Motivations for Public Economics

Introduction to Public Economics

Governments play a crucial role in much economic life.

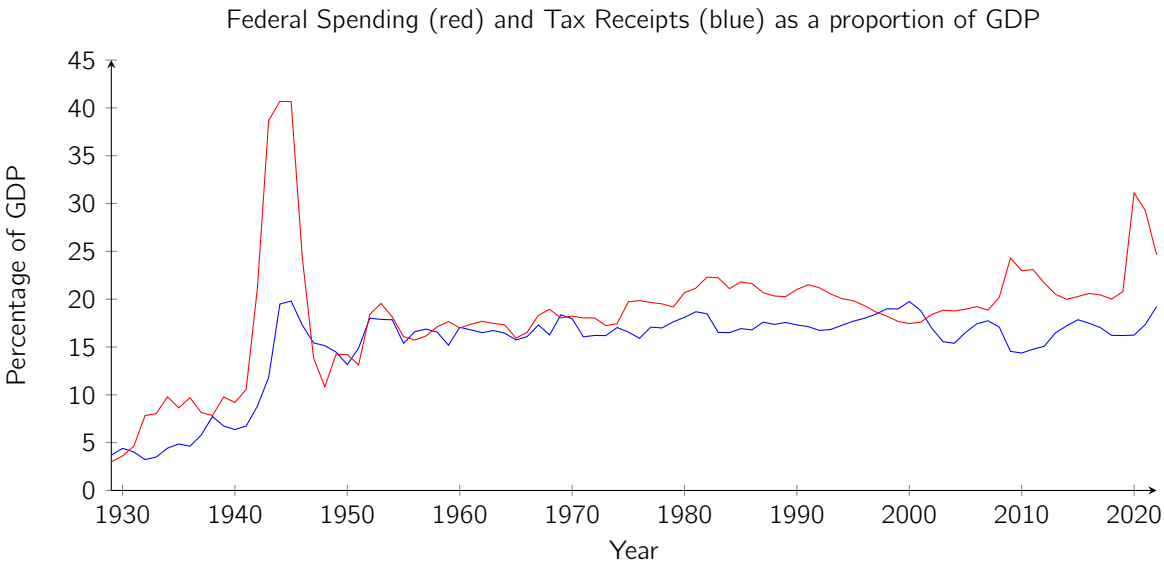
- Regulatory structure (financial markets, pharmaceuticals, labor markets, civil rights).
- Taxes.
- Public goods and social welfare spending.
- Macroeconomic stabilization.

Public finance is the study of the role of government in a market economy, primarily focusing on taxes and spending.

Reasons to study public economics:

- Governments have a lot of power in the realm of economic welfare.
- Nearly every economic transition is mediated by the government.
- It can inform debates about the appropriate role of government regarding taxes, healthcare, climate change, etc.
- The government is large.
  - It employs 1/6 of the US Workforce.
  - Tax revenue is approximately 27% of the United States's Gross Domestic Product.

The government (as measured by tax revenue/GDP) greatly increased in size between 1910 and 1940 (due to the establishment of the welfare state and various wars).



Two Motivations for Government Intervention

- Market Failure
- Redistribution

The First Welfare Theorem states that *in the absence of market failure*, markets will yield a result along the **utility possibilities frontier** (i.e., the set of all maximized utilities given the current market).
However, there are a lot of market failures:

- Externalities (pollution, network effects from vaccination)
- Public Goods (public safety)
- Asymmetric Information (market for lemons)
- Individual Mistakes (failure to save)
- Imperfect Competition (oligopoly, cartelization)

Policymakers also have to consider the *equity-efficiency tradeoff* in redistribution (i.e., some redistributive acts might reduce total utility)

Government as Social Cooperation

- Economists tend to have a narrow view of human behavior, but social cooperation undergirds much of the levels of societal coordination beyond individuals (i.e., families, communities, countries, global superstructures)
- Human societies of old depended on social cooperation for protection and taking care of the young, sick, and old.
- Modern states are the primary form of coordination today.
- Humans reveal their social nature (or social solidarity) via the size of the government (informal and formal).

Introduction Activity

Activity: Introduction to Public Economics

Econ 308

Brandon Lehr

Dani Rodrik and Stefanie Stantcheva (2021) begin their recent working paper as follows:

> One of the biggest challenges that countries face today is the very unequal distributions of opportunities, resources, income and wealth across people. Inclusive prosperity — whereby many people from different backgrounds can benefit from economic growth, new technologies, and the fruits of globalization — remains elusive. To address these issues, societies face choices among many different policies and institutional arrangements to try to ensure a proper supply of productive jobs and activities, as well as access to education, financial means, and other endowments that prepare individuals for their participation in the economy. In this paper we offer a simple, organizing framework to think about policies for inclusive prosperity.

1. This framework is illustrated with the 3x3 matrix below. Where might the following public policies fit in this matrix?

   (a) cash transfers to low income people
   (b) estate/inheritance taxation (based on estate of deceased)
   (c) healthcare; primary education
   (d) minimum wage; apprenticeships
   (e) on-the-job training; labor laws; protectionist trade policy
   (f) progressive income taxation; wealth taxes; corporate taxes
   (g) public higher education
   (h) R&D tax credits; antitrust policy
   (i) social insurance (e.g., unemployment insurance, disability insurance, Social Security)
   (j) universal basic income

		At what stage of the economy does policy intervene?		
		Pre-Production Stage: shape the endowments with which people enter the workforce	**Production Stage:** shape the employment, investment, and innovation decisions of firms	**Post-Production Stage:** redistribute income and wealth after they have been realized
Which income group is the target of the policy?	**Bottom**	c	d	a f i j
	Middle	c g	d e	f i j
	Top	c	h	b f i j

2. At which stage do you think it is most important for policy to intervene? Why?

   Pre-production: helping to equalize endowments to equalize opportunity

Microeconomic Foundations: Consumer Theory

Utility function $u(X, Y)$ translates consumption quantities into utility.

Indifference curve A graphical representation of all bundles of goods that make an individual equally well

off. Mathematically, an indifference curve is the set of all bundles (X, Y) such that $u(X, Y) = U$ for some utility level U.

Marginal Rate of Substitution MRS_{XY} is the negative slope of the indifference curve — it's the rate at which the consumer will trade Y for X.

$$MRS_{XY} = \frac{\partial u/\partial X}{\partial u/\partial Y}$$

Budget Constraint the set of all bundles for which the total amount spent equals income

- Let I indicate income and P_X and P_Y represent the prices of goods X and Y respectively.
- The budget constraint is the line segment $P_X X + P_Y Y = I$.
- The slope of the budget constraint is $-\frac{P_X}{P_Y}$.

Utility Maximization A rational consumer maximizes utility subject to the budget constraint via the parallel conditions of tangency ($MRS_{XY} = \frac{P_X}{P_Y}$) and the budget constraint ($P_X X + P_Y Y = I$).

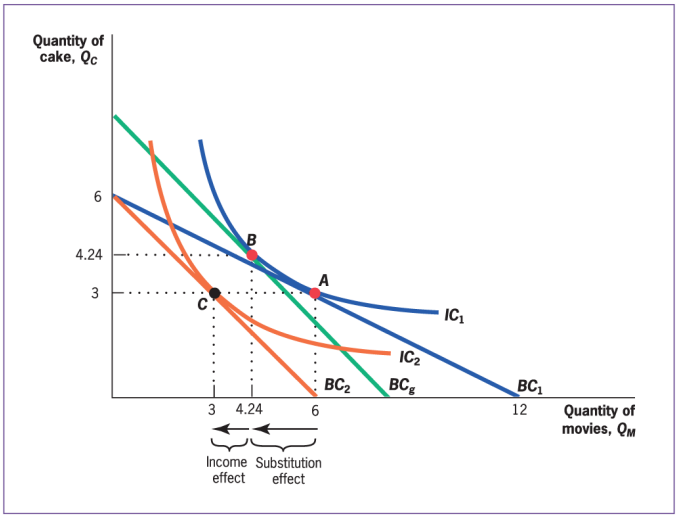
Demand Functions Utility maximization generates demand functions (quantity in terms of price) $X(P_X, P_Y, I)$ and $Y(P_X, P_Y, I)$. There are two primary canonical utility functions.

- Cobb-Douglas: $u(X, Y) = A\ln(X) + B\ln(Y)$, or $u(X, Y) = X^A \cdot Y^B$. The demand function for this utility function yields that P_X has no effect on Y and P_Y has no effect on X.
- Quasilinear: $u(X, Y) = v(X) + BY$ where $v'(X) > 0 > v''(X)$ (i.e., concave down, sloping up). The demand function for this utility function yields that I has no effect on X assuming an interior solution.

Price Effects The impact of a change in P_X on demand for X is composed into two effects:

- Substitution Effect: change in consumption due to the change in relative prices, with utility held equal. When the price of a good increases, the substitution effect is always negative, and vice versa.
- Income Effect: change in consumption due to a change in purchasing power as a result of the price change, where relative prices are held constant at the final price ratio. Income effects can be positive or negative depending on the type of good.
- The total effect is equal to the income effect and the substitution effect.

Income and Substitution Effects



Gruber, *Public Finance and Public Policy*, 6e, © 2019 Worth Publishers

Price Elasticity The price elasticity of demand is the % change in demand caused by a 1% change in price of a good.

$$E^D = \frac{dD}{dP}\frac{P}{D}$$

Elasticities are *unit-free*, typically negative, and tend not to be constant along a demand curve.

Game Theory

Some decision problems involve strategic interactions between individuals.

- For example, Antonia and Bruno might care about giving to a local charity, and give G_A and G_B respectively.
- Their utility functions depend on each other $u_A(G_A, G_B)$, $u_B(G_A, G_B)$.
- The **Nash Equilibrium** yields each individual choosing an action that maximizes their utility *given the other person's behavior*.

Social Welfare

Economists incorporate distributional concerns by use of social welfare functions.

$$SWF = f(U_1, U_2, \ldots, U_n)$$

We have two canonical social welfare functions:

- Utilitarian SWF: $SWF = U_1 + U_2 + \cdots U_n$.
  - Marginal utility decreasing in income $\rightarrow$ redistribution from the rich to the poor.
  - Taking \$1 from a rich person decreases their utility by a small amount, but transferring to a poor person increases their utility by a large amount.
- Rawlsian SWF: $SWF = \min\{U_1, U_2, \ldots, U_N\}$
  - Social welfare is maximized by maximizing the well-being of the worst-off person.
  - Rawlsian social welfare is more redistributive than utilitarian social welfare.

There are a few other philosophies regarding the fairness of economic distribution in society.

Just deserts Individuals should be compensated in line with their contributions.

Commodity egalitarianism Society should ensure that individuals meet a set of basic needs, but beyond that point income distribution is irrelevant.

Equality of opportunity Society should ensure that all individuals have equal opportunities for success.

Present Discounted Value

The present discounted value of a future value of money F that is received and spent in n periods is:

$$PDV = \frac{F}{(1+r)^n}$$

For the **discount rate** r, typically the interest rate.

For a stream of future expenses F_i, we use the following formula:

$$PDV = \sum_{i=1}^{n} \frac{F_i}{(1+r)^i}$$

If the values of F_i are equal, then $PDV = \frac{F}{r}$, via the geometric series formula.

Theoretical Tools Activity

Activity: Theoretical Tools of Public Economics

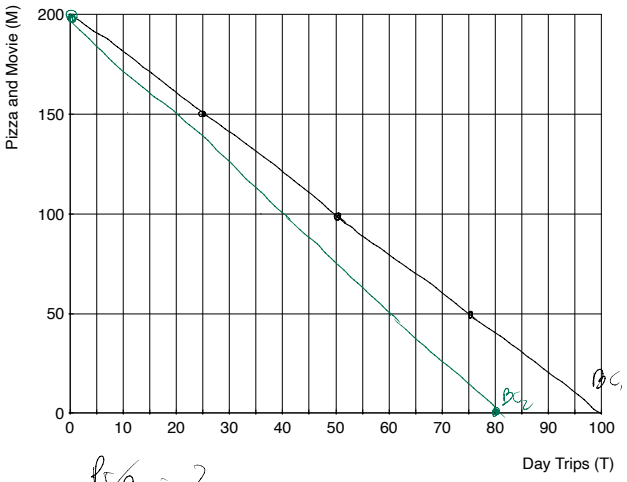
Econ 308

Brandon Lehr

1 Gruber 2.14: Consumer Choice

You have \$4,000 to spend on entertainment this year (lucky you!). The price of a day trip (T) is \$40 and the price of a pizza and a movie (M) is \$20. Suppose that your utility function is $u(T, M) = T^{3/4} \times M^{1/4}$.

a. Draw the budget constraint below. What combination of T and M will you choose?



$P_T/P_m = 2$

$\frac{\partial u}{\partial m} = \frac{1}{4}\left(\frac{T}{M}\right)^{3/4}$

$\frac{\partial u}{\partial T} = \frac{3}{4}\left(\frac{M}{T}\right)^{1/4}$

$\frac{\frac{\partial u}{\partial T}}{\frac{\partial u}{\partial m}} = 3\frac{M}{T}$

$2 = \frac{3M}{T}$

$3M = 2T$

$40T + 20M = 4000$

$60M + 20M = 4000$

$80M = 4000$

$M = 50$

$T = 75$

b. Suppose that the price of day trips rises to \$50. Draw the new budget constraint in the same plot in part (a). What combination of T and M will you now choose?

$P_T/P_M = 3/2$

$MRS_{TM} = 3M/T$

$3T = 6M$

$50T + 20M = 4000$

$60M + 20M = 4000$

$M = 50$

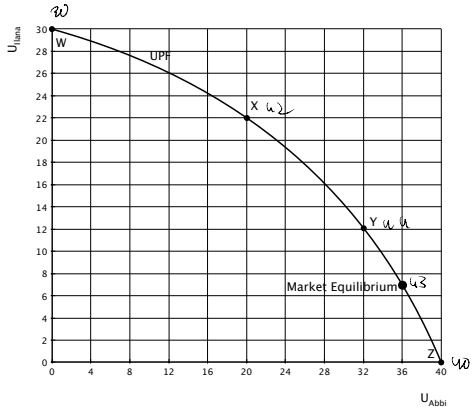
$T = 60$

2 Bonus: Social Welfare Functions

The utility possibilities frontier (UPF) drawn below corresponds to an economy with only two individuals, Abbi and Ilana. There is a high level of inequality at the market equilibrium outcome, i.e., Abbi is rich, while Ilana is poor. Which of the four labeled points (W, X, Y, Z) is most preferred by a social planner with:

a. a Utilitarian social welfare function? (Y)

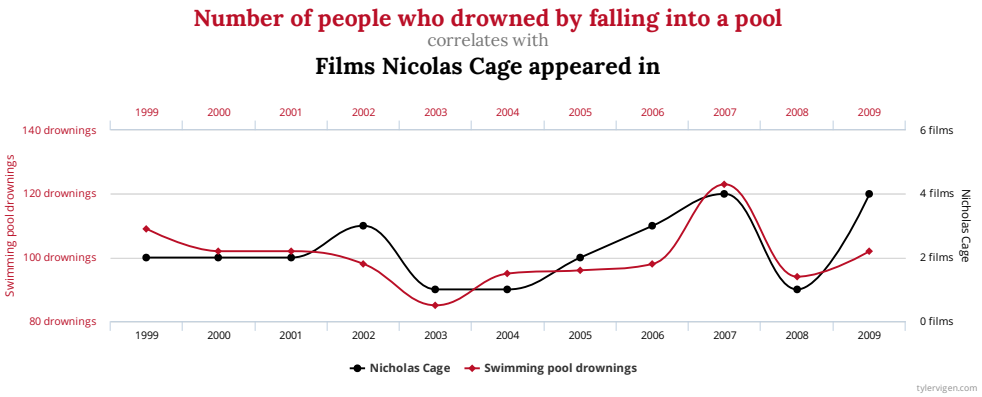
b. a Rawlsian social welfare function? (X)



Empirical Tools of Public Economics

Identification Problem: When two variables, A and B, are correlated with each other, how do you identify which is causing which (or if both are being caused by something else)? Could the correlation be spurious?

Spurious Correlation



Randomized Controlled Trials

The gold standard of experimentation.

- One group obtains the treatment, one gets the placebo.
- The two groups must have close to identical traits aside from the treatment, which is part of the "control."

Randomized controlled trails tend to be common in science, and have grown in popularity in the social sciences, but for economics, RCTs are often difficult to carry out (after all, you cannot split up the timeline).

Limitations of RCTs

External validity: it's difficult to apply the outcomes of a randomized controlled trial to other contexts.

Attrition: unequal loss of participants (i.e., people in the control group drop out more than the treatment group, or vice versa).

Unethicality: little feedback on the part of participants, and the interests of participants aren't necessarily taken into account.

Observational Data

Data generated by individual behavior observed in the real world, not in specially-designed experiments.

- Time Series analysis: changes in two different stats over time and trying to find results from said changes.
- Cross-Sectional Regression Analysis: finding a relationship between two variables with many data samples.

$$Y_i = \alpha + \beta X_i + \epsilon_i$$

The OLS estimate, $\hat{\beta}$, is the slope of the data. In order for this relationship to be causal, the error term, ϵ_i, must be uncorrelated with X_i.

Quasi-Experiments

Changes in the economic environments that create nearly identical treatment and control groups for studying the effect of that environmental change, allowing economists to take advantage of quasi-randomization created by external forces.

Empirical Tools Activity

Activity: Empirical Tools of Public Economics

Econ 308

Brandon Lehr

1 Air Quality and Covid-19

There have been a number of recent studies that estimate a positive association between regions with poor air quality and Covid-19 mortality rates.[1]

a. Write-ups about such studies in the popular press often suggest that poor air quality increases the severity of illness from Covid-19. Why might this inference be unwarranted?

There are other confounders like poverty, population density, etc. that affect the rate of mortality

b. Suppose you were tasked with empirically testing the hypothesis that poor air quality impacts Covid-19 mortality rates.

i. Suggest a randomized controlled trial to test this hypothesis. Assume away any legal, financial, or ethical constraints.

Take two groups, put one in a polluted area and one in a clean area, and give both groups COVID

ii. Suggest a plausible quasi-experiment to test this hypothesis.

LA County mortality rate in days with vs. without wildfires.

[1] https://www.lung.org/blog/covid-19-mortality-and-air-pollution

2 Gruber 3.14: Tax Cut Effects

Your state introduced a tax cut in 2017. You are interested in seeing whether this tax cut has led to increases in personal consumption within the state. You observe the following information:

Year	Consumption in Your State
2012	330
2014	350
2016	370
2018	430

a. Your friend argues that the best estimate of the effect of the tax cut is an increase in consumption of 60 units, but you think that the true effect is smaller because consumption was trending upward prior to the tax cut. What do you think is a better estimate?

40 units (trend before vs. trend after)

b. Suppose that you find information on a neighboring state that did not change its tax policy during this time period. You observe the following information in that state:

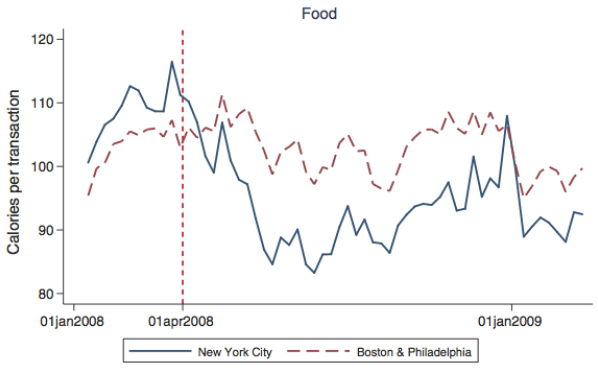
Year	Consumption in Neighboring State
2012	300
2014	320
2016	340
2018	350

Given this information, what is your best estimate of the effect of the tax cut on consumption? What assumptions are required for that to be the right estimate? Explain.

50 units (assumes similar trend would otherwise happen)

3 Bonus: Starbucks Calorie Posting

The Affordable Care Act (Obamacare) included nutritional labeling requirement for restaurants with at least 20 locations. A similar policy was enacted on April 1, 2008 in New York City (but nowhere else). Bollinger et al. (2011) study the effect of the New York City law on caloric purchases at Starbucks, summarizing their data in the figure below:



The average food calories per transaction were:

	New York City	Boston and Philadelphia
Before Law	107	102
After Law	87	96

a. What is the difference-in-differences estimate of the causal effect of the NYC labeling law?

-14

b. Do you find this to be a surprising result? Why or why not?

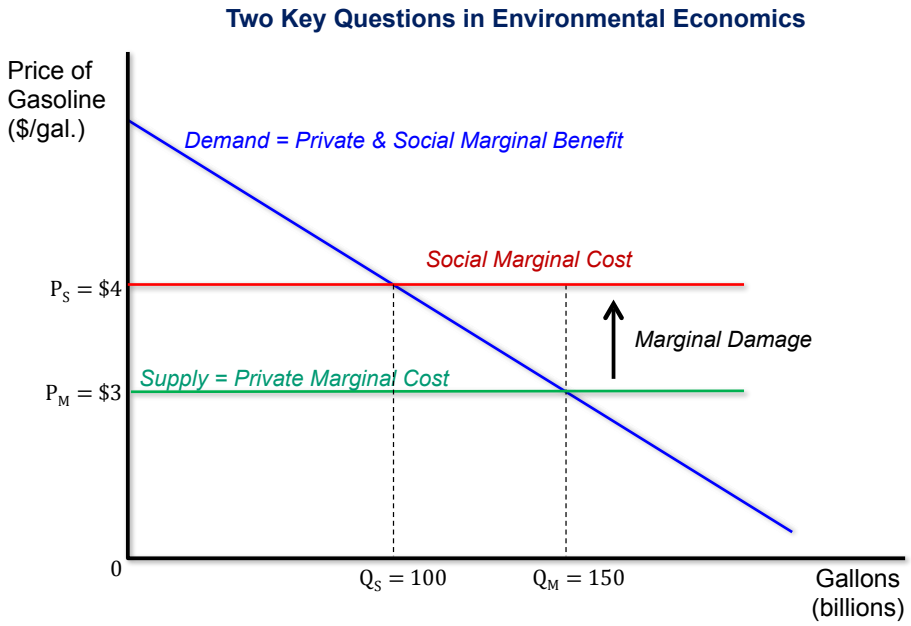
No, people tend to be self-conscious.

Market Failures

Market Failures: Externalities

The most classic market failure is *externalities*: when the fully private costs/benefits and the social costs/benefit are misaligned. For example, possibly the most important example of a *negative* externality is climate change from CO_2 emissions, a byproduct of industrial development.

Economists focus on balancing the total costs (private + social) of pollution with the total benefits (private + social) from pollution.



The top triangle represents the deadweight loss because the problem of underpriced externalities is one of *overproduction*, not of underproduction. We are facing two questions:

- How do we calculate the marginal social cost of pollution?
- What is the best way to reach the marginal social cost of pollution?

Estimating the social cost of carbon

Step 1: How does one extra ton of CO_2 impact the climate?

Step 2: How does a marginal change in climate affect various human outcomes?

Step 3: Calculate the current social cost by converting future costs into current dollars via discounting.

For Step 2, economists control for the particular location, and do comparisons over time, and check the average effect across all the locations.

After calculating or estimating future costs, economists need to discount these costs toward the present value.

$$PDV = \frac{C_1}{1+r} + \frac{C_2}{(1+r)^2} + \cdots$$

where r is the *social discount rate*, the rate at which society is willing to trade off between consumption today and consumption tomorrow.

Changes in r matter a lot:

- If r is large (i.e., we don't care much about future generations), then carbon costs are not large.
- If $r = 0$ (i.e., we care equally about all generations), then carbon costs can be infinite.

Various economists calculated different social discount rates:

- Giglio, Maggiori, and Stroebel calculated 2.6%, by using differences in price between perpetual ownership and 100 year leases.
- The Obama administration used a 3% discount rate, implying the social cost of carbon was \$42/ton.
- The Trump administration used a 7% discount rate, implying the social cost of carbon was \$5/ton
- The Biden administration uses an estimate of \$51/ton.

Policy Solutions

We can use *pigouvian taxes* to force market participants to internalize the cost of the externality.

Alternatively, we can use *industrial policy*, to restructure the market in order to pursue a public goal.

- Subsidizing clean energy.
- Research in clean energy technologies both public and private.
- Phase out carbon in various economic sectors and weaken fossil fuels.

We have seen great dividends so far: solar and wind have become extremely cheap to produce, and replacing coal with natural gas has yielded emissions reductions that were greater than targets. The Biden administration focused on making clean energy even cheaper via the Inflation Reduction Act.

Externalities Activity

Activity: Externalities: The Case of Climate Change

Econ 308

Brandon Lehr

1 Gruber 6.1: Environmental Policy and Justice

There is concern that California's cap-and-trade system for greenhouse gases, implemented in 2012, has led to "hot spots" of pollution — localized areas with very high concentrations of hazardous pollutants such as air toxics and particulate matter. A recent study by Cushing et al. (2018) showed that socioeconomically disadvantaged communities have been disproportionately exposed to these air pollutants under California's cap-and-trade program. This has happened despite the fact that overall greenhouse gas emissions and exposure to pollutants in the state have been reduced significantly under the program.

a. How might a cap-and-trade system lead to such "hot spots"?

Changes to set up a market use/pollution source in poorer areas + richer environmentalists can buy permits.

b. How would a utilitarian and a Rawlsian each evaluate overall social welfare from this cap-and-trade system?

Utilitarian: pro - overall harm net pollution decreases

Rawlsian: poorer people worse off → anti - cap and trade

c. And what are some policy solutions to mitigate the disproportionate harm?

Local govt. buys trade permits in poorer areas, performance standard for all factories.

2 Gruber 6.18: Getting Pigouvian Taxes Right

Suppose that the (external) damage done by pollution is known to be $MD = 300 + 5Q$, and the (private) cost and benefit are given by $MC = 100 + 2Q$ and $MB = D_0 - 2Q$, where D_0 is not precisely known.

a. If $D_0 = 1,000$, what would be the optimal quantity? What tax would be necessary in order for that to be the equilibrium quantity?

$100+2Q = 1000-2Q$

$4Q = 900$

$Q = 225$

$P = 550$

$4Q + 7Q = 1000 - 2Q$

$9Q = 600$

$Q = 200/3$

$T = \frac{1000}{3}$

b. (Bonus) Suppose that, based on the result from part (a), a tax is imposed to allow the optimal quantity of pollution to be produced. If $D_0 = 900$, what would be the deadweight loss associated with having the wrong tax level?

$400+7Q = 900-2Q$

$9Q = 500$

$Q = \frac{500}{9}$

$T = 300 + 5\left(\frac{500}{9}\right) = \frac{5200}{9} = \frac{1300}{3}$

$H = 200$

$B = \frac{100}{9}$ → $DWL = \frac{10,000}{9}$

Public Goods

A public good is not something that's merely good for the public, but it is specifically a good that is rival and non-excludable (essentially, the same quantity of the good has to be available to every person.)

The optimal level of publi goods is when the *vertical* sum of individual demand curves equals the

marginal cost of providing the public good. The reason the demand curves are added vertically is because market participants can share the public goods.

Recall that MRS_{GX} of a public good G for a private good X is how much an individual values an additional unit of G in terms of unit of X.

$$MRS_{GX} = \frac{MU_G}{MU_X}$$

The total number of units of X society is willing to give up for 1 more unit of G is the sum of all MRS_{GX}.

Assuming that $P_X = 1$, this is a measure of how many dollars society is willing to pay for 1 more unit of G. The societal value is the sum of the different MRS_{GX} values:

$$MC_G = \sum_{I=1}^{N} MRS_{GX}^I$$

This is the *Samuelson Rule*.

Private Provision of Public Goods

- The private outcome is the Nash equilibrium of a game where individuals choose how to allocate income between G and X, taking into account spending on G by others.
- The Nash equilibrium does not satisfy the Samuelson Rule.
- Public goods problems can be described as **free-rider** problems (i.e., underproduction).
- There is not necessarily zero private provision of public goods. Private provision works well in the following cases:
  - Some people have much higher marginal rate of substitution (i.e., they care more than others).
  - Altruism: people care purely about giving to others.
  - Warm Glow: people get utility from giving to others.

There is experimental evidence of free-riding; for example, Marwell and Ames (1981) had an experiment where they tested whether subjects were willing to contribute to a public good, where the Nash equilibrium was that people didn't contribute, and the social optimum was everyone contributing.

People are willing to cooperate at first, but then get upset as time goes on and retaliate.

Public Goods Example

Let there be 2 people, with the same utility functions over X, F, where X is a private good and F is a public good:

$$U_i(X_i, F) = 2\ln(X_i) + \ln(F) \qquad \text{for } i = 1, 2$$

Each participant has a budget constraint, where P_X and P_F are both 1.

$$X_i + F_i = 100 \qquad \text{for } i = 1, 2$$

Where $F = F_1 + F_2$.

Maximize Person 1's utility: (i.e., finding the best response function)

$$
\begin{aligned}
0 &= \frac{\partial U_1}{\partial F_1} \\
0 &= -\frac{2}{100 - F_1} + \frac{1}{F_1 + F_2} && \text{recall that } X_1 = 100 - F_1 \\
2(F_1 + F_2) &= 100 - F_1 \\
F_1 &= \frac{100 - 2F_2}{3}
\end{aligned}
$$

Since this is a symmetric game, we know that the best response function for Person 2 is $F_2 = \frac{100-2F_1}{3}$.

The Nash equilibrium is when all players are playing their best response to each other. Since the game is symmetric, we know that $F_1^* = F_2^*$

$$
\begin{aligned}
F_1^* &= \frac{100 - 2F_1^*}{3} \\
F_1^* &= 20 \\
F_2^* &= 20
\end{aligned}
$$

The social optimum (following the Samuelson Rule) is as follows:

$$
\begin{aligned}
MRS_{FX}^1 + MRS_{FX}^2 &= 1 \\
\frac{MU_F}{MU_{X_1}} + \frac{MU_F}{MU_{X_2}} &= 1 \\
\frac{1}{2}\frac{X_1}{F} + \frac{1}{2}\frac{X_2}{F} &= 1 \\
2 &= \frac{X_1 + X_2}{F} \\
F &= \frac{X_1 + X_2}{2} \\
F &= \frac{200 - F}{2} \\
F &= \frac{200}{3}
\end{aligned}
$$

Public Provision of Public Goods

The primary problem of public provision of public goods is *crowding out* (i.e., reduced private contributions to public goods). There are two key assumptions of one-to-one private contributions:

- Individuals do *not* have warm glow preferences.
- Private actors are contributing to the public good.

There is empirical evidence of partial crowd-out (i.e., reduction in charitable giving as government spending is increased). However, since people tend to give to charity due to warm glow preferences and social pressures, the crowd-out is not one-to-one. Hungerman (2005) estimates that the crowd-out effect is 20–40 cents per dollar of welfare spending.

Public Goods Activity

Activity: Public Goods
Econ 308

Brandon Lehr

1 Gruber 7.9: Estimating Crowd-Out

Let's consider why it is difficult to empirically determine the degree to which government spending crowds out private provision of public goods.

a. Discuss some challenges with interpreting a correlation between government spending and private provision of a public good in observational data.

b. Experimental investigations of crowding out can isolate the causal effect, but these experiments typically take place in laboratory environments. What are some drawbacks to this type of evidence?

2 Gruber 7.14: Private vs. Optimal Provision of Public Goods

The town of Springfield has two residents: Homer (H) and Bart (B). The town currently funds its fire department solely from the individual contributions of these residents. Each of the two residents has a utility function over private goods (X) and total firefighters (M) of the form:

$$U_H = 6\ln(X_H) + 2\ln(M)$$
$$U_B = 6\ln(X_B) + 2\ln(M)$$

The total provision of firefighters hired, M, is the sum of the number hired by each of the two persons: $M = M_H + M_B$. Homer and Bart both have income of \$100, and the price of both the private good and a firefighter is \$1. Thus, they are each limited to providing between 0 and 100 firefighters.

a. How many firefighters are hired if the government does not intervene? How many are paid for by Homer? By Bart?

$\max \; 6\ln(100 - M_H) + 2\ln(M_H + M_B)$

$0 = \frac{-6}{100 - M_H} + \frac{2}{M_H + M_B}$

$\frac{6}{100 - M_H} = \frac{2}{M_H + M_B}$

$6M_H + 6M_B = 200 - 2M_H$

$M_H = 25 - \frac{3}{4}M_B$, $M_B = 25 - \frac{3}{4}M_H$

Symmetric $\rightarrow M_H^* = M_B^*$

$M_H^* = 25 - \frac{3}{4}M_H^*$

$M_H^* = \frac{100}{7} = M_B^*$

$M = M_H^* + M_B^* = \frac{200}{7}$

b. *Bonus.* What is the socially optimal number of firefighters? If your answer differs from (a), why?

$MRS_{(MX_H)} = \frac{1}{3}\frac{X_H}{M}$

$MRS_{(MX_B)} = \frac{1}{3}\frac{X_B}{M}$

$\frac{200-M}{3M} = 1$

$M = 50$

Free rider problem

Three Questions in Public Economics

Descriptive: What are the effects of interventions and policies (empirical)?

Normative: What is the optimal policy (theoretical)?

Public Choice: Why do governments choose the policies they do (mixture of theory and empirics)?

Rules of Social Decision

There are at least 2 individuals with transitive preferences over at least 3 options. A **social decision rule** aggregates these preferences into a social preference over these options.

Suppose we want our social decision rule to satisfy the following properties:

Transitive: if a is ranked above b by our rule, and b is ranked above c by our rule, then a has to be ranked above c by our rule.

Pareto Efficiency: if everyone prefers a to b, then our rule should rank a above b.

Independence of Irrelevant Alternatives: the ranking of any two options depends only on how individuals rank the options (and not the ranking of other alternatives).

Non-dictatorship: There is no individual whose preference ranking of any two options matches the social ranking (no matter the preferences of others).

All of these seem reasonable, so they must be possible, right? Well...

Arrow's Impossibility Theorem

Arrow's Impossibility Theorem is the following:

> There is no social decision rule that satisfies the properties of Universality (i.e., transitivity), Pareto Efficiency, Independence of Irrelevant Alternatives, and Non-dictatorship.

The implication of Arrow's Impossibility Theorem means that the only voting method that satisfies the first three properties of our social rule is a dictatorship. However, since we don't want a dictatorship, we can use the following:

- Restrict preferences (i.e., transitive is insufficient)
- Relax Independence of Irrelevant Alternatives, and let intensity of preferences play a role (for example, the Samuelson rule).

Case Study: Majority Voting

Pairwise majority voting is a mechanism to aggregate individual votes into a social decision. Since Majority Voting is obviously Pareto efficient and non-dictatorial, and satisfies IIA because the ranking of a over b only depends on how many people vote a vs. how many people vote b. Therefore, we must have that majority voting fails transitivity.

Consider an election for funding public schools:

Majority Voting "Working"

Preferences	Type of Voter		
	Parents (1/3)	Elders (1/3)	Young Couples (1/3)
First	H	L	M
Second	M	M	L
Third	L	H	H

- $M >_s L$
- $L >_s H$
- $M >_s H$

Therefore, $M > L > H$, and so M is the social choice made by pairwise majority voting.

Majority Voting "Failing"

Preferences	Type of Voter		
	Public School (1/3)	Private School (1/3)	Young Couples (1/3)
First	H	L	M
Second	M	H	L
Third	L	M	H

- $M >_s L$
- $L >_s H$
- $H >_s M$

We are now stuck in a cycle $M > L > H > M > \cdots$. This election fails to successfully aggregate the preferences of the populace.

A way to get out of the trap of transitivity is to rule out preferences that are not single-peaked (i.e., one local maximum).

However, when we have single-peaked preferences, we have that the preferences of the median voter is that which is preferred by society.

Implications of the Median Voter Theorem

If we restrict our analysis to single-peaked preferences, majority voting is a social decision rule that satisfies all desirable properties.

However, this means majority voting does not imply efficiency. For example, if a public good is efficient, but because a minority has a large marginal benefit and the majority has a small marginal benefit, the public good will get rejected. What matters for efficiency is the *average* marginal benefit across individuals, not the *median* marginal benefit.

However, the median voter theorem doesn't really hold in real life (i.e., Democrats close to the median still vote similar to their caucus, and Republicans close to the median still vote similar to their caucus).

Political Economy Activity

Activity: Political Economy
Econ 308

Brandon Lehr

1 Jail and Voter Participation

a. Suppose that people who receive short jail sentences vote less than people without previous jail sentences. Does this evidence imply that jail sentences reduce voter participation? Explain.

Not necessarily — people who receive jail sentences also tend to be younger or have other factors that may mean they have less frequency in voting.

b. Read the summary of Ariel White's research on the next page. How does she identify the causal effect of short jail sentences on voter participation?

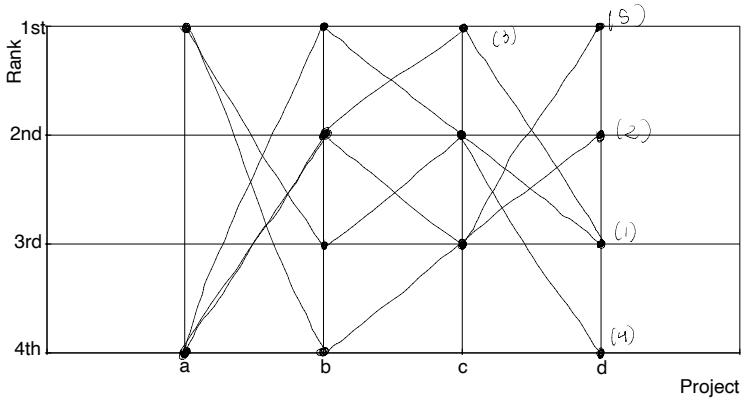
Randomised courtrooms → differences in severity → similar defendants sentenced to different punishments

2 Majority Voting

Consider five people $\{1, 2, 3, 4, 5\}$ with preference rankings over four projects $\{a, b, c, d\}$ as follows:

	1	2	3	4	5
First	b	a	c	a	d
Second	c	d	b	c	b
Third	d	c	d	b	c
Fourth	a	b	a	d	a

a. Draw the preferences by ranking the preferences in the figure below.



b. Who has single-peaked preferences? And who does not?

(1), (3)

c. Which project will be selected by (pairwise) majority voting? If none is selected, explain why.

c > b > a

(c) > b > d > a

→ c is the project that will be selected

3 Bonus: Plurality Voting and Borda Counts

Consider four people $\{1, 2, 3, 4\}$ with preference rankings over three projects $\{a, b, c\}$ as follows:

	1	2	3	4
First	a	a	b	c
Second	b	b	c	b
Third	c	c	a	a

Assume that voters cast their votes sincerely.

a. Plurality voting is a social decision rule that chooses the option that receives the most votes among the alternatives. Under plurality voting, which project wins?

a

b. The Borda count is a social decision rule in which individuals submit their preference ranking, each rank is assigned a point value (higher points for higher ranks), and the option with the most points is selected. Suppose that 3 points are assigned to the first choice, 2 points to the second choice, and 1 point to the third choice. Which project wins?

a: 6 c: 7

b: 9

c. Now find a Borda point system where a wins.

d. What general lessons have you learned from this exercise?

Local Public Goods: the Tiebout Hypothesis

Tiebout (1956) asks the following: What is it about the private market that guarantees optimal provision of private goods that is missing in the case of public goods?

Proposed answer: **shopping** and **competition**. However, we can see this somewhat resolved in the local

level:

- Individuals can vote with their feet (i.e., move between different cities)
- The threat of exit can induce competition in the provision of public goods.

  Just as the consumer may be visualized as walking to a private market place to buy his goods,... we place him in the position of walking to a community where the prices (taxes) of community services are set. Both trips take the consumer to the market. There is no way in which the consumer can avoid revealing his preferences in a spatial economy.

Modeling the Tiebout Hypothesis

Suppose there are $2N$ families, each with identical income Y, and 2 towns with N homes each.

Towns 1 and 2 supply levels G_1 and G_2 of local public goods at $MC_G = 1$.

- N families with kids, with $U^K(C, G)$ over private consumption C and public schools G.
- N elderly families, with $U^E(C)$ over only private consumption C.

The allocation of families across towns is a **Tiebout Equilibrium** if and only if:

(1) No two families want to exchange locations across towns; and

(2) In each town, G is decided by the median voter and financed *equally* by town residents.

Family Budget Constraint: $Y = C + G/N$, where $P_C = 1$ and $P_G = 1/N$

- If the majority in the town is elderly, then $G = 0$, as that maximizes $U^E(Y - G/N)$.
- If the majority in the town is families with kids, then G^* is that which maximizes $U^K(Y - G/N, G)$ such that $MRS^K_{GC} = \frac{1}{N}$.

Tiebout Theorem

Part 1 (Sorting): In equilibrium, families will sort themselves in towns according to their taste for the public good (1 town with elderly only, one town with families with kids).

Part 2 (Efficiency): In each town, the level of local public good is efficient.

- In the elderly town, $G = 0$, which is efficient as nobody values G.
- In the kids town, $\sum MRS^K_{GC} = \frac{1}{N} = 1 = MC_G$, which is the Samuelson Rule for efficiency in public goods provision.

A Tiebout equilibrium may not be socially desirable if there is a public interest in integration within towns and for reducing social distance across groups (i.e., intergroup contact theory).

Assumptions of the Tiebout Hypothesis

(1) High Mobility: may be impeded by (artificial or natural) barriers to entry, preferences over other qualities.

(2) Perfect information: lack of knowledge about the quality of public goods.

(3) Scale: inability to fund public goods for one's preferences.

(4) Variety: enough options of types of towns.

(5) No externalities or spillovers: otherwise, public good will be underproduced.

(6) Equal financing of the public good (poll tax): local public finance generally tends to be based on property or sales tax.

Financing Local Public Goods

- Towns finance their local public goods through property taxation, where the rich pay more than the poor.
- Property taxes induce the poor to chase the rich, while the rich want to segregate themselves from the poor.
- The rich tend to implement mechanisms to prevent free-riding: making houses expensive.
  - Zoning laws restrict supply of housing.
- Discriminatory practices (such as deed restrictions) to stop poorer people from living near rich people.

Tiebout Implications for Redistribution

It's very hard for a local government to redistribute from the rich to the poor (easy exit).

In localities, to avoid migration, public goods financing needs to have strong **tax-benefit linkages** (i.e., the relationship between taxes and the government benefits need to be very explicit and visual).

Higher levels of government can redistribute across communities using taxes to directly or indirectly incentivize public goods spending.

For example, state governments can match, in which case the price of public goods goes from 1 to $\frac{1}{1+m}$, for m being the value of the match. However, with any intergovernmental transfers (either matching or intergovernmental transfer), there is a potential for crowd-out.

Flypaper Effect

Empirical evidence suggests that the crowd-out of state spending by federal spending is low and often close to zero — "the money sticks where it lands."

More recent studies show that while there is a flypaper effect in the short run, there is substantial crowd-out from block grants in the long run.

Local Public Goods Activity

Activity: Local Public Goods

Econ 308

Brandon Lehr

1 Gruber 10.10: Assessing Tiebout

Suppose that appetite for education is uncorrelated across generations: children with a thirst for learning can be born into families that don't share that bent, just as readily as a child born into a highly educated family may have no interest in school. Consider the public good of education from the perspective of students who will use their local education system as a way to invest in their future human capital. What potential implications will a mismatch in parent and student appetite for education have for the efficiency of a Tiebout equilibrium in the funding of schools through local property taxes?

In eq'm, there will be a match among parents who care about education & are among parents who don't, but 50% of the students will be misallocated.

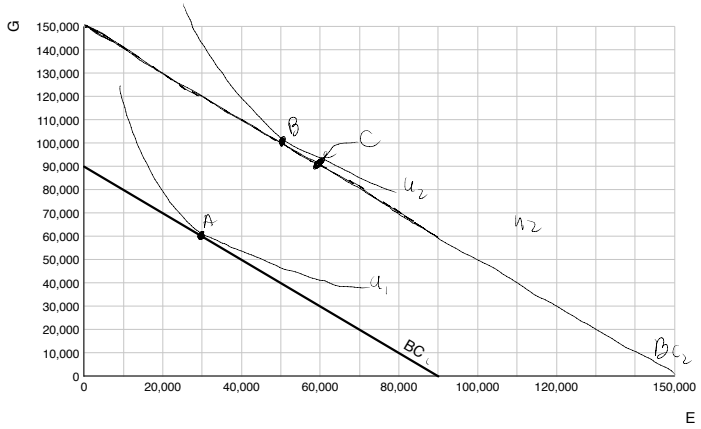
Property taxes will thus be allocated inefficiently for the purposes of education.

2 Block Grants

The city government of Grantsburg must choose how to allocate the city budget of \$90,000 between units of educational quality (E) and all other public goods (G). The city's preferences (i.e., those of the median voter) over E and G are given by:

$$U(E,G) = \ln(E) + 2 \cdot \ln(G)$$

For simplicity, let the unit prices of E and G both equal \$1 so that we can interpret E and G as levels of spending. The city's budget constraint, BC, is depicted below.



a. Determine Grantsburg's choice of E and G. Label this choice as point A in the figure.

$\frac{2E}{G} = 1$ $\quad G = 2E$ $\quad E + 2E = 90K$ $\quad E = 30K$ $\quad G = 60K$

b. If the city is given a \$60,000 unconditional block grant from the state, what is Grantsburg's new choice of E and G? Draw the new budget constraint BC' and label the new choice as point B in the figure.

$E = 50K$, $G = 100K$

c. Finally, suppose that the \$60,000 block grant is conditional — it must be spent on education. Label Grantsburg's choice with this policy as point C in the figure.

$B = C$

Education: Motivations for Government Intervention

Education is one of the largest public goods provided by the government: 6% of GDP and 12.7% of total government expenditure.

However, education is not a pure public good:

- Excludable: private education that charges money, or only allow a particular enrollment area.
- Rival: there are a limited number of seats in a school (capacity constraints, quality degrades if there are too many students).
- Private returns: your investment in education primarily benefits you.

So, why should the government be involved?

- Positive Externalities: a more educated workforce might lead to more productivity and technology.
- Family Failures: if the parental units are unable to provide for education that the child desires, the government involvement is welfare-improving.
- Borrowing constraints: education can be very expensive, and loans are less available (no collateral).
- Behavioral Mistakes: maybe you don't know what's in your best interest.

Education Reform

Supply-Side: Improving the process of education:

- Smaller classes
- Improving teachers
- Charter Schools

Demand-Side: Improving individual funding:

- Vouchers
- Subsidies/Loans/Grants

Class Size

We cannot simply compare outcomes for students in small vs. large classes because of omitted variables (like income/family education).

In Sweden, class size cuts off at 30 students — this yields quasi-experimental variation, which we can use to see causality. Frederiksson, Ockert, and Oosterbeek (2013) found that there was a 4% jump in wages for students in a smaller class.

Teacher Quality

One measure of teacher quality: teacher value-added or test score-based metrics of teacher performance.

- How much does a teacher raise their students' test scores on average (adjusting for noise and controlling for differences in students)?

How do we measure the impact of teacher quality?

- The ideal RCT randomizes students to teachers with different levels of value-added.
- Quasi-Experiment: use the turnover in teachers across school years.

Chetty, Friedman, and Rockoff (2014) use data on all kids who went to NYC public schools in 1989 and link to tax records.

They found an increase of 1.5% in earnings and 1.5% reduction in teenage births going from a 5th percentile teacher to a 95th percentile teacher.

Most school districts do not use any performance measures to evaluate teachers.

Charter Schools

Charter schools are schools financed with public funds that are not usually under the direct supervision of local school boards or subject to state regulation.

To measure effectiveness, one cannot simply compare outcomes at charters with public schools due to different types of students.

Quasi-experiment: compare charter lottery winners with charter lottery losers (for schools that use a lottery).

In Massachusetts, Angrist, Pathak, and Walters (2013) found that urban charter schools boost achievement while non-urban charters reduce achievement.

Charters tend to constitute a market-based approach to education, but there are some limitations:

- Demand-side limitations: Families may not have enough information to choose the best charter schools for their children.
- Supply-side limitations: Schools have an incentive to reject less-qualified applicants (nicknamed "cream-skimming").

Vouchers

With free public education, there is an incentive to *cluster* their total education spend around where the public school's funding is. However, if the money for public education is given equally to all families, it's akin to a conditional block grant.

Vouchers are meant to induce every household to increase their spending on education.

However, there are mixed results from using vouchers:

- Angrist et al. (2002), using lottery assignment of vouchers in Colombia, show positive effects on education.
- Abdulkadiroglu, Pathak, and Walters (2018) used randomized lotteries to evaluate the Louisiana Scholarship program, and found lower achievement by those who win vouchers compared to those who did not.

Government Involvement in Higher Education

- State Provision (about $82 billion)
- Pell Grants ($22.4 billion)
- Direct Student Loans ($34.4 billion)
- Other reliefs ($10.3 billion)

Supply trends:

- Priate non-profit universities have inelastic supply (fixed student bodies).
- Historically, supply for higher education has been administered by public universities (creating new UCs/Cal States/etc.)
- Supply has been provided by for-profit schools, although for-profit schools tend to provide little in benefits to students.

Reduced (direct) state funding has led to higher tuition and student debt.

Additionally, women have been earning many more higher education degrees than men.

Effect of Higher Education on Mobility

Chetty et al. (2017) study parental income and student earning outcomes by college.

Certain schools (Harvard and Berkeley) tend to have high quantities of students in the top 20% of the income distribution, while others like SUNY-Stony Brook tend to have more equal student populations.

They found that education tends to have a flattening effect (i.e., poor kids and rich kids tend to have more similar outcomes when educated).

The mobility rate for a given school can be found by (i.e., the rate of children born in the bottom quintile ending up in the top quintile) can be calculated by finding the success rate (probability that a child moves from the bottom quintile to the top quintile) and multiplying by the proportion of parents in the bottom quintile.

The top college by mobility rate was Cal State LA, and the top ten tended to be open-access public universities (i.e., Glendale Community College, City University of New York, etc.).

Insurance Theory

The **expected utility model** helps study decision under uncertainty.

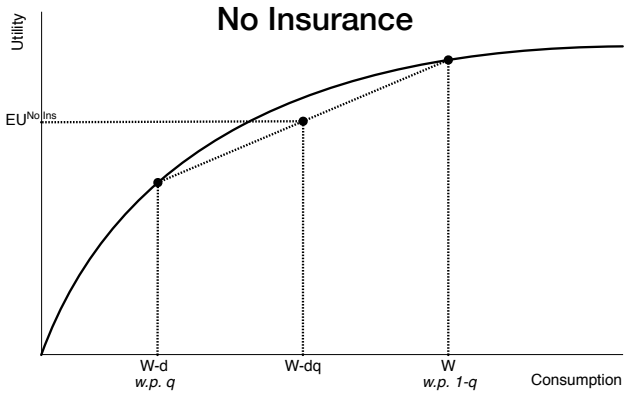
$$E(u(I)) = \sum p_k u(I_k)$$

A **risk-averse** individual is someone who prefers a certain given income to a risky income with the same expected value.

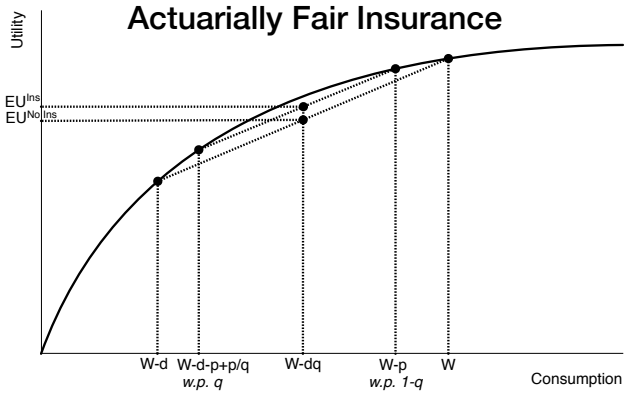
Risk-averse individuals will *demand* insurance in order to smooth their consumption across uncertain states of the world.

Consider health insurance and assume the following:

- $U(c)$ is increasing and concave.
- The person has income W
- The person is sick with probability q
- If sick, the person incurs medical cost d
- The expected utility without insurance: $EU^{\text{No Ins}} = (1-q)U(W) + qU(W-d)$



- In an insurance contract, a person pays premium p always and receives b only if sick.
- The insurance company's expected profit is $E_\pi = p - qb$.
- In perfect competition, $E_\pi = 0$, implying that $b = \frac{p}{q}$, which is when the insurance contract is **actuarially fair**.



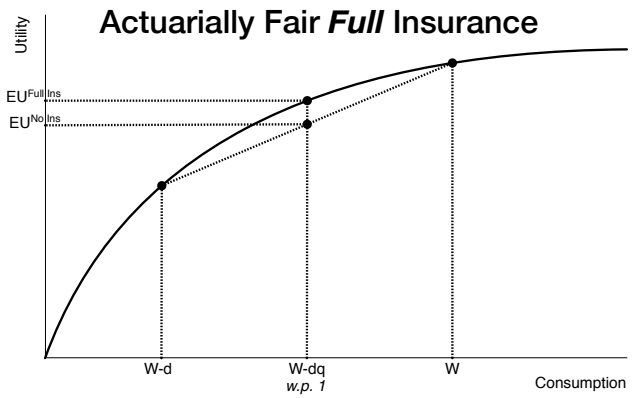
- The expected utility with actuarially fair insurance: $EU^{\text{Ins}} = (1-q)U(W-p)+qU(W-d-p+p/q)$
- If an individual can choose p to maximize expected utility, and they are risk averse:

$$\frac{\partial EU^{\text{Ins}}}{\partial p} = 0$$

$$p = d \cdot q \qquad \text{Premium = Expected Damage}$$

$$b = d \qquad \text{Benefit = Damage}$$

- A risk-averse individual will choose actuarially fair *full* insurance.



- The overall consumption in both states is $W - d \cdot q$, meaning $EU^{\text{Full Ins}} = U(W - dq)$
- The maximum that an individual is willing to pay is p' such that $EU^{\text{No Ins}} = U(W - p')$ (essentially, they have to be at least indifferent to buying insurance with premium p' and remaining uninsured). If $p' > dq$, then the individual will be fully insured, else they will be partially insured at p'.

Social Insurance

- The government provides social insurance for a range of events:
  - Social Security (money in retirement)
  - Unemployment Insurance (job loss)
  - Disability Insurance
  - Workers' Compensation (for job accidents)
  - Medicare (healthcare for the old)
  - Income supports (poverty)
- The growth in government in the past 100 years was primarily due to the growth in social insurance.

However, as we saw in the insurance theory section, we might wonder why the government would get involved in creating social insurance.

The answer is **asymmetric information**.

Case 1 — Symmetric Information: Insurance companies and individuals can observe types.

- If this is allowed, then the insurance companies will offer two actuarially fair policies, $(p_s, b_s = p_s/q_s)$ for sick people, and $(p_h, b_h = p_h/q_h)$ for healthy people.
- Each type i will choose to buy full insurance, $p_i = dq_i$ and $b_h = b_s = d$.
- Private insurance does not equalize incomes across types, only within types: each consumes $W - dq_i$ in each state.

- Pre-existing conditions will lead to inequality in insurance premia and welfare, but no market failure.

Case 2 — Asymmetric Information: Insurance companies cannot observe or price different types, but individuals know their risk.

- If insurance companies offer the same two policies but everyone wants to buy the healthy insurance with lower premia, the insurance company will make losses.
- Insurance companies could offer a single full insurance contract at the *average* actuarially fair price, which is a bad deal for low risk people and they may not buy.
- **Adverse selection** occurs when individuals know more about their risk level than the insurer, meaning higher risk people are more likely to buy insurance.
- Adverse selection can cause the insurance market to unravel in a death spiral, leading to no insurance contract even if everyone would benefit from full insurance.

- The private market can avoid a death spiral in two ways:

  Pooling: The insurance companies may be able to offer one contract that is a good deal for sick people but mediocre (but still net good) for healthy people.

  Separating: Insurance companies offer one full insurance contract for the sick and one partial insurance contract for healthy, which each type self-selects into. This equilibrium is inefficient as the healthy are still under-insured.

- The government's solution is to require insurance.
  - This leads to redistribution from the healthy to the sick.
  - If society views health as primarily determined by luck, such redistribution can receive strong public support.
  - All OECD countries have universal health insurance (and the United States is quite close to universal health insurance)

There are other reasons for public insurance too:

Redistribution: May not be the side effect of addressing adverse selection, but the motivating rationale.

Externalities: Lack of insurance can be a cause of illness for me, thereby exerting a negative externality.

Behavioral Mistakes: Individuals may be myopic or inattentive, leading them to not insure themselves.

Administrative Costs: Administrative burden tends to be lower in public insurance than private insurance.

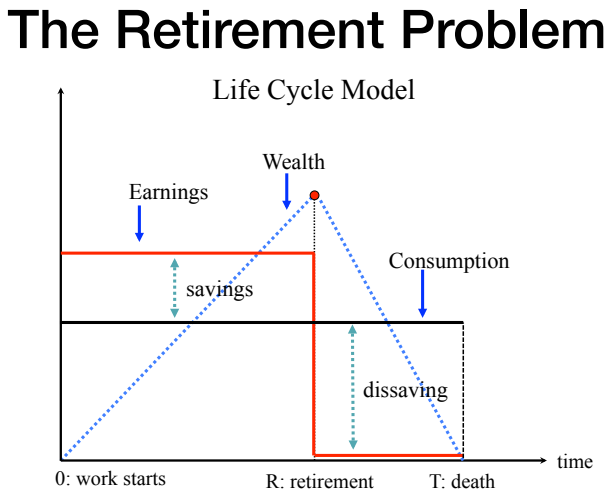
However, there is the problem of **moral hazard**, where people take adverse actions in response to taking on an insurance contract.

- Reduced precaution against entering the adverse state (i.e., driving more recklessly when one has auto insurance)
- Increased odds of staying in the adverse state (remaining unemployed longer after one receives unemployment insurance)
- Increased expenditures when in the adverse state (overspending on healthcare procedures when one has health insurance)

Social insurance systems should partially, but not fully insure individuals (to properly balance consumption smoothing benefits with moral hazard costs).

The Retirement Problem

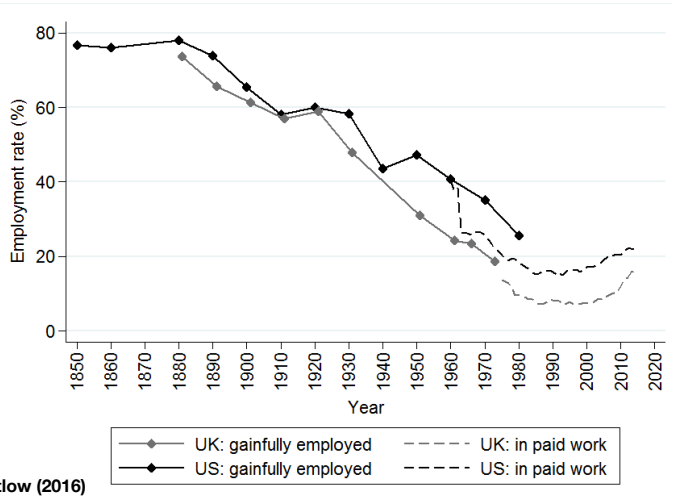
During earning years, people save up money via working in order to smooth their consumption, and wealth tends to peak right around retirement age — if people live for too long, however, their wealth might be "used up" before they die.



Before the advent of the welfare state (starting in the late 1800s, then ramping up in the 1930s and 40s), people above the age of 65 worked a lot more than they do now.

Before the Welfare State, Don't Retire

Figure 2.6: Employment rate of men aged 65+ in the UK and the US



Blundell, French, and Tetlow (2016)

- Most OECD countries provide government-funded retirement programs around 5–10% of GDP.
- Individuals pay payroll taxes while working and receive annuity until death (annuities are essentially equivalent to insurance against living too long)
- It's no longer working kids who take care of you in old age, but all the workers in the country.
- In the U.S., public retirement is called **Social Security**, and is the largest single expenditure of the federal government.

Social Security Details

The payroll tax is a 12.4% tax on earnings up to a cap of $160,000.

Workers who have worked and paid for 40 quarters and are at least 62 years old are eligible.

The annuity payment is a function of the person's 35 highest earning years, where each month's earnings are expressed in today's dollars (AIME: average indexed monthly earnings).

The function for PIA in terms of AIME is expressed as follows:

$$P(I) = \begin{cases} 0.9I, & I \leq 895 \\ 806 + 0.32(I - 895), & 895 < I \leq 5397 \\ 2246 + 0.15(I - 5397), & 5397 < I \leq 6316 \end{cases}$$

One way in which social security is regressive is that rich people tend to live longer than poor people, and will generally receive more in benefits. One of the risks with increasing the retirement age is that doing so may lock out lower income people.

Sources of retirement income:

(1) Social Security: for 2/3 of retirees, SS is >50% of income, 1/3 of elderly households depend almost entirely on SS.

(2) Homeownership: 75% of elderly households are homeowners.

(3) Pensions: 40–45% of elderly US households have pensions:

- Defined Benefit pensions: employer carries full risk.
- Defined Contribution pensions: 401(k)s, employee carries risk.

(4) Savings: about 10% of retirees have significant extra savings.

Most of the wealth in the bottom 90% of retirees is housing + pensions - debts (mortgage, consumer credit, student loans).

Types of Programs

There are two primary forms of retirement programs:

- Unfunded (pay-as-you-go): Benefits of current retirees are paid out of contributions from current workers (also known as a generational link).

  current benefits = current contributions

- Funded: worker contributions are invested in financial assets and will pay for benefits when they retire.

  current benefits = past contributions + market returns on past contributions

- Social Security has traditionally been unfunded, but the trust fund helps finance current benefits out of treasury bonds it purchases with payroll taxes.
- Private retirement plans are funded.

Evaluating Social Security

Recall that optimal social insurance balances consumption-smoothing benefits with moral hazard costs.

People tend not to smooth their own consumption for two main reasons:

- Adverse selection in the annuities market:
  - People with shorter life expectancy tend to be less likely to buy private annuities, increasing

their price.
  - Market could unravel.

- Behavioral mistakes:
  - people tend not to save for their retirement due to myopia, inattention, self-control issues, etc.
  - Social Security's popularity suggests that people understand their own failure to align savings intentions with savings actions.

We model this via a two period life-cycle model with c_1 while working and c_2 in retirement.

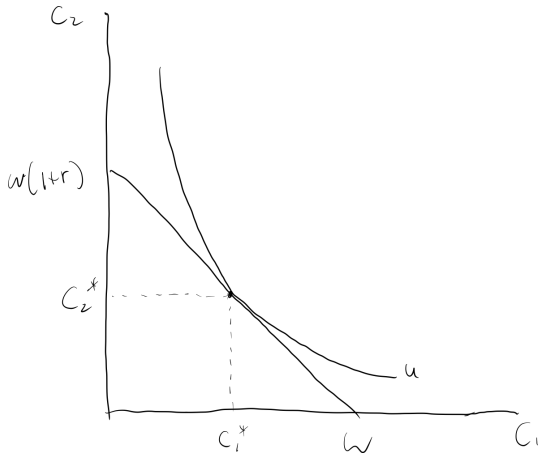
- Period 1 budget constraint: $c_1 + s = W$, where W equals earnings.
- Period 2 budget constraint: $c_2 = s(1 + r)$, where r is the interest rate.
- Our intertemporal budget constraint:

$$c_1 + c_2/(1 + r) = W$$

- We want to find the following:

$$\max_{c_1, c_2} u(c_1) + \delta u(c_2)$$

  such that $c_1 + c_2/(1 + r) = W$, where $\delta \in [0, 1]$ refers to the discount factor.



- The solution is as follows:

$$\begin{aligned} MRS_{c_1,c_2} &= \frac{u'(c_1)}{\delta u'(c_2)} \\ \frac{u'(c_1)}{\delta u'(c_2)} &= \frac{P_1}{P_2} \\ &= (1 + r) \\ u'(c_1) &= \delta(1 + r)u'(c_2) \end{aligned}$$

- Standard assumption: $\delta = \frac{1}{1+r}$

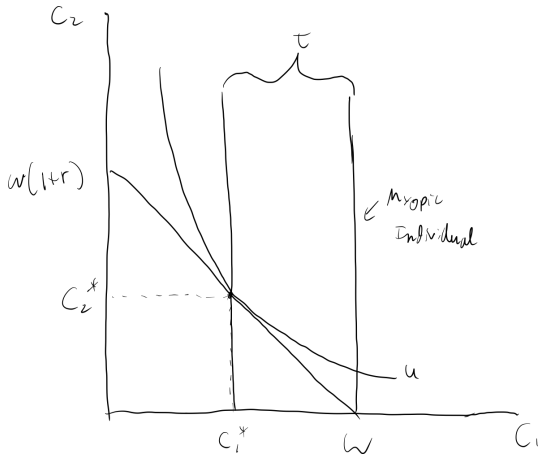
$$\begin{aligned} \frac{u'(c_1)}{u'(c_2)(1 + r)} &= (1 + r) \\ u'(c_1) = u'(c_2) & \\ c_1^* &= c_2^* \\ s^* &= W - c_1^* \end{aligned}$$

- (Extreme) Myopia: $\delta = 0$

$$c_1 = W$$
$$c_2 = 0$$

Introduce a social security program:

- Forced savings tax τ equal to the amount saved by standard savers
- Retirement benefit: $b = \tau(1 + r)$.
- This tax does not affect the myopic saver (0% crowd-out of private savings), and fully crowds out the rational saver.



Evidence suggests that social security has about 30–40% crowd-out, implying that people are predominantly myopic.

Consumption, on average, drops substantially for all quartiles of wealth, except for those who also had top quartile income.

However, social security has likely contributed to a substantial drop in poverty for the elderly.

Moral Hazard

- Key efficiency cost of SS is the moral hazard of inducing workers to retire early.
- Benefits available at Early Eligibility Age (62), before Full Benefit Age (67).
- If a 62 year old worker works until 63, three things happen:
  - Cost 1: One extra year of payroll tax.
  - Cost 2: One year less of SS benefits.
  - Benefit: Higher adjusted SS benefit (8% extra per year of delay)
- The adjustment is actuarially fair if the benefit outweighs the two costs perfectly (as is basically true in the United States)
- Two big issues:
  - EEA might encourage early retirement.

  - Non-actuarially fair benefits may create a large implicit tax on work.

Retirement rate increases substantially around EEA, suggesting substantial moral hazard.

The actuarial adjustment among most countries tends to disincentivize work relative to the United States.

Current Problems and future outlook of Social Security

- Increase in life expectancy at retirement.
- Decrease in birth rates (fewer workers in the generational link)
- Increasing number of elderly per working age adult.
- Slower productivity growth since 1970.

Social Security requires adjusting taxes and/or benefits to remain in balance. The first set of reforms were implemented in 1983, and attempted to solve the SS budget shortfall via the following:

- Increased payroll taxes to create the trust fund.
- Increased retirement age for full benefit (from 65 to 67)

The trust fund peaked recently, and will be exhausted in 2034 — taxes will cover 75% of promised benefits.

Reform options:

(1) Increase payroll tax now to 15.84%, or increase payroll tax to 16.55% in 2035

(2) Increase base of wages (immigration, tax wages above 160K)

(3) Reduce benefits:
   - Index retirement age to life expectancy (likely regressive)
   - Index benefits to chain-price CPI rather than CPI
   - Make benefits taxable for income tax

(4) Means-testing

(5) Invest trust fund in higher-yield assets

(6) Privatization (very unpopular)

Unemployment Insurance

- Unemployment Insurance, like other social insurance programs, is triggered by an adverse event (namely, involuntary job loss).
- UI tends to spend about $20–30 billion/month in normal times.
- Plays a role in macroeconomic stabilization — Congress often extends unemployment insurance to stabilize the economy.
- Need to balance consumption smoothing benefits with moral hazard cost.

Different countries tend to have different structures for the generosity of unemployment insurance and length of term.

Optimal UI Theory

We want to define UI as that which maximizes utility.

$$\begin{aligned} EU &= (1-p)u(c_e) + pu(c_u) \\ &= (1-p)u(w-t) + pu(b) \end{aligned}$$

- p: probability of being unemployed
- c_e: consumption when employed
- c_u: consumption while unemployed
- w: wage while working
- t: taxes used to finance program
- b: unemployment insurance benefit

The actuarially fair taxes are as follows:

$$\begin{aligned} (1-p)(t) &= (p)(b) \\ t &= \frac{p}{1-p}b \end{aligned}$$

If there is no moral hazard, then p is not affected by b:

$$EU = (1-p(b))u\left(w - \frac{p}{1-p}b\right) + pu(b)$$

The optimal benefit b^* is that where $c_u = c_e$ (i.e., full insurance with no moral hazard).

However, if there is moral hazard, and p is a function of b, we would have to change the expected utility as follows:

$$EU = (1-p)u\left(w - \frac{p(b)}{1-p(b)}b\right) + pu(b)$$

The optimal benefit now is that which satisfies the following equation:

$$\begin{aligned} \frac{u'(c_u) - u'(c_e)}{u'(c_e)} &= \frac{\varepsilon_{p,b}}{p-1} \\ &= \frac{1}{1-p}\frac{b}{p}\cdot\frac{dp}{db} \end{aligned}$$

where $\varepsilon_{p,b}$ denotes the elasticity of the unemployment rate with respect to benefits. This equation sets the consumption smoothing benefit from UI equal to the moral hazard cost.

The implication of this equation is that partial insurance is optimal: $0 < c_u < c_e < w$.

It is difficult to estimate the optimal benefit using the previous formula: consumption smoothing benefits require knowledge of the utility function, wages, and other savings that are unobservable.

To do this, Raj Chetty uses a Taylor approximation: $u'(c_u) - u'(c_e) = u''(c_e)(c_u - c_e)$.

$$\begin{aligned}
\frac{\varepsilon_{p,b}}{1-p} &= \frac{u'(c_u) - u'(c_e)}{u'(c_e)} \\
&\approx \frac{u''(c_e)(c_u - c_e)}{u'(c_e)} \\
&= \underbrace{\left(-\frac{u''(c_e)}{u'(c_e)} c_e\right)}_{\text{coeff. of relative risk aversion}} \left(\frac{c_e - c_u}{c_e}\right) \\
&= \gamma \frac{\Delta c}{c_e}(b^*)
\end{aligned}$$

It is possible to measure γ, $\varepsilon_{p,b}$, and consumption, meaning we can empirically find b^*.

This equation extends to models with arbitrary savings, borrowing constarints, private insurance, and leisure benefits of unemployment.

Moral Hazard in Unemployment Insurance: Empirical Evidence

Since, in the United States, unemployment policy differs by state, we have quasi-experimental variation.

Meyer (1990) finds that $\varepsilon_{p,b} \approx 0.5$, average.

Consumption Smoothing in Unemployment: Empirical Evidence

Consumption falls by 10–15% when people lose their job.

A 10 percentage point increase in the UI replacement rate causes a 2.8% reduction in the rate of consumption drop.

- The replacement is much less than one to one, because UI tends to crowd out other self-insurance behaviors: spouses work less and individuals save less.

The current consensus is that the optimal UI replacement rate is around 50%, paid in a constant path.

Disability Insurance

- Disability is close to retirement: some people become unable to work before old age.
- All rich countries offer a form of public disability insurance — almost always linked to the public retirement system.
- Key issue that makes disability different from UI is *screening*.
- Moral hazard: incentive to exaggerate or make fraudulent claims.
- Screening challenge: back injuries and mental health conditions especially.
- Audits of Bureau of Disability Insurance assessments reveal substantial rates of error.
  - Oftentimes there are different acceptance thresholds for men and women.
  - Suggests that disability insurance applications should be gender blind or use machine learning algorithms.

- Disability Insurance has increased substantially over time. Two views:
  - Lenient system leads to inefficiency.
  - Program helping people with legitimate needs.

Evidence and Optimality of Disability Insurance

Quasi-experimental evidence: Maestas, Mullen, and Strand (2013) exploit random variation in DI examiners' stringency, find that labor force participation would have been 28 percentage points higher without DI.

Meyer and Mok (2019) find that food and housing consumption fall by 24% with disability and 2/3 never return to work.

Evidence suggests that current DI benefits may be too low.

Workers' Compensation

WC is insurance for temporary injuries on the job, covers medical costs and approximately 2/3 of lost wages.

Meyer, Viscusi, and Durbin (1995) find that a 10% increase in WC benefit raises out-of-work duration due to injury by 4%.

There are not a lot of empirical studies on the consumption smoothing benefits of WC.

Health Insurance: Overview

There are multiple problems with healthcare and health insurance in the United States:

- Cost: The United States spends almost twice the OECD average (and approximately the OECD average on *just* private healthcare).
- Despite this spending, the United States has a higher than average mortality, especially impacting native-born Americans without a high school education.
- Additionally, the United States has a relatively high uninsurance rate.
- The life expectancy problem in the United States is also made worse because the United States does not have above-average risk factors.

Primary policies of other countries: mandated universal coverage.

- **Beveridge Model:** single-payer insurance with government-run hospitals, private care, or mix. Primarily used in Anglophone countries and Sweden
- **Bismarck Model:** all-payer rate setting, non-profit insurance companies, mandated participation, and private care. Primarily used in Northern Europe, Israel, and Japan.

Governments control cost and limit overutilization by regulating for cost-effectiveness, setting prices, or rationing care, as well as cost-sharing with customer (coinsurance).

In the United States, the plurality of health insurance (49%) comes from employers, and another 34% of people get health insurance through public health insurance. Only 6% of people buy individual coverage.

- Most full-time employees are offered insurance through their employer, who contracts with an insurance company.

- For insurers, workers at firms are a large risk pool that avoids adverse selection.
- Employees participate because compensation in the form of health insurance is not taxed. More than $270 billion in taxes are forgone through this process.
- The "employer contribution" is actually fully incident on the employee (i.e., it comes out of lost wages).
  - Other drawbacks include job lock (can't get health insurance on the private market).
- Individuals can buy insurance directly from an insurer rather than through a group plan.
- Before Obamacare, this market was plagued by adverse selection (thanks to everyone who had group insurance subsidized via the tax code through their employer).
- Obamacare exchanges forbid price discrimination on the basis of pre-existing conditions and provide subsidies for buying individual insurance.
- Additionally, Obamacare expanded Medicaid to cover everyone under the poverty line and imposed a penalty on those who did not purchase insurance through the exchanges.
- However, after the Tax Cuts and Jobs Act (2017), the penalty was set to zero; many were worried about unraveling of the individual insurance market.

Health Insurance: Evaluation and Optimal Policy

The consumption-smoothing benefits of health insurance depend on the magnitude of the potential loss.

- Small shocks lead to small differences in marginal utility in the healthy and unhealthy states.
- Small predictable losses can be self-insured via saving.

The moral hazard from health insurance manifests in overconsumption because insured people pay only a fraction of healthcare costs when sick (also known as a co-pay or co-insurance). The question is the price elasticity of demand.

In the late 1970s, the RAND corporation randomly assigned 6000 people different types of health insurance with different co-pays plans to study moral hazard. They found three key things:

- Contrary to popular belief, demand for medical care is somewhat price sensitive. Individuals in the free plan used 1/3 more healthcare than those paying 95%. The researchers found an elasticity of -0.2.
- Lower co-pays did not lead to a significant improvement in a participant's health.
- For those who are chronically ill and could not easily cover co-pays, there was some deterioration in health.

For small health shocks, there are low benefits and some moral hazard, meaning we should avoid insuring them.

For large, unpredictable shocks, there are high benefits and less moral hazard.

These results suggest that the optimal health insurance has a high deductible (so patients bear a large share of cost initially) and full coverage for catastrophes.

However, ACA exchange data shows that people dislike plans with high deductibles.

- Brot-Goldberg et al. (2017) show that high deductible plans lead to large and likely inefficient cuts in healthcare use.

Evaluation of Medicare and Medicaid

Medicare Part D (which covers prescription drugs) was not optimal: there is 100% coinsurance between \$0 and \$275, 25% coinsurance between \$275 and \$2510, 100% coinsurance between \$2510 and \$5726, and then 7% coinsurance above \$5726.

A study on Medicare Part D found that many people clustered their total annual expenditure around where the 25% coinsurance segment ended, near \$2500. The ACA changed the \$0 – \$5726 bracket to be 75% coinsurance.

We can examine the direct effects of health insurance on health directly rather than looking at consumption-smoothing. However, merely comparing health outcomes for people who are on Medicaid vs. those who are not would yield biased results.

Currie and Gruber (1996) used Medicaid expansions as quasi-experiment to evaluate the value of the program.

- Utilization increased: early prenatal care visits rose by more than 50%
- Infant mortality declined by 8.5% due to the expansions in Medicaid for pregnant women.
- Policy was highly effective, only \$1.4m/life saved (value of a statistical life is around \$10m)

Finklestein et al. (2012) used a Medicaid lottery to estimate causal effects of coverage, specifically in Oregon. Lottery losers were the control, while lottery winners were the treatment group. Expanding access to Medicaid led to

- higher healthcare utilization
- lower out-of-pocket medical expenditures and medical debt
- slightly better self-reported physical health, and large improvements in mental health

Using regression discontinuity, Card, Dobkin, and Maestas (2008) found that Medicare availability led to increased utilization and lower post-hospitalization mortality.

Health Insurance Benefits

How do we reconcile the RAND result that found little impact on increased coverage on health? The studies examine different parts of the **health effectiveness curve**.

- Moving individuals from uninsured to some insurance leads to strong positive effects on health.
- Adding to generosity of insurance does not cause significant changes in health.

The United States leaves many uninsured, but provides overly generous care to the insured.

Provider-Side Evaluation

Consider a model where payment for physician services is $P = \alpha + \beta \cdot c$, where α is a fixed payment while β is payment proportional to physician costs c.

Physician contracts are primarily of two types:

- **Fee for Service**: $\alpha = 0, \beta > 1$, no fixed payment for practice, but insurance company pays full cost of visits + surcharge.
- **Diagnosis-based Payment**: $\alpha > 0, \beta = 0$, varying payment by type and number of patients, but not by services rendered.

If a physician is compensated for all costs, it is in their interest to do lots of procedures even if they aren't cost-effective.

The general trend has been toward higher α and lower β.

- The private market has shifted toward **Health Maintenance Organization (HMO) capitation** contracts, where the insurer pays a fixed amount per patient regardless of health costs.
- Kaiser physicians receive a flat payment per person enrolled based on age/gender.
- However, lower β provides incentives for doctors to provide fewer services, and maybe even too few services.
- In 1983, Medicare moved from fee for services to diagnosis-based payment. Cutler (1993) found that this shift led to:
  - Reduction on treatment intensity — doctors place some weight on profits.
  - No adverse impact on patient outcomes — doctors under fee for service were practicing non-cost-effective service.
  - Cost growth slowed dramatically in the first five years, but then accelerated as hospitals upcoded treatments.

Inflation Reduction Act: Effects on Healthcare

Medicare and Medicaid have been instructed to negotiate down prices of 10 high-cost drugs,

Fee imposed on companies that raise price of Rx drugs faster than inflation.

Out of pocket spending capped at $35/month.

Medicare and Medicaid recipients can get all vaccines for free.

Medicare Part D cost-sharing reforms:

- In 2024, copay above catastrophic threshold drops to 0%
- Starting in 2024, out of pocket costs capped at $2K/year.

Estimated effects of these policies is savings of $115 billion over next 10 years.

Taxation and Redistribution

Motivations for Redistribution

The primary motivation for government intervention in the economy is for correcting market failures, as we had analyzed in the previous section.

- Taxes and subsidies for negative/positive externalities

- Funding public goods

These interventions attempt to move the market equilibrium *toward* the Utility Possibilities Frontier.

Meanwhile, the objective of taxation and redistribution is to move *along* the Utility Possibilities Frontier. However, there is a potential cost — using taxes and transfers may move the market equilibrium outcome away from the Utility Possibilities Frontier.

Labor and Capital Income

Individuals derive pre-tax income from *labor* and *capital*:

$$z = wl + rk$$

where w is the wage, l is the labor supply (in hours), k is capital, and r is the rate of return on capital.

Labor income inequality is due to differences in:

- Work abilities reflected in w (education, experience, talent, physical ability, etc.)
- Work effort reflected in l (hours in work)
- Institutions (minimum wages, unions, etc.)
- Social norms (such as household division of labor) and discrimination (race and gender discrimination)

Capital income inequality is due to differences in:

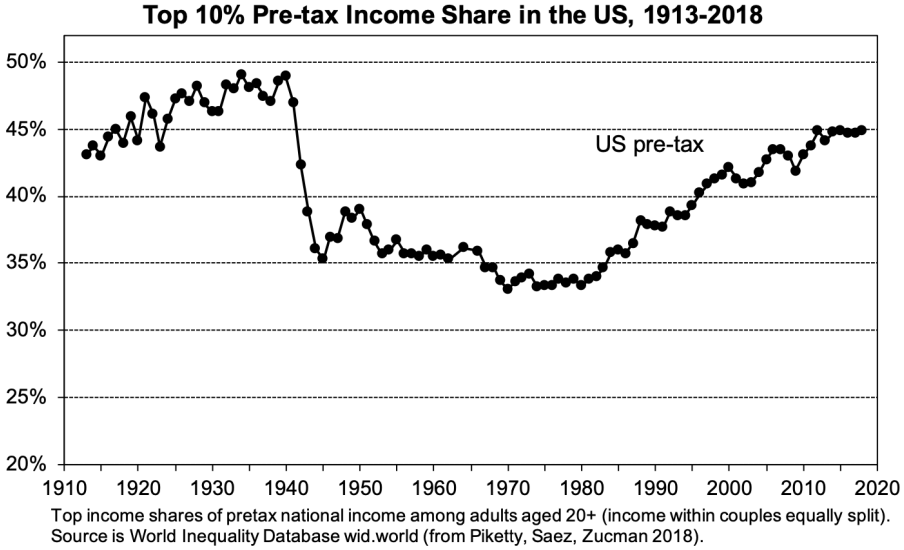
- total capital stock k (due to past saving and inheritance)
- rates of return r

In the aggregate, wl makes up about 75% of z, while rk makes up about 25% of z.

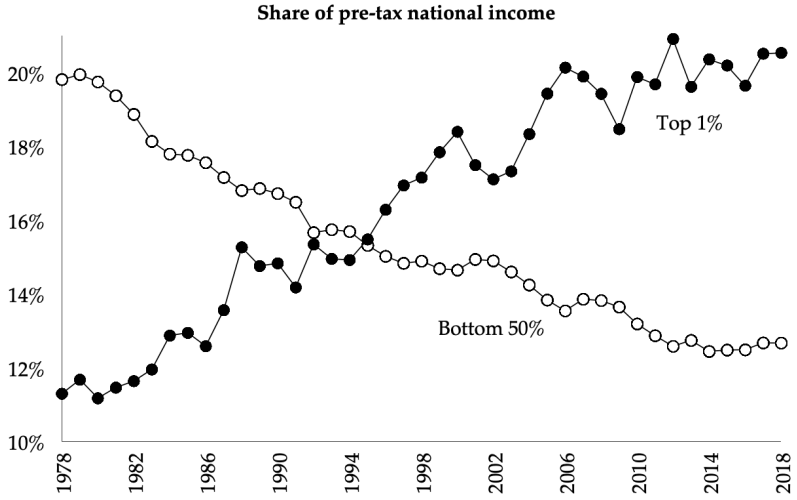
The capital stock is about 400–500% of z, with a rate of return r of around 6%.

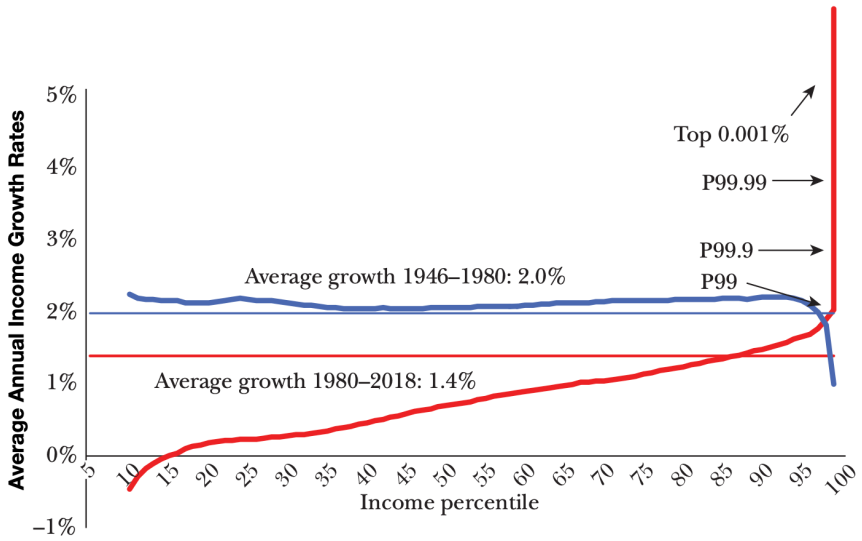
Empirical Facts about US Income Inequality

Fact 1: Top income share is U-shaped, but has leveled off in the past decade.

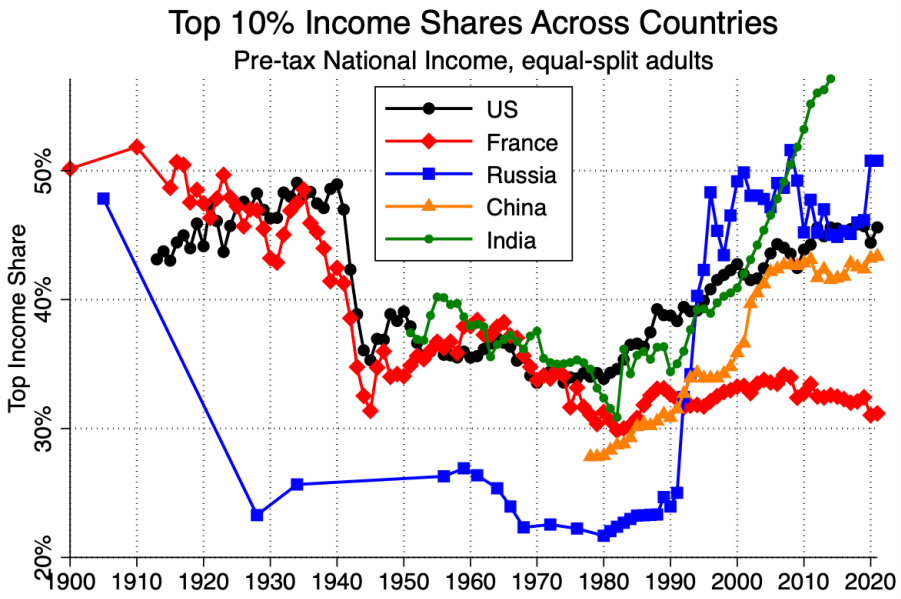


Fact 2: Income stagnation at the bottom of the income distribution.





Fact 3: Top 10% share of income is U-Shaped in Anglophone countries, but L-shaped in Continental Europe and Japan. There is a large increase in inequality in many developing countries



Measuring Poverty

We define the poverty rate as the fraction of a population with disposable income normalized by household size below the threshold z^*.

- Absolute Poverty: z^* is fixed *in real terms*. The World Bank uses a measure of \$2.15/person/day in 2017 dollars. There will not always be absolute poverty.
- Relative Poverty: z^* is fixed *relative to the median*. The European Union uses 60% of median. There will always be relative poverty.

We can also measure poverty based on consumption c, not pre-tax market income.

$$c = z - T(z) + B(z) + K(z) + E - s$$

where

- $T(z)$ is income taxes and credits
- $B(z)$ is cash transfers
- $K(z)$ is in-kind government transfers
- E is net private transfers
- s is net savings

However, consumption is difficult to measure.

How the US Measures Poverty

- Definition developed in 1963 by Molly Orshansky at SSA as 3 times the amount required to buy a "thrifty food plan"
- Based on annual *family* **money income**

$$z_m = z + B(z) + E$$

- Excludes taxes, tax credits, and in-kind transfers.
- Threshold adjusted annually using CPI.

Over time, the poverty rate has decreased from around 20% in 1960 down to around 12% in 2021.

Redistribution Methods

The government taxes income and consumption and provides transfers:

$$y = z - T(z) + B(z) + K(z)$$

If inequality in y is more than inequality in z, then the tax and transfer system is regressive — if inequality in y is less than in z, then the tax and transfer system is progressive.

- If $y = z(1 - \tau)$, then the tax and transfer system is neutral.
- If $y = z(1-\tau)+R$, where R is a universal transfer, then the tax and transfer system is progressive. Actual tax and transfer systems in rich countries work roughly like this.

Overall, the United States's tax and transfer scheme is progressive — while the system's raw progressivity (in terms of effective tax rates) has declined, the level of redistribution has increased despite declining effective tax rates for the rich.

Redistribution Motivations Activity

Activity: Redistribution Motivations and Foundations

Econ 308

Brandon Lehr

1 Gruber 17.1: It's Absolutely Relative

As Table 17-2 shows, members of the poorest fifth of U.S. households have a much smaller share of total U.S. income than is typical in other developed countries. Does this mean that the poorest fifth of U.S. households are worse off in the United States than are the poorest fifth of households elsewhere? Why or why not?

Table 17.2, Income Share in Total Income for OECD Nations, Page 490

	Income Share in Total Income					
Country (2018)	**Bottom 10%**	**Bottom 20%**	**Bottom 40%**	**Top 40%**	**Top 20%**	**Top 10%**
Austria	3.1	8.5	22.7	59.4	36.6	22.3
Belgium	3.8	9.2	23.3	58.1	34.8	20.7
Canada	2.9	7.8	21.0	61.1	37.8	22.9
Czech Republic	4.1	9.9	24.3	57.5	34.7	20.5
Denmark	3.8	9.5	23.6	58.3	35.7	21.9
finland	3.9	9.4	23.4	58.8	36.3	22.3
france	3.4	8.6	21.9	60.8	39.0	25.0
Germany	3.3	8.5	22.1	60.1	37.5	23.2
Greece	2.9	7.7	20.9	61.5	38.2	23.4
Hungary	3.2	8.5	22.2	60.2	37.5	23.1
Italy	2.0	6.6	19.4	63.0	39.7	24.5
Korea	2.2	6.2	18.4	64.4	40.7	25.0
Luxembourg	2.7	7.6	20.5	62.2	39.5	24.6
Mexico	2.0	5.6	15.9	69.4	47.9	32.3
New Zealand	2.9	7.3	19.2	64.5	42.2	27.5
Norway	3.3	8.9	23.6	57.9	35.2	21.4
Poland	3.2	8.5	22.3	59.6	36.6	22.2
Portugal	3.0	7.8	20.7	62.2	39.8	25.1
Slovak Republic	3.5	9.4	24.5	56.5	32.8	18.5
Sweden	3.5	8.7	22.6	59.4	36.6	22.7
Turkey	2.4	6.2	17.1	67.7	46.3	31.3
United Kingdom	2.4	6.7	18.5	65.4	43.6	29.0
OECD	2.9	7.7	20.6	62.3	39.6	24.8
United States	1.6	5.3	16.2	67.6	44.5	28.5

Total income in the US is much higher than many other countries, so bottom 20% is likely richer than other countries's bottom 20%

2 Measuring Poverty

The US poverty threshold increases with household size:

2023 POVERTY GUIDELINES FOR THE 48 CONTIGUOUS STATES AND THE DISTRICT OF COLUMBIA	
Persons in family/household	**Poverty guideline**
1	$14,580
2	$19,720
3	$24,860
4	$30,000
5	$35,140
6	$40,280
7	$45,420
8	$50,560

This is because a larger household needs more income to achieve a given level of utility, but the question is how much more income? **Equivalence scales** are the economist's way of answering this question and provide the means of adjusting measured incomes into comparable quantities.

a. Observe that the US poverty threshold for a family of size $N \geq 2$ is not simply N times the poverty threshold for a family of size 1. This is because there are economies of scale in family size. Provide some examples of such economies of scale.

N person family does not need n× square footage for one family, family plans (bulk → cheaper

b. There are two senior advisors to the government, A an B, who both agree that the poverty line is at $12,000 for a single person. However, they have different equivalence scales. A believes that the scale factor in determining income should be 0.25 for each additional family member (i.e., an additional $3,000/person). B suggests that the scale factor should be 0.5 for each additional family member.

   i. Find the poverty threshold for a family of two, three, and four under both values of the scale factor 0.25 and 0.5.

Scale Factor	Single	Two	Three	Four
0.25	12,000	15000	18000	21 000
0.5	12,000	18000	24000	30 000

ii. Which advisor believes that the economies of scale in family size are larger?

Advisor A

iii. Suppose the government is committed to provide welfare eligibility to every family below the poverty threshold. If this government wishes to keep total welfare spending to a minimum, which of the two advisors should it listen to?

Advisor A

Bonus. Assume a family utility function $U = \ln\left(\frac{M^2}{d}\right)$ where M is family income and d is family size. And suppose that the government wants to set poverty thresholds so that at each threshold by family size, family utility is the same. Determine the implied equivalence scale: the ratio of the poverty threshold for a family of size d to the poverty threshold for a family of size 1. How much larger is the poverty threshold for a family size of 2, or 4?

$\ln(M^2) - \ln(d)$

$U = k$ $\quad m^2 = 2$

$k = \ln\left(\frac{M^2}{1}\right)$ $\quad \frac{1}{\sqrt{2}}$

Tax Incidence

Tax incidence is the effect of tax policies on prices and the economic welfare of individuals. Incidence is analyzed at different levels:

- Producer vs. Consumer: effect of the tax on income/welfare of consumers vs. shareholders vs. producers of intermediate goods.

- Source of income: taxes on labor income vs. capital income vs. land income.
- Income level: progressivity of the labor income tax.
- Spatial/Regional: effect of a property tax increase on residents of one area vs. another area.
- Intergenerational: effect of the social security payroll tax on income/welfare of young vs. old.

Tax incidence is a **positive analysis**, we only analyze the impact of taxes.

Key Results of Tax Incidence

Key Result 1: The statutory incidence of a tax is *not* equal to the economic incidence of a tax. Taxes can shift relative prices.

- For example, liberals tend to favor increased capital income taxation because capital income is concentrated at the high end of the income distribution.
- However, if people save less because of the increased capital taxes, then capital stock goes down, and marginal product of *labor* (i.e., wages) can go down due to reduced capital stock.

Key Result 2: Equilibrium is independent of who nominally pays the tax.

We will focus on **partial equilibrium** in a **perfectly competitive** market.

- Assume that the government levies an excise tax t on a good — excise taxes are levied on a quantity (e.g., a \$1/pack tax on cigarettes). This is contrasted with *ad valorem* taxes, which are on a particular fraction of prices (e.g., a 9% sales tax).
- The supply function $S(p)$ depends on the producer price p, and the demand function $D(q)$ depends on the consumer price q.
- In equilibrium, $S(p) = D(q)$, and $t = q - p$.
- Therefore, with a producer tax, we have $S(q - t) = D(q)$, and for a consumer tax, we have $S(p) = D(p + t)$.

Key Result 3: The more inelastic factor bears more of the tax.

$$\begin{aligned}\frac{dp}{dt} &= \frac{\varepsilon_D}{\varepsilon_S - \varepsilon_D} \in [-1, 0] \\ \frac{dq}{dt} &= \frac{dp + dt}{dt} \\ &= \frac{\varepsilon_S}{\varepsilon_S - \varepsilon_D} \in [0, 1]\end{aligned}$$

If ε_D is close to 0, then consumers bear more of the tax. Meanwhile, if ε_S is close to 0, then producers bear more of the tax.

In the limiting cases, if $\varepsilon_D = -\infty$, then consumers bear 0% of the tax, while if $\varepsilon_D = 0$, then consumers bear 100% of the tax.

Tax Incidence in General Equilibrium

So far, we have focused on taxation in partial equilibrium (where we focus on the effect of the tax on one market in isolation).

General equilibrium models consider the effects on related markets of a tax imposed on one market. For example, imposing a tax on cars may reduce demand for steel, which affects prices for other products beyond that of the car market.

General Equilibrium Analysis: Berkeley Soda Tax

Consider the market for soda in Berkeley, California.

- Berkeley imposed a soda tax in 2015 of $0.01 per ounce.
- The goal of the tax was to reduce soda consumption, and thereby improve health outcomes.
- If soda demand in Berkeley is inelastic, consumers bear more of the burden, but demand for soda *in Berkeley* is likely to be elastic — consumers consume less soda or buy soda in Oakland.

Consider the extreme case: perfectly elastic demand for soda.

- As a result, Berkeley soda sellers bear the full burden of the tax.
- However, soda sellers are made up of land, capital (buildings, kitchen equipment, etc.) and labor (cashiers, cooks, waitstaff, etc.)
- These two factors must bear the loss as a result of the tax — incidence is "shifted backward" to land, capital, and labor.

Short Run Incidence:

- Assume that labor is perfectly elastic (workers can go to Oakland if they get paid less in Berkeley).
- However, given that restaurants/shops have fixed leases, we can also assume that capital is perfectly inelastic. Additionally, land in Berkeley is perfectly inelastic, seeing as you can't move land or make more land in Berkeley.
- Therefore, capital and land bear the tax in the short run.

Long Run Incidence:

- However, in the long run, capital is highly elastic — restauranteurs can close, sell, and take their money to invest elsewhere, or sign a different lease.
- Therefore, in the long run, we assume that (physical) capital is perfectly elastic.
- So, we can assume that landowners are those that bear the incidence of the soda tax in the long run, assuming full elasticity of soda demand, labor, and capital are fully elastic.

This is an idealized example — in practice, soda demand, labor, and capital are not fully elastic, so incidence is shared in the long run.

Federal vs. Total Tax Rates

The federal tax system is very progressive — higher income and wealth individuals are taxed more than lower income and wealth individuals. However, state and local taxes are less progressive.

Saez and Zucman (2019) found that most people across the income distribution pay around 25–30%. Poorer individuals pay more in consumption and payroll taxes, while middle–high income individuals pay more in individual income taxes, and extremely wealthy individuals pay more of their income in corporate and property taxes.

Tax Incidence Activity

Activity: Equity Implications of Taxation

Econ 308

Brandon Lehr

1 Gruber 19.4: Computing Tax Incidence

The demand for football tickets is $D(P) = 360 - 10P$ and the supply of football tickets is $S(P) = 20P$. Calculate the gross price paid by consumers after a per-ticket tax of \$4. Calculate the after-tax price received by ticket sellers.

$360 - 10(P+4) = 20P$

$320 - 10P = 20P$

$P_S = \frac{32}{3}$

$P_D = \frac{44}{3}$

2 Gruber 19.7: Tax Incidence Factors

You have determined that producers, rather than consumers, will bear the lion's share of the burden associated with a new tax. How does the elasticity of labor supply influence whether this tax burden will, in turn, be borne more by workers or more by property owners?

Labor supply inelastic → more incidence on workers

Labor supply elastic → more incidence on landowners

3 Incidence of US Federal Taxes

The Congressional Budget Office (CBO) analysis considers the incidence of the full set of taxes levied by the *federal* government. The CBO assumes that:

- **Individual Income taxes** are borne fully by the households that pay them.
- **Payroll taxes** (i.e., Social Security and Medicare taxes) are borne fully by workers, regardless of whether these taxes are paid by the workers or by the firm.
- **Excise taxes** (e.g., gas, alcohol, tobacco taxes) are fully shifted to consumer prices and so are borne by individuals in proportion to their consumption of the taxed item.
- **Corporate taxes** are borne by both workers and owners of capital (but more so for the latter). *Note: This is subject to debate.*

a. What do these assumptions imply for what the CBO must think about the relative elasticities for supply and demand in the labor market? In the markets for goods subject to excise taxes? And the relative elasticities for the supply of labor and capital to the corporate sector?

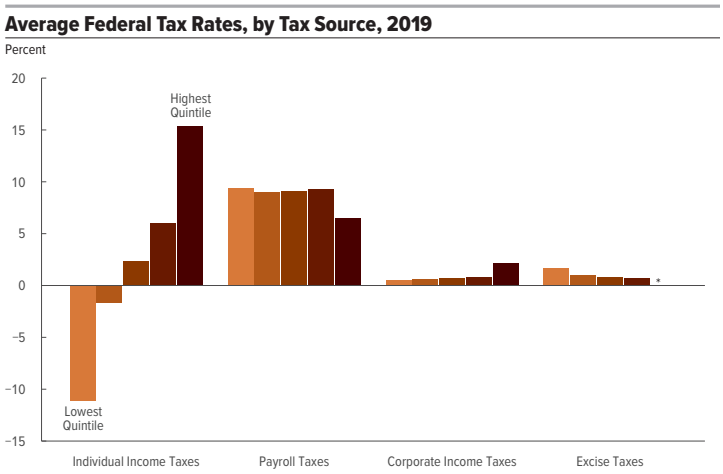
FIT: fully relastic labor supply

Payroll: fully relastic labor supply

Excise: fully relastic consumer demand

Corporate: relatively relastic capital, relatively elastic labor.

b. The results of the CBO analysis are pictured below. Which federal taxes are progressive, regressive, and neutral?



IIT, Corporate: progressive

Payroll: ~ neutral

Excise: regressive

Efficiency Costs of Taxation

Incidence is concerned with how taxes affect equilibrium prices and the *distribution* of the proverbial economic pie.

A second set of general questions is to understand how taxes affect the *size* of the proverbial economic

pie. Governments impose taxes to:

- Raise revenue to fund public goods and social insurance
- Redistribute income

However, raising tax revenue has an efficiency cost — to generate \$1 of revenue, welfare of those taxes falls by more than \$1 due to behavior distortion.

Deadweight Loss (DWL), or excess burden, is defined as the welfare loss created by a tax over and above the tax revenue generated by the tax.

Insights from Deadweight Loss

Assuming no income effects and competitive production, we calculate the DWL of a tax by finding the area of the *Harberger Triangle*:

$$\begin{aligned} \text{DWL} &= \frac{1}{2}(-dQ)(dt) && \text{recall that } dQ \text{ is negative} \\ &= \frac{(\varepsilon_S)(-\varepsilon_D)}{2(\varepsilon_S - \varepsilon_D)} \cdot \frac{Q}{p} \cdot (dt)^2 \end{aligned}$$

Insight 1: DWL increases with the absolute size of elasticities (i.e., it is more efficient to tax relatively inelastic goods).

Insight 2: DWL increases with the *square* of the tax rate (i.e., more efficient to have lower rates and broader bases than higher rates with smaller bases).

Optimal Commodity Taxation

Ramsey (1927) was asked Pigou to solve the following problem:

- Consider one consumer who consumes K different goods.
- What are the tax rates $t_1, t_2, \ldots, t_K$ on each good that raise a given amount of revenue while minimizing the welfare loss to the individual?

Obviously, uniform rates are not optimal if there are more elastic demands for certain goods.

Ramsey Rule: the optimal tax rates are such that the marginal DWL for the last dollar of tax collected is the same across all goods:

$$\frac{M\text{DWL}_k}{MR_k} = c, \qquad \text{for all } k = 1, 2, \ldots, K$$

where c is the marginal value of government revenue.

Implication: Tax more the goods that have inelastic demand (and tax less the goods that have elastic demand)

Limitation: Ramsey's result abstracts from redistribution and focuses solely on efficiency

Note on Income Effects

If we don't assume away income effects, the Harberger triangle *overstates* DWL, since income effects do not create an efficiency cost.

If there is a $100 per person lump sum tax, consumers buy less and achieve lower welfare, but the DWL of the tax is zero because the amount the government would need to spend to arrive back at original welfare is $100 — the exact amount of revenue gained from the lump sum tax.

Tax Salience

Standard economic models assume that individuals are fully aware of what they pay, but this is not necessarily so.

Chetty, Looney, and Kroft (2009) test this assumption and develop a theory of taxation with inattentive consumers.

Chetty et al. (2009) Empirical Strategy

Randomized Field Experiment

- In a treatment store, they display new price tags showing the levels of sales tax and total price on a subset of products.
- Researchers compare shopping behavior for treated products vs. control products in treated store, before and after new tags are implemented (DD estimator).
- They then repeat analysis in control stores as a *placebo* DD estimator.

Quasi-Experiment

- Use variation in beer excise and sales taxes across states.
- Excise tax is salient because they're built into the posted price.
- Sales tax is not salient because it is not included in the posted price.

RFE Results: Posting sales taxes reduces demand for particular goods by 7.6%.

Quasi-Experiment Results: Beer consumption is elastic to excise tax rates but not to sales tax rates.

Key result: tax salience matters — if tax is not salient, then demand is less elastic, and consumers bear more of the tax burden.

A number of empirical studies show that individuals are not fully informed and/or attentive; important consequences for policy.

Commodity Taxation Activity

Activity: Tax Inefficiency

Econ 308

Brandon Lehr

1 Gruber 20.7: The ABCs of Commodity Taxes

You are a consultant to the government of Buttony. The government has decided to cut taxes on one of the following: apples, bananas, or cantaloupe. The government wants your input on which fruit would be the best choice for a tax cut. It provides you with the following information. What is your recommendation, and why?

Good	Unit Price	Sales (thousands)	Unit tax	Marginal tax revenue	Marginal DWL
Apples	\$1	100	\$0.10	20	5
Bananas	\$2	100	\$0.25	30	20
Cantaloupe	\$4	50	\$0.15	10	20

Note: marginal tax revenue and marginal DWL are measured in thousands of dollars per \$1 of additional tax.

Cantaloupe – its MR is lower than the MDWL.

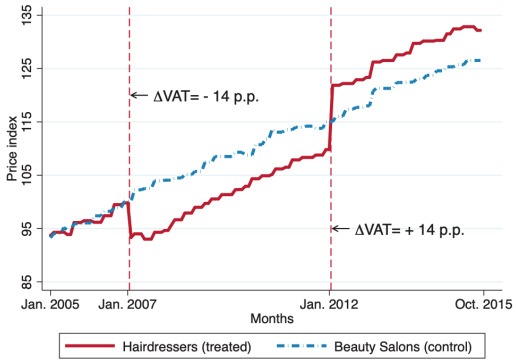
2 Gruber 20.13: Efficient Commodity Taxation

Schmeezle and Schmoozle are two advisors for the government of Feldspar. Schmeezle says that since the elasticity of demand for granite countertops is -3 and the elasticity of demand for sinks is -1.5, taxes should be raised entirely from granite countertops. Schmoozle argues that it is better to levy taxes on both goods anyway. Which advisor should the Feldspar government listen to? Why?

Feldspar govt. should raise tax entirely on sinks/as elasticity lower → lower DWL, so it should listen to Schmoozle.

3 What Goes Up May Not Come Down

European countries have large taxes on consumption known as a Value Added Tax (VAT). Normal VAT rates are high (15%-25%) but some goods/services have lower rates (or are exempt). Benzarti et al. (2017) study the effects of VAT rate increases and decreases. As part of their analysis, they present the following illustrative case study: hairdressers in Finland got a VAT cut of 14 percentage points (p.p.) in Jan 2007 that was repealed in Jan 2012. Their graphical analysis is below.



a. What is the empirical strategy that the authors are using?
DD estimator

b. Does this empirical strategy appear to be valid for making causal inferences about the effect of the VAT changes for Finnish hairdressers?
Yes, parallel trends assumption appears to hold — similar services but comparable price trends

c. How much of the 14 p.p tax cut is passed on to consumers? How much of the 14 p.p. tax increase is passed on to consumers?
5 pp of tax cut, but ~ 10 pp to tax increase

d. Why is the answer to part (c) surprising for standard economic analysis?
We would expect elasticity to be equal → equal share borne in either direction

e. Suggest a possible explanation for the empirical finding.
Consumers are more sensitive to high prices for some reason (?)

Taxation across Countries

As a share of GDP, OECD economies on average:

- Collect 34% of GDP in tax revenue
- Spend 20% of GDP on social transfers

Should governments mitigate inequality using transfers and taxes? If so, how?

Transfers in the United States

- Universal Transfers: public education, healthcare, social insurance (retirement, disability, unemployment)
- Means-tested transfers:
  - In-Kind transfers: Medicaid, public housing, SNAP
  - Cash transfers: TANF, Supplemental Security Income

  Means-tested transfers tend to have high take-up costs, and (relatively) low utilization.
- Refundable Tax Credits: EITC; managed by the IRS, paid as lump-sum, low take-up costs and high utilization.

Types of Taxes

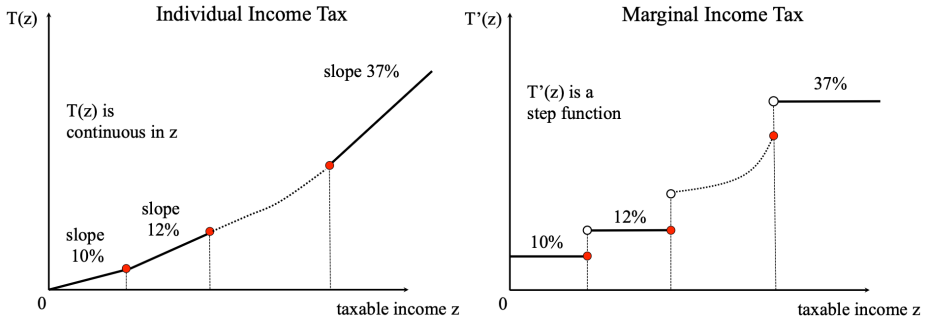
- Consumption Taxes: sales taxes and excise taxes (value-added taxes in other countries)
- Payroll Taxes: taxes on labor earnings, funds Social Security and Medicare
- Individual Income Tax: taxes on a broader income base; includes labor income (wages), capital income (interest, dividends, and business income), and land income (rent, not including imputation from owner-occupancy)
- Corporate income tax: taxes on corporate profits net of reinvestment.
- Wealth taxes: taxes primarily on property (more common locally) and estates (e.g. inheritance)

Definition of Income

- Income tax $T(z)$ is a function of annual household taxable income, z.
- z is defined as **adjusted gross income** (AGI), net of standard or itemized deductions.
  - AGI = Gross Income net of "above-the-line" deductions (healthcare premiums, pensions, etc.)
  - AGI base is approximately 70% of national income
  - Standard Deduction is \$13850 for singles and \$27700 for couples in 2023
  - Itemized deductions are taken in place of the standard deduction: state/local taxes (up to \$10K), interest on mortgages, or charitable giving are the most common itemized deduction.

Marginal Tax Rates and Brackets

$T(z)$ is a continuous piecewise linear function with slope equal to MTR by taxable income bracket. $T'(z)$, which reflects the marginal tax rate, is a step function.



Tax Credits

Tax credits are additional reductions in taxes above and beyond a deduction.

- Non-refundable credits (cannot reduce net tax burden below zero): foreign tax credits (offset taxes paid abroad), childcare expenses, residential energy efficient credits, etc.
- Refundable credits (can reduce net taxes below zero/act as a transfer): Child tax credit ($2K per child, partially refundable), EITC (up to $3.6K, $5K, $6.7K, based on family size)

Tax and Transfer Credits

- Draw budget constraint $c = z - T(z)$ that integrates taxes and transfers
- Demogrant: $T(0)$, net transfer with zero earnings
- Marginal tax rate: $T'(z)$: individual keeps $1 - T'(z)$ for every additional $1 of earnings (intensive margin of labor supply)
- Break-even earnings point z^*: point where $T(z^*) = 0$
- Participation tax rate: $\tau_p = [T(z) - T(0)]/z$: individual keeps $1 - \tau_p$ of earnings when moving from zero earnings to z (extensive margin of labor supply) — average slope of post-tax income

Optimal Income Taxation: No Behavioral Response

What is the optimal labor income tax function $T(z)$?

Suppose everyone has the same concave utility function $u(c)$ over after-tax income $c = z - T(z)$, and there are two individuals with fixed incomes $z_1 < z_2$.

We want to find $T(z)$ that maximizes Utilitarian SWF:

$$f(z_1, z_2) = u(z_1 - T(z_1)) + u(z_2 - T(z_2))$$
$$0 = T(z_1) + T(z_2)$$

Both of which imply

$$z_1 - T(z_1) = z_2 - T(z_2)$$

This implies 100% redistribution — perfect equalization of after-tax income — mathematically equivalent to full insurance with risk aversion and no moral hazard. However, this is unrealistic:

- 100% redistribution would remove all incentive to work, so the assumption that z is fixed is unrealistic.

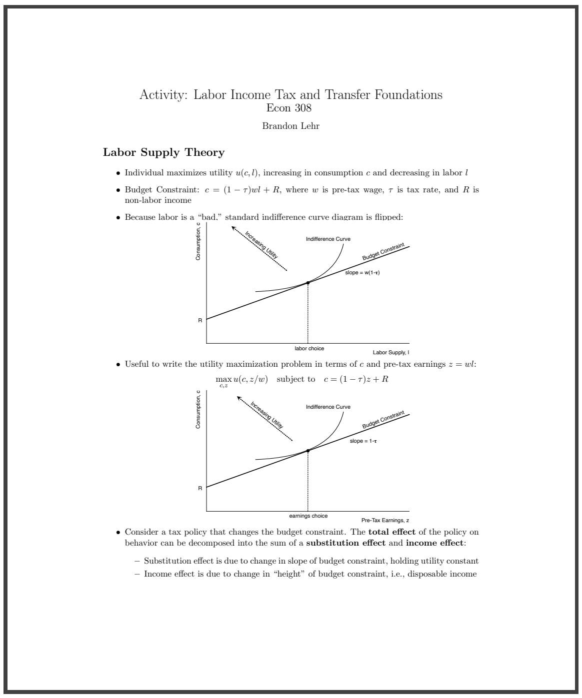
- With utilitarianism, behavioral responses are the only factor preventing complete redistribution. However, not everyone is a utilitarian, and citizens' views on fairness bound the ability of the government to redistribute.

Maybe we can use the Second Welfare Theorem?

- Second Welfare Theorem: Any Pareto efficient outcome can be reached via a suitable set of lump sum taxes (based on intrinsic earning ability), then letting markets work freely
- There is no conflict between efficiency and equity, right?

However, governments must base taxes and transfers based on actual earnings, rather than intrinsic earnings ability — there is a real conflict between efficiency and equity.

Tax Effect Activity



Activity: Labor Income Tax and Transfer Foundations
Econ 308

Brandon Lehr

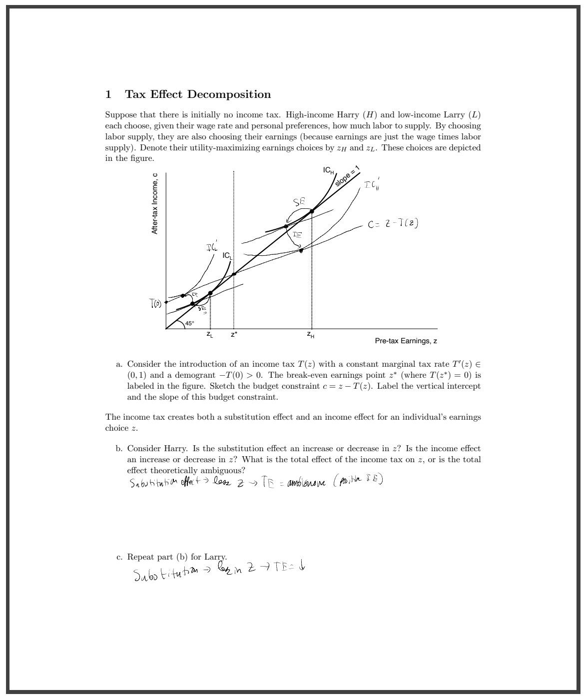
Labor Supply Theory

- Individual maximizes utility $u(c, l)$, increasing in consumption c and decreasing in labor l
- Budget Constraint: $c = (1-\tau)wl + R$, where w is pre-tax wage, τ is tax rate, and R is non-labor income
- Because labor is a "bad," standard indifference curve diagram is flipped:

- Useful to write the utility maximization problem in terms of c and pre-tax earnings $z = wl$:

$$\max_{c,z} u(c, z/w) \quad \text{subject to} \quad c = (1-\tau)z + R$$

- Consider a tax policy that changes the budget constraint. The **total effect** of the policy on behavior can be decomposed into the sum of a **substitution effect** and **income effect**:
  - Substitution effect is due to change in slope of budget constraint, holding utility constant
  - Income effect is due to change in "height" of budget constraint, i.e., disposable income



1 Tax Effect Decomposition

Suppose that there is initially no income tax. High-income Harry (H) and low-income Larry (L) each choose, given their wage rate and personal preferences, how much labor to supply. By choosing labor supply, they are also choosing their earnings (because earnings are just the wage times labor supply). Denote their utility-maximizing earnings choices by z_H and z_L. These choices are depicted in the figure.

a. Consider the introduction of an income tax $T(z)$ with a constant marginal tax rate $T'(z) \in (0,1)$ and a demogrant $-T(0) > 0$. The break-even earnings point z^* (where $T(z^*) = 0$) is labeled in the figure. Sketch the budget constraint $c = z - T(z)$. Label the vertical intercept and the slope of this budget constraint.

The income tax creates both a substitution effect and an income effect for an individual's earnings choice z.

b. Consider Harry. Is the substitution effect an increase or decrease in z? Is the income effect an increase or decrease in z? What is the total effect of the income tax on z, or is the total effect theoretically ambiguous?

c. Repeat part (b) for Larry.

Optimal Labor Income Taxation

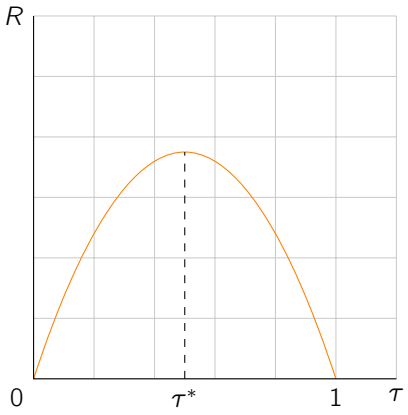
Assume individuals pay a linear tax rate τ on earnings z and receive a fixed universal transfer R. There are N individuals, with each person i choosing their utility-maximizing labor supply l_i by maximizing

$$u^i(c_i, l_i) = u^i\left((1-\tau)w_i l_i + R, l_i\right)$$

The government budget is

$$R = \tau \cdot Z$$
$$Z = \sum_i z_i / N,$$

meaning average tax revenue is $R(\tau) = \tau \cdot Z(1-\tau)$. Therefore, at $\tau = 0$, no tax revenue (meaning no one works), and at $\tau = 1$, no one works due to no incentive to do so.



Known as the Laffer curve, although the idea dates back to Dupuit (1844), revenue is maximized in this simplistic model at $\tau = \tau^*$. It is inefficient to have $\tau > \tau^*$. Under the Rawlsian framework, we want to maximize R, meaning that $\tau = \tau^*$.

To determine τ^*, we consider the effect of an increase in the tax rate by $d\tau > 0$.

- Mechanical Effect: increases average tax revenue (by increasing tax rates) — $dM = Z \cdot d\tau > 0$
- Behavioral Effect: reduces average tax revenue (by reducing labor supply) — $dB = \tau \cdot dZ < 0$

At the optimum, $dM + dB = 0$ (i.e., maximizing tax rates).

$$\tau^* = \frac{1}{1+e}$$

where

$$e = \frac{1-\tau}{Z}\frac{dZ}{d(1-\tau)}$$

is the elasticity of average income with respect to the net-of-tax rate. There are different estimates of e, but $e = 0.25$ implies that $\tau^* = 80\%$.

Under a utilitarian framework, we want to maximize

$$\begin{aligned} SWF &= \sum_i u^i(c_i, l_i) \\ &= \sum_i u^i\left((1-\tau)w_i l_i + \tau \cdot Z(1-\tau), l_i\right) \end{aligned}$$

taking into account that labor supply responds to tax policy. In this scenario, the optimal linear income tax rate is

$$\tau = \frac{1-\overline{g}}{1-\overline{g}+e}$$

where $\overline{g} \in [0,1]$ measures the degree of pre-tax earnings equality.

This formula captures the equity-efficiency tradeoff:

- τ is decreasing in behavioral elasticity e
- τ is decreasing in pre-tax earnings equality $\overline{g}$

The formula is very general and applies if:

- People respond to taxation by dropping out of the labor force instead of adjusting hours.
- People choose education based on tax rate.
- Earnings are generated by a combination of ability and luck.

Optimal *Top* Income Tax Rate

Consider a constant marginal tax rate τ above fixed z^*. Denote z_m to be the *mean* income of top bracket earners. Let elasticity of z_m with respect to net-of-tax rate be

$$e = \frac{1-\tau}{z_m} \cdot \frac{dz_m}{d(1-\tau)}$$

Assume a small $d\tau > 0$ tax increase. Then, we have

- Mechanical Effect: $dM = d\tau \cdot (z_m - z^*)$
- Behavioral Effect: $dB = \tau dz_m$

We find that the optimal top tax rate is

$$\tau = \frac{1}{1+ae}$$
$$a = \frac{z_m}{z_m - z^*}$$

where a denotes the Pareto coefficient — i.e., the average income of people who make above the income threshold. In the United States, $a \approx 1.5$ — i.e., $z_m = 3z^*$.

Estimating e is very hard:

- If $e = 0.25$, then $\tau_m^* = 73\%$
- If $e = 0.5$, then $\tau_m^* = 57\%$