2 Supply and Demand

Supply and demand are the most fundamental tools of economic analysis. Most applications of economic reasoning involve supply and demand in one form or another. When prices for home heating oil rise in the winter, usually the reason is that the weather is colder than normal and as a result, demand is higher than usual. Similarly, a break in an oil pipeline creates a short-lived gasoline shortage, as occurred in the Midwest in the year 2000, which is a reduction in supply. The price of DRAM, or dynamic random access memory, used in personal computers falls when new manufacturing facilities begin production, increasing the supply of memory.

This chapter sets out the basics of supply and demand, introduces equilibrium analysis, and considers some of the factors that influence supply and demand and the effects of those factors. In addition, quantification is introduced in the form of elasticities. Dynamics are not considered, however, until Chapter 4, which focuses on production, and Chapter 5 introduces a more fundamental analysis of demand, including a variety of topics such as risk. In essence, this is the economics "quickstart" guide, and we will look more deeply in the subsequent chapters.

2.1 Supply and Demand

2.1.1 Demand and Consumer Surplus

Eating a French fry makes most people a little bit happier, and we are willing to give up something of value – a small amount of money, a little bit of time – to eat one. What we are willing to give up measures the value – our personal value – of the French fry. That value, expressed in dollars, is the *willingness to pay* for French fries. That is, if you are willing to give up three cents for a single French fry, your willingness to pay is three cents. If you pay a penny for the French fry, you've obtained a net of two cents in value. Those two cents – the difference between your willingness to pay and the amount you do pay – is known as *consumer surplus*. Consumer surplus is the value to a consumer of consumption of a good, minus the price paid.

The value of items – French fries, eyeglasses, violins – is not necessarily close to what one has to pay for them. For people with bad vision, eyeglasses might be worth ten thousand dollars or more, in the sense that if eyeglasses and contacts cost $10,000 at all stores, that is what one would be willing to pay for vision correction. That one doesn't have to pay nearly that amount means that the consumer surplus associated with eyeglasses is enormous. Similarly, an order of French fries might be worth $3 to a consumer, but because French fries are available for around $1, the consumer obtains a surplus of $2 in the purchase.

How much is a second order of French fries worth? For most of us, that first order is worth more than the second one. If a second order is worth $2, we would still gain from buying it. Eating a third order of fries is worth less still, and at some point we're unable or unwilling to eat any more fries even when they are free, which implies that at some point the value of additional French fries is zero.
We will measure consumption generally as units per period of time, e.g. French fries consumed per month.

Many, but not all, goods have this feature of *diminishing marginal value* – the value of the last unit consumed declines as the number consumed rises. If we consume a quantity q, it implies the marginal value $v(q)$ falls as the number of units rise.[1] An example is illustrated in Figure 2-1. Here the value is a straight line, declining in the number of units.

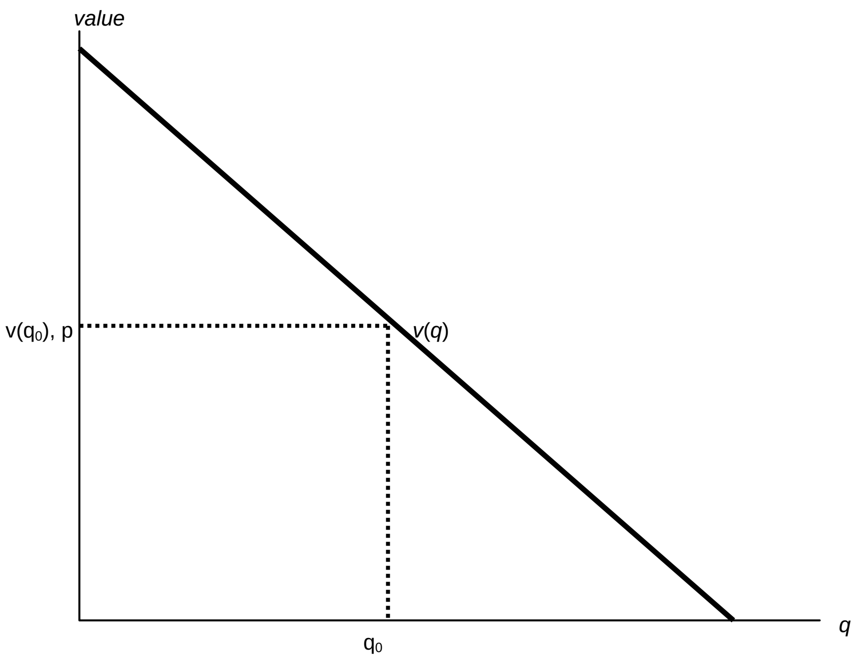


Figure 2-1: The Demand Curve

Demand need not be a straight line, and indeed could be any downward-sloping curve. Contrary to the usual convention, demand gives the quantity chosen for any given price off the horizontal axis, that is, given the value p on the vertical axis, the corresponding value q_0 on the horizontal axis is the quantity the consumer will purchase.

It is often important to distinguish the demand curve itself – the entire relationship between price and quantity demanded – from the quantity demanded. Typically, "demand" refers to the entire curve, while "quantity demanded" is a point on the curve.

Given a price p, a consumer will buy those units with $v(q)>p$, since those units are worth more than they cost. Similarly, a consumer should not buy units for which $v(q)<p$. Thus, the quantity q_0 that solves the equation $v(q_0)=p$ gives the quantity of units the consumer will buy. This value is also illustrated in Figure 2-1.[2] Another way of

[1] When diminishing marginal value fails, which sometimes is said to occur with beer consumption, constructing demand takes some additional effort, which isn't of a great deal of consequence. Buyers will still choose to buy a quantity where marginal value is decreasing.

[2] We will treat units as continuous, even though in reality they are discrete units. The reason for treating them as continuous is only to simplify the mathematics; with discrete units, the consumer buys those units with value exceeding the price, and doesn't buy those with value less than the price, just as before. However, since the value function isn't continuous, much less differentiable, it would be an accident for

summarizing this insight is that the *marginal value* curve is the inverse of demand function, where the demand function gives the quantity demanded for any given price. Formally, if $x(p)$ is the quantity a consumer buys given a price of p, then $v(x(p)) = p$.

But what is the marginal value curve? Suppose the total value of consumption of the product, in dollar value, is given by $u(q)$. That is, a consumer who pays $u(q)$ for the quantity q is just indifferent to getting nothing and paying nothing. For each quantity, there should exist one and only one price that exactly makes the consumer indifferent between purchasing it and getting nothing at all, because if the consumer is just willing to pay $u(q)$, any greater amount is more than the consumer should be willing to pay.

The consumer facing a price p gets a net value or consumer surplus of $CS = u(q) - pq$ from consuming q units. In order to obtain the maximal benefit, the consumer would then choose the level of q to maximize $u(q) - pq$. When the function CS is maximized, its derivative is zero, which implies that, at the quantity that maximizes the consumer's net value

$$0 = \frac{d}{dq}\big(u(q) - pq\big) = u'(q) - p.$$

Thus we see that $v(q) = u'(q)$, that is, the *marginal value* of the good is the derivative of the total value.

Consumer surplus is the value of the consumption minus the amount paid, and represents the net value of the purchase to the consumer. Formally, it is $u(q)$-pq. A graphical form of the consumer surplus is generated by the following identity.

$$CS = \max_q \big(u(q) - pq\big) = u(q_0) - pq_0 = \int_0^{q_0} \big(u'(x) - p\big)dx = \int_0^{q_0} \big(v(x) - p\big)dx.$$

This expression shows that consumer surplus can be represented as the area below the demand curve and above the price, as is illustrated in Figure 2-2. The consumer surplus represents the consumer's gains from trade, the value of consumption to the consumer net of the price paid.

---

marginal value to equal price. It isn't particularly arduous to handle discreteness of the products, but it doesn't lead to any significant insight either, so we won't consider it here.

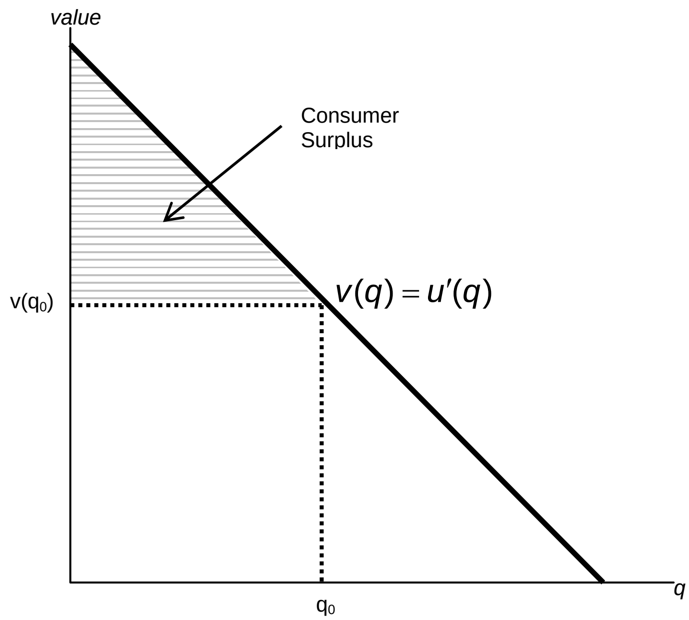


Figure 2-2: Consumer Surplus

The consumer surplus can also be expressed using the demand curve, by integrating from the price up. In this case, if $x(p)$ is the demand, we have

$$CS = \int_{p}^{\infty} x(y)\,dy \,.$$

When you buy your first car, you experience an increase in demand for gasoline because gasoline is pretty useful for cars and not so much for other things. An imminent hurricane increases the demand for plywood (to protect windows), batteries, candles, and bottled water. An increase in demand is represented by a movement of the entire curve to the northeast (up and to the right), which represents an increase in the marginal value v (movement up) for any given unit, or an increase in the number of units demanded for any given price (movement to the right). Figure 2-3 illustrates a shift in demand.

Similarly, the reverse movement represents a decrease in demand. The beauty of the connection between demand and marginal value is that an increase in demand could in principle have meant either more units demanded at a given price, or a higher willingness to pay for each unit, but those are in fact the same concept – both create a movement up and to the right.

For many goods, an increase in income increases the demand for the good. Porsche automobiles, yachts, and Beverly Hills homes are mostly purchased by people with high incomes. Few billionaires ride the bus. Economists aptly named goods whose demand doesn't increase with income *inferior* goods, with the idea that people substitute to better quality, more expensive goods as their incomes rise. When demand for a good increases with income, the good is called *normal*. It would have been better to call such goods superior, but it is too late to change such a widely accepted convention.

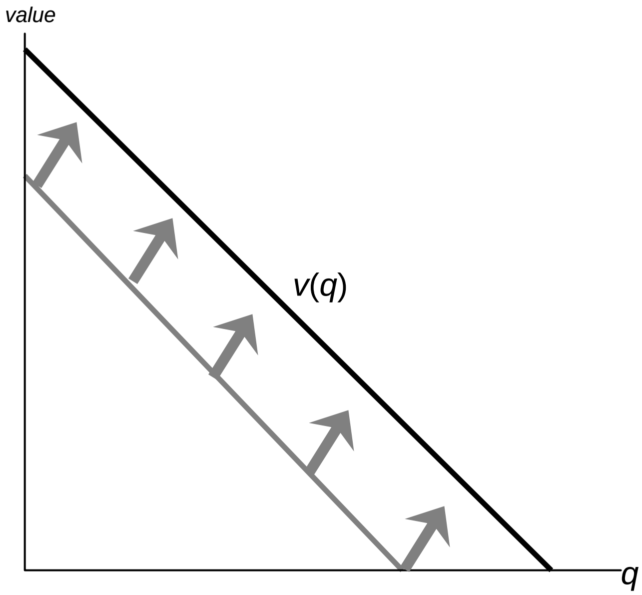


Figure 2-3: An Increase in Demand

Another factor that influences demand is the price of related goods. The dramatic fall in the price of computers over the past twenty years has significantly increased the demand for printers, monitors and internet access. Such goods are examples of *complements*. Formally, for a given good X, a complement is a good whose consumption increases the value of X. Thus, the use of computers increases the value of peripheral devices like printers and monitors. The consumption of coffee increases the demand for cream for many people. Spaghetti and tomato sauce, national parks and hiking boots, air travel and hotel rooms, tables and chairs, movies and popcorn, bathing suits and sun tan lotion, candy and dentistry are all examples of complements for most people – consumption of one increases the value of the other. The complementarity relationship is symmetric – if consumption of X increases the value of Y, then consumption of Y must increase the value of X.[3] There are many complementary goods and changes in the prices of complementary goods have predictable effects on the demand of their complements. Such predictable effects represent the heart of economic analysis.

The opposite case of a complement is a *substitute*. Colas and root beer are substitutes, and a fall in the price of root beer (resulting in an increase in the consumption of root beer) will tend to decrease the demand for colas. Pasta and ramen, computers and typewriters, movies (in theaters) and sporting events, restaurants and dining at home, spring break in Florida versus spring break in Mexico, marijuana and beer, economics

---

[3] The basis for this insight can be seen by denoting the total value in dollars of consuming goods x and y as $u(x, y)$. Then the demand for x is given by the partial $\partial u / \partial x$. The statement that y is a complement is the statement that the demand for x rises as y increases, that is, $\partial^2 u / \partial x \partial y > 0$. But then with a continuous second derivative, $\partial^2 u / \partial y \partial x > 0$, which means the demand for y, $\partial u / \partial y$, increases with x.

courses and psychology courses, driving and bicycling are all examples of substitutes for most people. An increase in the price of a substitute *increases* the demand for a good, and conversely, a decrease in the price of a substitute decreases demand for a good. Thus, increased enforcement of the drug laws, which tends to increase the price of marijuana, leads to an increase in the demand for beer.

Much of demand is merely idiosyncratic to the individual – some people like plaids, some like solid colors. People like what they like. Often people are influenced by others – tattoos are increasingly common not because the price has fallen but because of an increased acceptance of body art. Popular clothing styles change, not because of income and prices but for other reasons. While there has been a modest attempt to link clothing style popularity to economic factors,[4] by and large there is no coherent theory determining fads and fashions beyond the observation that change is inevitable. As a result, this course, and economics more generally, will accept preferences for what they are without questioning why people like what they like. While it may be interesting to understand the increasing social acceptance of tattoos, it is beyond the scope of this text and indeed beyond most, but not all, economic analyses. We will, however, account for some of the effects of the increasing acceptance of tattoos through changes in the number of firms offering tattooing, changes in the variety of products offered, and so on.

2.1.1.1 (Exercise) A *reservation price* is the maximum willingness to pay for a good that most people buy one unit of, like cars or computers. Graph the demand curve for a consumer with a reservation price of $30 for a unit of a good.

2.1.1.2 (Exercise) Suppose the demand curve is given by $x(p) = 1 - p$. The consumer's expenditure is $px(p) = p(1 - p)$. Graph the expenditure. What price maximizes the consumer's expenditure?

2.1.1.3 (Exercise) For demand $x(p) = 1 - p$, compute the consumer surplus function as a function of p.

2.1.1.4 (Exercise) For demand $x(p) = p^{-\varepsilon}$, for $\varepsilon > 1$, find the consumer surplus as a function of p. (Hint: recall that the consumer surplus can be expressed as $CS = \int_p^\infty x(y)\,dy$.)

2.1.2 Supply

The supply curve gives the number of units, represented on the horizontal axis, as a function of the price on the vertical axis, which will be supplied for sale to the market. An example is illustrated in Figure 2-4. Generally supply is upward-sloping, because if it is a good deal for a seller to sell 50 units of a product at a price of $10, then it remains a good deal to supply those same 50 at a price of $11. The seller might choose to sell

[4] Skirts are allegedly shorter during economic booms and lengthen during recessions.

more than 50, but if the first 50 weren't worth keeping at a price of $10, that remains true at $11.[5]

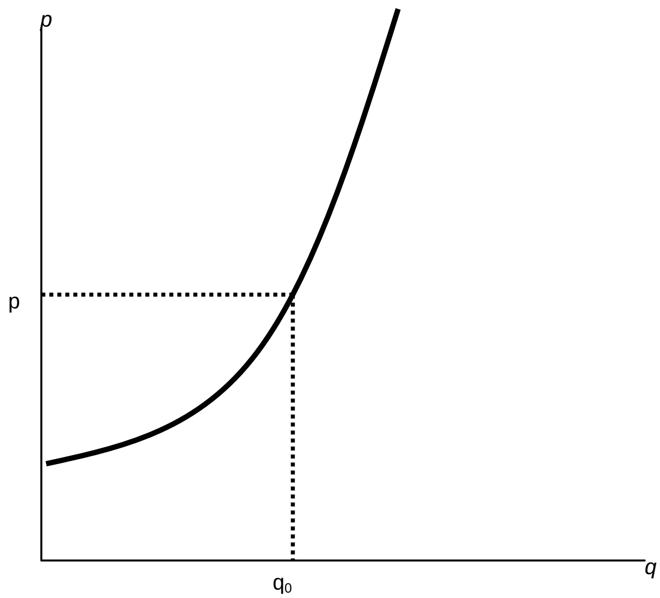


Figure 2-4: The Supply Curve

The seller who has a cost $c(q)$ for selling q units obtains a profit, at price p per unit, of $pq - c(q)$. The quantity which maximizes profit for the seller is the quantity q^* satisfying

$$0 = \frac{d}{dq} pq - c(q) = p - c'(q^*).$$

Thus, price equals marginal cost is a characteristic of profit maximization; the seller sells all the units whose cost is less than price, and doesn't sell the units whose cost exceeds price. In constructing the demand curve, we saw that the demand curve was the inverse of the marginal value. There is an analogous property of supply: *the supply curve is the inverse function of marginal cost.* Graphed with the quantity supplied on the horizontal axis and price on the vertical axis, the supply curve is the marginal cost curve, with marginal cost on the vertical axis.

Exactly in parallel to consumer surplus with demand, profit is given by the difference of the price and marginal cost

[5] This is a good point to remind the reader that the economists' familiar assumption of "other things equal" is still in effect. If the increased price is an indication that prices might rise still further, or a consequence of some other change that affects the sellers' value of items, then of course the higher price might not justify sale of the items. We hold other things equal to focus on the effects of price alone, and then will consider other changes separately. The pure effect of an increased price should be to increase the quantity offered, while the effect of increased expectations may be to decrease the quantity offered.

$$\text{Profit} = \max_q pq - c(q) = pq^* - c(q^*) = \int_0^{q^*} \left(p - c'(x)\right) dx.$$

This area is shaded in Figure 2-5.

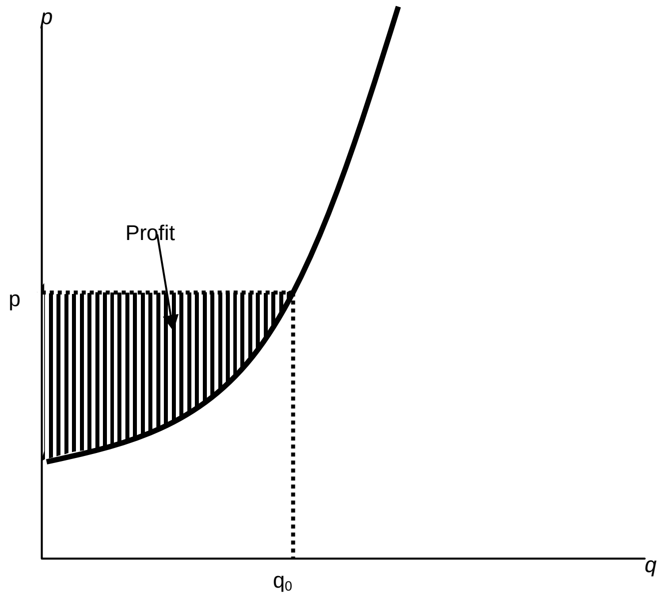


Figure 2-5: Supplier Profits

The relationship of demand and marginal value exactly parallels the relationship of supply and marginal cost, for a somewhat hidden reason. Supply is just negative demand, that is, a supplier is just the possessor of a good who doesn't keep it but instead offers it to the market for sale. For example, when the price of housing goes up, one of the ways people demand less is by offering to rent a room in their house, that is, by supplying some of their housing to the market. Similarly, the marginal cost of supplying a good already produced is the loss of not having the good, that is, the marginal value of the good. Thus, with exchange, it is possible to provide the theory of supply and demand entirely as a theory of *net* demand, where sellers are negative demanders. There is some mathematical economy in this approach, and it fits certain circumstances better than separating supply and demand. For example, when the price of electricity rose very high in the western United States in 2003, several aluminum smelters resold electricity they had purchased in long-term contracts, that is, demanders became suppliers.

However, the "net demand" approach obscures the likely outcomes in instances where the sellers are mostly different people, or companies, than the buyers. Moreover, while there is a theory of complements and substitutes for supply that is exactly parallel to the equivalent theory for demand, the nature of these complements and substitutes tends to be different. For these reasons, and also for the purpose of being consistent with common economic usage, we will distinguish supply and demand.

An *increase in supply* refers to either more units available at a given price, or a lower price for the supply of the same number of units. Thus, an increase in supply is graphically represented by a curve that is lower or to the right, or both, that is, to the south-east. This is illustrated in Figure 2-6. A decrease in supply is the reverse case, a shift to the northwest.

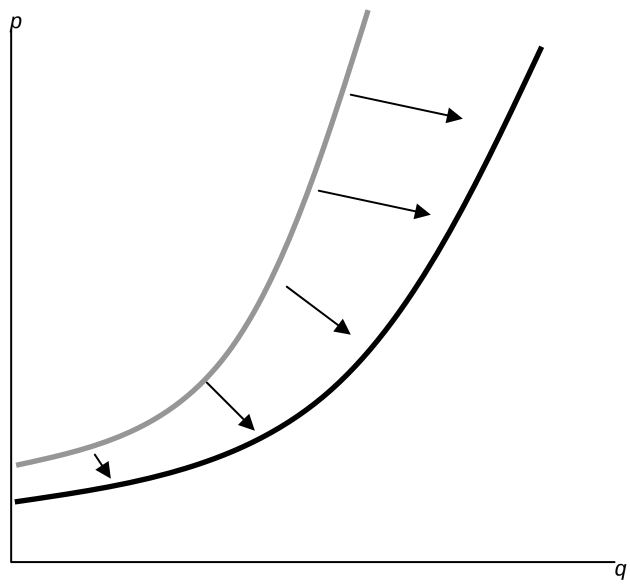


Figure 2-6: An Increase in Supply

Anything that increases costs of production will tend to increase marginal cost and thus reduce the supply. For example, as wages rise, the supply of goods and services is reduced, because wages are the input price of labor. Labor accounts for about two-thirds of all input costs, and thus wage increases create supply reductions (a higher price is necessary to provide the same quantity) for most goods and services. Costs of materials of course increase the price of goods using those materials. For example, the most important input into the manufacture of gasoline is crude oil, and an increase of $1 in the price of a 42 gallon barrel of oil increases the price of gasoline about two cents – almost one-for-one by volume. Another significant input in many industries is capital, and as we will see, interest is cost of capital. Thus, increases in interest rates increase the cost of production, and thus tend to decrease the supply of goods.

Parallel to complements in demand, a *complement in supply* to a good X is a good Y such that an increase in the price of Y increases the supply of X. Complements in supply are usually goods that are jointly produced. In producing lumber (sawn boards), a large quantity of wood chips and sawdust are also produced as a by-product. These wood chips and saw dust are useful in the manufacture of paper. An increase in the price of lumber tends to increase the quantity of trees sawn into boards, thereby increasing the supply of wood chips. Thus, lumber and wood chips are complements in supply.

It turns out that copper and gold are often found in the same kinds of rock – the conditions that give rise to gold compounds also give rise to copper compounds. Thus, an increase in the price of gold tends to increase the number of people prospecting for gold, and in the process increases not just the quantity of gold supplied to the market, but also the quantity of copper. Thus, copper and gold are complements in supply.

The classic supply-complement is beef and leather – an increase in the price of beef increases the slaughter of cows, thereby increasing the supply of leather.

The opposite of a complement in supply is a *substitute in supply*. Military and civilian aircraft are substitutes in supply – an increase in the price of military aircraft will tend to divert resources used in the manufacture of aircraft toward military aircraft and away from civilian aircraft, thus reducing the supply of civilian aircraft. Wheat and corn are also substitutes in supply. An increase in the price of wheat will lead farmers whose land is reasonably well-suited to producing either wheat or corn to substitute wheat for corn, increasing the quantity of wheat and decreasing the quantity of corn. Agricultural goods grown on the same type of land usually are substitutes. Similarly, cars and trucks, tables and desks, sweaters and sweatshirts, horror movies and romantic comedies are examples of substitutes in supply.

Complements and substitutes are important because they are common and have predictable effects on demand and supply. Changes in one market spill over to the other market, through the mechanism of complements or substitutes.

2.1.2.1 (Exercise) A typist charges $30/hr and types 15 pages per hour. Graph the supply of typed pages.

2.1.2.2 (Exercise) An owner of an oil well has two technologies for extracting oil. With one technology, the oil can be pumped out and transported for $5,000 per day, and 1,000 barrels per day are produced. With the other technology, which involves injecting natural gas into the well, the owner spends $10,000 per day and $5 per barrel produced, but 2,000 barrels per day are produced. What is the supply? Graph it.

(Hint: Compute the profits, as a function of the price, for each of the technologies. At what price would the producer switch from one technology to the other? At what price would the producer shut down and spend nothing?)

2.1.2.3 (Exercise) An entrepreneur has a factory the produces L^{α} widgets, where $\alpha<1$, when L hours of labor is used. The cost of labor (wage and benefits) is w per hour. If the entrepreneur maximizes profit, what is the supply curve for widgets?

Hint: The entrepreneur's profit, as a function of the price, is $pL^{\alpha} - wL$. The entrepreneur chooses the amount of labor to maximize profit. Find the amount of labor that maximizes, which is a function of p, w and α. The supply is the amount of output produced, which is L^{α}.

2.1.2.4 (Exercise) In the above exercise, suppose now that more than 40 hours entails a higher cost of labor (overtime pay). Let w be \$20/hr for under 40 hours, and \$30/hr for each hour over 40 hours, and $\alpha = ½$. Find the supply curve.

Hint: Let $L(w, p)$ be the labor demand when the wage is w (no overtime pay) and the price is p. Now show that, if $L(20,p) < 40$, the entrepreneur uses $L(20,p)$ hours. This is shown by verifying that profits are higher at $L(20,p)$ than at $L(30,p)$. If $L(30,p) > 40$, the entrepreneur uses $L(30,p)$ hours. Finally, if $L(20,p) > 40 > L(30,p)$, the entrepreneur uses 40 hours. Labor translates into supply via L^{α}.

2.1.2.5 (Exercise) In the previous exercise, for what range of prices does employment equal 40 hours? Graph the labor demanded by the entrepreneur.

Hint: The answer involves $\sqrt{10}$.

2.1.2.6 (Exercise) Suppose marginal cost, as a function of the quantity q produced, is mq. Find the producer's profit as a function of the price p.

2.2 The Market

Individuals with their own supply or demand trade in a market, which is where prices are determined. Markets can be specific or virtual locations – the farmer's market, the New York Stock Exchange, eBay – or may be an informal or more amorphous market, such as the market for restaurant meals in Billings, Montana or the market for roof repair in Schenectady, New York.

2.2.1 Market Demand and Supply

Individual demand gives the quantity purchased for each price. Analogously, the *market demand* gives the quantity purchased by all the market participants – the sum of the individual demands – for each price. This is sometimes called a "horizontal sum" because the summation is over the quantities for each price. An example is illustrated in Figure 2-7. For a given price p, the quantity q_1 demanded by one consumer, and the quantity q_2 demanded by a second consumer are illustrated. The sum of these quantities represents the market demand, if the market has just those two-participants. Since the consumer with subscript 2 has a positive quantity demanded for high prices, while the consumer with subscript 1 does not, the market demand coincides with consumer 2's demand when the price is sufficiently high. As the price falls, consumer 1 begins purchasing, and the market quantity demanded is larger than either individual participant's quantity, and is the sum of the two quantities.

Example: If the demand of buyer 1 is given by $q = \max\{0, 10 - p\}$, and the demand of buyer 2 is given by $q = \max\{0, 20 - 4p\}$, what is market demand for the two-participants?

Solution: First, note that buyer 1 buys zero at a price 10 or higher, while buyer 2 buys zero at a price of 5 or higher. For a price above 10, market demand is zero. For a price between 5 and 10, market demand is buyer 1's demand, or $10 - p$. Finally, for a price between zero and 5, the market quantity demanded is $10 - p + 20 - 4p = 30 - 5p$.

Market supply is similarly constructed – the market supply is the horizontal (quantity) sum of all the individual supply curves.

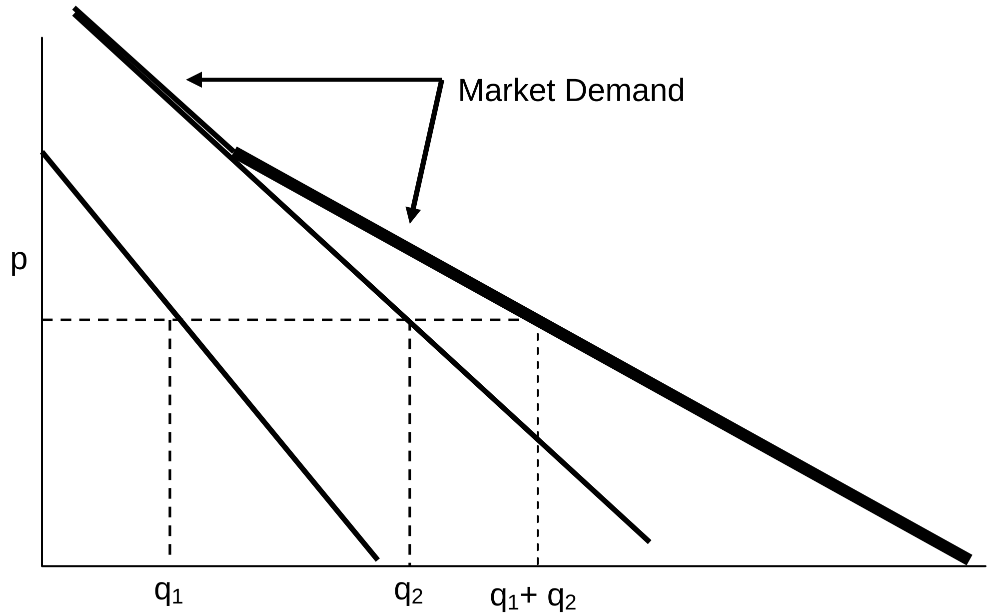


Figure 2-7: Market Demand

Example: If the supply of firm 1 is given by $q = 2p$, and the supply of firm 2 is given by $q = \max\{0, 5p - 10\}$, what is market supply for the two-participants?

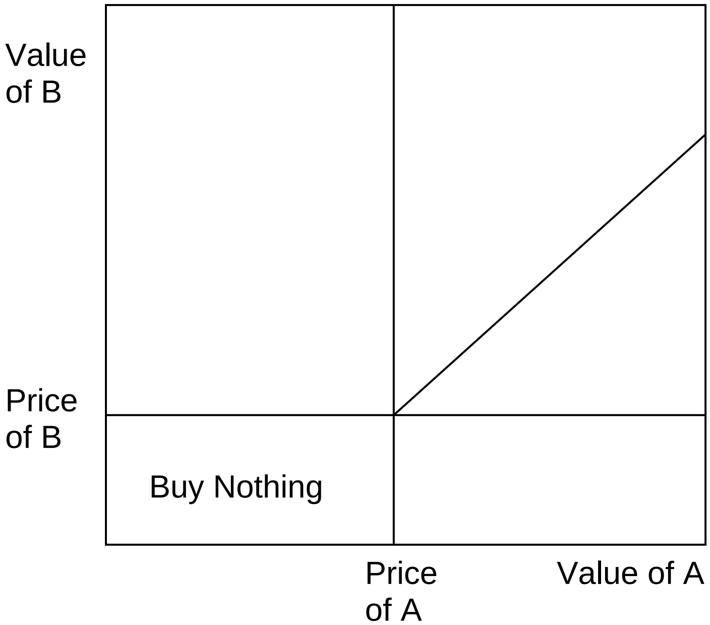
Solution: First, note that firm 1 is in the market at any price, but firm 2 is in the market only if price exceeds 2. Thus, for a price between zero and 2, market supply is firm 1's supply, or $2p$. For $p>2$, market supply is $5p - 10 + 2p = 7p - 10$.

2.2.1.1 (Exercise) Is the consumer surplus for market demand the sum of the consumer surpluses for the individual demands? Why or why not? Illustrate your conclusion with a figure like Figure 2-7.

2.2.1.2 (Exercise) Suppose the supply of firm i is $\alpha_i\, p$, when the price is p, where i takes on the values 1, 2, 3, ... n. What is the market supply of these n firms?

2.2.1.3 (Exercise) Suppose consumers in a small town choose between two restaurants, *A* and *B*. Each consumer has a value v_A for *A* and a value v_B for *B*, each of which is a uniform random draw from the [0,1] interval. Consumers buy whichever product offers the higher consumer surplus. The price of *B* is 0.2. In the square associated with the possible value types, identify which consumers buy from firm *A*. Find the demand (which is the area of the set of consumers who buy from *A* in the picture below). Hint: Consumers have three choices: Buy nothing (value 0), buy from A (value $v_A - p_A$) and buy from *B*, (value $v_B - p_B$

= v_B – 0.2). Draw the lines illustrating which choice has the highest value for the consumer.



2.2.2 Equilibrium

Economists use the term *equilibrium* in the same way as the word is used in physics, to represent a steady state in which opposing forces are balanced, so that the current state of the system tends to persist. In the context of supply and demand, equilibrium refers to a condition where the pressure for higher prices is exactly balanced by a pressure for lower prices, and thus that the current state of exchange between buyers and sellers can be expected to persist.

When the price is such that the quantity supplied of a good or service exceeds the quantity demanded, some sellers are unable to sell because fewer units are purchased than are offered. This condition is called a *surplus*. The sellers who fail to sell have an incentive to offer their good at a slightly lower price – a penny less – in order to succeed in selling. Such price cuts put downward pressure on prices, and prices tend to fall. The fall in prices generally reduces the quantity supplied and increases the quantity demanded, eliminating the surplus. That is, a surplus encourages price cutting, which reduces the surplus, a process that ends only when the quantity supplied equals the quantity demanded.

Similarly, when the price is low enough that the quantity demanded exceeds the quantity supplied, a *shortage* exists. In this case, some buyers fail to purchase, and these buyers have an incentive to accept a slightly higher price in order to be able to trade. Sellers are obviously happy to get the higher price as well, which tends to put upward pressure on prices, and prices rise. The increase in price tends to reduce the quantity demanded and increase the quantity supplied, thereby eliminating the shortage. Again, the process stops when the quantity supplied equals the quantity demanded.

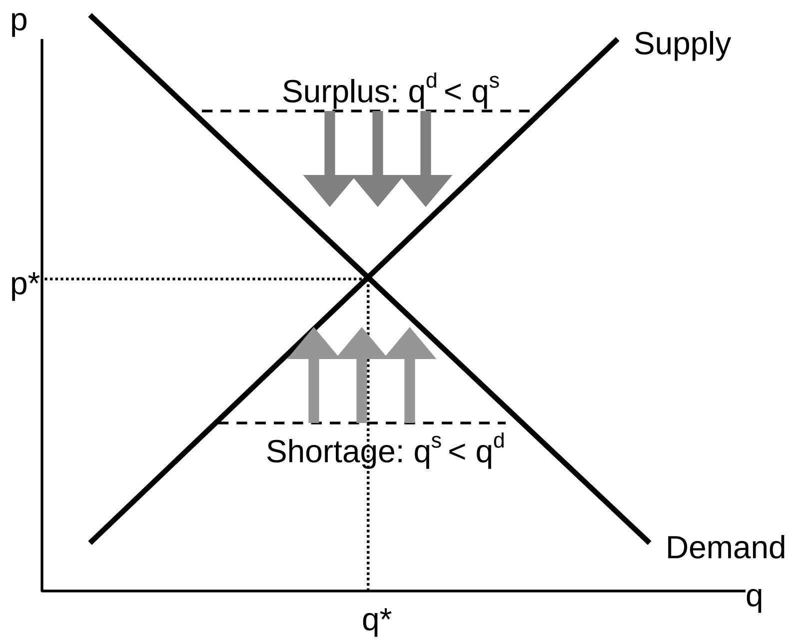


Figure 2-8: Equilibration

This logic, which is illustrated in Figure 2-8, justifies the conclusion that the only equilibrium price is the price in which the quantity supplied equals the quantity demanded. Any other price will tend to rise in a shortage, or fall in a surplus, until supply and demand are balanced. In Figure 2-8, a surplus arises at any price above the equilibrium price p^*, because the quantity supplied q^s is larger than the quantity demanded q^d. The effect of the surplus – leading to sellers with excess inventory – induces price cutting which is illustrated with three arrows pointing down.

Similarly, when the price is below p^*, the quantity supplied q^s is less than the quantity demanded q^d. This causes some buyers to fail to find goods, leading to higher asking prices and higher bid prices by buyers. The tendency for the price to rise is illustrated with the arrows pointing up. The only price which doesn't lead to price changes is p^*, the *equilibrium price in which the quantity supplied equals the quantity demanded*.

The logic of equilibrium in supply and demand is played out daily in markets all over the world, from stock, bond and commodity markets with traders yelling to buy or sell, to Barcelona fish markets where an auctioneer helps the market find a price, to Istanbul gold markets, to Los Angeles real estate markets.

2.2.2.1 (Exercise) If demand is given by $q^d(\mathrm{p}) = a - bp$, and supply is given by $q^s(\mathrm{p}) = cp$, solve for the equilibrium price and quantity. Find the consumer surplus and producer profits.

2.2.2.2 (Exercise) If demand is given by $q^d(p) = ap^{-\varepsilon}$, and supply is given by $q^s(p) = bp^{\eta}$, where all parameters are positive numbers, solve for the equilibrium price and quantity.

2.2.3 Efficiency of Equilibrium

The equilibrium of supply and demand balances the quantity demanded and the quantity supplied, so that there is no excess of either. Would it be desirable, from a social perspective, to force more trade, or to restrain trade below this level?

There are circumstances where the equilibrium level of trade has harmful consequences, and such circumstances are considered in Chapter 6. However, provided that the only people affected by a transaction are the buyer and seller, *the equilibrium of supply and demand maximizes the total gains from trade.*

This proposition is quite easy to see. To maximize the gains from trade, clearly the highest value buyers must get the goods. Otherwise, if there is a potential buyer that doesn't get the good with higher value than one who does, the gains from trade rise just by diverting the good to the higher value buyer. Similarly, the lowest cost sellers must supply those goods; otherwise we can increase the gains from trade by replacing a higher cost seller with a lower cost seller. Thus, the only question is how many goods should be traded to maximize the gains from trade, since it will involve the lowest cost sellers selling to the highest value buyers. Adding a trade increases the total gains from trade when that trade involves a buyer with value higher than the seller's cost. Thus, the gains from trade are maximized by the set of transactions to the left of the equilibrium, with the high value buyers buying from the low cost sellers.

In the economist's language, the equilibrium is *efficient*, in that it maximizes the gains from trade, under the assumption that the only people affected by any given transaction are the buyers and seller.

2.3 Changes in Supply and Demand

2.3.1 Changes in Demand

What are the effects of an increase in demand? As the population of California has grown, the demand for housing has risen. This has pushed the price of housing up, and also spurred additional development, increasing the quantity of housing supplied as well. We see such a demand increase illustrated in Figure 2-9, which represents an increase in the demand. In this figure, supply and demand have been abbreviated S and D. Demand starts at D_1 and is increased to D_2. Supply remains the same. The equilibrium price increases from p_1* to p_2*, and the quantity rises from q_1* to q_2*.

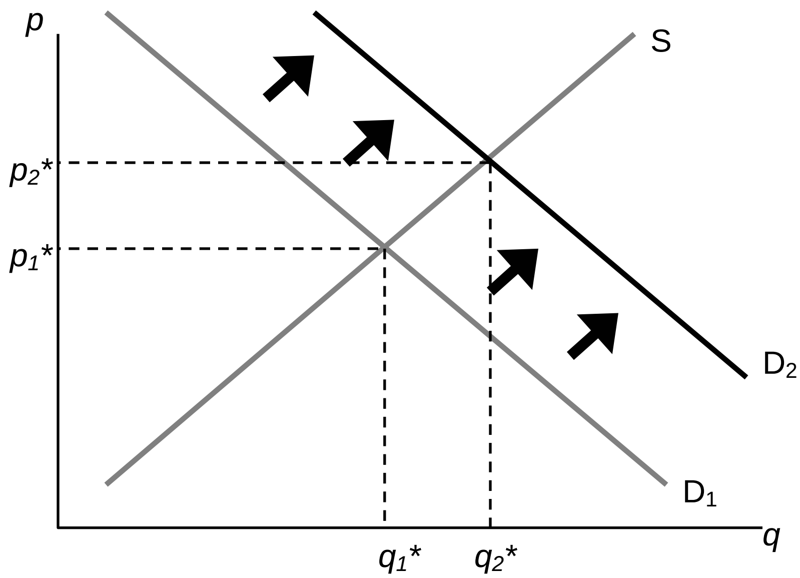


Figure 2-9: An Increase in Demand

A decrease in demand – such as occurred for typewriters with the advent of computers, or buggy whips as cars replaced horses as the major method of transportation – has the reverse effect of an increase, and implies a fall in both the price and the quantity traded. Examples of decreases in demand include products replaced by other products – VHS tapes were replaced by DVDs, vinyl records replaced by CDs, cassette tapes replaced by CDs, floppy disks (oddly named because the 1.44 MB "floppy," a physically hard product, replaced the 720KB, 5 ¼ inch soft floppy disk) replaced by CDs and flash memory drives, and so on. Even personal computers experienced a fall in demand as the market was saturated in the year 2001.

2.3.2 Changes in Supply

An increase in supply comes about from a fall in the marginal cost – recall that the supply curve is just the marginal cost of production. Consequently, an increased supply is represented by a curve that is lower and to the right on the supply/demand graph, which is an endless source of confusion for many students. The reasoning – lower costs and greater supply are the same thing – is too easily forgotten. The effects of an increase in supply are illustrated in Figure 2-10. The supply curve goes from S_1 to S_2, which represents a lower marginal cost. In this case, the quantity traded rises from q_1* to q_2* and price falls from p_1* to p_2*.

Computer equipment provides dramatic examples of increases in supply. Consider Dynamic Random Access Memory, or DRAM. DRAMs are the chips in computers and many other devices that store information on a temporary basis.[6] Their cost has fallen dramatically, which is illustrated in Figure 2-11.[7] Note that the prices in this figure reflect a logarithmic scale, so that a fixed percentage decrease is illustrated by a straight line. Prices of DRAMs fell to close to 1/1000th of their 1990 level by 2004. The means

[6] Information that will be stored on a longer term basis is generally embedded in flash memory or on a hard disk. Neither of these products lose their information when power is turned off, unlike DRAM.
[7] Used with permission of computer storage expert Dr. Edward Grochowski.

by which these prices have fallen are themselves quite interesting. The main reasons are shrinking the size of the chip (a "die shrink"), so that more chips fit on each silicon disk, and increasing the size of the disk itself, so that more chips fit on a disk. The combination of these two, each of which required the solutions to thousands of engineering and chemistry problems, has led to dramatic reductions in marginal costs and consequent increases in supply. The effect has been that prices fell dramatically and quantities traded rose dramatically.

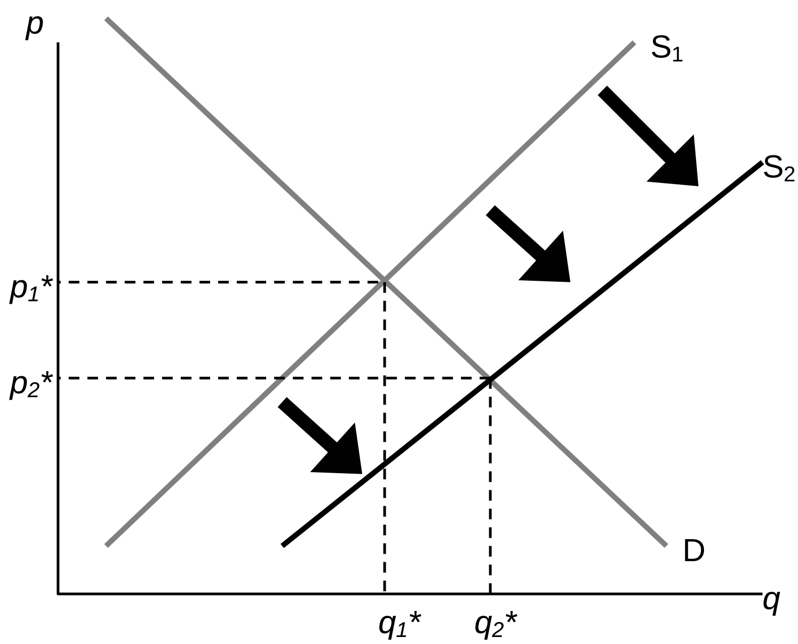


Figure 2-10: An Increase in Supply

An important source of supply and demand changes are changes in the markets of complements. A decrease in the price of a demand-complement increases the demand for a product, and similarly, an increase in the price of a demand-substitute increases demand for a product. This gives two mechanisms to trace through effects from external markets to a particular market via the linkage of demand substitutes or complements. For example, when the price of gasoline falls, the demand for automobiles (a complement) overall should increase. As the price of automobiles rises, the demand for bicycles (a substitute in some circumstances) should rise. When the price of computers falls, the demand for operating systems (a complement) should rise. This gives an operating system seller like Microsoft an incentive to encourage technical progress in the computer market, in order to make the operating system more valuable.

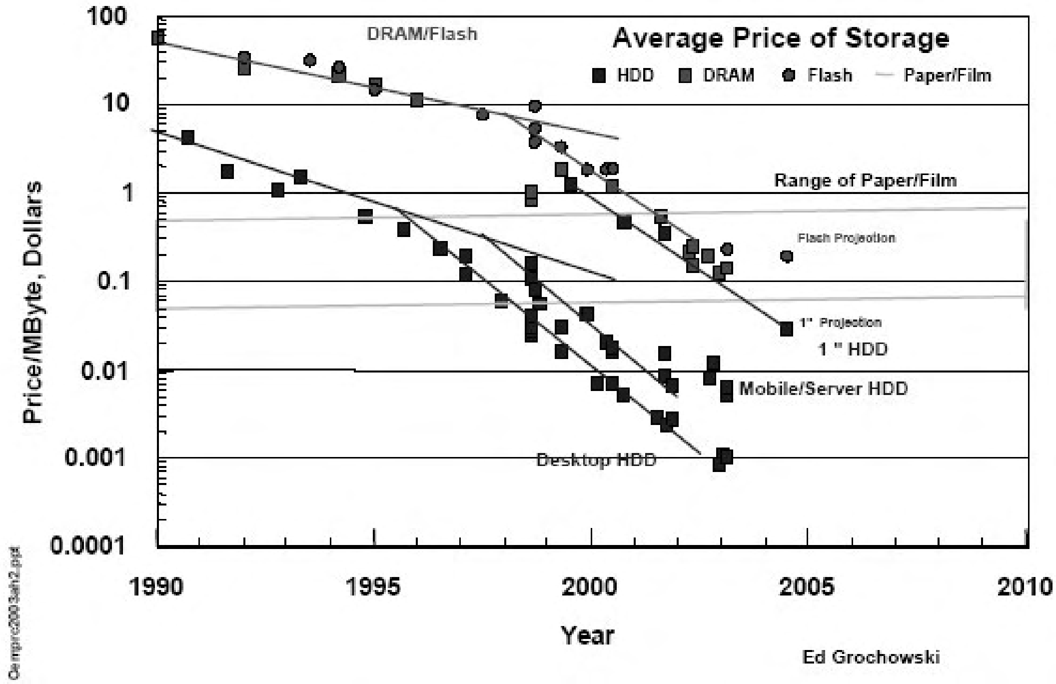


Figure 2-11: Price of Storage

An increase in the price of a supply-substitute reduces the supply of a good (by making the alternative good more attractive to suppliers), and similarly, a decrease in the price of a supply complement reduces the supply of a good. By making the by-product less valuable, the returns to investing in a good are reduced. Thus, an increase in the price of DVD-R discs (used for recording DVDs) discourages investment in the manufacture of CD-Rs, which are a substitute in supply, leading to a decrease in the supply of CD-Rs. This tends to increase the price of CD-Rs, other things equal. Similarly, an increase in the price of oil increases exploration for oil, tending to increase the supply of natural gas, which is a complement in supply. However, since natural gas is also a demand substitute for oil (both are used for heating homes), an increase in the price of oil also tends to increase the demand for natural gas. Thus, an increase in the price of oil increases both the demand and the supply of natural gas. Both changes increase the quantity traded, but the increase in demand tends to increase the price, while the increase in supply tends to decrease the price. Without knowing more, it is impossible to determine whether the net effect is an increase or decrease in the price.

2.3.2.1 (Exercise) Video games and music CDs are substitutes in demand. What is the effect of an increase in supply of video games on the price and quantity traded of music CDs? Illustrate your answer with diagrams for both markets.

2.3.2.2 (Exercise) Electricity is a major input into the production of aluminum, and aluminum is a substitute in supply for steel. What is the effect of an increase in price of electricity on the steel market?

2.3.2.3 (Exercise) Concerns about terrorism reduced demand for air travel, and induced consumers to travel by car more often. What should happen to the price of Hawaiian hotel rooms?

When the price of gasoline goes up, people curtail their driving to some extent, but don't immediately scrap their SUVs and rush out and buy more fuel-efficient automobiles or electric cars. Similarly, when the price of electricity rises, people don't replace their air conditioners and refrigerators with the most modern, energy-saving models right away. There are three significant issues raised by this kind of example. First, such changes may be transitory or permanent, and people reasonably react differently to temporary changes than to permanent changes. The effect of uncertainty is a very important topic and will be considered in section 5.2.6, but only in a rudimentary way for this introductory text. Second, energy is a modest portion of the cost of owning and operating an automobile or refrigerator, so it doesn't make sense to scrap a large capital investment over a small permanent increase in cost. Thus people rationally continue operating "obsolete" devices until their useful life is over, even when they wouldn't buy an exact copy of that device, an effect with the gobbledygook name of *hysteresis*. Third, a permanent increase in energy prices leads people to buy more fuel efficient cars, and to replace the old gas guzzlers more quickly. That is, the effects of a change are larger over a larger time interval, which economists tend to call the *long-run*.

A striking example of such delay arose when oil quadrupled in price in 1973-4, caused by a reduction in sales by the cartel of oil-producing nations, OPEC, which stands for the Organization of Petroleum Exporting Countries. The increased price of oil (and consequent increase in gasoline prices) caused people to drive less and to lower their thermostats in the winter, thus reducing the quantity of oil demanded. Over time, however, they bought more fuel efficient cars and insulated their homes more effectively, significantly reducing the quantity demanded still further. At the same time, the increased prices for oil attracted new investment into oil production in Alaska, the North Sea between Britain and Norway, Mexico and other areas. Both of these effects (long-run substitution away from energy, and long-run supply expansion) caused the price to fall over the longer term, undoing the supply reduction created by OPEC. In 1981, OPEC further reduced output, sending prices still higher, but again, additional investment in production, combined with energy-saving investment, reduced prices until they fell back to 1973 levels (adjusted for inflation) in 1986. Prices continued to fall until 1990, when they were at all-time low levels and Iraq's invasion of Kuwait and the resulting first Iraqi war sent them higher again.

Short-run and long-run effects represent a theme of economics, with the major conclusion of the theme that substitution doesn't occur instantaneously, which leads to predictable patterns of prices and quantities over time.

It turns out that direct estimates of demand and supply are less useful as quantifications than notions of percentage changes, which have the advantage of being unit-free. This observation gives rise to the concept of elasticity, the next topic.

2.4 Elasticities

2.4.1 Elasticity of Demand

Let $x(p)$ represent the quantity purchased when the price is p, so that the function x represents demand. How responsive is demand to price changes? One might be tempted to use the derivative x' to measure the responsiveness of demand, since it measures directly how much the quantity demanded changes in response to a small change in price. However, this measure has two problems. First, it is sensitive to a change in units. If I measure the quantity of candy in kilograms rather than pounds, the derivative of demand for candy with respect to price changes even when demand itself has remained the same. Second, if I change price units, converting from one currency to another, again the derivative of demand will change. So the derivative is unsatisfactory as a measure of responsiveness because it depends on units of measure. A common way of establishing a unit-free measure is to use percentages, and that suggests considering the responsiveness of demand in percentage terms to a small percentage change in price. This is the notion of *elasticity of demand*.[8] The elasticity of demand is the percentage decrease in quantity that results from a small percentage increase in price. Formally, the elasticity of demand, which is generally denoted with the Greek letter epsilon ε (chosen to mnemonically suggest elasticity) is

$$\varepsilon = -\frac{dx/x}{dp/p} = -\frac{p}{x}\frac{dx}{dp} = -\frac{px'(p)}{x(p)}.$$

The minus sign is included to make the elasticity a positive number, since demand is decreasing. First, let's verify that the elasticity is in fact unit free. A change in the measurement of x cancels because the proportionality factor appears in both the numerator and denominator. Similarly, if we change the units of measurement of price to replace the price p with $r=ap$, $x(p)$ is replaced with $x(r/a)$. Thus, the elasticity is

$$\varepsilon == -\frac{r\frac{d}{dr}x(r/a)}{x(r/a)} = -\frac{rx'(r/a)\frac{1}{a}}{x(r/a)} = -\frac{px'(p)}{x(p)},$$

which is independent of a, and therefore not affected by the change in units.

How does a consumer's expenditure, also known as (individual) total revenue, react to a change in price? The consumer buys $x(p)$ at a price of p, and thus expenditure is $TR = px(p)$. Thus

$$\frac{d}{dp}px(p) = x(p) + px'(p) = x(p)\left(1 + \frac{px'(p)}{x(p)}\right) = x(p)(1-\varepsilon).$$

Therefore,

[8] The concept of elasticity was invented by Alfred Marshall, 1842-1924, in 1881 while sitting on his roof.

$$\frac{\frac{d}{dp}TR}{\frac{1}{p}TR} = 1 - \varepsilon.$$

Table 2-1: Various Demand Elasticities[9]

Product	ε
Salt	0.1
Matches	0.1
Toothpicks	0.1
Airline travel, short-run	0.1
Residential natural gas, short-run	0.1
Gasoline, short-run	0.2
Automobiles, long-run	0.2
Coffee	0.25
Legal services, short-run	0.4
Tobacco products, short-run	0.45
Residential natural gas, long-run	0.5
Fish (cod) consumed at home	0.5
Physician services	0.6
Taxi, short-run	0.6
Gasoline, long-run	0.7
Movies	0.9
Shellfish, consumed at home	0.9
Tires, short-run	0.9
Oysters, consumed at home	1.1
Private education	1.1
Housing, owner occupied, long-run	1.2
Tires, long-run	1.2
Radio and television receivers	1.2
Automobiles, short-run	1.2-1.5
Restaurant meals	2.3
Airline travel, long-run	2.4
Fresh green peas	2.8
Foreign travel, long-run	4.0
Chevrolet automobiles	4.0
Fresh tomatoes	4.6

[9] From http://www.mackinac.org/archives/1997/s1997-04.pdf; cited sources: *Economics: Private and Public Choice*, James D. Gwartney and Richard L. Stroup, eighth edition 1997, seventh edition 1995; Hendrick S. Houthakker and Lester D. Taylor, *Consumer Demand in the United States*, 1929-1970 (Cambridge: Harvard University Press, 1966,1970); Douglas R. Bohi, *Analyzing Demand Behavior* (Baltimore: Johns Hopkins University Press, 1981); Hsaing-tai Cheng and Oral Capps, Jr., "Demand for Fish" *American Journal of Agricultural Economics*, August 1988; and U.S. Department of Agriculture.

In words, the percentage change of total revenue resulting from a one percent change in price is one minus the elasticity of demand. Thus, a one percent increase in price will increase total revenue when the elasticity of demand is less than one, which is defined as an *inelastic* demand. A price increase will decrease total revenue when the elasticity of demand is greater than one, which is defined as an *elastic* demand. The case of elasticity equal to one is called *unitary elasticity*, and total revenue is unchanged by a small price change. Moreover, that percentage increase in price will increase revenue by approximately 1-ε percent. Because it is often possible to estimate the elasticity of demand, the formulae can be readily used in practice

Table 2-1 provides estimates on demand elasticities for a variety of products.

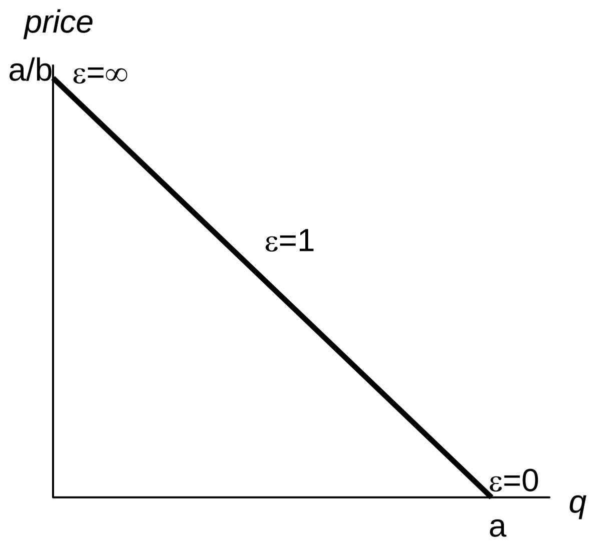


Figure 2-12: Elasticities for Linear Demand

When demand is linear, $x(p)=a-bp$, the elasticity of demand has the form

$$\varepsilon = \frac{bp}{a-bp} = \frac{p}{a/b - p}.$$

This case is illustrated in Figure 2-12.

If demand takes the form $x(p)=ap^{-\varepsilon}$, then demand has *constant elasticity*, and the elasticity is equal to ε.

2.4.1.1 (Exercise) Suppose a consumer has a constant elasticity of demand ε, and demand is *elastic* (ε > 1). Show that expenditure increases as price decreases.

2.4.1.2 (Exercise) Suppose a consumer has a constant elasticity of demand ε, and demand is inelastic (ε < 1). What price makes expenditure the greatest?

2.4.1.3 (Exercise) For a consumer with constant elasticity of demand $\varepsilon > 1$, compute the consumer surplus.

2.4.2 Elasticity of Supply

The elasticity of supply is analogous to the elasticity of demand, in that it is a unit-free measure of the responsiveness of supply to a price change, and is defined as the percentage increase in quantity supplied resulting from a small percentage increase in price. Formally, if $s(p)$ gives the quantity supplied for each price p, the elasticity of supply, denoted η (the Greek letter "eta", chosen because epsilon was already taken) is

$$\eta = \frac{ds/s}{dp/p} = \frac{p}{s}\frac{ds}{dp} = \frac{ps'(p)}{s(p)}.$$

Again similar to demand, if supply takes the form $s(p)=ap^{\eta}$, then supply has *constant elasticity*, and the elasticity is equal to η. A special case of this form is linear supply, which occurs when the elasticity equals one.

2.4.2.1 (Exercise) For a producer with constant elasticity of supply, compute the producer profits.

2.5 Comparative Statics

When something changes – the price of a complement, the demand for a good – what happens to the equilibrium? Such questions are answered by *comparative statics*, which are the changes in equilibrium variables when other things change. The use of the term "static" suggests that such changes are considered without respect to dynamic adjustment, but instead just focus on the changes in the equilibrium level. Elasticities will help us quantify these changes.

2.5.1 Supply and Demand Changes

How much do the price and quantity traded change in response to a change in demand? We begin by considering the constant elasticity case, which will let us draw conclusions for small changes in more general demand functions. We will denote the demand function by $q_d(p)=ap^{-\varepsilon}$ and supply function by $q_s(p)=bp^{\eta}$. The equilibrium price p^* is given by the quantity supplied equal to the quantity demanded, or the solution to the equation:

$$q_d(p^*) = q_s(p^*).$$

Substituting the constant elasticity formulae,

$$ap^{*-\varepsilon} = q_d(p^*) = q_s(p^*) = bp^{*\eta}.$$

Thus,

$$\frac{a}{b} = p^{*\varepsilon+\eta},$$

or

$$p^* = \left(\frac{a}{b}\right)^{1/(\varepsilon+\eta)}.$$

The quantity traded, q^*, can be obtained from either supply or demand and the price:

$$q^* = q_S(p^*) = bp^{*\eta} = b\left(\frac{a}{b}\right)^{\eta/(\varepsilon+\eta)} = a^{\eta/(\varepsilon+\eta)} b^{\varepsilon/(\varepsilon+\eta)}.$$

There is one sense in which this gives an answer to the question of what happens when demand increases. An increase in demand, holding the elasticity constant, corresponds to an increase in the parameter a. Suppose we increase a by a fixed percentage, replacing a by $a(1+\Delta)$. Then price goes up by the multiplicative factor $(1+\Delta)^{1/(\varepsilon+\eta)}$ and the change in price, as a proportion of the price, is

$$\frac{\Delta p^*}{p^*} = (1+\Delta)^{1/(\varepsilon+\eta)} - 1.$$

Similarly, quantity rises by

$$\frac{\Delta q^*}{q^*} = (1+\Delta)^{\eta/(\varepsilon+\eta)} - 1.$$

These formulae are problematic for two reasons. First, they are specific to the case of constant elasticity. Second, they are moderately complicated. Both of these issues can be addressed by considering small changes, that is, a small value of Δ. We make use of a trick to simplify the formula. The trick is that, for small Δ,

$$(1+\Delta)^r \approx 1 + r\Delta.$$

The squiggly equals sign $\approx$ should be read "approximately equal to."[10] Applying this insight, we have that:

[10] The more precise meaning of $\approx$ is that, as Δ gets small, the size of the error of the formula is small even relative to Δ. That is, $(1+\Delta)^r \approx 1 + r\Delta$. means $\frac{(1+\Delta)^r - (1+r\Delta)}{\Delta} \xrightarrow[\Delta \to 0]{} 0$.

For a small percentage increase Δ in demand, quantity rises by approximately $\eta\Delta/(\varepsilon+\eta)$ *percent and price rises by approximately* $\Delta/(\varepsilon+\eta)$ *percent.*

The beauty of this claim is that it holds even when demand and supply do not have constant elasticities, because the effect considered is local, and locally, the elasticity is approximately constant if the demand is "smooth."

2.5.1.1 (Exercise) Show that for a small percentage increase Δ in supply, quantity rises by approximately $\varepsilon\Delta/(\varepsilon+\eta)$ percent and price falls by approximately $\Delta/(\varepsilon+\eta)$ percent.

2.5.1.2 (Exercise) If demand is perfectly inelastic, what is the effect of a decrease in supply? Apply the formula and then graph the solution.

2.6 Trade

Supply and demand offers one approach to understanding trade, and it represents the most important and powerful concept in the toolbox of economists. However, for some issues, especially those of international trade, another related tool is very useful: the production possibilities frontier. Analysis using the production possibilities frontier was made famous by the "guns and butter" discussions of World War II. From an economic perspective, there is a tradeoff between guns and butter – if a society wants more guns, it must give up something, and one thing to give up is butter. That getting more guns might entail less butter often seems mysterious, because butter, after all, is made with cows, and indirectly with land and hay. But the manufacture of butter also involves steel containers, tractors to turn the soil, transportation equipment, and labor, all of which either can be directly used (steel, labor) or require inputs that could be used (tractors, transportation) to manufacture guns. From a production standpoint, more guns entail less butter (or other things).

2.6.1 Production Possibilities Frontier

Formally, the set of production possibilities is the collection of "feasible outputs" of an individual, group or society or country. You could spend your time cleaning your apartment, or you could study. The more of your time you devote to studying, the higher your grades will be, but the dirtier your apartment will be. This is illustrated, for a hypothetical student, in Figure 2-13.

The production possibilities set embodies the feasible alternatives. If you spend all your time studying, you would obtain a 4.0 (perfect) grade point average (GPA). Spending an hour cleaning reduces the GPA, but not by much; the second hour reduces by a bit more, and so on.

The boundary of the production possibilities set is known as the *production possibilities frontier*. This is the most important part of the production possibilities set, because at any point strictly inside the production possibilities set, it is possible to have more of everything, and usually we would choose to have more.[11] The slope of the production possibilities frontier reflects opportunity cost, because it describes what must be given up in order to acquire more of a good. Thus, to get a cleaner apartment, more time, or capital, or both, must be spent on cleaning, which reduces the amount of other goods and services that can be had. For the two-good case in Figure 2-13, diverting time to cleaning reduces studying, which lowers the GPA. The slope dictates how much lost GPA there is for each unit of cleaning.

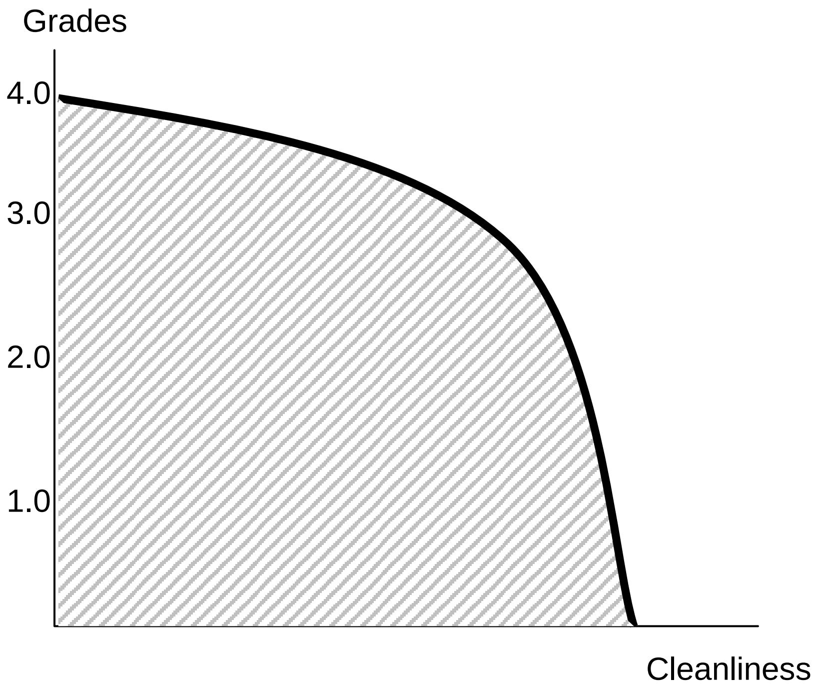


Figure 2-13: The Production Possibilities Frontier

One important feature of production possibilities frontiers is illustrated in the Figure 2-13: they are concave toward the origin. While this feature need not be universally true, it is a common feature, and there is a reason for it that we can see in the application. If you are only going to spend an hour studying, you spend that hour doing the most important studying that can be done in an hour, and thus get a lot of grades for the hour's work. The second hour of studying produces less value than the first, and the third hour less than the second. This is the principle of *diminishing marginal returns*. Diminishing marginal returns are like picking apples. If you are only going to pick apples for a few minutes, you don't need a ladder because the fruit is low on the tree; the more time spent, the fewer apples per hour you will pick.

[11] To be clear, we are considering an example with two goods, cleanliness and GPA. Generally there are lots of activities, like sleeping, eating, teeth-brushing, and the production possibilities frontier encompasses all of these goods. Spending all your time sleeping, studying and cleaning would still represent a point on a three-dimensional frontier.

Consider two people, Ann and Bob, getting ready for a party. One is cutting up vegetables, the other is making hors d'oeuvres. Ann can cut up two ounces of vegetables per minute, or make one hors d'oeuvre in a minute. Bob, somewhat inept with a knife, can cut up one ounce of vegetables per minute, or make two hors d'oeuvres per minute. Ann's and Bob's production possibilities frontiers are illustrated in the Figure 2-14, given that they have an hour to work.

Since Ann can produce two ounces of chopped vegetables in a minute, if she spends her entire hour on vegetables, she can produce 120 ounces. Similarly, if she devotes all her time to hors d'oeuvres, she produces 60 of them. The constant translation between the two means that her production possibilities frontier is a straight line, which is illustrated in the left side of Figure 2-14. Bob's is the reverse – he produces 60 ounces of vegetables, or 120 hors d'oeuvres, or something on the line in between.

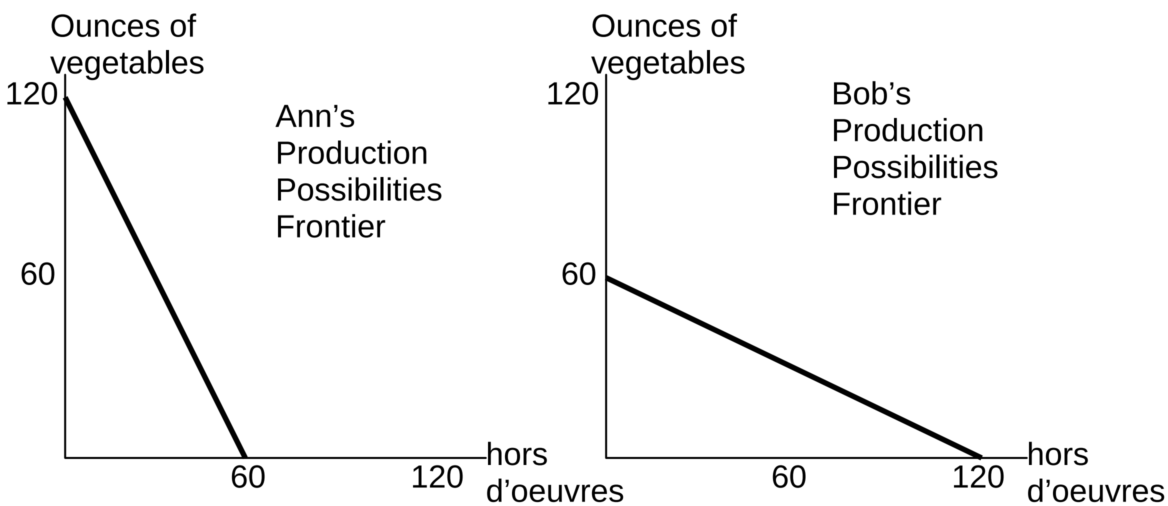


Figure 2-14: Two Production Possibilities Frontiers

For Ann, the opportunity cost of an ounce of vegetables is half of one hors d'oeuvre – to get one extra ounce of vegetable, she must spend 30 extra seconds on vegetables. Similarly, the cost of one hors d'oeuvres for Ann is two ounces of vegetables. Bob's costs are the inverse of Ann – an ounce of vegetables costs him two hors d'oeuvres.

What can Bob and Ann accomplish together? The important insight is that they should use the low cost person in the manufacture of each good, when possible. This means that if fewer than 120 ounces of vegetables will be made, Ann makes them all. Similarly, if fewer than 120 hors d'oeuvres are made, Bob makes them all. This gives a joint production possibilities frontier illustrated in the Figure 2-15. Together, they can make 180 of one and none of the other. If Bob makes only hors d'oeuvres, and Ann makes only chopped vegetables, they will have 120 of each. With fewer than 120 ounces of vegetables, the opportunity cost of vegetables is Ann's, and is thus half an hors d'oeuvre, but if more than 120 are needed, then the opportunity cost jumps to two.

Now change the hypothetical slightly – suppose that Bob and Ann are putting on separate dinner parties, each of which will feature chopped vegetables and hors d'oeuvres in equal portions. By herself, Ann can only produce 40 ounces of vegetables and 40 hors d'oeuvres if she must produce equal portions. She accomplishes this by spending 20 minutes on vegetables and 40 minutes on hors d'oeuvres. Similarly, Bob can produce 40 of each, but using the reverse allocation of time.

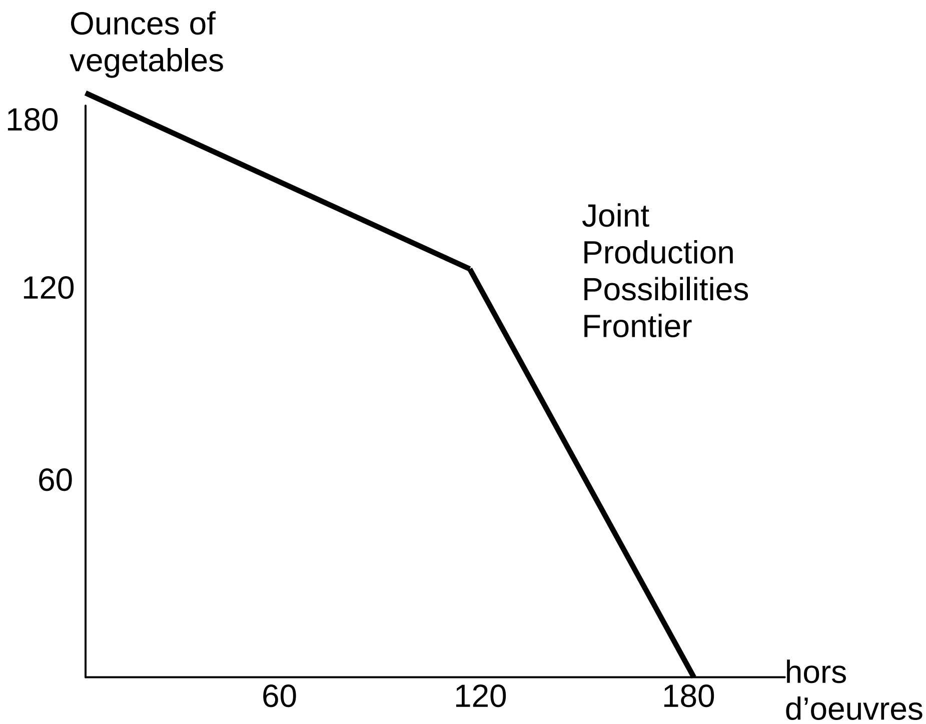


Figure 2-15: Joint PPF

By working together, they can collectively have more of both goods. Ann specializes in producing vegetables, and Bob specializes in producing hors d'oeuvres. This yields 120 units of each, which they can split equally, to have 60 of each. By specializing in the activity in which they have lower cost, Bob and Ann can jointly produce more of each good.

Moreover, Bob and Ann can accomplish this by trading. At a "one for one" price, Bob can produce 120 hors d'oeuvres, and trade 60 of them for 60 ounces of vegetables. This is better than producing the vegetables himself, which netted him only 40 of each. Similarly, Ann produces 120 ounces of vegetables, and trades 60 of them for 60 hors d'oeuvres. This trading makes them both better off.

The gains from specialization are potentially enormous. The grandfather of economics, Adam Smith, writes about specialization in the manufacture of pins:

> "...One man draws out the wire; another straights it; a third cuts it; a fourth points it; a fifth grinds it at the top for receiving the head; to make the head requires two or three distinct operations ; to put it on is a peculiar business; to whiten the pins is another ; it is even a trade by itself to put them into the paper ; and the important business of making a pin is, in this

> manner, divided into about eighteen distinct operations, which, in some manufactories, are all performed by distinct hands, though in others the same man will sometimes perform two or three of them."[12]

Smith goes on to say that skilled individuals could produce at most twenty pins per day acting alone, but that with specialization, ten people can produce 48,000 pins per day, 240 times as many pins.

2.6.1.1 (Exercise) The Manning Company has two factories, one that makes roof trusses, and one that makes cabinets. With m workers, the roof factory produces $\sqrt{m}$ trusses per day. With n workers, the cabinet plant produces $5\sqrt{n}$. The Manning Company has 400 workers to use in the two factories. Graph the production possibilities frontier. (Hint: Let T be the number of trusses produced. How many workers are used making trusses?)

2.6.1.2 (Exercise) Alarm & Tint, Inc., has 10 workers working a total of 400 hours per week. Tinting takes 2 hours per car. Alarm installation is complicated, however, and performing A alarm installations requires A^2 hours of labor. Graph Alarm & Tint's production possibilities frontier for a week.

2.6.2 Comparative and Absolute Advantage

Ann produces chopped vegetables because her opportunity cost of producing vegetables, at ½ of one hors d'oeuvre, is lower than Bob's. A lower opportunity cost is said to create a *comparative advantage*. That is, Ann gives up less to produce chopped vegetables than Bob, so in comparison to hors d'oeuvres, she has an advantage in the production of vegetables. Since the cost of one good is the amount of another good foregone, a comparative advantage in one good implies a *comparative disadvantage* in another. If you are better at producing butter, you are necessarily worse at something else, and in particular the thing you give up less of to get more butter.

To illustrate this point, let's consider another party planner. Charlie can produce one hors d'oeuvre, or one ounce of chopped vegetables, per minute. His production is strictly less than Ann's, that is, his production possibilities frontier lies inside of Ann's. However, he has a comparative advantage over Ann in the production of hors d'oeuvres, because he gives up only one ounce of vegetables to produce a hors d'oeuvres, while Ann must give up two ounces of vegetables. Thus, Ann and Charlie can still benefit from trade if Bob isn't around.

2.6.2.1 (Exercise) Graph the joint production possibilities frontier for Ann and Charlie, and show that collectively they can produce 80 of each if they need the same number of each product. Hint: First show that Ann will produce some of both goods, by showing that if Ann specializes, there are too many ounces of vegetables. Then show, if Ann devotes x minutes to hors d'oeuvres, that

$$60 + x = 2(60 - x).$$

---

[12] Adam Smith, "An Inquiry into the Nature and Causes of the Wealth of Nations," originally published 1776, released by the Gutenberg project, 2002.

When one production possibilities frontier lies outside another, the larger is said to have an absolute advantage – more total things are possible. In this case, Ann has an absolute advantage over Charlie – she can, by herself, have more – but not over Bob. Bob has an absolute advantage over Charlie, too, but again, not over Ann.

Diminishing marginal returns implies that the more of a good that a person produces, the higher is the cost (in terms of the good given up). That is to say, diminishing marginal returns means that supply curves slope upward; the marginal cost of producing more is increasing in the amount produced.

Trade permits specialization in activities in which one has a comparative advantage. Moreover, whenever opportunity costs differ, potential gains from trade exist. If person 1 has an opportunity cost of c_1 of producing good X (in terms of Y, that is, for each unit of X that person 1 produces, person 1 gives up c_1 units of Y), and person 2 has an opportunity cost of c_2, then there are gains from trade whenever c_1 is not equal to c_2 and neither party has specialized.[13] Suppose $c_1 < c_2$. Then by having person 1 increase the production of X by Δ, $c_1 \Delta$ less of the good Y is produced. Let person 2 reduce the production of X by Δ, so that the production of X is the same. Then there is $c_2 \Delta$ units of Y made available, for a net increase of $(c_2\text{-}c_1)\,\Delta$. The net changes are summarized in Table 2-2.

Table 2-2: Construction of the Gains From Trade

	1	2	Net Change
Change in X	$+\Delta$	$-\Delta$	0
Change in Y	$-c_1 \Delta$	$c_2 \Delta$	$(c_2\text{-}c_1)\,\Delta$

Whenever opportunity costs differ, there are gains from re-allocating production from one producer to another, gains which are created by having the low cost producers produce more, in exchange for greater production of the other good by the other producer, who is the low cost producer of this other good. An important aspect of this re-allocation is that it permits production of more of all goods. This means there is little ambiguity about whether it is a good thing to re-allocate production – it just means we have more of everything we want.[14]

How can we guide the reallocation of production to produce more goods and services? It turns out that under some circumstances, the price system does a superb job of creating efficient production. The price system posits a price for each good or service, and anyone can sell at the common price. The insight is that such a price induces efficient production. To see this, suppose we have a price p which is the number of units

[13] If a party specialized in one product, it is a useful convention to say that the marginal cost of that product is now infinite, since no more can be produced.
[14] If you are worried that more production means more pollution or other bad things, rest assured. Pollution is a bad, so we enter the negative of pollution (or environmental cleanliness) as one of the goods we would like to have more of. The reallocation dictated by differences in marginal costs produces more of all goods. Now with this said, we have no reason to believe that the reallocation will benefit everyone – there may be winners and losers.

of Y one has to give to get a unit of X. (Usually prices are in currency, but we can think of them as denominated in goods, too.) If I have a cost c of producing X, which is the number of units of Y that I lose to obtain a unit of X, I will find it worthwhile to sell X if $p > c$, because the sale of a unit of X, nets me $p - c$ units of Y, which I can either consume or resell for something else I want. Similarly, if $c > p$, I would rather buy X (producing Y to pay for it). Either way, only producers with costs less than p will produce X, and those with costs greater than p will purchase X, paying for it with Y, which they can produce more cheaply than its price. (The price of Y is $1/p$ – that is the amount of X one must give to get a unit of Y.)

Thus, a price system, with appropriate prices, will guide the allocation of production to insure the low cost producers are the ones who produce, in the sense that there is no way of re-allocating production to obtain more goods and services.

2.6.2.2 (Exercise) Using Manning's production possibilities frontier in 2.6.1.1 (Exercise), compute the marginal cost of trusses in terms of cabinets.

2.6.2.3 (Exercise) Using Alarm & Tint's production possibilities frontier in 2.6.1.2 (Exercise), compute the marginal cost of alarms in terms of window tints.

2.6.3 Factors and Production

Production possibilities frontiers provide the basis for a rudimentary theory of international trade. To understand the theory, it is first necessary to consider that there are fixed and mobile factors. *Factors of production* are jargon for inputs to the production process. Labor is generally considered a fixed factor, because most countries don't have borders wide open to immigration, although of course some labor moves across international borders. Temperature, weather, and land are also fixed – Canada is a high-cost citrus grower because of its weather. There are other endowments that could be exported, but are expensive to export because of transportation costs, including water and coal. Hydropower – electricity generated from the movement of water – is cheap and abundant in the Pacific Northwest, and as a result, a lot of aluminum is smelted there, because aluminum smelting requires lots of electricity. Electricity can be transported, but only with losses (higher costs), which gives other regions a disadvantage in the smelting of aluminum. Capital is generally considered a mobile factor, because plants can be built anywhere, although investment is easier in some environments than in others. For example, reliable electricity and other inputs are necessary for most factories. Moreover, comparative advantage may arise from the presence of a functioning legal system, the enforcement of contracts, and the absence of bribery, because enforcement of contracts increases the return on investment by increasing the probability the economic return to investment isn't taken by others.

Fixed factors of production give particular regions a comparative advantage in the production of some kinds of goods, and not in others. Europe, the United States and Japan have a relative abundance of highly skilled labor, and have a comparative advantage in goods requiring high skills, like computers, automobiles and electronics. Taiwan, South Korea, Singapore and Hong Kong have increased the available labor skills, and now manufacture more complicated goods like VCRs, computer parts and the like. Mexico has a relative abundance of middle-level skills, and a large number of assembly plants operate there, as well as clothing and shoe manufacturers. Lower skilled Chinese workers manufacture the majority of the world's toys. The skill levels of China are rising rapidly.

The basic model of international trade was first described by David Ricardo (1772-1823), and suggests that nations, responding to price incentives, will specialize in the production of goods in which they have a comparative advantage, and purchase the goods in which they have a comparative disadvantage. In Ricardo's description, England has a comparative advantage of manufacturing cloth, and Portugal in producing wine, leading to gains from trade from specialization.

The Ricardian theory suggests that the United States, Canada, Australia and Argentina should export agricultural goods, especially grains that require a large land area for the value generated (they do). It suggests that complex technical goods should be produced in developed nations (they are) and that simpler products and natural resources should be exported by the lesser developed nations (they are). It also suggests that there should be more trade between developed and underdeveloped nations than between developed and other developed nations. The theory falters on this prediction – the vast majority of trade is between developed nations. There is no consensus for the reasons for this, and politics plays a role – the North American Free Trade Act vastly increased the volume of trade between the United States and Mexico, for example, suggesting that trade barriers may account for some of the lack of trade between the developed and the underdeveloped world. Trade barriers don't account for the volume of trade between similar nations, which the theory suggests should be unnecessary. Developed nations sell each other mustard and tires and cell phones, exchanging distinct varieties of goods they all produce.

2.6.4 International Trade

The Ricardian theory emphasizes that the relative abundance of particular factors of production determines comparative advantage in output, but there is more to the theory. When the United States exports a computer to Mexico, American labor, in the form of a physical product, has been sold abroad. When the United States exports soybeans to Japan, American land (or at least the use of American land for a time) has been exported to Japan. Similarly, when the United States buys car parts from Mexico, Mexican labor has been sold to the United States, and similarly when the Americans buy Japanese televisions, Japanese labor has been purchased. The goods that are traded internationally embody the factors of production of the producing nations, and it is useful to think of international trade as directly trading the inputs through the incorporation of inputs into products.

If the set of traded goods is broad enough, the value of factors of production should be equalized through trade. The United States has a lot of land, relative to Japan, but by selling agricultural goods to Japan, it is as if Japan had more land, by way of access to US land. Similarly, by buying automobiles from Japan, it is as if a portion of the Japanese factories were present in the United States. With inexpensive transportation, the trade equalizes the values of factories in the United States and Japan, and also equalizes the value of agricultural land. One can reasonably think that soybeans are soybeans, wherever they are produced, and that trade in soybeans at a common price forces the costs of the factors involved in producing soybeans to be equalized across the producing nations. The purchase of soybeans by Japanese drives up the value of American land, and drives down the value of Japanese land by giving an alternative to its output, leading toward equalization of the value of the land across the nations.

This prediction, known as *factor price equalization*, of modern international trade theory was first developed by Paul Samuelson (1915 –) and generalized by Eli Heckscher (1879 – 1952) and Bertil Ohlin (1899 – 1979). It has powerful predictions, including the equalization of wages of equally skilled people after free trade between the United States and Mexico. Thus, free trade in physical goods should equalize the price of haircuts, and land, and economic consulting, in Mexico City and New York. Equalization of wages is a direct consequence of factor price equalization because labor is a factor of production. If economic consulting is cheap in Mexico, trade in goods embodying economic consulting – boring reports, perhaps – will bid up the wages in the low wage area, and reduce the quantity in the high wage area.

An even stronger prediction of the theory is that the price of water in New Mexico should be the same as in Minnesota. If water is cheaper in Minnesota, trade in goods that heavily use water – e.g. paper – will tend to bid up the value of Minnesota water, while reducing the premium on scarce New Mexico water.

It is fair to say that if factor price equalization works fully in practice, it works very, very slowly. Differences in taxes, tariffs and other distortions make it a challenge to test the theory across nations. On the other hand, within the United States, where we have full factor mobility and product mobility, we still have different factor prices – electricity is cheaper in the Pacific Northwest. Nevertheless, nations with a relative abundance of capital and skilled labor export goods that use these intensively, nations with a relative abundance of land export land-intensive goods like food, nations with a relative abundance of natural resources export these resources, and nations with an abundance of low-skilled labor export goods that make intensive use of this labor. The reduction of trade barriers between such nations works like Ann and Bob's joint production of party platters: by specializing in the goods in which they have a comparative advantage, there is more for all.