

Lecture Notes Part 2
Economics of Sustainability
K Foster, CCNY, Spring 2016

Table of Contents

Short Review of Production	4
Firms Choosing How to Produce	4
One Input	5
More than One Input	6
Short Run vs Long Run	7
Profit Maximization	8
Profit Maximization with One Input	8
Profit Maximization with Multiple Inputs	9
Cost Minimization/Profit Maximization	10
Marginal Revenue	11
Hicks-Marshall Rules of Derived Demand:	12
Supplementary Material for Advanced students	13
Costs	13
Production Externalities	20
Marginal Abatement Costs	24
Regulation of Pollution	26
Fees & Tradable Permits	26
Tradable Permits	26
Pollution Over Time	31
When Costs & Benefits are Imperfectly Known (i.e. The Real World...)	32
Extreme Case 1: Threshold effects of pollution	33
Extreme Case 2: Constant Marginal Damages	36
Case 3: MD steeper than MC	39
Case 4: MD flatter than MC	41
Details of Fees and Permits	43
Regulation through Liability	43

Regulation through Insurance.....	44
Costs and Benefits	45
Valuation of Life	45
Risk & Uncertainty.....	45
Expected Surplus	46
Option Price	47
Irreversibility and Precautionary Principle	47
Actual Behavior of People making Choices under Uncertainty	47
The Asian Disease	47
Prospect Theory	48
Choice of Dumping or Safe Disposal	49
Background on Global Climate Change.....	50
Point 1: Human activity is changing the earth's climate	50
Point 2: The poorest people will bear much of the burden and the costs of climate change..	53
Point 3: As poor people become richer, this worsens the pace of climate change	55

Short Review of Production

Firms Choosing How to Produce

Assume that firms want to maximize profit, π , which is Revenue minus Cost. This is far from a perfect description of the world of course but it's a start.

Split the production decision into two parts: first, if a firm wants to make a particular quantity of output, what is the cheapest way to make it; second, how much output will a firm choose to make. This division allows us to focus on particular pieces first.

The first question – to make a particular quantity of output, what is the cheapest way to make it? – gives us the single essential number: the cost of that amount of output. The cost of this output is the only important item that the firm, when choosing amount of output to produce, needs to know. It does not need to know the quantities of inputs or relative costs. This split can also be thought of as reflecting a firm's organization: there is the corporate level that makes the decisions about how much product to make, if those output levels have a particular cost. Then these decisions are communicated to the plants that make the output, where each plant manager is told to make a particular amount of output, using the cheapest input mix possible. The plant manager doesn't need to know how a particular quantity of output was chosen; the corporate level doesn't need to know details of how that output is made, just the cost. Subcontracting works the same way.

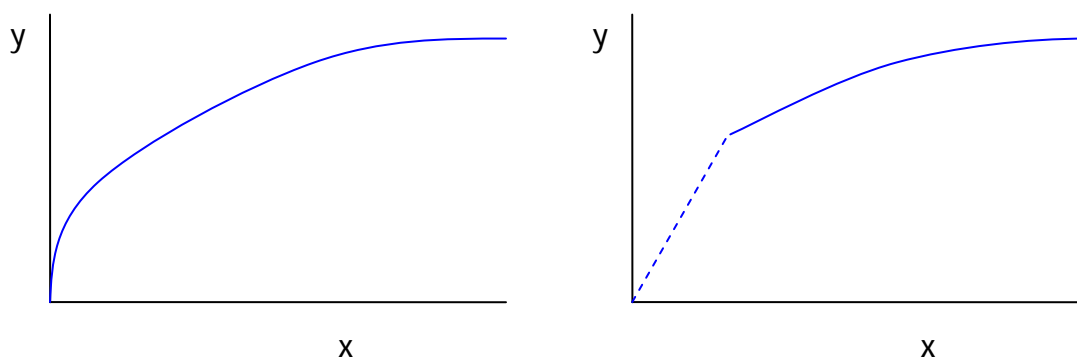
At this level we are not paying attention to questions of corporate structure. Given the decision structure from above, we might think of the plant managers as being a separate firm, outsourcing production. (A brand-name computer maker buys chips from a separate company; it doesn't need to know details of how the chips are made, indeed that might be a close-held secret. All it needs to know is the cost.) Our modern economy has many such firms providing corporate services, from high-level research down to the company cafeteria. The informational savings are immense: a firm doesn't need to know the details of how each input is made, it just needs to know how much they cost. If they want paper, they don't need to ask about how many trees grew for how long, they just need the cost per ream. (This is why central planning fails, since there are no prices and so no informational savings – see Oskar Lange, Hayek.)

We begin our analysis at the base, at the level of the plant, which is given an order for a particular quantity of output and must choose how to most cheaply make it. Again we divide the decision into two parts: first asking what is physically possible (what inputs can make the output) and then asking which combination is cheapest.

One Input

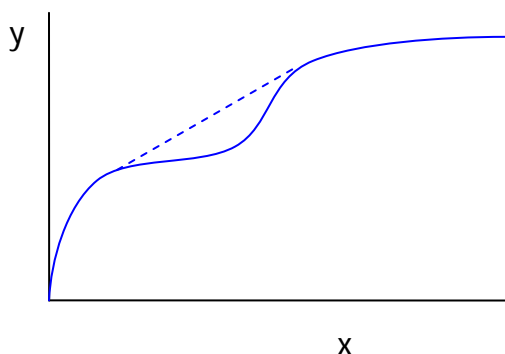
The simplest case is where one input makes one output, so we simply have $y = f(x)$. The marginal product of the input is how much additional output is made by adding more input, or $\frac{\Delta y}{\Delta x}$, which is the slope of the graph. Assume that it is increasing, continuous and convex.

The assumption of $f(x)$ being an increasing function (i.e. that $\frac{\Delta y}{\Delta x} > 0$) is anodyne. Just as with utility, if output actually falls when inputs rise then you don't need an advanced degree to figure out that you should cut back. The interesting problem occurs when output could still be increased and you want to figure out if it is profitable.



The assumptions of continuity and convexity don't seem as obvious. But they can be solved if we think of the firm's problem over a slightly longer period. Suppose that a firm's underlying physical process of production is discontinuous: it takes at least 100 units of input in a day to make 100 output units, but less than 100 of the input just won't even start up the machine. Is the firm's production function to be considered discontinuous? Well what if the firm got orders for 50 units of output per day – what would it do? Clearly it could just run the machine every other day, and average 50 units of output with 50 units of input per day. If orders run at 80 per day then the machine is run on 4 out of 5 days, and so forth. Of course this assumes that the output is storable and that the time over which we are speaking is relatively short (more on this later). But the assumption is not too bad.

The convexity assumption comes by the same assumption. If the firm can make 100 output with 100 input then it could make at least half as much output with half the input. (On the graph, any chord drawn between 2 points will lie on or beneath the production function.) If there were non-convexities in the underlying physical process then, again, production could be structured to avoid these.



The convexity assumption is also why we often talk about a "Law" of Diminishing Marginal Product. It is reasonable to assume that the Marginal Product, $\frac{\Delta y}{\Delta x}$, is diminishing (or at least not increasing) because if it were increasing then, as in the graph above, the firm would want to figure out ways to exploit this.

Clearly, assuming just one input to the production function is restrictive. I can't think of too many things that are produced in that way (except for the world's oldest profession). We want to consider multiple inputs.

More than One Input

We commonly limit ourselves to two inputs because that allows easy graphing and still gets to most of the complexities. But you should be able to see how the number of inputs could be increased.

Now describe the production plant with a production function where inputs (x_1, x_2) are transformed into outputs by way of a production function: $y = f(x_1, x_2)$. We could imagine a wide variety of production functions but we assume that it has some basic properties. Note that, whereas in the consumer problem, we were reluctant to make restrictions directly to the utility function and instead discussed assumptions about the underlying preferences, that was because utility was un-measurable and only a convenient descriptive device. Production is more easily measured as long as there is some physical output: tons of steel or pairs of sneakers or casks of beer. So we make assumptions directly about the production function.

We again assume that the production function is increasing (so more inputs lead to more output), continuous, and convex (or something like convex, in higher-dimension analogue).

Now define each input's Marginal Product: $MP_1 = \frac{\Delta y}{\Delta x_1} = \frac{\Delta f(x_1, x_2)}{\Delta x_1}$, where we use the function

notation to remind ourselves that the MP for each input is likely to be different, for different levels of each input. This is important – there are likely to be complementarities in production.

The Marginal Product of one input is likely to depend on the levels of other inputs as well. (For example we often hear statistics that workers in third-world countries are not as productive as US workers – this doesn't mean they're any worse, just that they have different levels of other inputs.) Someone with a jackhammer is more productive (of destruction) than someone with a sledgehammer.

Again we assume diminishing marginal products, that MP_1 falls as input 1 rises (holding constant input 2) and vice versa. This "holding constant" part is particularly important since, while in the long run we might be able to increase output by increasing both inputs, in the short run one input or the other is usually less flexible. Consider the typical office worker nowadays, who usually gets one computer. If a company hires more people without buying more computers, then the productivity of the new people (whatever their talent!) will be limited as they have to jostle for computer time. Similarly if the company got new computers without hiring new people – a few people might get multiple computers on their desks, and some might be more productive with those new computers, but not very much.

This is distinct from returns to scale, which asks what happens to output if all of the inputs are increased. Hiring people without getting more computers might not raise output much; getting computers without hiring more people might not raise output much; but hiring more workers and giving each a computer might still raise output. Diminishing Marginal Products for each input alone does not imply diminishing returns to scale.

To more formally define returns to scale, suppose a firm doubled its inputs, and ask what would happen to outputs? If output doubled exactly then a firm would have constant returns to scale (CRS). If output increased by more than double then the firm has increasing returns to scale (IRS). If output increased by less than double then there are decreasing returns to scale (DRS). To put this a bit more abstractly, we compare the output from doubling the inputs, $f(2x_1, 2x_2)$, with twice the original output, $2f(x_1, x_2)$. If $f(2x_1, 2x_2) = 2f(x_1, x_2)$ then production is CRS; if $f(2x_1, 2x_2) > 2f(x_1, x_2)$ then IRS; if $f(2x_1, 2x_2) < 2f(x_1, x_2)$ then DRS. Or, more generally, for any scale factor t , if $f(tx_1, tx_2) = tf(x_1, x_2)$ then production is CRS; if $f(tx_1, tx_2) > tf(x_1, x_2)$ then IRS; if $f(tx_1, tx_2) < tf(x_1, x_2)$ then DRS.

There is a moving company called "2 Guys and a Truck" which pretty much delineates their production function to move stuff. Could 4 guys & 2 trucks move twice as much stuff? 6 guys & 3 trucks? That's returns to scale. How much more could be moved by 3 guys & 1 truck? 2 guys and 2 trucks? That's marginal product.

Short Run vs Long Run

Often one input is more flexible than the other. This means that our analysis should distinguish between the short run (when one input is fixed) and the long run (when both inputs are flexible). Often we assume that labor is flexible and capital (the machines) are fixed since

building, say, a new assembly line takes time. But other firms might have different rankings – universities have tenured faculty, many of whom have been there longer than some of the buildings on campus!

Profit Maximization

A firm's profits are revenues minus costs, so a firm selling n different output goods, each for price p_i and using m different inputs, each with cost w_j , would have profit $\pi = \sum_{i=1}^n p_i y_i - \sum_{j=1}^m w_j x_j$.

First note that the costs must all be put into the same units – dollars per time unit. Which raises the question, if a firm buys, say, a truck that is expected to last for 5 years, how is this cost compared against the daily wage of the person driving it? To answer this we suppose that another company were set up that just rented out trucks: it goes to the bank, gets a loan to buy the truck, and then charges enough per day to pay off the loan per day. We consider that, even if a company doesn't actually rent the truck but actually buys it, that it *could have* rented the truck. So the rental rate is the correct cost of that capital good. In the real world more and more companies are separating their daily operations from their loan portfolio and renting equipment. If you work at an office, you know that most photocopiers are rented. Airlines rarely own their own jets, they rent them. Offices are usually rented space. (Employees are rented, too!)

The companies have figured out that correctly measuring costs allows them to make better decisions. Capital goods which are owned and given away internally as if they had zero cost are not efficiently used.

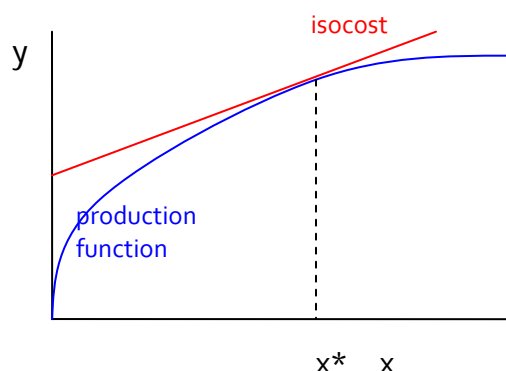
Economists also measure costs differently from the way that accountants do regarding payments to shareholders/owners. If a public company has an IPO and sells its shares for \$100 each, then those shareholders expect something in return. They expect that the dividends (and/or capital gains) will return them as much or more money as if they had invested their \$100 in some other venture. So the company had better return \$8 per year if the investors could have gotten 8% returns. An accountant would count this \$8 per share as a "profit" but economists see that as a cost to be paid to shareholders for the use of their money (their capital). If the firm were to return just \$6 then the shareholders would be angry and the firm would be in trouble; if the firm returns \$12 then the shareholders are delighted.

So economists often talk about "zero profits" being a general case, which makes people wonder how much economists know about the real world since any business newspaper daily reports companies making "profits". But we're just counting different things. If the regular return to capital is 8% then, if the firm makes \$8 that the accountants call profit, we call it a cost and report that the firm made zero economic profit.

Profit Maximization with One Input

This means

$\max_x \pi = py - wx$ subject to $y = f(x)$. Hiring one more unit of input will raise the firm's cost by w ; this one additional unit of input will raise output by MP and so revenues will rise (assuming no market power) by $p \cdot MP$, which is the value of the marginal product. If $p \cdot MP > w$ then the firm should hire more inputs; if $p \cdot MP < w$ then the firm should hire fewer inputs; so in equilibrium $p \cdot MP = w$.



Note that, if one input is fixed, then even the two-input model becomes, in the short run, just this problem of profit maximization with one input.

Firms of course look at the price that they actually pay, not necessarily the price that someone else thinks they ought to pay. For instance, before laws against pollution, disposal of a firm's waste was free – it just dumped the waste into a river or something. (Disposal is, in a sense, an input into production since the firm can't make more stuff until it cleans up from the day before.) So of course a firm would have no incentive to reduce waste. But once there were laws about pollution (e.g. Clean Air Act, Clean Water Act, RCRA Resource Conservation & Recovery Act, CERCLA Comprehensive Environmental Response Compensation and Liability Act - Superfund), a firm might have to pay a specialist disposal firm to take it away. Now there is a price on this input so the firm has an incentive to limit the amount disposed.

Profit Maximization with Multiple Inputs

Consider a firm which has multiple inputs available for making the output, each of which is useful and productive. Each input has a cost (or wage, if we extrapolate from the case of hiring workers) denoted w_i .

As with the consumer's diminishing marginal utility, we agreed that the firm faces diminishing marginal productivity; the production function is $y = f(x_1, x_2, \dots, x_N)$ and the marginal product of each input is the partial derivative, $MP_i = \frac{\Delta y}{\Delta x_i} = \frac{\partial y}{\partial x_i} = \frac{\partial f(x_1, x_2, \dots, x_i, \dots, x_N)}{\partial x_i}$. The firm is to

$\max_x py - w_1x_1 - w_2x_2$ subject to $y = f(x_1, x_2)$.

Also as noted previously, the fact that each individual marginal product is diminishing does not mean that production overall has diminishing returns to scale – where 'scale' refers to a case where all of the relevant inputs are increased. As a simple example, most offices generally operate with each employee getting a computer. Buying more computers without hiring more people might increase output, but at a diminishing rate; the same would hold true for hiring more people without getting more computers. But getting more of both could allow the business to expand.

The firm will maximize profits by choosing inputs such that (in the long run), the ratios of $\frac{MP_i}{w_i}$, marginal productivity per cost of each input, is equal. The explanation should, by now, be typical: if spending \$1 more on input i increased output by more than spending \$1 more on input j, then the firm should decrease spending on input i while increasing spending on input j. This will not only allow the firm to make more output more cheaply but also tend to bring down the marginal productivity of input j while increasing the marginal productivity of input i, so that in equilibrium we have $\frac{MP_i}{w_i} = \frac{MP_j}{w_j}, \forall i, j$.

If one input has a price which is increased (say, by some environmental regulation) then this input will be used less. This is the substitution effect (see from marginal condition).

There is also a Scale Effect. As the cost of production rises, the quantity of output demanded will fall, so fewer of all types of input will be demanded.

Also, if that input is non-excludable like polluted air or water, then as one industry shrinks, other industries could see their costs fall, so input used more – a different cross-industry substitution effect. Also a different scale effect depending on conditions in that industry.

Cost Minimization/Profit Maximization

The firm's problem to maximize profits generates a dual (sometimes easier) problem, which is how to minimize costs subject to a constraint of making a particular amount of output. (The constraint is important, though -- if the firm wanted to minimize costs without that constraint, clearly setting $y=0$ would be best!)

If the firm wanted only to maximize revenue (or if the input were costless) then the firm would $\max_x \pi = py$ subject to $y = f(x)$. A "costless" input sounds crazy but remember that Mankiw says, "rational people make decisions on the margin." Zero marginal cost is not that unusual: it's the usual condition for media companies – they pay a big fixed cost (to record a song or make a movie or TV show), but then their marginal costs are about zero: one more iTunes download does not cost them much!

Marginal Revenue

To figure out this problem of zero marginal cost but changing revenue, we just need another definition. Marginal Revenue, MR, is the change in revenue per change in output, $MR = \frac{\Delta Rev}{\Delta y}$.

If price is a function of the level of output (e.g. the firm has monopoly power) then MR can be a complex function. If the firm operates in a competitive market then the price is outside its control, so the increase in revenue from selling one more unit of output is the price, p .

But many firms have monopoly power. Consider a fashion label selling handbags. They want to sell more because that means more revenue. But they also know that scarcity has value: some designers sell in only a few select boutiques and can charge very high prices; other designers sell bulk quantities in department stores and cannot charge high prices. This is a very general problem.

A firm that wanted just to maximize revenue would expand production as long as $MR > 0$ and only stop when $MR = 0$, when producing more output would no longer raise its Revenue.

Most firms, however, do not simply care about maximizing revenue; they want to maximize profits. (Particular parts of firms, however, might want to maximize revenue: for instance, most sales people are paid commissions on the sales they generate not necessarily the profits. Countrywide got paid per mortgage regardless of quality.)

A firm that wants to maximize profits will also have to take account of Marginal Cost. It is also convenient to figure other cost definitions.

In a perfectly competitive environment the firm's demand curve is a horizontal line – the market price for this homogenous commodity. The firm can sell as much quantity as it can produce at that price. Lowering the price will not increase the amount sold; raising the price will stop all sales. This is an extreme assumption but, as you think about it, not completely unrealistic. One of the important points of any business plan for a new company is 'the competition' – what other firms sell and at what price. If my firm offers the same product then I can't charge a higher price. Of course my firm might charge a higher price and offer a better product, but this means that I'm selling a different output.

The business press sometimes discusses the "commodification" of different markets: for example at one point, when computers were new, the chips were quite different and there were many different prices; now that chips are standardized there is much less variability in price. Financial markets are commodity markets: nobody offers a "20% sale" on stocks or bonds! Markets are organized in order to standardize and "commoditize" certain products: e.g. the CBOT offers corn futures to deliver "5000 bushels of No. 2 yellow corn"; or "100 troy ounces of refined gold not less than .995 fineness cast as one bar or 3 one-kilogram bars"; the NYMEX trades gasoline "reformulated gasoline blendstock for oxygen blending (RBOB) futures contract trade in units of 42,000 gallons (1,000 barrels)" for delivery in New York.

This is not to say that the market demand curve is flat, just that the curve for the particular firm is flat. Another way of thinking of it is that the firm is such a tiny player in the whole market that it sees just a tiny piece of it, which is essentially flat. But even if there are a limited number of companies then a firm might still face a flat demand curve for its product based on the other's price. A final example: whatever I sell nowadays, I have to know what Amazon or eBay charges for it – most all of my customers will!

Hicks-Marshall Rules of Derived Demand:

Demand for input is more elastic when

1. technical substitution is easy
2. input cost share is high
3. input substitutes are supplied elastically
4. demand for output is elastic

So putting the Scale effects and Derived Demand effects all together gets complicated. What is the impact of pollution restrictions on a firm, hindering its use of a particular dirty input? Clearly this adds to costs of the dirty input, but the impact of this cost change could be small or large depending on application of the Hicks-Marshall rules. Then what is the impact on other inputs (usually labor, i.e. jobs)? If the cleaned-up input is more labor intensive then this could mean a net rise in jobs; oppositely if the cleaned input is more capital intensive. If the cleaned input is more labor intensive, but the rise in cost greatly diminishes the demand for the product, then net jobs could fall if industry output falls. If the avoided pollution makes other factors more productive, then there could be further effects.

Might, for example, want to know the impact of a carbon tax on power plants. Here the share of inputs in total cost is clearly substantial; demand for output is inelastic – that's easy. Technical substitution is happening in the long term (few new coal plants, many more natural gas plants that are dramatically cleaner) but in the short term is limited. Input substitution is complicated in long/short run: natural gas production from domestic sites is increasing while facilities for imports are limited; oil is not much better; nuclear plants are tough to build; new hydro or geothermal is limited; other power sources are available but with limits (e.g. uranium production, solar panels now often imported, biomass facilities, etc.). Natural gas shows the inter-relationship of #1 and #3 in Hicks-Marshall rules: new gas turbines are relatively easy to install but if there is a rush of new construction then natural gas prices (which have recently moderated) will rise again and the cost advantage of natural gas over coal would erode.

Supplementary Material for Advanced students

Cobb Douglas Example

Consider a numerical example of a firm with a very simple Cobb-Douglas production function, so $y = \sqrt{x_1 x_2}$

so the marginal products are $MP_1 = \frac{1}{2} \frac{x_2}{\sqrt{x_1 x_2}} = \frac{1}{2} \frac{\sqrt{x_2}}{\sqrt{x_1}} = \frac{1}{2} \frac{y}{x_1}$ (the last equation comes from a convenient simplification; it's a bit of a trick that's not obvious the first time you see it but you should be able to verify that it is, indeed, correct) and $MP_2 = \frac{1}{2} \frac{y}{x_2}$. The firm's cost is $c(y) = w_1 x_1 + w_2 x_2$. So put these expressions for

the marginal products into the firm's marginal conditions that $\frac{MP_1}{w_1} = \frac{MP_2}{w_2}$ so $\frac{\frac{1}{2} \frac{y}{x_1}}{w_1} = \frac{\frac{1}{2} \frac{y}{x_2}}{w_2}$ so

$w_1 x_1 = w_2 x_2$ or $x_2 = \frac{w_1}{w_2} x_1$. Put this into the production function to solve for $y = \sqrt{x_1 \frac{w_1}{w_2} x_1} = x_1 \sqrt{\frac{w_1}{w_2}}$ so

$x_1 = y \sqrt{\frac{w_2}{w_1}}$ and $x_2 = y \sqrt{\frac{w_1}{w_2}}$ -- these are the input demand functions, giving the firm's demand for each

input as a function of both its output level and the relative price of the input. Put these into the cost function to find that the long-run cost is $c(y) = y \sqrt{w_1 w_2} + y \sqrt{w_1 w_2} = 2y \sqrt{w_1 w_2}$.

The long-run average cost is thus $\frac{c(y)}{y} = \frac{2y \sqrt{w_1 w_2}}{y} = 2\sqrt{w_1 w_2}$ which sets the price in the market.

This allows us to easily see the scale and substitution effects. If the cost of an input, say w_1 , rises, then this will mean that directly less of that input is used since $x_1 = y \sqrt{\frac{w_2}{w_1}}$; this is the substitution effect. However it also

raises the firm's average costs, $\frac{c(y)}{y} = 2\sqrt{w_1 w_2}$ in the long run so the amount sold must decrease. The size of this decrease in output depends on demand elasticity: if the output is elastically demanded then the price rise will produce a sizeable downward shift in quantity demanded.

Costs

The cost function is $c(w_1, w_2, y)$. It is determined, in the long run, by the wages of each input and by the level of output chosen. In the short run it is also determined by the amount of the fixed factor.

Marginal Cost is the change in cost per change in output, $MC = \frac{\Delta c}{\Delta y}$. Marginal cost is not generally constant but is commonly considered to vary with output.

We define several other costs:

Average cost, AC, is the cost per unit of output, $AC = \frac{c(y)}{y}$. In the short run, some costs are fixed (F) and some are variable ($c_v(y)$).

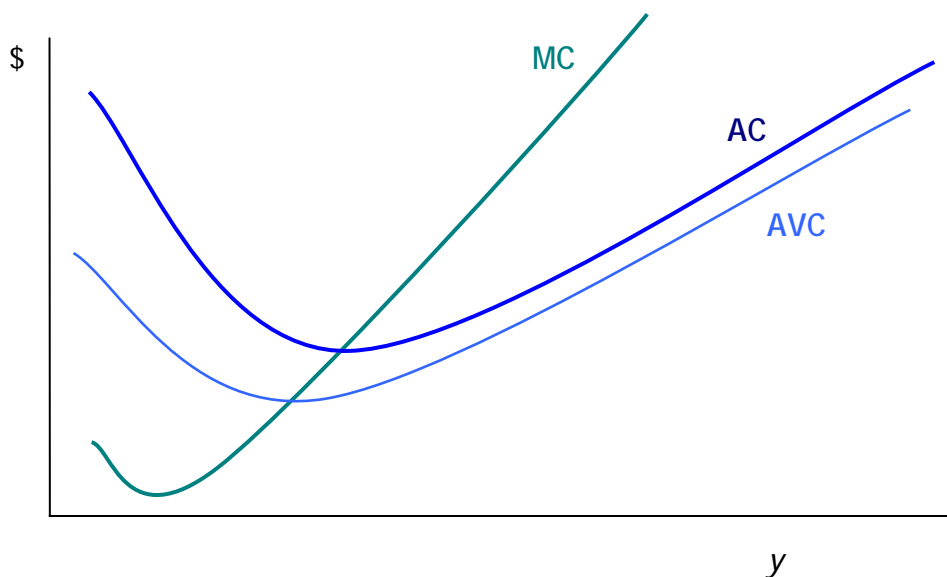
Average variable cost, AVC, is the variable cost per unit of output, $AVC = \frac{c_v(y)}{y}$.

Average fixed cost, AFC, is $\frac{F}{y}$, but since F does not change, this is just a rectangular hyperbola and doesn't change much – so we rarely pay much attention to AFC. However we note that $AC = AVC + AFC$.

Also, from the definition of marginal cost and of fixed cost, we note that there is no need to define both marginal total cost and marginal variable cost – since fixed costs don't change, marginal fixed costs are always zero so marginal total cost and marginal variable cost are

always identical: $MC = \frac{\Delta c(y)}{\Delta y} = \frac{\Delta c_v(y)}{\Delta y}$.

These SR curves are typically graphed as:



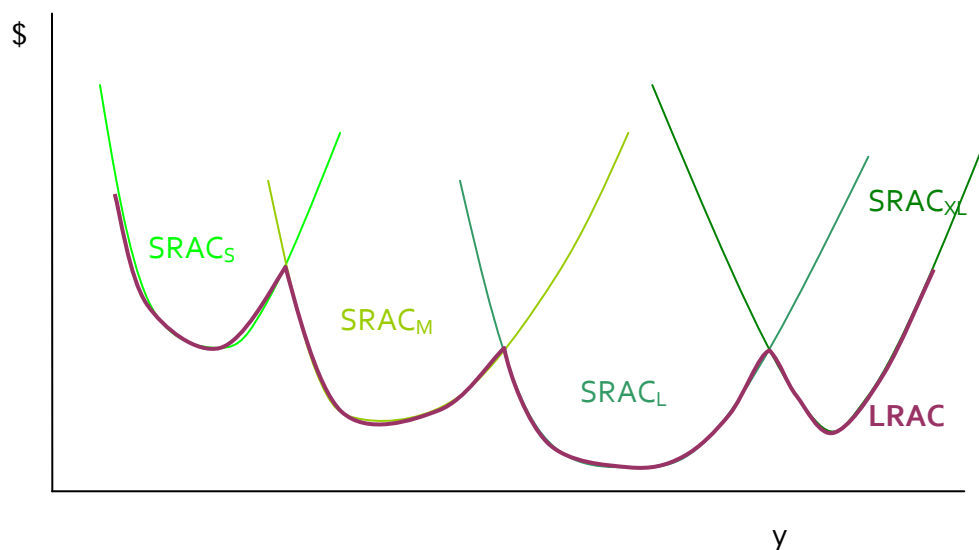
Where we note that MC must intersect AC at the minimum point of AC; also MC must intersect AVC at the minimum point of AVC. To show this, note that by definition if $MC > AC$ then AC must rise; if $MC < AC$ then AC must fall. The minimum point of AC is where it turns from falling

to rising, where it, for at least a short (infinitesimal) time it is neither rising nor falling so $MC=AC$. Same argument goes for AVC .

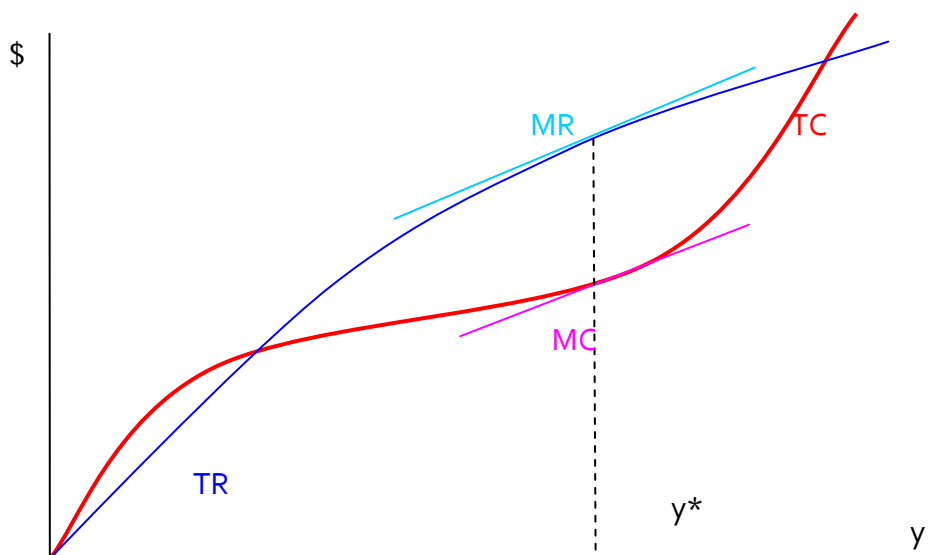
In the long run, there are no fixed costs, so long-run average costs (LRAC) are equal to long-run variable costs. LRMC are defined analogously to the short run.

LRAC can never be greater than the short-run AC curves – having more choices can never hurt profits!

In Long Run there are no fixed costs (can always choose zero output at zero cost). LR AC curve is envelope of SR AC curves – with a scalloped edge if there are discrete plant sizes but, as plant sizes become continuously variable, the LRAC becomes a smooth curve.



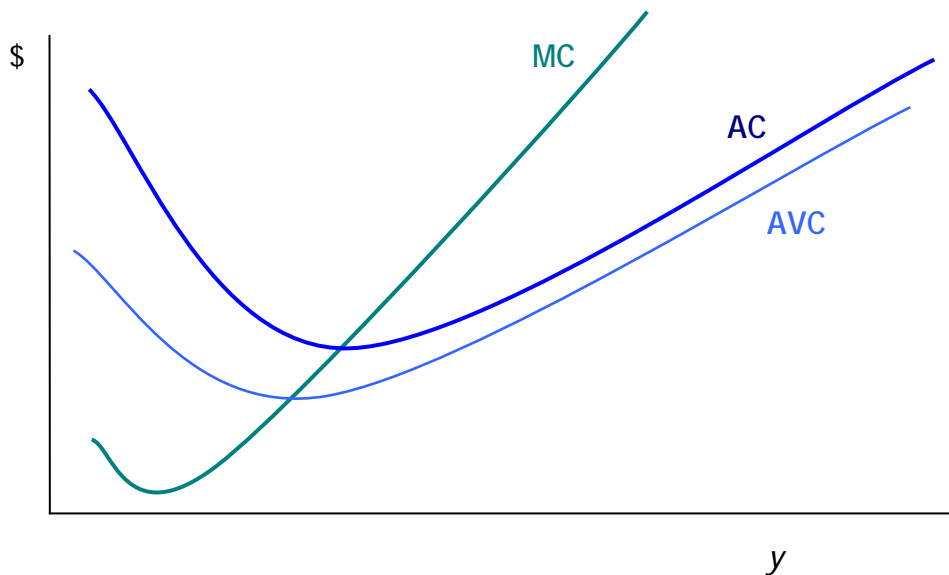
To maximize profit the firm will set $MR = MC$. Consider this graph, where MR is allowed to vary as well as MC . To maximize profits, the firm wants to find where TR is farthest away from TC . The usual argument applies: making and selling one more unit of output raises revenue by MR ; making one more unit of output raises costs by MC . If $MR > MC$ then this was a good choice for the firm and it should raise output more. If $MR < MC$ then this increased output was not a good choice and it should decrease output. It will stop this changing and reach equilibrium where $MR = MC$. In perfect competition where $P = MR$, this gives us the equilibrium condition $P = MR = MC$.



Profit Maximization in the Short Run

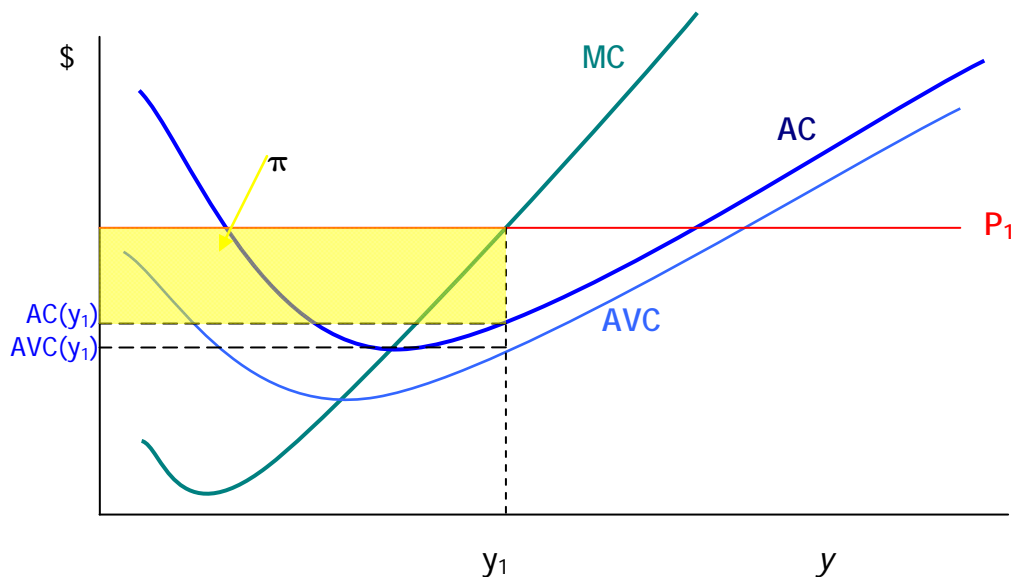
In the short run the firm must account separately for fixed costs. The profit maximization becomes: $\max_y py - c_v(y) - FC$. Note that variable costs, $c_v(y)$, are a function of the level of output but "Fixed Costs are fixed." Fixed Costs don't change depending on the level of output, they're fixed. So the lowest level of profits that the firm could make are $-FC$ if output is set to zero (revenue and variable cost both become zero in that case). Beyond this, however, the usual rules apply. Recall that the marginal variable cost is exactly equal to marginal total cost. MC is how much cost increases when output increases. MR (which we assume to be p in this case, for simplicity) is the amount by which revenue increases when output rises. Again, if $MR > MC$ then the firm will produce more; if $MR < MC$ then less.

So a firm with these cost curves (which we describe as canonical):



could face prices that lie in 3 separate regions: (A) either price intersects MC where MC is above AC; or (B) price intersects MC where MC is below AC but above AVC; or (C) price intersects MC where MC is below AVC.

Consider (A):



in this case, when price is at P_1 , the firm will choose y_1 level of output to maximize profits.

Profits are $\pi = py - c(y) = py - c_v(y) - FC$ but can be more easily seen graphically as

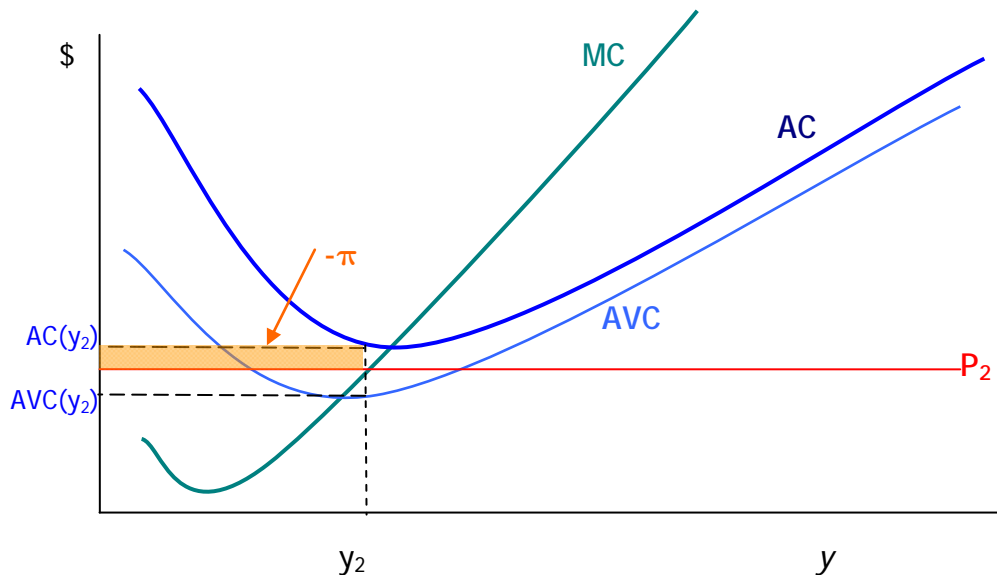
$$\pi = y(p - AC(y)) = y\left(p - AVC(y) - \frac{FC}{y}\right) = y(p - AVC(y)) - FC. \text{ So profits are drawn as}$$

the area of the rectangle with height $(p - AC(y_1))$ and width y_1 , marked yellow in the graph.

The decomposition of costs into Variable Costs and Fixed Costs can also be seen: VC are the area of the rectangle with height $AVC(y_1)$ and width y_1 ; FC is the area of the rectangle with

height $AC(y_1) - AVC(y_1) = \frac{c(y_1) - c_v(y_1)}{y_1} = \frac{FC}{y_1}$ and width y_1 .

With price at (B), where it intersects MC where MC is below AC but above AVC, we find:

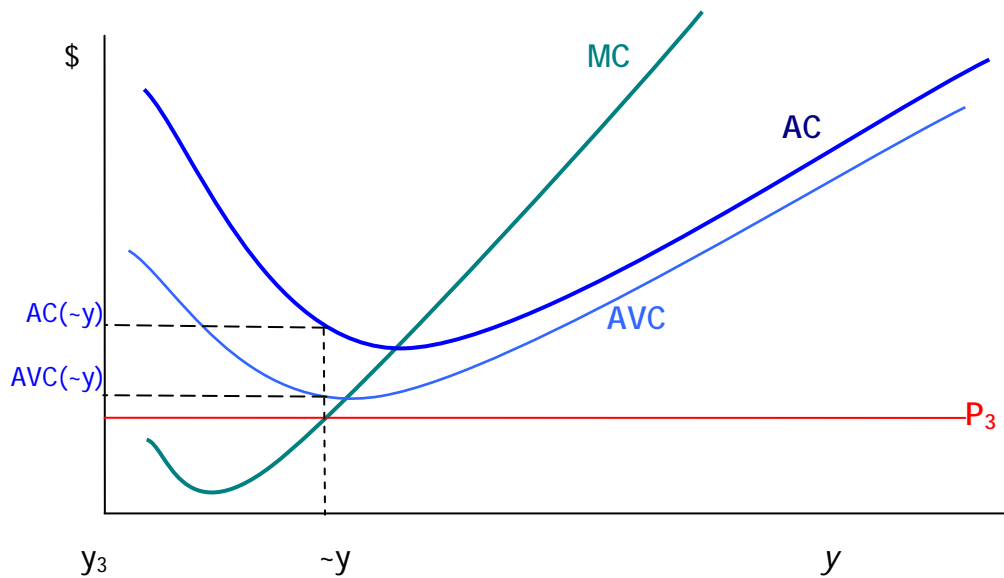


Again the firm chooses the point where $p = MC$, which is y_2 . Profit here is actually negative, since $AC(y_2) > p_2$, the firm's average costs are greater than average revenue. So the question arises, is the firm really maximizing profit? Well, what else could the firm do? It could shut down completely, but in this case it would lose $-FC$, which, in the diagram, is again the area

of the rectangle with height $AC(y_2) - AVC(y_2) = \frac{c(y_2) - c_v(y_2)}{y_2} = \frac{FC}{y_2}$ and width y_2 -- and

this rectangle is clearly bigger than the actual profit lost. Basically, since operating costs (AVC) are below the price, it makes sense to operate even if the firm doesn't cover all of its fixed costs. It covers some amount of its fixed costs and so reduces its losses. This is just another manifestation of the old rule: sometimes the best that we can do still isn't very good. The firm is maximizing profits but still losing money.

Only in the case of (C) would the firm actually shut down. Consider:



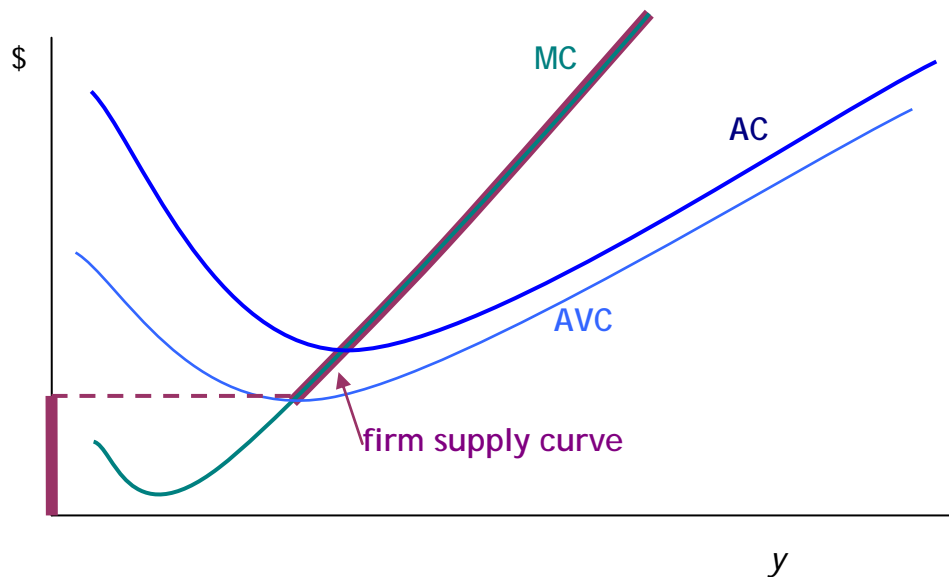
Now with price at p_3 , the firm could choose to continue producing where $p = MC$, at the point labeled $\sim y$ (from computer programming, the tilde means "Logical Not"). But at this point, the losses to operating, totaling $y(AC(\sim y) - p_3)$ would be larger than the losses to just closing down and losing fixed costs, the rectangle of area $y(AC(\sim y) - AVC(\sim y))$. So the firm chooses $y_3 = 0$ whenever the price is less even than average variable costs.

This tripartite division has many real-life ramifications. This is why hotels and airlines are willing to give last-minute deals: a butt in an airplane seat paying even \$20 is more than the extra cost for the jet fuel to haul that little bit more weight. They try and try to charge more, to cover their fixed costs, but when it comes right down to the end they know that their variable costs are low.

Another ramification is seen in the housing bust. Driving through certain neighborhoods, there were still houses being constructed – why? Clearly the answer is that, since the builder has bought the land (usually the most expensive part), that FC is lost now. If a house sits half-built then the construction company loses all the costs put into it so far. But even if completing the building takes more expense, it might still be worthwhile – the builders will lose money, but not as much as if they walked away.

Firm Supply Curve

The firm's supply curve is then the locus of price and quantity choices, which is the firm's MC curve above AVC, then quantity jumps to zero if the price falls below this point. In the graph, this is:

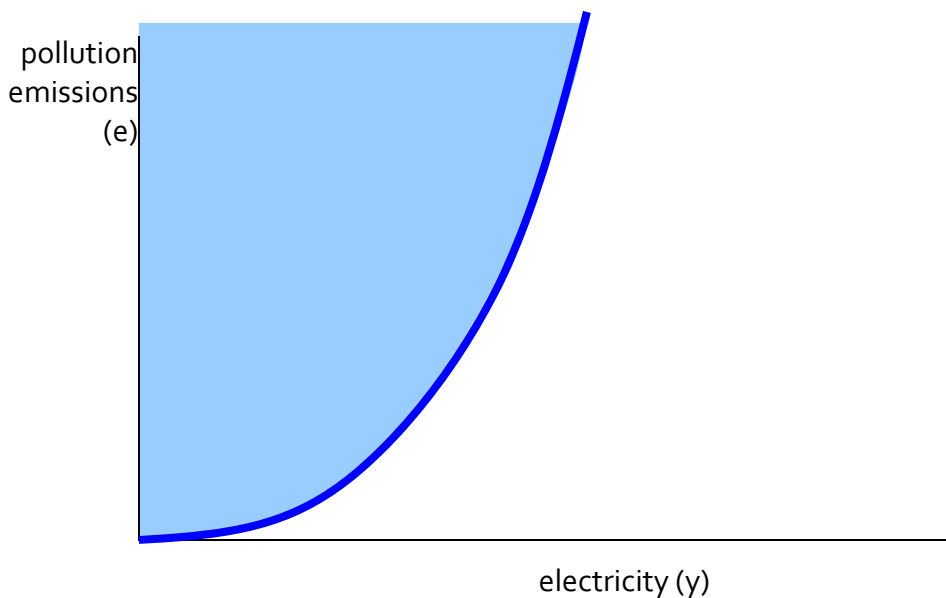


Examples of tax affecting either FC or MC depending whether it is per production plant or per unit of output. If pollution is per output but tax/regulation is per plant, this could have mixed effects.

Production Externalities

In the simplest case, we can examine a firm making a single private (rival and excludable) output and incidentally a single public (nonrival and nonexcludable) output (for now, we assume that this public good is disliked). An easy example could be a power plant which makes electricity and pollution. (Actually a variety of sorts of pollution, which affect different groups of people: carbon, mercury, NOX, and sulphur dioxide are the main ones.)

In this case the production can be shown as being like a production possibility frontier but with the pollution increasing along with the output, something like:

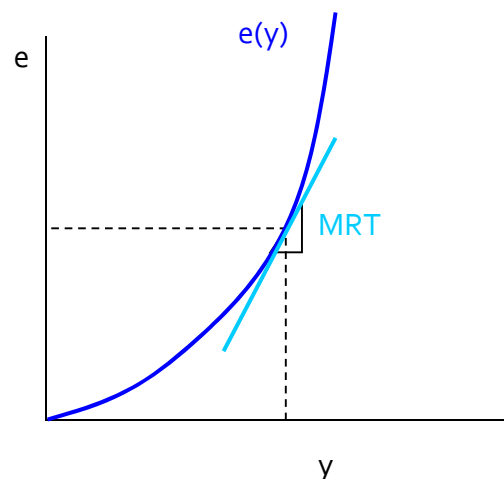


The firm can choose any combination of electricity & pollution within the light blue area. Clearly, however, the firm would be foolish to choose a point inside the area; the points at the dark blue line are efficient. These are the production possibility frontier. They are efficient because there is no way to increase the output of electricity without also increasing the output of pollution (this would not be true for points in the interior).

At any point along the frontier of production possibilities, we can define the marginal rate of transformation as the change in output of pollution per change in output of electricity – the slope of the line. With the notation of e for pollution emissions and y for the output of the firm, the marginal rate of transformation,

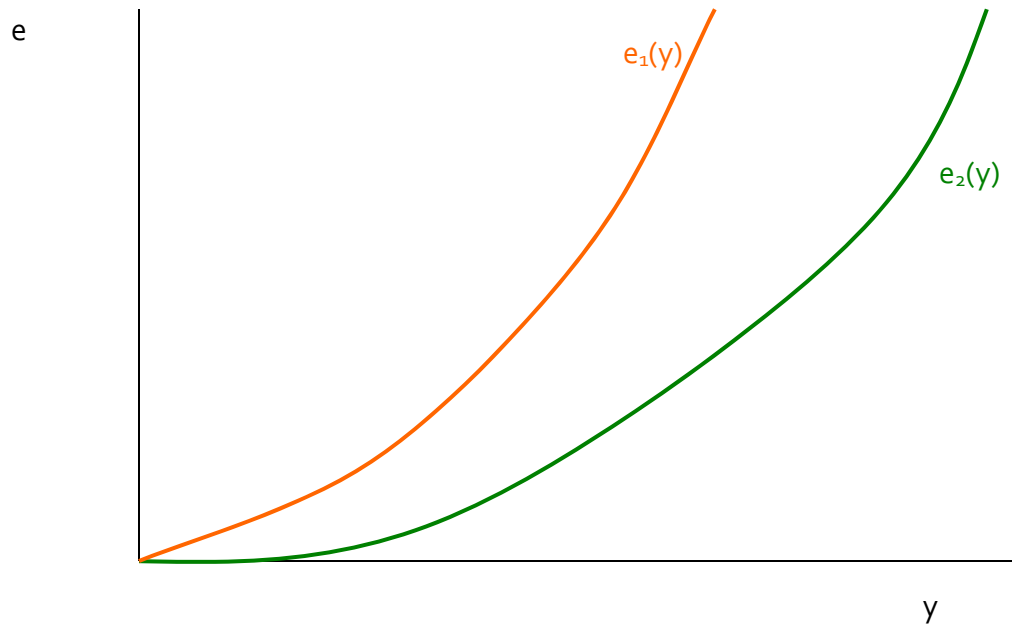
MRT, here is $MRT = \frac{\Delta e}{\Delta y} = \frac{de(y)}{dy}$, where

$e(y)$ is the function linking the amount of emissions generated as determined by the amount of output produced. We can think of electricity generation as transforming some amount of a public good (in this case clean, unpolluted air) into a private good (electricity).

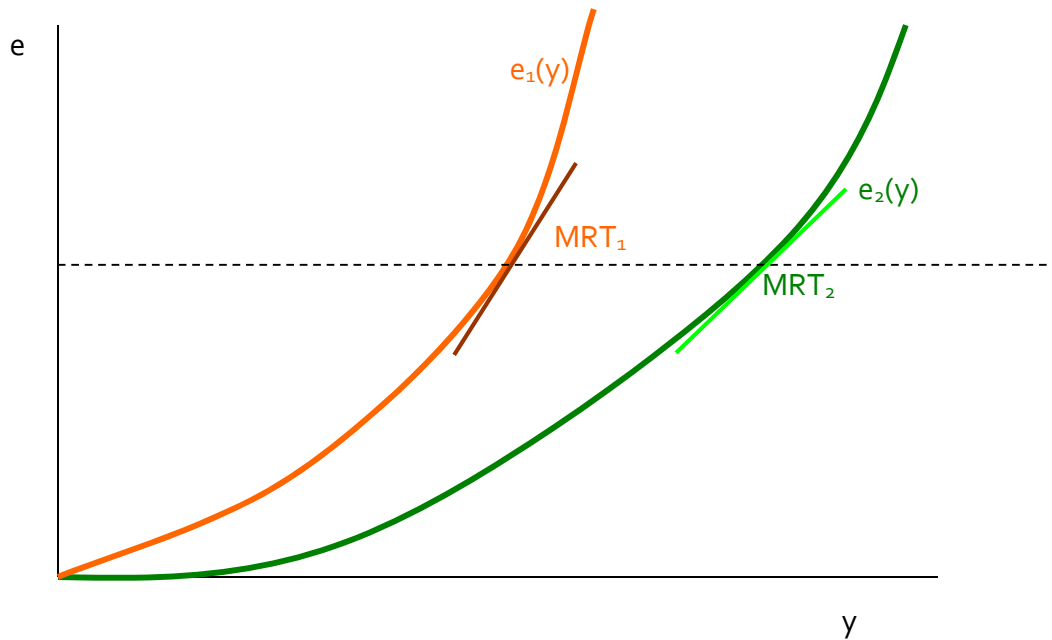


This interpretation of the choice along the production possibility frontier as representing a choice of marginal rate of transformation allows us to compare firms and make statements about the relative efficiency.

Suppose there are two firms which, for some reason or another, have different emissions per unit of output. Graphically this would be represented as:

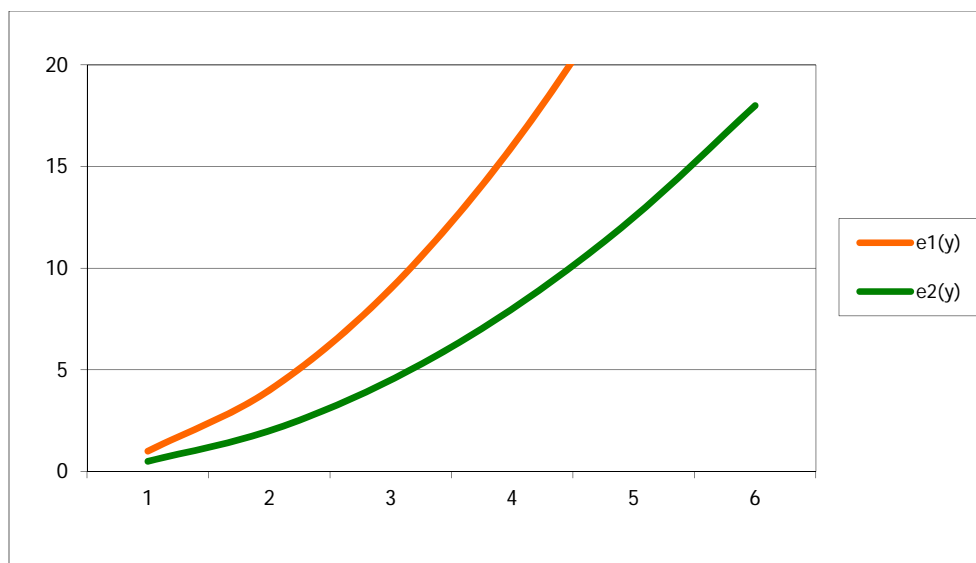


If they each produced the same amount of emissions, they would of course be able to generate different output levels, but their marginal rates of transformation would also be different.



Clearly the marginal rate of transformation of firm 2 is lower than the marginal rate of transformation of firm 1. This means that when firm 2 generates one more unit of output, it creates fewer emissions than firm 1 does. This means that, if firm 2 were to make one more unit of output while firm 1 made one unit less – keeping the total output of the two firms at the same level, the increase in emissions from the second firm would be (in absolute value) less than the decrease in emissions from the first firm. So total emissions would be smaller even though the output was kept constant.

Consider a simple numerical example, where $e_1(y) = y^2$ but $e_2(y) = \frac{1}{2}y^2$. This is plotted as:



If emissions of each firm are 16, then firm 1 is producing 4 units of electricity while firm 2 is producing 5.66 units of electricity. If firm 2 produced one more unit of electricity its emissions would rise to 22.16, an increase of 6.16. If firm 1 produced one less unit of electricity its emissions would fall to 9, a decrease of 7. So if, instead of both firms producing 16 units of emissions, firm 1 produced less and firm 2 produced more, the overall production of electricity could remain constant while emissions fall.

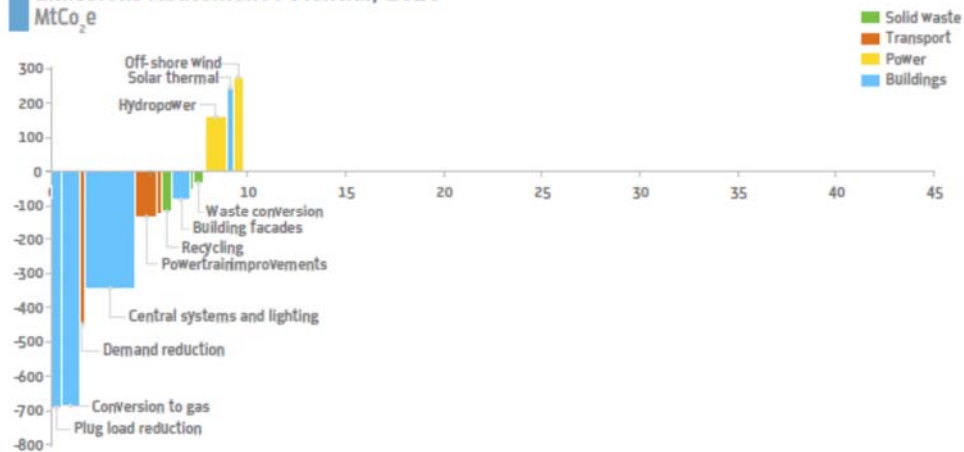
We can continue this trade-off as long as the marginal rates of transformation are unequal. It is only when the marginal rates of transformation are equal that there will be a total efficient way of getting the most output with the least amount of harmful emissions. With a bit more math, we can find the point where the MRTs for each firm will be equal.

When we get to policy (next), we return to this idea: at the most efficient point, the marginal rates of transformation will be equal – which will not necessarily be the point where emissions are equal.

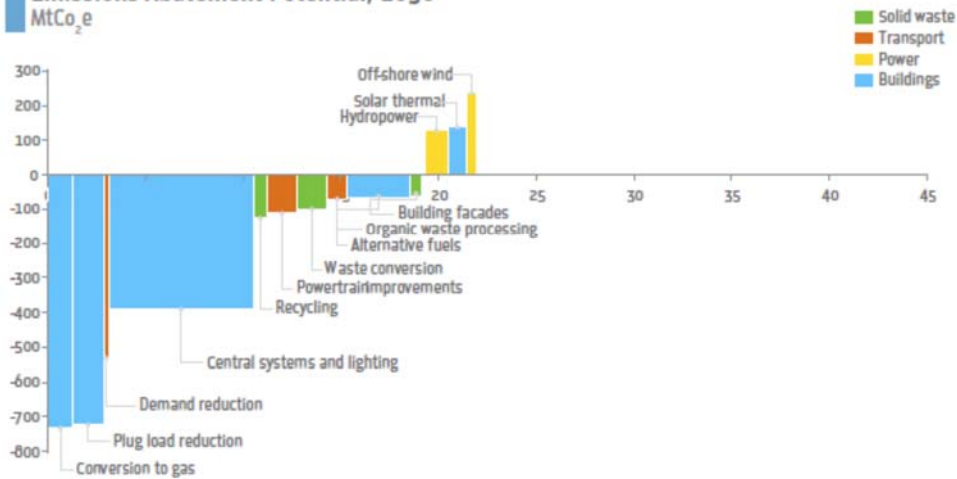
Marginal Abatement Costs

Example from PlaNYC

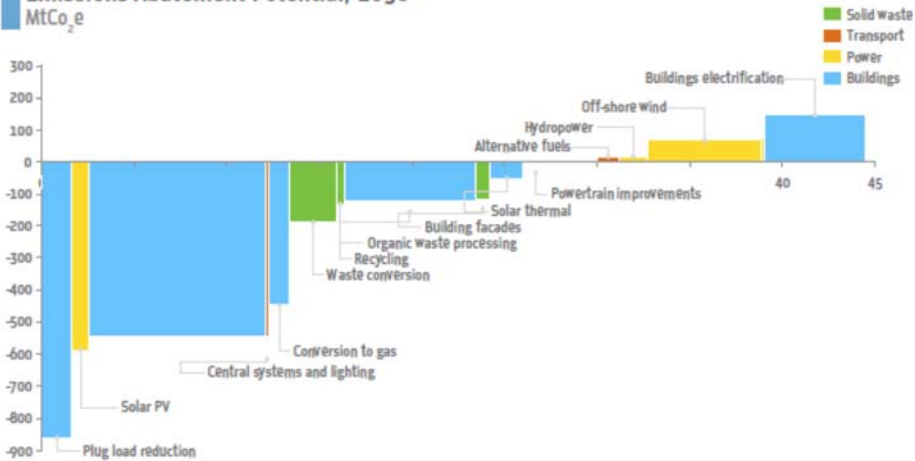
Emissions Abatement Potential; 2020 MtCo₂e



Emissions Abatement Potential; 2030 MtCo₂e



Emissions Abatement Potential; 2050 MtCo₂e



Source: NYC Mayor's Office

From p.117 of City of New York, *New York City's Pathway to Deep Carbon Reductions*, Mayor's Office of Long-Term Planning and Sustainability, New York, 2013.

Regulation of Pollution

Command & Control

good because:

- flexible in complex processes (law of unintended consequences)
- more certainty for producers

bad because:

- need so much information
- low incentives for innovation
- inefficient since generally violate equimarginal principle

Other refinements & policies:

subsidies might occasionally have some "bonus" or increasing returns provision, as with land use: a landowner who converts land to park gets more subsidy if it borders on an existing park; this is useful if the marginal benefits (to species habitat) are increasing in contiguous land area

from Law & Econ, we know that 100% monitoring is inefficient, better to catch some portion of offenders (e.g. 1 in 10) but fine them extra (e.g. 10 times as much)

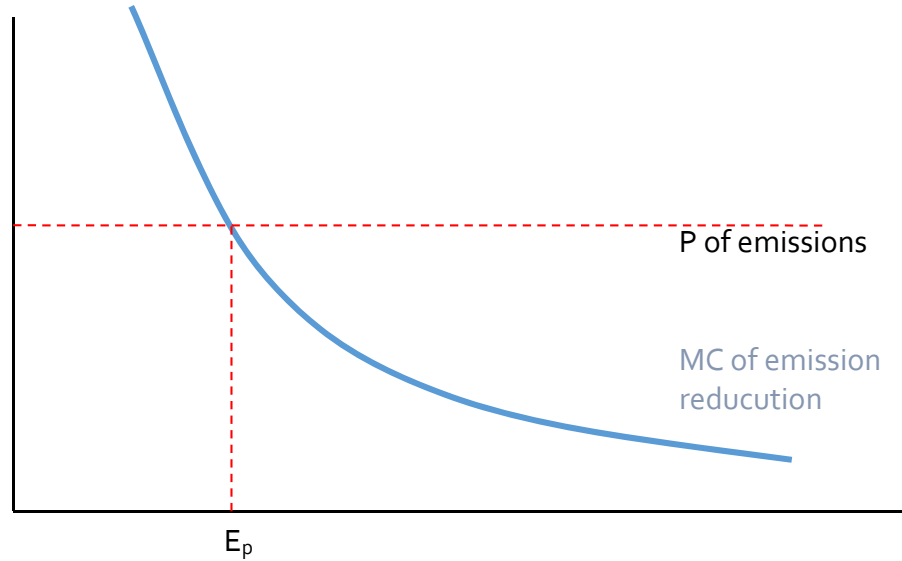
could use "performance bonds" but these need careful monitoring (since generally forfeiture of bond involves lengthy legal proceedings); sometimes used on surface mining; also impose liquidity costs (extra financing needed) and open 'moral hazard' for regulators – 'pay-to-play'

Fees & Tradable Permits

A fee or tax per unit of emission is the Pigouvian solution – set the price and let firms decide. Tradable permits can give an equivalent outcome – permits are sold for a price; this price is essentially a tax.

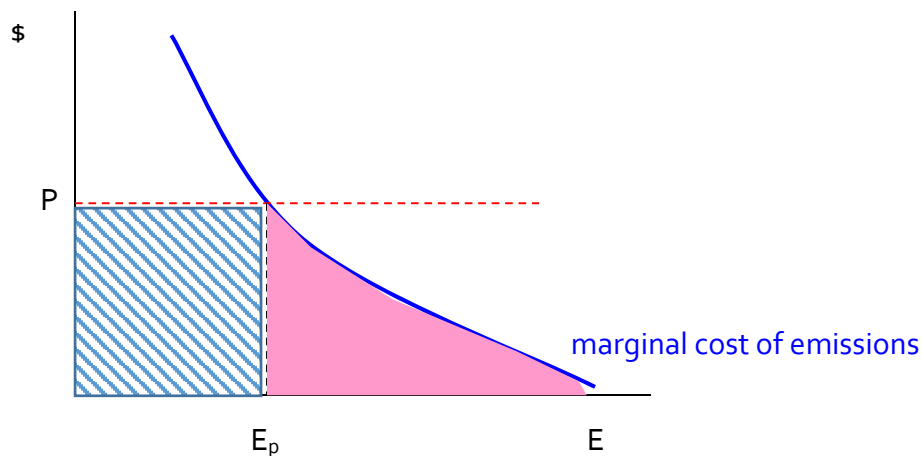
Tradable Permits

Can easily show the financial burden on firms. Consider first the simple case: tradable permits sell for a price, P . At this price the firm chooses emissions of E_p .



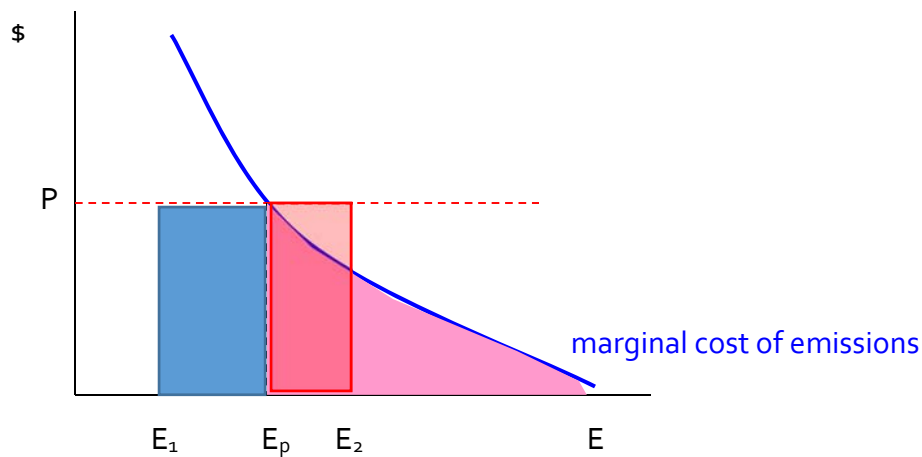
Because, if the firm emitted more than E_p , it could cut emissions at low cost and sell permits at a higher cost; if the firm emitted less than E_p , it is cheaper to buy permits than cut emissions itself.

If the firm were given no permits at all, it would cost



Where the pink triangle is the cost of compliance: the emissions that the firm cuts back in response to the permit regime. The striped box shows the cost of buying E_p permits at price P . If the firm were given a few permits, E_1 , insufficient for its needs, then it would face the same cost of compliance but now:

Under this regime the firm gets E_1 permits and so only needs to buy the remaining permits, $E_p - E_1$. Or the firm might even get extra permits, which could reduce its costs below the cost of compliance:



Now the firm gets E_2 permits, of which it sells off $(E_2 - E_p)$ for a profit, which mitigates some of the cost of compliance.

If the permits are allocated only once (for example, when the policy is begun), then the dynamic effects of keeping unprofitable firms going (noted in the section on subsidizing firms not to pollute) will be small. If there are regular allocations of permits (yearly, for example) then these dynamic effects will be larger.

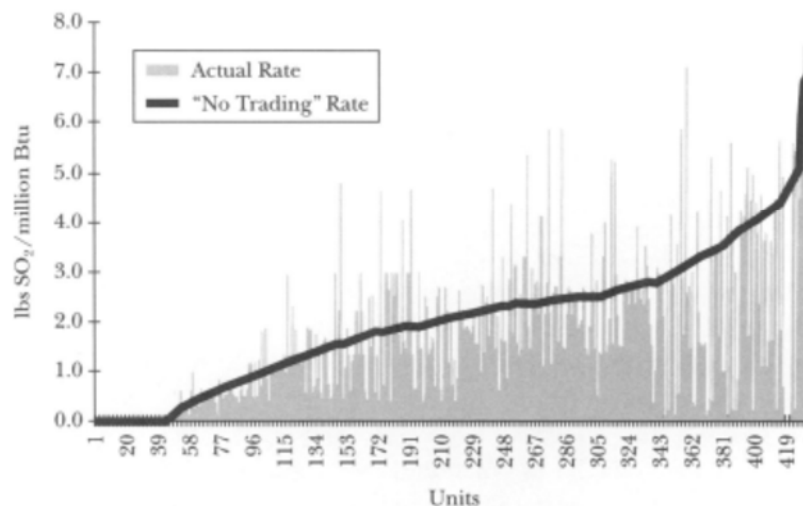
If the permits are given out in proportion to past emissions then firms will have an incentive to raise emissions just before the law goes into effect. Since many laws are debated for quite a while before taking effect, this is relevant. A law might have a multiyear lookback period.

Giving out permits in proportion to past emissions is also discriminatory to new entrants. If we consider policies like carbon permits to mitigate global climate change then this would mean that emerging economies would get fewer permits relative to richer countries.

Nonetheless giving out permits is common because it might make the program politically feasible: existing firms are given these valuable permits to get them to accept the new standards.

These worked really well, when US instituted tradable permits for sulfur dioxide (SO_2) in 1995 (see Schmalensee et al, Journal of Economic Perspectives 1998). They show this graphic, where the heavy line shows historical emissions per plant (sorted by level) while the light bars show actual emissions.

Figure 2
Effect of Trading: 1996



Source: Derived from Pechan (1995), EPA's EMS, and EPA (1996).

Clearly there was enormous variation: some plants drastically reduced emissions, far below what was required; others increased. The variation gives an idea of the scope of DWL from regulation.

Tradable Permits are usually allocated by either

- giving them away to current polluters (usually proportional to current pollution amounts, sometimes called 'grandfathering')
- auctioning them off to the highest bidder

Further details of cap-and-trade emission control systems:

- Only optimal for uniformly distributed pollutants – CO₂ leading to Global Climate Change is a perfect example
- For pollutants where the distribution is uneven, cap-and-trade *could* lead to more harm. If the plant with the highest cost to cleaning buys many permits, then its neighborhood will be highly polluted. However US experience with SO₂ has been reasonably successful.
- For non-uniform pollutants, the trading could be moderated by transfer coefficients (below)
- Depends on all actors being profit-maximizing – many polluters are government agencies, so this assumption might poorly fit particular cases
- If there are only a few firms, then the assumption of perfect competition becomes risible. One dominant firm could either get its permits cheaply or force competitors to pay extra.
- Transactions costs can also reduce the efficiency of the market
- So trading among zones or with complicated transfer coefficients has problems of both high transactions costs AND market power

- System design is really important, we have now accumulated examples of cap-and-trade systems that work well (SO₂ for first 20 years) and those that fail (SO₂ after 2005; Euro carbon emissions). Legal framework is crucial.

A bit more Algebra

basic model distinguishes sources and receptors, spread over space
emissions, e_i , from source any of I sources (plus background levels, B)
cause pollution, p_j , at any of J receptors

$$p_j = f(e_1, e_2, \dots, e_I) + B_j$$

linearize for each j to get transfer coefficients, $p_j = \sum_{i=1}^I a_{ij} e_j + B_j$ (valid in neighborhood

if function is differentiable)

distinguish marginal damage of receiving p versus marginal savings of emitting e
Marginal Damage Cost, MDC, then is

$$MDC_i = \frac{\Delta \text{Damage}}{\Delta e_i} = \frac{\Delta \text{Damage}}{\Delta p_j} \frac{\Delta p_j}{\Delta e_i} = MD_j \frac{\Delta p_j}{\Delta e_i} = a_{ij} MD_j$$

Firms save money by emitting freely, so Marginal Control Cost

Social Optimum, (assume one pollutant) for each emission, i:

$$MCC_i = MDC_i = a_i MD(p)$$

Pigouvian Optimal tax policy implies set $t_i = MCC_i$ therefore firms will choose emissions such that

$$t_i = a_i MD(p) \text{ or } \frac{t_i}{a_i} = MD(p)$$

Suppose EPA issues L permits for pollution and each firm gets L_i

Firm emissions are ℓ_1, ℓ_2 and $L = \ell_1 + \ell_2$

Pollution, p, is $a_i e_i$, so $\ell_1 = a_1 e_1, \ell_2 = a_2 e_2$

price of pollution is C(e)

So firm's total cost is

$$TC_i = C_i(e_i) + \alpha(\ell_i - L_i)$$

$$TC_i = C_i(e_i) + \alpha(a_i e_i - L_i)$$

to minimize these costs set MTC=0 so

$$MC_e + a_i \alpha = 0 \text{ therefore}$$

$$-\alpha = \frac{MC_1}{a_1} = \frac{MC_2}{a_2}$$

thus equimarginal principle is met

if there are multiple receptor standards, all of which must be met, then firms will choose e_i to meet the lowest limit

3 Results:

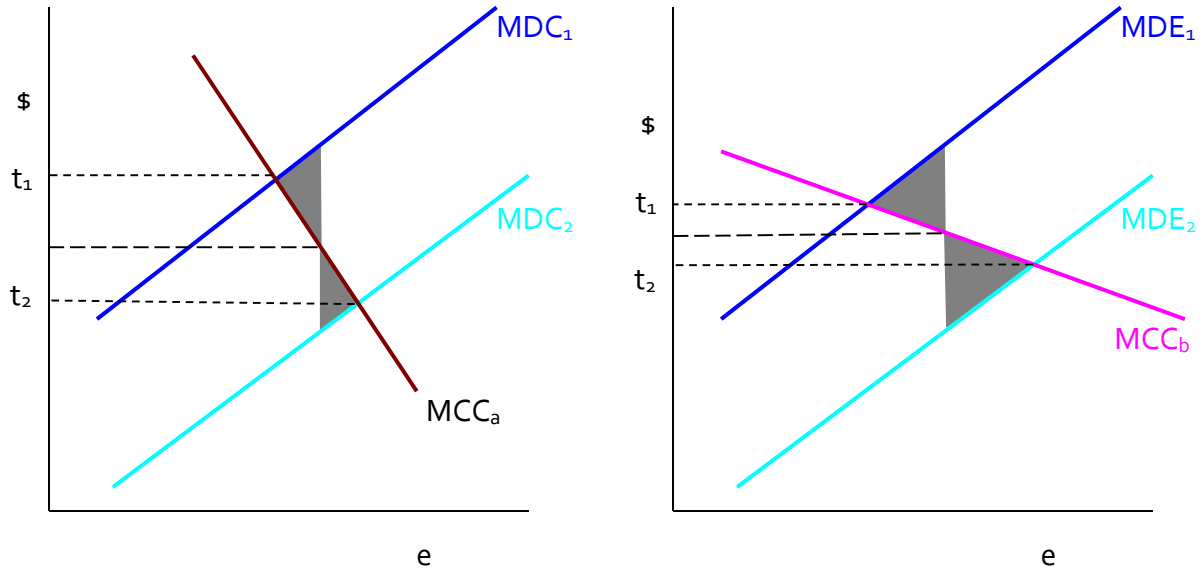
Equilibrium exists for any initial allocation of permits

Emissions from each source are efficient (no matter initial allocation)

If price equals marginal damage then equimarginal principle holds

What if different firms (firms with different transfer coefficients) pay same tax? Some inefficiency; size depends on elasticity of demand

If t_1^*, t_2^* are the optimal amounts that emitters 1 and 2 ought to pay,
but instead the tax is set at \bar{t} , then the DWL is:



So efficiency depends on relative elasticities again.

Pollution Over Time

This analysis is for pollution that is transitory. However much pollution is cumulative: current emissions will pollute for a lengthy time period.

Model with stock of pollutants, S_t , the stock at time t , increased by current emissions, e_t , while some fraction (δ) of previous pollutants decay.

$$S_t = \delta S_{t-1} + e_t$$

The Net Cost, NC is the present discounted value of all future costs of lowering emissions and all future damages from the stock of pollutants:

$$NC = \sum_{t=0}^{\infty} \beta^{t-1} [C_t(e_t) - D_t(S_t)]$$

so the marginal net cost per level of emission is

$$\frac{\Delta NC}{\Delta e_1} = \sum_{t=0}^{\infty} \beta^{t-1} \left[\frac{\Delta C}{\Delta e_1} - \frac{\Delta D}{\Delta S_t} \frac{\Delta S_t}{\Delta e_1} \right]$$

where, from $S_t = \delta S_{t-1} + e_t$, we can re-write $S_t = e_t + \delta e_{t-1} + \delta^2 e_{t-2} + \dots + \delta^t e_0$ so that $\frac{\Delta S_t}{\Delta e_1} = \delta^{t-1}$

$$\text{so } \frac{\Delta NC}{\Delta e_1} = \sum_{t=0}^{\infty} \beta^{t-1} \left[\frac{\Delta C}{\Delta e_1} - \delta^{t-1} \frac{\Delta D}{\Delta S_t} \right].$$

To minimize Net Cost, NC, set Marginal Net Cost equal to zero (the present discounted value of costs and damages equal to each other), which implies, notating $\frac{\Delta C}{\Delta e_1} = MC_1(e_1)$ and

$$\frac{\Delta D}{\Delta S_t} = MD_t(S_t), \text{ means setting } MC_1(e_1) + \sum_{t=0}^{\infty} \beta^{t-1} \delta^{t-1} MD_t(S_t) = 0.$$

This can be interpreted as the marginal savings today should be equal to the sum of marginal damages in the future, where the future damages are discounted both by time preference and the persistence of the pollutant.

Of course for the case where the pollutant is completely transitory, so $\delta = 0$, this gives us the same formula as before.

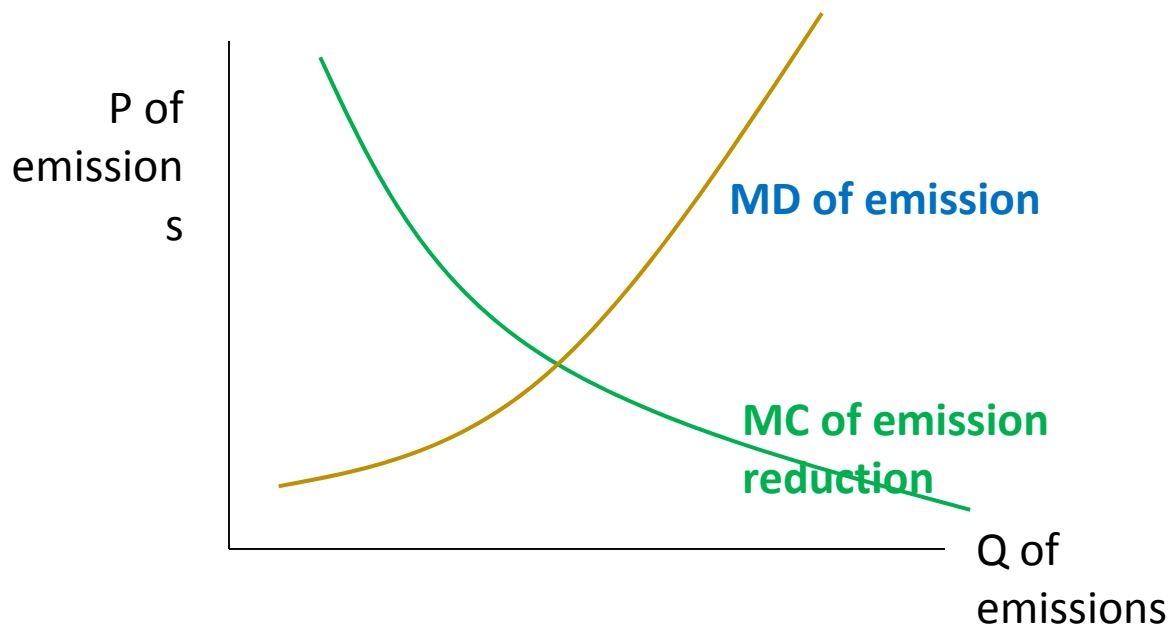
(More complicated cases of dynamic optimal control, use Bellman Equation.)

When Costs & Benefits are Imperfectly Known (i.e. The Real World...)

Can think of regulation as either controlling quantity or price. Issuing a particular number of permits is setting the quantity and letting market determine the price of emission; a fee or tax sets the price of emission and lets the market determine quantity.

Need to consider:

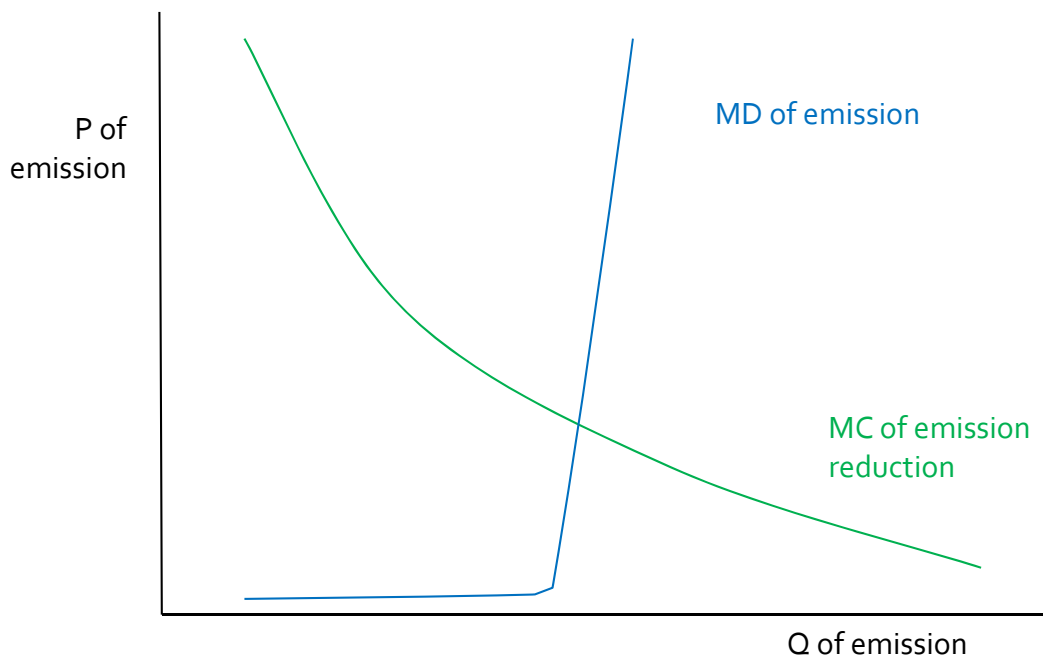
- whether there is relatively more uncertainty about marginal savings or marginal damages
- whether marginal savings or marginal damages are relatively steep



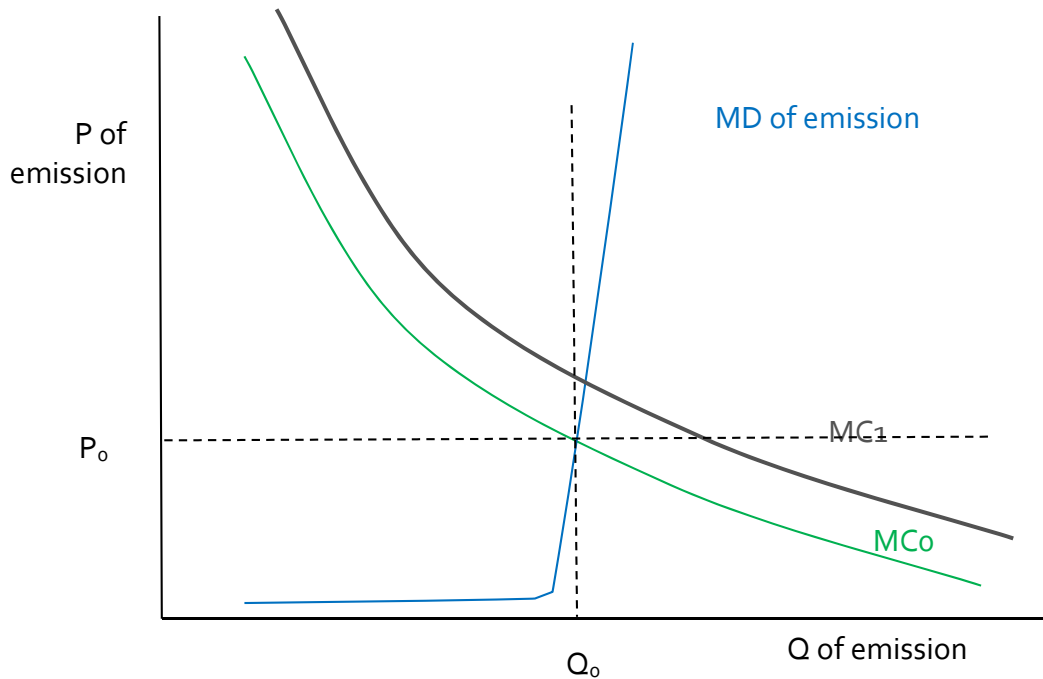
Marginal Damages increase with the amount of pollutant, while Marginal Cost of emission reduction falls, since the first cleanup is easiest.

Extreme Case 1: Threshold effects of pollution

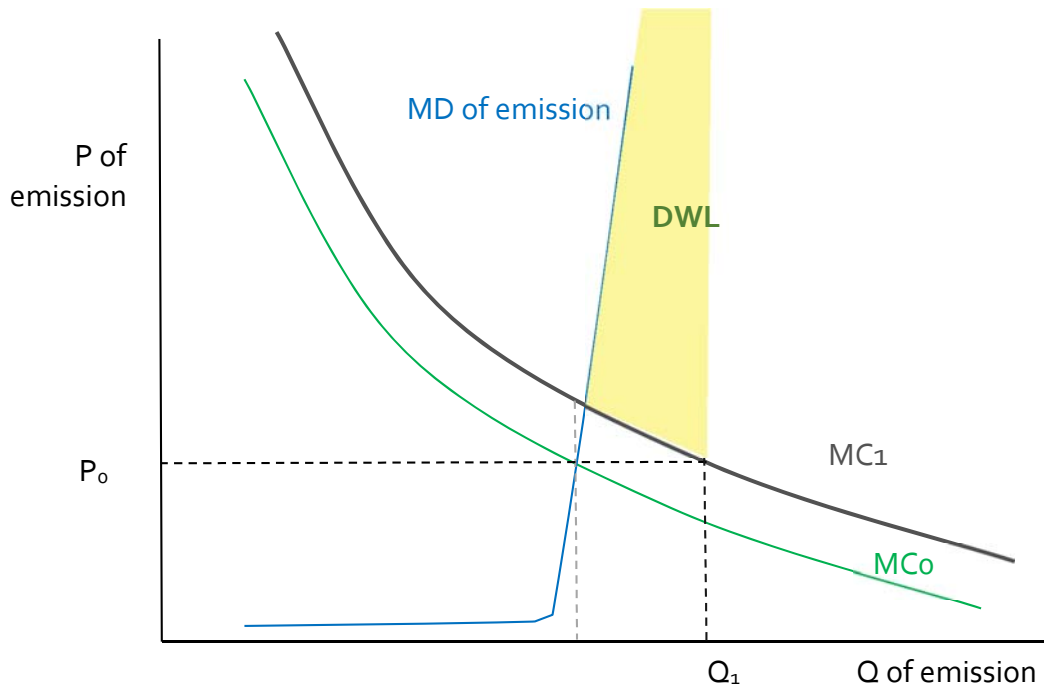
- virtually no damage below some level then damages jump to higher level once threshold boundary is crossed



Thought that MC was low at MC_0 but actually is higher at MC_1 . A tax would have given P_0 ; tradable permits would have set Q_0 .

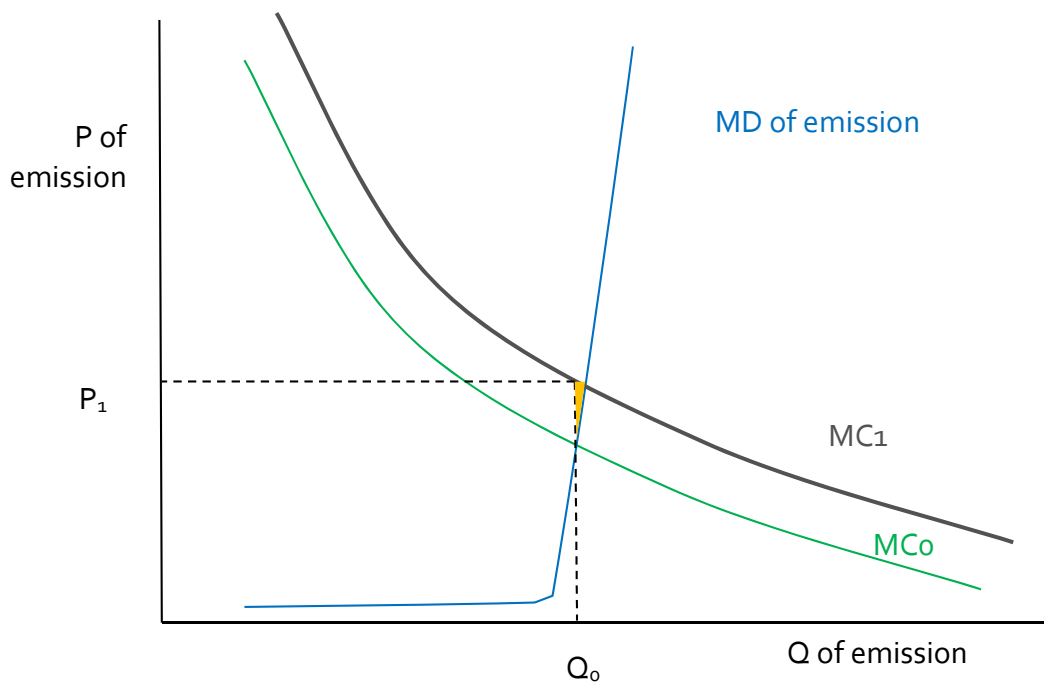


So if costs of emission avoidance had originally been estimated to be MC_0 but were actually higher at MC_1 , and if a tax had been set at level P_0 , then firms, seeing this price, would choose Q_1 of emissions and we would have this case:



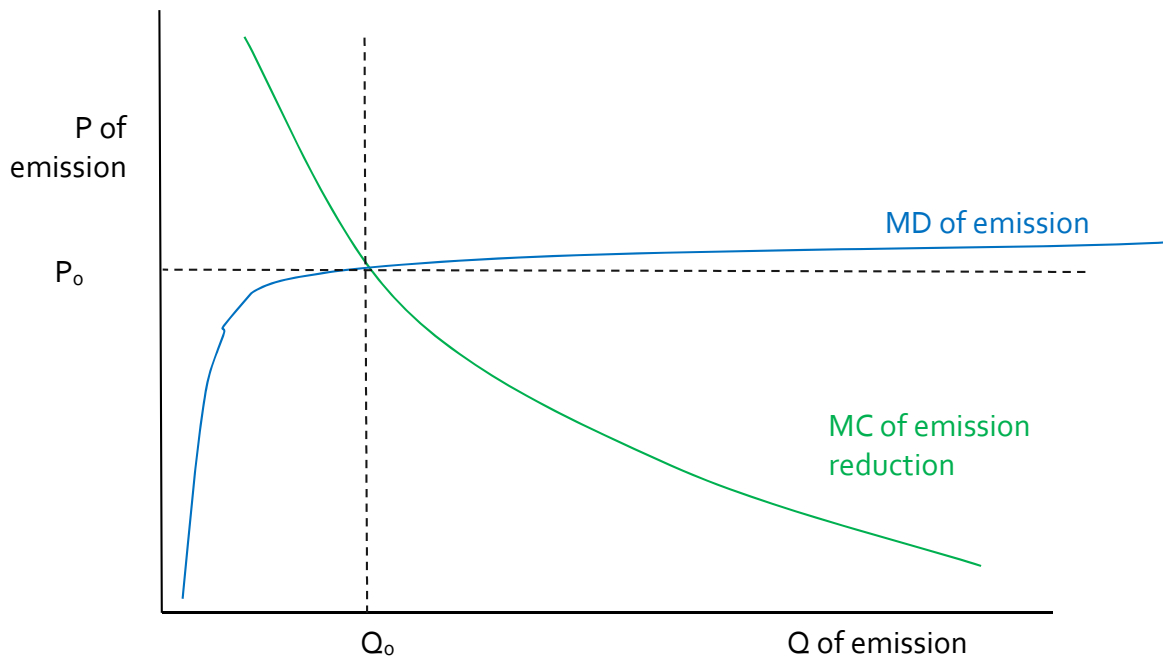
where there is now a HUGE deadweight loss: emissions are so high that marginal damages rocket upward – the yellow area shows the DWL of taxes set to the wrong level.

On the other hand, if marginal costs had been incorrectly estimated but policy had set a quantity target (number of tradable permits) at Q_0 , then firms would pay a somewhat higher price and there would be a small DWL since the number of permits is slightly smaller than optimal:

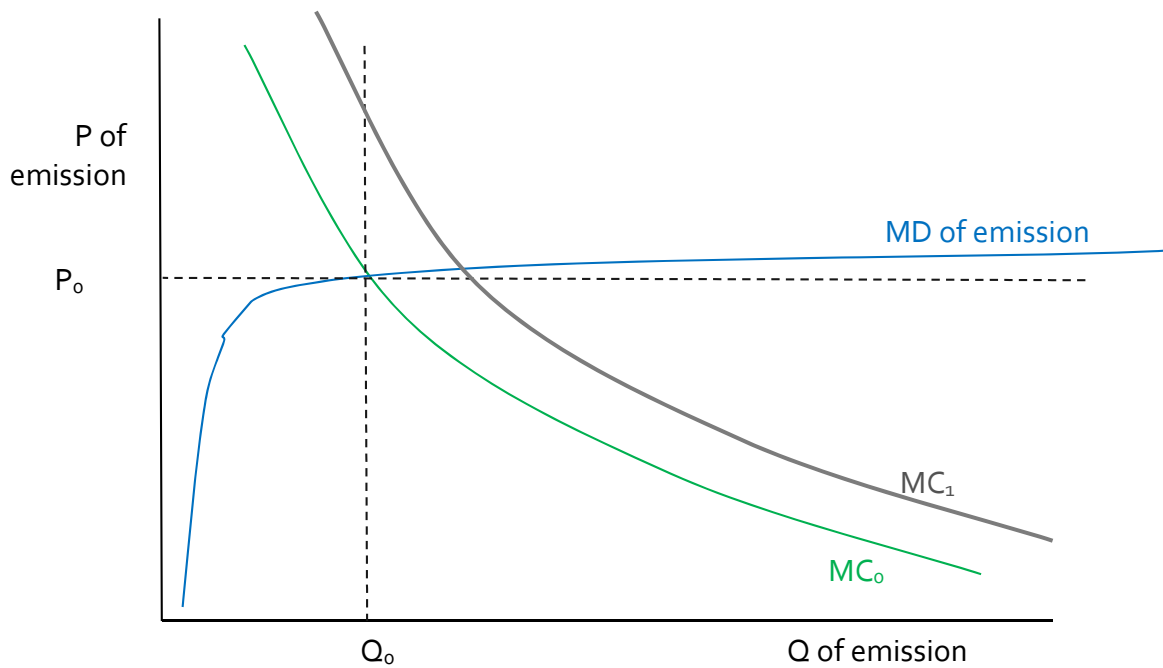


Extreme Case 2: Constant Marginal Damages

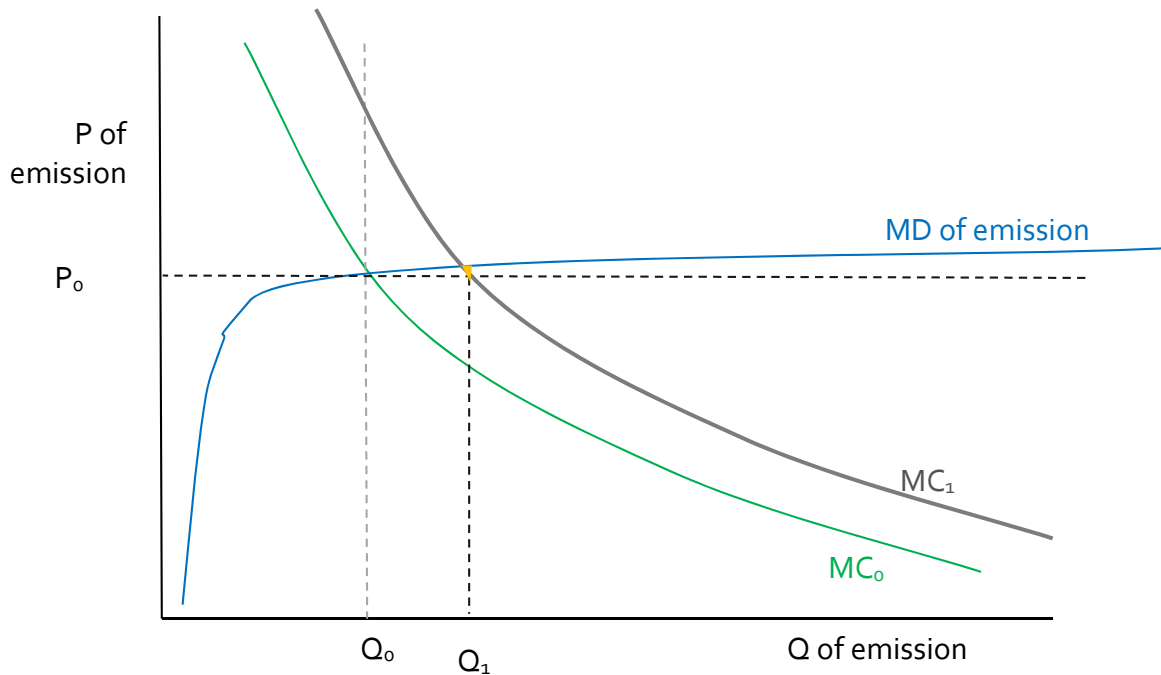
- level of damage is nearly linear in amount emitted
- probably reasonable in cases where, for example, damage is limited in geographical scope:
ruining 200 acres of habitat is twice as bad as ruining 100 acres



So policymakers can regulate either the quantity (through issuing tradable permits) or price (with a tax on emissions). Again we ask, what if they are wrong in estimating MC? Suppose it's MC_1 not MC_0 ?

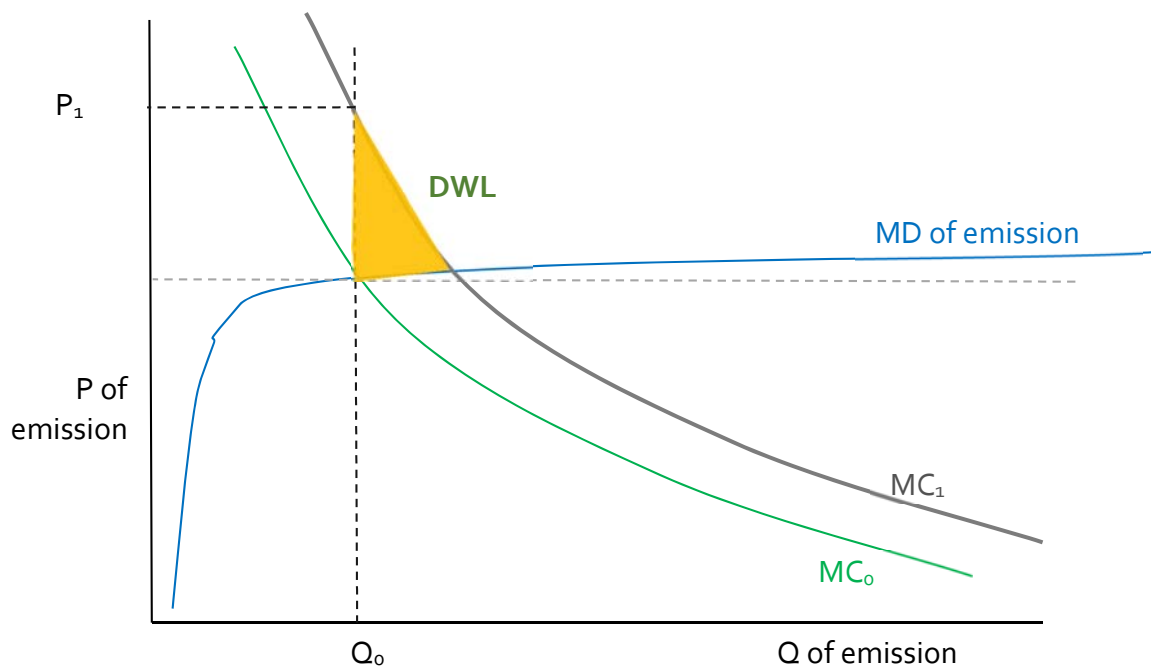


Then if regulators had set a level of tax at P_0 , then the quantity chosen would be Q_1 with the new higher costs, so we'd get this situation:



Where there is a tiny bit of DWL (since MD might be slightly above the tax level, P_0) but this is quite small.

On the other hand, if the quantity had been regulated, then the prices of tradable permits would be bid up very high to P_1 since emissions abatement was much larger than anticipated, so we would have this case:

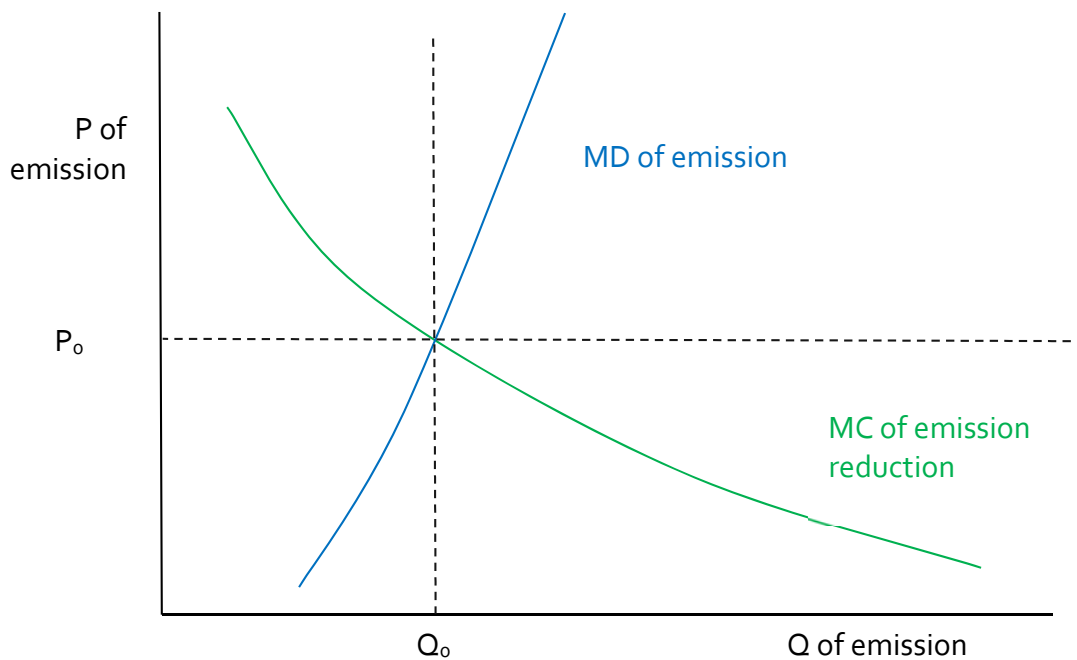


So here the DWL is much larger: firms ought to be allowed to emit more pollutant since their costs of abatement are so much higher than anticipated, but the stringent rule allows a sub-optimal level of pollution.

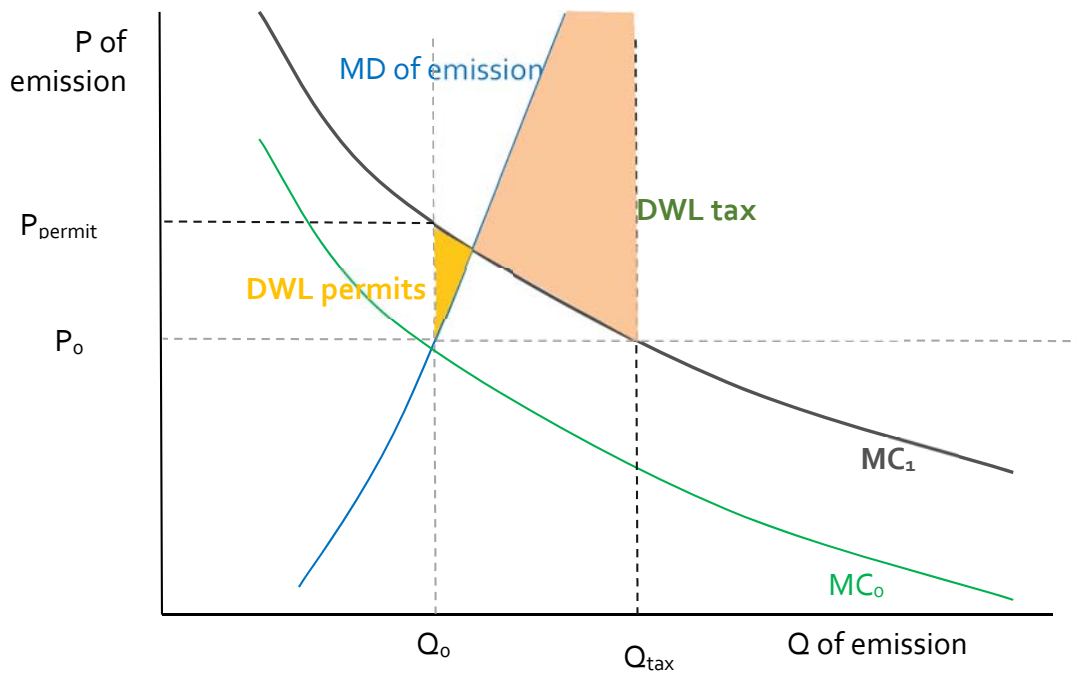
Neither of the extreme cases above might be likely in the world, but they're useful to pin down our thinking. When policymakers have a better idea of what is the optimal quantity (i.e. threshold), then they ought to regulate the quantity; if they have a better idea of the optimal price (linear damages) then they ought to regulate the price.

Case 3: MD steeper than MC

If the Marginal Damages curve is relatively steeper (less elastic) than the marginal costs of emission reduction, then we have a situation like this:



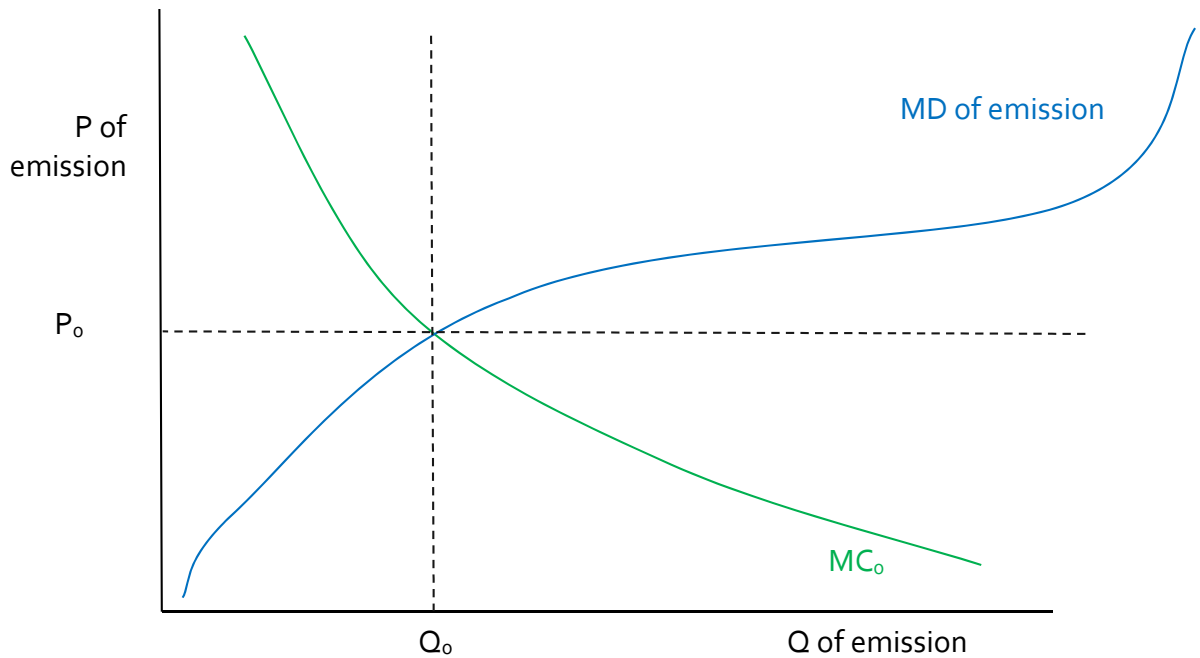
Again, if MC of emission reduction was incorrectly estimated and compliance is actually higher, then we have this case:



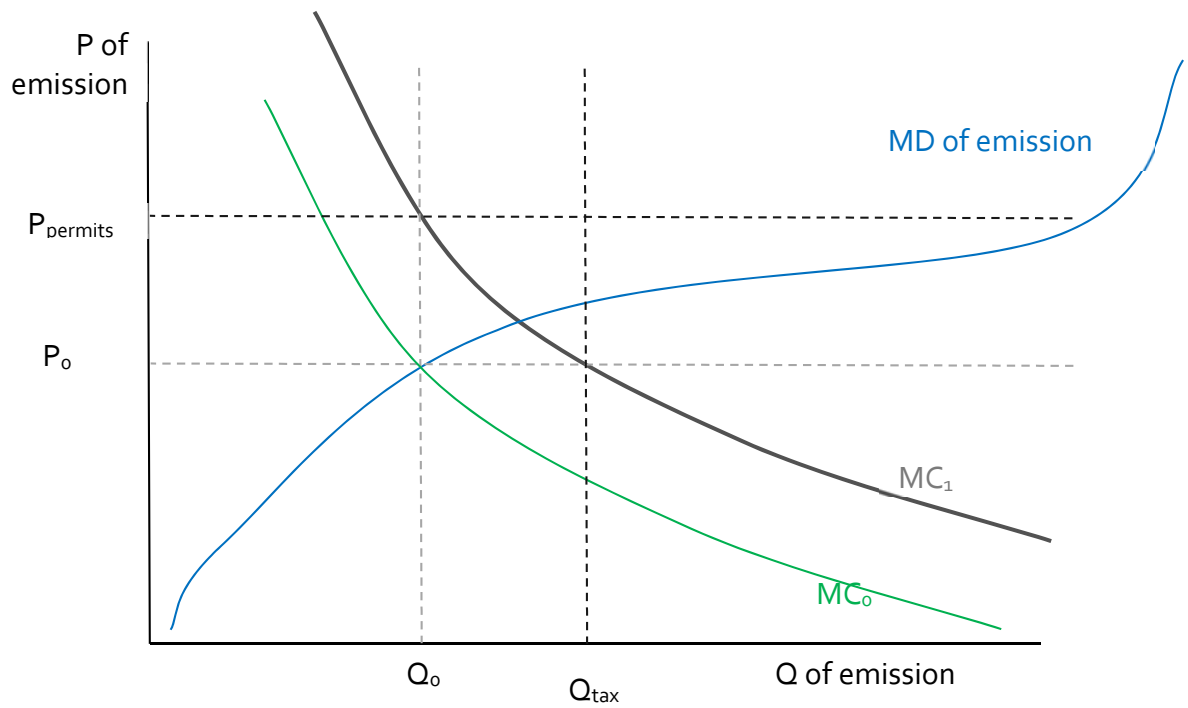
So if a tax had been set, the DWL would be the large pink area (from being too dirty); had tradable permits been set, then DWL would be the smaller orange area (being too clean).

Case 4: MD flatter than MC

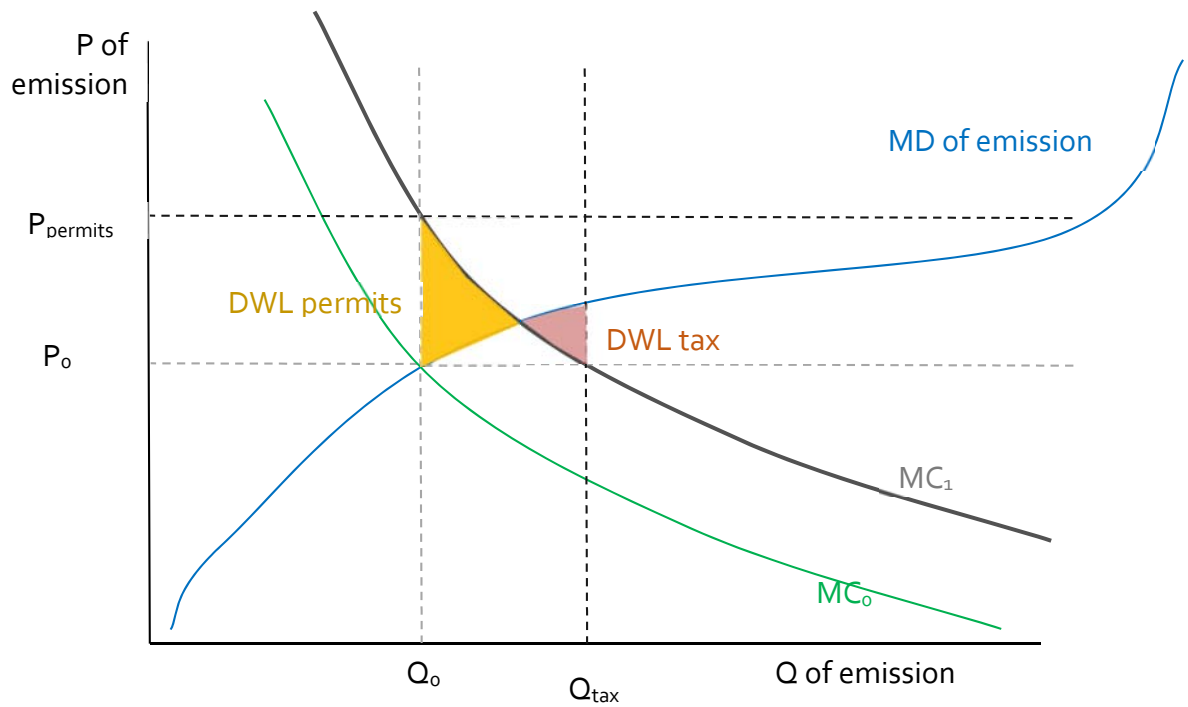
If the marginal damage curve is less elastic than marginal costs, then we would have this case:



So if policymakers set quantity (Q_0) then they also get P_{permits} ; if policy sets the price at P_0 then they get the quantity Q_{tax} .



So policy gets either of these two DWL areas:



So in this case, a tax would give a preferable result.

You might wonder what would happen in each case if the MD curve instead of MC were to shift, or if MC were to move down rather than up – I encourage you to sketch those out for yourself!

The general principle that we can deduce is that if policymakers are relatively surer about the optimal level of pollution, then tradable permits are good; if policymakers are relatively surer about the level of damage then a tax is good. Making good policy choices is tough!

Hybrid Price/Quantity controls might be better, to replicate the marginal net social damage, so for example a tax could be set but with an 'escape valve' that if the quantity of pollutants rose above some level then the tax would step up; or permits could be issued but again with an escape clause that if the price of permits rose above some level then more permits would be issued.

Details of Fees and Permits

Regulation of Quantity, in real world, must specify:

- time of permit duration (daily, annually, etc)
- information required
- monitoring data to be provided
- inspection schedule and costs of non-compliance (review Law & Econ result)
- how often permit/fee will be updated
- whether firms can bank permits

For dynamic efficiency, price regulation can be more effective than quantity regulation since there are better incentives for innovation – the price target is known so the marginal benefit to efficiency is known, which lowers the uncertainty of investment in efficiency.

Regulation through Liability

- more law & econ

Tort law (liability) involves the state setting rules to govern the behavior of 2 individuals, the injurer and the victim (technically, the potential injurer and potential victim, but for now we use the shorthand terms). Both may take a certain amount of care in their activities. For instance a manufacturer of a toy should take care that it not be dangerous; buyers should take care that it not be used in a dangerous way. Denote x as the care taken by the injurer, at a cost $c(x)$. Often we might assume that the cost rises with x . Then denote $L(x)$ as the loss to the victim; presumably it would be a decreasing function of x . The social objective function is to provide incentives for people to choose x to $\min c(x) + L(x)$. A typical analysis would show that the optimal level of care, x^* , is where the marginal cost equals the marginal loss.

There are at least three possible sets of rules that would set incentives to each party:

- No liability would give the injurer an incentive to minimize x , without regard to L .

- Strict liability has the injurer paying all costs, so their costs are $c(x) + L(x)$, so the injurer would take the optimal amount of care
- Negligence, where the injurer is liable for all costs if he/she did not exercise “reasonable care”. If the level of reasonable care, x' , is set equal to the optimal level, x^* , then this would provide the proper incentives again. If the injurer takes less care then they are back in the “strict liability” world where they pay the full costs; if they take more care then they have no gain; so they should set $x = x^*$.

Up to now we have discussed care as taken only by the injurer. But now introduce care by the victim, y , so again there is $c(y)$ the cost of the victim taking care and now $L(x,y)$, where the loss to the victim depends on the care taken by each party. Now society wants to min $c(x) + c(y) + L(x,y)$. Now there are two optimal levels, x^* and y^* , that set the marginal cost of care equal to the marginal diminution of loss.

Now the liability standards are:

- No liability, so $x=0$ and y is too high.
- Strict liability, where now the injured party has no costs so $y=0$ and x is too high.
- Strict division of losses, where each side pays some fraction of the loss, f . In this case the injurer will min $c(x) + fL(x,y)$ and so choose x to set $MC(x) = f \cdot ML$, so x will be too low. Similarly for the victim, who will min $c(y) + (1 - f)L(x,y)$ and will choose a y that is too low.
- Negligence, where the injurer is liable if their care is below some x' value. This must be analyzed as a game, since the outcome depends on the other actor's behavior. We can see that the Nash equilibrium has each side choosing x^* and y^* . Suppose that the victim chose y^* and the injurer is choosing. He/she knows that if care is too low then he/she will pay the full costs, so just as in the simpler case the injurer will choose x^* . The victim will face the same choice: if the injurer chooses x^* then the victim will be liable for the losses if y is too low; again the victim will choose y^* .
- Strict liability with defense of contributory negligence, where the injurer is fully liable unless the victim's care was below some y' level. Again, if y' is set to y^* then this gives an optimal outcome.
- Double liability, where each side bears the full costs. The problem with strict division was that neither side took due account of the loss. If both sides pay the full cost, however, then both will take due care. This is useful in cases where the level of care is difficult to observe. It is the logic behind “no fault” car insurance where each party's car insurance pays the bills and the traffic courts separately determine fault.

Regulation through Insurance

Insurance is intimately tangled with liability since often a firm, which is legally liable for some action, will buy insurance against that outcome (for instance, Director's & Officer's Insurance, which covers the company management from personal liability for their decisions at work). Workman's compensation is often used to cover liability to hazardous working conditions.

Insurance, to work well, needs six factors:

- risk pooling of
- clear losses
- over a well-defined time period
- that are frequent enough
- with a small moral hazard and
- small problems of adverse selection.

Note that "pooling" problems gave rise to reinsurance.

Costs and Benefits

In many policy choices, economists estimate the costs and benefits of different policies. It is worth looking under the hood at how these numbers are derived. Figuring out clever ways of putting a dollar value on things that are not obviously measured that way can be useful. But it remains difficult and uncertain. We'll discuss some of these subtleties.

Valuation of Life

See Viscusi (1993) "The Value of Risks to Life and Health," Journal of Economic Literature.

Risk & Uncertainty

Risk is based on probabilities and can be treated mathematically

Uncertainty cannot be easily represented.

People are lousy even at evaluating risks, with little ability to differentiate between risks of different magnitude. That's why casinos can exist.

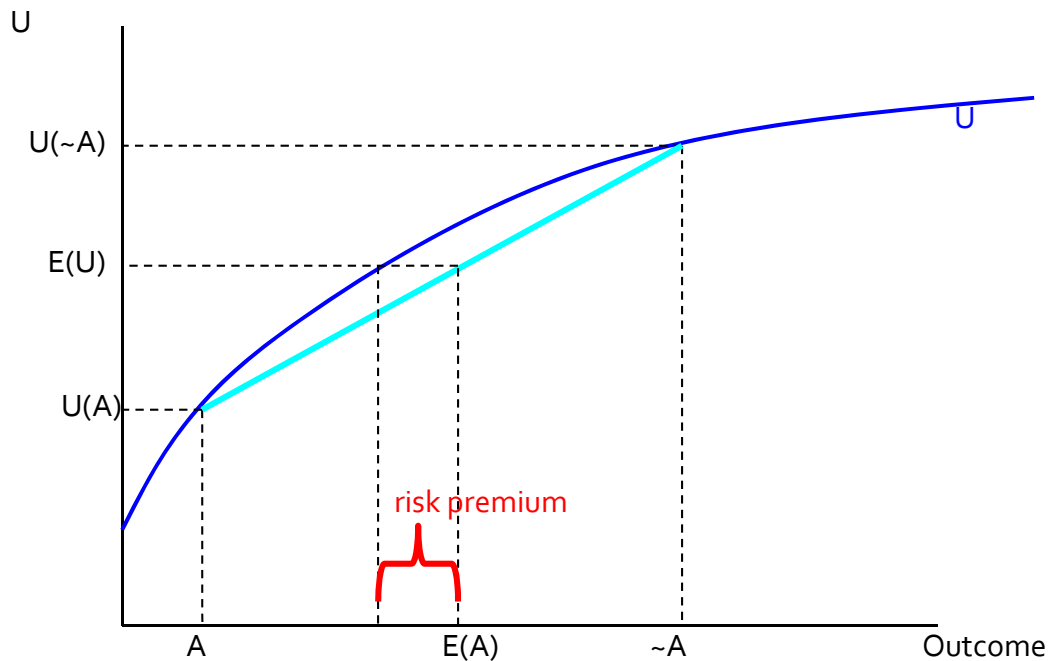
Behavioral economics has formalized some of these observations about how people are systematically irrational.

For a rational decision maker it is usually convenient to assume that an individual has von Neuman-Morgenstern utility. This means that a person's expected utility can be represented as:

$E(U) = u(x_A, x_B, \pi_A, \pi_B) = \pi_A U(x_A) + \pi_B U(x_B)$. If a person's instantaneous utility function is concave then $E(U(x)) \leq U(E(x))$ where we define the expectation operator as

$E[X] = \pi_A X_A + \pi_B X_B$, X_i is the value that X takes in each case, and π_i is the probability in each case. The risk premium is the difference by which $E(U)$ exceeds $U(E)$.

This graph shows the case where there is either an accident (A) or not an accident ($\sim A$):



Often is convenient to suppose that an individual's instantaneous utility depends on two factors, income and some other event, some disaster. We assume that income is at level Y no matter which of the two outcomes occurs; then the other event is either A occurs or it does not ($\sim A$). Then utility is $U_A(Y, A)$ or $U_{\sim A}(Y, \sim A)$. Evidently if the probability of A occurring is π_A then the probability of A not occurring is $\pi_{\sim A} = (1 - \pi_A)$.

Various measures of how valuable it is, to eliminate the uncertainty.

Expected Surplus

Define V as the amount of income which a person would sacrifice to be indifferent between having the full income and the disaster or having less income but no disaster. (This is analogous to the Hicks measure of income effect from way back in baby Micro.) You might recall that this measure of income is not generally the same depending on where you start. In this case we must differentiate between utility starting from a disaster, U_A , and utility starting from no disaster, $U_{\sim A}$. So define:

V_A such that $U_A(Y - V_A, \sim A) = U_A(Y, A)$ and

$V_{\sim A}$ such that $U_{\sim A}(Y - V_{\sim A}, \sim A) = U_{\sim A}(Y, A)$.

So the Expected Surplus, ES, is defined as the expected value of these valuations,

$$ES = \pi_A V_A + (1 - \pi_A) V_{\sim A}$$

Option Price

Or the option price is the amount that a person would pay now to eliminate the uncertainty, so as to be indifferent between $E(U(A))$ and $E(U(\sim A))$. This option price, OP, is such that

$$\pi_A U_A(Y - OP, \sim A) + (1 - \pi_A) U_{\sim A}(Y - OP, \sim A) = \pi_A U_A(Y, A) + (1 - \pi_A) U_{\sim A}(Y, A).$$

Generally the option price will be larger than the Expected Surplus due to risk aversion.

In general financial markets (even prediction markets) can give significant insight about risks and probabilities.

Irreversibility and Precautionary Principle

Some decisions about the environment are irreversible, whether developing a wild "untouched" natural area or climate change that melts glaciers or loss of habitat that causes extinction of species. Additionally, there may be uncertainty about the valuation of these stocks: how much is a species worth, if we haven't even studied it yet?

This is called a "real option" in corporate finance (businesses confront these questions all of the time, investing in technologies with wildly uncertain outcomes). Waiting to make a decision becomes an investment in lowering the uncertainty of the outcome. A lower level of uncertainty has a value (from finance, people regularly trade off risk versus return, choosing for example between high-risk and high-return investment strategies or lower-risk and lower-return investments). So although making a development decision today increases the return (since the reward is closer to the present), delaying it brings a benefit of less uncertainty.

Actual Behavior of People making Choices under Uncertainty

People don't actually make choices in a way that adheres to these models; they're more complicated and irrational. In figuring costs and benefits we usually make assumptions of rationality (that a person would choose the option with greatest margin of benefits over cost) but these assumptions often do not hold in the real world.

Kahneman and Tversky give these examples (from Kahneman's 2002 Nobel Lecture):

The Asian Disease

Imagine that the United States is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows:

- If Program A is adopted, 200 people will be saved
- If Program B is adopted, there is a one-third probability that 600 people will be saved and a two-thirds probability that no people will be saved

Which of the two programs would you favor? Majority choose A.

Alternate Statement:

- If Program A' is adopted, 400 people will die
- If Program B' is adopted, there is a one-third probability that nobody will die and a two-thirds probability that 600 people will die

Which of the two programs would you favor? Majority choose B.

This is the "Framing Effect" and it has even been shown to affect the choices of experienced physician, depending whether treatments had a "90% survival rate" or "10% mortality rate".

Prospect Theory

Problem 1

Would you accept this gamble?

- 50% chance to win \$150
- 50% chance to lose \$100

Would your choice change if your overall wealth were lower by \$100?

Problem 2

Which would you choose?

lose \$100 with certainty

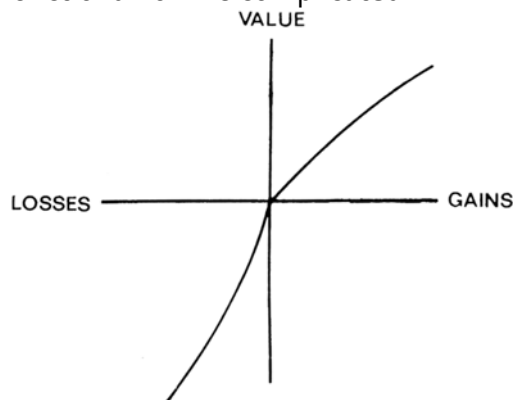
or

- 50% chance to win \$50
- 50% chance to lose \$200

Would your choice change if your overall wealth were higher by \$100?

The choices are clearly identical but most people switch choices.

Can model people's utility as not from absolute level but from changes in wealth, and functional form is complicated:



"It is worth noting that an exclusive concern with the long term may be prescriptively sterile, because the long term is not where life is lived. Utility cannot be divorced from emotion, and emotion is triggered by changes. A theory of choice that completely ignores feelings such as

the pain of losses and the regret of mistakes is not only descriptively unrealistic. It also leads to prescriptions that do not maximize the utility of outcomes as they are actually experienced – that is, utility as Bentham conceived it."

Most people make decisions based on simple *heuristics*, which are often approximately correct and are useful in minimizing the total mental effort of making a choice. Most people make choices about, say, what restaurant to choose for dinner tonight – not worth spending a great deal of time thinking about! They're not inclined to think much more deeply about bigger problems.

When college students are asked, on a survey, "How happy are you with your life in general?" and "How many dates did you have last month?" there is almost zero correlation; however if the survey asks them in the opposite order, the correlation jumps to 0.66!

The immediate corollary is that people can be cued to respond in a more statistically sound (rational or logical) manner, in ways as simple as just reminding them to "think like a statistician."

But it complicates the question of how we, as a society, ought to come to conclusions about complex issues involving a range of tradeoffs in the face of uncertain possible outcomes.

Choice of Dumping or Safe Disposal

(from Kolstad's Environmental Econ textbook)

Sometimes polluters face a choice of either cheaply dumping emissions (in a way that is socially harmful) or safely disposing of them. An example is garbage – it may be optimal to try to give households incentives to reduce their garbage by charging per pound or per bin, but this might also provide incentives for people to dump garbage in some deserted area. Consumers' batteries and electronic devices should be properly disposed of. This is especially relevant for producers who use various chemicals that need proper disposal.

One option is subsidizing the safe disposal, but this is costly and, of course, encourages more disposal. Suppose households were paid some amount of money per pound or per bin of garbage! That sounds crazy – but it might not be, if we can change other costs.

Assume a firm can dispose of waste, w , either, safely at cost $C_s(w)$ or not safely at cost $C_N(w)$. Clearly there is only a problem if $C_s(w) > C_N(w)$. So the government might introduce a subsidy per unit of waste, so changing the cost of proper disposal to $C_s(w) - sw < C_N(w)$ (if the subsidy does not lower the cost enough then we're back to the original problem).

So how can we change other costs? Suppose the regulator observed the total volume of waste and taxed that, so now the costs for the firm are increased by tw but this is charged regardless of whether the firm dumps or disposes safely. The firm will not dispose safely if

$C_s(w) - sw + tw < C_N(w) + tw$ which clearly is no different. So the subsidy rate should be set to give an incentive to dispose safely while the net tax on waste, $(t - s)$, should now equate to marginal damages.

A classic example of this is the deposit-refund system on bottles of soda, where consumers pay extra at purchase but then get a subsidy for proper disposal.

There are other examples of narrow-focused disposal charges, where buyers initially pay more and then can claim back some of this charge if the item is properly disposed (e.g. household electronics).

Background on Global Climate Change

I expect most students know most of this but I give a quick review just to make sure. (I wrote this part for an undergrad class so please just skim it.)

The three most important aspects of global climate change are:

1. Human activity is changing the earth's climate.
2. The poorest people will bear much of the burden and the costs of this climate change.
3. As poor people become richer, this worsens the pace of climate change.

Note that #2 and #3 are contradictory – this is one of the reasons why there is no clear or easy solution. Part #3 also helps understand why predictions of the future are so difficult: different development paths can mean huge swings in carbon output.

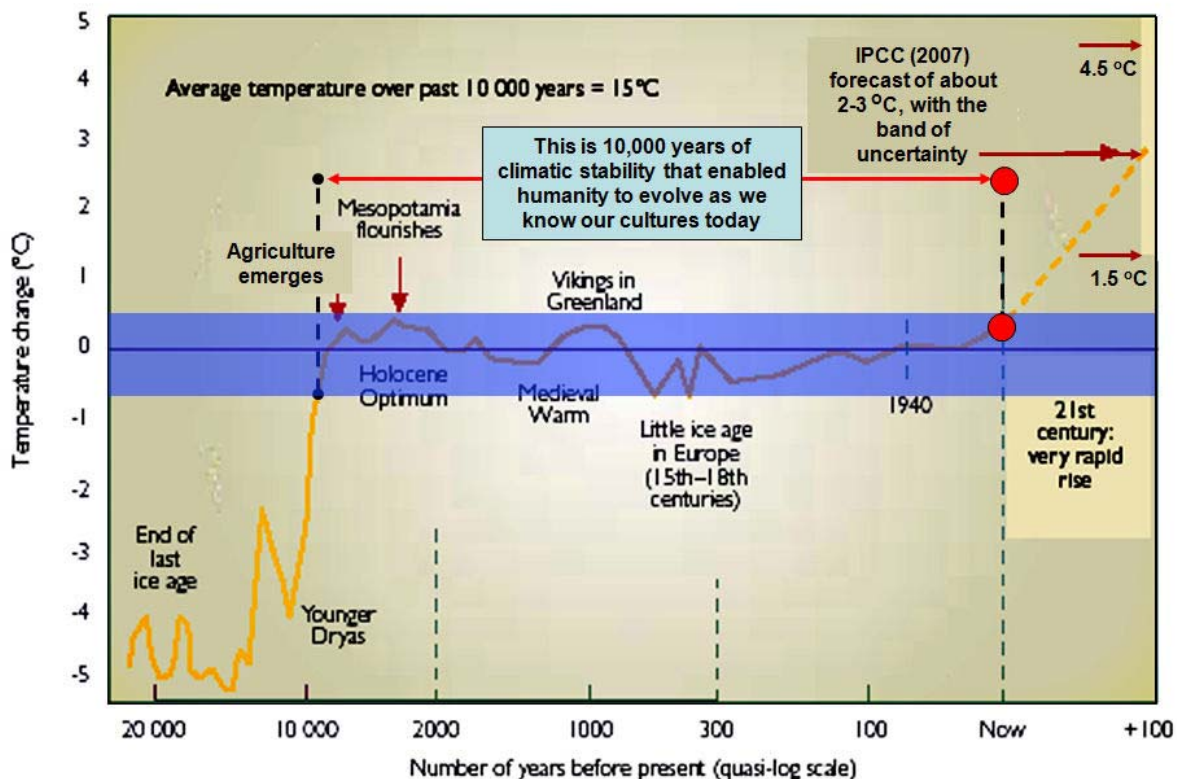
For point #3, note that the easiest way to 'solve' GCC problems would be to reincarnate Mao and put him back in charge of China to impoverish a billion people; also reincarnate Nehru to put him back into India to impoverish another billion – poor people have a small carbon footprint! If you think that those seem like bad policies then we need to consider alternatives, of how people can become better off without polluting so much.

Point 1: Human activity is changing the earth's climate

Climate science gives insight about the role of carbon dioxide (CO₂) in regulating our planet's temperature. In earth's lengthy history (over 4bn years) the planet's temperature has fluctuated; glaciers have advanced (covering even New York City's current location) or tropical conditions have spread outward from the equator. Life has continued but in different forms with varied species expanding, diminishing, or becoming extinct.

All of human history has occurred during a recent period of relative cooling – we humans are attached to the particular climate that we've gotten used to in our short time on this planet.

The Last 20,000 Years seems to have been Ideal for the Development of Human Societies. Is this a Historic “Sweet Spot” that Enabled Humans to Flourish?



from climateprogress.org, "Must have PPT #1: The narrow temperature window that gave us modern human civilization," August 27, 2008

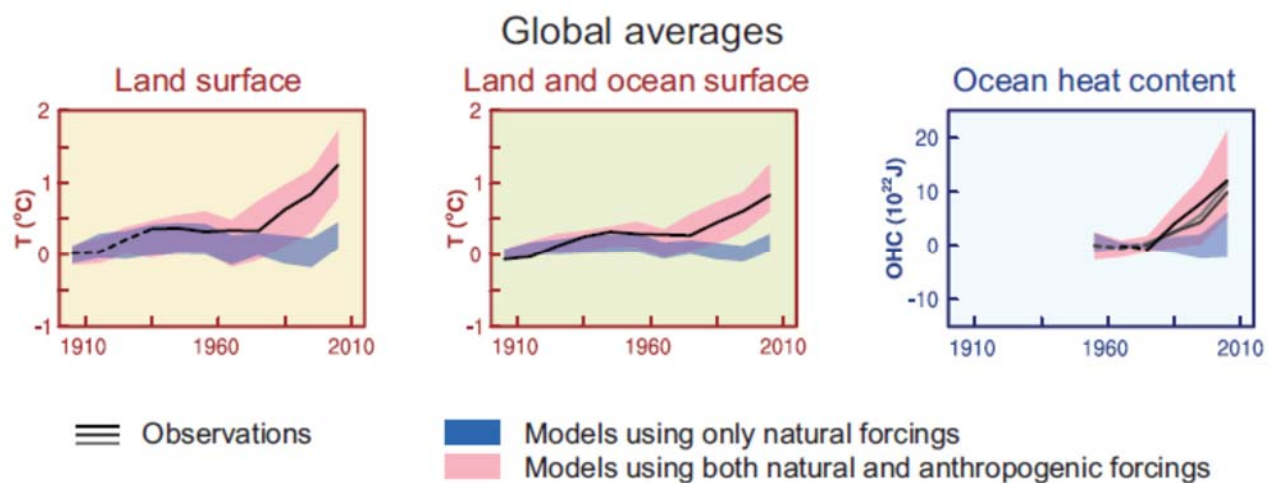
The earth's temperature has been influenced by the quantity of gases such as carbon dioxide, through the "greenhouse effect." The greenhouse effect traps some of the Sun's heat within the Earth's atmosphere. The atmosphere is transparent to incoming solar radiation (in certain wavelengths); much of this is absorbed by the planet and some is re-emitted, usually at lower wavelengths such as infrared. Greenhouse gases (GHG) absorb some of this infrared radiation so the greenhouse gases can trap heat within the planet. Therefore as the amount of greenhouse gases such as carbon dioxide rises, the amount of heat trapped within the planet rises.

There is a positive feedback loop (which is very complicated to model) where warmer temperatures can lead to more CO₂ emissions. Recall from basic science that CO₂ is helpful to plants: they take in CO₂, use photosynthesis to release energy from it, take the carbon (the C) to make the plant itself, then emit the oxygen (the O₂ part). Kids wonder how a tiny acorn becomes a giant tree – it's carbon. Trees are mostly made of carbon that has been taken out of the air. This is one reason why planting more trees can help global climate change, since they take carbon out of the air. But eventually when the tree dies, the carbon is usually released back into the atmosphere.

But not always; some of the carbon might sink into a swamp, for example. Oil deposits are the carbon residue of ancient life: carbon was taken out of the atmosphere millions of years ago and hidden away deep inside the earth. Hidden, until humans extracted it.

Where does the increase in CO₂ come from? Much of it is "Anthropogenic," thus the term Anthropogenic Greenhouse Gases – gases which are created by humans (have genesis from us anthro's). There is debate over precisely how much of the increase in CO₂ is due to humans. However there is no longer any debate over the basic fact of rising average temperatures and that anthropogenic greenhouse gases make the temperature rise worse.

Temperatures, showing what models show would have been natural versus what is due to humans:



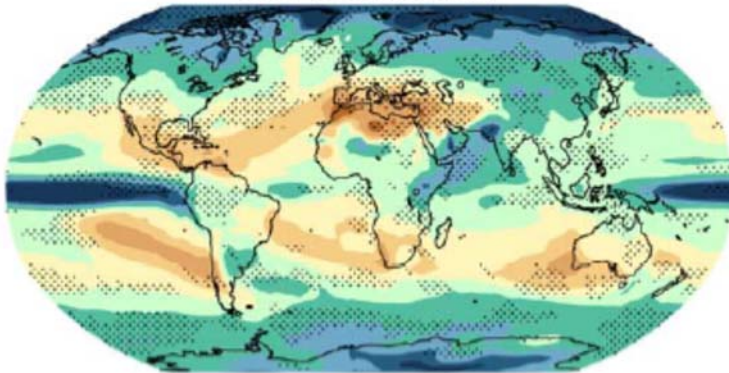
From IPCC AR5 Summary for Policymakers, showing 90% confidence intervals.

Note, however, that although greenhouse gases tend to increase global average temperature, we have stayed away from using the term "Global Warming" since not all parts of the earth are getting warmer. The climate is changing, and on average there is more warming than cooling, but there are some places that are getting colder. These parts are not necessarily the parts that need it, though! It seems that much of Antarctica is getting colder. For this reason scientists prefer the more precise label of "Global Climate Change" (GCC) instead of global warming.

Just as an increase in global average temperature does not mean that every location gets warmer, an increase in local average temperature does not simply mean that every day will be a bit warmer. There is a greater likelihood of high-stress events such as heat waves and drought. There also seems to be a greater likelihood of harsher winters in some areas.

An example of the complicated effects caused by GCC is the pattern of rainfall. On average higher temperatures mean more evaporation and more capacity for the atmosphere to hold moisture. This chart shows the areas of the globe that are predicted to get more rain (in blue) or less rain (in red).

Precipitation scaled by global T (% per °C)
CMIP3 : 2080-2099



From IPCC AR5 Working Group I Report

Looking closely shows that many of the areas that will get less rain are the areas that are already dry: the Mediterranean and northern Sahara, southern Africa, the US Southwest, and southern Australia. The human effects are disparate: in the US Southwest, people in LA and Las Vegas might have to conserve water better; in Africa many people who are already in poverty might face starvation as crops fail. There is some evidence that the conflict in Darfur was worsened by drought and competition for scarce water.

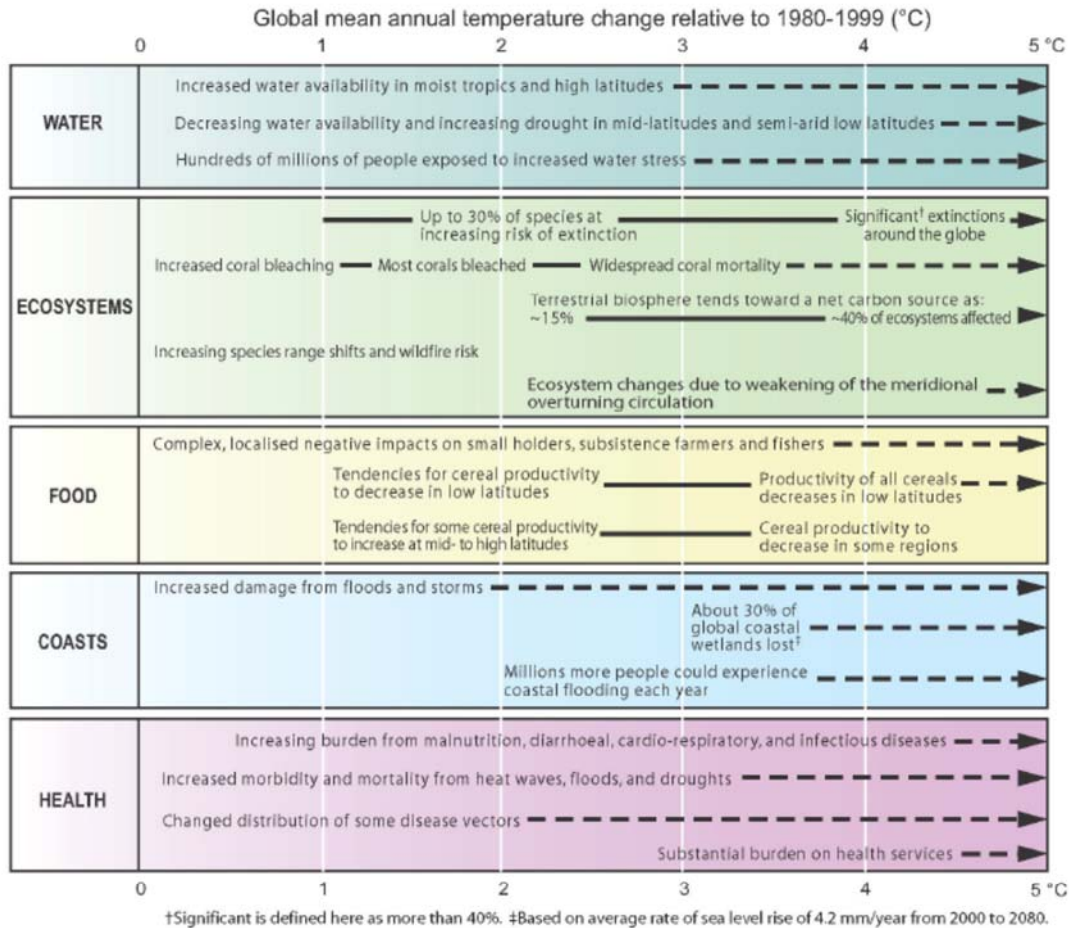
Point 2: The poorest people will bear much of the burden and the costs of climate change

As with just about everything else in the world, poor people will get the worst of it. This is true both across countries and within countries.

An example of the human impact of weather events was provided by Hurricane Katrina in New Orleans (although it was not directly caused by GCC). The victims were overwhelmingly those with the least income, the least political influence, and the lowest social standing. The hurricane caused significant economic losses but the wealthier people usually had insurance (which diffuses the costs). The worst effects and most deaths were among the poorest.

There are a host of other effects. The chart below shows the impacts on water availability, on ecosystems, food availability, coasts, and human health. Below the chart is a second graphic which shows the range of outcomes predicted by different models, with most showing a 2°-4° C rise in global average temperature (for Americans, this is about a 4°-7° F rise).

Examples of impacts associated with global average temperature change
(Impacts will vary by extent of adaptation, rate of temperature change, and socio-economic pathway)



Warming by 2090-2099 relative to 1980-1999 for non-mitigation scenarios

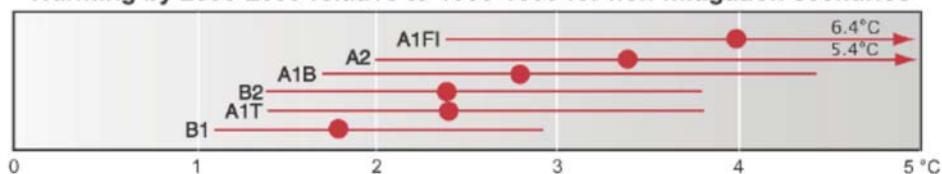


Figure SPM.7 from IPCC AR4 SYR SPM.

Water stress and the risk of drought is predicted to increase. The rise in average temperature not only changes patterns of rainfall and snowfall but also reduces the size of mountain glaciers. These glaciers act as reservoirs that release water into rivers slowly throughout an entire season rather than triggering flooding at spring thaw. Among the catastrophic outcomes would be if the Himalayan glaciers decreased significantly, since these feed rivers stretching through Pakistan, India, Bangladesh, Myanmar, and China. (Andean glaciers face similar threats.)

Ecosystems will begin to see a rising number of extinctions. Since the ocean absorbs some of the increased carbon dioxide, this change in its chemistry will have negative impacts particularly on coral reefs.

For food the picture is more complicated: colder regions at higher altitudes are likely to see increases in production for modest temperature rises – farmers in Canada and Russia win. However food production in countries nearer the equator will fall – African and India will lose. Again, this means that some of the poorest people on the globe will face threats to their basic ability to scrape a bare living.

Coastlines will see significant problems from the rising sea level combined with more rain and possibly more frequent or stronger hurricanes. The rise in sea level hits coastal wetlands which have an important role in water quality as well. Many of the world's cities (therefore a high fraction of population) are on low-lying areas susceptible to floods. Richer cities will build levees and dykes; poorer cities have fewer resources available. Within cities the richer people can move to higher land or commute, leaving the poor behind.

Finally there is the direct effect of disease changes from climate change. Again colder areas will see a modest improvement from fewer wintertime deaths but this is likely to be more than offset by more deaths from tropical diseases like malaria as well as heat waves.

Point 3: As poor people become richer, this worsens the pace of climate change

As the previous part of the course has set out in detail, richer people buy more stuff and nearly all of this 'stuff' has a significant carbon footprint, whether eating more meat, driving cars, living in bigger nicer houses, or jetting to holidays in warm places.

The IPCC report notes that human societies have two main strategies against climate change: Adaptation & Mitigation.

Adaptation is taking steps to survive and prosper under new climactic conditions: society's stocks of capital were designed for particular conditions and rebuilding these is expensive. If roads near the coast need to be redirected, if subways and tunnels need new pumping systems, if water systems need more pipes, if flood protection needs higher walls – all of these substantial public infrastructure projects are expensive. While these projects are counted as adding to future GDP, they come at the cost of investing in other areas or direct consumption.

Humans have been successful in adapting to many different climates. But again the important note is that the rich and powerful will likely adapt easily; the poor and powerless will face huge problems. Poverty reduction goals set by the UN and other development agencies will be much more difficult to meet.

Mitigation is taking steps to reduce carbon emissions so that the Global Climate Change is smaller. Climate change has already begun; even if humans stopped emitting carbon now there is enough inertia that the average temperature would still climb more before stabilizing.

The choice is not between Climate Change or no; the choice is how much Climate Change we are willing to accept.

Mitigation strategies must identify which industries and human activities emit the most greenhouse gases and then figure out which of these emissions can be reduced most easily.

One of the most important mitigation strategies is for governments to create the proper incentives, often to stop subsidizing harmful activities and to begin discouraging them (for example, coal mining is often subsidized by government policies).

The Stern Report clearly summarizes the common position of many economists that a carbon tax or a cap-and-trade policy, which can have identical effects, would be the most effective way to reach a targeted level of emission reduction.

These policies to reduce climate change will reduce GDP now and in the future. Larger reductions mean larger costs. Given the complexities of estimation of climate and economy, a cost-benefit analysis seems implausible. How complex? Consider the records of economic forecasters and weather forecasters; combine them. Seriously, it is very complicated because any scenario of future carbon emissions has to take account of how much carbon is emitted by the economy, which is going to be chosen by the society over the next century, as well as the future path of poverty reduction. (Poor people have few resources and so don't pollute much but a reasonable policymaker would want more richer less-polluting people.)

The particularities and the complications are innumerable; the field of climate science is still developing rapidly. Nevertheless the uncertainties in the science are relatively small compared with the uncertainties about human behavior in the future – the global climate a century from now will be most affected by policy choices.

The paper, "The Economics of Global Climate Change: A Historical Literature Review," by Stern, Jotzo, and Domes, is a useful overview. Also the Stern Report and Nordhaus' replies.

See additional readings