A First Course in Advanced Microeconomic Theory

Josh Evan Barker

July 31, 2026
Version 1.1

First Edition: July 2026

Contents

List of Figures

Introduction to this Book

This book is designed to provide a rigorous and advanced introduction to microeconomics. This book is aimed at advanced undergraduates or early-stage master's level students. Knowledge of Calculus, basic differentiation, and first-order conditions is assumed. A student comfortable with the material traditionally taught in intermediate-level microeconomic theory will be well-prepared.

The goal is to cover many of the foundational topics with an eye towards institutionalism and bridge a gap between intermediate-level and graduate-level Economics. It is not designed to be a comprehensive guide to these topics, but rather to review many key topics. This volume serves as my own bridge from undergraduate to graduate-level Economics, as I completed version 1.1 in July 2026, just prior to my arrival at Clemson University to begin my doctoral studies. This initial version is limited to a core foundation of Price Theory—institutions, consumption, exchange, production, and equilibrium, though I hope to expand upon it when the benefit of devoting more time to this project exceeds the cost of doing so.

The bibliography of each chapter is an excellent source for the eager student seeking to dig in deeper—as much of them are already standard reading for PhD-level theory courses. For the student seeking additional guidance, much of this book aligns with my short-form social media content on Microeconomic Theory. On my YouTube Channel, you can find a video compilation corresponding to each chapter.

Artificial Intelligence was used to review and critique drafts and conversations with AI were useful to help solidify my own knowledge, in addition to the sources cited throughout. AI tools were also extremely helpful with debugging issues with figures and other TEX coding issues. I must acknowledge my appreciation for Perplexity and Gemini in particular, as well as NotebookLM, Claude, and ChatGPT, which were used less frequently. While AI was used for these tasks, AI did not, however, write a single sentence in this book and was not used to generate any written content.

If you identify any errors—typographical, conceptual, or factual—or have other suggestions for improvement, please let me know, as I work to build upon and improve this resource.

Acknowledgements

I draw heavily on the work of economists of both the University of Chicago and UCLA, some of the key people being Friedman, Coase, Alchian, Demsetz, Buchanan, Becker, and Hirshleifer.

I want to thank my numerous teachers over my entire educational journey. I want to give particular thanks: to my undergraduate economic theory professors at Hillsdale College, Michael Clark, Christopher Martin, and Ivan Pongracic who all laid an incredibly solid foundation of economic knowledge and provided me with significant exposure to the "Masonomics" approach; to my first Economics professor, Gary Wolfram, who first exposed me to the Economic way of thinking and is responsible for truly drawing me in; to my professor of the History of Economic Thought, Charles Steele; and to my professors of this study less directly including of Math, Politics, History, Econometrics, and Data Science, Jamie Andrews, David Azerrad, Adam Carrington, Mickey Craig, Jeff Eicher, John Grant, Khalil Habib, Barry Monk, Paul Moreno, Kevin Portteus, Kevin Slack, Christopher Scott Vaughen, and Jennie Zhang.

In my own studies and pursuit of knowledge outside of the classroom, I am heavily indebted particularly to the Liberty Fund for creating and supporting the Online Library of Liberty, Library of Economics and Liberty, and EconTalk. I am also incredibly grateful for my former employers and coworkers, particularly those with a research focus, including the American Enterprise Institute and the Federal Reserve Bank of Philadelphia, who have continued to shape my thinking outside of the classroom.

Chapter 1

Introduction to Economics

1.1 What is Economics?

Economics is a way of thinking—a method for analyzing the world around us. As Paul Heyne summarized it, the economic way of thinking suggests that "all social phenomena emerge from the choices of individuals in response to expected benefits and costs to themselves."

At its root, Economics is about decision-making under constraints. No decision is completely costless. We face scarcity: our wants exceed the means available to satisfy those wants. Our time and energy are limited. If we pursue one project, we must give up time that could be spent on something else. In light of constraints and trade-offs, economics assumes only that individuals make choices consistent with their desires and values. This assumption can be broken down into self-interest—that an individual is seeking to pursue his goals, whatever those goals are—and rationality—that individuals choose based on a reasoned method and they use data to update their mental (cause-and-effect) models of the world. Individuals never see the full picture of all the various tradeoffs they face. Much of decision-making is made under a significant degree of uncertainty. Experiences shape individuals and allow them to update their expectations of the world and understand how the world around them operates.

Heyne (1995) writes that Economics, then, looks at both human actions—"economizing... [that is] to allocate available resources in a way that extracts from those resources the most of whatever the economizer wants"—and human interactions—how "a process of continuous, ongoing interaction among suppliers and demanders... an exchange process" operates. Heyne writes:

> If the core problem for economic *actions* is scarcity, the core problem for economic *interactions* is a multiplicity of diverse and incommensurable projects. The solution to the scarcity problem is economizing; the solution to the problem of diverse projects is coordination.

In "The Economic Way of Looking at Life" (1993), Chicago economist Gary Becker writes that economics "is a *method* of analysis, not an assumption about particular motivations."

> The analysis assumes that individuals maximize welfare *as they conceive it*, whether they be selfish, altruistic, loyal, spiteful, or masochistic. Their behavior is forward-looking, and it is also assumed to be consistent over time. In particular, they try as best they can to anticipate the uncertain consequences of their actions.

The so-called *homo economicus* does not have perfect foresight nor is he perfectly selfish. Rather, using the information he has, he makes the most of his resources to predict the best way to accomplish his goals. People learn through doing and trial and error. It is through a feedback mechanism that individuals' actions are moved in the direction of "optimality."

As a discipline, Economics is generally concerned with consumption, production, and exchange within markets. Thus, there is a strong focus on the business world and production and consumption of various goods and services. Yet the Economic method is applicable to all of human behavior, even outside of traditional markets.

Economics is not Psychology, as it does not seek to explain the mental processes that lead people to certain conclusions, values, or decisions. Economics focuses on the trade-offs that an individual faces when making decisions. In fact, it is not a challenge to Economics that individuals rarely explicitly think in economic terms. As Alan Blinder writes, "a good pool player makes excellent use of the laws of physics without understanding them, and certainly without being able to articulate them." In the same way, many people do not recognize how their decision-making processes are influenced by economic considerations.

1.1.1 How Did We Get "Economics"?

The first usage of a term like economics comes from Ancient Greece with *Oeconomicus*—a combination of *oikos* meaning house and *nemein* meaning management. The latinized term *Oeconomus* would refer to the manager or

treasurer of the household. The first usage of *yconomy* in English came in 1440 in reference to the management of the monastery.

In the seventeenth century, the term was expanded from management of the wealth and resources of a household to the management of the wealth and resources of the nation as a whole with the introduction of the term political economy generally attributed to Antoine de Montchrestien in 1615. A century and a half later, Adam Smith in *An Inquiry into the Nature and Causes of the Wealth of Nations* (1776) wrote that

> Political oeconomy, considered as a branch of the science of a statesman or legislator, proposes two distinct objects: first, to provide a plentiful revenue or subsistence for the people, or more properly to enable them to provide such a revenue or subsistence for themselves; and secondly, to supply the state or commonwealth with a revenue sufficient for the public services. It proposes to enrich both the people and the sovereign.

Just over a century later, the language shifted again, as thinkers, led by Alfred Marshall, began to drop the *political* from political economy and shift to using Economics. Marshall opens his 1890 *Principles of Economics*:

> Political Economy or Economics is a study of mankind in the ordinary business of life; it examines that part of individual and social action which is most closely connected with the attainment and with the use of the material requisites of well-being.

However, Marshall seeks to transform the focus of the study from being "a science and an art" useful to the politician and rather to "a science, pure and applied." Marshall's preference for "Economics" and "Economic Science" over "Political Economy" was that it was a broader term that recognized that it can help determine not only the ends of the statesman, "but also what are the best methods of a broad policy devoted to that end." Especially as the field has adopted rigorous mathematical methods for modeling and adopted a broader approach to the topics it studies, the term economics has stuck.

While Marshall includes both a positive and normative element in Economics, today, the vast majority of Economics is Positive Economics. As shorthand, one can think of positive and normative economics as Economic Science and Political Economy, respectively. However, today, much of the normative work and Political Economy takes place under Political Science and Philosophy rather than under the discipline of Economics itself. But to the extent that this is true, that also means that economists cannot offer normative advice, except upon a pre-determined end. Advice must presuppose a given goal.

1.1.2 Economic Methodology

As a social science, Economics deals with the extremely complicated reality of purposive human activity. Because Economics is not Psychology, economic theory is not seeking to recreate the extremely complex mental activity behind the human action. Rather, it is seeking to model the outcomes that result. For this reason, in "Methodology of Positive Economics" (1953), Milton Friedman writes that a theory "is to be judged by its predictive power for the class of phenomena which it is intended to 'explain.'" The focus should not be on assumptions and "the conformity of these 'assumptions' to 'reality' [as] a test of the validity of a hypothesis..." Any assumptions for human activity will be extreme simplifications. Any model will be unrealistic. In fact, the entire point of modeling activity is that the model is simpler than the real world. Thus, "to be important, therefore, a hypothesis must be descriptively false in its assumptions."

Positive Economic Theory is composed of both a language (or structure of tautologies) and a body of substantive hypotheses (or testable theory) designed to abstract essential features of complex reality. It is critical to distinguish between the two. Hypotheses can be tested and accepted or rejected—though never proven—while tautologies are necessarily true. In "The Problem of Verification in Economics" (1955), Machlup reminds us that verification may find a hypothesis compelling but cannot prove a theory's "correctness, adequacy, or applicability." When it comes to our assumptions, there are many assumptions that "require no independent empirical tests but may be significant steps in arguments reaching conclusions which are empirically testable."

Richard E. Wagner (2020) writes that Economics is built on "mythological presuppositions" in the sense that "myths are fictions that nonetheless encapsulate mixtures of belief and knowledge that render the world of experience sensible and apprehensible to those who embrace those myths." At times, there will be fundamental disagreement due to a difference in presuppositions alone. For instance, Wagner writes that assuming a state of perfect competition suggests that reality is Pareto efficient, while assuming a state of imperfect competition suggests that it is inefficient. "It is, however, impossible to decide between these two possibilities without invoking mythological presuppositions." Sometimes, the myth takes the form of analogies. "Walras escapes this combinatorial arithmetic by invoking mythology to postulate the existence of an auctioneer who would continually cry out prices until equilibrium was established throughout society. This was mysticism at work and not plausible reasoning... [But the] embrace had a remote tinge of plausibility because market processes bear some family resemblance to

auctions."

While Economics can be tested empirically, as Leeson writes in "Economics Is Not Statistics and Vice Versa" (2019), fundamentally, "economic analysis is a theoretical approach, not an empirical one—a way of thinking... not a way of testing. When applied to observed phenomena, that way of thinking delivers context-specific predictions which may then be tested." This testing need not be with quantitative data. "In some cases, qualitative evidence is the best evidence we've got."

In his Nobel Prize acceptance speech, Hayek (1974) notes that the desire to quantify things can, at times, be useful for showing mathematical relationships, but we must recognize the limits of this activity. Despite "possess[ing] a fairly good 'qualitative' knowledge of the forces by which a correspondence between demand and supply in different sectors of the economic system is brought about... we can never produce statistical information which would show how much the prevailing prices and wages deviate from those which would secure a continuous sale of the current supply of labor." Hayek warns that the mathematical method "has led to the illusion, however, that we can use this technique for the determination and prediction of the numerical values of those magnitudes; and this has led to a vain search for quantitative or numerical constants." But even the founders of the mathematical method, such as Vilfredo Pareto, recognized that such calculation was not possible.

Within these parameters, there remain two types of reasoning. First is deductive reasoning where the goal is to develop logically sound arguments—conclusions are guaranteed to follow from premises—so that, if the premises are true, the conclusions are necessarily true. Thus, the key question is the plausibility of the assumptions. When engaging in deductive reasoning, theory simplifies reality to only the relevant facts. Second is axiomatic reasoning, which more common today under mathematization. Under axiomatic reasoning, the focus is on the internal coherence of the system rather than a focus on the realism of the assumptions. Equilibrium is a key way that economists ensure explanations are mutually compatible.

However, it is critical to not simply avoid reasoning altogether following an informal process that aligns with a gut feeling or expectation rather than building a model. This can, in practice, mean one is assuming *a priori* or as self-evident, the entire system of relationships—far more complex than rudimentary assumptions used to build models. This, in essence, relaxes the axiomatic requirement for logical consistency, but the result is that your conclusions do not *necessarily* follow from your premises. It is tempting when a model fails empirical verification to assume the opposite conclusion is true. But this is a faulty approach of reasoning. Perhaps confounders make the

model impossible to use when considering the real world, but this means that the conclusion is impossible to verify.

Some of the most important and most fundamental assumptions in economics relate to optimization: that individual agents optimize their utility and firms maximize profit. While some recent work seeks to deny that this happens in practice, this misses the feedback loop that creates conditions so that agents and firms do act *as if* they optimize. Due to competition, firms that do not maximize profit eventually die. Agents learn over time cause-and-effect, so, even if they don't have the mental computational power and knowledge to maximize, repeated actions ensure they are, at the very least, moving closer to achieving their goals and aligning with their preferences.

Buchanan (1969) argues that it is critical to *separate* the pure logic of choice from the scientific theory of behavior. The pure logic of choice, in general, seeks to explain results, while the science of behavior seeks to predict results.

The pure logic of choice is not concerned with empirical measurement, but seeks to explain why choices are made. Such logic is rooted in individual's subjective valuations of both costs and benefits, thus it is unobservable. The economist may consider a market where price ratios appear to differ from production ratios—for instance, the hunting of two animals takes the same amount of time, in general, but they do not exchange on a one-for-one ratio. A possible solution is that agents "have some non-pecuniary or non-economic arguments in their utility functions [so] marginal costs, as these affect choice behavior, may then not be the same as the simply observed ratios of labor time," for instance. But Buchanan notes that "in resorting to non-economic arguments in the utility function to rectify his falsified predictions [that price = marginal cost], however, the economist has shifted the whole analysis from a predictive to a non-predictive and purely logical theory."

On the other hand, the scientific theory of behavior can be used to make predictions. In predictive theory, we suggest that individuals act as if they respond to observable variables like prices and quantities. But this model "is not a theory of *choice* at all. Individuals do not choose; they behave predictably in response to objectively measurable changes in their environment." Buchanan's point here is that, when we say prices equal marginal cost, "*only prices have objective, empirical content*; neither the marginal evaluations of the demanders nor the marginal costs of the suppliers—the marginal evaluations of foregone alternatives—can be employed as a basis for determining prices. ... Prices are not brought into equality with some objectively determinable and empirically measurable phenomena, on either the demand or the supply side of the market." In predictive theory, we cannot predict *what prices "should" be* in an objective sense, because we

cannot measure the costs. What we can do is predict how changes in the environment shift the actions of given actors. The attempt by neoclassical price theory to conflate utility into this predictive theory is to confuse the distinction.

Even the use of utility functions is to cover over the actual action of choice. Buchanan (1964) writes that

> if the utility function of the choosing agent is fully defined in advance, choice becomes purely mechanical. No 'decision,' as such, is required; there is no weighing of alternatives. On the other hand, if the utility function is not wholly defined, choice becomes real, and decisions become real, and decisions become unpredictable mental events.

Thus, Buchanan suggests that it is better "to concentrate on 'exchange' rather than on 'choice.'"

1.1.3 Evolution and Ecological Rationality

Alchian (1950) suggests thinking in terms of a different paradigm altogether. Since economic actors live in a world of uncertainty, they do not know how to maximize profits (or utility), because all of the possibilities and the probabilities of outcomes associated with possible choices are not available to them. He argues that "where foresight is uncertain, 'profit maximization' is *meaningless* as a guide to specifiable action." Even in the simplest case of choosing between two probability distributions of profits, there is no "maximum" to choose from. We can only know the maximum *ex post*.

Instead, Alchian suggests considering activity within economic systems as analogous to the functioning of natural selection within biological systems. He supports a "method which treats the decisions and criteria dictated by the economic *system* as more important than those made by the individuals in it"—that is, how "successful and surviving firms are selected." The role of the economist under such a method is to "predict the more adoptable or viable types of economic interrelationships that will be induced by environmental change, even if individuals themselves are unable to ascertain them." The economist need not assume that individual actors have unrealistic perfect foresight or even that they are aware of their supply and demand curves. "These are concepts for the economist's use and not necessarily for the individual participant's, who may have other analytic or customary devices which, while of interest to the economist, serve as data and not as analytic methods."

In a free market system, it is the *realization* of profits that determines which firms survive and which ones disappear—those with losses, or negative profits. Given these survival conditions, so long as there is a group of participants that submits various combinations and organizations of firms and production for the system's selection and adoption, we will end up with the survival of the lower-cost firms and elimination of the higher-cost firms. Certainly, chance and luck play an important role in the *who* survives—and what ends up being the lowest-cost methods. Yet, on net, even if individual firms do not adjust their behavior, in light of changes in prices and technology, firms with "relatively lower cost positions" will have "a higher probability of survival." Even *without* individual foresight, the economist can predict how changes lead to adjustments in "the characteristics of the new *set* of firms"—whether actually new firms or simply the set of old firms that survive. Through the process of trial and error, the successful firms will be imitated. It is because natural selection is at work in the marketplace that "the economist can predict the consequences of higher wage rates, taxes, government policy, etc. like the biologist... predicts the effects of environmental changes on the surviving class of living organisms."

Too much focus on the individual's choice alone obscures how important the environment is for individual action. Compared to the study of Robinson Crusoe's decision making, "the fact of association requires that a wholly different, and wholly new, sort of behavior take place, that of 'exchange,' 'trade,' or 'agreement,'" writes Buchanan (1964). General Equilibrium and perfect competition's "flaw lies in its conversion of individual choice behavior from a social-institutional context to a physical-computational one."

> A market is not competitive by assumption or by construction. A market *becomes* competitive, and competitive rules *come to be* established as institutions emerge to place limits on individual behavior patterns. It is this *becoming* process, brought about by the continuous pressure of human behavior in exchange, that is a central part of our discipline... A general solution, if there is one, *emerges* as a result of a whole network of evolving exchanges, bargains, trades, side payments, agreements, contracts which, finally at some point, ceases to renew itself.

In fact, humans create institutions to solve collective action problems. The market itself is "the institutional embodiment of the voluntary exchange processes that are entered into by individuals in their several capacities." This evolutionary capacity to change the environment makes the discussion of so-called "market failures" absurd. While we might say that draining a swamp that is a breeding ground for mosquitos may suffer from a collective

action problem, "if the market institutions, defined so narrowly, will not work, they will not meet individual objectives [so] individual citizens will be led, because of the same propensity to search voluntarily for more inclusive trading or exchange arrangements. A more complex institution may emerge to drain the swamp." The environment shapes man, but man also shapes the environment.

Vernon Smith (2003), then, suggests that, when looking at economic history, we view human behavior through the lens of **Ecological Rationality**. This approach seeks the behavior of individuals based on their experience and folk knowledge. In this view, truth is discovered in the form of the intelligence embodied in rules and traditions that have formed out of the ancient history of human social interactions. The ecological concept of rationality asks from whence came the structure observed, and how they serve the fitness needs of those who unintentionally created them through their interactions.

Smith suggests that uncertainty is not the only challenge facing individual agents. There is also the serious cost of computation. Our unconscious mind does an excellent job at economizing on computational power. Our brains cannot evaluate preferences for every combination of tens of thousands of grocery items, because these mental processes are enormously opportunity-costly, so our conscious mind does not incur the cost. We "possess natural mechanisms for reminding ourselves of the brain's off-line activities and accomplishments." We can also learn very important things subconsciously. It is how we first learn language and social interaction. Before we assume that people are acting irrationally, we should reexamine the maintained hypotheses and ask what the subject's perception is of the problem that they are trying to solve.

Because thought resources are scarce, our brains delegate most decision-making to automatic processes that do not require conscious attention. Emergent arrangements must have survival properties that incorporate opportunity costs and environmental challenges. Such an order is necessarily based on trial and error in both cultural and biological evolution and rules emerge and a "social mind" is born that incorporates the interactions among all individuals that have survived to date.

1.2 Property Rights

1.2.1 Bundles of Rights

All economic analysis revolves around property rights. Broadly speaking, **Property Rights** are the rules of the game that define the actions available

to an actor. Property rights are appropriately plural, as property rights encompass a bundle of activities. Possessing one right does not imply possessing all rights. Among this bundle of rights is control of condition, use, net profit, and ownership. One can have the power to make decisions about the physical condition of a good, the use of said good, receiving income or bearing loss as a result of said use, and sell or give the ownership right to others. Property rights are instruments of society and help man form expectations which he can reasonably hold in his dealing with others. They specify how persons may be benefited and harmed and who must pay whom to modify the actions taken by persons.

All societies have some form of property rights. The question is not over whether property rights exist, but rather the extent to which they exist and who owns them. In some communities, there is a communal right to use the land or water and take for oneself the rewards of one's foraging. In the Soviet Union, individuals still owned personal property, but all capital goods—productive resources used to produce goods—were owned by the political authority. In the United States, one can possess the right to control the ownership of their land and general use and condition, but with a binding agreement to allow an outside authority, such as a Home Owners Association, to have a veto on decisions regarding the use and condition of the land (or part of it—such as the exterior and yard of a home).

Even within the bundle of rights, there are generally some limits and exceptions. In an emergency, one can take or use another's property to prevent a greater harm than the harm of violating one's property rights. There is also generally a rule against perpetuities. Individuals are allowed to bestow an inheritance—choosing how their property is distributed at their death—however, an individual cannot place a permanent restriction on all future uses of that wealth, such as land usage restrictions.

While property rights may seem like a concrete bundle, there is, in truth, almost an infinite partitioning of specific rights that one can have. For instance, consider a single space. Rights to use the physical land can be divided by height and space. If there is a building, one person can have a right to a first story and another to a second and third. A person can maintain the right of transferring ownership and general control of the land, but rent it out to tenants who then have a right to use the land for their purposes—yet the tenants cannot sell the land itself. Planes can have the right to fly over this land. There are also further usages. Consider the rights to broadcast radio frequencies across the space or to produce sound, smell, or other pollution that extends across the space. If someone possesses some property rights upstream a river, they may or may not have the right to take,

treat, or expel waste into the water. All of these actions, ultimately, impact those with property rights downstream.

Property rights can also include intangible property such as intellectual property, for which there are patents and copyright preventing use without permission.

The key focuses of Economics are consumption, production, and exchange. Exchange is necessarily the exchange of property rights. Production is taking things to which one has a property right and—to use the words of John Locke—mixing one's labor with it to produce something different for the purpose of using it oneself or exchanging it with others for something of greater value. Consumption involves using things for which one has the right of usage.

We generally think of property rights as bundled—a person with property rights has the rights to: make decisions about its condition, use, net profit, and ownership—and exclusive—only the person with the property right has the control over these facts. Sometimes we use ownership or property right as shorthand for this. Someone who owns a car has the right to paint it whatever color they want, drive it as frequently or infrequently as they want, receive income from renting it out or providing a service with it, and to rent it out or sell it. Nobody else has the right to modify or drive your vehicle or sell it without your permission. Similarly, someone who owns the right to pollute a certain quantity of emissions into the air may possess only this right—and nothing with land usage—but this involves the decisions of how these emissions go into the air, whether they do, when they do, and the ability to sell this right to someone else.

1.2.2 Enforcement of Rights

How are property rights enforced? In some cases, property rights are merely socially recognized conventions and enforced by etiquette and custom or ostracism. In other cases, they are legally established and enforced by the state. In some cases, there is a third option. Writing about the developing world in South America, Hernando de Soto discusses an "extralegal sector" which "represent[s] combinations of rules selectively borrowed from the official legal system, *ad hoc* improvisations, and customs brought from their places of origin or locally devised, and they are held together by a social contract supported by the community as a whole and enforced by authorities the community has selected." Even if the state itself is incapable, unable, or unwilling to enforce rights, people can come together to ensure this takes place. De Soto writes that in highly regulated environments where it could

take a decade to get legal approval to start a business or build a house, many have willingly turned into criminals to meet their needs and adopted an extra-legal system instead. He writes that "massive extralegality... is what always happens when governments fail to make the law coincide with the way people live and work."

Consider the recent developments in the United States, particularly in urban areas, towards locking up everyday items behind plexiglass at major retailers. For centuries, the lack of theft was due primarily to the cultural disgrace of being a thief—including religious, moral, and social condemnation. These cultural institutions allowed for goods to remain out in the open where they easily *could* be taken, but were rarely actually at risk of being stolen. Today, in areas where theft is common and these institutions have broken down—regardless of the specific cause—rather than turning exclusively—or even primarily—to the government to enforce property rights, these stores are investing in capital to prevent theft in the first place. Across the developing world where the state is weak and formal enforcement of property rights is minimal, protection racketeering—where a criminal organization is paid for agreeing not to harm individuals and to protect them from violence of others—is common.

1.2.3 Property Rights and Collective Action Problems

Property rights are not pre-existent, but emerge out of human actions and interactions. They form the environment in which individuals act; however, there is a mutual relationship: men create and modify property right arrangements. If the institutions do not serve their goals, people will devise alternatives. Buchanan (1964) foreshadowed the work of Elinor Ostrom (2010) suggesting that humans can solve collective action problems by adjusting their institutional arrangements.

Externalities are one type of collective action problem. An **Externality** is a cost or benefit not born by the agent making a decision. For instance, if I produce electricity but pollute the air with sulfur as a byproduct, the sulfur is not entering my monetary costs as a business-person, but the air impacts the quality of life for all those around me. In this case, there is a social cost or external cost not recognized by me, as the decision-maker when deciding how much electricity to produce and what price to charge my customers.

In "The Problem of Social Cost," Coase (1960) wrote that virtually all externality problems are merely ill-defined property rights. In the example of the electric generation polluter, if everyone around me is harmed significantly by the pollution, they would be willing to pay me to not pollute—paying up

to the external cost. While this may seem like a moral problem of one person harming others, Coase disagrees:

> The question is commonly thought of as one in which A inflicts harm on B and what has to be decided is: how should be restrain A? But this is wrong. We are dealing with a problem of a reciprocal nature. ... The real question that has to be decided is: should A be allowed to harm B or should B be allowed to harm A.

Simply banning me from producing electricity absolutely would pose a harm to me. Coase cites a real-life case that appears far more ambiguous of a doctor who added on to the back of his facility, which backed up to a baker who ran a loud confectionary machine. Before the doctor added on to his facility, there was no noise externality that bothered anyone. However, once he added on, the doctor complained that the baker should be liable for his noise. Yet, Coase notes, the doctor is likewise imposing a harm on the baker. Coase suggests that, absent transaction costs, in each case, the property rights will go to the party who values it more. If the doctor values the quiet more than the baker values the machine that is loud, he will be willing to pay the baker to stop or to use a different machine. If the baker is able to produce more value with the machine than the doctor values the quiet, then he will not accept the payment and the noise will continue, though perhaps the doctor might be willing to pay to insulate his room and try and remedy the problem somewhat.

Coase's key conclusion is that, when there are not transaction costs, the original assignment of property rights does not matter. If the baker has the right to produce whatever noise he wants or if the doctor has a right to peace and quiet, so long as they are permitted to negotiate, whoever values the right to noise more will end up with it. This will happen either by paying the other party for it—such as the baker paying the doctor for the cost of the noise or doctor paying baker to not make the noise—effectively moving the property right from the original owner to the other party or by not accepting the other party's offer leaving the property right where it was initially assigned. In the real world, there are transaction costs—sometimes, costs so big that the initial assignment does matter. The goal of policy, however, should be to reduce the costs of Coasean bargaining between two individuals.

So, in light of a society, which will necessarily have *some* property rights, a key question is the initial assignment. Who has property rights? Which things are publicly owned—whether by the state or by the community—or privately owned? As Alchian (1965) writes, the key difference between public and private ownership is the inability of a public owner to sell his share of

public ownership. Alchian writes that the lack of transferability has massive consequences. Private ownership makes each person's wealth more dependent upon his own activities, so people are incentivized to move towards acquiring more ownership in the area for which they can do the most to increase its profits and concentrate where they have a comparative advantage. For this reason, privately owned companies rarely need to impose restrictions like civil service, nepotism restrictions, tenure, or sealed bids for contracts, but public agents who bear less direct burden need to be disciplined by such rules.

Demsetz (1967) notes that it is important to distinguish between communal and state ownership. Communal ownership is rights exercised by all members of the community—like walking on a city sidewalk. Private ownership implies the right to exclude others from exercising the owner's private rights. State ownership implies that the state may exclude anyone from the use of a right.

Demsetz (1967) argues that property rights arise when the gains from defining and enforcing them is greater than the cost. Thus, "property rights develop to internalize externalities when the gains of internalization become larger than the cost of internalization." Communal ownership can lead to the externality commonly dubbed the "tragedy of the commons": overuse, overworking, and a lack of caring by others. This will tend to lead towards a focus on the present with little concern for the future, because resources are open to all. If an oil field is freely available, drillers will be focused on extracting as much as possible today before others get to it. Private ownership, however, can maximize the present value of the land and take into account alternative future streams of costs and benefits. "In effect, an owner of a private right to use land acts as a broker whose wealth depends on how well he takes into account the competing claims of the present and the future. But with communal rights, there is no broker, and the claims of the present generation will be given an uneconomically large weight in determining the intensity with which the land is worked." For this reason, the communal owners will not work to improve their resources.

Private ownership not only internalizes the tragedy of the commons, but it also reduces transaction costs between affected parties. Under communal land right, there is no reason for two individuals to bargain about land usage, but under private ownership with the right of excludability, this negotiation is possible. Demsetz writes,

> Suppose an owner of a communal land right, in the process of plowing a parcel of land, observes a second communal owner constructing a dam on adjacent land. The farmer prefers to have the stream as it is, and so he asks the engineer to stop his construction/ The engineer says, "Pay me to stop." The farmer

> replies, "I will be happy to pay you, but what can you guarantee in return?" The engineer answers, "I can guarantee you that I will not continue constructing the dam, but I cannot guarantee that another engineer will not take up the task, because this is communal property; I have no right to exclude him." What would be a simple negotiation between two persons under a private property arrangement turns out to be a rather complex negotiation between the farmer and everyone else. This is the basic explanation, I believe, for the preponderance of single rather than multiple owners of property.

The structure of ownership has likewise also evolved in this way. There are economies of scale in operating enterprises, but not all owners can negotiate every decision. They must delegate some power to "a small management group [that] becomes the *de facto* owners." However, the laws of partnership, making each partner-owner to a legal entity jointly and severally liable for the corporation, can be extremely unappealing, because a business failure makes the owners personally liable for the entire debt, even if the fault lay with management and not the owners themselves. Demsetz suggests that, in this way, "managerial *de facto* ownership can have considerable external effects on shareholders... [and] make it exceedingly difficult for entrepreneurs to acquire equity capital from wealthy individuals." For this reason, limited liability arose so liabilities can only cost the equity owner as much equity as they invested and no more. Likewise, unlike partnership law, modern corporations do not require the approval of all other owners to sell one's share, making it easier to both join and remove oneself from corporate ownership.

But Leeson and Harris (2018) argue that privatization is not necessarily always beneficial to society. One of the more important factors is whether decision makers are residual claimants—that is, whether they have "a personal claim to net changes in the aggregate money value of an organization's assets." When property decision makers are residual claimants, they bear the costs of their decisions and will only privatize when the effort to define and enforce property rights is perceived to be less than their benefits. However, when the property decision makers are not the residual claimants, "privatization occurs if it benefits decision makers personally, even if privatization destroys social wealth." The reason that the rules of private property vary across societies and histories is because of different tradeoffs.

> In a pastoralist society that inhabits an arid or semi-arid region, for example, it is often cheaper to realize scale economies when land is held in common than through costly market transactions under private property. Similarly, in such a society, common

> land holdings may insure individuals against drought risk more cheaply than arranging insurance through markets when land is owned privately. Like all choices, the choice between property regimes therefore involves tradeoffs. A property regime's effect on social wealth—the aggregate money value of all assets in a community—depends on how it negotiates these tradeoffs.

For this reason, top-down management of property rights systems pose a great risk.

> Suppose, for example, that state fiat transfers land decision authority to government officials in the country in which our society of pastoralists is located. Like most government officials... [they] are not residual claimants... [but] land privatization... presents administering officials a new, potentially lucrative source of rents to which they do have personal claims: payments by pastoralists, and perhaps others, in exchange for choice parcel allocations. Bribes paid to officials, for example, are pocketed by officials and benefit them directly. Further, land privatization may personally benefit administering officials indirectly, for example, by expanding the budgets of the government agencies that oversee land privatization.

Turning to a real-world example of the Maasai pastoralists in Kenya, they see this story play out and result in "economic decay" following land privatization. They speak favorably of recent developments in Kenya with the Community Land Act of 2016 which forbade the government from redistributing community land unilaterally and required two-thirds support of the community before privatizing. Beyond this, the empirical evidence of colonial imposition of property rights is dubious. "From Ghana to Afghanistan, Kenya to Madagascar, Rwanda to Uganda, Cambodia, and Peru, the creation of private land rights has led to questionable economic benefits at best and economic losses at worst." There can be great wisdom in the existing customs, and they very well may maximize social welfare. There is not necessarily a one-size-fits-all solution for all times and places.

1.2.3.1 Economic Systems as Different Rules

Capitalism is not unique in having property rights; every system maintains property rights. The key difference between capitalism, socialism, and communism is not the *existence* of property rights but who has property rights and what particular rights do they have.

Under capitalism, the rights to the vast majority of goods and, in particular,

the means of production, are held privately by individuals. Under socialism, these rights are held by the state, who sets rules regarding use and control. Under communism, these rights are held communally.

Some Marxists hold to a distinction between ownership of the means of production (private property) and articles of consumption (personal property), the latter of which they find defensible. Certainly it is true in all realized socialist economies, there exists a degree of personal property, even if the property originates under the state who then dictates how each specific article is to be distributed. While Engels foreshadowed this difference in *Anti-Dühring* (1878), Marx himself did not actually make such a distinction. Instead, for Marx, the state should own everything that could distinguish an individual by class and everything else could remain in private hands. But consider, any luxury, any good with higher quality than the lowest common denominator under such a standard could not survive—at least, not under *private* ownership.

1.2.4 Public and Common Goods

Some goods do not fit as neatly into the property right paradigm. **Public Goods** are commodities that are non-rival in consumption—that is, one person's consumption does not reduce another person's consumption of the good—and non-excludable—it is not possible to prevent anyone from consuming the good. A purely public good can fall victim to the free rider problem. Since it is not possible to exclude people, some people who benefit from consuming the good will nonetheless not pay, because they are not required to. Thus, the burden will be born by those who are willing to cover a disproportionate cost.

An example of a non-rival and non-excludable good is fireworks. Once one sets off fireworks, they can be seen and enjoyed from miles around. One person's enjoyment does not prevent another person from enjoying it. The same fireworks are available to be seen by all. A standard example of a non-excludable public good is national defense and particularly defensive missiles. Missiles protect an entire area, and if one person in that area does not pay, it is not possible to exclude a single person's home from protection. For a given area, such defense is also non-rival.

Goods generally fall on a spectrum between public and private. In "Public Good Definitions and their Institutional Context," Cowen (1985) argues that "nearly every good can be classified as either public or private depending upon the institutional framework surrounding the good and the conditions of the good's production." When considered across a broad area, national

defense is a public good, but "a single anti-ballistic missile... can only service a limited number of individuals... [it] cannot protect both New York and San Francisco." Likewise, the space in a particular lane of a particular road is certainly rival—as traffic shows—and yet, in light of the entire roadway system of the country, one driver in New Jersey has no impact on a driver in Nebraska. Cowen concludes that "the more limited and narrow our perception of what the relevant marginal unit of the good is, the closer it approaches being a private good."

Similarly, Cowen writes that the characteristic of "exclusion is always *possible*, it is just more or less costly. However, the costliness of exclusion is not a function of the nature of the good, but rather depends on how the good is supplied and at what levels it is produced and consumed." Originally, radio and later television were considered a non-excludable good. Today, cable and satellite radio and television have created technology to make the radio and television frequencies excludable. For a movie theater operator, it is possible to stop people from sneaking it, but it is easier to stop a few people than a large crowd.

Ostrom takes it one step further to separate out the rivalrous and excludable natures of public goods to separate into four different types of goods. Goods are rivalrous if there is a high subtractability of use. Something like food—where, if I eat it, you cannot—has a high subtractability of use. Something like a road, swimming pool, and fireworks all have low subtractability of use—though, at some level, each can become congested. If too many people are trying to come see the fireworks from the limited space they can be seen, it will be too crowded. A swimming pool is suitable for multiple people, but eventually the crowd becomes undesirably large. Goods that are easy to exclude include many tangible physical goods, while those that are more difficult to exclude include natural resources, like groundwater basins, the air, which can become polluted or filled with radio frequency signals.

The four-type division is as follows.

	Difficulty to Exclude		
Subtractability of Use		Low	High
	Low	Club Goods	Public Good
	High	Private Good	Common Pool Resource

Figure 1.1: Types of Goods

An example of a Club (or Congestion or Toll) Good is a road, which may be more difficult to exclude but have a low enough subtractability of use that we would expect multiple users of the same good. The most common

examples of Common Pool Resources are natural resources like groundwater, lakes, and forests.

If the Club Good is desired, we can expect the development of technology to make it easier to exclude individuals. Then, the questions become: how much of the good to produce and how many individuals to exclude. When it comes to Common Pool Resources, the classic view was the tragedy of the commons and overuse and overworking of the land. However, in her analysis of real world Common Pool Resources, Elinor Ostrom found that communities found ways to avoid this tragedy and develop rules—formal or informal—governing behavior of users and boundaries of the resource and tools to monitor and enforce these rules.

1.2.5 Philosophy on Property Rights

As property rights have been defined above, they constitute a social recognition of legitimate control and usage. Any appropriation of any natural resource for any reason—including survival—entails control and usage, so to outright deny the legitimacy of any property rights would be to deny the possibility of humanity's very existence. However, this does not resolve all questions. After all, why allow private property—the existence of some natural resources that are available to some individuals but excluded from others—instead of communal (or state) ownership? Is this exclusion a ground for injustice? What have the key philosophers held about property rights?

In Aristotle's *Politics*, written in the late fourth century BC, Aristotle holds that some sort of property "is evidently given by nature itself to all animals, both immediately from birth and when they have reached completion," since "it is impossible to live without sustenance," and in order to sustain oneself, one must appropriate for one's sole use resources. As mankind has developed from nomadic life to living in a settled community, there arises "the art of acquiring those goods a store of which is both necessary for life and useful for the community of a city or household... these things that make up genuine wealth." Once humans begin making goods that can hold value, Aristotle holds that "every possession has a double use. ... In the case of shoes, for example, one can wear them or one can trade them." All items become useful as either items of barter or use. Money arises to make commerce easier, but, in his view, this is a natural outgrowth only in so far as the money is used as an intermediary between goods and not as an end in itself.

Aristotle notes that every settled community shares some things. "Now it is evident that to share in nothing is impossible; for the regime is a certain sort of community, and it is necessary in the first instance to share a location..."

But the question arises: should people share everything or "is it better to share in some but not in others?" After all, in Plato's *Republic*, there is a proposal "that children and women [wives] and property should be [held in] common." Yet Aristotle suggests that this would be to "seek to unify the city excessively." Foreseeing the problem identified millennia later as the tragedy of the commons, Aristotle adds that "What belongs in common to the most people is accorded the least care: they take thought for their own things above all, and less about things common, or only so much as falls to each individually." For Aristotle, communally held property is likely to encourage conflict and private property allows for the virtue of liberality.

> In general, to live together and share in any human matter is difficult... This is clear in communities of fellow travelers, most of whom are always quarreling as a result of friction with one another over everyday and small matters. ... Having property in common involves, then, these and other similar difficulties... Dividing the care of property will cause them not to raise these accusations against one another, and will actually result in improvement, as each applies himself to his own; and it will be through virtue that 'the things of friends are common,' as the proverb has it , with a view to use. ... [I]n finely administered cities... everyone has his own property, but he makes some of it useful to his friends, and some he uses as common things... It is evident, then, that it is better for property to be private, but to make it common in use. ... Moreover, it is a very pleasant thing to help or do favors for friends, guests, or club-mates; and this requires that property be private. ... For it is precisely those who possess things in common and share whom we see most at odds, not those who hold their property separately.

Writing in the mid-13th century AD, Thomas Aquinas expressed agreement with Aristotle's view in his *Summa Theologica*. Aquinas adds that "man has a natural dominion over external things, because, by his reason and will, he is able to use them for his own profit, as they were made on his account." He writes,

> it is lawful for man to possess property. Moreover this is necessary to human life for three reasons. First because every man is more careful to procure what is for himself alone than that which is common to many or to all: since each one would shirk the labor and leave to another that which concerns the community, as happens where there is a great number of servants. Second, because human affairs are conducted in more orderly fashion if

> each man is charged with taking care of some particular thing himself, whereas there would be confusion if everyone had to look after any one thing indeterminately. Third, because a more peaceful state is ensured to man if each one is contented with his own. Hence it is to be observed that quarrels arise more frequently where there is no division of the things possessed.

Nonetheless, Aquinas suggests that "the division of possessions is not according to the natural law, but rather arose from human agreement which belongs to positive law." Such positive law can arise "in two ways: first by private agreement, as that which is confirmed by an agreement between private individuals; second, by public agreement, as when the whole community agrees that something should be deemed as though it were adjusted and commensurate to another person, or when this is decreed by the prince who is placed over the people, and acts in its stead."

For David Hume, man is a unique animal, whose special characteristics give rise to the development of society and, then, property.

> [for] the sheep and ox... their appetites are moderate, and their food is of easy purchase. In man alone, this unnatural conjunction of infirmity, and of necessity, may be observed in its greatest perfection. not only the food, which is required for his sustenance, flies his search and approach, or at least requires his labor to be produced, but he must be possessed of clothes and lodging, to defend him against the injuries of the weather, though to consider him only in himself, he is provided neither with arms, nor force, nor other natural abilities, which are in any degree answerable to so many necessities. Tis by society alone he is able to supply his defects... [and] all his infirmities are compensated.

Once men come together in society, however, Hume suggests that "the principal disturbance in society" arises among goods that are able to quickly and easily "transition from one person to another." A remedy naturally arises by convention as people come to agree to "leave every one in the peaceable enjoyment of what he may acquire by his fortune and industry," providing for the "stability in his possessions." People agree to these rules of property, because "I observe that it will be for my interest to leave another in the possession of his goods, *provided* he will act in the same manner with regard to me."

John Locke, on the other hand, holds that property is not merely conventional. Since men need sustenance, how is it that the world, which, in the beginning, belongs to mankind in common, can be appropriated by each

person individually? For Locke, this begins with self-ownership: "every man has a property in his own persons... the labor of his body and the work of his hands, we may say, are properly his. Whatsoever, then, he removes out of the state that nature hath provided, and left it in, he hath mixed his labor with , and joined to it something that is his own, and thereby makes it his property. It being by him removed from the common state nature hath placed it in, it hath by this labour something annexed to it that excludes the common right of other men." For instance, a deer roaming the forest or fruit on a wild tree were both theoretically held in common, but once killed by a hunter or picked by the forager, it belongs to them. At the same time, in nature, goods spoil quickly, so one is not to take more than one can use. If you pick all the fruit off of a patch of trees and can only eat a few, you have ruined and wasted what was commonly available. This type of hoarding is impermissible. Likewise, when appropriating natural resources, they must ensure that "there is enough, and as good, left in common for others." No one person can seek to appropriate an entire continent, which would likewise be to "hoard up more than he could make use of." However, the invention of money as a non-perishable store of value and then investment in non-perishable goods has allowed for the accumulation of wealth without concern of depriving others of the subsistence needed for life.

Locke suggests that, outside the state of nature, some restrictions can be made, however. "In land that is common in England or any other country where there is plenty of people under government... no one can enclose or appropriate any part without the consent of all his fellow commoners, because this is left common by compact, i.e., by the law of the land." Thus, while "labour, in the beginning, gave a right of property wherever anyone was pleased to employ it upon what was common," absent a state, there is no authority to objectively vindicate claims of justice and punish violations. Each individual is his own judge, which is difficult, and *de facto* might makes right. So, people join together to "take sanctuary under the established laws of government and therein seek the preservation of their property."

1.3 Knowledge, Coordination, Prices, and Cost

1.3.1 The Price System

Economics has been defined as the study of constrained decision-making. In general, both individuals and broader societies must decide how to allocate

limited resources among a near-infinite quantity of competing uses. An economic system must communicate both information about the things—primarily goods and services—desired by individuals and the real constraints on the availability of those things. Put differently, the economic system must solve for the appropriate output in light of desires and relative scarcity. Thus, the fundamental economic problem is one of knowledge.

As Hayek writes in "The Use of Knowledge in Society" (1945),

> The economic problem of society is not merely a problem of how to allocate 'given' resources... It is rather a problem of how to secure the best use of resources known to any of the members of society, for ends whose relative importance only these individuals know—to put it briefly, a problem of the utilization of knowledge which is not given to anyone in its totality.

In a society where the economy is governed by the state, output can be determined by voting and administrative setting, but it would be next to impossible for this to meet the optimum in a single period, much less across time. In a free market economy with property rights, the price system guides the economy.

A **Price** is a relative trade-off. In a barter economy, it is how much of one good you must give up to receive another. In a money economy, it is how much money you must give up for (or receive for) a good. Prices are necessarily, then, ratios. Prices have nothing to do with some vague intrinsic value possessed by a good, rather they communicate relative scarcities—the difference between what is available and what is desired. Given there is only so much of a given good, prices ensure that goods go to individuals with the highest willingness to pay, which is a key recognition of efficiency. Prices direct resources towards their most valuable uses.

The Price System solves three problems. First, it solves a communications problem. Ideas of how to innovate and use a given set of resources best is dispersed throughout the population. Only each individual knows their own desires. These ideas must be communicated in some way. By having producers bid for resources and consumers bid for final outputs, each individual subjective valuation is communicated. There is also a significant amount of tacit knowledge that could not be fully articulated or transmitted, even if one wanted to do so. The circumstances on the ground can also change rapidly and very locally leaving "circumstances of time and place" in which there must be some decisions "left to the 'man on the spot.'" However, each "man on the spot cannot decide solely on the basis of his limited but intimate knowledge of the facts of his immediate surroundings. There still remains

the problem of communicating to him such further information as he needs to fit his decisions into the whole pattern of changes of the larger economic system."

Second, it solves a dynamic computational problem. As technology advances and as real shocks—from weather and drought to war and pandemics—affect resources and our abilities to use them, the economy is in constant flux, even if, at the aggregate, our output statistics seem relatively static. Multiple inputs can be used to produce the same or similar output in different ways. Which input should be used for what task? How much of each input? As circumstances change, even if all the relevant data could be quantified and communicated to a central authority, the next step would require actually computing values for production.

Finally, the market is solving a generative problem. The knowledge that is required for us to negotiate trade-offs between various activities and uses of resources is only generated within the process of competitive bids and asks. It is not merely that all of the possible knowledge needed to be accessed by a central planner is so difficult for the planner to access; Hayek suggests that much of this knowledge does not *exist* without the market process. For instance, when the price of an input rises significantly, this signals that innovators should investigate alternative ways to meet their needs for this input. These alternative ways may not—and probably will not—already be known, but the price signal will induce individuals to find them. Consumers did not already know that they wanted smartphones until they were invented, and it is often experience with the product that gives consumers a better idea of what features they want and appreciate and those to which they are indifferent.

For Hayek, the price system is ultimately a coordinating system. It allows us to attach to each scarce resource a numerical index in which the relevant information is concentrated. For instance, consider that tin is made more scarce—it does not matter whether it is due to new demanders of tin or an elimination of some tin supply—all the users of tin need to know is that some of the tin they used to consume is now more profitably employed elsewhere, so they must economize tin. The change in price communicates all that they need to know. The alternative to price coordination is to ration goods based on force, violence, political power, or first-come first-served.

1.3.2 Cost

A **Cost** is the foregone opportunity. If I spend my money on a new house, I cannot spend that money on a new car, clothes, or food. If I spend my time

hunting, I can't spend that same time building tools or fishing. The cost of an hour hunting might appear to be 2 hammers and 5 fish, then.

An **Opportunity Cost** looks at the second-best option as evaluated by the decision-maker. The reason spending money or time is a cost is because of the potential valuable alternative. If you were not spending it doing one thing, what *would* you be doing? What's the next best alternative? In this sense, cost is clearly subjective, even when we're dealing with what appears to be objective numbers, like a money price. How I would use the money, if not on this particular good, will probably be different than how you would use it. Thus, our analysis is rooted in preferences, which is where studies of microeconomics often begin. If we think of costs in this opportunity cost sense, Coase (1938) writes that "to cover costs and to maximize profits are essentially two ways of expressing the same phenomenon."

Prices are useful at communicating (relative) costs. The classical economists dating back to Adam Smith noted that the exchange ratio between two goods should be related to the opportunity cost of production. If the exchange ratio differed from the opportunity cost—for instance, if 1 beaver traded for 2 deer, but I am able to catch 5 beaver in the same amount of time as it takes to hunt 2 deer—then arbitragers would take advantage of this and produce the scarcer good until the opportunity costs were equal. For this reason, for most goods, exchange value tends to equal the cost of production. This doesn't work when supply is fixed and new production is not possible.

After the Marginal Revolution, which theorized that value is ultimately subjective, cost *also* should be recognized as subjective in the pure logic of choice. Thirlby (1946) writes:

> By deciding to take the preferred course, he [an agent] incurs cost—he displaces the alternative opportunity. The cost is not the things—e.g., money—which will flow along certain channels as a result of the decision; it is the loss, prospective or realized, to the person making the decision, of the opportunity of using those things in the alternative courses of action. This cost cannot be discovered by another person who eventually watches and records the flow of those things along those channels. Cost is not something which is objectively discoverable... it is something which existed in the mind of the decision-maker before the flow began, and something which may quite likely have been but vaguely apprehended.

Buchanan (1969) writes that there is a "subtle but essential distinction between cost in the act of choice and cost in the predictive theory of economic

behavior." He suggests that choice-influencing cost—cost under the pure logic of choice—is composed of 6 components:

1. Cost must be borne exclusively by the decision-maker, and cannot be shifted to or imposed on others.
2. Cost is subjective, existing in the mind of the decision-maker and nowhere else.
3. Cost is based on anticipations and is necessarily forward-looking.
4. Cost can never be realized, because that which is given up cannot be enjoyed.
5. Cost cannot be measured by someone other than the decision-maker, because there is no way to directly observe that subjective experience.
6. Cost can be dated at the moment of decision or choice.

Choice-influencing cost is defined in the utility space. After making a choice, an agent learns and updates their model of the world and their future behavior by analyzing costs. If the agent subjectively evaluates that the sacrificed enjoyment was greater than the realized enjoyment, this will influence later assessments and future choices. Buchanan notes that prices themselves are not cost. They are useful in calculation of costs and evaluating trade-offs, but cost is fundamentally subjective. Choice-influencing cost is "based on anticipations," and "represents the anticipated utility loss upon sacrifice of a rejected alternative. Because utility functions are necessarily personal, cost is tied directly to the chooser and cannot exist independently of him." This also means that since "costs carry no objective meaning [they] cannot be employed as criteria for determining prices in some welfare or efficiency sense."

As Buchanan has noted, we can also use cost in predictive theory of behavior. In this context, we have a choice-influenced cost which is *ex post* and "not personally tied to the chooser or decision-maker." Choice-influenced cost can appear more objective and be defined in the commodity space. For instance, money receipts represent the cost stream—the resulting physical consequences—that is realized by choices to pursue a particular decision. But this doesn't explain why an individual makes the decision that they do; it's the effect, not the cause. The choice to pursue a particular project leads to money outlays.

In predictive theory, often this "cost is measured directly by prospective money outlays" and that "the market value of the alternate product that might be produced by rational reallocation of resource inputs to uses other than observed... is reflected in the market prices for resource units."

This cost is simply the aggregation of subjective valuations of "potential alternatives as judged by others rather than by the chooser himself." When making predictions, we can model individuals as acting *as if* they maximize wealth and minimize cost to predict the direction of behavior in light of an environmental change.

We can only use money cost and money revenue under the assumptions that people economize, which is how equilibrium is defined—when each agent is economizing and, to use Hayek's equilibrium terminology, "when the plans of participants in the economic interaction process are mutually satisfied." But this means that money outlays can only indirectly represent costs and only in equilibrium with no uncertainty.

> If the whole economy is not operating at full competitive equilibrium, profits-losses may occur and, hence, observed outlays cannot be taken to reflect foregone opportunities of actual decision-takers in any general setting. In full equilibrium, on the other hand, observed outlays directly represent the maximum contribution of resources in different uses. ... If, on the other hand, the individual incorporates non-pecuniary or non-economic considerations in his decision, if there are profits to be secured elsewhere than in the activity in question, if discrete rather than marginal adjustments are possible, then objectively measured marginal outlay is not a veritable expression of genuine opportunity costs, because these "ifs" may represent inhibitions upon choice behavior, which are not susceptible to objective measurement. ... In the face of uncertainty, the evaluation of alternatives by the actual decision-taker may differ from the evaluations of any external observer, even if the qualifying conditions are met.

We cannot, then, use this objective sort of cost of predictive theory in a normative sense. Predictive theory suggests that people behave as if they economize. "There is no implied presumption that men *should* behave economically." Thus, Buchanan argues that much of welfare economics is mistaken in its normative claims which rely on some objective metric of cost.

> the applied economist and the welfare theorist alike accept the behavior of *Homo economicus* as a value criterion. In their zeal to apply economic theory not to an analysis of institutional interactions but to real choice, they indirectly propose that decision-takers, singly or in the aggregate *should* minimize objectively measurable outlays.

They confuse the assumptions of choice that people act as if they minimize

costs and maximize utility with the normative claim that they should do so.

1.4 Tying It All Together: Two Approaches to Economic Analysis

In Kant's *Critique of Pure Reason* (1781), he formalized the analytic/synthetic distinction. In the traditional distinction, analytic propositions are true by definition—e.g. "all bachelors are unmarried"—while synthetic propositions are statements about the world. Analytic propositions relate ideas to each other, while synthetic propositions connect ideas and facts in the real world. Likewise, Economic analysis includes the realm of the analytic and the synthetic.

Analytic economics focuses on choice—how individuals are making decisions. This is the realm of explaining why various institutions or interactions take place. Individuals acting with subjective valuations are assumed to maximize. Under analytic economics, as Leeson (2019) writes that "every observed institution is efficient." If individuals maximize, the result is maximization by definition. Perhaps there are transaction costs or cognitive costs that must also be economized. Following Buchanan's (1969) critique of welfare economics, Leeson writes that much of the criticisms of market failures or inefficiencies by economists "is phantom, conjured only by judging what's achieved under *actual* constraints against what's achievable under *fictitious* constraints that are less severe." Oftentimes, the truth is economists "*disapprove*" of various real-world situations and "want economics to ratify [their] disapproval," but analytic economics is not a tool permissible for this sort of critique.

This is not to say the real world lacks possibility of improvement. Opportunities for economic and political entrepreneurship exist. However, such opportunities require entrepreneurs to shift the relative costs and benefits—perhaps by reducing information, transaction, or switching costs. Positive sum change and gains-from-trade opportunities must be sought. Analytic economics provides no normative standard for evaluating policy. The only objective way to measure whether improvement was possible is whether it is adopted voluntarily by all, suggesting a Pareto improvement.

Under analytic economics, there is no person with a willingness to pay greater than another person's willingness to accept once accounting for *all* costs—information, transaction, bargaining, etc—but entrepreneurship is extremely powerful by bearing these costs or identifying and implementing ways to lower these costs.

Some key works in this area include the above section on property rights and transaction costs including Coase (1960), Alchian (1965), Demsetz (1967). Analytics can extend knowledge as the logic is applied to more situations, such as explaining why property rights may develop, why the firm may develop (Coase 1937), why there is irrational behavior (Becker 1962, Caplan 1999), and why institutions may change over time. The best applied use of analytic economics is comparative institutional analysis.

Synthetic economics focuses on human behavior—rather than choice—and how it manifests itself in the real world. Synthetic economics looks to the emergent order. How does nature and nurture, real constraints and artificial institutional constructions shape behavior?

Looking to individual behavior, Vernon Smith (2003) writes that people have ecological rationality using "intelligence embodied in the rules, norms, and institutions of our cultural and biological heritage that are created from human interactions but not by deliberate human design. People follow rules without being able to articulate them, but they can be discovered." Thus, he warns against accusing humans of irrationality, but suggesting we first ask "the subjects' perception of the problem that they are trying to solve."

Looking to the dynamics of aggregated behavior, Alchian (1950) points to the incentives created by the institution environment. Over time, the firms making profits will succeed and those with losses will fail. As individuals find positive-sum arbitrage opportunities (Pareto improvements), more plans can become coordinated and mutually satisfying. Based on this incentive structure, we can use supply and demand curves for instance to model the outcomes of agent's adjustments, even if the actual adjustments and the actual activity of the individual behavior is rooted in far different terms. This is a testable theory. For this reason, we can focus on whether the theory is actually predictive, as Friedman (1953) writes. In analyzing policy, are we considering the appropriate variables? We can seek to estimate values such as elasticities and various coefficients. However, we must be cautious, since measuring parameters that hold constant under alternative policies is quite a difficult task, as Lucas (1976) notes.

References

Alchian, Armen A. 1950. "Uncertainty, Evolution, and Economic Theory" *Journal of Political Economy*, Vol. 58, No. 3, pp. 211-221.

Alchian, Armen A. 1965. "Some Economics of Property Rights," *Il Politico*, Vol. 30, pp. 816-829.

Aristotle. 2013. *Politics*, translated and edited by Carnes Lord, University of Chicago Press.

Aquinas, Thomas. 2020. "Part Two of the Second Part, Questions 1 to 91," *Summa Theologica*, translated by Laurence Shapcote, Aquinas Institute.

Becker, Gary S. 1993. "The Economic Way of Looking at Life," *Coase-Sandor Institute for Law & Economics Working Paper*, No. 12.

Buchanan, James M. 1964. "What Should Economists Do?" *Southern Economic Journal*, Vol. 30, No. 3, pp. 213-222.

Buchanan, James M. 1979. "Is Economics the Science of Choice?" *What Should Economists Do?*, Liberty Fund, pp. 39-63.

Buchanan, James M. 1979. "Natural and Artifactual Man," *What Should Economists Do?*, Liberty Fund, pp. 93-112.

Buchanan, James M. 1999 [1969] *Cost and Choice: An Inquiry in Economic Theory*, Liberty Fund.

Coase, Ronald H. 1973 [1938]. "Business Organization and the Accountant," *L.S.E. Essays on Cost*, edited by J.M Buchanan and G.F. Thirlby, London School of Economics and Political Science, pp. 95-132.

Coase, R.H. 1960. "The Problem of Social Cost," *Journal of Law and Economics*, Vol. 3, pp. 1-44.

Cowen, Tyler. 1985. "Public Goods Definitions and their Institutional Context: a Critique of Public Goods Theory," *Review of Social Economy*, Vol. 43, No. 1, pp. 53-63.

Demsetz, Harold. 1967. "Toward a Theory of Property Rights," *American Economic Review*, Vol. 57, No. 2, pp. 347-359.

Friedman, Milton. 1953. "Methodology of Positive Economics," *Essays in Positive Economics*, University of Chicago Press.

Hayek, Friedrich August von. 1989 [1974]. "The Pretence of Knowledge," *American Economic Review*, Vol. 79, No. 6, pp. 3-7.

Hayek, F.A.. 1945. "The Use of Knowledge in Society," *American Economic Review*, Vol. 35, No. 4, pp. 519-530.

Heyne, Paul. 2008. *"Are Economists Basically Immoral?" and Other Essays on Economics, Ethics, and Religion.* edited by Geoffrey Brennan and A.M.C. Waterman, Liberty Fund.

Hume, David. 1896 [1740]. "Book III, Part II: Of Justice and Injustice," *A Treatise of Human Nature*, edited by L.A. Selby-Bigge, Clarendon Press, pp. 477-573.

Machlup, Fritz. 1955. "The Problem of Verification in Economics," *Southern Economic Journal*, Vol. 22, No. 1, pp. 1-21.

Leeson, Peter T and Colin Harris. 2018. "Wealth-Destroying Private Property Rights," *World Development*, Vol. 107, pp. 1-9.

Leeson, Peter T. 2019. "Logic is a Harsh Mistress: Welfare Economics for Economists," *Journal of Institutional Economics*, Vol. 16, No. 2, pp. 145-150.

Leeson, Peter T. 2020. "Economics is Not Statistics and Vice Versa," *Journal of Institutional Economics*, Vol. 16, pp. 423-425.

Locke, John. 2005. *The Selected Political Writings of John Locke*. edited by Paul E. Sigmund, Norton Critical Edition.

Lucas, Robert E. 1976. "Econometric Policy Evaluation: A Critique," *Carnegie-Rochester Conference Series on Public Policy*, Vol. 1, No. 1, pp. 19-46.

Lucas, Robert E. 2011. "What Economists Do," *Journal of Applied Economics*, Vol. 14, No. 1, pp. 1-4.

Ostrom, Elinor. 2010. "Beyond Markets and States: Polycentric Governance of Complex Economic Systems," *American Economic Review*, Vol. 100, No. 3, pp. 641-672.

Smith, Adam. 1904 [1776]. *An Inquiry into the Nature and Causes of the Wealth of Nations*, Volume 1. edited by Edwin Cannan. Methuen & Co.

Smith, Vernon. 2003. "Constructivist and Ecological Rationality in Economics," *American Economic Review*, Vol. 93, No. 3, pp. 465-508.

Thirlby, G.F. 1973 [1945]. "The Subjective Theory of Value and Accounting 'Cost'," *L.S.E. Essays on Cost*, edited by J.M Buchanan and G.F. Thirlby, London School of Economics and Political Science, pp. 137-161.

Wagner, Richard E. 2020. "Public Policy as the Political Calculation of Economic Value," *Macroeconomics as Systems Theory*, Palgrave MacMillan.

Chapter 2

Consumption

What do we mean by consumption? As we open this chapter on consumption, it is important to note what we are referencing. Consumption can, of course, mean taking and using various resources physically. Cotton and wool can be gathered and transformed into clothing that is used. Crops can be eaten and provide nutrition. Wood, bricks, and mortar can be used to build a structure used as a shelter from the elements. However, consumption is not *merely* about this.

Boudreaux (2023) writes that the focus of consumption is on ends.

> On economists' understanding, consumption includes not only the likes of eating, watching the Super Bowl, and bedecking oneself with jewelry. Consumption includes also the satisfaction we get from living in safe and thriving communities, from enjoying our jobs, from interacting pleasantly with our neighbors, our fellow church members, and our co-workers. It includes the joy and knowledge derived from traveling, reading, and attending public lectures. Consumption, as economists understand it, is that wide array of actions the experience of which we anticipate will directly make our lives more fulfilling.

The study of consumption, then, is the study of how humans make the most of the opportunities to pursue their given ends. Thus, we begin our analysis with formalizing the role of preferences in human decision-making.

2.1 Preferences

2.1.1 Defining Preferences

Individuals have preferences between some different outcomes. If the individual truly had no preferences at all, action would be impossible. When an individual acts, we say that they have *revealed* their preferences—to prefer the action they took as opposed to the alternative.

Preferences are fundamentally a qualitative ordering. Different individuals may come up with a different rank ordering. We can compare preferences between two single goods or, more commonly, baskets or bundles of goods. Consider that even things that may appear to be single goods can have bundled elements. For instance, when you purchase a car, the car has a specific exterior color, interior color, stereo, radio and screen, seats, features like blind spot monitoring, adaptive cruise control, a backup camera, keyless start, automatic start-stop, heated seats, and status—whether as a luxury brand or not. More practically, it has a specific miles per gallon or charging range, drivetrain, and reputation for reliability. When you purchase a car, all of these features and facts are bundled into one good and preference relations between cars will aggregate all of these individual features—of which you may prefer some to others. Of course, the actual preferences between different bundles may depend on numerous factors.

Lancaster (1966) argues that when consumers are making preference rankings and deciding which goods to acquire, what is really demanded is not the goods themselves, but the services they provide. He writes that "consumption is an activity in which goods, singly or in combination, are inputs and in which the output is a collection of characteristics." For instance,

> a meal (treated as a single good) possesses nutritional characteristics, but it also possesses aesthetic characteristics, and different meals will possess these characteristics in different relative proportions. Furthermore, a dinner party, a combination of two goods, nutritional, aesthetic, and perhaps intellectual characteristics different from the combination obtainable from a meal and a social gathering consumed separately.

Thus, Lancaster separates the good itself from the characteristics that arise from it (or the services that it provides).

There are three types of preference relations:

- Strong Preference – $x \succ y$ meaning x is preferred to y
- Weak Preference – $x \succsim y$ meaning x is at least as good as y

- Indifference – $x \sim y$ meaning the individual does not have a preference between x or y, but would be equally satisfied with either. $x \sim y$ is equivalent to *both* $x \succsim y$ and $y \succsim x$.

In order to come to any conclusions based upon preference relations, we must make a few assumptions:

1. Completeness – an individual can compare any two baskets of commodities with the potential outcomes of $x \succsim y, y \succsim x$, or $x \sim y$.
2. Reflexivity – each good is only as good as itself $x \sim x$.
3. Continuity – small changes in bundles do not result in drastic jumps in preferences.
4. Transitivity – for an individual, if $x \succsim y$ and $y \succsim z$, then $x \succsim z$.

Preferences that follow these rules can be modeled according to a utility function.

Consider some preferences that fall outside of these rules. Lexicographic preferences in which consumers strictly prioritize one good (or attribute) over another, regardless of the quantities, cannot be mapped with a utility function. With such preferences, no amount of a second good could compensate for any loss in the first good. If an agent is searching for a job, but only looks for work in a specific location, such that *no* salary will induce him to consider a job outside of his specified region. In this case, preferences are not complete. Another slightly different scenario to think about this is that the agent looks at the location first when choosing a job, *then*, if there are competing offers with equally good locations, he considers salary (and other benefits), but he *strictly* considers location first. Such preferences cannot be mapped with a real-valued function.

2.1.2 Preferences as Static

When individual action changes, it can be tempting to suggest that the new behavior is due to a preferences shift in preferences. On the contrary, the thesis of Stigler and Becker (1977) is that "widespread and/or persistent human behavior can be explained by a generalized calculus of utility-maximizing behavior without introducing the qualification 'tastes remaining the same.'" We need not assume that addictive behavior *changes* preferences. For instance, a consumer purchasing illegal drugs, such as heroin, is likely demonstrating their preference for a state of "euphoria." However, "an increase in the consumption of current euphoria raises the cost of producing euphoria in the future," thus we need not assume changes in taste when we see the amount consumed grows. If demand for euphoria is relatively inelastic,

we would expect that "the use of heroin would grow with exposure at the same time that euphoria fell." A person may seek to appreciate music—a beneficially addictive good. Such a person, then, invests in human capital in the early stages of his consumption. Over time, the marginal costs of an additional unit of output falls "as skill and experience in the appreciation of music are acquired with exposure."

When individual action remains static in the face of changing constraints, it can be tempting to suggest that preferences might have changed here too? Stigler and Becker, however, point to search costs as a key factor. After a buyer has found a commodity that they like well enough and have experience with it, they have to judge whether the potential benefits of a cheaper and better alternative are worth the cost of searching to try and find it. We may see consumers continuing to buy from the same supplier even as their price rises, even above the market price. However, we would expect that as the price of the commodity increases, the number of purchases between each search will fall.

2.1.3 Preferences as Chosen

Buchanan (1979) writes that for both animals and humans we can design experiments to measure some preferences and try to predict future action and measure a utility function. However, "our predictions about man must always be less accurate than our predictions about animals," because, while animals are "natural," humans are both "natural and artifactual." Humans are natural in the sense that they are "bounded by natural constraints," but artifactual in the sense that "we are and will be, at least in part, that which we make ourselves to be [and w]e construct our own beings..." For instance, in education, a person is able to invest, not merely in future revenue streams, but also "in becoming the different person that he must become as he acquires knowledge and wisdom."

For Buchanan, an individual's "natural" limitations include both what is typically thought of as biology as well as "culturally evolved rules of conduct or modes of behavior" and an individual's own history and past experiences. However, through the veil of these natural limitations, "recognizing the possible prospects that remain open to him through his remaining life span, and ranking these prospects in some way, the individual remains capable of achieving the prospect that he chooses."

> I have used the term *artifactual* here precisely for the purpose of allowing some recognition of the basic constraints of human nature while, at the same time, allowing for wide areas of

> choice within these constraints, areas within which we can, and do, construct ourselves as individuals, from the base largely constructed for us by our forbearers.

The only way to account for people quitting smoking or dieting when holding preferences constant—and when constraints appear to also be constant—is by "allow[ing] for some recognition that persons imagine themselves to be other than they are and that they take action designed to achieve imagined states of being. The desire for liberty itself is rooted in this artifactual nature. "Man does not want liberty in order to maximize his utility, or that of the society of which he is a part. *He wants liberty to become the man he wants to become*." In fact, unless we allow for this artifactual nature of man, we cannot justly hold him responsible for his own actions, as they are merely responses to stimuli.

Cowen (1989) is critical of the approach of Stigler and Becker (1977) because he finds that it masks the true underlying motivations. "A change in the demand for X can be interpreted as an increase in X's ability to produce Z, where Z is the true object of desire. But how are we to explain a change in the demand for Z?" This is precisely what Buchanan is arguing here too. For Stigler and Becker, Z is that for which demand is fundamentally static. Consider within the light of Buchanan's framework: what would lead a person to desire to become a certain type of person? There is some base desire here that—perhaps is understood or pursued differently—does not radically shift within an individual over time. It may very well be that these desires are among the key traits that give continuity to the individual himself, making him the "same person" across time. The Stigler and Becker framework could accept Buchanan's artifactual paradigm as an intermediate step here—that is, a constant *base* or core desire leads to one seeking to shape various traits and characteristics about himself—but reject that the foundation(s) are malleable.

Nonetheless, both Buchanan and Stigler and Becker reject the idea of the environment as deterministic for the individual's tastes and preferences. For instance, Stigler and Becker (1977) write, advertising need not change tastes and preferences, but rather knowledge and prices. Individuals are not slaves to the whims of marketing. Advertising may increase awareness of a good for practical use and increase the prestige of a good, and, as people with varying preferences for prestige, social distinction, and fashion, such activities *do* change their constraints. "The commodity apparently produced by fashion goods is social distinction," but since "distinction is scarce and is to a large extent simply redistributed among persons... people are often 'forced' to conform to new fashions."

Caplan (2000) suggests that, contrary to popular opinion, many seemingly irrational behaviors or beliefs may be due to underlying preferences. The model of **Rational Irrationality** suggests that beliefs to which the individual has a specific attraction or from which they receive a psychological benefit for which the benefits are judged greater than the costs. We would expect such irrationality to thrive in low cost situations and see that various political, ideological, religious, or philosophical beliefs; however, citing other empirical work, while "in terms of both statistical and economic significance from full rationality is often large, but... divergence from income maximization is usually small..." Thus, "the agent might be irrational... because irrationality isn't costly enough." As the costs rise, we would expect individuals to seek more information and adjust their behavior or even their beliefs leading to less biased outcomes.

In Caplan (2006), he goes so far as to suggest that many various diagnoses of mental illness are actually just articulations of distinct and extreme preferences. Substance abuse can demonstrate a preference of "consuming alcohol to career success or family stability," and need not be considered "irrational." ADHD "in economic terms is a high disutility of work combined with a strong taste for variety." Anti-Social Personality Disorder (ASPD) diagnostic criteria "are exclusively about preferences—for narrow selfishness, high discount rates, and affinity for violence." When it comes to delusions or beliefs in conspiracy theories or absurdities, "just because a person believes [them] does not imply that he *cannot* believe otherwise, that changing his mind is outside his choice set. Instead, he may have preferences over beliefs. If individuals have to choose between maintaining a cherished worldview and giving the other side a fair hearing, many would rather forego the latter." Consider also that "for many delusions, the fact that you would try to feign recovery shows that your degree of irrationality—not just outward behavior—is incentive-sensitive." Caplan says that psychiatrist Thomas Szasz argues that "people largely become schizophrenics because they find reality too unpleasant to cope with." He also tells the story of John Nash, a Nobel Prize winning economist, who was also a paranoid schizophrenic and notes that "Nash's first-hand account is that his return to rationality was a *choice*." These extreme preferences, then, can be sensitive to costs and benefits.

Caplan writes the gun-to-the-head test is useful for considering whether a behavior is truly a challenge to *homo economicus*. If pointing a gun at the person could allow them to stop their behavior, then we can tell that the behavior was not pre-determined and an alternative was available in the choice set. Of course, the failure of the test does not guarantee, but it does make it more likely that the alternative is impossible. "Pointing a gun at

a paralyzed man will not enable him to walk... [and] conditions like mental retardation and Alzheimer's disease are also highly unlikely to pass the gun-to-the-head test."

Preferences are not merely a given fact about the world or exogenous data. At the meta-level, they are endogenous: shaped by the system through nature, culture, and experiences, but also cultivated by the chooser himself. It is critical to recognize that human action is not mechanical; people make real decisions and adjust their behavior—and even their beliefs and personalities—based on the perceived costs and benefits, even if weighed only implicitly. Ultimately, not all preferences are immutable *constraints*, even if, as Stigler and Becker argue, it is possible and useful to model preferences as static when conducting analysis.

2.2 Utility

2.2.1 "Measuring" Utility and Indifference Curves

The **Utility Function** used by the economist mathematically maps these preferences. The relative magnitude of utility represents the direction of an individual's preference. Utility is merely a quantitative measure of a qualitative rank ordering—that is, it is fundamentally ordinal. The units are arbitrary—inches, feet, centimeters, meters—but, as long as we are consistent with the numbers we assign to given units, we can measure something objective.

Utility functions are structured such that for

$$x \succsim y \Rightarrow u(x) \geq u(y)$$

However, there are not objective units for utility. Any utility function that preserves the rank-ordering is functionally equivalent; thus, any monotonic transformations are permissible—that is, multiplication by positive constants, addition of a constant, power functions with positive exponents, and logarithmic functions. All that matters is that the rank ordering is preserved.

As Alchian (1953) writes, measuring utility between "sure prospects"—where the agent is choosing between two distinct options and knows precisely what he is getting—the rank order is all that matters. Given completeness and transitivity assumptions, we can predict actions. "Utility for this purpose and by this method is measurable up to a monotonic transformation, *i.e.*, it is ordinal only."

However, most of the time, we are choosing between more than just two single goods, but rather *baskets of goods*. These baskets can be "groups of sure prospects"—where the agent knows exactly what he is getting, but there are multiple distinct items—or it can include "uncertain prospects" where the agent is not certain which of the possible outcomes he will receive by choosing the option.

When it comes to groups, Alchian writes that

> [if] we want to be able to order *groups* of entities according to their weights... we want to assign numbers to each of the component objects so that when we combine the objects into sets or bundles, we can order the weights of the composite bundles... by *merely adding* together the numbers assigned to each component. ... [However], although utility was measurable for the purpose of the kind of prediction in the preceding problem, it is not measurable in the sense that these components can be aggregated or combined in any known way to predict choices among *groups* of sure prospects.

For comparing groups, we must use a cardinal utility, which is measured on an interval scale. In this case, only a positive linear transformation is permissible for utility functions to maintain both their ordering and spacing of intensity.

$$v = au + b \text{ where } a > 0$$

However, cardinal utility is not equivalent to psychological intensity of a preference, but rather it is a measurement for determining the consistency of choices when looking either at baskets of sure prospects of probable outcomes.

We generally add two further qualifications on preferences:

5. Monotonicity – the agent prefers more of at least one good. For two bundles x and y where the quantity of the components of the bundles are such that

$$x_i \geq y_i \text{ for each } i \text{ and } x_i > y_i \text{ for some } i \Rightarrow x_i \succ y_i$$

6. Convexity – the agent may be indifferent between different goods, but, when comparing bundles, they generally prefer a bundle with a mixture of goods (so that the indifference curves are convex to the origin).

$$rx + (1 - r)y \succsim a \text{ where } 0 < r < 1$$

Does Monotonicity mean that we are assuming individuals can never be satisfied and are so greedy they can never have enough to themselves? Not

necessarily. Charitable giving, for instance, is a form of an individual's consumption. If the individual gains from charitable giving having more to donate will provide them with more satisfaction. They need not consume the additional goods selfishly. But it is also true that as living standards have risen, humans have demonstrated the ability to have seemingly insatiable wants. We can always live longer, slightly more comfortable, more conveniently, and more entertainingly.

To consider what monotonicity and convexity mean, we can graph baskets of goods between which the consumer would be indifferent. For simplicity, baskets are made up of two goods x and y. A curve representing baskets of goods for which $a \sim b$ and $u(a) = u(b)$ is an **Indifference Curve**.

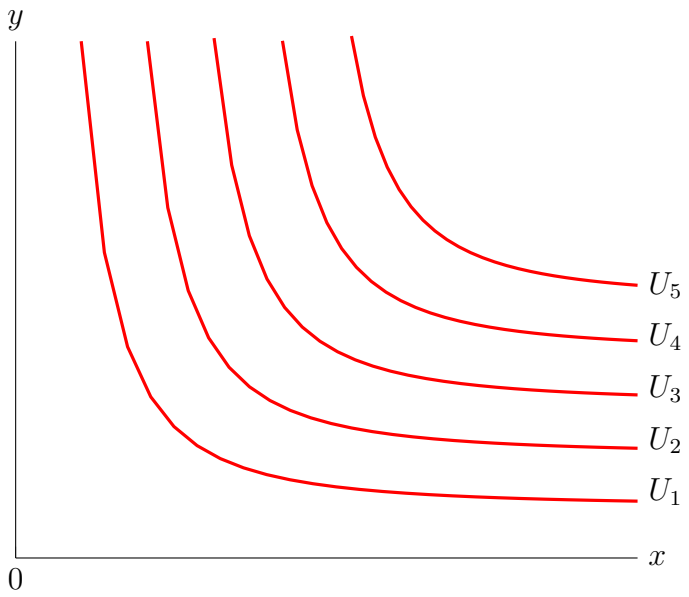


Figure 2.1: Indifference Curves

Monotonicity means that as the quantities of the goods increase, the agent's utility increases. In the graph below, $U_2 \succ U_1$ as the amount of utility represented by each point on U_2 is greater than the amount of utility represented by each point on U_1.

Convexity is what gives the indifference curves their shapes. The agent is willing to go from a basket with very few b to a mixture of a and b but a lower quantity overall. In the graph below, on the extremes, when b is plenty, the agent is willing to give up a significant amount of b $(b_1 - b_2)$ to receive one a, while when b is scarce, the agent would barely give up any b $(b_3 - b_4)$ to receive one more a. This ratio is called the Marginal Rate of Substitution.

Thus, the slope of the indifference curve decreases as the amount of one good increases relative to the other good. Mathematically, this is the derivative of the indifference curve.

$$MRS_{xy} = -\frac{\Delta y}{\Delta x}$$

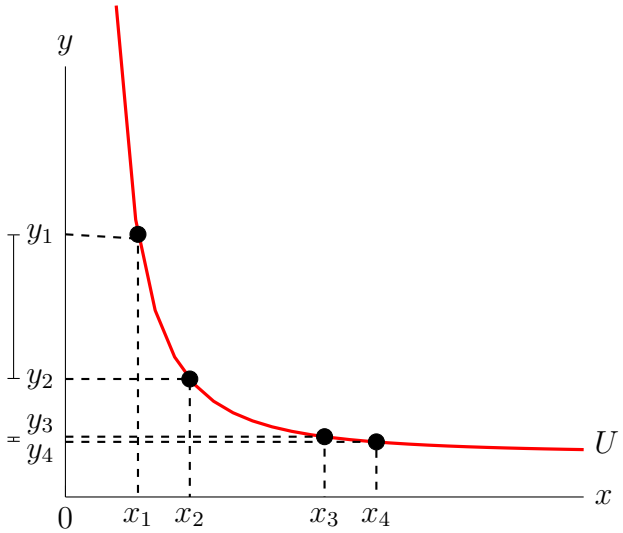


Figure 2.2: Shape of an Indifference Curve

Since the indifference curve is, by definition, a curve where utility is equal at each point, this rate is equivalent to the ratio of the utilities that the additional unit of each good provides. For

$$U(x, y) = c$$

it must be that

$$U'_x dx + U'_y dy = 0$$

The derivative of our utility function with respect to each good can be rewritten as the Marginal Utility of that good. It represents the change in utility from the next unit.

$$MU_x = U'_x = \frac{dU}{dx}$$

Thus:

$$MU_x dx + MU_y dy = 0$$

$$MU_y dy = -MU_x dx$$

$$MRS_{xy} = \frac{dy}{dx} = -\frac{MU_x}{MU_y}$$

Some things are "bads"—that is, they are undesirable—in which case*less* is preferred to more. For instance, agents must balance out the difficulty of labor with income and risk with potentially higher returns. In fact, any good beyond the point of satiation becomes a bad. We can draw indifferent curves between a good and a bad; however, their shape is slightly different. The individual prefers as little of the bad as possible, but they are willing to accept more of it, if they receive an increasing quantity of the good. Note that this is just a flipped version of the indifference curve between two goods, and we have an increasing marginal rate of substitution.

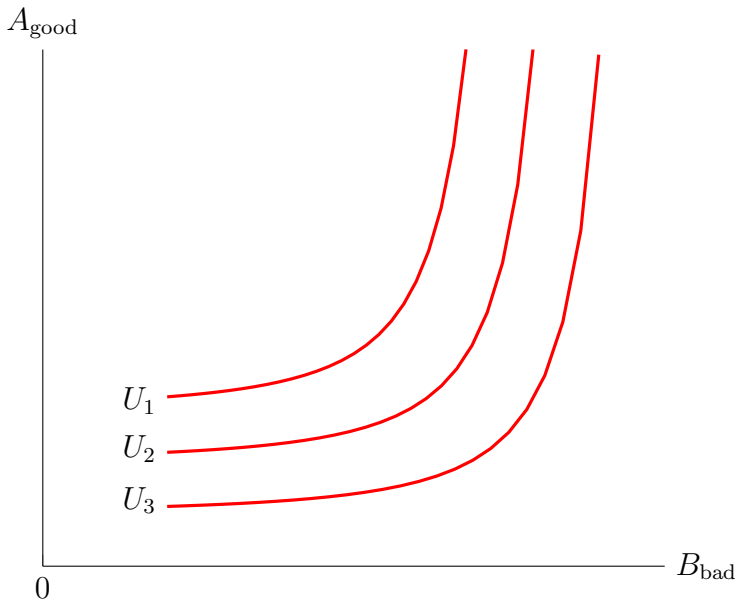


Figure 2.3: Indifference Curves with a Good and Bad

Some goods also have a relation to the consumption of the other good. For instance, the indifference curves between perfect substitutes would be straight lines with the slope of 1. One unit of x is equivalent to one unit of y, so there would be a constant marginal rate of substitution of 1. Perfect Complements can only be consumed with each other so more of one is—effectively—useless without the other, so the individual only has utility given an equal—or appropriate—amount of each and would not be willing to trade one for the other. A good example is the utility from a left shoe. People would only gain utility from one shoe if they also had its match to form a pair. For the shoes, what matters is how many of each the person has. Once the

person has more of both pairs, they gain utility, beyond that, we generally model the individuals as gaining no utility from 3 as compared to 4 or 5 left shoes, when they only have 3 right shoes, thus there is one point where the two are equal and then vertical and horizontal lines extend indefinitely up and to the right reflecting that more of one does not translate to a change in utility without more of the other.

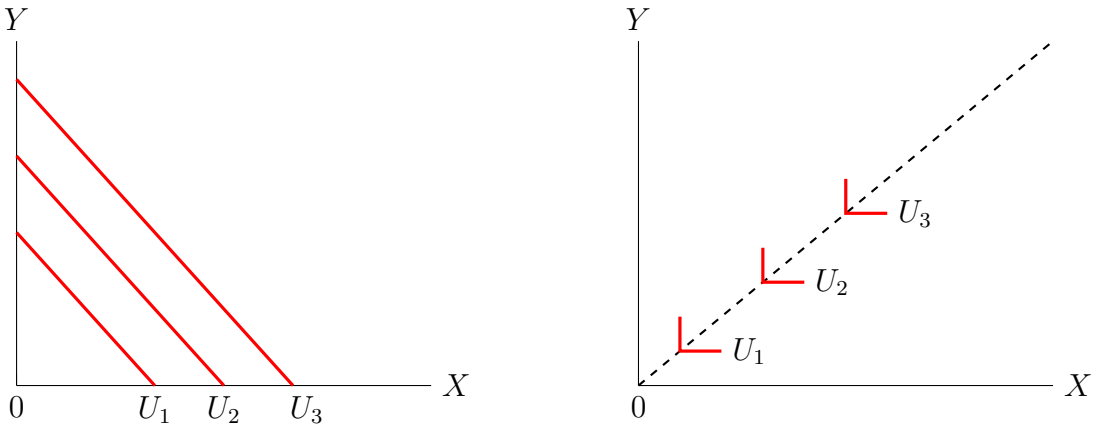


Figure 2.4: Indifference Curves of Perfect Substitutes (left) and Perfect Complements (right)

2.2.2 Utility Maximization and the Margin

Economic agents are assumed to be utility maximizers. This simply means that when agents are making decisions, they are seeking to fulfill as many of their preferences as possible. If agents choose the options they prefer, $a \succ b$, they will be choosing the options for which $u(a) > u(b)$. Maximization suggests that they choose the most preferred option, given their constraints—typically modeled as a budget constraint.

In order to get the most given monetary constraints, an individual must compare the bang-for-the-buck. They first buy the good that gives the highest utility given the price. Then, compare whether it makes sense to buy another unit of that good next or buy the first unit of another good.

Mathematically, if we are seeking to maximize a utility function, we use the Lagrangian method of optimization from multi-variable Calculus. For instance, for a given income I, given prices P_x and P_y and a given utility function $U(x, y)$:

$$\max_{x,y} U(x, y) \text{ subject to } P_x x + P_y y = I$$

This becomes the benefit function $U(x, y)$ plus a multiplier λ against the cost function set equal to 0.

$$\mathbb{L} = U(x, y) + \lambda(I - P_x x - P_y y)$$

Then, to find the optimal quantities of x and y, we take the first-order conditions. That is, we take the derivatives of the entire function with respect to each variable:

$$\frac{d\mathbb{L}}{dx} = MU_x - \lambda P_x \stackrel{\text{set}}{=} 0$$
$$\frac{d\mathbb{L}}{dy} = MU_y - \lambda P_y \stackrel{\text{set}}{=} 0$$

Solving, then for x:

$$\frac{d\mathbb{L}}{dx} = MU_x - \lambda P_x = 0 \Rightarrow MU_x = \lambda P_x$$
$$\frac{MU_x}{P_x} = \lambda$$

and for y:

$$MU_y = \lambda P_y$$
$$\frac{MU_y}{P_y} = \lambda$$

In the end, then, we get the condition:

$$\frac{MU_x}{P_x} = \frac{MU_y}{P_y} = \lambda$$

At the maximum, the consumer seeks to equalize the bang-for-the-buck—the marginal utility per unit of money—for each good. λ is commonly referred to as the "shadow price," as it represents this ratio.

Utility functions are generally represented as concave down—that is, the second derivative of the utility function is negative.

$$\frac{d^2u}{dx^2} < 0 \qquad \text{Diminishing Marginal Utility}$$

With diminishing marginal utility, it is easy to see that, as we buy more x, MU_x falls, leading to the equalization of the bang-for-the-buck ratios.

However, we need not assume diminishing marginal utility for this optimization solution. After all, the Lagrangian method makes no assumptions about the second order derivatives. Instead, given the diminishing Marginal Rate of Substitution, regardless of whether the change in utility—however measured—is diminishing in absolute terms, it is at least diminishing *relative to* all other goods.

2.2.2.1 Cobb-Douglas Utility Function

One of the most common forms of utility functions is the **Cobb-Douglas** Utility Function. It takes the form:

$$U(X,Y) = X^a Y^b$$

So, for instance, a single indifference curve could look like:

$$100 = X^2 Y^3$$

For this general utility function, the marginal utilities are:

$$MU_X = aX^{a-1}Y^b$$

$$MU_Y = bX^a Y^{b-1}$$

And the Marginal Rate of Substitution is:

$$MRS_{xy} = -\frac{MU_X}{MU_Y} = -\frac{aX^{a-1}Y^b}{bX^a Y^{b-1}} = -\frac{aY}{bX}$$

2.2.3 Utility and Uncertainty

When we consider the third type of analysis, choosing between "uncertain prospects," we add probabilities to our analysis. When choosing between a sure prospect and uncertain prospects or between two uncertain prospects, we use **von-Neumann–Morgenstern Expected Utility Theory**. We take the expected utility of each option and weight it by its probability.

$$EU_{\text{basket}} = \sum_{i=1}^{i} p(x) \cdot x_i$$

where $\sum p(x) = 1$. For example, for a basket composed of A with probability of p and C with a probability of $1 - p$:

$$EU_{AC} = p \cdot U(A) + (1-p) \cdot U(C)$$

In this setup, the process for how A and C are determined is irrelevant for the ranking, so long as the process does not change p.

The Archimedean Axiom states that for rank ordering

$$C \succ B \succ A$$

there is some probabilities α and β such that:

$$\alpha C + (1-\alpha)A \succ B \succ \beta C + (1-\beta)A$$

Even though C is strictly preferred to B, no matter how bad A is, some mixture of A and C will give us an option better than B.

Even consider if option A is death, note that every action you take carries a probability of death—driving or riding in a car, crossing a street, or eating food that could've been poisoned. Nonetheless, we still choose to engage in many of these activities knowing that there is *a* risk of death, because the probability (α) is small enough that we prefer that to not acting. Of course, sometimes the reverse is true. You probably would not want to pick a basket that includes a 10% chance of $1 million and 90% chance of death. No matter how good the monetary prize, if the weight of the bad or least preferred option is strong enough, you'll choose something else.

In comparison to the expected utility, the utility of the expected value of a basket is equivalent to

$$EV_{\text{Basket}} = U\left(\sum_{i=1}^{i} p(x)\cdot x_i\right)$$

This is easiest to understand when considering a money lottery. For instance 50% chance of $0 and 25% of $10 and 25% of 100:

$$EV_{\text{lottery}} = U\left(0.5\cdot 0 + 0.25\cdot 10 + 0.25\cdot 100\right) = U(\$27.50)$$

We can measure how risk-averse (or risk-preferring) an agent is based on the difference between the expected value and expected utility.

$$U(EV_{\text{basket}}) = E(U_{\text{basket}}) \qquad \text{Risk-Neutral}$$

$$U(EV_{\text{basket}}) < E(U_{\text{basket}}) \qquad \text{Risk-Seeking}$$

$$U(EV_{\text{basket}}) > E(U_{\text{basket}}) \qquad \text{Risk-Averse}$$

Under uncertainty, we say that consumers maximize their expected utility, but this includes their personal risk preferences and tolerances.

2.2.3.1 Subjective Probabilities

What happens when the agent does not know the probability distribution? Savage (1954) proposed the Subjective Expected Utility approach where probabilities are assessed subjectively by each agent. People do not have

to explicitly assign probabilities, but people generally recognize that improbable events are still possible. Relying on many of the standard utility assumptions—like completeness, transitivity, monotonicity, and continuity—as well as non-triviality—that is, the agent prefers something and is not indifferent to everything—Savage adds the sure-thing principle and independence of beliefs from tastes.

The sure-thing principle suggests that if there was certainty for an event and it would not affect the outcome of a decision, uncertainty about the event should not affect the outcome of the decision. Savage writes:

> A businessman contemplates buying a certain piece of property. He considers the outcome of the next presidential election relevant to the attractiveness of the purchase. So, to clarify the matter for himself, he asks whether he would buy if he knew that the Republican candidate were going to win, and decides that he would do so. Similarly, he considers whether he would buy if he knew that the Democratic candidate were going to win, and again finds that he would do so. Seeing that he would buy in either event, he decides that he should buy, even though he does not know which event obtains, or will obtain, as we would ordinarily say.

The Independence of Beliefs from Tastes suggests that a taste in the outcome should not affect the probability beliefs. For instance, the magnitude of a prize will likely influence a person's willingness to bet and how much they are willing to bet (with a taste for more), but the magnitude should not influence the belief about the probability of the outcome.

With these assumptions, Savage holds that a maximization problem is solvable—we can predict how agents will choose—even if there are no objective probabilities and only subjectively evaluated ones.

Subjective Probability Theory suggests that people update their beliefs in accordance with Bayes' Rule:

$$P(A|B) = P(B|A) * \frac{P(A)}{P(B|A) * P(A) + P(B|\neg A) * P(\neg A)} \qquad \text{Bayes' Rule}$$

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \qquad \text{Bayes' Rule}$$

The key argument of Bayes' Rule is that a conjunction can never be more probable than either of the components. A simplified form is the Conditional Probability Formula:

$$P(A|B) = \frac{P(A\&B)}{P(B)} \qquad \text{Conditional Probability Formula}$$

It can also be generalized beyond a binary with:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{n} P(B|A_j)P(A_j)}$$

2.2.3.2 Rational Expectations

The Rational Expectations hypothesis seeks to relate subjective probabilities to the real world in the aggregate. For group of people with rational expectations, their combined subjective probability distributions are identical to the objective probability distribution. On average, the error is zero—that is, judgments are unbiased—and mistakes are uncorrelated with available information. This allows for significant errors. It suggests that information—or the lack thereof—changes the *variance* of estimates, not the mean.

Muth (1961) laid out three claims that are the foundations of Rational Expectations. Quoting Muth:

1. Information is scarce, and the economic system generally does not waste it.
2. The way expectations are formed depends specifically on the structure of the relevant system describing the economy.
3. A public prediction... will have no substantial effect on the operation of the economic system unless it is based on inside information.

To the extent that we think experts know better than the average individual and have a more complex model, the second component of Rational Expectations suggests that the best models of reality as they exist today shape the formation of expectations. An economic theorist could not use his knowledge to consistently profit above the average market returns based upon this theoretical knowledge, because "the aggregate expectations of the firms is the same as the prediction of the theory," because the knowledge of this theory is available to them. A public prediction is a publicly announced forecast about future economic conditions.

Haltiwanger and Waldman (1985) distinguish between an aggregated and individual case of rational expectations. The original Rational Expectations hypothesis, they write, is that *in the aggregate* expectations are rational and "expectational deviations across agents will tend to cancel out." This is *not* the same as the assumption that "the expectations of each agent taken separately is, by itself, consistent with the predictions of the relevant economic theory." In fact, the two do not lead to identical outcomes, so we cannot necessarily model the former—an aggregated rational expectations world—with the use of a representative agent whose expectations are rational.

The two are not the same, for example, when there are network or bandwagon effects. In the environment that they label "synergism," bias is amplified as "rational agent arbitrageurs are incentivized to go with the flow [and] buy what the biased agents are buying, since payoffs rise with the scale of participation."

Fremling and Lott (1996) advocate for a weaker form of rational expectations that allows for some bias in aggregated decisions. The second component of Muth's hypothesis is that the best specified model shapes expectations of the general public and their mental model formation. However, this does not mean that each agent's model is appropriately specified. While they note that a *misestimation* error cancels out in the aggregate, "individual errors in *identifying the relationships* among variables cause a downward bias in the aggregate that would be equivalent to the public *underestimating* the strengths of the true relationships." They argue that even if only a fraction of the population omits a critical variable from their mental models, we can end up with biased results.

Caplan (2001) notes that the Rational Irrationality model also provides explanations for deviations from Rational Expectations. Many "people hold systematically biased, low-information, high certitude beliefs." So long as there is little to no cost for deviating from Rational Expectations, the market mechanism that leads to the aggregated Rational Expectations result does not exist. It is arbitrage opportunities that are exploited and discipline the market. But when there are political, ideological, or conspiratorial beliefs for which there is little to no cost, individuals revert to their most preferred "bliss belief." He writes that, "if the most pleasant belief for an individual differs from the belief dictated by rational expectations, agents implicitly weigh the hedonic benefits of deviating from rational expectations against the expected material costs of self-delusion."

2.2.4 Interpersonal Utility Comparisons

When modeling utility, a key question has been raised about whether we can make interpersonal comparisons of utility. Can we just add up all the utility of the people in the community to create a social welfare function and maximize that, just as we each maximize individual utility? Can we use utility as an objective way to evaluate changes in policy? When utility is merely a person's ordinal ranked preferences, it is clear that this is nonsensical to add these numbers to get anything. What could adding this ranking possibly mean? When utility is cardinal, however, this seems far more plausible. After all, cardinal utility displays some degree of preference, even if

not psychological intensity. But utility is also not the measure of demeanor or disposition—nor would we want it to be—since it is of little help to compare a poor, happy person with a miserable, rich person.

A common mantra is that utilitarian cost-benefit analysis can work by giving to each individual an "equal weight." Robbins (1938) writes that "the postulate of equal capacity for satisfaction [comes] from outside... rest[ing] upon ethical principle rather than upon scientific demonstration." Thus, we cannot make objective or measurable interpersonal utility comparisons. At the minimum, it is clear that we cannot use "the same methods of demonstration as were available in other fields of social judgment" to resolve differences. While "observation or introspection" are available to us for determining an individual's utility function, this is not true for *inter*-personal comparisons.

Often, what appears to be our interpersonal judgments are merely intuitions that express our own preferences and values. For instance, we judge how funds are used; redistributing money from a rich to a poor man is judged based on whether the money is used to objects deemed necessities or alcohol, drugs, or luxuries. We can also empathize with one person and express the position we would rather have or what our ideal society would look like, but these are not scientific claims.

Since modern economics holds only to a diminishing marginal rate of substitution and not diminishing marginal utility, which would suggest some diminishing utility of income or consumption—potentially allowing for interpersonal comparisons—it is not even analytically necessary. Because objective interpersonal utility comparisons are not possible, this is not a measure for good policy. Robbins writes that economics can "judge the consistency of different policies, but, in itself, it [could pass] no verdict of good or bad." To the extent that we engage in subjective utility comparisons, it makes more sense to go to the root and identify the philosophical and ethical principles underlying the value judgment rather than masking it in the language of utility.

Perhaps we cannot compare directly between one person and another, but could we somehow aggregate the entire community's well-being into a single metric using a Social Welfare Function, similar to the Utility Function at the individual level? Arrow (1950) proves that there is no social welfare function that can aggregate an entire community's preferences into a single utility function while satisfying 5 conditions simultaneously:

1. Completeness – SWF defined for all possible alternatives
2. Positive Correlation between each Individual and Social Value – "if

an alternative social state x rises or does not fall in the ordering of each individual without any other change in those ordering and if x was preferred to another alternative y before the change in individual orderings, then x is still preferred to y.

3. Independence of Irrelevant Alternatives – ranking between any two states of the world is the same when we add a third possible state.
4. Citizen Sovereignty – this precludes a function that is "imposed" where "the community can never express preference for y over x."
5. Non-Dictatorship – a single individual does not end up controlling each outcome.

Arrow (1950) concludes with the Possibility Theorem.

> If we exclude the possibility of interpersonal comparisons of utility, then the only methods of passing from individual tastes to social preferences which will be satisfactory and which will be defined for a wide range of sets of individual orderings are either imposed or dictatorial.

Thus, it seems that the individual is the appropriate unit of analysis here and that there is no rational way to aggregate social action to create a function like a utility function to consistently measure outcomes and preferences and have those outcomes and aggregated preferences appear rational without leading to a dictator.

2.3 The Budget Constraint

When an individual chooses among different possible activities and consumption possibilities, preferences are critical, but the other side of the equation is the opportunities. They are limited ultimately by scarcity and practically by their own output and income. Thus, all consumption spending must remain within this constraint. In the simplest model with two goods x and y, this looks like

$$P_x x + P_y y \leq I$$

When the constraint is non-binding—that is Consumption is less than Income ($C < I$)—there is slackness. When the constraint is binding, there is no slack.

2.3.1 The Budget Constraint and Indifference Curves

A utility-maximizing consumer will seek to get to the highest possible indifference curve. We generally assume that the consumer has not yet

reached the point of satiation—that is, their wants exceed their income—so the consumer uses all of his income and the constraint holds.

$$P_x x + P_y y = I$$

The slope of the budget line is:

$$\frac{\Delta Y}{\Delta X} = -\frac{I/P_y}{I/P_x} = -\frac{P_x}{P_y}$$

When maximizing, the agent will seek to find a point tangent to their indifference curve. Only this will allow the consumer to achieve the highest possible indifference curve—which means the most utility available. Recall that the slope of the indifference curve is $-\frac{MU_x}{MU_y}$, so:

$$\frac{MU_x}{MU_y} = \frac{P_x}{P_y} \qquad \text{Utility Maximization}$$

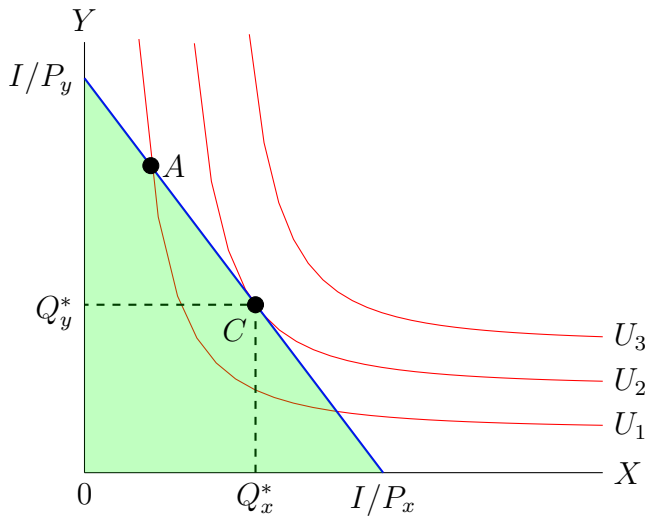


Figure 2.5: Indifference Curves with Budget Constraint

For instance, in this graph, the green area is all of the feasible combinations of goods available to the consumer. The blue line represents the budget constraint. Its intercepts are the point at which all of the income would be spent on one good. If the agent bought none of x, they would buy as many y as their income could buy—that is, $\frac{I}{P_y}$. But being on the budget line isn't enough. We can see that if the consumer was at point A, he would rather

be at point C, giving up some y in exchange for more x. We know $C \succ A$, because $U_2 \succ U_1$ by definition.

As we have seen, at maximization, the ratio of marginal utilities is equal to the price ratio. However, we can rewrite this to get a more familiar equation:

$$\frac{MU_x}{MU_y} = \frac{P_x}{P_y}$$

$$\frac{MU_x}{P_x} = \frac{MU_y}{P_y}$$

Once again, we have equality between the bang-for-the-buck ratio.

Now, not all goods will necessarily be chosen. Some goods will have a quantity consumed of 0. This is the case of a corner solution.

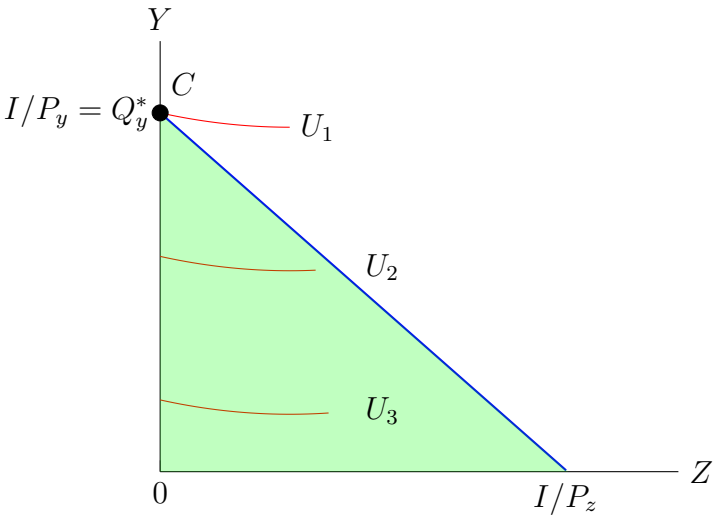


Figure 2.6: Indifference Curves with Corner Solution

For all goods consumed in positive quantities at the optimum, the bang-for-the-buck ratio is the same. For any good that is not consumed, the bang-for-the-buck ratio must be less, even for the very first unit consumed.

$$\frac{MU_x(x>0)}{P_x} = \frac{MU_y(y>0)}{P_y} > \frac{MU_z(z=0)}{P_z}$$

2.3.2 Shifting the Budget Constraint

The Budget Constraint can shift in two ways: either through changes in overall income and the purchasing power of the given set of goods *or* through changes in *relative* prices.

2.3.2.1 Changes in Income

A change in income results in a change in the intercept, but not the slope, of the budget constraint. When income expands, the consumer is able to move to a higher indifference curve. The **Income-Consumption Curve** (sometimes also called the Income Expansion Path) shows the optimal consumption bundles at each level of income for a given (unchanging) price ratio—that is, the budget constraint's slope remains constant.

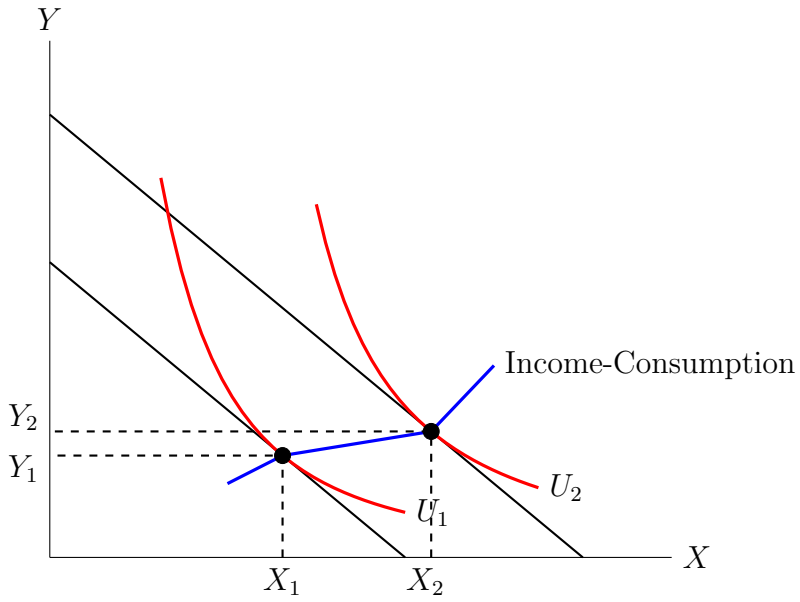


Figure 2.7: Income-Consumption Curve for a Normal Good

Now, the Income-Consumption Curve need not be—and likely is not—perfectly linear. This is easier to think of when your income falls. You will not buy all the same goods in lower quantities. You will still spend on food, shelter, and transportation, but will likely forego various purchases that appear to be less necessary. Likewise, when your income increases, you will not merely buy all of the same goods in higher quantities.

A Normal Good is such that, as your income increases you will buy more—though, not necessarily *proportionately* more and the Income-Consumption Curve has a positive slope. A Superior Good (or Luxury Good) is a type of normal good. As your income increases, the proportion of your income spent on a superior good will increase. At a very low level of income, the proportion of income spent on jewelry or vacations is probably close or equal to 0; however, at higher levels of income, the spending on these categories

will increase both in absolute terms but also in relative terms, as a proportion of the income.

On the other hand, an Inferior Good, however, is a good of which you buy *less* as your income increases, giving the Income-Consumption Curve a negative slope. For instance, one's consumption of used goods—clothing or cars—or generic brands is likely to decrease with a rising income as one substitutes in their stead consumption of new clothing and cars and name-brand products.

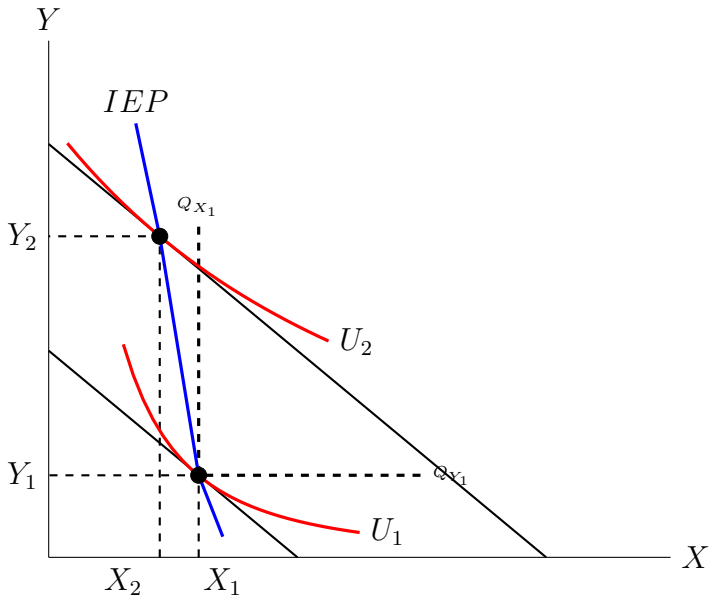


Figure 2.8: Income-Consumption Curve for an Inferior Good

Most goods are inferior only over a certain range. Consider a person who purchases ramen noodles to eat for food: if they receive an additional dollar, they might purchase another few cups of ramen noodles—thus, the good is normal at this range. However, if they receive an additional hundred dollars, they will likely choose to purchase other food instead—thus, at that level of income expansion, the good is inferior.

We can graph the quantity of a good that an individual would demand against the level of income using the **Engel Curve**. In the graph below, the blue curves represent normal goods, while the red curve represents an inferior good. Note that inferior good I is only inferior beyond a certain point S. The luxury good, likewise, is only consumed beyond point S and the quantity consumed is increasing at an increasing rate $\frac{d^2}{dx^2} > 0$.

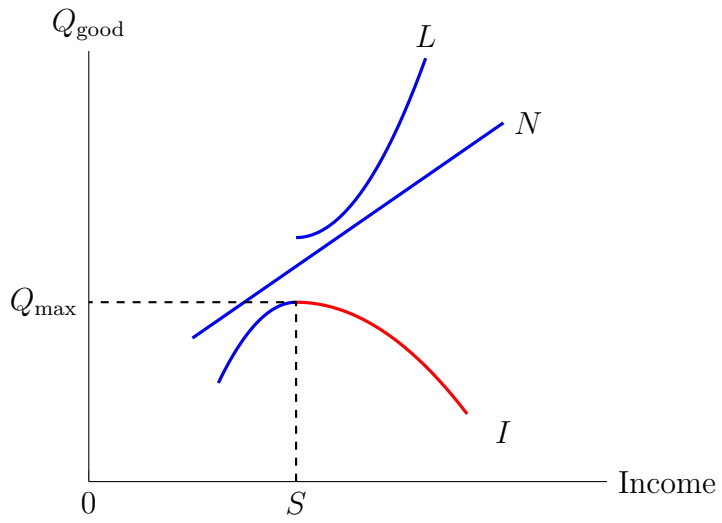


Figure 2.9: Engel Curve of Normal Good N, Luxury Good L, and Inferior Good I

Finally, note that a *proportionate* change in prices is equivalent to a change in income—that is, a change in the intercept, but not the slope, of the budget constraint. For instance, if all prices doubled, that would be the equivalent of your income being cut in half.

2.3.2.2 Changes in Relative Prices and Hicks Decomposition

The other type of potential change in the budget constraint is a shift in relative prices. This changes the slope, but not the intercept, of the budget constraint, given one of the prices remains constant. A change in the relative prices will lead to changes in the quantity of both goods consumed. The **Price-Consumption Curve** or Offer Curve runs through each point of tangency between the various budget constraint slopes and the agent's indifference curve. Each point along the curve, then, is the consumer's optimal consumption budget given the price ratio. As the price of X is increasing (from P_X^1 to P_X^2 and P_X^3), the consumer can buy fewer of it and chooses to buy more Y instead.

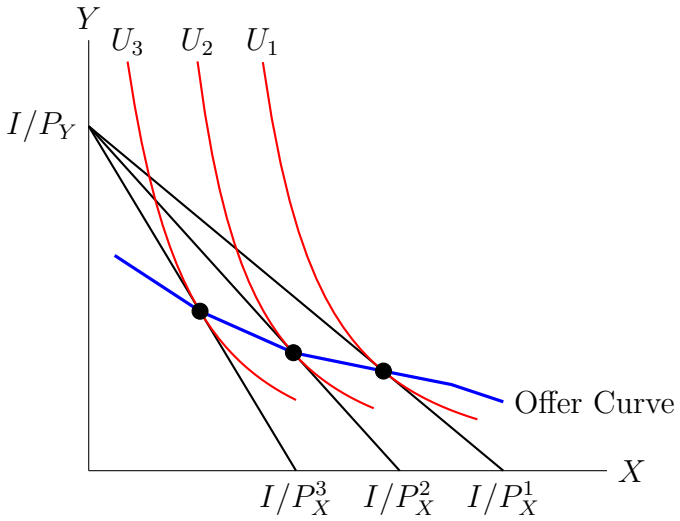


Figure 2.10: Price-Consumption Curve

When there is a shift in relative prices, there are two effects that are taking place.

First, there is an effective shift in the consumer's real income or purchasing power—an Income Effect—due to the change in the price of X. This change to the consumer's effective income moves him up to a higher indifference curve with the fall in the price of X. Now, the consumer must adjust the basket of goods based on his new wealth. If X is a normal good, given a fall in price, the quantity demanded will rise; however, if X is an inferior good, given a fall in price, the quantity demanded will fall.

Second, given the change in the price ratio leads the consumer to rebalance their basket so that the bang-for-the-buck equation is equal. This is called the Substitution Effect. Even if the consumer's income had remained constant, the consumer would want to readjust their basket of goods, since X is now cheaper. Without adjusting:

$$\frac{MU_X}{MU_Y} = \frac{P_X^1}{P_Y} > \frac{P_X^2}{P_Y}$$

where P_X^1 is the initial price, but P_X^2 is the new, lower, price. The result would be giving up units of Y in favor of more X until $\frac{MU_X}{MU_Y} = \frac{P_X^2}{P_Y}$.

The purpose of **Hicks Decomposition** is to separate out the Income and Substitution Effects. For a change in the price of X from P_1 to P_2, we can decompose a consumer's shift using the graph below. A is the initial optimizing point given the price ratio $\frac{P_1}{P_Y}$. C is the final optimizing point

given the price ratio $\frac{P_2}{P_Y}$. To identify the substitution effect, we take the new price ratio and find the optimal point for these relative prices on the original indifference curve U_1. In the graph below, this is point B.

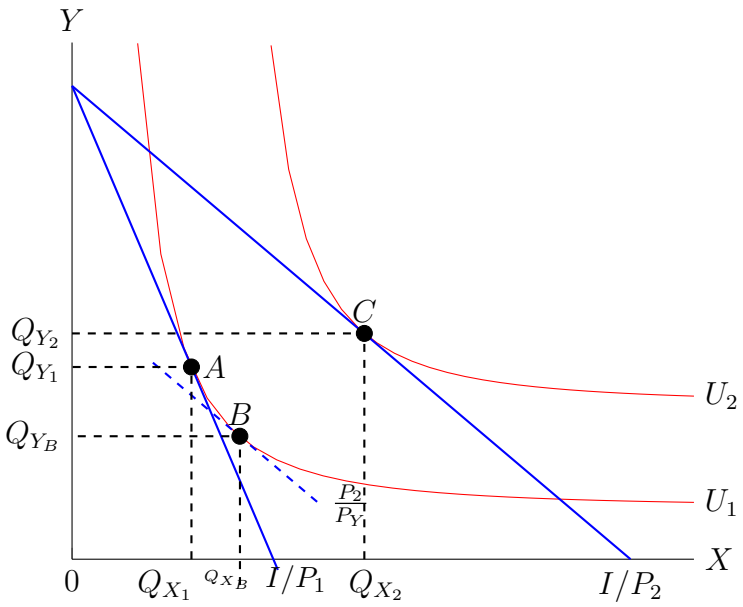


Figure 2.11: Hicks Decomposition of the Income and Substitution Effects

For a decrease in the price of X, the Substitution Effect is the move from A to B. Because X has become relatively cheaper, this is a substitution away from Y (see the decrease from Q_{Y_1} to Q_{Y_B}) and towards X (see the increase from Q_{X_1} to Q_{X_B}). The Income Effect, then is the move from B to C. Since both X and Y are normal goods, this is an increase from Q_{X_B} to Q_{X_2} and from Q_{Y_B} to Q_{Y_2}.

Thus, the total effect is:

$$\text{Total Effect of Price Change} = \text{Substitution Effect} + \text{Income Effect}$$

Using our graph's discrete values, this is:

$$(Q_{X_2} - Q_{X_1}) = (Q_{X_B} - Q_{X_1}) + (Q_{X_2} - Q_{X_B})$$

2.4 Demand

2.4.1 Deriving Marshallian Demand

We can change our analysis from considering the trade-off between two goods to simply the trade-off between good X and all other goods. Then, we

can create a graph where the y-axis represents the absolute price of X, maintaining the quantity of X as the x-axis. This second graph is the individual's Marshallian demand curve.

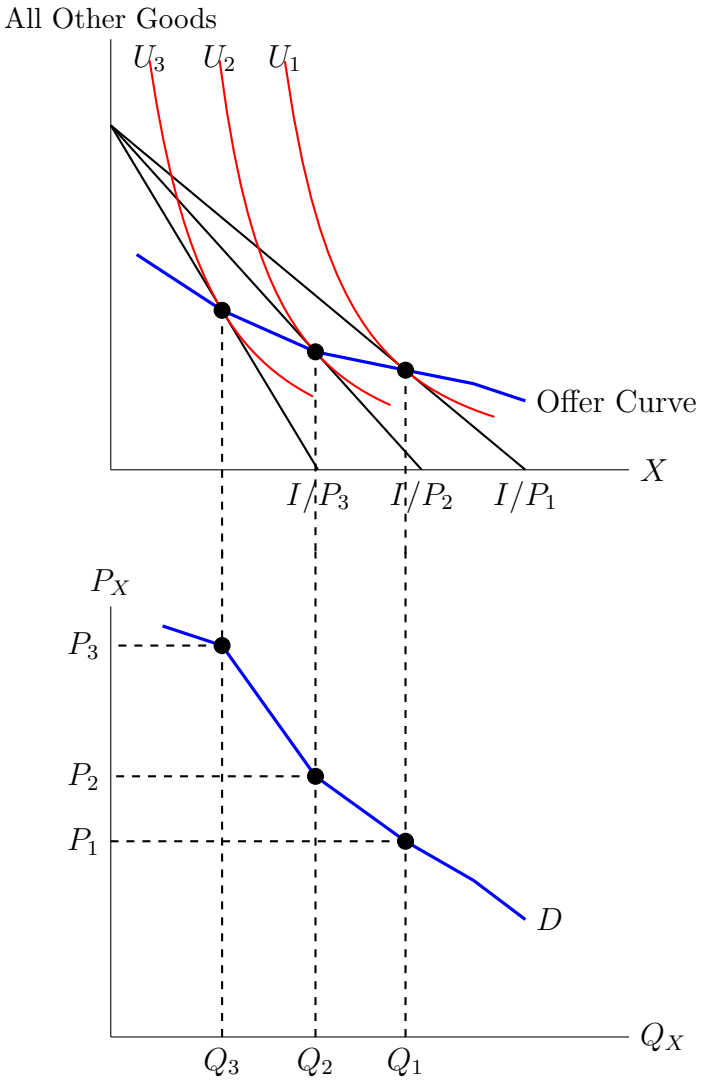


Figure 2.12: Price-Consumption Curve (top) and the Marshallian Demand Curve (bottom) for Good X

The **Marshallian Demand Curve** holds constant the consumer's nominal income—the y-intercept of the budget constraint does not change—and shows the total change in the quantity demanded of X including both the income and substitution effects. Movements along the curve assume that tastes and preferences and the nominal (money) income of consumers remain

the same as well as the price of every other commodity.

2.4.1.1 Elasticity of Demand

The elasticity of demand ϵ_x represents the percentage change in the quantity demanded of X (Q_X) for a 1% change in the Price of X ($P_X = P$). It will vary at different points along the demand curve.

$$\epsilon_x = \frac{\%\Delta Q_X}{\%\Delta P_X} = \frac{\frac{dX^M}{X}}{\frac{dP}{P}} = \frac{P}{Q} \cdot \frac{dQ^M}{dP}$$

This will be equivalent to a scaling factor $\frac{P}{X}$—the ratio of the price to the quantity demanded—multiplied by the Marshallian derivative $\frac{dX^M}{dP}$. (Note that X^M is equivalent to Q_X along a Marshallian demand curve.)

- When $\epsilon = 0$, demand is perfectly inelastic
- When $-1 < \epsilon < 0$, demand is relatively inelastic – an increase in price results in lower total revenue
- When $\epsilon = -1$, demand is unit elastic – a change in price does not impact total revenue
- When $\epsilon < -1$, demand is relatively elastic – an increase in price results in an increase in total revenue
- When $\epsilon = -\infty$ — that is, the demand curve is completely flat—demand is perfectly elastic, so an increase in price results in no revenue.
- When $\epsilon > 0$, demand is *upward* sloping making this a Giffen Good—as the price increases, the quantity demanded *increases*.

For demand function $Q = F(P)$:

$$\epsilon_x = \frac{F'(P) \cdot P}{F(P)}$$

Thus, we can see that elasticity is not *just* the slope of the demand curve $F'(P)$ but rather it varies at each point along the curve because the quantity demanded $F(P)$ and the price P change at each point. Thus, as we move along the demand curve, even given a linear demand function (constant F'), we go from undefined at the top—when $Q = 0$, ϵ_x = undefined and $\lim_{x \to 0} \epsilon_x = -\infty$—and elastic (a high price, the more responsive demand is to a fall in price). Then, eventually, the curve becomes inelastic until it reaches the x-axis at which point $P_X = 0$ so $\epsilon_x = 0$. For linear demand functions, the halfway point along the curve is unit elastic.

For demand function

$$Q = -mP + b \ , m, b, P, Q \geq 0$$

when $Q = \frac{b}{2m}$, $\epsilon = 1$.

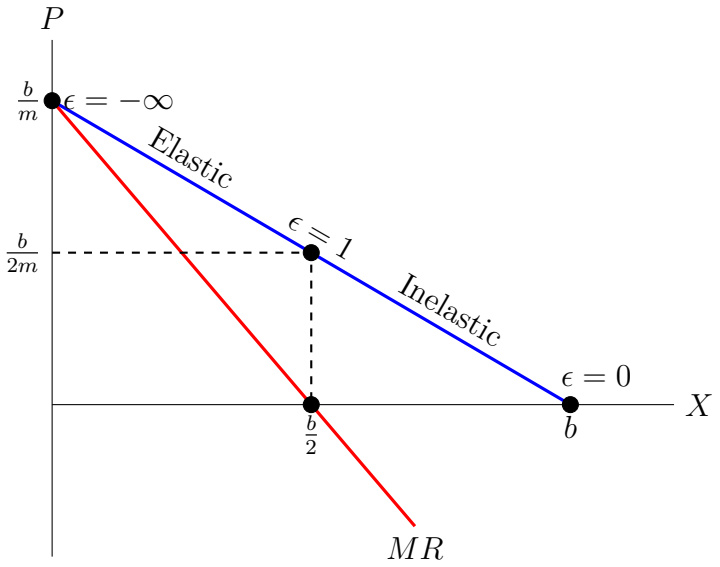


Figure 2.13: Elasticity Along a Demand Curve

For a firm facing this demand curve, absent any additional constraints, the unit elastic point is the revenue maximizing point. Note, this is not the *profit*-maximizing point, as this does not take into account any costs, which can and often does shift the optimal (profit-maximizing) point into the elastic region of the demand curve. Above the unit elastic point, the firm can *increase* its total revenue by lowering the price. Below this point, the firm can *increase* total revenue by raising its price.

Note that the Marginal Revenue curve MR is equivalent to the derivative of Total Revenue with respect to quantity.

$$TR = P \cdot Q$$

$$MR = TR\frac{d}{dQ}$$

For the demand function given above of $Q = -mP + b$:

$$\frac{b - Q}{m} = P$$

So:

$$TR = \left(\frac{b-Q}{m}\right) \cdot Q = \frac{b}{m}Q - \frac{1}{m}Q^2$$

$$MR = \frac{d}{dQ}\left(\frac{b}{m}Q - \frac{1}{m}Q^2\right) = \frac{b}{m} - \frac{2}{m}Q$$

Thus, we find that the marginal revenue curve has a slope twice as steep as the demand curve itself. At the unit elastic point, $MR = 0$.

This leads to a very important finding. If a firm is raising prices to increase revenue, they are on the inelastic segment of the demand curve when marginal revenue for the last unit of output is *negative*. Since $MR < 0$, the firm would maximize revenue by producing fewer units at higher prices. By raising prices, they are moving closer to the unit elastic point where $MR = 0$. As we will see, the optimum for the firm is $MR = MC$. This means that the firm will produce up to the point where it is indifferent between producing one more unit or not—the gain MR is equal to the cost MC. So, clearly, the firm is producing far more than the optimum if $MR < 0$ and $MC \geq 0$.

2.4.1.2 Cross-Price Elasticity

The cross-price elasticity ϵ_{YX} gives the percent increase in the demand for good Y given a change in the price of good X.

$$\epsilon_{yx} = \frac{\%\Delta Q_Y}{\%\Delta P_X} = \frac{P_X}{X} \cdot \frac{dY}{dP_X}$$

When cross-price elasticity is positive, goods x and y are substitutes. When cross-price elasticity is negative, goods x and y are complements.

This highlights an important fact: a demand *curve* can shift due to changes in the prices of other goods—like substitutes or complements—but a change in the price of the good the demand of which is being graphed is a move *along* the curve.

2.4.2 Deriving a Compensated Demand Curve

2.4.2.1 Hicksian Demand

While Marshallian Demand can be observed—as it suggests the total quantities that we should expect to be purchased given a change in price—Hicksian demand functions or **Compensated Demand** functions seek to measure the substitution effect alone and keep the consumer on the same indifference curve. While Marshallian Demand is seeking to maximize utility given a

budget constraint, Hicksian Demand is seeking to minimize cost given a utility constraint. Hicksian demand removes the income effect leaving only the substitution effect.

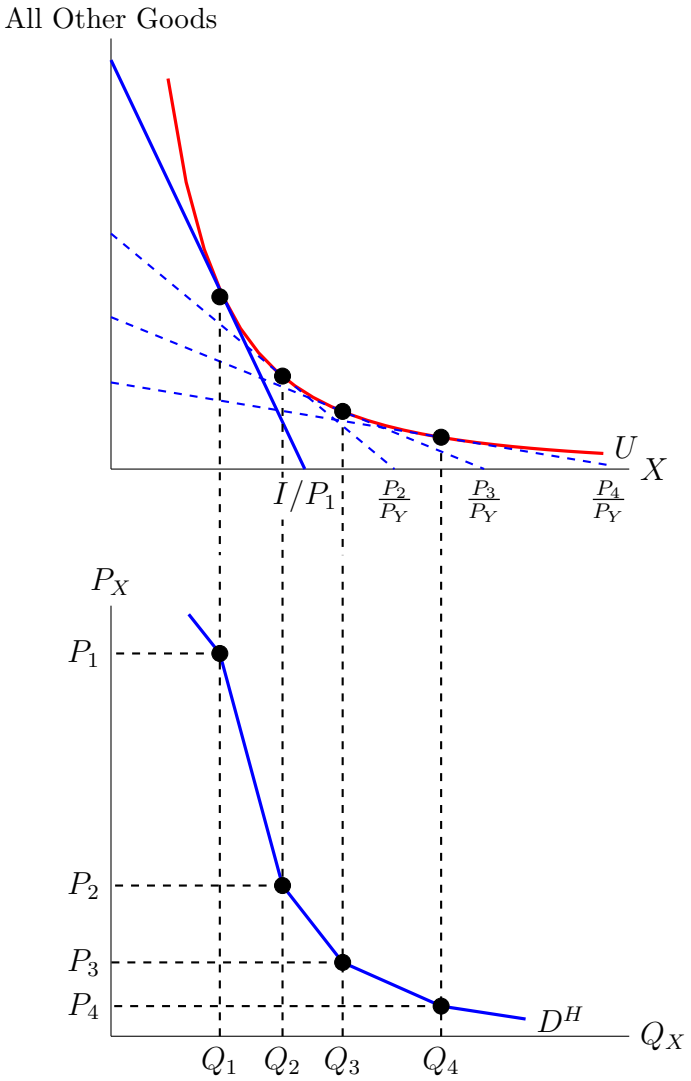


Figure 2.14: Utility-Constant Price-Consumption Curve (top) and the Hicksian Demand Curve (bottom)

The quantity of Hicksian Demand *necessarily* decreases as the good's price increases. Since utility is constant, if this were not the case, that would suggest that the consumer had already reached the point of satiation of *all* goods including the good in question.

2.4.2.2 Friedman's Alternative Compensated Demand Curve

Friedman (1949) proposes a compensated demand curve that holds real income or purchasing power—rather than utility—constant. Friedman argues that such a demand curve is better able to account for changes in welfare. For instance, he suggests that such a compensated curve will better predict the outcome of government subsidies. For instance, modeling using a Marshallian curve D^M will suggest a lower change in price and higher change in quantity than the constant real income demand curve D^F. A subsidy for a particular good that shifts out the supply curve of that good to the right requires the new suppliers of the good to bid up the prices of resources—given no slack in the economy—and will raise the prices of all other goods, which will then shift the Marshallian Demand curve back down.

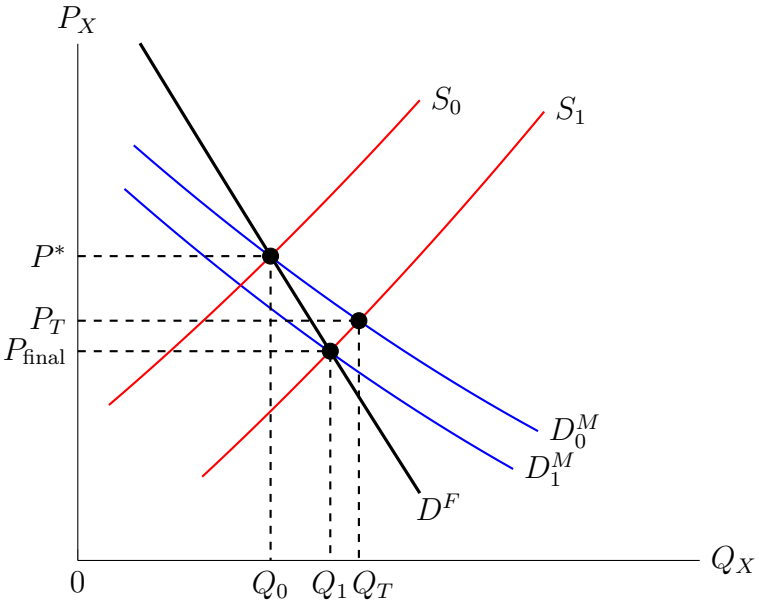


Figure 2.15: Supply and Demand with Government Subsidy

Bailey (1954) suggests that Friedman's curve has little superiority, as he is *really* seeking to model a production-frontier demand curve that "shows what in fact the community will take when the repercussions on the production of other commodities are taken into account." Friedman, in essence, argues that his curve "is a better approximation" to this production-frontier curve than the standard Marshallian "other-things-equal" curve. But Friedman's curve is an analytical nightmare. "For an arbitrary change in the price of commodity Y, what price change of X would keep *two* individuals on the same levels of real income?" Bailey argues these curves are undefined unless

all consumers have identical tastes and, thus, homogeneous indifference curves.

2.4.3 Relating Marshallian and Hicksian Systems

Each point on the Marshallian Demand curve represents a different utility for the consumer and thus its own Hicksian Compensated Demand curve. In the graph below, if the price of good X goes from P_1 to P_2, we can find the substitution effect along the Hicksian Compensated Curve with the constant utility of U_2 and the consumer moves from point A to point B as the consumer begins shifting more of the consumption bundle to X. Then, due to the income effect, assume nothing else has changed, he moves from B to C able to consume more of all goods. At point C, because of the consumer's increased purchasing power, he has moved to a higher indifference curve with a higher level of utility (U_3).

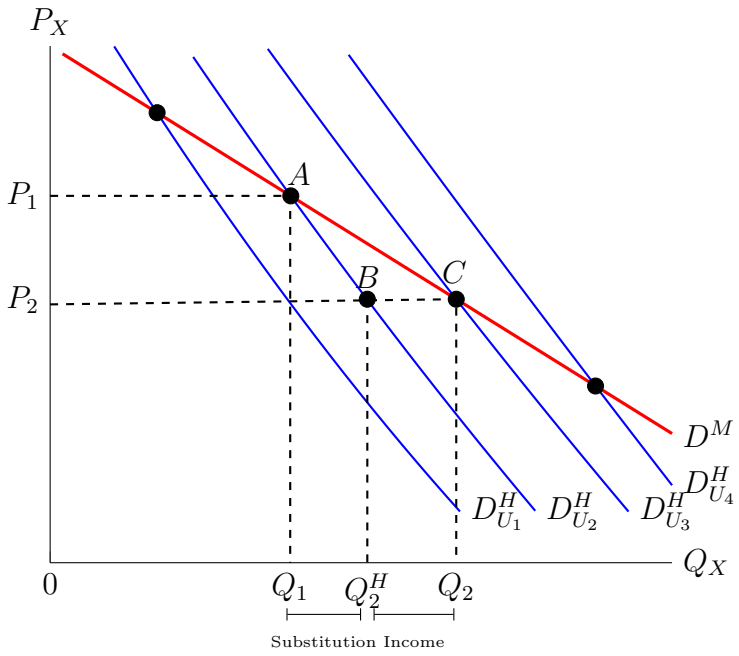


Figure 2.16: Marshallian and Hicksian Demand Curves

The Slutsky Equation mathematically relates the two systems to each other and, thus, also expresses a decomposition of the income and substitution effects. (Recall that the Hicksian Demand function isolates the substitution effect, thereby removing the Income Effect by holding utility constant.) It uses an indirect utility function $v(\bullet)$ which measures the maximum utility

attainable for given prices and income, rather than mapping directly onto the quantities of goods consumed:

$$v(\mathbf{p}, I) = \max_{\mathbf{p},I} U(\mathbf{x})$$

Where the Marshallian Demand function for X is M_X with arguments for the price vector $\mathbf{p}$ for all prices in the economy and income I, Hicksian Demand function for X is H_X, and indirect utility function id v:

$$\frac{\partial M_X(\mathbf{p}, I)}{\partial P_X} = \frac{\partial H_X(\mathbf{p}, v(\mathbf{p}, I))}{\partial P_X} - \frac{\partial M_X(\mathbf{p}, I)}{\partial I} \cdot M_X(\mathbf{p}, I)$$

Total Change in Marshallian Demand = Substitution Effect + Income Effect

The left side of the equation looks at the change in Marshallian Demand given a constant nominal income I for the consumer and holding all relevant prices *other than* the price of X (P_X) fixed. The middle term measures the change in Hicksian Demand when P_X changes while holding constant all other prices and maintaining the consumer at a constant level of utility, $v(\mathbf{p}, I)$.

The final term looks at the change in Marshallian Demand given the change in purchasing power (or Income). $\frac{\partial M_X(p,I)}{\partial I}$ looks at how the price change results in a change in income that affects the quantity demanded of X. For a price increase, $\frac{\partial M_X(p,I)}{\partial I} > 0$ for normal goods and $\frac{\partial M_X(p,I)}{\partial I} < 0$ for an inferior good. This partial derivative is then multiplied by the Marshallian quantity demanded as it scales the change in wealth (decrease in wealth for a price increase and increase in wealth for a price decrease) based on the initial quantity of goods that were purchased.

Why is this last term negative? This is because of the inverse relationship between the change in price and purchasing power. An increase in P_X *reduces* real income by this quantity. We can see the equation is equivalently written as:

$$\partial M_X(\mathbf{p}, I) = \frac{\partial H_X(\mathbf{p}, U(\mathbf{p}, I)}{\partial P_X} \partial P_X - \frac{\partial M_X(\mathbf{p}, I)}{\partial I} \cdot M_X \partial P_X$$

Thus, when there is a decrease in P_X, that is $dP_X < 0$, this final term is multiplied by that negative which is canceled out by the subtraction, and this change is an *addition* to the change in quantity demanded.

2.4.4 Laws of Demand

Regardless of which system we use, we generally see a downward sloping demand curve. The Three Laws of Demand are:

1. Demand is Downward Sloping. As the price rises, the quantity demanded falls.
2. The Price Elasticity of Demand is greater in the longer-run (as consumers are able to make more adjustments to a change of price).
3. As a fixed cost increases, the quantity demanded of higher quality goods relative to lower quality goods will increase.

2.4.5 Market Demand Curve

The Market Demand Curve is simply the horizontal sum of the individual demand curves for a given product. Since the y-axis is our price, this horizontal sum is asking: for a given price, how many units of the good are demanded by each individual? For the individual demand curves in this figure below, a price P_1, individuals a, b, and c demand their respective quantities Q, so the market demand at P_1 is $\sum Q_N = Q_a + Q_b + Q_c$.

As the price rises, only individual C demands the good above P_3, so the market demand curve consists exclusively of his curve beyond P_3. Between P_2 and P_3, individual a does not demand any of the good, so only individual b and c demand curves are summed.

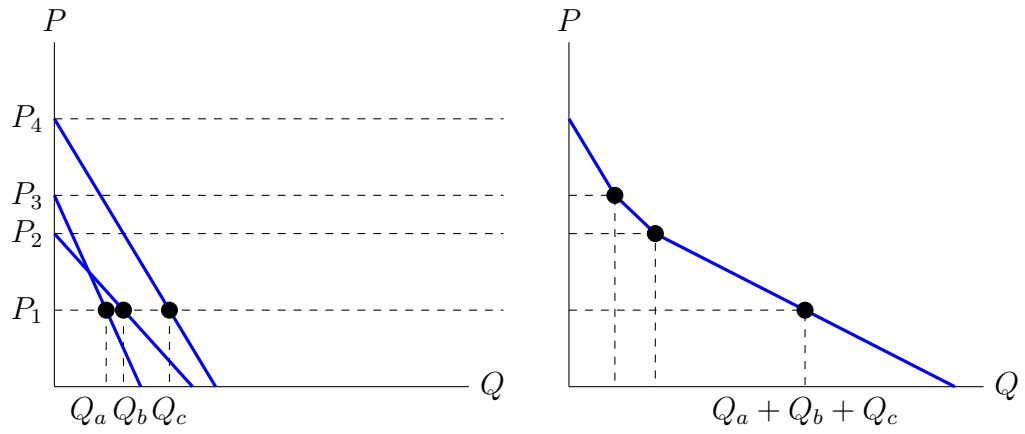


Figure 2.17: Individual Demand Curves (left) and Market Demand Curve (right)

Because of the challenge of negative numbers, each individual demand curve is technically, $Q = \max(P(Q), 0)$, so if we take $P(Q)$ and rearrange the variables to get $Q(P)$, we can only sum the $Q(P)$ equations *if* we exclude negative values.

2.4.6 Consumer Surplus

Consumer Surplus measures the area under a demand curve but above the price paid. It is the difference between the amount that consumers would be willing to pay and the amount they actually paid. For instance, if all consumers pay the same price, for P_1, only Q_1 units of X will be purchased and the green shaded area will represent the Consumer Surplus. If the price is lowered to P_2, now the same consumers who would've purchased Q_1 units have more surplus—the light blue area. They were willing to purchase the units at P_1 (or higher) but now only must pay P_2. The orange shaded area reflects surplus of new customers who are now willing to purchase (more) units of X because they valued the next unit between P_1 and P_2. At P_2, the total consumer surplus is all three shaded areas.

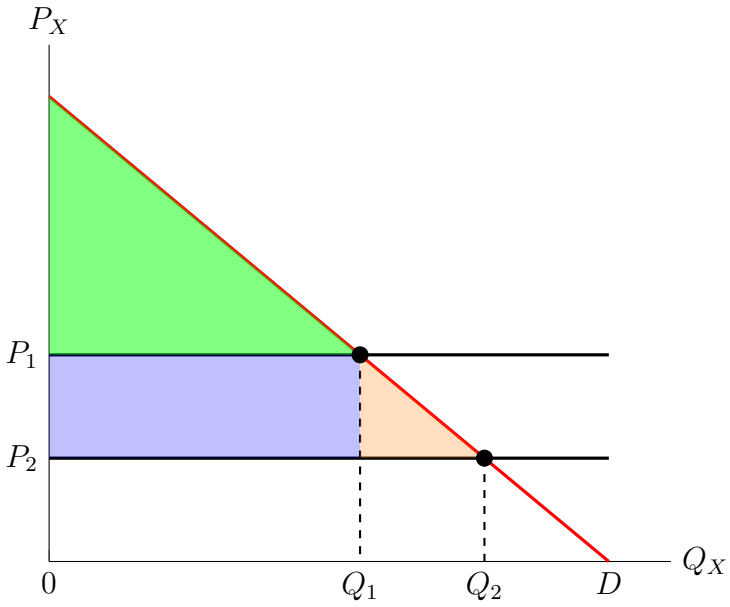


Figure 2.18: Consumer Surplus

Mathematically, Consumer Surplus CS is:

$$CS(P) = \left(\int_0^{Q(P)} D(q)\delta q \right) - P \cdot Q(P)$$

where $D(q) = P$. This is equivalent to the amount consumers value the good minus the price paid.

Waugh (1944) notes that Consumer Surplus is a convex function. For a function to be convex, the second derivative must be positive ($f''(x) \geq 0$).

For Consumer Surplus, the first derivative is:

$$CS' = \frac{d}{dP}\left(\int_0^{Q(P)} D(q)\delta q\right) - \frac{d}{dP}\left(P \cdot Q(P)\right)$$

Per Leibniz Integral Rule:

$$\frac{d}{dP}\left(\int_0^{Q(P)} D(q)\delta q\right) = D(Q(P)) \cdot \frac{dQ}{dP} = P \cdot \frac{dQ}{dP}$$

So:

$$CS' = P \cdot \frac{dQ}{dP} - \frac{d}{dP}(P \cdot Q(P)) = -Q(P)$$

Thus, the derivative of consumer surplus is the negative of the demand curve.

Then the second derivative is:

$$CS'' = \frac{d}{dP} - Q(P) = -\frac{dQ}{dP}$$

This is simply the negative of the slope of the demand curve. Since the demand curve is downward sloping, this double negative becomes positive.

For a convex function, Jensen's inequality says that $E[g(X)] \geq g(E[X])$—that is, a weighted average of the function values is greater than the function of the weighted average. Thus, Waugh (1944) notes that price instability, where the price fluctuates around the optimal price—sometimes higher, sometimes lower—actually *increases* consumer surplus compared with a steady price at the optimal value P^*, even if the *average price* (the price's expected value) is equal in both scenarios.

$$E[P] = \frac{P_L + P_H}{2} = P^*$$

$$CS_{AVG} > CS_{P^*}$$

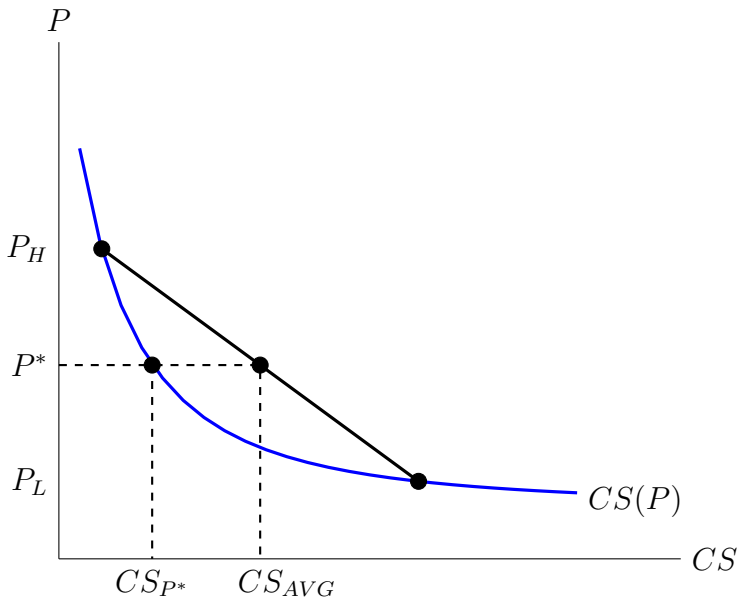


Figure 2.19: Consumer Surplus Function

References

Alchian, Armen A. 1953. "The Meaning of Utility Measurement," *American Economic Review*, Vol. 43, No. 1, pp. 26-50.

Arrow, Kenneth J. 1950. "A Difficulty in the Concept of Social Welfare," *Journal of Political Economy*, Vol. 58, No. 4, pp. 328-346.

Bailey, Martin J. 1954. "The Marshallian Demand Curve," *Journal of Political Economy*, Vol. 62, No. 3, pp. 255-261.

Becker, Gary S. 1962. "Irrational Behavior and Economic Theory," *Journal of Political Economy*, Vol. 70, No. 1, pp. 1-13.

Boudreaux, Donald. 2023. "Economists' Meaning of 'Consumption,'" *The Daily Economy*.

Buchanan, James M. 1979. "Natural and Artifactual Man," *What Should Economists Do?*, Liberty Fund, pp. 93-112.

Caplan, Bryan. 2000. "Rational Irrationality: A Framework for the Neoclassical-Behavioral Debate," *Eastern Economic Journal*, Vol. 26, No. 2, pp. 191-211.

Caplan, Bryan. 2001. "Rational Ignorance versus Rational Irrationality," *Kyklos*, Vol. 54, No. 1, pp. 3-26.

Caplan, Bryan. 2006. "The Economics of Szasz: Preferences, Constraints, and Mental Illness," *Rationality and Society*, Vol. 18, No. 3., pp. 333-366.

Cowen, Tyler. 1989. "Are All Tastes Constant and Identical?: A Critique of Stigler and Becker," *Journal of Economic Behavior & Organization*, Vol. 11, No. 1, pp. 127-135.

Fremling, Gertrud M. and John R. Lott Jr. 1996. "The Bias Towards Zero in Aggregate Perceptions: An Explanation Based on Rationally Calculating Individuals," *Economic Inquiry*, Vol. 34, No. 2, pp. 276-295.

Friedman, Milton. 1949. "The Marshallian Demand Curve," *Journal of Political Economy*, Vol. 57, No. 6, pp. 463-495.

Haltiwanger, John and Michael Waldman. 1985. "Rational Expectations and the Limits of Rationality: An Analysis of Heterogeneity," *American Economic Review*, Vol. 75, No. 3, pp. 326-340.

Lancaster, Kelvin J. 1966. "A New Approach to Consumer Theory," *Journal of Political Economy*, Vol. 74, No. 2, pp. 132-157.

Muth, John F. 1961. "Rational Expectations and the Theory of Price Movements," *Econometrica*, Vol. 29, No. 3, pp. 315-335.

Robbins, Lionel. 1938. "Interpersonal Comparisons of Utility: A Comment," *The Economic Journal*, Vol. 48, No. 192, pp. 635-641.

Stigler, George J. and Gary S. Becker. 1977. "De Gustibus Non Est Disputandum," *American Economic Review*, Vol. 67, No. 2, pp. 76-90.

Savage, Leonard J. 1954. *The Foundations of Statistics*, John Wiley.

Waugh, Frederick V. 1944. "Does the Consumer Benefit from Price Instability," *Quarterly Journal of Economics*, Vol. 58, No. 4, pp. 602-614.

Chapter 3

Exchange

3.1 Efficiency

Efficiency is the key metric of evaluation in economic analysis. A primary use of efficiency is to measure the overall consequences—gains and losses to each individual—of various policies in terms of willingness-to-pay.

3.1.1 Pareto Efficiency

Pareto Efficiency is a key measure of efficiency. A situation is Pareto Efficient if no person can be made better off without making someone else worse off. A state of the world that is Pareto Efficient has no missed opportunities. A Pareto Improvement moves us from an inefficient state closer to a Pareto Efficient state by making at least one person better off without making anyone worse off. A Strict Pareto Improvement makes *all* individuals better off and none worse off, while a Weak Pareto Improvement makes at least one person better off while not making anyone else worse off—that is, everyone is at least as well-off, but some may be no better than before.

In the real world, almost any change hurts *someone*. For instance, if you move across the country and agree to sell your old home to me, then the house guests that you would have had at the old home will be made worse off, unless I am also willing to house and host them. Even if I host them, if I make changes to the decor such that they find less attractive, they have been made worse off. In instances like this, it is incredibly difficult to make sure that everyone is compensated and this has led many to write that the real world is Pareto Efficient.

3.1.2 Potential Pareto Efficiency: The Kaldor-Hicks Approach

If everything is Pareto Efficient, then it is difficult to evaluate improvements to a state of the world based on efficiency; thus, many economists turn to a weaker version: Potential Pareto Improvements. Potential Pareto Improvements or Kaldor-Hicks Improvements allow for both winners and losers, but—hypothetically—the winners would be willing to compensate the losers and would still gain on net—or, at least, be equally well-off. Because such a compensation need not happen, a Kaldor-Hicks Efficient state is not necessarily a Pareto Efficient state.

In the Kaldor-HIcks Efficient State, the dollar value of resources is maximized. Thus, when measuring Kaldor-Hicks Efficiency, one is conducting a cost-benefit analysis. In this analysis, pure transfers are neutral. After all, the benefit recipient receives precisely what the donor gives up. Yet the methods of transfer often involve some various distortionary effects. Rarely are actual transfers merely a pure transfer of resources between two parties. *Taxes* are not costs—and tax cuts are not benefits—rather the activities to enforce, comply with, and evade tax law are the costs, while the benefit is the gained surplus.

When a state is not Kaldor-Hicks Efficient, the gains from the missed opportunities—where at least one person's willingness to pay *exceeds* another person's willingness to accept—are said to be deadweight loss. Deadweight loss is the potential surplus that would be recognized *if* the mutually-beneficial trade took place. For instance, a per unit tax on goods would create deadweight loss, because there is a difference between the price paid and price received. Trades will take place where

$$WTA - T \geq WTP$$

So the lost surplus is where the distance between the willingness to pay and willingness to accept is less than the tax.

$$WTA - WTP \leq T$$

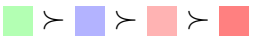
For instance, I may be willing to receive \$1,000 in exchange for my car and you may be willing to pay \$1,100 for the car. This is a Potential Pareto Improvement and a sale price will fall within this range of \$1,000 to \$1,100. However, if there is a 20% tax on the sale price of the car, a wedge is created between our willingness to pay and willingness to accept. If I must pay the tax, I will not be willing to accept less than \$1,250 (with a 20% tax being

$250 making the final after-tax receipt of $1,000). If you must pay the tax, you will not be willing to purchase for more than $917 (with a 20% tax of $183 and the final money price for the vehicle of $1,100).

The Scitovsky Paradox notes that, after a redistribution of the ownership of real resources, relative prices and incomes will change and could result in circularity. Prior to redistribution, we might see that state of the world $N \succ M$ where the winners could compensate the losers. However, after redistribution, due to shifts in the relative prices and incomes, we might find that $M \succ N$ and that the winners could still compensate the losers. To avoid this paradox, Scitovsky argued that, to be meaningful, Kaldor-Hicks Improvements must pass both a Forward and Backward Test. The Forward Test requires that a move from the current state of the world to a hypothetical alternative be a potential Pareto Improvement. The Backward Test requires that a move from such a hypothetical state of the world back to the original state must *not* be a potential Pareto Improvement.

3.1.3 Examples of Changes in Efficiency

To illustrate the two types of hypothetical improvements, consider a train with a $3x2$ seating arrangement. We assume identical preferences. Each customer prefers to have easy access to the aisle to leave at their stop. However, they strongly dislikes sitting immediately next to another person, due to the small seat size. Being *both* blocked and squished—the middle seat or a window seat with someone in the adjacent seat—are disliked. However, no person benefits from more than two seats to themselves—thus, on a three-seater, a person sitting in an aisle seat is not bothered by a person sitting at the window. We can illustrate these preferences as:



In the figure below, we can see that given the initial arrangement as the left-hand side, the person sitting in 1B, 3B, and 5D would benefit from moving to 1C, 3E, and 4A, respectively. In fact, given the known preferences, we would never expect such an arrangement as in the left-hand side to appear.

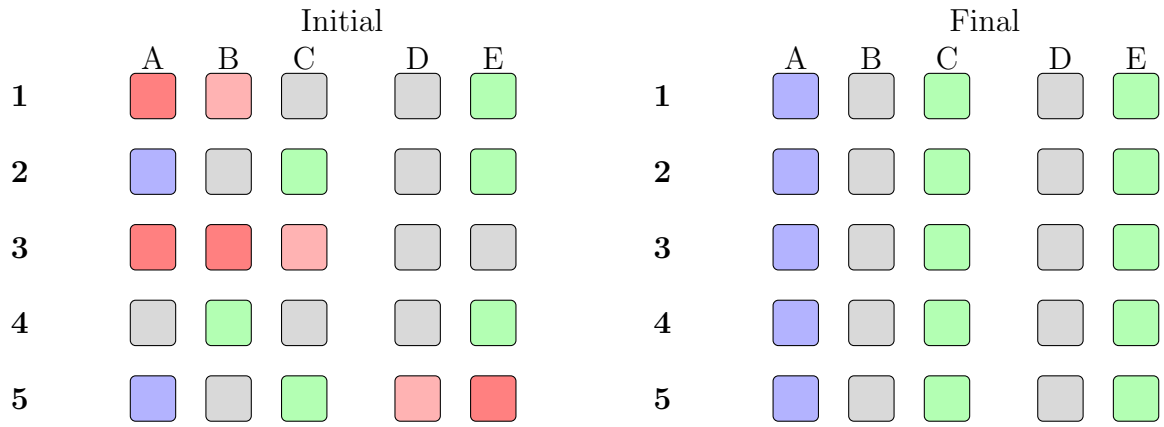


Figure 3.1: Pareto Improvement on a Train

Once the right-hand configuration is achieved, no more improvements are possible. We can engage in transfers, switching the individual in 3A and 3E, improving the lot of the former occupant of 3A, giving him easy access to the aisle, while making the former occupant of 3E equally worse by the occupants of 3C blocking him in. This would be making someone better off, while making someone else worse off. The potential gains have been exhausted.

Now, for evaluating a Kaldor-Hicks Improvement to this scenario, we must assign some willingness to pay associated with the preferences. For simplicity, dark red seats are valued at \$1, red valued at \$2, blue at \$3, and green at \$5. We can see that the Pareto Improvement in Figure 3.1 is also a Kaldor-Hicks Improvement. The original configuration is valued at \$46, while the final configuration in Figure 3.1 is valued at \$65. For instance, the individual sitting in seat 1A would be willing to pay the individual in 1B up to \$2 to move to seat 1C—and the individual in 1B *already* values seat 1C \$1 more than seat 1B. This Pareto Efficient state happens to be the Kaldor-Hicks Efficient state. As already mentioned, moving any single individual would make another worse off.

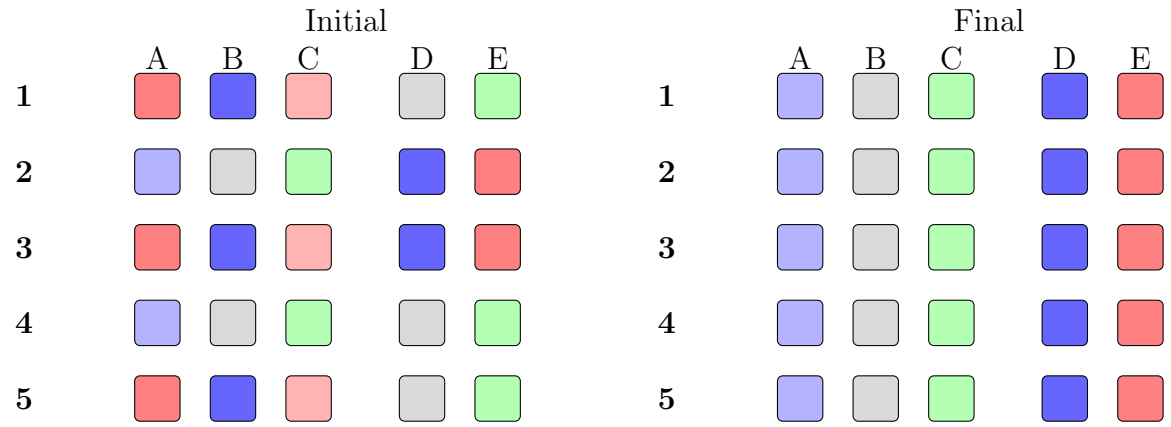


Figure 3.2: Kaldor-Hicks Improvement on Train

Let us consider, however, that 5 more individuals get on the train. These five new individuals are indifferent to any seat which they value at \$3 (which is already being represented by the blue square, but they will be distinguished by the darker blue), thus they make the left configuration. This has a total value of \$54. However, we could raise this by moving our three individuals out of the middle seats and into the aisle seats of the two-seaters. This improvement is *not* a Pareto Improvement, because there are both winners and losers. This makes those sitting in 1E, 4E, and 5E worse off, but those sitting in 1A and 1C, 3A and 3C, and 5A and 5C better off by more. In fact, the two individuals who were already sitting in the three-seaters could be willing to pay the new individuals to not sit in their middle seat. The individuals in the A seats gain \$2 in value and those in C gain \$3 in value from the movement of the individual out of the B seat. Thus, they would be willing to pay up to \$5. The individual sitting in the E seat would be indifferent between an occupant in the D seat and \$4. Thus, (for rows 1,3, and 5) the individuals in A and C can pay the individual in E somewhere between \$4 and \$5 and everyone ends up better off. When compensated in this way, all individuals would prefer the right-hand state of the world while making or receiving the payment to the left-hand state of the world.

But note, in order to call this a Kaldor-Hicks Improvement, these individuals need not *actually engage* in this exchange; the exchange need only be *possible*. It is a *Potential*, but not Actual Pareto Improvement unless such an exchange takes place. This KH Improved configuration results in a total value of \$60.

However, we can do one better. Because those sitting in the E seat are their least preferred seat, while those in D are indifferent to their seat, we could switch each individual in E with each individual in D, gaining \$1 of value for

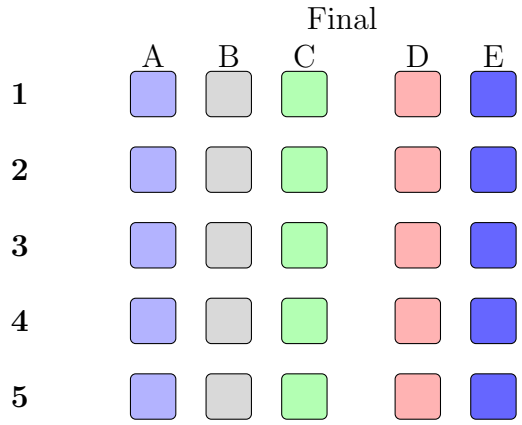


Figure 3.3: Kaldor-Hicks Efficient Train

each of the original occupants of an E seat (moving from valuing the seat at \$1 to \$2). This is only a Pareto Improvement because there is no loss in this change.

Moving from the right-hand configuration in Figure 3.2 to the configuration in Figure 3.3 is both a Kaldor-Hicks *and* Pareto Improvement. Those who go from D to E are made no worse off—as they are indifferent between the seats—but the occupants of E to D are made better off. However, note, in comparison to the initial configuration of Figure 3.2, this is *not* a Pareto Improvement, as the occupant of 5E, for instance, is made worse off as the value of his seat is initially \$5 and he ends with a seat he values at \$2.

It is at this point, no further Kaldor-Hicks Improvements are possible and the resources are maximized at a value of \$65, which happens to be the same value as the train carrying our 15 original individuals that is now more crowded and transporting 20 individuals. This also means, however, that our individuals are indifferent between the final configuration of Figure 3.1 and Figure 3.3. Consider, for instance, that the occupants of the D seats would be willing to pay up to \$3 for the new individuals to avoid taking the train, while the occupants of seat E would accept as little as \$3 to give up their seat on the train. At the exact \$3 valuation, each individual is just indifferent between the options: suffering in a squished seat vs a monetary cost with the benefit of more space or taking the train vs a monetary benefit and not taking the train. But it need not work out this way.

3.1.4 The Limits of Efficiency Analysis

Efficiency is a distinctive feature of economic analysis and is critical for identifying missed opportunities and unintended consequences. Despite its benefits, it is inappropriate to use efficiency alone as a judge against real-world conditions. Demsetz (1969) notes that the choice facing policymakers is never one "between an ideal norm and an existing 'imperfect' institutional arrangement," but rather between "alternative real institutional arrangements." Demsetz calls this fallacious analysis the *Nirvana Approach*.

Central to this analysis is typically at least one of three fallacies. The Grass is Always Greener Fallacy anthropomorphizes government actions as being capable of fixing problems identified in a free market setting, but, when comparing the "free enterprise solution" to an alternative outcome, "the political or non-profit forces that are substituted for free enterprise must be analyzed and the outcome of the workings of these forces must be compared to the market solution before any such conclusions can be drawn." The Fallacy of the Free Lunch examines only some of the costs and benefits and calls the real-world inefficient because these real costs exist. Among the biggest problems is the failure to recognize the real cost of knowledge. Finally, the People Could be Different Fallacy fails to recognize that the same people operating in the market also operate in government and are, thus, subject to the same biases and knowledge limitations. The problem is that it fails to account for real-world conditions, which means it does not distinguish between what can be changed and what can't be changed.

> To say that private enterprise is inefficient because indivisibilities and imperfect knowledge are part of life, or because people are susceptible to the human weaknesses subsumed in the term moral hazards, or because market commodity options is not costless, or because persons are risk averse, is to say little more than that the competitive equilibrium would be different if these were not the facts of life. But they are the facts of life [since] they cannot be erased from life at zero cost...

In each case, various judgments "say little more than that the competitive equilibrium would be different if these were not the facts of life."

This also highlights a very important component of all good cost-benefit analysis. Bastiat (1850) too reminds us how easy it is to only remember some costs. The benefits of government spending, he writes, are easily seen, while the costs of taxation are not. "The expenditure of the civil servant *is seen* because it is carried out, whereas that of the taxpayer *is not seen* as, alas! he is prevented from carrying it out." Likewise, any legitimate

efficiency analysis must include everyone. For instance, when discussing the effects of immigration restrictions on efficiency, there is often a focus on the domestic gains and losses, but this excludes the largest gains for those would-be immigrants. While efficiency analysis treats all considered individuals equally, in terms of their willingness to pay, good efficiency analysis includes *everyone* who is relevant.

3.2 Simple Two-Person Exchange

We can model simple two-person, two-good exchange with an Edgeworth Box. The Edgeworth Box puts two different individuals' indifference curves on top of each other but with flipped axes.

The boundaries of the Edgeworth Box are the real resource limits and the bottom left and top right corners represent the two individuals owning all of both goods, respectively. At each point in the Edgeworth Box, the sum of the goods is the same; each point represents a separate distribution of the given goods. As we move along the x-axis, moving towards the right means giving more of good X to Individual A, which necessarily means taking that X from Individual B. Thus, the reverse is true for moving to the left along the x-axis gives Individual B more of good X.

Individual A prefers indifference curves up and to the right—as this represents a higher utility. On the other hand, Individual B's higher indifference curves are down and to the left.

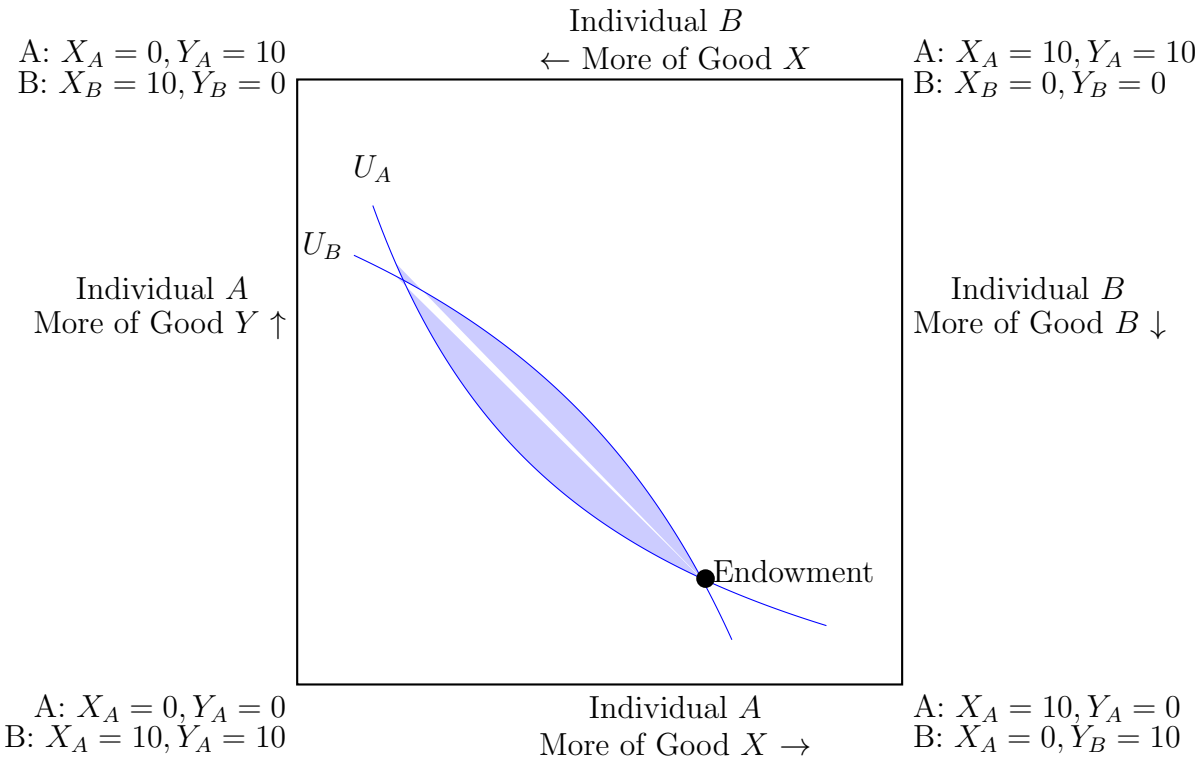


Figure 3.4: Edgeworth Box

For an initial endowment in the figure above and the two given indifference curves, we can see that there are potential gains from trade—a Pareto Improvement—the area shaded between the two indifference curves. Both A and B could gain at any point within the shaded area and move to a higher indifference curve. The only way to eliminate this shaded area of potential gains is for the indifference curves to be just tangent to each other. Until then, we can make both individuals better off.

Note that at each point in tangency, the Marginal Rate of Substitution between the two individuals is necessarily equal.

$$MRS_A = \frac{MU_X^A}{MU_Y^A} = MRS_B = \frac{MU_X^B}{MU_Y^B}$$

The Contract Curve passes through all of these Pareto Efficient points of tangency and shows the potential outcomes. In a market setting—that is, absent redistribution—the initial endowment matters for what points on the contract curve are Pareto improvements, and, thus, what will be available to a given actor. Only the points within the eye of the two initial indifference curves will be available.

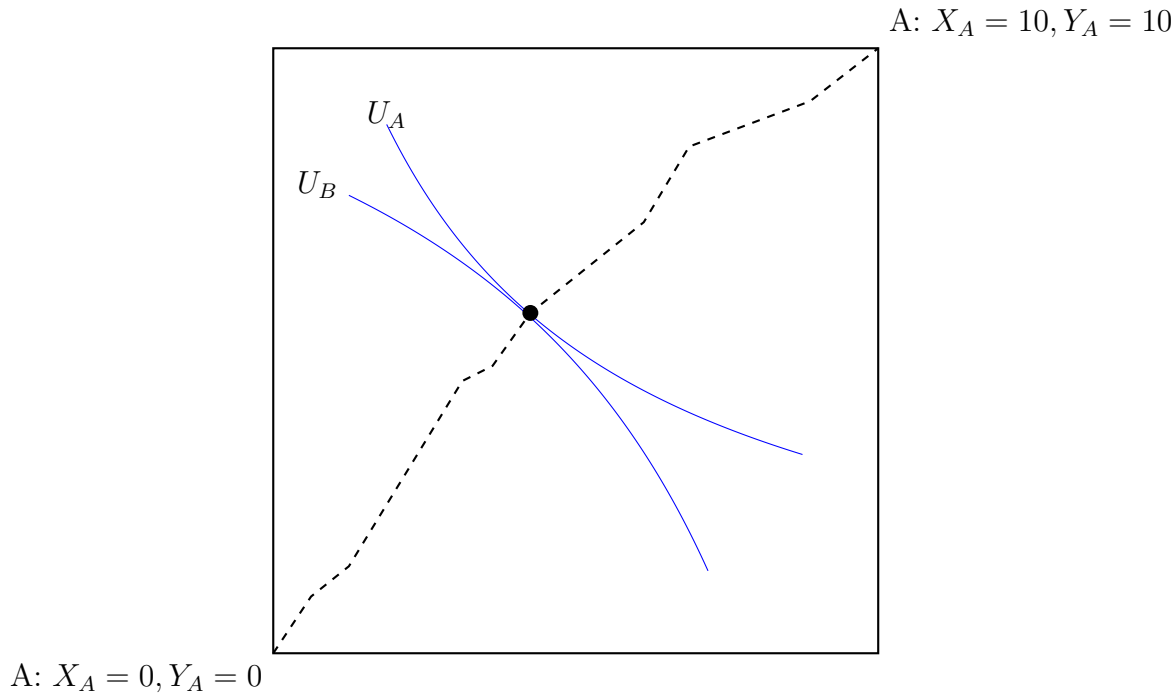


Figure 3.5: Edgeworth Box with Contract Curve

References

Bastiat, Frederic. 2016 [1850]. "What Is Seen and What Is Not Seen, or Political Economy in One Lesson," *Economic Sophisms and 'What is Seen and What Is Not Seen'*, translated by Jane Willems and Michel Willems, edited by Jacques de Guenin, Liberty Fund, pp. 401-452.

Demsetz, Harold. 1969. "Information and Efficiency: Another Viewpoint," *Journal of Law & Economics*, Vol. 12, No. 1, pp. 1-22.

Chapter 4

Production

Consumption is a study of choice among various ends. Production seeks to evaluate the means used to get to these possible states of consumption. Production is not an end in and of itself; it is how humans take given resources and transform them to provide services that they desire to consume. The value of production comes from the consumption opportunities it creates, and this value only exists if what is produced is what people want to consume.

We've looked at the choice of consumption bundles and the benefits of exchange, but from where do we get the budget line of chapter 2 or the endowment of goods for a given period in chapter 3? It is not dropped down from Heaven. We must turn to this study of production to see *how* certain consumer wants are supplied.

A fundamental question is how production decisions are made. For this, we turn to cost. As we have explored earlier, cost is simply a foregone opportunity. Resources have a near-infinite quantity of uses. When resources are used as inputs for one good, they cannot be used as inputs for another good instead. So, the key is ensuring that the inputs go to their highest valued use.

4.1 Comparative Advantage

4.1.1 Individual Production Possibilities

In a very simple model, a single individual has two potential goods that he can make with his labor alone. Think, for instance, of procuring food—whether by hunting or planting—or making crafts out of natural resources. The same time can only be spent on one activity.

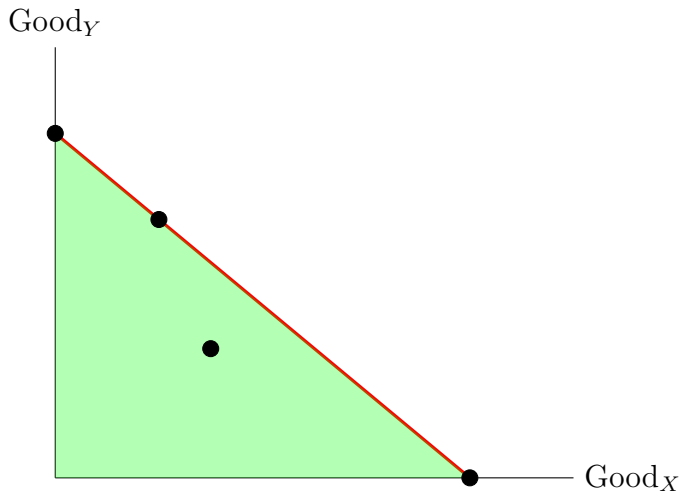


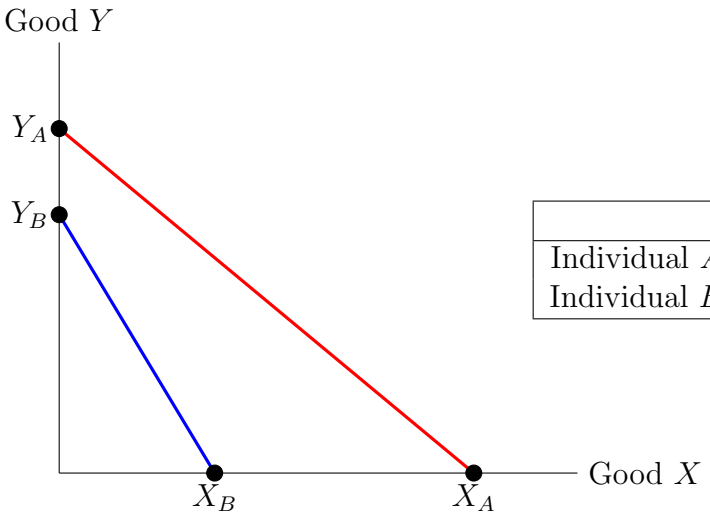
Figure 4.1: Individual's Production Possibilities Frontier

For given human capital—knowledge and skills—and other resources, the individual's trade-off between the two activities looks like this with the red line representing his Production Possibilities Frontier—the possible combinations of the two activities for a given period of time. He can spend all of his time and resources on one or the other or a combination of the two. Any point within the shaded region is feasible, but if the individual is using his resources most effectively, he is on the outer edge of this line.

We can add a second individual with a slightly different set of skills and resources—so that his trade-off between the goods is slightly different and the slope of his production possibilities frontier is different.

In this case, in a single day, person A is better at making both goods—if he specialized in Y or X we would make more than person B, if he specialized in the same. The table shows us the numerical trade-offs. On the surface then, it may seem like A should produce everything or that A could not gain from doing business with the "less efficient" B, but this is not the case.

We can see that, while A may be more skilled across the board, they do not have the same relative trade-off between goods. For person A, the cost of 1 unit of Good X is 1 unit of Good Y. For person B, the cost of 1 unit of good X is only $\frac{1}{2}$ a unit of Good Y. Person B has a comparative advantage to Person A at producing Good X, then.



	Good X	Good Y
Individual A	8	8
Individual B	6	3

Figure 4.2: Two Individuals' Production Possibilities Frontiers and Production Capacity

4.1.2 Joint Production Possibilities

In considering person A and B's joint production possibilities frontier, we can add their curves together. We do this by starting with both individuals specializing only in good Y and then ask which person has the lower cost for making one unit of good X. This is the question of the lower opportunity cost. Since our simple model has a static trade-off between the goods, we can see the comparative advantage remains constant. We can see this by looking down the columns of the table below and seeing the cost in terms of the other good. Thus, person B has the lower opportunity cost for producing good X

	Good X	Good Y
Individual A	1	1
Individual B	2	0.5

Table 4.1: Opportunity Cost Calculation for Two Individuals with Price in Terms of Other Good

and person A has the lower opportunity cost for producing good Y.

Thus, when we add the curves together, we get a Y-intercept of 14. Then, we see that person B has the lower cost of producing good X. For producing 1 unit of X, he only loses half of a unit of Y. For producing a second unit of X, we only lose another half of a unit of Y. So, since the trade-off is constant, we lay out A's curve first. Then, we reach the point of specialization, where

A is producing only good X with the production of 8 units, and person B is producing only good Y with a production of 6 units. At this point, we add person B's curve since, if we want more X, he must begin producing X and A's cost of producing more X at this point is effectively ∞—for all intents and purposes, outside our range of possibilities. The final curve, then, is:

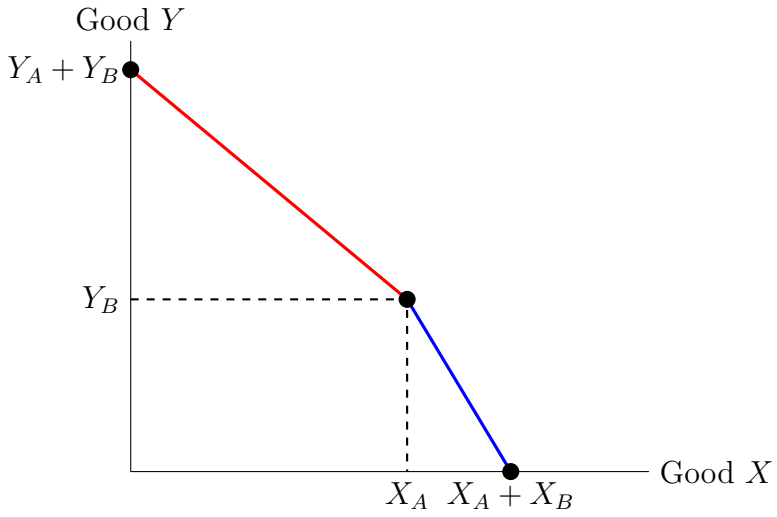


Figure 4.3: Two Individuals' Combined Production Possibilities Frontiers

What happens when we add these curves together is that we actually push out the possibilities of the baskets of goods available to the individuals. Autarky is the condition of no exchange, where each individual only consumes what he produces. Regardless of the given preferences between the goods, we find that both individuals will be better off. Assuming continuous goods—that is, that fraction of units are producible and consumable:

	Autarky			Joint		
Relative Preference	A	B	$\sum$	$\sum$	A	B
1:1	$(4, 4)$	$(2, 2)$	$(6, 6)$	$(7, 7)$	$(7, 1)$	$(0, 6)$
2:1	$(2\frac{2}{3}, 5\frac{1}{3})$	$(1\frac{1}{2}, 3)$	$(4\frac{1}{6}, 8\frac{1}{3})$	$(4\frac{2}{3}, 9\frac{1}{3})$	$(4\frac{2}{3}, 3\frac{1}{3})$	$(0, 6)$
1:2	$(5\frac{1}{3}, 2\frac{2}{3})$	$(2.4, 1.2)$	$(7\frac{11}{15}, 3\frac{13}{15})$	$(8.8, 4.4)$	$(8, 0)$	$(0.8, 4.4)$
Production Ratio	$(4, 4)$	$(1.5, 3)$	$(5.5, 7)$	$(8, 6)$	$(8, 0)$	$(0, 6)$

Table 4.2: Possible Baskets of Goods in Production (Q_X, Q_Y)

Other than at the end points, we find that the quantity produced of both goods is greater than under autarky. If individuals have the same preferences, we can find these points geographically by finding where the line that passes

through the origin with the slope of their preferences k intersects their individual PPFs and the joint PPF.

A final point that is worth noting is that, when an individual specializes, they learn from doing. With more time devoted to a more narrow task, the specialist is more likely to improve their skills expanding out the production possibilities frontier.

4.1.3 Comparative Advantage and Exchange

What will be the relative price of the goods? It will fall somewhere between the two individual's own cost. If the price of good Y was 1 X, then A would be indifferent between making his own X and Y and only person B would gain. Likewise, if the price of good Y was 2 X, then B would be indifferent between making his own X and Y and only person A would gain. So:

$$\frac{P_X^A}{P_Y^A} < \frac{P_X^*}{P_Y^*} < \frac{P_X^B}{P_Y^B}$$

$$1 < \frac{P_X}{P_Y} < 2$$

The farther from one's own price ratio, the more that individual gains. If the settled price ratio in terms of Y is 1.25 (that is, 1 X trades for 1.25 Y and 1 Y trades for 0.8 X), more of the increased production will go to person B. On the other hand, if the settled price ratio is 1.75 (that is, 1 X trades for 1.75 Y and 1 Y trades for approximately 0.57 X), then person A gains more. There is a wide variety of prices that will be acceptable to both individuals, then and result in these net gains, but how they are distributed depends on the precise price.

For instance, if the price settled upon is precisely in between the two ratios A's $1:1$ and B's $1:2$ and is $1:1.5$, this means that if one wants another unit of X, he must give up 1.5 Y and if he wants another unit of Y, we must give up $\frac{1}{1.5}$ or $\frac{2}{3}$ a unit of X.

For the joint production, of $(7,7)$, for instance, B will completely specialize in producing good Y and A will produce one unit of Y and 7 of X. Then, if the surplus is equally divided, the price ratio will be $1:1.167$ as A will be willing to give up 3 of X for 3.5 Y. For A, the per unit cost of each Y is approximately 0.86 X, while, if he made his own, it would be a cost of 1. For B, the per unit cost of each X is 1.167 of Y, while, if he made his own, it would be a cost of 2. We see, both individuals consume more of each good.

	Autarky			Joint		
Relative Preference	A	B	$\sum$	$\sum$	A	B
1:1	$(4,4)$	$(2,2)$	$(6,6)$	$(7,7)$	$(4.5,4.5)$	$(2.5,2.5)$
2:1	$(2\frac{2}{3},5\frac{1}{3})$	$(1\frac{1}{2},3)$	$(4\frac{1}{6},8\frac{1}{3})$	$(4\frac{2}{3},9\frac{1}{3})$	$(2\frac{11}{12},5\frac{5}{6})$	$(1\frac{3}{4},3\frac{1}{2})$
1:2	$(5\frac{1}{3},2\frac{2}{3})$	$(2.4,1.2)$	$(7\frac{11}{15},3\frac{13}{15})$	$(8.8,4.4)$	$(5\frac{13}{15},2\frac{14}{15})$	$(2\frac{14}{15},1\frac{7}{15})$
Production Ratio	$(4,4)$	$(1.5,3)$	$(5.5,7)$	$(8,6)$	$(5.25,3.5)$	$(2.75,2.5)$

Table 4.3: Possible Baskets of Goods in Consumption (Q_X, Q_Y) with Equal Surplus

For a strong preference towards Y—the $2:1$ preference ratio—both A and B make some Y, but A also makes some X. A makes more X than he wants and thus is willing to trade with B who has specialized in Y. A and B would both be willing to trade $2\frac{1}{6}$ units of X in exchange for $2\frac{14}{15}$ units of Y. The terms of trade are 1 X for 1.38 Y or 1 Y for 0.727 X.

The point of pure specialization on the graph where $X = X_A$ and $Y = Y_B$ is the kink point. The relationship of the preference ratio to the kink point will determine where joint production takes place. For A as the relatively lower cost producer of X and B as the lower cost producer of Y and a given preference ratio k that is the same between two individuals of $\frac{Y}{X}$ (or $Y:X$):

- If $\frac{Y}{X} > \frac{Y_B}{X_A}$, B will specialize but A will produce both goods.
- If $\frac{Y}{X} < \frac{Y_B}{X_A}$, A will specialize but B will produce both goods.
- If $\frac{Y}{X} = \frac{Y_B}{X_A}$, both A and B will specialize and the kink point (X_A, Y_B) is the optimal point.

If preferences are heterogeneous, that is $k_A \neq k_B$, the decision rule is complicated.

- If $k_A, k_B > \frac{Y_B}{X_A}$, B will specialize but A will produce both goods.
- If $k_A, k_B < \frac{Y_B}{X_A}$, A will specialize but B will produce both goods.

However, if $k_i > \frac{Y_B}{X_A}$ and $k_j < \frac{Y_B}{X_A}$, the result will depend on individual wealth.

4.2 The Nature of the Firm

One of the key objects of analysis in production is the **Firm**. A firm is simply an artificial organization designed to take advantage of team production. It is a formalized arrangement between various individuals for the purpose of producing final goods and services. Coase (1937) writes that "a firm,

therefore, consists of the system of relationships which comes into existence when the direction of resources is dependent on an entrepreneur." But Alchian and Demsetz (1972) write that a firm "has no power of fiat, no authority, no disciplinary action any different in the slightest degree from ordinary market contracting between two people."

Why does the firm arise? Uncertainty is a key reason. The firm owners take on risk and uncertainty in exchange for a share of profit, while the employees take advantage of a stable income and the information from others who gather information about conditions and determine a production plan. Coase (1937) adds that use of the price mechanism for coordination is not costless. The firm is also able to lower the costs of organizing. Instead of each stage and piece of production involving separate contracts for each exchange, the entrepreneur takes control of the factor inputs, including labor. By having a single contract across a longer span of time, the entrepreneur saves the costs of many separate contracts. By maintaining general contracts in terms of what service laborers, for instance, are expected to produce, the entrepreneur has flexibility in how to use the inputs and respond to changing conditions rather than contracting out individual and highly specific tasks.

Coase (1988) summarizes his argument writing, "A firm had therefore a role to play in the economic system if it were possible for transactions to be organized within the firm at less cost than would be incurred if the same transactions were carried out through the market." The fundamental argument is rooted in a comparative institutional analysis:

> the gains which accrue through the existence of the firm... come from a reduction in transaction costs... those which would otherwise have been incurred in market transactions between the factors now cooperating within the firm. It is the *comparison* of these costs with those that would have to be incurred to operate a firm which determines whether it would be profitable to establish a firm.

Alchian, Crawford, and Klein (1978) argue that vertical integration is the solution to the moral hazard of some longer-run contracts. For instance, some machinery or other inputs into the production process may be so specialized that, "once the investment in [a particular] asset is made, the asset may be so specialized to a particular user that monopoly or monopsony market power, or both, is created." The owner of the capital good input could renege on his initial contract, such as by alleging high maintenance costs. While long-term contracts are a possible solution, legal remedies can be quite costly. They cite the real world example of General Motors, which had contracted with a third party to make their closed auto bodies. But, after believing that the third

party was abusing its power by charging higher quasi-rents—the return for a current use greater than its value in its next best use to another renter—and so GM eventually took them over in a corporate merger and vertical integration was the solution to the (potential) exploitation of quasi-rents.

Coase (1937) suggests that there is eventually diminishing returns to organization, which is a key reason why the entire economy's production activities are not subsumed into a single firm. Coase (1988) writes that, perhaps, "people working in a large firm may find the conditions of work less attractive than in a small firm and therefore will require a higher remuneration to compensate them for this."

Alchian and Demsetz (1972) argue that firms can also reduce the costs of monitoring. Absent costless monitoring, "each input owner will have more incentive to shirk when working as part of a team than if his performance could be monitored easily or if he did not work as a team." They argue that firms arise in industries where innovation makes team-oriented production more efficient. Thus, the firm allows for a system of input organization and policing of individual working. It is very difficult to put out for a contracting bid tasks that require joint effort on the part of more than one party that must come together to produce the output. While classical theory suggests paying factors in relation to their marginal productivity, how do you measure this when dealing with two individuals lifting heavy cargo?

Williamson (1981) lays out four types of relationships of labor within a firm based on the specificity of the human capital and skills and the difficulty of ascertaining—or metering—productivity.

		Human Assets	
		Nonspecific	Highly Specific
Metering	Easy	Spot Market	Obligational Market
	Difficult	Primitive Team	Relational Team

Figure 4.4: The Governance of Internal Organizations

When it is easy to measure productivity and human capital is non-specific, we expect workers to easily move between firms and little to no special structure for this relationship, since the employers have no training start-up costs, for instance. Migrant farm workers, custodians, and even engineers can fit this criteria, leading to a spot market. When productivity is difficult to measure but the actual skills and human capital is non-specific, a team is necessary, but compensation is difficult to determine on an individual basis. When metering is easy, but human capital is highly specific

requiring firm-specific learning, both firms and workers want to maintain continuity in the employment relationship. In this setting, we expect various safeguards against arbitrary dismissal (protecting the worker) or non-vesting in retirement and accrual of benefits across time to discourage unwanted quitting (protecting the employer). Finally, when productivity metering is difficult *and* the human capital is highly specific, the firm will be more invested in ensuring that the employees are working, so they will turn to social conditioning seeking to ensure employees are dedicated to the work of the firm, while, at the same time, due to the highly specific nature of their human capital, employees will want considerable job security.

Alchian and Demsetz (1972) likewise suggest that different types of work will lead to different structures of a firm. Where behavior monitoring is easy and self-policing, they expect firms to be mostly profit-sharing where all team members share in the profits. Where much capital is necessary, the corporation is likely to arise with shareholders who provide capital but delegate decision-making authority to a management group whose membership and key decisions are subject to revision by the shareholders. Where the cost of monitoring is high, such as artistic and professional employments, we would expect partnerships rather than employer-employee relationships, because the cost of management is higher than its benefit. Within the role of the firm, unions are another possible monitor for both employees and employer performance.

Alchian and Demsetz (1972) write that non-profit firms are likely to have more shirking, because the future consequences of improved management are not capitalized into present wealth of stockholders, making it more difficult for competition by new would-be monitors. Socialist firms, where the state dictates employee ownership, there is likewise a need to reduce shirking by a central monitor. In countries like Yugoslavia, we find worker committees can terminate a manager's contract, likely due to the incentives for excessive shirking.

4.3 Marginal Cost

4.3.1 Rising Marginal Cost

There are two reasons that marginal cost is *necessarily* rising at some point, as we seek to produce more units of a particular good.

Central to a discussion of cost is that cost is the foregone alternative. If there is a diminishing marginal rate of substitution, as you move along the budget constraint of your own production possibilities, there is a rising cost—that

is, absent a corner solution, you eventually move to lower indifference curves because you have so few of the other good. Cost as subjectively evaluated by the individual must rise. Even given a constant trade-off—a linear budget constraint (or individual production possibilities frontier), the indifference curves leave you with an increasing cost. Producing the maximum Y leads to a lower utility than producing a mix of X and Y.

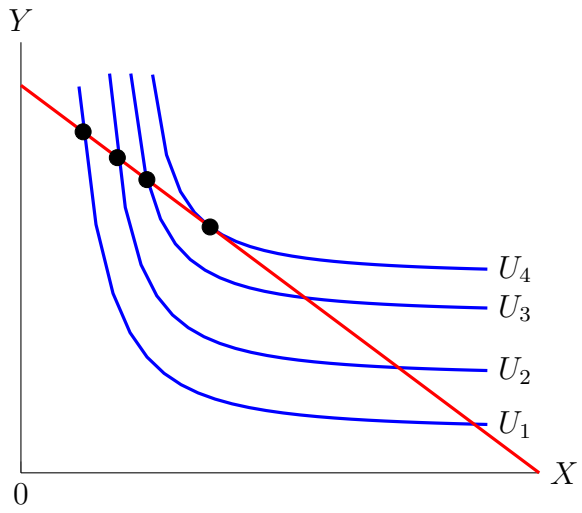


Figure 4.5: Indifference Curves and Rising Cost

We can generalize this beyond a single good. The production of X is compared against the production of every other good. What is most relevant is the opportunity cost: if you were not producing X, what would you be producing? What is the *next, best* alternative?

When you are producing for something other than your own consumption, your trade-off is no longer limited to your own production possibilities. Now, it is dependent upon the terms of trade for the good(s) which you produce in terms of every other good produced in the economy. An individual likely wouldn't want to *consume* only units of good Y, however, *producing* only units of good Y may get him the most in terms of purchasing power compared to his alternative of producing good X.

For simplicity, it is easiest to calculate your opportunity cost in a common good. This is especially true when we move beyond two goods. The prices are no longer a ratio between two goods but between one good and all other goods. Instead of calculating a trade-off between producing art and producing food, furniture, clean water, and entertainment, the producer picks one as the unit of account.

For instance, individual B in the example above could choose good Y as his unit of account. His trade-off between Y and X is $2:1$. Under autarky, his cost of X is 2. So, if $P_X < 2$, B can arbitrage by producing Y and trading it for X increasing his consumption, as we saw with comparative advantage. He will be willing to trade for any price under 2, though the quantity traded will depend on how much lower, his overall wealth, and the preferences of others.

However, once we trade, in addition to the subjective cost increase, there is an objective change in the trade-offs. Beyond the kink point, there is a marginal cost greater than prior to the kink point. Consider that the total cost in terms of foregone Y is the flipped version of the production possibilities frontier.

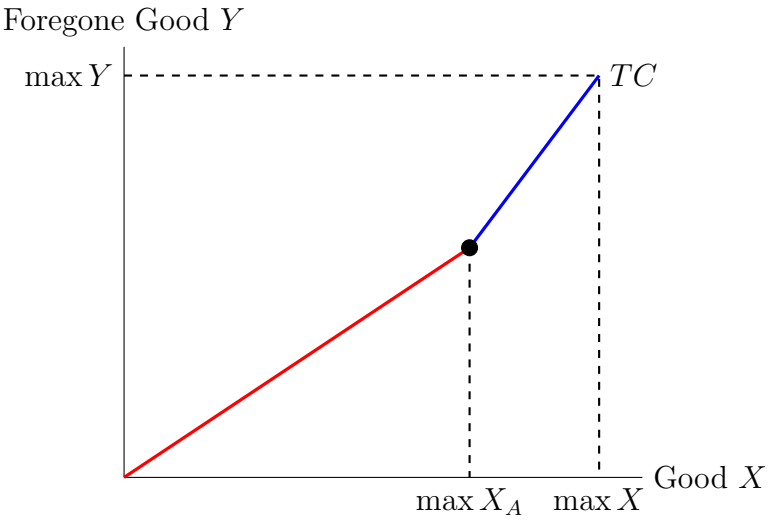


Figure 4.6: Two Individuals' Total Cost Curve

When we move beyond the lowest cost producer, the marginal cost is higher. Consider that the Marginal Cost is the derivative of the Total Cost Curve.

$$MC = \frac{d}{dC}TC$$

For this linear function where the trade-off is constant between X and Y for each individual, marginal cost is constant for the length of that individual's production, then discontinuous when switching individuals.

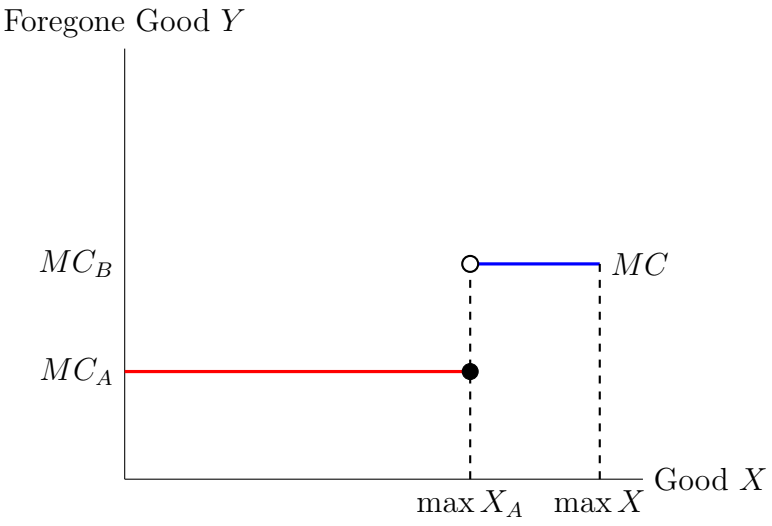


Figure 4.7: Two Individuals' Marginal Cost Curve

As we begin adding more potential producers (and more goods), we see that the marginal cost curve is steadily rising. When adding four individuals, for instance, we see the same process continues with a rising marginal cost curve of X in terms of foregone Y (P_X^Y).

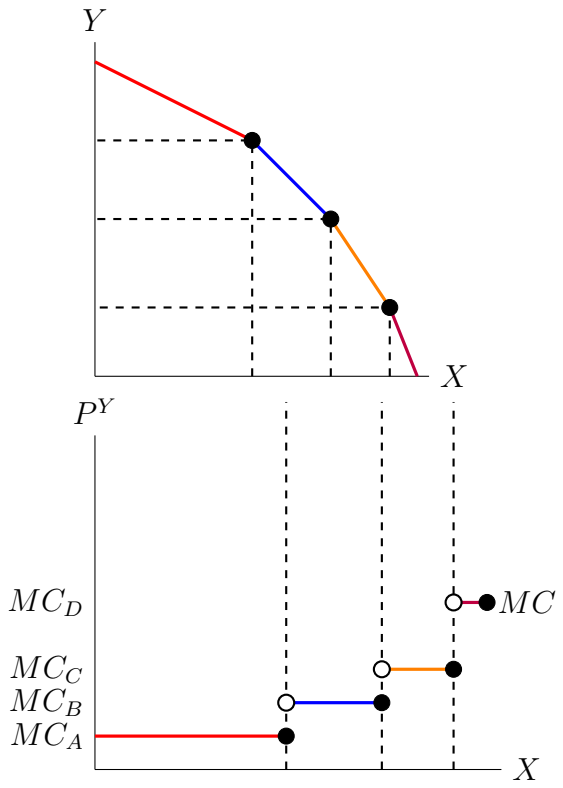


Figure 4.8: Marginal Cost Curve of X

As we generalize to add more goods, the y-axis becomes foregone "all other goods" in terms of some unit of account. This is where money arises. To simplify calculation (unit of account) and to ease transactions (medium of exchange), money can serve as the y-axis quantity for each individual good. Yet, the nature of the price has not changed—it still represents the foregone alternatives. Now, it is the *money value* of the foregone alternatives.

4.4 Cost Functions

4.4.1 The Production Function

The Production Function maps how a producer turns inputs into output. A key variable in the production function is returns to scale. Returns to scale measures the responsiveness of output to varying all inputs proportionately. For production function $f(\bullet)$, constant $t > 1$ and inputs x:

- Constant Returns to Scale – $f(tx) = t \cdot f(x)$. As the inputs increase, output increases proportionately. Double the inputs, and you double the outputs.
- Increasing Returns to Scale – $f(tx) > t \cdot f(x)$. As the inputs increase, output increases *more than* proportionately. By scaling up your inputs to produce more units, you can produce more output per a given input. This is usually when increasing the scale of production involves changing the methods of production, such as through increasing the division of labor or simplifying the production process. Industrialization is often thought of as increasing returns, but this is ultimately because it changes the production function.
- Decreasing Returns to Scale – $f(tx) < t \cdot f(x)$. As inputs increase, output increases, but by *less than* 1. For instance, if the inputs are doubled, output increases by only 80%. Why might we have decreasing returns to scale? If all factors are variable, it is possible that the increasing management layers within a firm lead to more challenges that result in decreasing returns. Thus, this is decreasing returns within a fixed management structure which remains unchanging. However, the decreasing returns scenarios are extremely counter-intuitive—surely we could just repeat the identical process again to get the same result, right?—because many so-called decreasing returns to scale are actually a misrepresented presentation of diminishing marginal returns.

Diminishing Marginal Returns is explicitly focused on the returns of increasing one input while holding all other inputs constant. Note that the assumption of scarcity requires that at *some* point, inputs are limited, so we will end up with a scenario where at least one input cannot be increased at the same cost. At some point, not *all* inputs can be doubled. For instance, the ownership oversight is certainly an input into production, and yet this does not double when a single firm increases all *other* inputs. A firm at such a point would almost universally be said to have decreasing returns. Yet, so long as we treat ownership oversight as an input, this is *really* a case of diminishing marginal returns. More obviously, we can increase all of out physical inputs and labor inputs for a good, but if we remain on a fixed plot of land or within a fixed structure, eventually crowding will lead to what may appear to be decreasing returns. Thus, mathematically, many of these decreasing returns scenarios can be better thought of as a composite function F with constant returns to scale production function f and a special input z that is not on the relevant range (or cannot be) increased: $F(z, x) = zf(\frac{x}{z})$.

To the extent that there remain unused—or even undiscovered—methods of production, because of the existing scale of the market mean that the removal

of trade barriers, by extending the market, could have disproportionately high returns. Buchanan (1992) also notes that this leads to an externality for labor, whose returns are then used to demand productions of others. When someone works more or works harder, "this extension in demand, which is matched by the initiating increase in supply that made it possible, will create the potential for the exploitation of further specialization *in some production*, that remained just below the margin of economic viability in the pre-change equilibrium [so] the real cost of production in the affected industry falls" even if the industry had not been able "to identify in advance... increasing returns." Buchanan concludes, then, "if the extent of the market limits the potential for deriving the advantages of specialization, at least one industry in the totality that describes the production-exchange nexus must exhibit increasing returns..." For Buchanan, the development of strong work ethics has been critical for internalizing this externality.

However, as mentioned above, when the function is considered globally, decreasing returns are necessary. Eventually, not all inputs can be increased. Those which are not scarce cannot provide even constant returns. Eventually, we use up all the farmable land and more seed and fertilizer is not enough to produce proportionately more crops. Eventually, we use up all the possible glass, and steel—inputs for greenhouses—and fertilizer, dirt, and seeds themselves and no inputs are even left. Of course, this is what remains true in theory; whether we reach these points is beside the point. Far before the absolute limits of real resources are reached, competition for the limited resources readily available increases and their cost increases.

For many businesses and goods, what matters is not whether planet earth or the solar system or universe have certain inputs available for use. Instead, their focus is, given a fixed investment of a factory, facility, plot of farmland, or other investments, how to vary their per unit inputs. From this perspective, diminishing marginal returns to scale are always inevitable. When making marginal decisions, the current known methods of production are what are being considered—though, of course, the firm can engage in a search problem, estimating the expected gain and search costs for finding a new and more efficient method of production. If the firm is a massive player in the market, it may be true that they also must take note of how their own demand will shift the prices of their inputs. If its inputs have decreasing returns—if they consist of raw materials that must be grown, mined, or harvested, for instance—an increase in demand will ultimately increase the price.

A key question, then, becomes the elasticity of scale (ϵ), which measures the increase in output at a given point for an increase in all inputs—or all

inputs being considered. For t as our scaling variable—that is, if $t > 1$ we are increasing operations, if $t = 1$ we are maintaining the same operations, and $t < 1$ we are scaling down operations—and $f(\bullet)$ as our production function:

$$y(t) = f(tx)$$

$$\epsilon = \frac{\frac{dy(t)}{y(t)}}{\frac{dt}{t}} = \frac{df(tx)}{dt} \cdot \frac{t}{f(tx)}\bigg|_{t=1}$$

So, at the current point $t = 1$, we can use ϵ to measure whether the current method of production is in the constant ($\epsilon = 1$), increasing ($\epsilon > 1$), or decreasing ($\epsilon < 1$) range.

4.4.2 Cost Curves

If the inputs are all available in increasing quantities at a fixed price, the total cost curve is simply the inverse of the production function multiplied by the price per unit of input. For production function $f(\bullet)$, input vector $\mathbf{x}$, the prices associated with the inputs $\mathbf{P}$, and quantity of output Q:

$$Q = f(\mathbf{x}) \qquad \text{Production Function}$$

$$x = g(Q) \qquad \text{Inverse Production Function}$$

$$TC = \mathbf{P} \cdot \mathbf{x} \qquad \text{Total Cost Function}$$

If the firm is optimizing for its given input to get the most output, it is on the two respective curves. Otherwise, it will be somewhere in the shaded area above the curves.

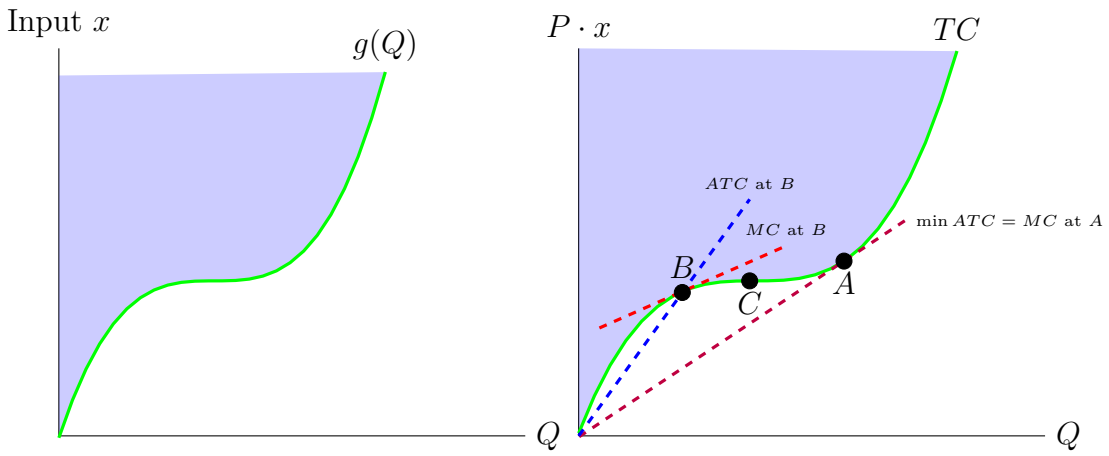


Figure 4.9: Relationship between the Production Function (left) and Total Cost Curve (right)

From the total cost curve, we can derive the per unit costs. The Marginal Cost is the cost of the next unit and is the derivative of the total cost curve. The Average Cost is the total cost divided by the number of units and is equal to the slope of the line running from the origin through each point on the total cost curve. The lowest point on the Average Cost curve is when the average cost line from the origin to the point is just tangent to the total cost curve. At this point, the average cost is equal to marginal cost. When marginal cost is less than average cost, it continues to bring average cost down, but once it is larger than average cost, it begins to bring the average cost up.

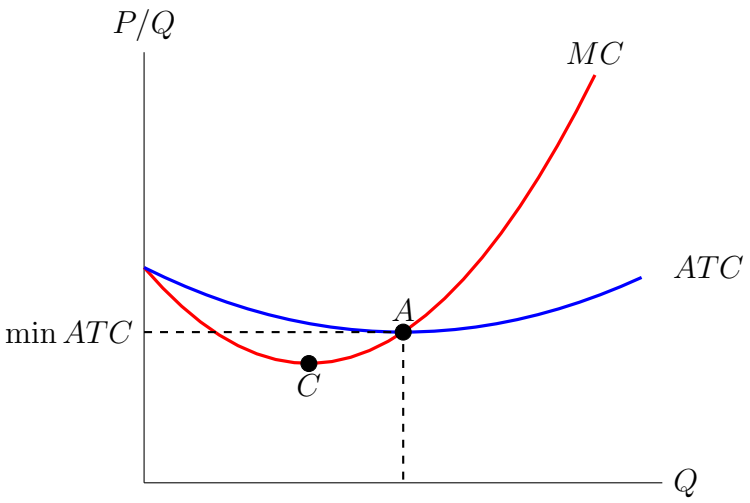


Figure 4.10: Average and Marginal Cost

We have already seen from comparative advantage and scarcity why marginal cost is necessarily rising, but it is worth considering why mathematically we must have a rising marginal cost *after some threshold* for a solution to be attainable. If marginal cost was constantly decreasing, eventually it would reach 0 allowing us to produce an infinite quantity of goods. If marginal cost was even merely constant, a firm facing a horizontal (perfectly elastic) demand curve would either produce nothing (if $MC > P$) or an infinite quantity (if $MC \leq P$). For a firm facing a downward-sloping demand curve, the elasticity of demand limits it from leading to an infinitely large quantity.

In our discussion of diminishing returns, we looked at inputs x and z of which x could be varied and z could not. We can also think of these as representing variable and fixed costs, respectively. Fixed costs z are the initial set-up cost that do not *necessarily* vary proportionately to a unit of output. For instance, to produce more cars, you need more steel, glass, rubber, plastic, labor hours, and electricity that goes into the making of each car; however, some costs like building a factory, buying machines and assembly lines, and hiring management, human resources, and finance staff do not vary with output. They are generally thought to be more difficult to vary as well. This seems obviously to be the case in terms of large investments like a factory or any capital good—a machine input to the production process; but, if including management and any labor not directly producing a specific good—or who is on staff regardless of how much is being produced—it seems likely that there is more wiggle room for variance.

Variable costs are all of the costs, generally of direct inputs, that vary

directly with output. If output increases by any amount, these costs *necessarily* increase. If we increase one more car, management and the factory investment can stay the same, but this is not true of the steel, glass, and rubber inputs.

4.4.3 The Firm's Choice Problem

4.4.3.1 Cost Minimization

To even derive cost functions, the firm must find the cost function that minimizes the cost of producing a certain quantity of output given certain constraints. For many goods, there is not a unique formula. Some inputs are substitutable. A structure can be built with brick, concrete, cinderblock, or wood. A cake could have cane sugar, brown sugar, syrup, honey, stevia, or an artificial sweetener. In the most general form, we speak of capital and labor. For instance, a sandwich shop could invest in capital to build a robot to take orders, build sandwiches, clean the shop, and maintain supplies. Alternatively, they could hire labor to do all of these things. We could invest in building a refrigerator or in hiring a person to fan the ingredients to keep them cool. Perhaps some of this sounds absurd, but it is precisely because we naturally engage in this trade-off analysis internally that we see the cost of buying a refrigerator is almost certainly less than hiring people to fan ingredients and keep them equally cool.

Firms recognize that there is not a unique solution of capital and labor for a given quantity of output; rather, we have level curves. We can produce 100 sandwiches, for instance, with a very high investment in capital and little in labor—with the entire process nearly automated—or with very little capital investment and much labor investment—with the entire process being manual—or a mix. For such a product, it seems clear that a mix will be superior to either extreme.

We have already seen that, at some level, there must be diminishing returns. When considering the trade-off between inputs, we *always* see diminishing returns when we increase one input but hold all other inputs constant. Thus, if we have a given quantity of output that we are trying to produce, as we substitute capital for labor, the ratio of substitution must increase—as we continue to use less labor, we will need to use proportionately *more* capital to make up for the lost labor and keep output constant.

We can graph this using the **Isoquant**. An Isoquant maps all of the possible combinations of our input to give us the same output. For instance, see the curves below. For instance, at each point along the blue curve, output is equal

to 200 units. The slope of the isoquant is the Marginal Rate of Technical Substitution (or the Marginal Rate of Substitution in Production). In the figure below,

$$MRTS_{LK} = \frac{\Delta K}{\Delta L} = \frac{MP_L}{MP_K}$$

it represents how much capital would be sufficient to maintain constant production given a one unit decrease in labor.

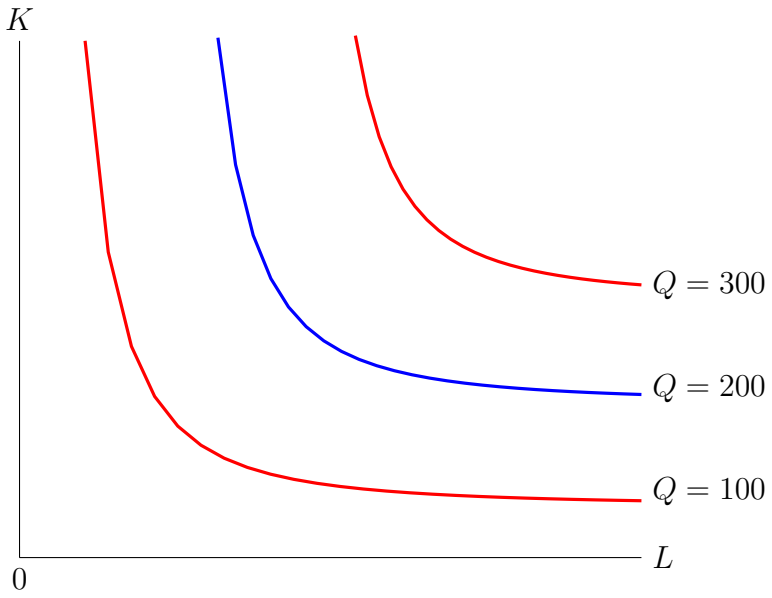


Figure 4.11: Isoquants

So how will the firm choose which combination of capital and labor—or of any other more specific inputs—to use? For this, we must add a budget constraint. The **Isocost** Curve shows a combination of the inputs for which the total outlay is the same. When there is a constant price for units of labor and capital, this is a straight line. For w as the wage rate (price of labor) and r as the rental rate of capital, the Isocost curve takes the form with each line representing the same total expenditure C:

$$C = wL + rK$$

Thus, the slope of the Isocost curve can be derived by setting the total differential of cost equal to 0—since, along the Isocost line, the total cost remains constant:

$$dC = wdL + rdK \stackrel{set}{=} 0$$

$$wdL = -rdK$$

$$-\frac{w}{r} = \frac{dK}{dL}$$

The slope of the isocost line ($\frac{dK}{dL}$) is, thus, the marginal cost of labor w divided by the marginal cost of capital and could be rewritten as:

$$\frac{MC_L}{MC_K} = \frac{dK}{dL}$$

The lowest cost method of production depends on the price of the inputs. See how two different price ratios lead to two different points of tangency on the isoquant. A cost-minimizing firm uses the basket of inputs at the point of tangency between the Isoquant and Isocost curves at which point the slopes of the two will be equal:

$$MRTS_{LK} = \frac{MP_L}{MP_K} = \frac{MC_L}{MC_K}$$

At this point, the ratio of the marginal products of the inputs will be equal to the exchange (or price ratio) of the marginal costs of the inputs.

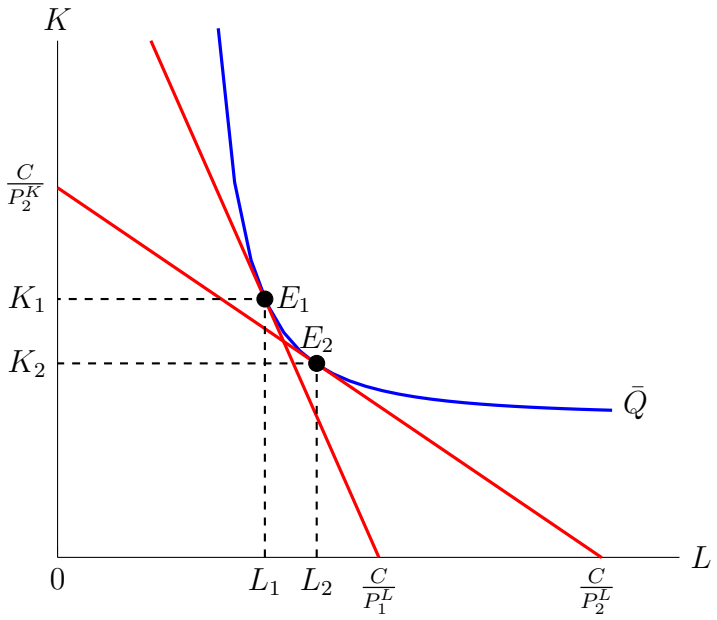


Figure 4.12: Isocost Curves and Isoquants

So, when building a total cost function, each combination of total cost and total output represents a point of tangency between a unique Isocost line and an Isoquant.

4.4.3.2 Profit Maximization

Economic profit is the creation of economic value *above* the opportunity cost. Economic profit is zero when these resources would be used for equally good satisfaction of consumer willingness-to-pay if employed in an alternative (the next-best) use. Clearly, this is different from a firm's book (or accounting) profit. Accounting profit only compares the money spent and the revenue earned. Accounting costs include only the explicit monetary costs that are paid—costs the firm seeks to minimize. It does not also include the implicit costs of the potential revenue brought by other uses of that money. Economists, on the other hand, look at opportunity cost and relative productive abilities to measure profit. Economic profits represent the creation of value. If all of the same inputs were put towards another use, the result would be something that satisfy fewer or weaker desires valued less.

Graphically, economic profit is the gap between the average cost and the average revenue. While the area is between these two averages, its maximum is *determined*—not by the averages—but at the margin by the relations between marginal cost and marginal revenue.

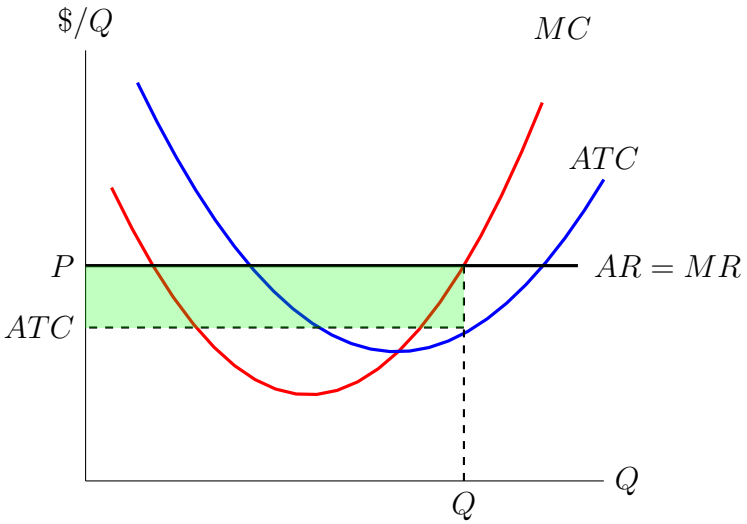


Figure 4.13: A Firm with Economic Profit

Machlup (1946) reminds us that we cannot "test" for profit maximization, given the subjectivity of the quantities. He writes that "cost, revenue, [and] profit," for instance are all "subjective–that is, perceived or fancied by the men whose decisions or action are to be explained." None of them are objective in the sense that they could be "calculated by disinterested men who are observing these actions from the outside and are explaining

them." Thus, "the producer's actual decision is based on what he himself thinks"—not a hypothetical outside observer—and "based on 'subjective' cost expectations." Likewise, when we "derive statistical cost curves from accounting data—which of necessity refer to conditions of the past," these *may* be useful for "evaluating prospective and hypothetical cost and revenue figures," but they are not equivalent to the firm's *anticipations* which "alone are the relevant variables in the marginal calculus of the firm."

The firm is assumed to profit maximizing. Thus, the firm's primary choice problem is how to adjust the price and quantity of its output to maximize profit. Since

$$\pi = TR - TC$$
$$\pi = pQ - (TFC + VCq)$$

to maximize profits, we need to find the largest gap between the total revenue and total cost curves, which will be the quantity of output q that solves the first-order condition

$$0 = \frac{dTR}{dq} - \frac{dTC}{dq} = \frac{dPQ}{dq} - \frac{dVCq}{dq}$$
$$\frac{dPQ}{dq} = \frac{dVCq}{dq}$$
$$MR = MC$$

Thus, our optimal condition is that marginal revenue is equal to marginal cost. But we also want to ensure that this quantity is such that our revenue is greater than our costs, so, for our second-order conditions, we can check to ensure that the Marginal Cost curve is approaching the Marginal Revenue curve from below rather than from above:

$$\frac{d^2TR}{dq^2} < \frac{d^2TC}{dq^2}$$

When $MR = MC$ after coming from below, we are at the optimum. To solve for profit maximization, no other conditions—such as average cost—matter. At the unit before, it is possible to receive more in revenue than you spend by producing another unit. At the next unit, it costs us more to produce the next unit than we receive in exchange for it, so we would *lose* money on it. Thus, at the optimum, producing more would destroy value—it would be to consume resources worth more than consumers are willing to pay for its output. To produce less would be to miss out on turning a bundle of inputs into something valued by consumers more than the inputs—value-creation.

Profit maximization does not mean that, in practice, business owners only care about monetary considerations. Machlup (1946) writes

> there is nothing essentially "uneconomic " in the conduct of a business man who chooses to pay higher prices for raw material purchased from a fraternity brother or to sell at a special discount to members of his church, or who refrains from embarking on a promising expansion of his business because he prefers an easier life.

The trade-off may lie between "social prestige or a certain measure of inner satisfaction" as opposed to higher profits. But this is why Friedman (1970) so vehemently argued that non-owner managers should focus on increasing profits and allowing the owners to make the trade-off between profits and other non-pecuniary benefits rather than forcing the owners to accept lower profits, resulting in lower value creation across the economy, in favor of these non-pecuniary benefits.

The evolutionary tendency of economic systems is to eliminate higher cost firms who must charge higher prices in favor of lower cost firms who then are closer to maximization. But, if society truly values something—that is, they have a willingness to pay for it—maximization does not mean ruthless and reckless actions. If consumers are concerned about certain pesticides or chemicals in food products, they will pay more for products without them; thus, consumer willingness-to-pay acts as a regulator. If consumers care about supporting local businesses, they will buy from them and be willing to pay a higher price than they would elsewhere; thus, consumer willingness-to-pay acts philanthropically. If consumers care about frequenting businesses that employ or are owned by people of a certain gender, race, religion, or political affiliation, they will buy from them and be willing to pay higher prices for their goods and services than from the goods and services from the less or disfavored gender, race, religion, or political affiliation. However, while it is common for people to speak about their desires for these various things, actions speak louder than words. If people are not willing to actually pay more to support these particular businesses, their demonstrated preference shows that their utility functions are not attuned to these factors and that they do not actually prefer these things.

4.4.3.3 Operation

While we do not look to average costs to determine the profit maximization optimum in the short-run, the average cost does play an important role in the firm's choice problem in the longer-run. The difference between average costs and revenue lead to entry and exit in an industry.

For instance, when $P > ATC$ and a firm is earning economic profit, new

firms will be incentivized to enter the market—so long as they have identical or very similar cost and revenue curves—because they can also earn this economic profit. We would expect firms to enter until $P = ATC$ and a firm would be indifferent between entering this industry or producing in any other industry, because the next-best alternative (opportunity cost) of the inputs is equally profitable.

On the other hand, when $P < ATC$, a firm is taking a loss. Now, in the short-run, a firm may continue to operate. In the graph below, we can see that the firm will minimize its losses (which is maximizing its profit) by producing when it has acquired a fixed cost so long as the price is greater than the average variable cost. If $P > AVC$, each unit produces more value than the per unit variable costs and this can cover *some of*—but not all of—the fixed costs. If $P < AVC$, then the firm would lose money for producing any units, so it will shut down and simply have to eat the fixed cost.

Note that, as the quantity increases, the Average Fixed Cost (for a given fixed cost) decreases, because this given cost is unchanging and being divided by more units. Thus:

$$\lim_{Q\to\infty} AFC = \frac{FC}{Q} = 0$$
$$\lim_{Q\to\infty} ATC - AVC = 0$$

So, the AVC approaches the ATC over higher quantities.

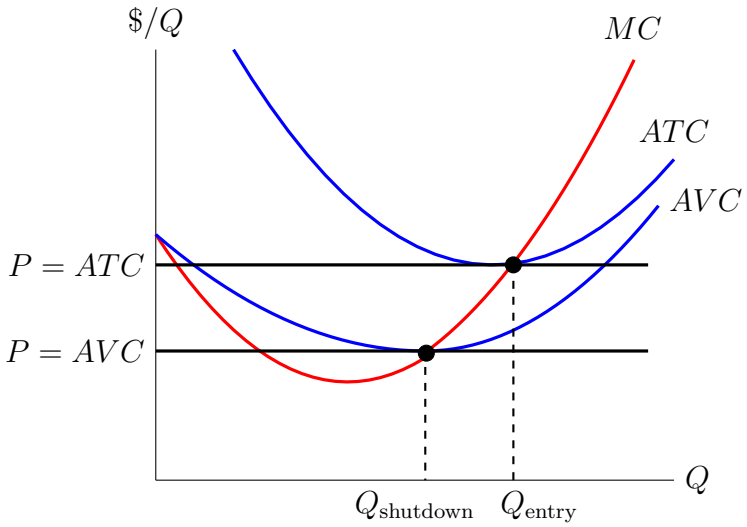


Figure 4.14: Cost Functions of the Competitive Firm

The dynamics of the market, then, lead to the dissipation of economic profit to the point of zero profit—costs are equal to benefits. As conditions change

or new inventions are made, economic profits are available, but as more entrepreneurs recognize the opportunity, more entry will lead to an increase in the overall quantity supplied, putting downward pressure on the price until $P = ATC$. Thus, the condition of zero economic profit simply means that an industry has no entry nor exit, as there is no incentive to leave or enter.

4.4.3.4 Investment

While we often speak of a single average total cost curve with a specific fixed cost and variable costs, there are generally many production opportunities available to a firm. A firm has a wide variety of options for the fixed cost. How many factories to build and how large to build them? How many employees should be kept on retainer, ready to produce when the season or demand warrants it?

Thus, there is a long-run average total cost curve (sometimes called an "envelope curve"). It is made up of a bunch of short-run average total cost curves, each with their own specified fixed cost. The $LRATC$ includes the lowest possible average cost for each level of output—the points of tangency between $LRATC$ and individual $SRATC$ curves. Thus, along the $LRATC$ curve, the fixed costs vary along the curve depending on which $SRATC$ curve the point belongs to.

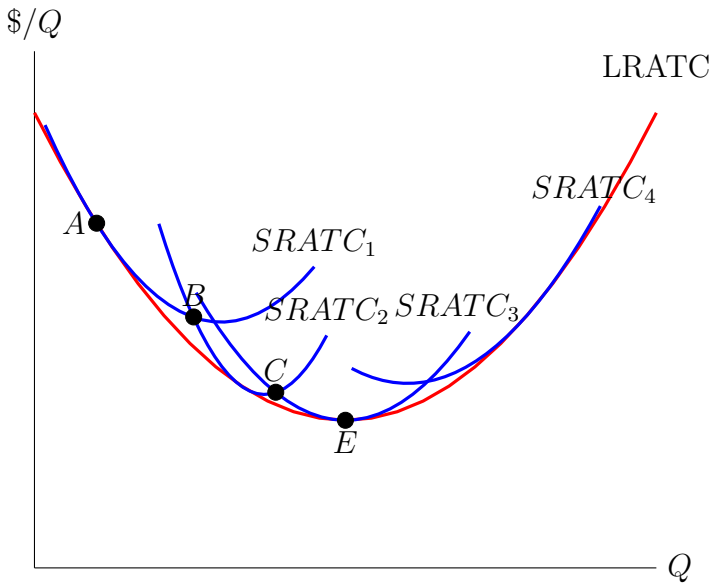


Figure 4.15: The Long-Run Average Total Cost Curve

The structure of the Long-Run Average Total Cost Curve can shape the

typical firm's size in an industry and its optimal investment. If the firm starts at point A, for instance, with a low fixed cost, they are still on the increasing returns section of the short-run average total cost curve and will be incentivized to move to point B, the minimum of that $SRATC_1$. But this point has an equal cost between the fixed costs of $SRATC_1$ and $SRATC_2$ and, with the fixed cost of $SRATC_2$ a lower minimum average cost of C is possible. Yet, none of these points are even on the $LRATC$ curve, meaning that these are not the lowest cost methods of production for that given quantity of output. Again, the same scenario plays out until we end up at point E, the minimum of the long-run average total cost curve. The lowest cost firms, then, will have the fixed cost associated with $SRATC_3$ giving them a better likelihood of survival than smaller (or larger) firms.

4.5 The Supply Curve

4.5.1 Firm's Supply

Since we have seen that the firm's profit maximization rule is to produce up to the point where $MC = P$, *but* only when $P > AVC$, the firm's short-run supply curve will be equivalent to its marginal cost curve above this point, given the fixed costs that it has incurred.

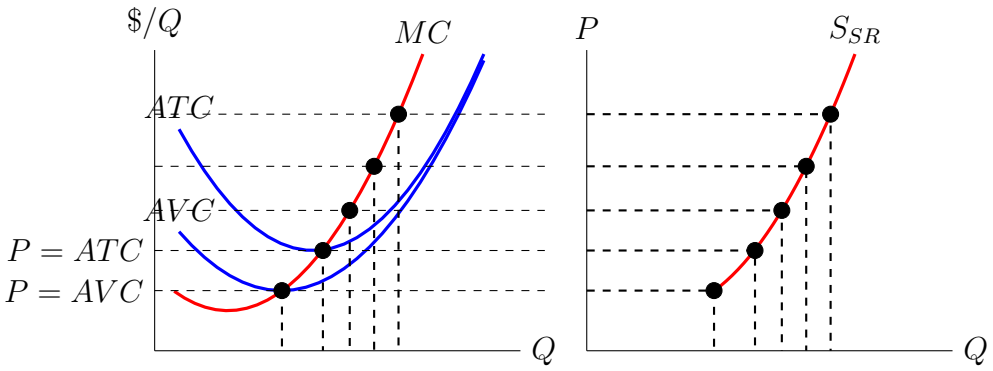


Figure 4.16: Firm's Short-Run Cost Curves (Left) and Short-Run Supply Curve (Right)

On the other hand, the firm's long-run supply curve will be the marginal cost curve that goes through the lowest point of its Long-Run Average Total Cost curve. And, in the long-run, the firm will only produce where $P \geq ATC$, so the firm's supply curve is equivalent to the marginal cost curve at and above this point.

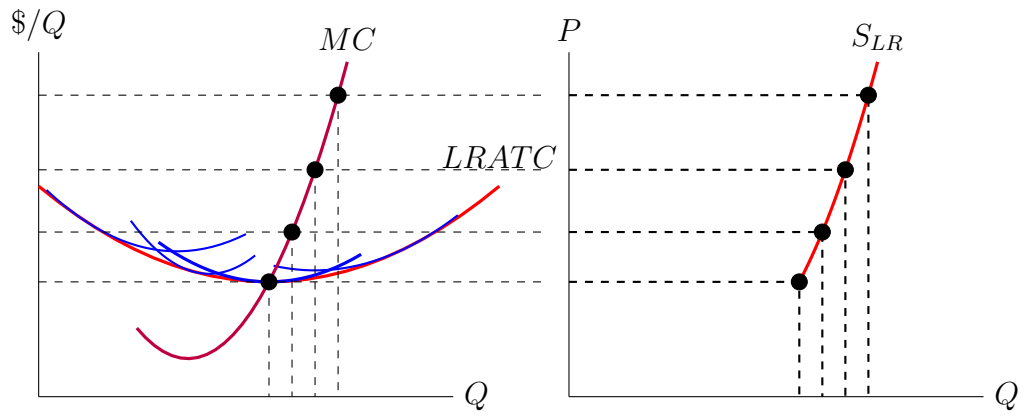


Figure 4.17: Firm's Long Run Cost Curves (Left) and Long-Run Supply Curve (Right)

The firm's supply curve shifts in light of changes in the underlying conditions. If technology improves and more efficient methods of production are discovered, the curve will shift out to the right—firms will be willing to produce more goods at the same price. If the prices of inputs rise, the curve will shift left—firms will be willing to produce fewer goods at the same price. Even if there are improvements in *other* sectors, the shift in opportunity cost will require firms to adjust—whether they move industries to pursue more lucrative opportunities elsewhere or raise their bids on inputs and eventually their output prices to not lose out to these other sectors. Baumol (1967) wrote that as the economy experienced productivity improvements in the production of goods, the service sector did not see this as much, so the result is "cumulatively increasing costs [in the service sector and other sectors with no productivity growth] relative to those of other economic activities."

4.5.2 Market Supply

Just as the Market Demand curve is the horizontal sum of the quantities demanded by each individual at a given price, the Market Supply curve, likewise, is the horizontal sum of the quantities supplied by each individual producer (or firm) at a given price. Thus, we take all of the individual firm supply curves and sum them.

If there is free entry and exit and all firms have roughly the same cost curves and similar entry prices—points distributed around $ATC = P^*$, then the supply curve will be relatively elastic, almost horizontal—relatively flat. For the number of firms able to enter N, the supply curve will look like this.

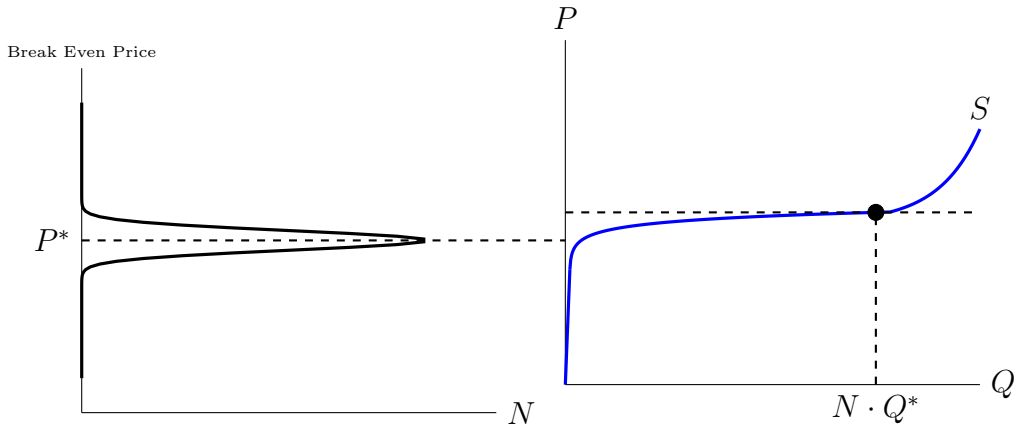


Figure 4.18: Relatively Homogeneous Firms Entry Prices (left) and Market Supply Curve (right)

If not all firms are the same and face different costs (heterogeneous firms), then the market supply curve could look like this where the break-even point for firm 1 is P_1 and the break-even for firms 2 and 3 is P_2.

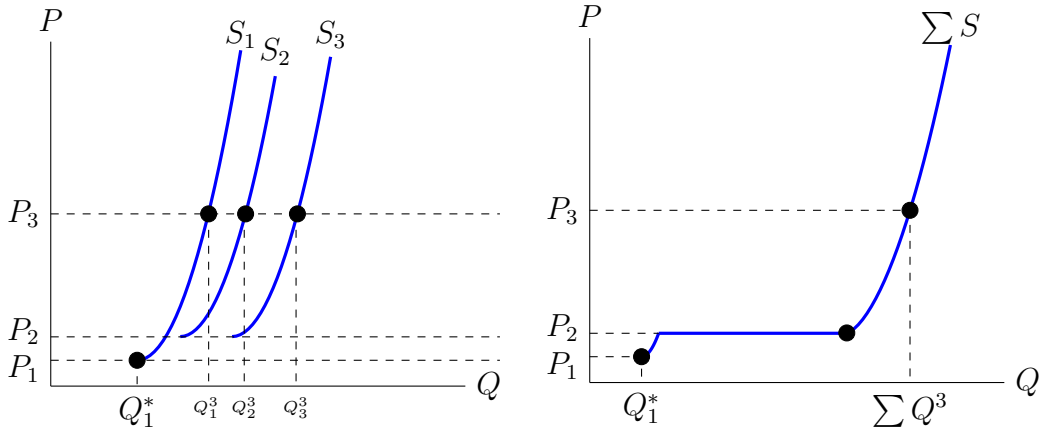


Figure 4.19: Heterogeneous Individual Firm Supply Curves (left) and Market Supply Curve (right)

Just like the firm has short- and long-run supply curves, the market supply curve also has a short- and long-run. In comparison with the short-run supply curve, the long-run supply curve is far more elastic (flatter). This means that the firm is *more* responsive to price changes in the longer-run than in the shorter run, in large part due to the entry (or exit) of firms. The short-run curve also allows for some production at lower prices than the long-run

curve—accounting for the short-run fixed costs of the existing firms in the industry.

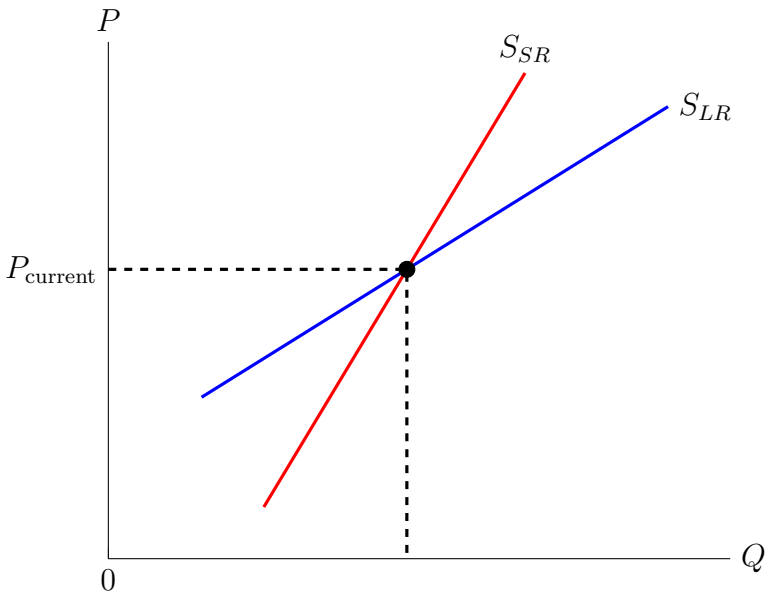


Figure 4.20: Short-Run and Long-Run Market Supply Curves

4.5.3 Producer Surplus

Producer Surplus area above the supply curve below the price received. It is the producers' gains from trade—the difference between what they would have been willing to receive to produce the units of the good compared to what consumers are willing to pay them. At a price of P_1, producers are willing to produce Q_1 and receive the green area in surplus. If the price increases to P_2, then the producers of the first Q_1 units receive the blue area *in addition to* the green area as surplus. The orange area represents the surplus of the new units supplied at the new price.

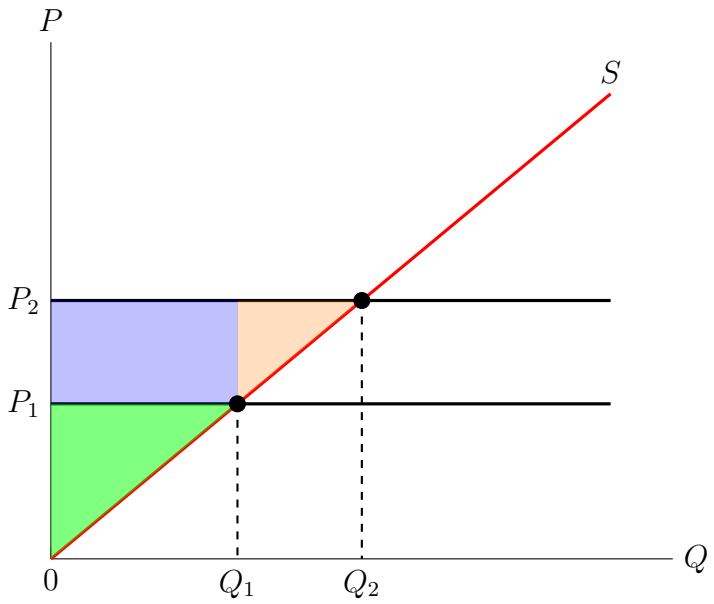


Figure 4.21: Producer Surplus

References

Alchian, Armen A and Harold Demsetz. 1972. "Production, Information Costs, and Economic Organization," *American Economic Review*, Vol. 62, No. 5, pp. 777-795.

Baumol, William J. 1967. "Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis," *American Economic Review*, Vol. 57, No. 3, pp. 415-426.

Buchanan, James M. 1994 [1992]. "The Supply of Labour and the Extent of the Market," *The Return to Increasing Returns*, edited by James M. Buchanan and Yong J. Yoon, University of Michigan Press.

Coase, R. H. 1937. "The Nature of the Firm," *Economica*, Vol. 4, No. 16, pp. 386-405.

Coase, R.H. 1988. "The Nature of the Firm: Meaning," *Journal of Law, Economics, & Organization*, Vol. 4, No. 1, pp. 19-32.

Friedman, Milton. 1970. "The Social Responsibility of Business Is to Increase Its Profits," *The New York Times*, September 13, 1970, p. 17.

Machlup, Fritz. 1946. "Marginal Analysis and Empirical Research," *American Economic Review*, Vol. 36, No. 4, pp. 519-554.

Klein, Benjamin, Robert G. Crawford, and Armen A. Alchian.

1978. "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process," *Journal of Law & Economics*, Vol. 21, No. 2, pp. 297-326.

Schumpeter, Joseph A. 1934 [1911]. "Entrepreneurial Profit," *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle*, translated by Redvers Opie, pp. 128-156.

Williamson, Oliver E. 1981. "The Economics of Organization: The Transaction Cost Approach," *American Journal of Sociology*, Vol. 87, No. 3, pp. 548-577.

Chapter 5

Partial Equilibrium

5.1 Equilibrium

Equilibrium comes from the Latin words for a balanced scale. The state of equilibrium arises when people are optimizing given their constraints. Of course, the conditions and underlying objective functions can be shifting and adjusting, which means that the equilibrium positions will likely fluctuate along with them.

A **Nash Equilibrium** is a state when people's strategies are the best responses to each other. In a Nash Equilibrium, given how everyone else is acting, you cannot improve your situation by unilaterally changing only your decisions. The most common type of equilibrium in microeconomic models is **Walrasian Equilibrium**. This is a subset of Nash Equilibria where all people's plans are consistent with each other leading to no shortages or surpluses. If you look around at everyone else's actions and want to *change* your decision, this is not a Walrasian equilibrium.

If there are various constraints that hinder market adjustment, such as the cost of repricing goods (menu costs), search frictions, or even government mandates, we can have a **Constrained General Equilibrium** where there may be some surpluses and shortages that is optimal, because the cost of adjusting.

When we are conducting a Partial Equilibrium analysis, we are looking only at one market and how it adjusts holding everything else—the prices of all other goods are held constant.

5.2 Comparative Statics

Comparative Statics looks at the initial and final equilibrium conditions for a specific given change. Comparative statics can tell us how the quantity supplied, quantity demanded, equilibrium price, or surplus will change given a shift in either supply or demand curves or changes in the environment that effectively shift the curves.

For instance, if consumers have more income for some reason, their demand curve for a given good will shift out to the right—more people will demand the good at a given price. The result will be an increase in the quantity demanded, which will be partially met by an increase in supply and partially counteracted by an increase in price, which reduces demand.

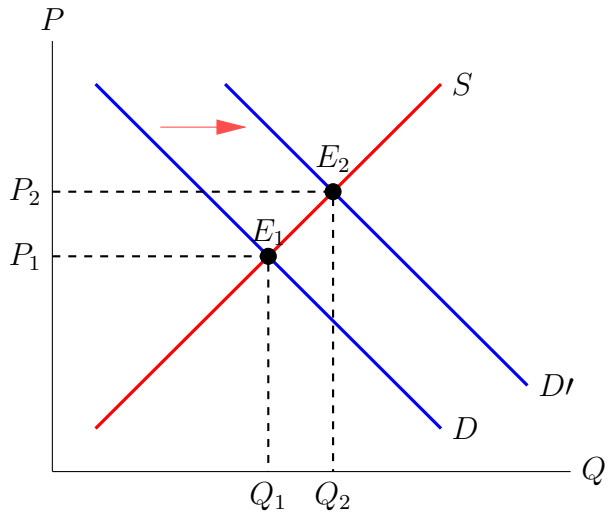


Figure 5.1: Equilibrium After an Increase in Demand

When conducting such an analysis, we make the *ceteris paribus* assumption—that is, literally "other things equal." We can only speak about how a given event or shift will affect output and prices, if we assume other factors that are shifting output and prices are unchanging. Of course, during the adjustment period, as part of the adjustment, many variables may change, and this is *part of* the analysis. However, if multiple shocks hit at once, we may not be able to decipher the ultimate movement of variables. Consider, if, in addition to the increase in demand, there is a fall in supply due to an unrelated supply shock. Now, the quantity may be greater than it initially was, it may remain the same, or it may fall, on net. To get a more precise estimate, we would need to estimate the elasticity of supply and of demand and the magnitude

of both shocks. This does not mean that our preceding analysis was wrong—only that it assumed that the production function and cost functions were unchanged. It is important to be clear about our assumptions.

While perfect information has become an assumption of certain types of equilibrium, Morgenstern (1935) writes that this is not the historic view. "Neither Walras nor Pareto unequivocally stated that full foresight is one of their assumptions." He goes so far as to suggest "that they introduced the precedent of the gradual formation of equilibrium *without* perfect foresight," as both wrote about the idea of successive adjustments and such adjustments are "irreconcilable with perfect foresight." At a hypothetical equilibrium, there would be no unexploited profit opportunities—no value-creating exchanges that have not already taken place. It is understandable why many economists discuss the need for all agents to have complete knowledge of every other agent's opportunities to ensure that this is the case. However, Morgenstern writes that this is not merely mistaken but incoherent. He gives the example of Sherlock Holmes pursued by Moriarity on a train. If both agents have perfect information, however, they will be unable to act.

> Sherlock Holmes, pursued by his opponent, Moriarity, leaves London for Dover. The train stops at a station on the way, and he alights there rather than traveling on to Dover. He has seen Moriarity at the railway station, recognizes that he is very clever and expects that Moriarity will take a faster special train in order to catch him in Dover. Holmes' anticipation turns out to be correct. But what if Moriarity had been still more clever, had estimated Holmes' mental abilities better and had foreseen his actions accordingly? Then, obviously, he would have travelled to the intermediate station. Holmes again, would have had to calculate that, and he himself would have decided to go on to Dover. Whereupon, Moriarity would again have "reacted" differently. Because of so much thinking they might not have been able to act at all or the intellectually weaker of the two would have surrendered to the other in the Victoria Station, since the whole flight would have become unnecessary. Examples of this kind can be drawn from everywhere. However, chess, strategy, etc. presuppose expert knowledge, which encumbers the example unnecessarily.... [A]lways there is exhibited an endless chain of reciprocally conjectural reactions and counter-reactions. This chain can never be broken by an act of knowledge but always only through an arbitrary act—a resolution. ... Unlimited foresight and economic equilibrium are thus irreconcilable with

one another.

5.2.1 Taxes and Equilibrium

We can also use comparative statics to compare how government policies will affect equilibrium. For instance, we can compare the effect of a per unit flat tax and a proportional tax on a particular good. The graphs below show taxes raised out of former consumer surplus (teal) and producer surplus (purple). They also show a welfare loss in the shaded orange area. This is Harberger's Triangle or Deadweight Loss. It represents *potential* consumer and producer surplus that is lost entirely. While the other shaded areas represent transfers—former surplus that becomes government revenue—this triangle is representative of mutually beneficial exchanges that would take place without a tax but now do not take place at all. The potential units and the potential economic value-add is lost entirely.

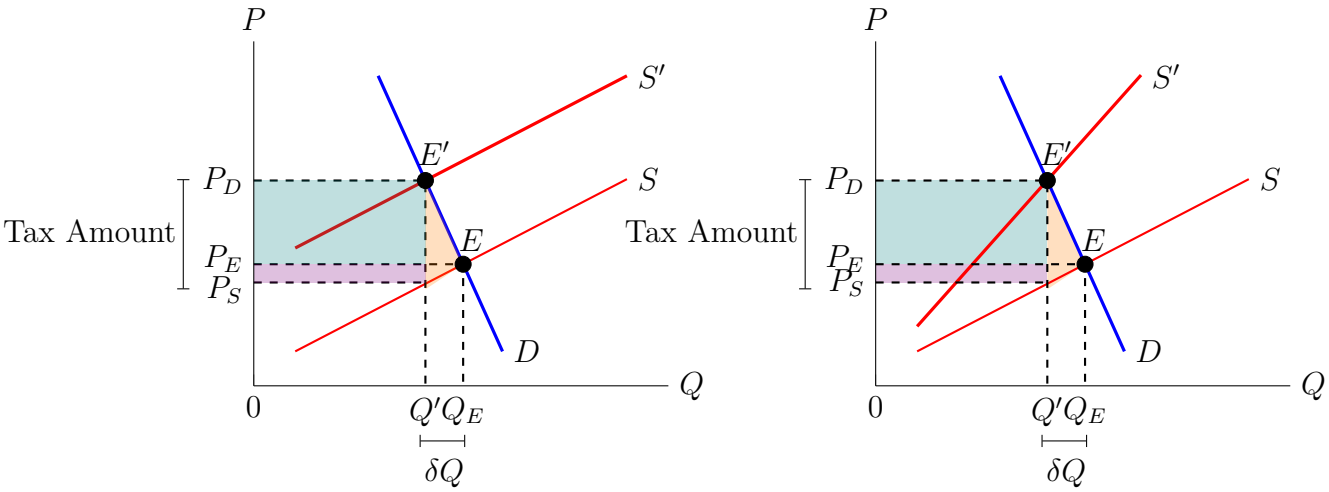


Figure 5.2: The effect of per unit (left) and proportional (right) taxes

We can see that, in this case, demand is highly inelastic—it is relatively unresponsive to changes in price—so consumers bear most of the effective burden; that is, lost consumer surplus exceeds lost producer surplus. In this graph, the left-hand side is a tax of $3 per unit, while the right-hand side is a percentage of the price of each unit that raises the same revenue for the given supply and demand curves—a 73% tax.

In drawing this graph, it would be equivalent to put the tax on either the supply or the demand curve. The supply curve would shift up by the amount of the tax, since this effectively adds a marginal cost to the producers by

the amount of the tax. S continues to represent the marginal cost to the producer and P_S is the cost received by the supplier after paying the tax. On the other hand, if we choose to show the tax on the demand curve, the consumer willingness-to-pay remains unchanged. Instead, the amount of the tax causes the curve to shift left, again, because $D\prime$ is their willingness to pay *producers* and D remains their willingness-to-pay overall.

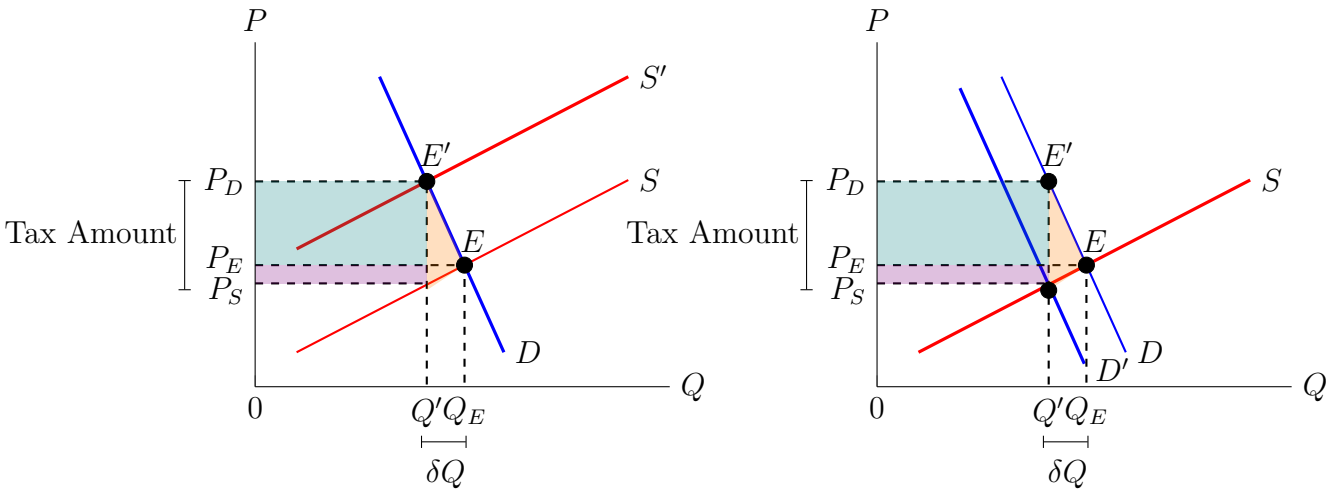


Figure 5.3: Legal Incidence Supply vs Demand

In this graph, the legal incidence of the tax does not matter. Whether the consumers or the producers are legally liable to pay the tax, the distribution of the burden remains the same with more lost consumer surplus than producer surplus, due to the inelastic demand curve.

5.3 Supply and Demand Dynamics

5.3.1 Shifts Along the Curves and Shifts of the Curves

When considering changes in equilibrium, it is important to differentiate between two causes: shifts *along* a curve and shifts *of* a curve.

When the equilibrium conditions change, at least one curve must shift. The other curve experiences a movement *along* the curve. A supply curve suggests, over a given time period, for each specific market price of a good, how many units will be supplied by producers (or potential producers) of the good. A (Marshallian) demand curve represents the quantities of a given good that would be demanded by all of the individuals in a market for various different prices holding constant all the other prices in the economy.

Recall that following Lancaster (1966) and Stigler and Becker (1977) the assumption of constant prices of all goods in the economy can be read expansively as thinking about the services that the goods provide. The vast majority of the time, the demand for the underlying services are unchanged (pleasure, enjoyment, status, safety, health, nutrition, shelter, etc), but what shifts are the relationships of specific goods to these services and alternative means of satisfying these desires. So, what we often call shifts in "tastes" are simply changes in underlying constraints. When a clothing trend is said to be "out of style," it no longer provides the same service of status, being considered fashionable, or fitting in. Thus, the demand for fashion or status can remain unchanged, but the service is no longer provided by these specific clothes, so the demand curve for these particular articles of clothing *shifts* to the left. Since the clothes no longer provide the valuable service of fashionability, fewer people are willing to buy them at the same prices. When it is often said that changing tastes shift demand curves, what is meant is that the prices for the underlying *services* being demanded.

But, lest this give us an air of objectivity, we must be reminded that it is individual's *subjective* valuations that are at play—what do you perceive to be high status, fashionable, etc. We can't forget, too, that a good number of aesthetic preferences that are clearly subjective—one's preferred iPhone color or the taste of food that people prefer—and are well-represented by the idea of "tastes."

So some key facts that can shift demand curves include a change in the prices or quantities of other related goods, income, and tastes. If the price of a complement rises, for instance, the quantity demanded of its related good will likely fall for a given price, since the overall price of the bundle has risen. If the price of a substitute rises, however, the quantity demanded of its related good will likely rise for a given price as consumers substitute in favor of it, since it has become relatively cheaper. Likewise, changes in the number of buyers—the number of individual demand curves—can shift a demand curve. Finally, shifts in expectations can play a key role. If people are worried about a shortage or potential shortage of a good, they will seek to buy more now even at a higher price. If people expect a price to rise, they will buy more now *before* the price rises, shifting out demand and ultimately forcing the price up anyway.

On the supply side, shifts in real conditions—changes in natural resources and inputs generally—are a key factor. If bad weather damages or destroys a set of inputs, fewer units of a good will be available at the same price. Likewise, changes in technology or, perhaps more broadly, the knowledge of how to convert inputs into outputs can shift the supply curve. Industrialization,

adoption of the assembly line, and invention of new products or methods of production can lead to a lower cost per unit shifting the supply curves out to the right. Finally, changes in legal rules—of course, taxes are a key example—are important for both supply and demand. Remember, people are focused on how various goods and services are able to serve their needs and desires, so changes in institutional constraints will simply change the trade-offs of various particular ways of satisfying those needs and desires.

5.3.2 From Partial to General: The Limits of Partial Equilibrium

Could a supply shock to one good cause multiple goods to move off to infinity? By examining two markets in isolation in partial equilibrium, we can end up with some crazy solutions.

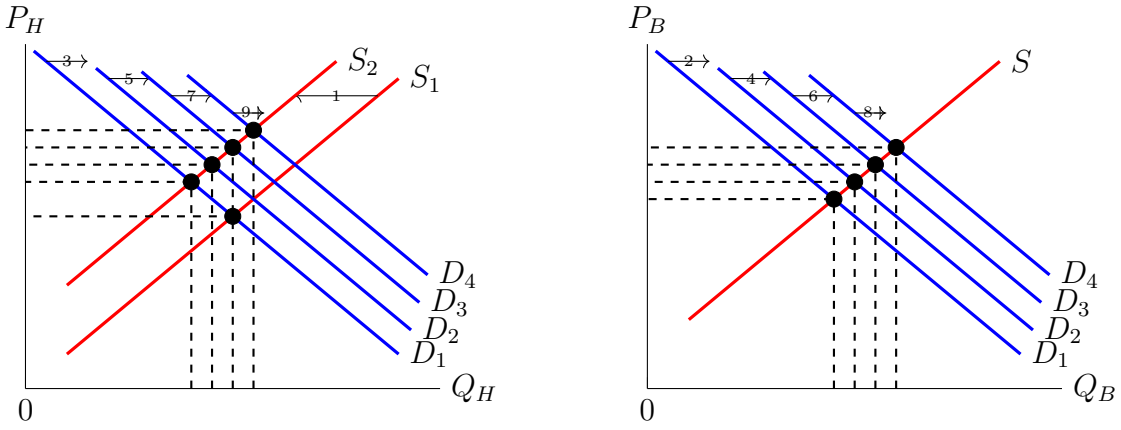


Figure 5.4: Problematic Partial Equilibrium Markets for Hot Dogs (H) and Burgers (B)

For example, consider the market for hot dogs and burgers. If some swine fever infects pork, leading to a reduction in the supply of hotdog meat available, the supply curve for hotdogs shifts to the left—less is supplied for a given price. However, one variable for the demand of burgers is the price of its substitutes. If the price of the substitute has risen, the demand curve for a good will shift to the right—more will be demanded for a given price—as individuals shift from hotdogs to burgers, which have been made relatively cheaper. This will result in a higher price for burgers and higher quantity supplied. Yet, when the price for a substitute rises, the demand curve for a good will shift to the right—more will be demanded for a given

price. This means the demand curve for hotdogs will shift to the right—as burgers have now become relatively more expensive, people will switch to hotdogs instead of burgers. However, after this change has taken place, the substitute of burgers has become more expensive, and suddenly the two demand curves both continue to shift right until the price and quantity of both goods is infinity.

Now, of course, when we go to do empirics and estimate the demand functions and cross-price elasticity, we would find that they would settle down to prevent such a situation from happening in the real world. (Note that this has also failed to allow for a return to S_1^H after the supply shock is ended. But this example can show how thinking in terms of just two markets in isolation can be problematic for analytical reasoning.

For this reason, it is critical to see how all the markets relate with one another and all others in real-time. Any time there is a feedback loop, such as a resource constraint or a spillover effect, we cannot just consider the two individual markets in isolation. We must solve for the optimal price and quantity jointly.

It is true that the demand function for burgers, for instance, includes a host of variables.

$$Q_B = f(P_B, P_H, I, \mu)$$

So when graphing the relationship between Q_B and P_B when the consumer's income I, the price of hot dogs P_H or any other relevant price or taste variable μ changes, we must shift the curve. However, this is not necessarily a sequential process.

General Equilibrium, which we will explore more in the next chapter, considers the interaction between markets. When consumers are making the decision between hot dogs and burgers, this decision is made *simultaneously*. When a consumer picks a hot dog over a burger, they are making a comparison in the moment. General Equilibrium reminds us to connect these markets in real time and how there is an instantaneous feedback between various markets. In fact, General Equilibrium recognizes that, ultimately, *every* market feedbacks into every other market. Thus, to truly solve for the optimal prices and quantities, we must take into account every other market.

5.4 The Market & the Market Process

Before moving on from partial equilibrium, a big question underlying much of this analysis is why we can even speak in terms of equilibrium. How we get to equilibrium? It is one thing to speak of curves shifting, but in the actual

world, when a tax is imposed, for instance, how do the individual actions of market actors move from point E to E'?

5.4.1 What is the Market?

First of all, it is helpful to define our terms. We use the language of "market," drawing upon the idea of a place to go to conduct trades, but when we speak of economic markets, we mean something much broader and inclusive of all purposive activity of individuals and the resulting institutions that facilitate this activity. When we think of the market as such an activity (and various institutional supports), it is clear that the market is not a static *thing* but rather is itself a process. Even the institutions that facilitate the market—whether it be formal institutions such as the civil courts that enforce contract law, business firms that direct activity, or clubs that facilitate cooperation towards a unified goal, or informal institutions such as customs and culture that create expectations surrounding how activity and trades take place—are not fixed and unchanging.

What is key about the market as opposed to alternative actions is the role of profit and loss in judging decisions. For instance, when a government builds a road, the decision of where to build the road and how is made by the government officials who have the power to seize the property necessary for the project, the expenditures are paid for by taxpayers, and then the benefits are intangibly enjoyed by the users of the road. We can maintain opinions about whether the benefits were "worth" the cost, but a freely available road makes no monetary profit that we can use as a metric to determine whether the value created was positive or negative. On the other hand, when individuals and organizations without the coercive power we usually attribute to government—e.g., forced taxation, eminent domain, and, in recent years for many large nations, the lack of a budget constraint due to the entangled relationship of the monetary printing press and the state—we have the profit and loss metric for evaluating actions.

The market is not simply the end states of equilibrium—in fact, we never reach a pure equilibrium where there are no more potentially beneficial opportunities. Changing conditions and new ideas constantly create new opportunities for profit. It is also important to not forget the other side of the ledger; after all, it is a profit and loss system. Losses happen because humans are fallible, foresight is imperfect, and the future can be difficult to predict. The market does not preclude losses, but it does ensure that they cannot perpetuate forever. Thus, the market is not always (and perhaps is never) *in* equilibrium, but the actions of individuals subject to the profit

and loss system generally work to move us *closer to* equilibrium in each given moment.

Buchanan and Vanberg (1991) caution against thinking about the market itself as goal-oriented. They write that "the market economy, as an aggregation, neither maximizes nor minimizes anything. It simply allows participants to pursue that which they value, subject to the preferences and endowments of others, and within the constraints of general rules of the game that allow and provide incentives for individuals to try out new ways of doing things." Of course, agents within markets have various goals, but this is the point: it is the agents—not the market itself—that have a *telos*.

For simplicity, we often model the market as the mechanism for solving an optimization problem: maximize value add given various resource and institutional (read as artificial, whether created by government or culture) constraints. The market is said to "distribute" resources to those who value them highest. There is certainly some truth in this, but this is a type of mythology. First, we have considered that it is an anthropomorphization—since "the market" does not "do" anything; it is the individual agents who take actions. But second, and perhaps more notable, there are no *given* goods to be distributed. It is within markets that knowledge, information, and preferences are developed and resources and their uses, methods of production, and products themselves are discovered and invented. Thus, Buchanan and Vanberg present the market as a game without goods. The actual goods are not pre-existing but created, invented, and discovered. Because of this, any focus on the market as merely a distributive mechanism misses the bigger picture and "any idealized omniscience on the part of a planner who might attempt to duplicate the market result would become patently absurd."

Does this mean that we can't use equilibrium concepts to speak about economic conditions and policies? Absolutely not. Many economic analyses that begin with equilibrium can be thought as simply a focus on such a policy's sustainability. Wagner (1983) writes that the common economic arguments against interventions like price controls really seek to articulate the unsustainability of such an approach. He writes that

> sustainability would require that people refuse to replace options they value less highly with options they value more highly. The rent control [for instance] creates an incentive for people to contract around it, because there are more people willing to rent apartments at [the rent-controlled price] than there are apartments available. ... Regardless of the particular features such contracting around takes, there will be a divergence between

> appearance and reality in that the adherence to the rent control will be maintained even though the reality is a higher price. ... Basic economic principle explains why such a policy measure can never have the consequence of *simply* making rental prices lower, and it also explains why the consequences of such legislation lead generally to poorer housing within the society.

When there is shortage, non-price rationing must take place, whether it is more discrimination of the chosen tenants or reduction in maintenance or other quality of the product. Government mandates like this invite people to find ways to reallocate resources to those who value them higher, whether through purely illegal sublets or loophole-like behavior such as requiring tenants to rent non-rent-controlled items or join an organization whose dues are not subject to the government-imposed ceiling, as a condition of apartment rental.

5.4.2 The Invisible Hand and Spontaneous Order

Through the market as it has evolved in much of the developed world, Buchanan and Vanberg (1991) write that "individuals... use their own imagination, their own assessment of the potential evaluations of others," then, they "seek[] to satisfy others through producing market value as an indirect means of producing value for themselves." Adam Smith was among the first to anthropomorphize the market using the myth of the invisible hand rooted in the choice problems of individual agents.

> As every individual, therefore, endeavours as much as he can both to employ his capital in the support of domestic industry, and so to direct that industry that its produce may be of the greatest value; every individual necessarily labours to render the annual revenue of the society as great as he can. He generally, indeed, neither intends to promote the public interest, nor knows how much he is promoting it... [H]e intends only his own security; and by directing that industry in such a manner as its produce may be of the greatest value, he intends only his own gain, and he is in this, as in many other cases, led by an invisible hand to promote an end which was no part of his intention.

The individual agents who take actions within markets also take actions to shape and create markets. This is what is meant by the term **spontaneous order**. The theory of spontaneous order suggests that many well-structured social patterns or institutions arise as an unintended consequence of human action. Just as much complexity has arisen in nature from a gradual

evolution, the same has happened within human societies. Only a few years after Smith, Scottish philosopher and sociologist Adam Ferguson would write more directly about the emergence of spontaneous order.

> What was in one generation a propensity to herd with the species, becomes in the ages which follow, a principle of natural union. What was originally an alliance for common defense, becomes a concerted plan of political force; the care of subsistence becomes an anxiety for accumulating wealth, and the foundation of commercial arts. Mankind, in following the present sense of their minds, in striving to remove inconveniences, or to gain apparent and contiguous advantages, arrive at ends which even their imagination could not anticipate; and pass on, like other animals, in the tract of their nature, without perceiving its end.
> ...
> Every step and every movement of the multitude, even in what are termed enlightened ages, are made with equal blindness to the future; and nations stumble upon establishments, which are indeed the result of human action, but not the execution of any human design.

More recently, Ullmann-Margalit (1978) writes that an invisible-hand explanation is one that explains "what looks to be the product of someone's intentional design, as not being brought about by anyone's intentions." For the very reason that the social order is unintended, there can be perversity in markets. Institutions matter and shape the tradeoffs of individual actors. Peter Boettke (2012) summarizes the invisible hand proposition as: "individuals pursuing their own self-interest within an institution setting of property, contract, and consent will produce an overall order that, although not of their intention, enhances the public good. Absent that institutional setting, self-interest may very well not produce publicly desirable outcome and, in fact, may produce the opposite." It is the desire of individuals to improve themselves and their ingenuity that led, gradually, to the development of property, contract, and consent as the foundation for many developed economies today, but the fact that this is not universal is concrete proof that the spontaneous order need not move in this direction.

5.4.3 The Market Process and the Entrepreneur

To make it easier to discuss, some separate the market and its institutions as they exist in a particular moment in time from the overall process which the market facilitates. Geloso (2026) helpfully defines the market process as

"the decentralized process through which individuals, responding to profit and loss, discover information, create institutions, reduce transaction costs, redefine property rights, and innovate in ways that continuously reshape the economy." The Market Process literature considers more concretely how it is that we move to (or towards) equilibrium. The key is the entrepreneur.

In Economics, individual agents are modeled as seeking to improve their condition and benefit themselves find ways to create mutually beneficial exchanges. For instance, when the population expands causing a shift in the demand curve of food to the right, the producers of food will more quickly sell out. When these existing producers do not have the demanded quantity of food available, it is the entrepreneur who sees that more food could be made available who then risks something to acquire the resources to produce more food and makes it available. The entrepreneur is the one who has foresight to see how things could be made better and then acts to pursue this. Baumol (1968) writes that it is "his job to locate new ideas and to put them into effect." While a manager "oversees the ongoing efficiency of continuing processes," the entrepreneur is in some way an "innovator" who "must lead, perhaps even inspire [and] cannot allow things to get into a rut."

For Kirzner (1997), the entrepreneur is both the mechanism for moving from disequilibrium towards equilibrium and a force disrupting a *status quo* equilibrium creating new possibilities and new conditions.

> The theory of entrepreneurial discovery sees the explanation of market phenomena in the way entrepreneurial decisions, taken under disequilibrium conditions, bring about changes in prices and quantities. The market process so initiated consists of continual entrepreneurial discoveries; it is a process of discovery driven by dynamic competition, made possible by an institutional framework which permits unimpeded entrepreneurial entry into both new and old markets... [T]he result [is] a powerful tendency for less efficient, less imaginative courses of productive action to be replaced by newly discovered superior ways of serving consumers—by producing better goods and/or by taking advantage of hitherto unknown, but available, sources of resource supply.

The reason the economy is not static is due to entrepreneurial alertness to these profit opportunities and the willingness to take action. In some ways, *every* individual engages in some entrepreneurial activity, searching for ways to better themselves (profit opportunities). There are unnoticed profit opportunities that, because of this alertness, "can be expected to be systematically whittled away by [their] spontaneous discovery..." The

entrepreneur must notice and work to coordinate markets. When they do so, they generally must advertise to "*alert consumers* to the availability and the qualities of goods." The market is not just a solution to optimal price and quantity vectors, but actually facilitating exchange between two real people who must find each other. This is how the market process can seek to reduce transaction costs. Advertising can also "awaken potential customers to new preferences." Profit is only available because of the existence of opportunities in need of discovery. The entrepreneur has right to the profit, because without his discovery, there would be no product and no profit, Kirzner suggests.

Kirzner writes that, when we conduct equilibrium analysis, we are really looking at a process of entrepreneurs at work:

> If one could, for purely analytical purposes, imagine consumer preferences, resource availabilities, and technical possibilities as frozen in time, then the entrepreneurial discovery processes will tend to ensure that the price of any given good or service will tend towards equality throughout the market; that resource-bundle prices will tend to equality with the prices of the respective commodities they can deliver through production; that, at the uniform prices so achieved, the market for each consumer good or service and for each resource service will tend to clear; and that all prospective buyers will find what they wish to buy at the price they expect and all prospective sellers will find buyers prepared to pay the prices which sellers are expecting and are prepared to accept. ... In the real world of incessant change in underlying consumer preferences, resource availabilities, and technical possibilities, these corrective tendencies may be partly or wholly frustrated or interrupted. ... But the direction of the powerful forces of entrepreneurial discovery will be shaped and moulded by the above-described systematic and corrective processes or error, disappointment, discovery, and surprise.

This leads Kirzner to suggest that we judge economic institutions as successful when they better serve " a coordinative process during which market participants become aware of mutually beneficial opportunities for trade and, in grasping these opportunities, move to correct the earlier errors."

What is the difference between the entrepreneur and the utility-maximizing consumer and profit-maximizing producer? Entrepreneurs are the instigators of the process that consumers and producers carry on. An entrepreneur first invented a use for natural resources, first convinced others to trade his bundle of resources for their bundle, and who first began a regular division of labor.

Schumpeter (1942) speaks of the revolutions brought about by entrepreneurs creating a cycle of creative destruction. The entrepreneur built our economy. Then, as we accept his innovations, we maximize based on what we know and what we know how to do. But, it is the entrepreneur who continues to come back with new and better ways. Sometimes, the entrepreneur gets it wrong and goes bust, but sometimes he surprises himself with how right he got it, as what he builds goes beyond his dreams and initial intentions and anticipations. The process of entrepreneurial activity, Schumpeter writes, "incessantly revolutionizes the economic structure *from within*, incessantly destroying the old one, incessantly creating a new one." While Schumpeter acknowledges the importance of entrepreneurial competition over price and quality what is even more important, in his view, is "the competition from the new commodity, the new technology, the new source of supply, the new type of organization" which "strikes not at the margins of the profits and outputs of existing firms, but at the foundations and their very lives."

When engaging in a comparative static analysis, we are assuming that entrepreneurs are at work adjusting the market between the two equilibria. At the same time, we also know that their further work, which we are generally holding constant, will continue to disrupt the market and find alternative ways to satisfy desires. For instance, entrepreneurs can make some bad policy better. A rent control policy enforced by the government will not result in as poor outcomes as the simplest supply and demand shortage graph shows, because illegal sublets will, rent bundling, or further ingenious ways of lessening the shortage will end up improving outcomes. At the same time, because entrepreneurs do not have perfect foresight and knowledge, sometimes the equilibrium analysis can paint too rosy of a picture making painful, difficult, and uncertain adjustment appear simple, easy, and costless. Absent barriers, the market encourages the best use of knowledge and profit-and-loss ensure in the long-run a move towards equilibrium, but it does not guarantee perfection. In light of a sudden, unexpected supply shock, will prices rise the perfect amount to ensure absolutely no shortages and absolutely no surpluses? Of course, adjustments will take place. Microeconomic analysis hinges on the fact that agents consistently respond to incentives. But entrepreneurs will do the best that they can with the knowledge they have. Eventually, a combination of time and the actions of entrepreneurs will allow this knowledge to be disseminated until it is known, so something like the equilibrium emerges out of an hazy adjustment period where some set prices too high, others too low—as Kirzner (1997) writes, some entrepreneurs are over-pessimistic and some are over-optimistic. A central planner could do no better—and, in fact, would fare far worse due to

perverse incentives, the lack of an incentive to innovate—or the disincentive of trying something new—and the decentralization of knowledge. But, within the profit and loss system, the errors are more easily identifiable, discovered, and corrected.

5.4.4 The Nature of Competition

After entrepreneurial discovery and action takes place, **competition** is the next step. This is what brings the entrepreneur's insight to consumers in the marketplace. Hayek (1979) writes that it is, in essence, the experimentation phase of the discovery process. Telser (1987) writes that competition is "the nature of the conditions under which individuals may trade property rights" and "assumes a definition of these rights and describes a trading process."

Competition is a tricky word. It is naturally descriptive of the process in which sellers seek to woo consumers to trade with them. In modern parlance, it is oftentimes used as short-hand to describe market structure—i.e., is there a monopoly, duopoly, or some other form of market power?—but it is important to separate the discussions of market structure from discussions of the process of competition.

The model of **perfect competition** looks to market structure. It considers whether the market in question has

- a commodity that is homogeneous
- a commodity that is supplied by a large number of sellers, none of whom individually have the power to influence the market price with the quantity that they supply
- a commodity that is demanded by a large number of buyers, none of whom individually have the power to influence the price with the quantity that they demand
- free entry and exit with the absence of restraints on prices
- complete knowledge among all market participants of the relevant facts

Under this structure, producers do not compete over quality—as all goods in the market are homogeneous with the same qualities—and they do not compete over price—as all goods sell for the market price. Producers do not have to inform consumers of their supply, because consumers (and other producers) have complete insight into everything available to them and the behavior of all other producers and consumers. Under perfect competition, the actual process of competition is entirely absent. McNulty (1968) writes that this model of competition is merely the opposite of a monopoly. A market that fits these conditions, he writes, "has evolved or progressed to

the point of equilibrium where no further competition within the industry is possible."

He adds that the defining characteristic of "perfect competition" as compared to various "imperfect" forms (monopoly, duopoly, or monopolistic competition), is the demand curves faced by the firms. Under the perfectly competitive model, the effective demand curve is a horizontal line. So long as you are able to produce, you can sell at the market price. While the math is far simpler to say it is a discrete line covering only the points on that line, but, of course, consumers would be more than happy to pay *less* than the market price for a good. But, in this model, if a particular firm lowers its price, the quantity demanded is the same as if it had kept the price the same, so it faces no incentive to lower its price below the market level. If a firm raises the price, even slightly, above the market price, however, since the exact same commodity is widely available to consumers elsewhere at the (lower) market price, it will lose *all* of its demand and sell nothing.

In its opposite, a monopoly, the firm faces a downward sloping demand curve. It can only increase the quantity it is able to sell by lowering the price. This is only the case between goods that are heterogeneous—that is, this precise good is not freely available elsewhere for a given price—and clearly the monopolist has some control over the market price with the quantity that it desires to sell.

It is often said that this type of market structure could only be possible due to barriers to entry, but Demsetz (1982) questions how pervasive such barriers truly are. A common example is a government mandate for possession of one of the limited number of available government-granted taxi medallions to operate a taxi in a local area. Demsetz notes that this is only creates differential costs between existing producers and would-be new entrants if there are fewer available than the number of taxis that want to operate *and* if it is costly to obtain; however, if they could be purchased at market prices and could be resold, both insiders (existing taxis) and outsiders (potential would-be taxis) would face the same costs. The existing taxis would be giving up the market value of the medallion by continuing to operate rather than selling. Still, it is certainly true that, under certain conditions, government mandates, for instance, *can* impose significant barriers—such as the unsaleable medallions granted first-come, first-serve available in lower quantity than demanded or a proscription on provision of various goods and services, such as a government-guaranteed monopoly. As Telser (1987), likewise notes, the competitive equilibrium will depend upon the nature of the property rights.

In the initial edition of their Economics textbook in 1964, Alchian and Allen

were the first to interpret the common perfect vs imperfect competition models into an alternate way of viewing firm's choice problems. They drew the distinction between price-takers and price-searchers. For a price-taker—who can map onto the perfect competition model—the primary question is the quantity to produce given the market price to maximize profit. For a price-searcher—who maps onto the monopoly or related monopolistically competitive firm—the primary question is what price to set to maximize profit.

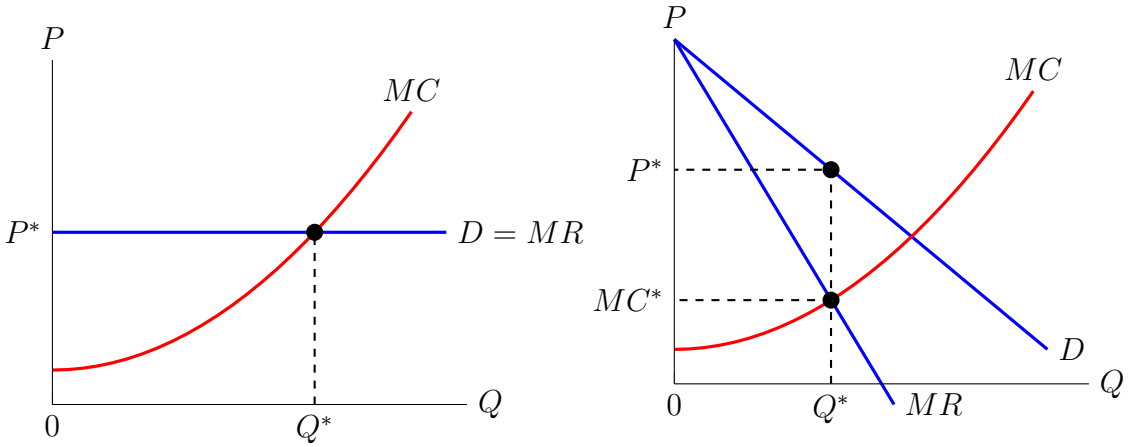


Figure 5.5: Price-Taker (left) vs. Price-Searcher (right)

Consider the market for fresh produce: with so many different farms producing very similar products, the market price is effectively a given, so the question to each farm is how much produce to grow. When it comes to the iPhone, it is a differentiated product. There are competitors to be sure, but no one else produces this identical product, so Apple must search for the appropriate price, and it does not always get it right on the first try. For instance, the very first iPhone released in the summer of 2007 was initially priced at \$499 for 4GB and \$599 for 8GB, but disappointing sales caused Apple to quickly realize that this was too high, inducing them to eliminate the 4GB model altogether just 2 months later and drop the price of the 8GB phone by a third to \$399 at the start of September 2007.

Hayek (1946) writes that the modern turn towards evaluating competition in light of market structure is entirely mistaken. When perfect knowledge is possessed and all goods are homogeneous, actual competition is far less important. Hayek argues that "nothing is solved when we assume everybody to know everything [when] the real problem is rather how it can be brought about that as much of the available knowledge as possible is used." Likewise,

since no two individuals are perfectly alike, so too are neither two producers identical. Because of this, we should not expect identical results, then. He argues that competition in the real world should not be judged against this "meaningless and impossible ideal," but rather against "the situation as it would exist if competition were prevented from operating.... a much bigger gulf divides competition from no competition than perfect from imperfect competition." Put differently, Hayek (1979) writes that competition should be judged by the *conditions*, not the *results*. "As is true of the results of scientific experimentation, we can judge the value of the results only by the conditions under which it was conducted, not by the results."

A model of the market is useful "because it tells how a system of this sort will work," but we cannot forget that these models do not give us the solution to our particular economic problems. A model can be useful in which we assume "that the theoretician possesses all the knowledge which guides all the several individuals whose interaction his model represents," but we cannot mistake this for real world conditions or the possibility of implementing some omniscient dictatorship.

> an omniscient dictator could indeed improve the use of available resources by requiring the firms to expand production until prices only just covered marginal costs. On this standard, habitually applied by some theorists, most of the markets in the existing world are undoubtedly very imperfect. For all practical problems, however, this standard is wholly irrelevant, because it rests on a comparison, not with some other state that could be achieved by some known procedure, but with one that might have been achieved if certain facts, which we cannot alter were other than they are. ... The real problem is how far we can raise efficiency above the pre-existing level, not how close we can come to what would be desirable if the facts were different.

Hayek (1979) presents a threefold defense of competition:

> Competition, if not prevented, tends to bring about a state of affairs in which: *first*, everything will be produced which somebody knows how to produce and which he can sell profitably at a price at which buyers will prefer it to the available alternatives; *second*, everything that is being produced is produced by persons who can do so at least as cheaply as anybody else who in fact is not producing it; and *third*, that everything will be sold at prices lower than, or least as low as, those at which it could be sold by anybody who in fact does not do so.

This is precisely what the general equilibrium model of economics seeks to show to which we will turn in the next chapter.

References

Baumol, William J. 1968. "Entrepreneurship in Economic Theory," *American Economic Review*, Vol. 58, No. 2, pp. 64-71.

Boettke, Peter J. 2012. "Economics for Yesterday, Today, and Tomorrow, *Living Economics: Yesterday, Today, and Tomorrow*, The Independent Institute, pp. 1-13.

Buchanan, James M. and Viktor J. Vanberg. 1991. "The Market as a Creative Process," *Economics & Philosophy*, Vol. 7, No. 2, pp. 167-186.

Demsetz, Harold. 1982. "Barriers to Entry," *American Economic Review*, Vol. 72, No. 1, pp. 47-57.

Ferguson, Adam. 1782. "The History of Political Establishments," *An Essay on the History of Civil Society*, T. Cadell in the Strand, pp. 203-225.

Geloso, Vincent. 2026. "Market Processes, Market Failures, and Government Failures," *Foundation for Economic Education*, July 21, 2026, `https://vincentgeloso.wordpress.com/wp-content/uploads/2026/07/slides_for_presentations-2.pdf`.

Hayek, F. A. 1948 [1946]. "The Meaning of Competition," *Individualism and the Economic Order*, University of Chicago Press, pp. 92-106.

Hayek, F.A. 2013 [1979]. "Government Policy and the Market," *Law, Legislation, and Liberty: A New Statement of the Liberal Principles of Justice and Political Economy*, Routledge Classics, pp. 404-433.

Kirzner, Israel M. 1997. *How Markets Work: Disequilibrium, Entrepreneurship, and Discovery*, Institute of Economic Affairs.

Lancaster, Kelvin J. 1966. "A New Approach to Consumer Theory," *Journal of Political Economy*, Vol. 74, No. 2, pp. 132-157.

McNulty, Paul J. 1968. "Economic Theory and the Meaning of Competition," *Quarterly Journal of Economics*, Vol. 82, No. 4, pp. 639-656.

Morgenstern, Oskar. 1976 [1935]. "Perfect Foresight and Economic Equilibrium," *Selected Economic Writings of Oskar Morgenstern*, edited by Andrew Schotter, translated by Frank H. Knight, New York University Press, pp. 169-183.

Schumpeter, Joseph A. 1942. "The Process of Creative Destruction," *Capitalism, Socialism, and Democracy*, Harper & Brothers Publishers, pp. 81-86.

Smith, Adam. 1979 [1776]. *An Inquiry into the Nature and Causes of the Wealth of Nations, Volume 1*, edited by R.H. Campbell, A.S. Skinner, and W.B. Todd, Liberty Fund.

Stigler, George J. and Gary S. Becker. 1977. "De Gustibus Non Est Disputandum," *American Economic Review*, Vol. 67, No. 2, pp. 76-90.

Telser, Lester G. 1987. "Competition, Cooperation, and Efficiency," *A Theory of Efficient Cooperation and Competition*, Cambridge University Press, pp. 47-73.

Ullmann-Margalit, Edna. 1978. "Invisible Hand Explanations," *Synthese*, Vol. 39, pp. 263-291.

Wagner, Richard E. 1983. "On Economics, Political Science, and Public Policy," *Policy Studies Journal*, Vol. 12, No. 2, pp. 251-257.

Chapter 6

General Equilibrium Theory

6.1 General Equilibrium

In conducting partial equilibrium analysis, we consider the solution for price and quantity for a single good or a particular firm or market. **General Equilibrium** looks at the broader economy as a whole and seeks to find price vectors for all goods and quantity vectors for all individuals or firms. Since each market is connected to each other due to substitutability and scarcity, the optimal result is mutually dependent.

General Equilibrium holds when all agents maximize utility without violating their budget constraints. In his 1874 work, *Elements of Pure Economics*, Leon Walras modeled the economy as a simultaneous auction in which the market works like an auctioneer crying out various prices until the market clears. He labeled the process of market clearing "trial-and-error"—often referred to using the original French word "tâtonnement." During this process, when one considers an economy moving toward market clearing, we hold constant production functions and preferences (and utility functions).

Arrow and Debreu (1954) famously proved mathematically the existence of general equilibrium under specific conditions. They write,

> Loosely speaking... if every individual has initially some positive quantity of every commodity available for sale, then a competitive equilibrium will exist.... [Likewise, there will exist a] competitive equilibrium if there are some types of labor with the following two properties: (1) each individual can supply some positive amount of at least one such type of labor; and (2) each such type of labor has a positive usefulness in the production of desired commodities.

We can guarantee the existence of a general equilibrium when we have

1. Completeness – there are utility functions for each consumer that span all possible goods such that all $U_i(k)$ is defined for each individual i and each good k.
2. Continuity – continuous functions
3. Convexity – Utility functions and indifference curves with a diminishing marginal rate of substitution and production returns to scale are constant or decreasing.
4. Local non-satiation – demand is non-zero and downward sloping: for a high enough price, each good k has a quantity demanded that falls short of the quantity available; for a low enough price, each good k has a quantity demanded that exceeds the quantity available.
5. Compactness – there are finite limits to prices and quantities, so using a Fixed Point Theorem, we can prove that there exists a point where $f(x) = x$, the point where excess demand for all goods is zero, meaning that markets clear.

Of course, while these conditions are sufficient for a general equilibrium, they are not mandatory. A general equilibrium may exist even if these particular conditions do not hold.

6.1.1 Pure Exchange General Equilibrium

We can consider a simple two commodity world with commodities x_1 and x_2 and no production. In this case, each agent i has a given utility function U.

$$U_i = x_1^{a_i} x_2^{b_i}$$

and for simplicity, we can take this as a Cobb-Douglas function where $a_i + b_i = 1$ for each individual i. Likewise, since utility is invariant to monotonic transformations, we can simplify this using natural logs to get rid of the exponents:

$$U_i = a_i \ln x_1 + b_i \ln x_2$$

Each agent is endowed ω of each good and their consumption is measured by x. In General Equilibrium, every good is consumed, so:

$$\sum x_1 = \sum \omega_1 \quad \sum x_2 = \sum \omega_2$$

For each individual agent, their wealth is equivalent to the market value of their endowments:

$$W_i = p_1 \omega_{1i} + p_2 \omega_{2i}$$

Each individual's total consumption is equivalent to the market value of the goods they consume:

$$C_i = p_1 x_{1i} + p_2 x_{2i}$$

So, each agent's choice problem is

$$\max_{x_{1i}, x_{2i}} U_i \text{ s.t. } C_i = W_i$$

$$\max_{x_{1i}, x_{2i}} (a_i \ln x_1 + b_i \ln x_2) \text{ s.t. } p_1 x_{1i} + p_2 x_{2i} = p_1 \omega_{1i} + p_2 \omega_{2i}$$

As a multivariate optimization problem, we can solve using a Lagrangian.

$$\mathcal{L} = U_i + \lambda(g(x) \stackrel{set}{=} 0)$$

So, solving our constraint for 0:

$$p_1 x_{1i} + p_2 x_{2i} = p_1 \omega_{1i} + p_2 \omega_{2i}$$

$$p_1 \omega_{1i} + p_2 \omega_{2i} - p_1 x_{1i} - p_2 x_{2i}$$

So, agent i's Lagrangian is:

$$\mathcal{L}_i(x_{1i}, x_{2i}, \lambda_i) = a_i \ln x_1 + b_i \ln x_2 + \lambda_i \left(p_1 \omega_{1i} + p_2 \omega_{2i} - p_1 x_{1i} - p_2 x_{2i}\right)$$

Then, taking the first order conditions and setting equal to zero:

$$\frac{\partial \mathcal{L}_i}{\partial x_{1i}} = \frac{a_i}{x_{1i}} - \lambda_i p_1 \stackrel{set}{=} 0$$

$$\frac{\partial \mathcal{L}_i}{\partial x_{2i}} = \frac{b_i}{x_{2i}} - \lambda_i p_2 \stackrel{set}{=} 0$$

$$\frac{\partial \mathcal{L}_i}{\partial \lambda_i} = p_1 \omega_{1i} + p_2 \omega_{2i} - p_1 x_{1i} - p_2 x_{2i} \stackrel{set}{=} 0$$

Simplifying, we find:

$$p_1 x_{1i} = \frac{a_i}{\lambda_i}$$

$$p_2 x_{2i} = \frac{b_i}{\lambda_i}$$

Substituting into our λ FOC and solving for $\frac{1}{\lambda}$:

$$p_1 \omega_{1i} + p_2 \omega_{2i} - \frac{a_i}{\lambda_i} - \frac{b_i}{\lambda_i} = 0$$

$$p_1 \omega_{1i} + p_2 \omega_{2i} = \frac{a_i}{\lambda_i} + \frac{b_i}{\lambda_i}$$

$$p_1\omega_{1i} + p_2\omega_{2i} = \frac{1}{\lambda_i}(a_i + b_i)$$

$$\frac{p_1\omega_{1i} + p_2\omega_{2i}}{a_i + b_i} = \frac{1}{\lambda_i}$$

Now, substituting back into our condition for x_{1i} and solving for the solution (x^*_{1i}):

$$p_1 x_{1i} = \frac{a_i}{\lambda_i} = a_i \left(\frac{p_1\omega_{1i} + p_2\omega_{2i}}{a_i + b_i} \right)$$

$$x_{1i} = \frac{a_i}{p_1} \cdot \frac{p_1\omega_{1i} + p_2\omega_{2i}}{a_i + b_i}$$

So, across the entire economy,

$$\sum x_{1i} = \sum \frac{a_i}{p_1} \frac{p_1\omega_{1i} + p_2\omega_{2i}}{a_i + b_i}$$

Since we have assumed Cobb-Douglas form, $a_i + b_i = 1$ for all i, so:

$$\sum x_{1i} = \sum \frac{a_i}{p_1} \cdot (p_1\omega_{1i} + p_2\omega_{2i})$$

$$\sum x_{1i} = \frac{p_1 \sum (a_i\omega_{1i}) + p_2 \sum (a_i\omega_{2i})}{p_1}$$

Since the markets clear at general equilibrium, $\sum \omega_{1i} = \sum x_{1i}$:

$$\sum \omega_{1i} = \frac{p_1 \sum (a_i\omega_{1i}) + p_2 \sum a_i\omega_{2i}}{p_1}$$

$$p_1 \sum \omega_{1i} = p_1 \sum (a_i\omega_{1i}) + p_2 \sum a_i\omega_{2i}$$

$$p_1 \sum \omega_{1i} - p_1 \sum (a_i\omega_{1i}) = p_2 \sum a_i\omega_{2i}$$

$$p_1 \left[\sum \omega_{1i} - \sum (a_i\omega_{1i}) \right] = p_2 \sum a_i\omega_{2i}$$

$$p_1 \left[\sum (1 - a_i)\omega_{1i} \right] = p_2 \sum a_i\omega_{2i}$$

$$\frac{p_1}{p_2} = \frac{\sum a_i\omega_{2i}}{\sum (1 - a_i)\omega_{1i}}$$

$$\frac{p_1}{p_2} = \frac{\sum a_i\omega_{2i}}{\sum b_i\omega_{1i}}$$

What is the result then? In this exchange economy, prices are a function of both individual preferences—a and b—and the endowments ω_k. As we

increase the quantity of good 2, the ratio of p_1 to p_2 rises—a given unit of good 1 is able to purchase *more* of good 2, holding all else constant. As we increase the quantity of good 1, the ratio of p_1 to p_2 falls—a given unit of good 1 is able to purchase *less* of good 2, holding all else constant. As the relative preference for good 1 a_i increases, it takes more of good 2 to acquire a unit of good 1. The reverse is also true.

6.1.2 Say's Law and Walras' Law

Nearly a century before Walras would exposit the idea we now think of as General Equilibrium, Say's Law served as the foundation for understanding the relation between markets. In Jean Baptiste Say's (1803) *Treatise on Political Economy*, he writes a chapter on "The Demand or Market for Products," in which he wrote:

> a product is no sooner created, than it, from that instant, affords a market for other products to the full extent of its own value. When the producer has put the finishing hand to his product, he is most anxious to sell it immediately, lest its value should diminish in his hands. Nor is he less anxious to dispose of the money he may get for it; for the value of money is also perishable. But the only way of getting rid of money is in the purchase of some product or other. Thus, the mere circumstance of the creation of one product immediately opens a vent for other products.

Sometimes markets do not clear, but that is because "precisely at the same time that one commodity makes a loss, another commodity is making excessive profit." Walras would later formalize this part of the principle. Walras' Law states that, so long as all agents are operating within their budget constraints, the total sum of excess demand in monetary terms across all markets in an economy is equal to 0. If there is excess demand for some commodity, there is necessarily excess supply of some other commodity (or commodities). Put differently, the rule states that for an economy with N markets, if $N-1$ markets are in equilibrium (with no excess demand nor excess supply in any of the individual markets), the final market is also in equilibrium.

Leijonhufvud and Clower (1973) write that Say's Principle is that "the net value of an individual's planned trades is identically zero... [B]roadly speaking, trade is a two-way street: i.e., individuals can expect to acquire commodities from other individuals only by giving commodities (or money) of equal market value in exchange." We only reach a complete equilibrium in which "all trades can be executed as planned" when there is no aggregate

excess demand for any commodity.

More recently, some economists, such as Yeager (1956), have noted the importance of noting the market for money. When we adopt a medium of exchange to facilitate exchange, we cannot forget that the medium of exchange itself can serve as the n^{th} market. If Walras's Law appears to be violated—that is, if there appears to be an excess supply of goods or excess demand for goods after summing up each market—this will be due to a change in the market for money. An excess demand for goods is likely to happen when the money supply has expanded yet prices have not (yet) risen. An excess supply of goods is likely when a shock causes people to hoard up money and seek to avoid spending on anything until the prices of individual goods fall such that people reach their desired savings in real (purchasing power) terms.

6.2 Markets and Efficiency

A key question to be considered is whether the market outcome at General Equilibrium is efficient. Using the Pareto criterion for efficiency, we ask: is it possible at equilibrium to make at least one person better off without making anyone else worse off? When analyzing efficiency in equilibrium, we are considering at a specific moment in time the state of the world given the resources (including knowledge of how to use said resources) and individual preferences.

6.2.1 Exchange Efficiency

If individual agents are utility maximizers, they will exchange with each other until they have exhausted all gains from trade. In the example below with a fixed bundle of two goods and two individuals, the initial position A is clearly not an equilibrium. Any trade within the shaded area would make both individuals better off—putting at least one individual on a higher indifference curve. Recall, for an Edgeworth Box, the origin for Individual 1 is the bottom left corner with indifference curves convex to his origin, while the origin for Individual 2 is the top right corner with his indifference curves convex to this origin.

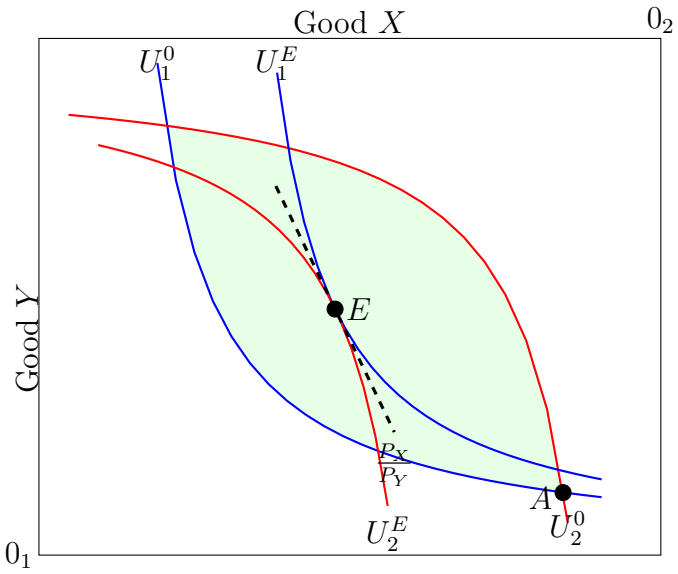


Figure 6.1: Exchange Efficiency

Until the indifference curves are tangent, there are Pareto Improvements: the agents could do better and the resources could be reallocated in such a way as to provide more value as described by *at least one* agent and as to leave *every* agent equally well off.

Consider the slope of the indifference curve. Using the total differential, in order to keep U constant:

$$\delta U = \frac{\partial U}{\partial X}\delta X + \frac{\partial U}{\partial Y}\delta Y \stackrel{set}{=} 0$$

$$\delta U = MU_X \delta X + MU_Y \delta Y = 0$$

$$MU_Y \delta Y = -MU_X \delta X$$

$$\frac{\delta Y}{\delta X} = -\frac{MU_X}{MU_Y}$$

Thus, the slope along the indifference curve, our Marginal Rate of Substitution, is the ratio of the marginal utilities of the two goods. At point A, for instance $MRS^1 < MRS^2$. For instance, person 1 might be willing to trade 5 units of X for 1 unit of Y, while person 2 (who has a lot of Y but not a lot of X) might be willing to trade 1 unit of X for 5 units of Y. Using this numbers:

$$MRS^1_{XY} = \frac{1}{5} < MRS^2_{XY} = 5$$

This is demonstrative of the potential for mutually beneficial exchange, so such a condition is *not* equilibrium. We only reach the state of rest when the Marginal Rates of Substitution reach equality. When the indifference curves are tangent to each other, their slope is equal, thus, for our two individuals and two goods scenario:

$$\frac{MU_X^1}{MU_Y^1} = \frac{MU_X^2}{MU_Y^2}$$

$$MRS_{XY}^1 = MRS_{XY}^2$$

The dashed line in the Edgeworth Box of Figure 6.1 represents the exchange ratio at this point of tangency. This is where the quantity of units of good Y each individual would be willing to exchange for one unit of good X is the same, so neither individual engages in further transactions.

We can also write our Marginal Rate of Substitution in terms of the Marginal Value—that is, the height of the consumer's individual demand curve. While the two are technically equivalent, Harberger (1971) writes that "the conversion of utility into money... greatly eases the aggregation problem." We cannot scientifically engage in interpersonal comparisons of utility, we can add up money values of willingness to pay. In the two good scenario below, the height of the demand curve is equivalent to the quantity of Y the individual is willing to give up for another unit of X. While we often think of the height of the demand curve in terms of money prices, the medium of exchange could be anything, so units of Y are equally valid as the numeraire.

The Marginal Rate of Substitution represents the consumer's marginal willingness to pay: how much of the numeraire (Y) that they would be willing to give up to receive the next unit of the good. We can also graph the individual's demand curve for the two-good scenario with good X on the x-axis and our numeraire (good Y) on the y-axis. However, typically, we don't think of exchange in terms of *any* good, but rather in terms of the primary medium of exchange: money. For this, we can use the term Marginal Value.

The Marginal Value is equivalent to the marginal rate of substitution between good X and money. There is fundamentally no difference between the scenario in which an individual has to choose between only two goods X and Y and either good X has a market price of \$5 a unit and good Y has a market price of \$10 a unit or if the two goods have a market exchange ratio of 1 unit of X for 2 units of Y.

If the individual would be willing to give up \$7 for one X, but only has to pay \$5, then we say $MV_X > P_X$ where the price of X. So long as, for the first unit of X, $MV_X > P_X$, the maximizing consumer will purchase unit(s) of good X until $MV_X = P_X$. This is representative of the maximizing consumer

moving down his demand curve until the added consumer surplus from the next unit is 0—that is, the consumer would be indifference between keeping his money (or medium of exchange good Y) and giving up the requisite quantity of money (or good Y) to receive an additional unit of good X. It is important to remember that this activity is identical to the consumer within an Edgeworth Box changing his allocation of goods to move to a higher indifference curve.

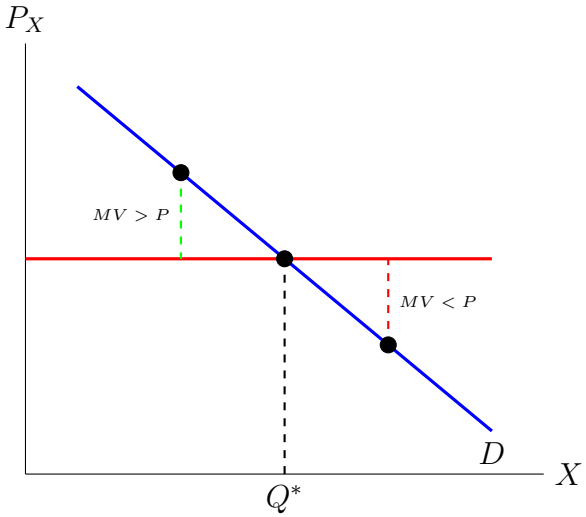


Figure 6.2: Consumer's Willingness to Pay

Since the Marginal Value is equivalent to the Marginal Rate of Substitution between good X and money:

$$MV \equiv \frac{MU_X}{\lambda}$$

where λ is the Marginal Utility of income (or money). So when we say

$$MV_X > P_X$$

this means that

$$\frac{MU_X}{\lambda} > P_X$$

When there are only two goods and the consumer is not satiated—that is, more of at least one good will provide more utility, the marginal utility of income comes from the marginal utility of the other potential good Y and the money price that matters is the relative price between the two goods:

$$\frac{MU_X}{MU_Y} > \frac{P_X}{P_Y}$$

We will only reach equilibrium and the agent will stop engaging in more mutually beneficial transaction when:

$$\frac{MU_X}{MU_Y} = \frac{P_X}{P_Y}$$

Recall that this is equivalent to the slope of an indifference curve being tangent to the budget constraint. Note that we can rewrite this condition as:

$$\frac{MU_X}{P_X} = \frac{MU_Y}{P_Y}$$

This means the individual is maximizing the bang for the buck—the utility per price. For a multi-commodity economy with N goods that are purchased with good Z not purchased, at equilibrium

$$\frac{MU_1}{P_1} = \frac{MU_2}{P_2} = \frac{MU_3}{P_3} = ... = \frac{MU_N}{P_N} \geq \frac{MU_Z}{P_Z}$$

Consider that the marginal utility of income λ will also be equal to this ratio, since $MU_{\text{Income}} \equiv \lambda$ and $P_{\text{Income}} = 1$. So:

$$\lambda = \frac{MU_1}{P_1} = ... = \frac{MU_N}{P_N}$$

When all consumers face the same price ratio, all individuals purchase until $MV_X = P_X$, so, for all individuals,

$$MRS^1_{XY} = MRS^2_{XY} = ... = MRS^K_{XY}$$

$$\frac{MU^1_X}{MU^1_Y} = \frac{MU^2_X}{MU^2_Y} = ... \frac{MU^K_X}{MU^K_Y}$$

In a competitive market where consumers face the same price ratio, we have efficiency in exchange and all goods are distributed to those who value them the highest. So long as we end up anywhere along the Edgeworth Box contract curve with tangent indifference curves, the marginal rates of substitution are equal and we have efficient exchange.

6.2.2 Input Efficiency

Are we using our inputs efficiently to generate outputs? This question asks whether it is possible to reallocate inputs such that no firm ends up producing less but some firms are able to produce more. If firms are cost minimizers, they will seek to use their inputs efficiently and produce using the mix of

inputs such that the Isoquant is tangent to their Isocost line. An Isoquant measures the varying levels of inputs to produce the same level of output. Thus, this represents the firm's indifference curve. At each point along the Isoquant, the firm is able to produce the same quantity of output.

For Q as the output, the slope of the Isoquant is:

$$\delta Q = \frac{\partial Q}{\partial L}\delta L + \frac{\partial Q}{\partial K}\delta K \stackrel{set}{=} 0$$

$$MPL\delta L + MPK\delta K = 0$$

$$MRTS_{LK} = \frac{MP_L}{MP_K} = -\frac{\delta K}{\delta L}$$

When the firm is setting their Isoquant equal to the cost ratios of inputs and the slopes are equal:

$$MRTS_{LK} = \frac{MP_L}{MP_K} = \frac{MC_L}{MC_K}$$

The ratio of the marginal products is equal to the ratio of the marginal costs.

So, for a given quantity of output that the firm is seeking to produce, the optimizing firm will produce at the point of tangency between the Isocost and Isoquant. When all firms face the same cost ratios, then for every firm j:

$$\frac{MP_L^j}{MP_K^j} = \frac{MP_L^1}{MP_K^1} = \frac{MP_L^2}{MP_K^2} = ... = \frac{MP_L^M}{MP_K^M} = \frac{MC_L}{MC_K}$$

Likewise, we can rewrite the firm's equality as:

$$\frac{MP_L}{MC_L} = \frac{MP_K}{MC_K}$$

this is that the marginal product per dollar of labor is equal to the marginal product per dollar of capital. At equilibrium, the marginal returns per dollar of input are equal for the individual firm. When all firms face the same marginal cost ratio, the returns per dollar of input equalize between all firms.

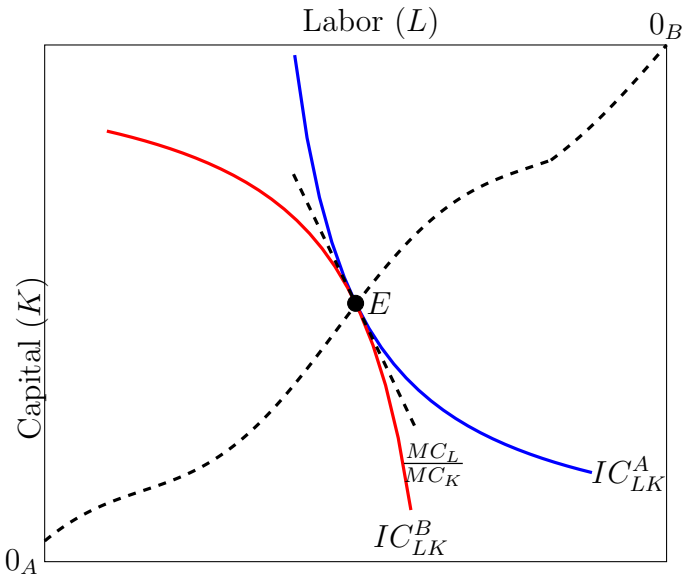


Figure 6.3: Input Efficiency

For a limited amount of capital and labor available in the world, firms must compete with each other to gain purchasing power of these resources. So we can think of firms as trading the resources amongst themselves—even if, for instance, we do not often think of the initial resource owners as other firms (such as labor).

At equilibrium, firms will continue to trade until every firm is on the contract curve line with their marginal product ratio equal to the marginal cost ratio of the inputs. For various firms that are producing different goods, there will be varying Isoquant functions—while exclusively convex to the origin, the substitution ratios along the curve will vary depending on the specifics of the production function. Thus, not every firm will use the same *mix* of inputs, because the production functions vary. Nonetheless, at the margin, they will mix resources until the *marginal* product of the inputs is equal among all firms at equilibrium.

When the marginal products are all equal—because each is set equal to the marginal cost ratio—there is no potential reallocation of resources that can allow any firm to produce more goods without making another firm produce fewer goods.

6.2.3 Output Efficiency

Finally, to achieve a Pareto efficient state of the world, we want to ensure that our productive resources are used to produce goods that consumers actually want. Firms have a near-infinite catalog of potential goods that they could produce. How do they choose which projects to pursue with their given resource inputs?

For analyzing this, we consider the Production Possibilities Frontier. The slope of the PPF between two goods shows many units of good X we can produce by reallocating the inputs used from producing one unit of good Y to the production of good X. This is called the **Marginal Rate of Transformation**. Rather than substituting between inputs, as measured by the Marginal Rate of Technical Substitution, the Marginal Rate of Transformation measures the marginal cost of Y in units of X.

$$MRT \equiv \frac{MC_X}{MC_Y}$$

A single individual who produces everything he consumes would maximize by finding the consumption bundle on his production possibility frontier that allows him to reach the highest indifference curve. In effect, he sets his indifference curve tangent to his PPF. If the individual produced, for instance, at A, he could receive a higher satisfaction from producing less Y and more X at point E moving from a lower indifference curve U_L to a higher curve U_H. At point A, $MRS_{XY} > MRT_{XY}$, so the optimizing individual will switch in favor of producing X until

$$MRS_{XY} = MRT_{XY}$$

$$\frac{MU_X}{MU_Y} = \frac{MC_X}{MC_Y}$$

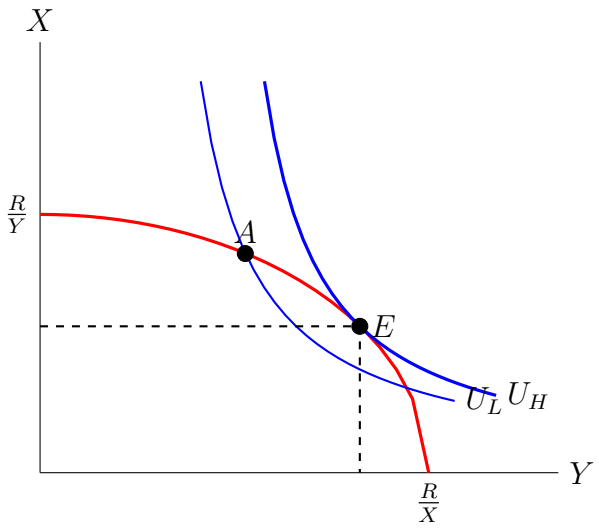


Figure 6.4: Output Efficiency

When we generalize beyond a single individual, it is no longer a single individual making both the productive and consumptive decisions. The condition would still hold at the optimum. If each member of the community was on indifference curve U_L, the situation is not Pareto efficient, because we can reallocate inputs towards producing less Y and more X to make everyone better off. The price system coordinates this equality. So long as consumers face the same price ratio and producers receive the same price ratio paid by the consumers:

$$MRT_{XY} = \frac{MC_X}{MC_Y} = \frac{P_X}{P_Y} = MRS_{XY} = \frac{MU_X}{MU_Y}$$

When this condition holds, this is **Output Efficiency** or sometimes referred to as Product-Mix Efficiency.

When we are dealing in terms of money—rather than in terms of a numeraire good—the Marginal Rate of Transformation is equivalent to the Marginal Cost of producing another unit of X in money terms. If firms are profit maximizers, they will produce until the cost of producing the next unit is equivalent to the revenue received from selling that unit—that is, $P = MC$. As we have already seen, the consumer will buy until their marginal willingness-to-pay equals the price of the next unit:

$$MV_X = P_X$$

When the price the producer receives is equal to the price the consumer pays:

$$MC_X = P_X = MV_X$$

Thus, the Marginal Rate of Transformation for firms is set equal to the Marginal Rate of Substitution for consumers.

At this equilibrium, no alternative mix of output (for the given resources) could move any consumer onto a higher indifference curve without moving someone else to a lower curve. This is a Pareto Efficient productive output.

6.2.4 First Fundamental Theorem of Welfare Economics

To review, in a competitive market where all consumers and producers face the same price ratios, we will have:

1. Exchange Efficiency – allocation of given goods among consumers is Pareto Efficient.
   - For each consumer,
   $$MRS_{XY} = \frac{MU_X}{MU_Y} \overset{set}{=} \frac{P_X}{P_Y}$$
   - For all consumers i (and N total consumers),
   $$MRS^i_{XY} = MRS^1_{XY} = MRS^2_{XY} = ... = MRS^N_{XY}$$
   - Because consumers are utility maximizers.
2. Input Efficiency – allocation of given inputs among producers is Pareto Efficient.
   - For each producer,
   $$MRTS_{LK} = \frac{MP_L}{MP_K} \overset{set}{=} \frac{MC_L}{MC_K}$$
   - For all producers j (and M total producers),
   $$MRTS^j_{LK} = MRTS^1_{LK} = MRTS^2_{LK} = ... = MRTS^M_{LK}$$
   - Because firms are cost minimizers.
3. Output Efficiency – allocation of productive resources amongst production for each given good is Pareto Efficient.
   - For each individual i and each firm j
   $$MRS^i_{XY} = \frac{MU_X}{MU_Y} = \frac{MC_X}{MC_Y} = MRT^j_{XY}$$
   $$MV^i_X = MC^j_X$$

- Because firms are profit maximizers.

The First Theorem of Welfare Economics is that if consumers are rational utility maximizers and producers are cost minimizers and profit maximizers, the General Equilibrium result of competitive markets is Pareto efficient. At General Equilibrium, exchange efficiency, production efficiency, and product-mix efficiency all hold. As Leeson (2019) notes, the key assumption is maximization. "Maximization implies efficiency, always and everywhere, because maximizers maximize, always and everywhere." However, this does not mean that the market is free from mistakes or errors.

> Unlike gods, people have limited cognitive abilities, limited abilities to execute their intentions, limited information. Ex post, many choices that maximizers make turn out to be "wrong." But the institutional results are no more "inefficient" than a marksman's wide shots. Pursuing a goal as best you can is no guarantee you'll achieve it. ... In a world of costly information, maximization *requires* "mistakes."

This First Fundamental Theorem means that, in such an equilibrium, it is impossible to make one person better off without making another person worse off. This is a type of formalization of Adam Smith's theory of the invisible hand, that, by pursuing their own interests, agents end up promoting social welfare, broadly speaking.

As we discussed in the last chapter, a key question is how one arrives at such an equilibrium. Note that all of the above conditions of efficiency require all profit opportunities to have been exploited. We reach equilibrium when our ratios are equal, but, until we get to this equality, there are profit opportunities. John Bates Clark (1889), along with his contemporaries of Leon Walras—for whom Walrasian General Equilibrium is named—and Vilfredo Pareto—for whom Pareto efficiency is named—argued that competition leads to the maximum satisfaction of human wants. Clark suggests, however, that this maximum is only possible because of the dynamic process of profit *prior to* equilibrium.

> The whole output of industry is at its maximum when the agents—[both] labor, and capital—are apportioned among the groups in a perfectly normal way; and that is when they have moved to the groups where profits have existed, till these gains have vanished and wages and interest have absorbed the whole social income. ... Profit is the lure that insures improvement, and improvement is the source of permanent additions to wages.

The mathematical mappings of the First Fundamental Theorem typically

involve the following three assumptions:

- perfect competition – all firms are price-takers
- perfect information – all agents costlessly know all relevant data about all prices, quantities, quality, and intentions of other agents.
- complete markets – all goods and services are tradable.

The first and last assumptions are fairly intuitive. If some firms are price-searchers and unable to price discriminate due to information costs, we have units for which individual's willingness to pay—points along the demand curve—lie above the marginal cost curve—since they represent negative marginal revenue for the firm, who must lower the price for *everyone* to sell these goods. If certain goods or service are not capable of exchange—the market is incomplete—we cannot guarantee efficiency, because such good or service remains with the individual to whom it is endowed, who may or may not value it more than someone else.

The information assumption is the most complicated and misunderstood. Critics of the First Fundamental Theorem and its relevance—of whom Stiglitz is likely the most famous—point to the reality of asymmetric information as being Pareto inefficient. The buyer's willingness to pay may be higher than the seller's willingness to accept, but since each individual only knows his own willingness and seeks to portray a lower (for the buyer) or higher (for the seller) willingness than he actually possesses to the other party, we may be able to make both parties better off, even if their interactions lead them to *think* this is not possible.

The Greenwald-Stiglitz Theorem formalized these assumptions concluding that "economies in which there are incomplete markets and imperfect information are not, in general, constrained Pareto efficient [and that] there exist government interventions (e.g., taxes and subsidies) that can make everyone better off." What Greenwald and Stiglitz (1986) show is that absent perfect information and complete markets, a Pareto efficient equilibrium is not *guaranteed* to exist. But take careful note to avoid falling prey to the fallacy of illicit contraposition. Imperfect information means that a Pareto optimal allocation is no longer guaranteed, but note that it does not mean a Pareto optimal allocation is impossible. We possess imperfect information all of the time and yet mutually beneficial exchanges take place all the time. Of course, such exchanges are certainly made *less* likely, depending on the degree of obfuscation and the availability of remedies.

However, is the willingness to pay the only factor in measuring efficiency? Critics of the First Theorem's applicability often seem to suggest so. If this is the case, transaction costs guarantee inefficiency, because one person's

willingness to pay may exceed another's willingness to accept, but not by the amount of the transaction costs. But such an exclusion is arbitrary, as Demsetz (1969) points out. Why would measurements of efficiency include anything short of *all* relevant costs? When we fail to consider all costs, we compare a hypothetical utopia to reality. A person in the midwest may be willing to pay more for New England crabs than a Maine fisherman is willing to accept at his fish stand, but this fact is trivial unless we factor in the costs of transporting the crabs and actually connecting the buyer and seller to each other.

The Greenwald-Stiglitz model points to potential externalities created by imperfect information—externalities, which, by definition, since agents are assumed to maximize, must exist due to the costs of transacting. For instance, as already discussed in considering the nature and structure of firms, monitoring of employees is costly in practice. If an omniscient being informed employers of which employees will work hard and which will shirk their responsibilities, we could improve outcomes for everyone. Instead, employers have to resort to management techniques or developing new firm structures and devoting resources towards monitoring rather than to producing more goods that can be consumed—or, following the Shapiro-Stiglitz efficiency wages model, they pay higher than market wages to prevent shirking, raising overall unemployment. Per Greenwald-Stiglitz and Stiglitz (2002), such a situation is Pareto inefficient. They suggest that subsidizing employment would be a Pareto improvement, assuming that the newly unemployed individuals would be able to compensate those who are taxed for their subsidy and still be better off. If this were the case, a profit opportunity remains available for firms and the unemployed and thus they would be incentivized to act. If individuals are maximizers and such an apparent profit opportunity remains *in equilibrium*, this must be merely an apparent—but not an actual—profit opportunity. The cost of acquiring information, revealing information, negotiating, or otherwise engaging in activities involved in such a transaction must be greater than the expected return.

6.3 Imperfect Competition and Market Efficiency

Now, we've seen the key assumptions that *guarantee* that the market outcome will be Pareto efficient, but as many point out, some of these assumptions do not hold in our real world markets. Turning to competition in particular, does the existence of a state other than perfect competition ensure that markets

are highly inefficient and that willingness-to-pay and willingness-to-accept (when accounting for *all* relevant costs)?

Consider the price-searching (monopolist) firm for whom marginal revenue is not equivalent to the price. Harberger (1954) identified the Deadweight Loss or Welfare Loss triangle. This represents units for which

$$MV_X > MC_X$$

That is, the willingness-to-pay measured by the height of the demand curve is greater than the cost to produce that unit. A monopolist produces an inefficiently small quantity when producing at the profit maximizing point $MR = MC$.

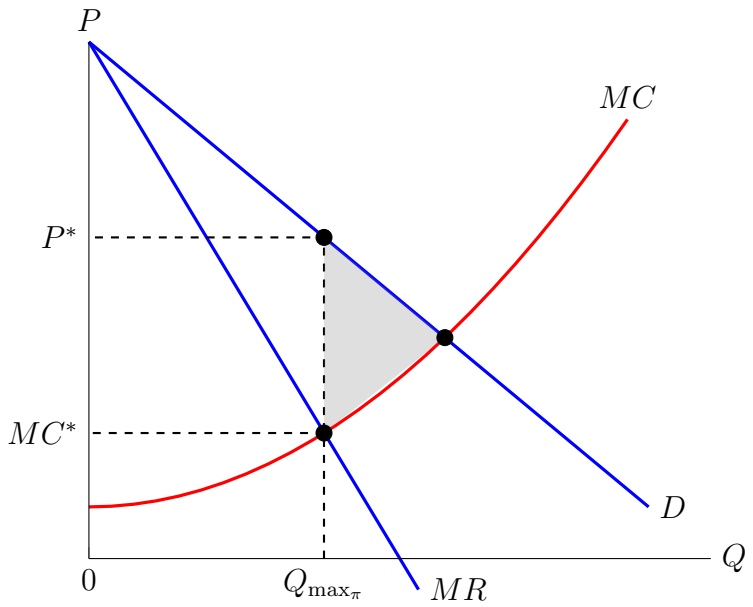


Figure 6.5: Monopolist and Efficiency Loss

Recall that Clark and others noted the importance of profit to lead us to this efficiency. When we consider the incentives facing producers, it turns out that the unrealistic assumptions of perfect competition such as an incredibly large number of firms and complete and perfect information are not necessary to avoid some of this inefficiency. Markets need not look like the perfect competition model to work well.

6.3.1 Cournot Model

The **Cournot Model of Oligopoly** suggest that firms choose quantities rather than prices when making production decisions and then market

clearing prices arise. Each firm has a reaction function—a curve mapping how many units it would want to produce based on the number of units its rival(s) produce. We can map the two-firm duopoly scenario, in the figure below with both firm 1 and firm 2 producing good x. At equilibrium, the reaction functions are equal: firm 1's choice of quantity is an optimal response to firm 2's choice in quantity and vice versa. For simplicity, we assume both firm 1 and 2 have the same cost function with a constant marginal cost.

For a simple linear demand curve:

$$P = a - bQ$$

with $Q = \sum q_i$ for all firms. So, the two profit-maximizing firms face

$$\max_{q_i}(Pq_i - MCq_i)$$

For $b = 1$ and a as the y-intercept of the Demand Curve—the lowest price for which $Q_D = 0$—then for the duopoly case:

$$P = a - q_1 - q_2$$

The profit for firm 1, then, is:

$$\pi_1 = Pq_1 - MCq_1 = (a - q_2 - q_1)q_1 - MCq_1$$

$$\pi_1 = (a - q_2 - q_1 - MC)q_1$$

Taking the first order condition:

$$\frac{\partial \pi}{\partial q_1} = (0-0-1)q_1+(a-q_2-q_1-MC) = -q_1+a-q_2-q_1-MC = a-q_2-MC-2q_1 \stackrel{set}{=} 0$$

$$2q_1 = a - q_2 - MC$$

So

$$q_1^*(q_2) = \frac{a - MC - q_2}{2}$$

Since firm 2 faces the same demand curve, the same math will result in

$$q_2^*(q_1) = \frac{a - MC - q_1}{2}$$

We can find the point of intersection between the reaction function using substitution:

$$q_1^*(q_2^*) = \frac{a - MC - \frac{a-MC-q_1}{2}}{2}$$

$$2q_1 = a - MC - \frac{a - MC - q_1}{2}$$
$$4q_1 = 2a - 2MC - a + MC + q_1$$
$$3q_1 = a - MC$$
$$q_1 = \frac{a - MC}{3}$$

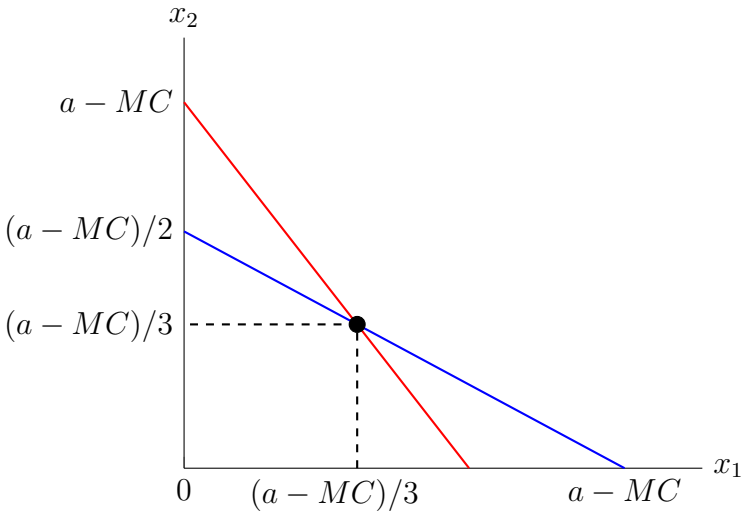


Figure 6.6: Cournot Duopoly Equilibrium

Thus, for the market as a whole:

$$Q = \sum q_i = 2 \cdot \left(\frac{a - MC}{3} \right)$$

$$P = a - q_1 - q_2 = a - \frac{2(a - MC)}{3} = \frac{a}{3} + \frac{2}{3}MC$$

However, since $P > MC$, this will induce more firms to enter the market to exploit this profit opportunity. Absent various barriers that prevent entry—whether it be heterogeneous firms with higher marginal costs than the existing firms or outright government prohibitions—such a duopoly will not be sustainable. If each of these new firms uses the same Cournot mechanism for quantity-setting, we can generalize the market result as:

$$Q = \sum_{i=1}^{N} q_i = \frac{N(a - MC)}{N + 1}$$

$$P = a - Q = \frac{a}{N+1} + \frac{N \cdot MC}{N+1}$$

Note, then, that

$$\lim_{n\to\infty} P = \lim_{n\to\infty} \frac{a}{N+1} + \frac{N \cdot MC}{N+1} = MC$$

and

$$\lim_{n\to\infty} Q = Q^*$$

Thus, as more firms enter, the lower the price goes, and the closer the price gets to the Marginal Cost. When we have an infinite number of firms entering, $P = MC$. This is the intuition of the Perfect Competition model. But even just a few firms gets us close to $P = MC$. Economic profit π remains positive—inducing new entry—until $P = MC$. If the Cournot Oligopoly degenerates into perfect competition, due to entry, we end up back at the conditions for the First Fundamental Theorem of Welfare Economics.

6.3.2 Bertrand Model

The **Bertrand Competition** model was developed in response to the Cournot Model. Also rooted in game theory, Bertrand theorized that we only need *two* firms to get the same outcome as a perfectly competitive market with price equal to marginal cost.

In the Bertrand Model:

- Firms set prices simultaneously
- If the firms set different prices, all customers buy from the lowest price firm.
- If firms set the same price, the customers randomly choose between the firms when buying the good.

In this set-up, if at least two firms can produce in the most productively efficient way—that is, more than one firm has the minimum marginal cost—the firm will be incentivized to set price equal to marginal cost. If one firm sets price greater than marginal cost, the other firm would have the incentive to set the price lower to undercut them and take all of the demand.

If, however, firms are not identical and do not face the same marginal cost curves, then we can expect a monopolist who takes command of the entire market by setting the price greater than his own marginal cost but less than his (potential) competitor's marginal cost:

$$MC_i < P < MC_j$$

Under this scenario, there are still some potential gains from trade, because $P > MC$; however, these are gains that are unique to this particular firm. Somehow, this firm has a unique cost function that is not replicable by other firms, which is what allows it to set its price *lower than* its potential competition. Thus, we need not even have another active competitor—merely a *potential* competitor to discipline market prices below the monopoly level.

Note that standard monopoly theory suggests that—so long as the marginal cost is greater than 0—the monopolist should always be operating on the elastic portion of the demand curve. Recall that the demand curve is unit elastic at the point $MR = 0$ and elastic when $MR > 0$. The monopolist is theorized to price goods where $MC = MR$, leading to a gap of inefficiency where units for which $MV > MC$—that is, an individual's willingness to pay is greater than the cost of production, but the marginal revenue ends up being less, because, by lowering the price of the good for everyone, revenue is reduced for all existing units sold. Yet, empirical evidence shows that industries from pharmaceuticals to energy and groceries generally appear to be priced on the inelastic demand of the demand curve. The Bertrand Model can provide one potential explanation for why this appears to be the case, even though the assumptions of perfect information and a near infinite number of small price-taking firms do not hold in these industries.

Baumol (1982) summarizes the general argument:

> A contestable monopoly offers us some presumption, but no guarantee, of behavior consistent with a second-best optimum, subject to the constraint that the firm be viable financially... [A] contestable monopoly has some reason to adopt the Ramsey optimal price-output vector, but it may have other choices open to it. But once each product obtains a second producer, that is, once we enter the domain of duopoly or oligopoly for each and every good, such choice disappears. The contestable oligopoly which achieves an equilibrium that immunizes it from the incursions of entrants has only one pricing option—it must set its price exactly *equal* to marginal cost and do *all* the things required for a first best optimum! In short, once we leave the world of pure or partial monopoly, any contestable market must behave ideally in every respect.

Baumol notes that markets that are more contestable—where entry is cheaper—will reach this point quicker. We need not have a perfectly competitive market, he argues—"firms need not be small or numerous or independent in their decision-making or produce homogeneous products."

What matters is that "there be no cost discrimination against entrants." A perfectly contestable market is one where entry and exit is entirely costless allowing for "hit-and-run entry" to take advantage of "even a very transient profit opportunity... before prices change, collect his gains, and then depart without cost, should the climate grow hostile."

6.3.3 The Coase Conjecture

Coase (1972) proposed that for a perfectly durable good, even *monopolists* for whom there is not even a potential second competitor will charge far less than the revenue-maximizing point on the demand curve. He gives the example of a monopolist land-owner. If a single person owned the entire American continent, but wanted to sell some land to make money—assuming the land is identical and too much to provide usefulness to the owner—he faces no marginal cost, so one would expect him to charge where $MR = 0$ at the unit-elastic, revenue-maximizing point on the demand curve.

However, after he makes these land sales, why should he continue to hold onto the rest of the land? It is also not providing him with value, so he would be incentivized to sell this land for a lower price. Coase suggests that buyers are smart enough to understand these incentives, and they know that, if the land owner expands supply, this will decrease the value of their land. Thus, he argues, that, far from facing a downward sloping demand curve typical for a monopolist, the demand schedule actually facing the landowner would be infinitely elastic at the competitive price. Coase suggests that the only way to charge a higher price would be through a contractual agreement not to sell land beyond the revenue-maximizing price or a contractually enforceable buy-back clause.

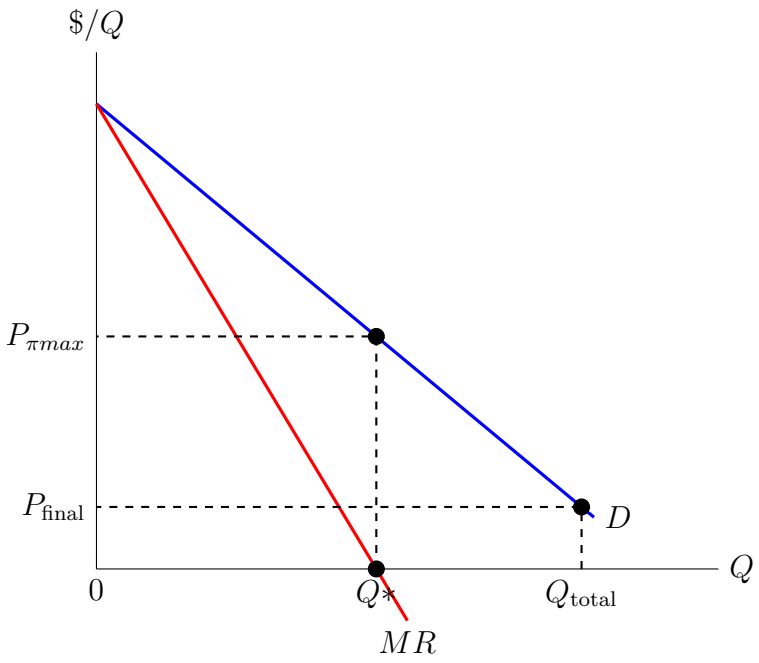


Figure 6.7: Coase (1972): Monopolist Land Owner

Coase notes that "if a less durable good is produced, a higher price can be charged, because consumers do not have to fear an increase in supply if they buy at the monopoly price." However, if costs are expected to lower over time, perhaps due to increased productivity, "the producer will have to consider the effect his actions will have on the expectations of consumers about his actions in future periods." Thus, the monopolist *does* have competition: not just with a hypothetical entrant, but with his future self.

6.4 Efficiency Distortions

While the conditions under which efficiency can be achieved are less constrained than the perfect competition model may appear upon first glance, there remain significant potential distortions that could prevent equality between $MRS^1 = MRS^2 = MV = MRT$ and $MRTS^A = MRTS^B$. Harberger's (1971) framework for conducting welfare analysis involves looking to the marginal values. He writes:

1. The competitive demand price for a given unit measures the value of that unit to the demander.
2. The competitive supply price for a given unit measures the value of

that unit to the supplier.

3. When evaluating the net benefits or costs of a given action (project, program, or policy), the costs and benefits accruing to each member of the relevant group (e.g., a nation) should normally be added without regard to the individual(s) to whom they accrue.

Practically speaking, this means looking at the height of the demand curve at the willingness-to-pay and height of the supply curve as the willingness-to-accept, and looking for gaps between them. This is where the Harberger Triangle or Deadweight Loss examined in the previous chapter comes into play. It represents an area of potential surplus, $WTP > WTA$ and yet, for various reasons, such trades both not take place.

Some of the distortions Harberger mentions include:

- Any taxes that are not *lump-sum* that create a division between the price paid by one agent and the price received by another agent whether income taxes (income received by employee $\neq$ income paid by employer), sales or consumption taxes (price paid by consumer $\neq$ price received by producer), traditional property taxes (where factor price ratio is not equivalent to the price received by factor owners, since real estate has an added cost).
- Monopoly profits due to restriction of sales where $P > MC$.
- Monopsony profits due to setting $P < MV$.
- Externalities in which the cost or benefit is not fully borne by the decision-making agents.

6.4.1 Externalities and Inefficiency

Externalities are the key argument for government intervention to remedy inefficiencies. For instance, public goods which are non-rival and non-excludable have an external benefit not captured by the decision-makers. Because once a public good is produced, it is not possible to exclude people from enjoying their benefit, there is a spillover value. Thus, the benefit is much greater than the individual benefit born by the person deciding whether or not to produce (or purchase) the good. We discussed this free rider problem in the context of property rights. Pollution is the example of a negative externality. The agent does not include the effect of noise or air pollution on the entire society around him when conducting a cost-benefit analysis and estimating the point at which $MC = MR$. As A.C. Pigou put it in his early twentieth century analysis, there is a Social Cost not included

in the Individual Cost—or, for a positive externality, a social benefit not included in the individual benefit curve.

Graphically, we can see that a positive externality results in the production of too few goods, because the price is less than the total benefit to all individuals, because the demand curve is aligned merely with *individual* benefits. See that the Total MB at the individual maximizing point Q_π is significantly greater than the MC at π (and the Price paid). On the other hand, a negative externality results in too much production of a good. See that at the individual maximizing point, the total cost is *greater* than the benefit. The gray area represents the inefficiencies. For a positive externality, it represents units for which $MB > MC$ and yet the competitive market does not produce them. For a negative externality, it represents units for which $MB < MC$ and yet the competitive market *does* produce them anyway. On the graph, the deadweight loss triangles point *toward* the optimum, pointing right with the positive externality (where the optimum is *more* production) and left with the negative externality (where the optimum is *less* production).

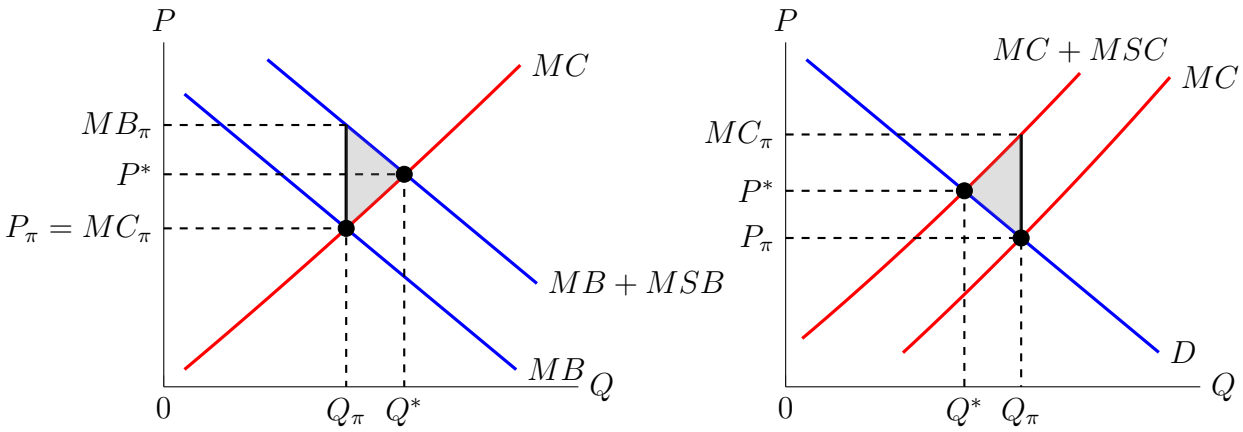


Figure 6.8: Positive Externality (left) and Negative Externality (right)

6.4.2 Resolving Externalities

Pigou proposed resolving externalities through a direct subsidization or taxation: subsidize the decision-maker facing a positive externality problem by the amount of the social benefit and tax the decision-maker facing a negative externality problem by the amount of the marginal social cost. However, while an elegant graphical solution, imposing such a tax would be difficult.

For Coase (1960), property rights allow for the internalization of externalities. In the presence of bargaining, these externalities would not exist. If the pollution resulted in a social cost, those who bear the social cost would be willing to pay up to the cost to prevent it. These payments for *not* producing certain units would result in a cost realized by the producer. If he produces these units, he does not receive these payments. For a positive externality, education is an oft-cited example. If there really is a large social benefit for having an educated society, the individuals across society who benefit from education of a given worker would be willing to pay him to get educated up to the social benefit, resulting in the would-be student facing a higher demand curve. In the presence of costless Coasean bargaining, the effective Pigouvian tax or subsidy would be realized in the form of a Coasean side payment, a negotiated voluntary transfer of resources in which one party pays the other to do (positive externality) or not do (negative externality) the action causing the externality. Albrecht, Hendrickson, and Salter write that a fundamental lesson of Coase (1960) is "that organizational costs, including any bargaining costs, are the sources of inefficient actions." The externality is not the *source* of the inefficiency, but rather transaction costs that prevent internalization of said externality.

Nye (2008) notes that Coase's findings actually lead to more problems for Pigouvian advocates, especially when it comes to Pigouvian taxes for negative externalities. While having the government act as a single bargaining unit in the face of high individual bargaining cost seems like a potential solution, estimating the actual magnitude of the problem is nearly impossible. Unless we can fully account for any existing Coasean transfers that already take place, a tax (or subsidy) based on the size of the externality (i.e., the social cost or benefit) itself, will be overstated. Likewise, the indirect effects of regulations and distortionary taxes—even if not designed to reduce the externality itself—will result in a lower production of a negative externality than the competitive market equilibrium assumed by Pigou. Finally, market imperfections like monopoly can also already lead to a lower production and higher cost for the good. Nye cites the OPEC market power over oil production must be factored in when estimating a Pigouvian tax on gas. In addition, it is critical to note that, even at the optimum, there is *some* production. In the graph above, the producer should not produce none of the negative externality, but rather only up to the point where $MC = MB$. Thus, Nye writes,

> Simply observing that cars continue to pollute or that rush hour traffic is becoming more congested are not sufficient grounds for an economist to argue for higher taxes, even if we obtained a

> precise measure of the pollution or congestion externalities. In all cases, advocating a Pigou tax requires that we know the relationship of existing Q to optimal Q^*, and not the size of the externality.

6.4.3 Externality, Transaction Cost, and Efficiency

Demsetz (2011) cuts deeper to the heart of externality analysis. Many have suggested that "the cost of transacting might be so high as to block the negotiation [for Coasean bargaining and] in this case, it appears that the economic system has failed to put this resource to its highest value use... [and] fail[ed] to allocate resources efficiently." But Demsetz argues that this "appearance of inefficiency is but an illusion." Several of the real-world cases cited by Coase involve the allocation of rights to parties by the court system. "We would not accuse the market of inefficiency if the government assigned a right to parties that put it to less valuable uses than would others. Why accuse the market of inefficiency if the court does the same?" The common law courts do not depend on profit—in fact, it is illegal for judges to take bribes. "Although there may be very good reasons for not creating a market-like legal system, we may note that if we suppose that courts are remade into market institutions whose survival depends on revenues secured from petitioners who purchase their services and decisions, ownership of a disputed resource would never go to the petitioner who is less capable or maximizing value from its use. Coase's imagined court decision would never have been made." The failure in these cases lay with government and not with markets. For Demsetz, "an efficient economic system is one that makes the most of scarce resources within whatever constraints are handed down to it by courts and legislatures." Thus, we cannot ignore the institutional environment within which the market operates.

However, Demsetz concludes by chastising economists for separating transaction costs from efficiency analysis arguing that transaction costs are no less real than any other cost.

> the cost of transacting may prevent some of these opportunities from being brought to a resource owner's attention by way of negotiated offers. However, those values that are not known will be only those for which the cost of acquiring price information is expected to exceed the value of the knowledge that is expected. to be obtained from this information. ... Put differently, there exists an efficient amount of ignorance in an economic system if the cost of acquiring information is positive. ... Ignorance not only may

> be bliss, it also may be efficient. One cannot claim that resources are wrongly allocated simply because information is not possessed or negotiation is absent; nor can one claim that resources are misplaced because a specific market does not exist. ... There is no difference between transaction cost and other costs... Indeed, one can rewrite the neoclassical model with transaction cost embedded in it and still deduce from it an efficient allocation of resources. Transaction cost just shifts supply curves upward (or demand curves downward, or some combination of both) as would an increase in any cost, and it carries no special implication of inefficiency at equilibrium values of price and output.

Returning to Leeson (2019), when we assume agents to be constrained maximizers, we assume they are maximizers. The lack of bargaining does not—and cannot—disprove this assumption. Under the maximization assumption, the lack of bargaining actually shows the opposite. It shows that the perceived costs of bargaining or of actually making a Coasean side-payment that would be acceptable to the other party is *higher* than the expected benefits of said bargaining. We explain the lack of Coasean bargaining the same way we explain other transactions that do not take place: the perceived cost must be higher than the perceived benefit. This is not to say agents can never make mistakes, overestimate costs, or underestimate benefits. However, if agents *are* misperceiving costs and benefits, this leads to a profit opportunity for a middleman to reduce transaction costs between the parties by undertaking negotiations for them and facilitating such a side-payment. If such a profit opportunity does not exist, then, by definition, there is no inefficient production or lack of efficient production taking place.

Perhaps one may argue that agent's underlying preferences are misplaced. For instance, some have argued that the preservation of our ecosystem is worth significantly higher costs than individuals are willing to pay. However, we must be clear that this goes beyond the realm of economic science, which does not judge individuals preferences. Instead, this paternalistic view that people do not really understand what is good for them is a normative philosophical judgment, not a statement of economic science. Whether people should have different preferences or different willingness-to-pay schedules has nothing to do with efficiency and the measures of whether people do the best to satisfy their desires given the resources that they have.

6.5 Distribution & the Second Fundamental Theorem of Welfare Economics

The Second Fundamental Theorem of Welfare Economics is that, maintaining the assumptions of the First Fundamental Theorem, any Pareto-efficient allocation can be achieved as a competitive market equilibrium after lump-sum transfers.

Consider the following Edgeworth Box:

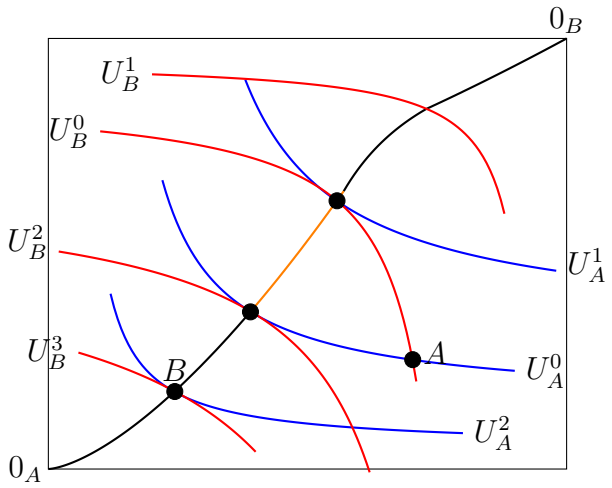


Figure 6.9: Second Fundamental Theorem of Welfare Economics

The contract curve shows all possible Pareto Efficient scenarios. For initial position A (and thus initial indifference curves U_A^0 and U_B^0), the orange segment of the contract curve shows all possible allocations that market trading could reach. However, the Second Theorem says that the entire contract curve is actually feasible *if* we allow for transfers. Now, the transfers that move us outside this orange segment would no longer be Pareto Improvements. For instance, B is a possible place if redistribution moved this point to be within the "eye" of the two curves—that is, if the new orange segment includes B, B is a possible solution. But B would make individual A *worse* off, putting him on a lower indifference curve, so this is not a Pareto Improvement. Nonetheless, the Second Theorem suggests that our final outcome after distribution remains plausibly a Pareto Efficient scenario. After all, at point B, neither individual could be made better off without making the other worse off. This is true for every point along the contract curve—the points of intersection between the indifference curves.

A different way of thinking about the Second Fundamental Theorem is that any Pareto efficient allocation of goods—any point along the contract curve—can be a general equilibrium. That means that any desired efficient allocation can be reached by merely redistributing income and then allowing competitive markets to work exploiting gains from trade until we reach the Pareto efficient equalities discussed with the First Fundamental Theorem.

6.6 Takeaways of This Study

Economics assumes that individual agents are optimizers. Individuals seek to maximize the fulfillment of their desires given the resources they have. As consumers, they desire more rather than less and seek to fulfill their preferences. As producers, they seek to minimize their costs and maximize their profits. Across history, institutions have evolved to facilitate market transactions. The societies that have flourished are those which are incentive-compatible: those for which everyone has the incentive to put resources to their highest valued use. Gains from trade are potentially enormous, so individuals in wealthy societies do not try to be self-sufficient. Rather, they take advantage of specialization and comparative advantage.

When there is a free market with property rights which allow individuals to internalize profit and loss of their decisions regarding that to which they have property rights, we approach the optimum which is: all goods which are valued higher than the next best alternative with the input resources are produced—put differently, if $MB > MC$, the good is produced, and if $MB < MC$, the good is not produced. Individuals do not have perfect knowledge or foresight. The market, however, serves to communicate each individual's imperfect knowledge, and the profit/loss function serves to ensure that the successful firms that satisfy consumer desires survive, while those that destroy value die. Bargaining is not costless and using the price mechanism results in transaction costs, but, due to the massive gains from trade, there are incentives to innovate to reduce such costs, whether this is through the adoption of a monetary medium of exchange or as merchants and facilitators of such transactions. This innovation leads to the expansion of markets and widening availability of all kinds of goods.

Circumstances are constantly changing. New knowledge, innovation, and ideas as well as "acts of God" affecting actual resource availability result in an equilibrium position that is always on the move. As constraints change, individuals, even with constant preferences, will eventually adjust their behavior. As the number of producers of each good becomes extremely large and as each buyer has a small market share, we gain

mathematical certainty that the markets will produce the optimum; however, we need not have these conditions to ensure this is the case. Because knowledge is imperfect, entrepreneurship is possible and necessary. Firm rivalry and market competition to satisfy consumer desires is what ensures that profit opportunities—though they may go unnoticed for a time—are being discovered and exploited. Yet such rivalry—with the protection of individual property rights—also serves as a safeguard against the exploitation of individuals.

The miracle of economic prosperity is that all this happens organically. The free market is a spontaneous order. There is no designer and no dictator telling people what and how much to produce. Rather, the system functions better than a micro-managed economy could, in large part because all of the relevant details are *subjective*. It is the individuals' own desires, preferences, and subjective valuations and estimations both as consumers and producers that create supply and demand. Cost and value are subjective. The magic of the market is to turn these subjective valuations from hypothetical utility into objective market prices allowing for the calculation of relative value and thus, profit and loss.

The twentieth century showed that attempts to centrally plan an economy—when the state sought to serve as the mythical "Walrasian auctioneer"—create economic chaos, mass starvation, and poverty. As the former members of the Soviet Union have liberalized and as China has emerged from its Maoist Era into a freer—though still far from free—society, billions across the world have been brought out of poverty as they have joined systems where production and consumption is determined by market forces, by the preferences of individuals and the ability of individuals and firms to satisfy those desires. But history is so easily forgotten, and this is why its study and the study of the economic way of thinking as a method for interpreting history is so critical. Continued liberalization will be key to maintaining prosperity, but it is far from guaranteed. We must not mistake the status quo for the free market outcome. Significant government interventions obscure the market realities leading to many of the populist complaints in recent years. For instance, so long as the government maintains a stranglehold on the healthcare industry through regulation and subsidization and prohibits individuals from using their private property to build housing, we can expect industries that satisfy no one and appear to be at a disequilibrium. But this also means there is an opportunity for policy entrepreneurs to identify and communicate these challenges and to show that the solution to these failures is not more control but more freedom.

We assume that agents maximize given their constraints, yet this need not

mean we take the Panglossian perspective of the world. Improvements are possible. Profit opportunities do exist. But these opportunities must be sought after. If they were well-known, they would have been taken advantage of. If a central authority knew of them, it could share this knowledge with the broader society and the market forces would validate whether or not such information is correct. Equilibrium Theory should give us humility to realize the vast expansive set of things we do not know and do not—and cannot—control. At the same time, it should encourage us to look for areas where there is deadweight loss or an unidentified opportunity. May your study of Economics become worthwhile because you take action in light of it.

References

Albrecht, Brian C., Joshua R. Hendrickson, and Alexander William Salter. 2022. "Evolution, Uncertainty, and the Asymptotic Efficiency of Policy," *Public Choice*, Vol. 192, pp. 169-188.

Alchian, Armen A and Harold Demsetz. 1972. "Production, Information Costs, and Economic Organization," *American Economic Review*, Vol. 62, No. 5, pp. 777-795.

Arrow, Kenneth J. and Gerard Debreu. 1954. "Existence of an Equilibrium for a Competitive Economy," *Econometrica*, Vol. 22, No. 3, pp. 265-290.

Baumol, William J. 1982. "Contestable Markets: An Uprising in the Theory of Industry Structure," *American Economic Review*, Vol. 72, No. 1, pp. 1-15.

Baumol, William J. 1999. "Retrospectives: Say's Law," *Journal of Economic Perspectives*, Vol. 13, No. 1, pp. 194-204.

Clark, John Bates. 1908. "Static Standards in a Dynamic Society," *The Distribution of Wealth: A Theory of Wages, Interest, and Profits*, The MacMillan Company, pp. 399-430.

Coase, R.H. 1960. "The Problem of Social Cost," *Journal of Law and Economics*, Vol. 3, pp. 1-44.

Coase, Ronald. 1972. "Durability and Monopoly," *Journal of Law and Economics*, Vol. 15, No. 1, pp. 143-149.

Greenwald, Bruce C. and Joseph E. Stiglitz. 1986. "Externalities in Economies with Imperfect Information and Incomplete Markets," *Quarterly Journal of Economics*, Vol. 101, No. 2, pp. 229-264.

Debreu, Gerard. 1959. *Theory of Value: An Axiomatic Analysis of Economic Equilibrium*, Yale University Press.

Demsetz, Harold. 1969. "Information and Efficiency: Another Viewpoint," *Journal of Law & Economics*, Vol. 12, No. 1, pp. 1-22.

Demsetz, Harold. 2011. "The Problem of Social Cost: What Problem?: A Critique of the Reasoning of A.C. Pigou and R.H. Coase," *Review of Law & Economics*, Vol. 7, No. 1, pp. 1-13.

Harberger, Arnold C. 1954. "Monopoly and Resource Allocation," *American Economic Review*, Vol. 44, No. 2, pp. 77-87.

Harberger, Arnold C. 1971. "Three Basic Postulates for Applied Welfare Economics: An Interpretive Essay," *Journal of Economic Literature*, Vol. 9, No. 3, pp. 785-797.

Lange, Oscar. 1942. "The Foundations of Welfare Economics," *Econometrica*, Vol. 10, No. 3, pp. 215-228.

Leeson, Peter T. 2019. "Logic is a Harsh Mistress: Welfare Economics for Economists," *Journal of Institutional Economics*, Vol. 16, No. 2, pp. 145-150.

McKenzie, Lionel W. 1981. "The Classical Theorem on Existence of Competitive Equilibrium," *Econometrica*, Vol. 49, No. 4, pp. 819-841.

Nye, John V.C. 2008. "The Pigou Problem," *Regulation*, Vol. 32, No. 2, pp. 32-37.

Jones, Ronald W. 1965. "The Structure of Simple General Equilibrium Models," *Journal of Political Economy*, Vol. 73, No. 6, pp. 557-572.

Say, Jean-Baptiste. 1855 [1803]. *A Treatise on Political Economy; or the Production, Distribution, and Consumption of Wealth*, edited by Clement C. Biddle, translated by C.R. Prinsep, Lippincott, Grambo, & Co.

Stiglitz, Joseph E. 2002. "Information and the Change in the Paradigm in Economics," *American Economic Review*, Vol. 92, No. 3, pp. 460-501.

Wicksell, Knut. 1977 [1934]. *Lectures on Political Economy Volume I: General Theory*, translated by E. Classen, edited by Lionel Robbins, Augustus M. Kelley.

Yeager, Leland B. 1956. "A Cash-Balance Interpretation of Depression," *Southern Economic Journal*, Vol. 22, No. 4, pp. 438-447.

Bibliography

In addition to the specific journal articles cited in each chapter, the following textbooks were used as resources in constructing this text.

1. **Alchian, Armen A. and William R. Allen**. (2018) *Universal Economics*, edited by Jerry L. Jordan. Liberty Fund.
2. **Becker, Gary S**. (1971) *Economic Theory*. Alfred A. Knopf.
3. **Friedman, Milton**. (1962) *Price Theory: A Provisional Text*. Aldine Publishing
4. **Hirshleifer, Jack, Amihai Glazer, and David Hirshleifer**. (2005) *Price Theory and Applications*. Seventh Edition. Cambridge University Press.
5. **Jaffe, Sonia, Robert Minton, Casey B. Mulligan, and Kevin M. Murphy**. (2019) *Chicago Price Theory*. Princeton University Press.
6. **Jehle, Geoffrey A. and Philip J. Reny**. (2011) *Microeconomic Theory*. Pearson Education.
7. **Layard, P.R.G. and A.A. Walters**. (1976) *Microeconomic Theory*. McGraw Hill.
8. **McCloskey, Donald N**. (1985) *Applied Theory of Price*. Second Edition. MacMillan Publishing Company.
9. **Stigler, George J**. (1987) *The Theory of Price*. Fourth Edition. MacMillan Publishing Company.
10. **Varian, Hal R**. (1992) *Microeconomic Analysis*. Third Edition. W.W. Norton & Company.