

Lecture Note 8a — General Equilibrium in a Pure Exchange Economy

David Autor, MIT Economics and NBER

14.03/14.003, Microeconomic Theory and Public Policy, Fall 2025

(compiled October 20, 2025)

1 Motivation

So far in 14.03/003, we have discussed one market at a time: labor, sugar, rental properties, etc. But this one-market-at-a-time approach is a convenient fiction—not always a badly misleading fiction, but it is still a fiction. Markets are always interrelated: Reducing sugar tariffs reduces sugar prices; this reduces employment of sugar cane workers in the U.S.; cane workers apply for other farm jobs reducing wages (slightly) among farm workers generally; arable land is freed for other uses; new crops are planted; the price of other farm products fall; real consumer incomes rise; rising consumer income increases demand for sweets; the dessert market grows and the cafe sector booms, etc. Literally, there is no end to this chain of events. But this raises the question: is there a *general equilibrium* where all of these connected markets equilibrate simultaneously? The answer is yes, and that's the topic we turn to now. All changes in quantities or prices ultimately feed back into the demand and/or supply for all other goods through several channels:

- Changes in the abundance/scarcity of resources
- Substitutability/complementarity of goods whose prices rise/fall
- Income effects: Changes in the real costs of goods also affect consumer wealth, which then affects uncompensated demand for other goods and services

To understand this richer story, we need a model that can accommodate the interactions of all markets simultaneously and allows us to determine the properties of the grand equilibrium. In other words, we need a *general equilibrium* (GE) model, in contrast to the *partial equilibrium* (PE) models we have used thus far this term.

2 The Edgeworth Box

To make the general equilibrium problem tractable, we want to reduce the dimensionality of the “all markets” problem to something manageable without sacrificing the essence of the problem. The eponymous Edgeworth box (after Francis Ysidro Edgeworth, 1845–1926) provides the tool we need. As it turns out, we require only two goods and two people to capture the essence of General Equilibrium. The Edgeworth box visually demonstrates the gains in welfare that may accrue from pure exchange of goods, and it perfectly expresses the economic concept of opportunity costs. Simple though it is, the Edgeworth Box allows us to intuitively demonstrate

(though not rigorously prove) two of the most fundamental results in economics: the First and Second Welfare Theorems.

[Note: We will not model or analyze the production of goods in this setting, only pure exchange. The extension of the GE model to production is fascinating in its own right and worth studying. I have regretfully concluded that 14.03/14.003 simply has too many important topics to cover to leave room for GE with production. If you would like to explore the rudiments of this topic on your own, please ask for my GE lecture notes from 14.03/14.003 Spring 2003 or consult a textbook.]

2.1 Edgeworth box, pure exchange: Setup

- There are two goods: call them food F and shelter S .
- There are two agents: call them A and B .
- The initial endowment is:

$$\begin{aligned} E_A &= (E_A^F, E_A^S) \\ E_B &= (E_B^F, E_B^S) \end{aligned}$$

- The consumption of A and B are denoted as:

$$\begin{aligned} X_A &= (X_A^F, X_A^S) \\ X_B &= (X_B^F, X_B^S) \end{aligned}$$

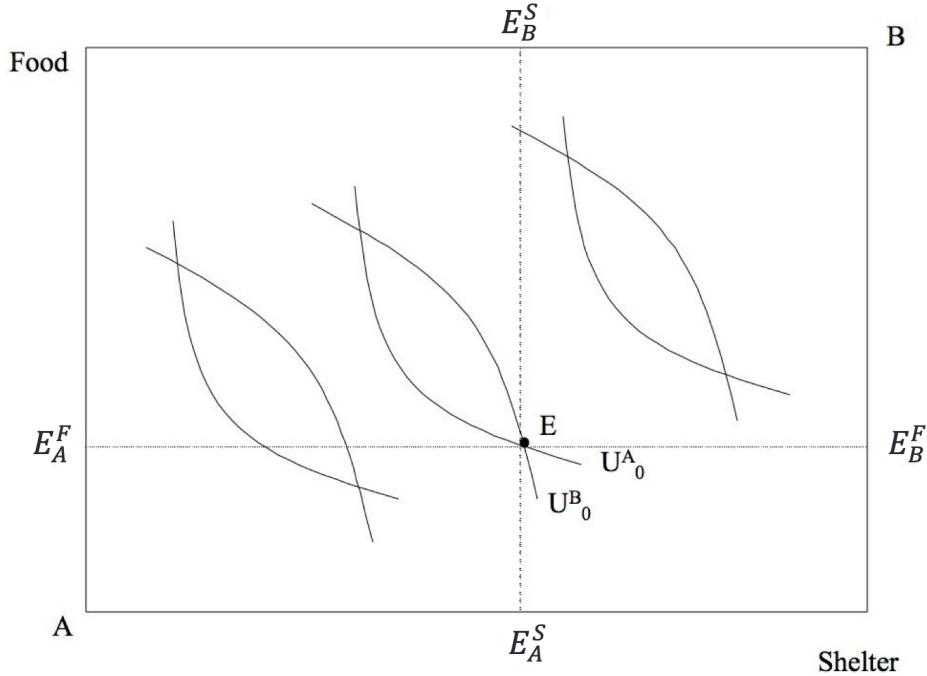
- *Without trade* between agents A and B , their consumption bundles will equal their endowments:

$$\begin{aligned} X_A &= E_A \\ X_B &= E_B \end{aligned}$$

- *With trade*, many exchanges between A and B become feasible, but the following equalities must always hold:

$$\begin{aligned} X_A^F + X_B^F &= E_A^F + E_B^F \\ X_A^S + X_B^S &= E_A^S + E_B^S \end{aligned}$$

That is, total consumption of each good is equal to the total endowment of each good.



- The figure above is called an Edgeworth box. It simultaneously depicts the preferences and consumption of two agents. The left and bottom edges of the figure frame the single consumer A's consumption bundles and preferences (i.e. the familiar figures from Lecture 3 and 4). The top and right edges of the figure show the rotated version of the same figure for consumer B (i.e. the familiar figure rotated 180 degrees).
- Note the elements of this figure:
 - All resources in the economy are represented – the x axis is $E_A^S + E_B^S$ and the y axis is $E_A^F + E_B^F$
 - The preferences of both parties are represented.
 - The notion of opportunity costs is clearly visible.

2.2 Market conditions

We assume that any trade that takes place between A and B satisfies the following four conditions:

(C1) *No transaction costs.* That is, neither F or S is consumed merely through the act of trading.

(C2) *No market power.* Although A and B are the only two agents in this market, we assume that each takes price as given and announces his demand for each good accordingly. That is, neither one strategically “withholds” his goods from the market to raise prices, nor does he anticipate that buying more of one good or another may raise its price. Although it seems a bit contrived for the agents in a two-person economy to act as price-takers, this assumption is realistic in an economy with many agents. Here, we impose the price-taking assumption because we don’t actually want to add another 100 agents to the model.

(C3) *No externalities.* A ’s utility depends only on his own consumption of F and S , and similarly for B . There is no sense in which A ’s consumption of F or S indirectly affects B or vice versa (e.g., through pollution, jealousy, etc.)

(C4) *Full information.* Both A and B are fully informed about the goods available for trade. This rules out the possibility that B sells A rotten food or A sells B shelter that happens to have a leaky roof.

(C5) *Property rights are complete.* All utility-relevant goods are owned by someone. (This condition is in fact weakly implied by the no externalities condition (C3), as we will discuss later in the semester in the context of the Coase Theorem.)

2.3 What happens when A and B trade?

- Starting from point E , the initial endowment, where will both parties end up if they are allowed to trade?
 - It is not fully clear because either or both could be made somewhat better off without making either worse off—that is, there are many Pareto-improving allocations that are feasible. But it’s clear that they need to be *somewhere* in the lens shaped region between U_A^0 and U_B^0 .
- How do we know this?
 - Because all of these points **Pareto dominate** E : One or both parties could be made better off without making the other worse off.
 - In other words, there are potential gains from trade: A would prefer more food and less shelter, B would prefer less food and more shelter.

- So hypothetically

A gives up $E_A^S - X_A^S$

A gains $X_A^F - E_A^F$

B gives up $X_A^F - E_A^F$

B gains $E_A^S - X_A^S$

- All points in the lens region are not equally beneficial. We use the concept of Pareto efficient to determine the point where the agents have the largest gains from trade.
- Q: What needs to be true at a Pareto efficient allocation?
 - A : The indifference curves of A, B are tangent. Otherwise, we could draw another lens.
 - So trading should continue until a Pareto efficient allocation is reached.

2.4 Pareto efficient allocations

1. At a Pareto efficient allocation, it is not possible to make one person better off without making at least one other person worse off
2. At a Pareto efficient allocation, all gains from trade are exhausted

How can you see a Pareto efficient allocation in the Edgeworth box? At a Pareto efficient allocation, the indifference curves of A, B will be tangent.¹ The set of points that satisfy this criterion comprise the Contract Curve (CC). All Pareto efficient allocations lie along this curve. Hence, after trade has occurred, the final allocation will lie somewhere on CC that passes through the lens defined by the points interior to U_A^0 and U_B^0 .

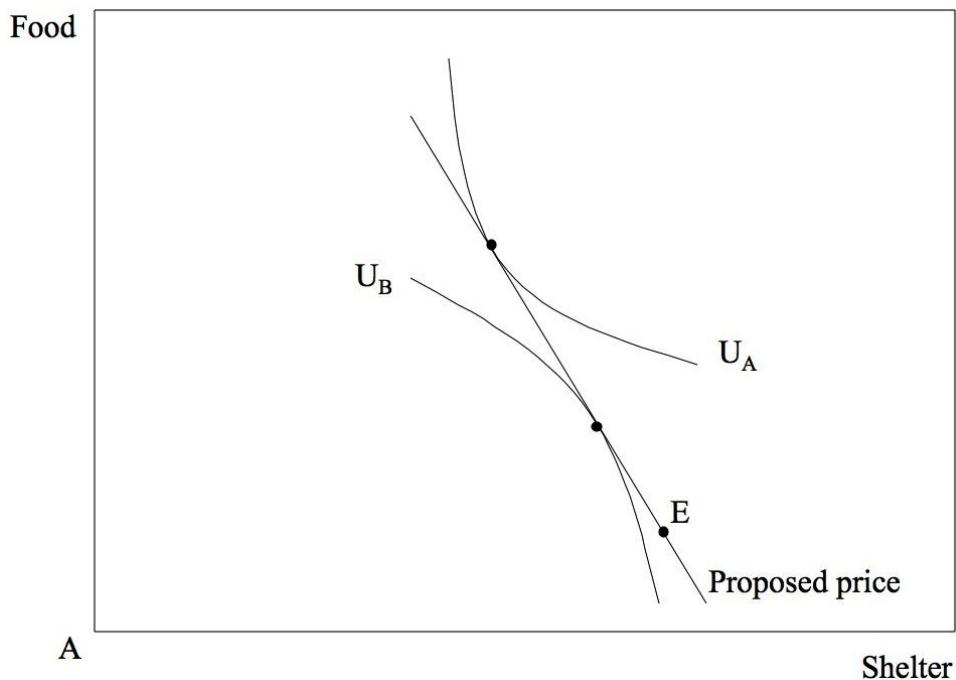
Note: In some examples, the Edgeworth box will not have a contract curve. That's because, for problems that yield a corner solution, there will likely be no points of tangency between the indifference curves of the two trading parties. But there may still be a set of Pareto efficient points (on the edges) that dominate the initial allocation. For example, if A values good X but not good Y and vice versa for B , there will be no tangency points and the only Pareto efficient allocation will involve giving the entire endowment of X to A and the entire endowment of Y to B .

¹Except in the case of a corner solution. Imagine if A didn't like shelter and B didn't like food. There is only one Pareto efficient allocation in this case, and it is at a corner.

2.5 How do we get from E to a point on the contract curve?

Famous analogy: Auctioneer (Leon Walras \rightarrow Walrasian auctioneer).

1. In the initial endowment, *the market clears (that is, all goods are consumed) but the allocation is not Pareto efficient.*
2. So, an auctioneer could announce some prices and then both parties could trade what they have for what they preferred at these prices.
3. Problem: *Choices would then be Pareto efficient but would not necessarily clear the market.*
4. It's possible there would be extra F and not enough S or vice versa.
5. So, must re-auction at new prices...



At proposed prices:

- A wants to reduce consumption of shelter and increase consumption of food
- B wants to increase consumption of shelter and decrease consumption of food
- But, A wants to increase consumption of food more than B wants to decrease

- A wants to decrease consumption of shelter more than B wants to increase
- So:

$$X_A^F + X_B^F > E_A^F + E_B^F \Rightarrow \text{Excess demand}$$

$$X_A^S + X_B^S < E_A^S + E_B^S \Rightarrow \text{Excess supply}$$

- What should the auctioneer do? Raise P^F/P^S .
- When the auctioneer gets the price ratio correct, the *market clears*. No excess demand or supply for any good. This is a market equilibrium (also known as a competitive equilibrium or a Walrasian equilibrium). In this equilibrium:
 - Each consumer chooses his most preferred bundle given prices and his initial endowment.
 - All choices are compatible so that demand equals supply.
 - There is Pareto efficient consumption (i.e. ‘Allocative Efficiency’):

$$\left(\frac{\partial U/\partial F}{\partial U/\partial S} \right)_A = \left(\frac{\partial U/\partial F}{\partial U/\partial S} \right)_B$$

- How do we know Allocative Efficiency will be satisfied?
 - Because both A, B face the same prices P^F/P^S .
 - Each person’s optimal choice will therefore be the highest indifference curve that is tangent to her budget set given by the line with the slope P^F/P^S that intersects E .
 - Because these choice sets (for A, B) are separated by the price ratio, we know they will be tangent to one another but will not intersect. (If we consider an economy with many goods, we can think of the equilibrium goods prices as forming a ‘separating hyperplane’—which is a generalization of a plane to more than two dimensions—that divides the indifference maps of consumers to create the desired tangency condition across all goods).
- This equilibrium price ratio *will exist* provided that:

- Each consumer has convex preferences (diminishing marginal rate of substitution) as we assumed during consumer theory.
- Or, each consumer is small relative to the aggregate size of the market so that aggregate demand is continuous even if individual preferences are not. (This is obviously not relevant in the two person case represented by the Edgeworth box.)

2.6 Aside: How do we know that both (all) markets clear simultaneously?

How do we know that both (all) markets clear simultaneously?

- Again consider two goods F, S (food and shelter) and two individuals A, B .
- As above, label A 's demand and supply (endowment) of each good as $X_A^F, X_A^S, E_A^F, E_A^S$ and similarly for consumer B .
- Consumer A 's budget constraint can be written as:

$$\begin{aligned} P^F X_A^F + P^S X_A^S &= P^F E_A^F + P^S E_A^S, \\ P^F (X_A^F - E_A^F) + P^S (X_A^S - E_A^S) &= 0, \\ P^F Z_A^F + P^S Z_A^S &= 0, \end{aligned}$$

where Z_A^F is A 's *excess demand* for food, and $Z_A^F = X_A^F - E_A^F$.

- The excess demand is the amount of a good that consumer A would like to consume relative to her current endowment.
- Excess demands can be positive or negative (so more precisely, there is either excess demand or excess supply).
- The above equation ($P^F Z_A^F + P^S Z_A^S = 0$) states that given an initial supply (endowment) of goods and a set of prices, an individual's total excess demand for goods is zero. Simply put, a consumer cannot buy more than the value of the goods she holds, since the value of these goods is her budget constraint.
- A similar budget holds for consumer B :

$$P^F Z_B^F + P^S Z_B^S = 0.$$

- Putting these excess demand functions together gives,

$$P^F(Z_A^F + Z_B^F) + P^S(Z_A^S + Z_B^S) = P^F Z^F + P^S Z^S = 0.$$

If, as we have established above, $P^F Z^F = 0$, this immediately implies that $P^S Z^S = 0$. *Which is to say, that there cannot be either excess demand or excess supply for all goods simultaneously.*

- This observation—that total excess demand must equal zero—is called Walras’ Law (after Leon Walras). If there are n goods, and there is no excess demand for $n - 1$ of these goods, then there is also *no excess demand* for the n^{th} good.
- Intuitively, we get the n^{th} solution for free because we have one more linear equations than we have unknowns (one more goods than we have price ratios). In this example, we have good F , good S , and one price ratio P^F/P^S . Since it is only the *price ratio*—not the absolute price level—that matters (reflecting the idea that all costs are opportunity costs), then with n goods, the matrix of demands has rank $n - 1$. So, if we solve for the market clearing prices of $n - 1$ goods, we have also obtained the market clearing price of the n^{th} good. Note that this also means that we can set the price of one good equal to a constant (usually 1) and all other prices are implicitly in units of that good’s price (the good with price one is called the *numeraire*). E.g. if we set the price of food to 1, then found a price of shelter of 1.7, then that means that our price of shelter is 1.7 times the price of food.
- In our two-good exchange economy above, this proves that if the market for food clears with no excess demand or excess supply, then the market for shelter clears simultaneously.

2.7 Aside #2: How do we solve these problems in practice?

- In general, we can solve for the equilibrium in a very similar fashion to the way that we solved utility maximization problems in partial equilibrium. After all, we are still assuming that agents are maximizing their utility, subject to constraints.
- For a two agent case, we can use our Lagrangian toolkit (constrained optimization) to solve:

$$\max_{X_A^F, X_A^S} u_A(X_A^F, X_A^S) \text{ s.t. } E_A^F P^F + E_A^S P^S \geq X_A^F P^F + X_A^S P^S$$

What’s the constraint? It’s the income constraint, now in terms of endowments. We can set up and solve the same problem for agent B.

- How do we combine our solutions for agents A and B (which will each be in terms of the prices)? We use our market clearing principle! If the market clears (i.e. the total endowment of each good equals the combined demands of the two agents), then that will pin down what the relative prices must be to avoid having excess demand.

3 How are equilibrium prices set? The First Welfare Theorem

You do not need the auctioneer.

- Leon Walras loosely proved that the market can reach this equilibrium without assistance from a central planner, that is, without an auctioneer (okay, Walras asserted this and couldn't actually prove it, but his conjecture was correct). This result—the existence of general equilibrium as a self-organizing outcome of the market—is fundamental. The description that Walras used was that the economy would reach equilibrium through a process of Tattonment (literally, “groping”). (See Banerjee Chapter 6 for a somewhat more formal derivation.)
- This equilibrium is a result of:
 1. Endowments of all consumers
 2. Preferences/tastes of all consumers (stemming from utility functions)
 3. In a model with production: technologies for turning factors (land, labor, capital) into goods
- Notice that we previously said in the Partial Equilibrium (PE) model of consumer choice that the consumer’s optimal consumption bundle is a function of three things:
 1. The consumer’s preferences
 2. The market price ratio
 3. The consumer’s budget
- In the General Equilibrium (GE) model, these three items each have a direct mapping:
 1. Preferences are a primitive in *both* models

2. The price ratio in the PE model is an emergent property of the GE model stemming from preferences, endowments and technologies. That is, while the PE model prices are exogenous, in the GE model, they are endogenous.
3. The budget in the PE model corresponds to the endowment in the GE model. However, there is an important difference between the two models. In the PE model, the consumer's budget set is taken as given. In the GE model, the budget is determined by the interaction between preferences and endowments. So, although the consumer has an exogenous endowment in the GE model, the corresponding budget set—that is, the bundles that the endowment can be traded for—is determined by the equilibrium of the model.

4 Efficiency

4.1 First Welfare Theorem: A free market, in equilibrium, is Pareto efficient

What Walras showed, and what is clear from the Edgeworth box, is that a competitive market will exhaust all of the gains from trade. That is, it will be Pareto efficient.

Note that the following stringent conditions must be satisfied for this result to hold:

- (C1) No externalities
- (C2) Perfect competition
- (C3) No transaction costs
- (C4) Full information

Under these conditions, the First Welfare Theorem guarantees that the market equilibrium will be Pareto efficient. A bit later in the semester, we will begin to examine what happens to market equilibria and market efficiency when these conditions are *not* satisfied. We will particularly focus on the market maladies that stem from externalities and imperfect information.

4.2 Another take on the First Welfare Theorem

- We can think of the General Equilibrium problem as a utility maximization subject to three constraints:

1. No actor is worse off in the market equilibrium than in the initial allocation. This will always hold because an agent could always refuse to trade and consume her original endowment. Thus, no party can be made worse off by trade *relative to her initial endowment*.
2. In equilibrium, no party can be made better off without making another party worse off (otherwise there are non-exhausted gains from trade).
3. No more goods can be demanded/consumed than the economy is endowed with (a resource constraint).

3a No goods are left unconsumed—that is, there is no excess supply. This is not truly a constraint—it's simply a property of any equilibrium, which follows from non-satiation.

- The *First Welfare Theorem* says that the free market equilibrium is the solution to the above problem. Simply by allowing unfettered trade among atomistic market actors, the market solution—that is, the price vector and resulting equilibrium choices—will satisfy the three constraints above.
- This is an important and non-obvious result. It implies that the decentralized market continually “solves” a complex, multi-person, multi-good maximization problem that would be difficult for any individual (or large government agency) to solve by itself due to the information requirements.
- Of course, markets are not always (or necessarily ever) “in equilibrium,” and conditions (C1) - (C4) for efficiency are not always (or necessarily ever) satisfied. So, the market solution may not be perfect. But one should also ask: would a “central planner” generally do better? We will discuss this question at various points throughout the semester.

4.3 Second Welfare Theorem

- Q: Does the First Welfare Theorem guarantee that the market allocation will be “fair” or equitable? Of course not! The First Welfare Theorem simply says that the market will enlarge the pie as much as possible; it has nothing to say about who gets what share. If, for example, the initial endowment had *A* consuming all goods and *B* consuming nothing, and assuming that *A* had standard preferences, then the initial allocation would be Pareto efficient; there are no further gains from trade available to *A* and *B*. Here, the Pareto efficient market allocation would also be maximally inequitable.

- This raises a fundamental question: Is there a trade-off between enlarging and dividing the pie—that is, between efficiency and equality?
- Stated rigorously, given a Pareto efficient allocation of resources, will there exist prices and an initial endowment such that this allocation is *also* an equilibrium? Concretely, can *any* Pareto efficient allocation be supported as a competitive equilibrium?
- If the answer to the above questions is **yes**, then there is no intrinsic trade-off between efficiency and equality. If the answer is **no**, then clearly there is a trade-off.
- The Second Welfare Theorem says that the answer to these questions is **yes**.

Second Welfare Theorem: Providing that preferences are convex and conditions C1-C4 are satisfied, any Pareto efficient allocation can be supported as a market equilibrium.

The reasons are self-evident in the Edgeworth diagram (though this is a far from a proof):

- Along the contract curve, every point represents the tangency point of two indifference curves
- This tangency point corresponds to a price ratio that separates the two tangent indifference curves
- This price ratio clearly must exist if the indifference curves are tangent and each is convex (so they don't recross at some later point)
- This price ratio is therefore the market price vector that will support that particular Pareto efficient allocation.

Hence, it is immediate from the Edgeworth box that all Pareto efficient distributions—that is, all points on the Contract Curve—are feasible as market equilibria. As long as the assumptions above are met, a competitive equilibrium will exist merely because each person is self-interestedly maximizing her own well-being.

The Second Welfare Theorem therefore implies that *there is no intrinsic trade-off between equity and efficiency.* [Notice that the converse is also generally true: non-Pareto efficient allocations cannot be attained in equilibrium.]

When we discussed partial equilibrium welfare analysis (as in case of the U.S. Sugar Program or the market for real estate brokers), we implicitly assumed that it was justifiable to maximize the sum of producer and consumer surplus, rather than worrying about their division. The Second Welfare Theorem is the result that justifies that approach.

4.4 If we don't like the distribution of wealth in the market equilibrium, how do we change it?

How do we get from one Pareto efficient allocation to another? It would seem that there are two tools available: lump-sum redistribution (i.e., where I reallocate food and shelter from *A* to *B*) and taxation to change the price ratio so that a different equilibrium obtains.

But these tools are not equivalent. What happens when we change the price ratio (by fiat) in this model to achieve some alternative equilibrium? The answer is clear from studying the Edgeworth box.

5 Interpreting the Fundamental Welfare Theorems

The fundamental welfare theorems provide some very basic policy guidance:

- The function of the price mechanism is to ensure that all resources are consumed in a Pareto efficient fashion—that is, all gains from trade are exhausted.
- Under assumptions C1 through C4, this occurs automatically as prices adjust to clear the market.
- Distorting the price system to achieve equity is generally not a good idea. Such distortions generally *do* create a trade-off between efficiency and equity—which is exactly what the Welfare theorems say we do *not* need to do.
- This does not mean we should ignore equity, however. We can achieve whatever equitable allocations of resources is desired through lump-sum distributions.

5.1 Is this dictum—don't distort prices—always correct?

- **No.** Because the strong assumptions underlying the Welfare Theorems are not always—or perhaps ever—satisfied.
- But the first and second welfare theorems do build a *prima facie* case that free market outcomes may be efficient—or at least hard to improve upon.
- When there is a case to be made for intervening in market outcomes (and there often is), this case should depend upon:
 - A reasoned diagnosis as to why the market allocation is not optimal.

- A policy prescription that builds on an analysis of how a specific intervention will remedy this fault.
- A careful accounting of the likely distortions (deadweight losses) that will result from tampering with the price system.
- Improving on market outcomes generally benefits from a rigorous analysis of why these outcomes are *not* desirable and, preferably, a proposed correction that harnesses the useful properties of markets to improve the outcome.

5.2 Are the welfare theorems non-obvious?

MIT economist and Nobel laureate Paul Samuelson once said, “There are few things in economics that are both universally true and non-obvious.” The fundamental welfare theorems are arguably one of those exceptional things. Why would anyone assume that prices are anything other than arbitrary social creations? This insight—that the free market system generates a Pareto efficient equilibrium through the endogenous emergence of prices—is one of the great insights of classical economics. Economic theory suggests that market equilibria (and *prices themselves*) have a fundamental logic that is an *emergent property* of the rational, atomistic actions of market participants.

The key insight: Blind pursuit of self-interest by autonomous actors in a market setting yields *collectively welfare maximizing behavior*. Under certain (strong) assumptions, this equilibrium *cannot be improved upon* without making at least one person worse off (Pareto efficiency).

Adam Smith published *The Wealth of Nations* in 1776. It’s clear that Smith intuitively understood the First Welfare Theorem when he wrote:

“It is not from the benevolence of the butcher, the brewer, or the baker, that we expect our dinner, but from their regard to their own interest. We address ourselves, not to their humanity but to their self-love, and never talk to them of our necessities but of their advantages...”

“Every individual necessarily labors to render the annual revenue of the society as great as he can. He generally indeed neither intends to promote the public interest, nor knows how much he is promoting it. ... He intends only his own gain, and he is in this, as in many other cases, led by an invisible hand to promote an end which was no part of his intention. ... By pursuing his own interest he frequently promotes that of the society more effectually than when he really intends to promote it. I

have never known much good done by those who affected to trade for the public good.”

Smith had apparently convinced himself of the first welfare theorem, though it’s not clear that he thought of the second. But it was 150 years until either welfare theorem was proved.

- Pareto and Barone proposed the 1st and 2nd welfare theorems formally in the 1930s.
- These theorems were proved graphically in 1934 by Abba Lerner.
- They were proved mathematically by Oskar Lange in 1942 and Maurice Allais in 1943 (for which Allais won the Nobel Prize in 1988).
- It was not until 1954 that papers by Lionel McKenzie and, independently, by Kenneth Arrow and Gerard Debrau, proved the existence of general equilibrium in a market economy.

Prior to Adam Smith—and long afterward—market behavior has been viewed with great suspicion. An example from Helibronner (1953), *The Worldly Philosophers* (New York: Touchstone).

In 1639 in Boston, the respected merchant Robert Keayne was charged with a crime: He had made over sixpence profit on the shilling, an outrageous gain. The Boston court debated whether to excommunicate him for his sin. In view of his spotless past, the court instead fined him 200 pounds (a huge sum!). Keayne was so distraught over his sin that he prostrated himself before the church elders and “with tears acknowledges his covetous and corrupt heart.”

The minister of Boston could not resist the opportunity to make an example of Keayne. In his Sunday sermon, he used the example of Keayne’s avarice to denounce “some false principles of trade:”

That a man might sell as dear as he can, and buy as cheap as he can. [Buying low, selling high.]

If a man loses by casualty of sea, etc., in some of his commodities, he may raise the price of the rest. [A reduction in supply may increase the market price.]

That he may sell as he bought, though he paid too dear. [Selling at a price that the market will bear.]

That free markets may produce socially desirable outcomes is a fundamental insight of economics. Two-hundred and forty years after Smith wrote *The Wealth of Nations*, this idea is not widely understood outside of the economics profession, though it has had a profound effect on the organization of modern economies.