rules and agreed upon standards for efficient trade. Coase goes on to say: "Of course, when trading takes place outside exchanges (and this is almost all trading) and where the dealers are scattered in space and have very divergent interests, as in retailing and wholesaling, such a private law would be difficult to establish and their activities will be regulated by the laws of the state." The character of these laws are for this reason crucial for the efficient functioning of commerce. In this sense, one can say that organized markets exist to reduce transaction costs and certain branches of law – such as corporate law, employment law, and tort law – exist partly to reduce transaction costs: they set standards and precedents that later parties can use with confidence and at low cost.

These ideas suggest that there will be a balance in an efficiently run economy between transactions that occur in the market and transactions that occur within firms. Some transactions will involve less transactions costs if done within firms, while others will involve less transactions costs if done on the market. As we have emphasized earlier in this chapter, relying on well-organized markets to at least some extent, has the advantage that such markets aggregate information through prices. They also provide discipline on firms that are not well run. An economy organized as a single firm lacks all such discipline. Finally, very large firms tend to have many layers of middle management. This then results in many layers of principle-agent problems which weaken incentives (see section 9 for more discussion of principle-agent problems).

8.2 The Coase Theorem

Twenty-three years after he published his paper about the nature of the firm, Coase wrote another paper which is, if anything, even more seminal. This paper is titled "The Problem of Social Cost" (Coase, 1960). In this paper, Coase critiques the standard economic analysis of harmful effects on others – which economists often refer to as externalities. The classic example of such harmful effects is that of a factory which pollutes its surroundings. Suppose the factory is not liable for harm associated with this pollution. In this case, the owners of the factory may not take this harm into account when they decide how to run the factory and this will typically lead them to pollute more than is efficient. Their private gain from polluting is larger than the social gain from polluting because they don't bear some of the harm that results from the pollution. In a nutshell, there is a divergence between private and social cost associated with polluting and this leads to inefficient decisions about pollution.

A standard argument made by economists is that this situation calls for some combination of the factory being made liable for harm and a tax on the pollution emitted by the factory. The idea is that these remedies will "internalize the externality", i.e., lead the factory to bear the cost of the pollution that it emits and in this way restore efficiency. In his article, Coase challenged the notion that such remedies are appropriate. He argued that "they lead to results which are not necessarily, or even usually, desirable" (page 2).

An important insight emphasized by Coase is that from an efficiency perspective situations like that of the factory that pollutes are reciprocal in nature. Suppose an

activity of A inflicts harm on B. Coase argues:

> To avoid the harm to B would be to inflict harm on A. The real question that has to be decided is, Should A be allowed to harm B or should B be allowed to harm A? The problem is to avoid the more serious harm. (page 2)

In the case of the polluting factory, taxing the factory harms the owners of factory and benefits those harmed by the pollution. In theory, there exists an appropriate level of such a tax that balances the marginal harm inflicted on the factory and the marginal reduction in harm resulting from the pollution. This tax level will bring about efficiency with respect to the pollution emitted by the factory.

A worry you may have with this solution is that the government may not be in a good position to assess what the appropriate level of the tax is or may not be able to administer the tax efficiently. Coase was quite critical of economist's casual attitude towards the notion that market failures could be rectified by taxes and subsidies imposed by the government. He called such analysis "blackboard economics." It is easy for a professor in front of a blackboard to demonstrate that an appropriate tax can solve all manner of problems. But this ignores the hard problem of figuring out the appropriate level of the tax in a practical situation and administering the tax in a low-cost manner.

The question then arises whether there is another solution to the externality problem. As a stepping stone towards answering this question, Coase showed that under the idealized assumptions typically made when economists engage in blackboard economics, there is actually no need at all for government intervention. This result has come to be called the *Coase Theorem* and can be stated (somewhat loosely) as follows:

Coase Theorem. *If agents are rational and transactions costs are zero, private negotiations between agents will result in resources being efficiently allocated independent of how the rights over the resources are initially allocated.*

The example Coase developed in detail to illustrate this result was one of cattle straying onto a neighboring farm and destroying crops. In this case, one can imagine two initial allocations of rights: one in which the cattle-rancher is liable for damages, and another in which he is not. If the cattle-rancher is liable for damages, he will limit the size of his herd or build a fence to limit his liability. He will do this to the

point where the marginal cost to him of additional such actions equals the marginal benefit to him of additional such actions in terms of lowering his liability.

If in contrast the cattle-rancher is not liable, it is in the farmer's interest to pay the cattle-rancher to limit the size of his herd or build a fence to limit the damages experienced by the farmer. The farmer will pay for such actions up to the point where the marginal cost to her of inducing the cattle-rancher to do more is equal to the marginal benefit she experiences from additional such actions in terms of lowering the damages from the cattle.

Notice that as long as the cattle-rancher and the farmer can negotiate costlessly and enforce their liability rights costlessly they will reach the same outcome in terms of the size of the herd and the size and strength of the fence whether of not the cattle-rancher is liable for the damages. The only difference between these two cases will be who incurs the costs associated with the mitigation actions. Furthermore, in both cases, the resulting allocation of resources will be efficient in the sense that the externality imposed by the straying cattle will be reduced to the point where additional mitigation actions are no longer worth the benefit they induce in terms of reducing harm to the farmer (i.e., the marginal benefit of additional mitigation actions is equal to its marginal cost).

In his article, Coase developed this example in much more detail and addressed various objections. Medema (2020) discusses a host of additional objections that have arisen in the subsequent literature. Medema shows that if zero transactions costs are defined appropriately, the Coase theorem survives all these objections. However, the lesson from this literature is that the conditions needed for the Coase theorem to hold are quite stringent. Bargaining must be costless. But in addition, information must be perfect: everyone must know everything everyone else knows. Furthermore, income effects must be zero. Having laid out these stringent conditions, Medema is able to provide a formal statement of the Coase theorem.

The Coase theorem is extremely controversial. On the one hand, it has had an enormous influence both within academia and in public discourse. There are branches of scholarship within economics and law where the Coase theorem has becomes a "rebuttable presumption" (i.e., analysis proceeds by default under the assumption that transactions costs are zero and the Coase theorem holds in a particular setting unless the researcher has provided strong evidence that this is not the case). Particularly among those that argue that markets work well, the Coase theorem is evoked frequently (either explicitly or implicitly). This is understandable given that the Coase theorem is a particularly expansive statement of the power of

voluntary exchange to yield efficient outcomes.

The Coase theorem is so expansive a statement of the power of the market that even the faculty of the economics department at the University of Chicago were initially skeptical. In 1960, members of the faculty of the economics department at the University of Chicago invited Coase to defend his then controversial position at a dinner party at one of their homes. At the end of the evening, Coase was said to have convinced those attending that the market could work in a sphere they had not previously thought it could work (i.e., in cases of externalities), something that was very hard to do at Chicago. It was George Stigler of the University of Chicago who coined the phrase "Coase theorem" to describe these ideas.

Many, however, question the relevance of the Coase theorem in practical situations, especially when the harm resulting from someones actions affects many others. The notion that all these people can costlessly negotiate an efficient solution seems ridiculous to many. But antagonism towards the Coase theorem goes beyond this. It arises additionally from moral objections many have to the conclusions that flow from the Coase theorem. As we discuss above, the Coase theorem emphasizes the reciprocal nature of many economic situations when it comes to attaining efficiency, and the fact that efficiency can be attained independent of the initial allocation of rights. In particular, when the Coase theorem holds, it doesn't matter from an efficiency standpoint whether injurers are liable for their actions or victims are liable: both the case of injurer liability and victim liability will result in the same outcomes (e.g., same amount of pollution or same safety standards). But even if this is true, the initial allocation of rights *does* affect who bears the cost of the harm. Many feel strongly that victims should not be made to bear the cost of harm imposed by others, i.e., that injurer liability is superior on moral grounds. As Medema points out, the mere possibility that the Coase theorem "could be used to justify making "innocent" victims liable for industrial pollution or torturous harm [has been] sufficient to generate vociferous opposition to the theorem" (Medema, 2020, p. 1079).

8.3 The Coase Theorem and the Real World

The zero transactions cost world of the Coase theorem is a very strange place. It is sufficiently strange that is it hard to fully grasp. If transactions costs were zero, firms, the law, the government, even property rights, would have no purpose. In such a world, all problems would be solved by private negotiations, everything

from one person buying a sack of grain from another to global limits on greenhouse emissions. Coase himself emphasized that the world of zero transactions costs is unrealistic. "It would not seem worthwhile to spend much time investigating the properties of such a world," he argues in Coase (1988, p. 15). Discussing this case, was to him meant to drive home "the need to introduce positive transactions costs explicitly into economic analysis so that we can study the world that exists." On this point, he was disappointed: "This has not been the effect of my article."

In a world with positive transactions costs, institutions can be viewed as devices to lower transactions costs. We discussed the role of firms and organized markets in this respect in section 8.1. But these are just a few examples of institutions. The law, more generally, can serve as a powerful tool to reduce transactions costs. One mundane – but hugely important – way the law does this is by setting standards. For example, a meter, a kilogram, an ounce, and a yard all have precise legal meaning. The founding fathers of the U.S. were conservative in the powers they granted Congress. But one of the powers they did grant Congress was the power to "fix the standard of weights and measures" (Article I, section 8). This was considered important because ambiguity in the definition of weights and measures and heterogeneity across locations had been a serious source of transactions cost since time immemorial.

Another way the law reduces transactions costs is by creating pre-specified contract types such as the corporation, the partnership, the condominium association, the marriage, etc. The standard nature of these contract types implies that the cost of dispute resolution is vastly lower than if all contracts were bespoke. When contracts are standardized through statute, a vast amount of legal precedent accumulates about these standard contracts which guides parties when disputes come up and also guides courts when these disputes lead to legal action. Boilerplate language in contracts serves the same purpose.

In addition to reducing the cost of dispute resolution, the existence of standard contracts allows parties to avoid the potentially hugely costly activity of writing bespoke contracts. Suppose for example you wish to hire a contractor to build a house. Without any standardization, this task would be incredibly difficult. There are innumerable contingencies that the parties may worry about. Writing a thorough contract involves negotiating about these contingencies; it also involves finding a common language to describe the nature of the contingencies and what is to happen in these contingencies; finally, this language must be understandable not only to both parties to the contract, but also to an outside authority (a court). In

countries with strong legal traditions, most people need not worry too much about this because the legal system along with boilerplate contracts has decided so many aspects of how their interaction with a contractor or some other professional will be handled. In this way, standardization and law is essential to making complicated economic transactions manageable.

My experience writing scientific papers has led me to be very sympathetic to the notion that the standardizing role of law and regulation is essential for the efficient functioning of the economy. I spend enormous amounts of time trying to convey what my papers are about as simply and clearly as I can. Invariably, however, I find that very smart people have huge amounts of trouble understanding what I am trying to say. Most of my colleagues report having similar trouble. The world of zero transactions costs completely assumes away this important element of reality.

An interesting literature exists within legal scholarship about exactly what can be accomplished through law that would be virtually impossible to accomplish through private contracting in our positive transactions cost world. One important example is property rights. Legal scholars emphasize the difference between property rights and contractual rights (Hansmann and Kraakman, 2002; Ayotte and Bolton, 2011). Property rights are rights that exist against all parties, while contractual rights only exist against other parties to the contract. For example, the owners of a beachfront property may agree to allow their neighbors to pass through the property to access the beach. However, if the owners then sell the beachfront property, this agreement does not bind the subsequent owners unless it has established a property right (an easement in this case). The law governs when such contractual rights become property rights. Establishing property rights by private negotiations is virtually impossible since it requires agreement of all parties in the economy (world). The law is therefore the only practical mechanism that exists to create property rights. Another important example is the role of organizational law in entity shielding, e.g., shielding business assets from the personal creditors of the owners of the firm (Hansmann and Kraakman, 2000).

Those on the right of the political spectrum sometimes say: "Government is not the solution, government is the problem." But in a world with positive transactions costs strong government institutions are arguably essential for private enterprise to work well. The Coase theorem suggests that all problems should be solved through private negotiation. But how would such private negotiations work in practice? Isn't the government the (imperfect) mechanism we have created to carry on these negotiations (at least in liberal democracies)?

Liberiterians will object to this last point due to an important difference between acts of legislatures and private agreements. In the case of a private agreement, all parties must agree. In contrast, legislatures can pass laws and regulations that bind all parties even when some oppose these acts. This is the basic reason why liberitarians often oppose expansive action by legislatures: such action typically involves some degree of coercion since those that oppose the action are compelled by the state to abide by it despite their opposition.

Liberitarians hold that no one should be forced to do anything against their will (except avoid harming others). To attain this goal would mean that any collective action by society would need to be agreed on unanimously by all members of society. In its purest form, this position is arguably utterly impractical. For this reason all societies develop institutions that allow a subset of the members of society to make decisions that bind all. Furthermore, most associations within society set up their bylaws in such a way that decisions can be made without unanimity (majority rule is common).

Neoclassical economics emphasizes competition and voluntary exchange as a means of generating prosperity. The Marxist intellectual tradition, however, emphasizes collective action of classes. One way to understand this focus of Marxist thought is as a response to the fact that we live in a world of positive transactions costs: collective action of the working class through unions and left-wing political parties can be viewed as a mechanism to reduce transaction costs. Adherents of the Coase theorem will often claim that each individual worker should simply negotiate with his or her employer. In a world with positive transactions costs, this obviously involves an grossly inefficient amount of transactions costs (just like everyone doing their own plumbing is highly inefficient). The efforts of organized labor to improve safety standards, increase wages, reduce working hours, improve benefits, etc. are thus an important part of our society's response to the unfortunate fact that we live in a world with positive transactions costs. Of course, it is not always the case that the institutions we create to facilitate transactions will result in an efficient outcome. The power of organized labor can be either too great or not great enough. Unfortunately, it is hard to tell which is the case or whether some other type of institutions would serve us better. Coase would have liked to see us economists spend more time pondering these types of questions.

8.4 Harmful Effects when Transactions Costs are Positive

Let's conclude this section by coming back to the example of the polluting factory and think through the economics of this situation and others like it in the presence of positive transactions costs. Coase argued that it is important to think of what is traded in the market as bundles of rights to perform certain actions. When one views economic activity through this lens, it becomes clear that the right to pollute is not fundamentally different from the right to eat an orange or the right to build a structure on a plot of land.

Coase further argued that rights "will tend to be acquired by those for whom they are most valuable either for production or enjoyment" (Coase, 1992, p. 12). The extent to which this is true will, however, depend on the degree of transactions costs involved in transferring the rights in question to those for whom they are most valuable. These transactions costs will rise with the complexity of the bundle of rights involved, but also with the number of parties involved. Transactions costs are therefore likely to be particularly high in cases having to do with pollution since those harmed are often extremely numerous. This suggests that leaving environmental matters to the private sector is less likely to work well than many other types of transactions that affect fewer parties.

In a world of positive transactions costs, there is generally speaking scope for the government to improve the allocation of resources by creating institutions, laws, or regulations that reduce the need for complex private negotiations. One important example of this is the default allocation of rights. When transactions costs are zero, whether injurers or victims are liable does not affect outcomes (other than the distribution of rewards). When transactions costs are positive, this is no longer the case. In this case, transactions costs will prevent many rights from being transferred through private means to those that values them the most. This implies that the initial default allocation of these rights will matter. If the default allocation of rights is "good", the outcome is more likely to be good.

In the real world, injurer liability will generally result in outcomes that are more favorable to the victim not only in terms of financial rewards but also in terms of the harm in question (e.g., pollution or workplace safety). Governments can go further and regulate or tax certain activities (think again of environmental or workplace safety regulations) so as to shift economic activity towards an efficient outcome that private contracting cannot achieve.

How much government action is beneficial depends importantly on how well

the government functions. Just as private agents face transactions costs, so do governments. It is not at all clear that governments can in general reach better solutions than the private sector. The government does have certain tools that are not at the disposal of the private sector such as the ability to compel. But this tool only improves outcomes if the government is able to figure out what a good outcome is and if it has the incentives and ability to implement such an outcome. Views differ widely on whether this is the case. Coase himself tended to be quite skeptical of government intervention. And as I mentioned earlier, he was very critical of economists' simplistic analysis of government intervention. In his view:

> [T]he fact that there are transactions costs and that they are large implies that many effects of people's actions will not be covered by market transactions. Consequently, "externalities" will be ubiquitous. The fact that governmental intervention also has its costs makes it very likely that most "externalities" should be allowed to continue if the value of production is to be maximized. This conclusion is strengthened if we assume that the government is not like Pigou's ideal but is more like his normal public authority – ignorant, subject to pressure, and corrupt. (Coase, 1988, p. 26)

Whether this perspective is correct is still an open question.

9 Market Failures

As we discussed in sections 7 and 8, even a cursory glance at the world reveals that the conditions needed for the First Welfare Theorem or the Coase theorem to hold are not satisfied. Many economists, however, believe that these theorems are useful because they help us organize our thinking about "market failures," i.e., the ways in which the world deviates from the conditions needed for markets to produce an efficient outcome. In this section, we will briefly discuss the main classes of market failures that have been emphasized in the economics literature. Some of these are more widely appreciated than others (and discussed in more detail in introductory economics texts). Many are discussed in detail elsewhere in this book in more specific contexts.

9.1 Transactions Costs, Externalities, Public Goods

We have actually already discussed in considerable detail two important sources of market failure, namely *transactions costs* and *externalities*. We did this in section 8. Another very relates source of market failure is *public goods*. Public goods are goods that provide benefits to all people in a particular group or location. Classic examples include national defense, police, fire departments, weather forecasts, and lighthouses. These examples vary somewhat in the extent to which the goods or services in question are excludable (i.e., whether it is possible to provide them to some without also providing them to others). National defense is a particularly good example of a public good since it is virtually impossible to provide it to some in a particular location (say a city) without providing it to essentially all in that location. Due to imperfect excludability, public goods give rise to a free rider problem. People have an incentive to contribute too little to the provision of public goods because they can free ride on the contributions of others. This tends to lead to under-provision of public goods by the private sector and provides a justification for government provision or these goods.

9.2 Market Power

In the perfectly competitive markets needed for the First Welfare Theorem to hold, there are many buyers and many sellers. Each buyer and each seller in such a market is sufficiently small relative to the overall size of the market that they believe that their own actions don't affect the price in the market. As we have discussed earlier in the chapter, this idealized situation is what we mean by the term perfect competition.

The market for most products is far from perfectly competitive. Any given product is typically sold by relatively few firms in a particular area and these firms are aware of the fact that if they try to sell more they will likely need to lower their prices. Firms that face such downward-sloping demand curves have weaker incentive to sell more than do firms in a perfectly competitive market. The reason for this is that they lose money on the inframarginal units (the ones they are already selling) when they try to sell more because they need to lower their prices on all units sold. This reduces the firms' incentive to sell more and therefore leads their prices to be set higher than they otherwise would be. We say that firms in such markets have *market power* and that this market power leads to higher prices than are efficient.

Market power is arguably ubiquitous in the economy. There are several reasons for this. One reason is returns to scale. The production of some products involves substantially increasing returns to scale. This is the case when production involves a large fixed cost but modest marginal cost. Classic examples of such products include many utilities such as water systems, electical systems, sewer systems, road systems, phone systems, and broadband internet. In some such cases, the fixed cost involved in supplying these goods and services are so large that the most efficient way to supply them is for there to be a single provider. In this case, we say that the market for that product is a natural monopoly. But even in cases where returns to scale are less severe, they may still be sufficiently large to result in an oligopolistic market with very few sellers.

Another reason why market power is ubiquitous is that products are extremely heterogeneous. Each firm typically produces products that differ in various large and small ways from the products produced by other firms (e.g., different models of cars, different brands of clothes, different types of yogurt, etc). Often, firms have intellectual property rights that mean that other firms cannot sell exactly same product that they sell. Since consumers differ in their tastes, firms have market power being the only ones that supply the exact product they produce.

A third source of market power is location. Even for standard products, there are often not many sellers of the same or similar products close by. This tends to be a more serious source of market power in smaller towns, areas that are sparsely populated, and locations where consumers are locked in (airports, sports events, etc.). But people's unwillingness to change the location in which they purchase goods can be surprisingly strong. I have worked in a building where a can of the same type of soda cost a different amount in vending machines on different floors.

Market power comes in two different flavors: *monopoly power* is market power by a seller of a good, while *monopsony power* is market power by a buyer of a good. The examples discussed above are examples of monopoly power in product markets. Firms with monopoly power in product markets use it to mark up the price of the products they sell above marginal cost. Monopsony power can also occur in product markets, but is arguably a more serious issue in labor markets. Firms with monopsony power in the labor market use it to mark down the wages that they pay their workers below the workers' marginal product. Competition in the labor market constrains firm's ability to push down wages in this way.

In many cases, there are technological reasons or reasons having to do with people preferences for why market power exists. But in other cases, laws and regu-

lations have been erected that limit competition. Tariffs and other import barriers are one important example of such policies. They limit competition from abroad. But many other examples exist. Producers will often lobby their government for protection against competition and in many cases such lobbying is successful partly because those that stand to gain from competition are a more diffuse bunch than those that stand to gain from any particular hindrance to competition. In an insightful book, economists Raghuram Rajan and Luigi Zingales discuss how we need to "save capitalism from the capitalists" who want to stifle competition (Rajan and Zingales, 2003).

In cases where market power occurs for technological reasons, an important question is how to best limit its detrimental effects. One solution is to have the government be the provider and set prices close to marginal cost. Another example is to have the government regulate the activities of a private corporation. These two solutions are employed quite widely in most economies (for example when it comes to utilities). However, those skeptical of the intentions and abilities of governments sometimes critique these solutions and argue that living with an unregulated private monopoly is a lesser evil than public monopoly or public regulation of private monopoly. A classic exposition of this argument appears in Friedman (1962, p. 27-32).

9.3 Asymmetric Information

Asymmetric information gives rise to two different types of market failures: *adverse selection* and *moral hazard*. To illustrate adverse selection, consider the market for health insurance. Suppose a (naive) entrepreneur has decided to offer health insurance and is thinking about how to price the insurance. As is typical of insurance companies, the entrepreneur doesn't know which potential buyers are relatively frail and which are in more robust health. All he knows is the distribution of people's health (what fraction is healthy, what fraction is frail, etc.). Being not too experienced in this business, the entrepreneur considers how costly it will be to offer people insurance on average if everyone (or a random sample) buys his insurance. Suppose this turns out to be $1,000 per year. The entrepreneur then offers insurance for $1,000 plus a small profit margin.

The trouble (for the entrepreneur) is that the people considering purchasing health insurance know more about their health than he does. Some of them know that it is unlikely that their medical bills will reach $1,000 in a year, while others

know that for them medical bills will likely be much higher than $1,000. Who buys the insurance? Not surprisingly, it ends up being disproportionately those that know that their medical bills will likely be high. The consumers self-select into buying the health insurance, and from the perspective of the entrepreneur, those that end up buying are adversely selected (they have higher than average expected costs). This implies that the costs of the entrepreneur end up being much higher than $1,000 per person insured over the course of the year and he looses money on the whole operation.

Not being too bright, the entrepreneur decides that the solution to this problem is to raise the price of the insurance. So, the next year he offers the insurance for $2,000 per year. Surely, this will be a high enough price for him to make a profit, he thinks to himself. But in response to this price increase some of his customers decide that now the health insurance he is offering is no longer worth their while. These tend to be the healthiest of his customers. They discontinue their policies and the entrepreneur is left with an even sicker customer base and even higher costs per customer. Again his insurance business makes a loss.

In year three, he again raises his price, this time to $4,000. But again, his healthiest customers leave and he is left only with customers who on average have higher medical bills than their insurance premia. As time passes, his price keeps rising and his pool of customers keeps shrinking. In this way the market for health insurance unravels. At some point, the entrepreneur realizes that there is no price at which he can make money and he simply stops offering health insurance.

In this example, the market for health insurance unravels due to adverse selection on the part of the customers who have more information about their health. This is a potentially important problem in all insurance markets but also in other markets such as the market for used cars and even in some cases for the medium of exchange as we will discuss in chapter XX [money and banking chapter].

Adverse selection does not in all cases lead to complete unraveling. In the case of insurance, risk aversion on the part of consumers can lead some to buy insurance even if its price is higher than the expected value of the payout. Risk aversion thus helps avoid complete unraveling. However, many insurance markets don't exist due to unraveling. For example, prior to the passing of the Affordable Care Act (Obamacare), about 45 million Americans did not have health insurance. For many of these people, no reasonably priced health insurance was available.

In the case of insurance, one way to solve an adverse selection problem is to mandate that everyone buy insurance. In this case, people are no longer able to self-

select. Such mandates are actually quite common when it comes to insurance. For example, all car owners are required to purchase car insurance and many banks require lenders to purchase fire insurance. Obamacare instituted a mandate for health insurance as well. This was one of the mechanisms the law employed to improve the health insurance market (although the mandate was effectively ended in 2019 when the penalty for violating the mandate was reduced to zero).

Adverse selection occurs when one party to a transaction knows more about the some aspect of what is being traded than the other. This is referred to as a hidden information problem. A different type of asymmetric information problem is one were one party to a transaction is not able to observe the actions of the other party. This type of problem is also important in insurance markets. The basic issue is that people that have purchased insurance put in less effort to guard against the hazard they are insured against. For example, insured drivers take less care not to damage their cars than uninsured drivers. The only reason they can get away with doing this is that the insurance company can't observe how carefully they drive. This is called a hidden action problem. It gives rise to moral hazard on the part of the insured party.

A particularly important context in which moral hazard arises is the employment relationship. Employment contracts ideally provide employees with incentives to work hard but also with insurance against bad luck. The problem is that employers are often not able to tell whether a bad outcome is due to bad luck or insufficient effort on the part of the employee. This gives rise to a trade-off between providing incentives and insurance. If the employment contract fully insures the employee, i.e., gives the employee a wage that is independent of the outcome, the worker's moral hazard is very extreme in that he or she has no incentive to put in effort. This will result in an inefficiently poor outcome. In order to induce the worker to put in an efficient amount of effort, the contract must thus reward the worker when the outcome is good. But this implies that the contract must scale back the amount of insurance it provides against bad luck.

The employment contract is actually only one example of a more general class of economic relationships called principle-agent relationships. In such relationships, one party (the principle) is asking another party (the agent) to perform some task. The problem of how to design an efficient contract between the principle and the agent when the principle can't observe the actions of the agent is called the *principle-agent problem*. Examples of principle-agent problems include an owner of a firm hiring a manager, an owner of land renting the land to a tenant farmer, an employer

hiring an employee, a home owner hiring a contractor, a client hiring a lawyer (or a doctor or a broker of some kind), and citizens electing politicians.

In all these cases, the agents have interests of their own which do not necessarily align with the interests of the principle. A major task of a contract between the principle and the agent is to align the interests of the agent with those of the principle. Since the principle can't observe the actions of the agent, the contract cannot simply specify what the agent should do. Instead, the contract typically gives the agent a financial reward that is larger if the outcome is good for the principle (e.g., a bonus to a CEO if the firm's profits are large). As we discuss above, there is typically a trade-off between providing the agent with strong incentives and offering him or her insurance against bad luck.

Another important example of moral hazard arises due to the fact that owners of firms typically have limited liability. Specifically, the combination of debt and limited liability gives rise to moral hazard. In such a situation, the owners of a firm have an incentive to take too much risk. If things go well, they reap the rewards. If things go badly, however, their losses are capped by limited liability (the most they can loose is the money they put into the company) and any additional losses are borne by their creditors. This type of moral hazard can be very extreme in a situation where a firm (or bank) with a lot of debt is close to going bankrupt. At such a point, the equity holders have already lost most of their equity. If they gamble the future of the firm by taking a great deal of risk, they face a "heads I win, tails you loose" situation, where "you" are the firm's creditors. The fact that firm owners take too much risk in this type of situation is called a debt "overhang problem," and the owners are said to "gamble for resurrection." To avoid this type of situation, creditors usually require firms to have a minimum amount of equity or "skin in the game" as a fraction of the funds they lend to the firm.

9.4 Rationality

The theoretical results discussed earlier in this chapter about the efficiency of markets – e.g., the First Welfare Theorem – depend heavily on the notion that people are rational. Recall that we define rationality as people being able to choose what is best for them from the set of choices they face (i.e., act in their own best interest). If people are not rational, there is no reason to believe that markets will result in a Pareto efficient outcome. In fact, there is every reason to believe that markets will result in a great deal of what Akerlof and Shiller (2015) call "phishing for phools." Akerlof

and Shiller argue that the same profit motive that yields efficiency when people are fully rational will spawn manipulation and deception if people are boundedly rational (i.e., not fully rational). If people can be tricked into paying too much for a product or buying products they have little use for, the market system gives sellers strong incentives to exploit these weaknesses. As a result, the market system will give rise to a great deal of "phishing," i.e., effort to get people to do things that are in the interest of the phisherman but not in their own interest.

To appreciate this better, it is useful to think about what full rationality implies. One implication of full rationality is that people have unlimited cognitive abilities and thus face no costs or trouble in understanding the choices that they face and the information that they receive. This implies that it is not possible to trick or confuse a fully rational person about anything. No matter how complicated the choices are that face such a person, they will never make a mistake. For example, choosing between the various health insurance options or retirement savings products their employer provides will be as easy as ABC. Likewise, deciding which type of mortgage to take out when buying a house, which add-ons to include then buying a car, which treatment option to choose when dealing with a medical problem, and the strength and credibility of evidence backing up claims made by the makers of various consumer products (such as beauty products and nutritional supplements) would all be as easy as a walk in the park.

Retirement savings, mortgage finance, and treatment choices at the hospital are complicated issues even if no one is trying to trick you. But much simpler matters – such as which snack to choose at the coffee shop – can become tricky when the seller *is* trying to trick you. For example, sellers of food routinely try to make their products seem healthier than they really are and make it difficult for consumers to assess the calorie contents of their products. Makers of beauty products routinely overstate the effectiveness of their products at reducing aging. And many products have "shrouded attributes" – attributes people tend not to notice or fully appreciate when buying (such as a resort fee at a hotel or high cost of replacement parts and maintenance). This is the case even though these products are often highly regulated, producers are required to include various information on labels, and false advertising is illegal. Examples of such phishing range from the relatively harmless to extremely serious matters such as the efforts of tobacco companies to mislead the public about the health consequences of smoking tobacco, the oil and gas industry's efforts to undermine research showing that leaded gasoline was a serious health hazard, and the efforts of the makers of Oxycontin to push their product on millions

of Americans despite its devastating effect on many that became addicted. None of these efforts would be effective if people were fully rational.

Research in psychology and behavioral economics has documented a long list of clear mistakes that people are prone to make. These include overconfidence (both overprecision and overoptimism), confirmation bias, hindsight bias, money illusion, availability bias, gamblers fallacy, hot-hand fallacy, sample-size neglect, vividness bias, and the list goes on (and on). Strategic settings give rise to a whole different list of mistakes: people play dominated strategies, they fail to reason correctly about how others will act strategically, and they fail to randomize properly when this is optimal. In other words, people don't play Nash equilibria, let alone subgame-perfect Nash equilibria, or more sophisticated rational strategies in complex settings. In addition, there is copious evidence that people are inattentive, are affected by framing, and more generally fail to understand the implications of their actions. All of this makes the real world a place ripe for phishing.

The economist Richard Thaler describes traditional economic theory as replacing humans – homo sapiens – with fictional creatures called homo economicus, or Econs for short (Thaler, 2015). Econs never make mistakes. But humans do. Econs are never phished for fools, but humans, unfortunately, are. One traditional defense of modeling people as Econs, put forth most forcefully by Milton Friedman in an essay titled The Methodology of Positive Economics, is that economic theory should not be judged by the realism of its assumptions but rather by the accuracy of its predictions (Friedman, 1953). Even if people are not rational, it may be that their behavior in markets yields outcomes that are *as if* they were rational. Friedman took the example of an expert billiard player, who plays as though he understands all the complicated mathematical formulas of classical physics even if he does not. Thaler's response is that most humans do not play billiard like an expert billiard player. Economic theory should give good predictions for ordinary people, not just experts.

In addition to the long list of clear mistakes people make, there is another long list of economic phenomena that are often considered mistakes but are less clear cut. These phenomena include the sunk cost fallacy, mental accounting, the endowment effect, loss aversion, addiction, procrastination, lack of self-control, a penchant for instant gratification. These phenomena are puzzling from the perspective of standard economic theory but can be explained if we assume that people have "non-standard" preferences. For example, prospect theory can explain the sunk cost fallacy, the endowment effect, various forms of mental accounting, and loss aversion

(Barberis, 2013; Thaler, 2015), and models with present-biased preferences can explain procrastination, lack of self-control, and a penchant for instant gratification (Loewenstein and Thaler, 1989; Angeletos et al., 2001).

Whether these phenomena result in failures of the First Welfare Theorem depends on whether they are considered mistakes. This is controversial. On the one hand, many economists believe strongly that people's preferences are unchallengeable axioms of their behavior. "De gustibus non est disputandum" is a Latin maxim that means "in matters of taste, there should be no disputes." According to this view, people's choices reveal their true, welfare-relevant preferences. A different view is that people's choices, in some cases, are not well thought out and that there is therefore an important distinction between the way people actually behave and the way they should behave so as to maximize their true welfare. In other words, there is a distinction between "normative" models of behavior – i.e., how people *should* behave – and "positive" or "descriptive" models of behavior – i.e., how people actually do behave.

This normative vs. descriptive distinction is completely standard in the field of ethics. A typical college ethics class will challenge students to think carefully about (and argue with each other about) how people should behave when ethical questions arise. A major objective of this exercise is to give students an opportunity to update their ethical views based on careful reflection. It is common that people will not agree in such classes even after much discussion. Yet many people that have taken such classes have experienced their views and actions changing as a consequence of these arguments (or as a consequence of reading books about ethics). One example is that many people have been persuaded by ethical reasoning that they should stop eating meat. On ethical matters our preferences, thus, seem malleable: they are typically not well thought out ex ante and can change based on reasoned argument. It seems likely that the same is true in other spheres.

The normative vs. descriptive distinction and the notion that our preferences are poorly thought out is, for some reason, much more controversial within economics than it is within ethics. Traditionally, it was a big 'no, no' within economics to view preferences as malleable and a poor guide to a person's true welfare, just as it was a big 'no, no' within economics to view people as boundedly rational. A common defense of these constraints has been that they provided the field with needed discipline (Becker, 1976). The idea is that if we economists allow ourselves to assume that preferences are malleable and people irrational, there will be any number of ways to explain all behavior and no way to distinguish between the different explanations.

This argument is (in my view) profoundly unscientific. In a science, the primary way to distinguish between competing explanations should be to gather empirical evidence on the implications of the different explanations. It should primarily be this empirical evidence that provides theorists with discipline rather than their own preconceptions about what type of theory is good or bad. Of course, empirical evidence does not exist on all matters. So, as a practical matter, theorists will need to resort to other means to choose which types of theories to focus on. But large amounts of empirical evidence does actually exist on people's rationality and (to a somewhat lesser degree) the fact that their preferences are responsive to reasoned argument. So, choosing to exclude explanations based on these notions categorically is not tenable. And considering explanations based on these notions does not mean that 'anything goes' since the empirical evidence provides discipline.

To illustrate further the potential usefulness of distinguishing between normative and descriptive models of behavior, let me discuss two examples. First, many introductory economics textbooks emphasize that sunk costs should be irrelevant to decision making, i.e., a rational decision-maker should ignore sunk costs when making decisions. In the real world, humans have quite a bit of trouble with this concept and quite frequently let sunk costs affect their behavior (e.g., by continuing with some loss making activity because of all the effort they have already expended on it even if stopping would avoid further losses). Economists refer to such behavior as the sunk cost fallacy. Most economists agree that falling prey to the sunk cost fallacy is a mistake and that people should be educated about the irrelevance of sunk costs so that they can make better decisions. But is this so clear cut? Prospect theory can rationalize the sunk cost fallacy. According to prospect theory, people have a reference level of utility and their preferences display loss aversion relative to this reference level of utility. In this case, following through on a course of action where costs have been sunk can avoid (or delay) realizing a loss and therefore be rational. For example, if you have tickets to a sports event, but getting to the event suddenly becomes very costly (for example due to bad weather) you may go to the event anyway to avoid realizing the psychic loss associated with wasting the ticket you have already bought.

Does the fact that one can rationalize the sunk cost fallacy with prospect theory mean that it is not a mistake? Or should one view prospect theory in this case as a descriptive model of behavior, but one that is not optimal? It seems relevant to consider whether people acknowledge sunk cost fallacy as a mistake after careful explanation and reflection. Given how prevalent it is for economists to teach their

students about the sunk cost fallacy, it seems that in this case they mostly (implicitly) agree with the normative vs. descriptive distinction.

A second example concerns self control. Standard economic theory implicitly posits that people have unlimited self control. They have no trouble sticking to plans they have previously made regarding saving for a rainy day and retirement, exercising, not eating too much, and eating a healthy diet. The fact that the checkout lane of the supermarket is lined with sweets and celebrity magazines, has no effect on their spending patterns (beyond what their dispassionate selves would want). They never break down and binge on some food, drink, or activity only to regret it the next day. Using the jargon of economics, standard theory implies that people's choices are time-consistent – i.e., they do not experience systematic preference reversals where they prefer A to B today but B to A tomorrow.

Real humans, of course, have limited self control. They are prone to overspend, procrastinate, and give in to various temptations that are presented to them. This weakness provides the sellers of products with ripe phishing grounds and potentially means that the market system ends up being highly inefficient if it is more profitable to exploit people's weaknesses than it is help them overcome their weaknesses (which seems to be the case). But again we can ask whether people's propensity to procrastinate and give in to temptations they later regret should be viewed as mistakes. Perhaps these are just people's true preferences. Perhaps people wish that they were more motivated to work, exercise, save for retirement, and do any number of other "good" things, but really the psychic cost of these things is simply larger than their benefits. If so, this type of behavior should not be viewed as mistakes and does not give rise to a failure of the First Welfare Theorem.

But another way to explain lack of self control is to think of people as having present biased preferences. People with present biased preferences attach a special extra weight to utility in the present. This leads them to have a penchant for instant gratification and to want to put off costly tasks until later (i.e., procrastinate). Present bias gives rise to time-inconsistent preferences. For example, a person with present biased preferences might prefer to receive two apples the day after tomorrow rather than one apple tomorrow, but when tomorrow rolls around they might reverse themselves and prefer one apple immediately to two apples a day later. These preference reversals imply that a person that realizes that they have present biased preferences may want to act strategically vis-a-vis their future selves: they may want to constrain the impatient behavior of their future selves. We will discuss present bias in more detail in chapter XX [Consumption-Savings].

The notion of present bias provides a rigorous basis for viewing self-control problems as mistakes. Since "the present" is relatively brief, "the future" will get much more weight in a person's calculations about welfare (even with the extra kick given to the present). If they had the option, a person with present biased preferences would therefore prefer to be forced to act as though they didn't have present biased preferences. In reality, many of us seek out commitment devices that tie the hands of our own future selves (e.g., by not having lots of junk food in the kitchen waiting to be eaten by our short-sighted future selves). But solving ones self-control problems on ones own is hard, especially in a world where sellers are constantly trying to play on the passions of ones impatient current self. This means that there is (arguably) ample scope for welfare improving interventions that limit the market's ability to exploit our self-control problems.

More generally, our imperfect rationality – the fact that we are humans, not Econs – is arguably one of the main rationales for government intervention in the economy. A very substantial share of law, regulation, and public policy is geared towards addressing market failures that arise from people's bounded rationality. For example, the purpose of a vast amount of law and regulation is to protect parties that are not sophisticated enough to protect themselves (i.e., consumers and workers). Prominent examples include food and drug safety laws, laws that ban false advertising, and laws that ban predatory lending practices. But we also have laws that explicitly combat present bias, such as the forced saving implicit in the government sponsored retirement savings systems that exist in many countries.

Much of this law, regulation, and public policy is paternalistic in nature. The government is trying to protect people from their own misguided choices. For many whose beliefs tend to the right of the political spectrum, this notion is hard to swallow for at least three reasons. First, they believe people should be free to make their own choices. Second, they believe in personal responsibility. Third, they tend to be skeptical of the intentions and ability of the government. The assumption that people are rational certainly simplifies analysis a great deal: if people are rational, one need not question their choices. Once one acknowledges that people are boundedly rational, things become much more complicated.

9.5 Society's Inability to Make Certain Commitments

A number of serious market failures arise from society's inability to make certain commitments. One example of this has to do with people living in hazardous ar-

eas such as areas prone to flooding, hurricanes, wildfires, earthquakes, or volcanic eruptions. If society could commit not to help people that live in such areas when disaster strikes, the risk of financial loss and loss of life would discourage settlement in such areas. But many societies understandably find it difficult to make such commitments. Once disaster strikes, they feel morally obliged to expend great effort to help those in harms way. This means that people that live in hazardous areas typically do not bear the full cost of their locational choice, which in turn means that too many people live in such areas. This type of market failure provides a justification for public policies that either restrict building in hazardous areas or mandate people that live in such areas to buy insurance for the hazard they face.

Another such market failure has to do with extreme poverty among the elderly. If society could commit not to help elderly people that face extreme poverty, this would provide younger people with an incentive to save for retirement. In reality, we feel morally obliged to provide the elderly with a minimal standard of living. This reduces everyone's incentive to save for retirement. Government pension systems that force those of working age to save a certain fraction of their income are partly a response to this commitment problem (but also a response to self-control problems and general myopia).

A third example relates to large banks. If society could commit not to "bail out" large banks in a financial crisis, these banks would have stronger incentives to avoid taking excessive risk. In reality, the banks understand that their own failure in a financial crisis would be catastrophic for society, which means that society will not let them fail under such circumstances. This encourages excessive risk taking on the part of banks since they face a "heads I win, tails you loose" situation where they reap the profits if their risky bets turn out well but society will bail them out if things turn sour.

In public discourse about bank bailouts, one sometimes hears commentators saying something to the effect of "why don't we simply commit not to bail out banks?" The reason is that we can't. Such commitments are not credible given the costs society would face if it were to let big banks fail in a financial crisis. This commitment problem and the resulting moral hazard problem provides a justification for prudential bank regulation such as capital adequacy requirements, minimum liquidity requirements, and requirements that bank portfolios be diversified.

9.6 Price and Wage Rigidity

For markets to allocate resources efficiently, prices and wages must adjust freely to equilibrate supply and demand. In reality, prices and wages are "sticky," i.e., they adjust infrequently and incompletely to changes in supply and demand. This can lead to situations where prices or wages are too high (say) which can lead output to be too low. In his seminal book *The General Theory of Employment, Interest, and Money*, economist John Maynard Keynes set out to explain how economies could experience large and persistent depressions (Keynes, 1936). His answer was that aggregate demand in the economy could be insufficient. The dominant formal model of Keynes' ideas has been to assume that prices and wages are sticky. If wages are stuck at a level that is too high, demand for labor will be insufficient leading to involuntary unemployment.

The sizable fluctuations in unemployment we see occur every so often and call recessions and depressions are due to this problem (according to the Keynesian view). Inefficient fluctuations in aggregate demand that arise due to sticky prices and wages provide a justification for macroeconomic stabilization policies such as monetary policy and fiscal stimulus. Chapters XX through XX will discuss business cycles, monetary policy, and fiscal policy from a Keynesian perspective in some detail.

9.7 Missing Markets, Search Frictions, Etc.

There are arguably a number of other "frictions" that result in market failures in the real world. Some of these overlap somewhat with the frictions we have already discussed. For example, markets for many goods and services are missing. It is simply impossible to purchase certain types of goods. For example, there are many insurance contracts that don't exist and more generally many financial contracts don't exist. Much of this is probably due to some combination of asymmetric information problems and transactions costs. But some of this is more fundamental. One example is that we can't trade with future generations who are not yet born. This potentially results in huge inefficiency. For example, it may be that future generations would like to pay us large amounts to impose a carbon tax or other policies that will limit global warming. Such a trade is not possible.

A large literature in economics has explored models in which it is costly to find someone to trade with. Such models are called search models. One way to think

about search models is as models that capture one aspect of transactions costs. These models arguably help us understand certain markets such as the housing market, the labor market, and perhaps also certain financial markets. Search models are one formal way of making sense of market not clearing – i.e., resources sitting idle. They thus arguably help us understand phenomena such as unemployment and vacant housing.

The discussion in this section has hopefully brought home the myriad ways in which the First Welfare Theorem doesn't hold in the real world. But I also hope that the chapter as a whole has made a convincing case that models in which markets work perfectly are useful both because they help us understand the power of competition as a force that limits economic exploitation and as a useful starting point for economic analysis that seeks to understand the limits of competitive markets.

Notes

[1]The derivation of equation (11) was for the case when η is given by equation (9). Is it possible that some other values for η may result in a Pareto improvement conditional on equation (11) holding? No. Equation (11) implies that if η is given by equation (9), both Robinson Crusoe and Friday are left indifferent by a marginal variation. (Equation (9) was derived so that Robinson Crusoe would be indifferent. A similar condition can be derived for Friday and this condition will hold given that equation (11) holds.) Now consider a variation with η being larger than equation (9) implies. In this case, Robinson Crusoe will be made worse off (he would be giving up more coconuts per unit of shelter than the amount that leaves him indifferent). If η is smaller than equation (9) implies, however, Friday would be made worse off by the variation (he would be getting fewer coconuts per unit of shelter that the amount that leaves him indifferent).

[2]You may ask: Why are there no $*$s in equation (23)? In section 5, we started off with a candidate efficient allocation P, and used $*$s to denote partial derivatives at that allocation. Here we are deriving the allocation that is implied by markets. This allocation will end up being the allocation P. But we have not yet finished showing this. As of now, we just know that this allocation must satisfy equation (23), which has the same form as equation (11), and therefore implies the same thing (exchange efficiency).

[3]Loosely speaking these are that production sets are convex, preferences are convex and locally non-satiated, and a few more conditions that guarantee that the economy is not at a boundary point. See, e.g., Mas-Colell et al. (1995, chapters 15 and 16), Malinvaud (1985, chapters 4 and 5), or Kreps (2013, chapters 14 through 16) for a rigorous statement of the theorem.

References

AKERLOF, G. A. AND R. J. SHILLER (2015): *Phishing for Phools*, Princeton, NJ: Princeton University Press.

ANGELETOS, G. M., D. LAIBSON, A. REPETTO, J. TOBACMAN, AND S. WEINBERG (2001): "The Hyperbolic Consumption Model: Calibration, Simulation, and Empirical Evaluation," *Journal of Economic Perspectives*, 15, 47–68.

AXELROD, R. (1984): *The Evolution of Cooperation*, New York, NY: Basic Books.

AYOTTE, K. AND P. BOLTON (2011): "Optimal Property Rights in Financial Contracting," *Review of Financial Studies*, 24, 3401–3433.

BARBERIS, N. C. (2013): "Thirty Years of Prospect Theory in Economics: A Review and Assessment," *Journal of Economic Perspectives*, 27, 173–196.

BATSON, C. D., E. R. THOMPSON, G. SEUFERLING, H. WHITNEY, AND J. A. STRONGMAN (1999): "Moral hypocricy: Appearing Moral to Oneself Without Being So," *Journal of Personality and Social Psychology*, 77, 525–537.

BECKER, G. S. (1976): "The Economic Approach to Human Behavior," in *The Economic Approach to Human Behavior*, Chicago, IL: University of Chicago Press, chap. 1.

COASE, R. H. (1937): "The Nature of the Firm," *Economica*, 4, 386–405.

——— (1960): "The Problem of Social Cost," *Journal of Law and Economics*, 3, 1–44.

——— (1988): *The Firm, the Market, and the Law*, Chicago, IL.: University of Chicago Press.

——— (1992): "The Institutional Structure of Production," *American Economic Review*, 82, 713–719.

DAWKINS, R. (1976): *The Selfish Gene*, Oxford, UK: Oxford University Press.

FRIEDMAN, M. (1953): "The Methodology of Positive Economics," in *Essays in Positive Economics*, Chicago, IL: University of Chicago Press, chap. 1.

——— (1962): *Capitalism and Freedom*, Chicago, IL: University of Chicago Press.

FRIEDMAN, M. AND R. D. FRIEDMAN (1980): *Free to Choose*, New York, NY: Harcourt.

HAIDT, J. (2012): *The Righteous Mind*, New York, NY: Panteon Books.

HANSMANN, H. AND R. KRAAKMAN (2000): "The Essential Role of Organizational Law," *The Yale Law Journal*, 110, 387–440.

——— (2002): "Property, Contract, and Verification: The Numerus Clausus Problem and the Divisibility of Rights," *Journal of Legal Studies*, 31, S373–S420.

HAYEK, F. A. (1945): "The Use of Knowledge in Society," *American Economic Review*,

35, 519–530.

KEYNES, J. M. (1936): *The General Theory of Employment, Interest, and Money*, London, UK: Palgrave Macmillan.

KLEVEN, H. J., M. B. KNUDSON, C. T. KREINER, S. PEDERSEN, AND E. SAEZ (2011): "Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark," *Econometrica*, 79, 651–692.

KREPS, D. M. (2013): *Microeconomic Foundations I: Choice and Competitive Markets*, Princeton, NJ: Princeton University Press.

KRUGMAN, P. (2009): "How Did Economists Get It So Wrong?" *New York Times*, September 6, 2009.

LIPSEY, R. G. AND K. LANCASTER (1956): "The General Theory of Second Best," *Review of Economic Studies*, 24, 11–32.

LOEWENSTEIN, G. AND R. THALER (1989): "Anomalies: Intertemporal Choice," *Journal of Economic Perspectives*, 3, 181–193.

MALINVAUD, E. (1985): *Lectures on Microeconomic Theory*, Amsterdam, The Netherlands: Elsevier Science Publichers B.V.

MARX, K. (1867/1887): *Capital: A Critique of Political Economy, Volume 1*, London, UK: Swan Sonnenschein, Lowrey, & Co, translated from the 3rd German ed. by Samuel Moore and Edward Aveling and edited by Frederick Engels, 1887.

MAS-COLELL, A., M. D. WHINSTON, AND J. R. GREEN (1995): *Microeconomic Theory*, Oxford, England: Oxford University Press.

MEDEMA, S. G. (2020): "The Coase Theorem at Sixty," *Journal of Economic Literature*, 58, 1045–1128.

RAJAN, R. AND L. ZINGALES (2003): *Saving Capitalism from the Capitalists*, New York, NY: Crown Business.

ROBERTSON, D. (1928): *Control of Industry*, London, UK: Nisbert & Co.

SANDEL, M. J. (2009): *Justice: What's the Right Thing to Do*, New York, NY: Farrar, Straus and Giroux.

SMITH, A. (1776/2000): *An Inquiry into the Nature and Causes of the Wealth of Nations*, New York, NY: The Modern Library.

STIGLITZ, J. E. (2000): *Economics of the Public Sector*, New York, NY: W.W. Norton & Co.

THALER, R. H. (2015): *Misbehaving*, New York, NY: W.W. Norton & Co.

TRIVERS, R. L. (1971): "The Evolution of Reciprocal Altruism," *Quarterly Review of Biology*, 46, 35–57.