

Undergraduate Public Economics

Danny Yagan
UC Berkeley

Lecture 1 Introduction

PUBLIC ECONOMICS DEFINITION

Public Economics (or public finance) = Study of the Role of the Government in the Economy

Government is instrumental in most aspects of economic life:

- 1) Government in charge of huge **regulatory** structure
- 2) **Taxes:** governments in advanced economies collect 35-50% of National Income in taxes
- 3) **Expenditures:** tax revenue funds **traditional public goods** (infrastructure, public order and safety, defense) and **welfare state** (Education, Retirement benefits, Health care, Income Support)
- 4) Macro-economic **stabilization** through central bank (interest rate, inflation control), fiscal stimulus, bailout policies

POLICYMAKERS SEEK ECONOMISTS' ADVICE



Four questions of public finance

- 1) When should the government intervene in the economy?
- 2) How might the government intervene?
- 3) What is the effect of those interventions on economic outcomes?
- 4) Why do governments choose to intervene in the way that they do?

When should the government intervene in the economy?

1) Market Failures: Market economy sometimes fails to deliver an outcome that is efficient \Rightarrow Government intervention may improve the situation

2) Redistribution: Market economy generates substantial inequality in economic resources across individuals \Rightarrow Government intervention may help reduce inequality by redistributing resources through taxes and transfers

First part of the class focuses on Redistribution

Second part of the class focuses on Market Failures

Main Market Failures

1) Externalities: (example: greenhouse carbon emissions) \Rightarrow require govt interventions (Pigouvian taxes/subsidies, public good provision)

2) Imperfect competition: (example: monopoly) \Rightarrow requires regulation (typically studied in Industrial Organization)

3) Imperfect or Asymmetric Information: (example: adverse selection in health insurance may require mandatory insurance)

4) Individual failures: People are not always rational. This is analyzed in behavioral economics, field in huge expansion (example: myopic people may not save enough for retirement)

Inequality and Redistribution

Even if market outcome is efficient, society might not be happy with the market outcome because market equilibrium might generate very high economic disparity across individuals

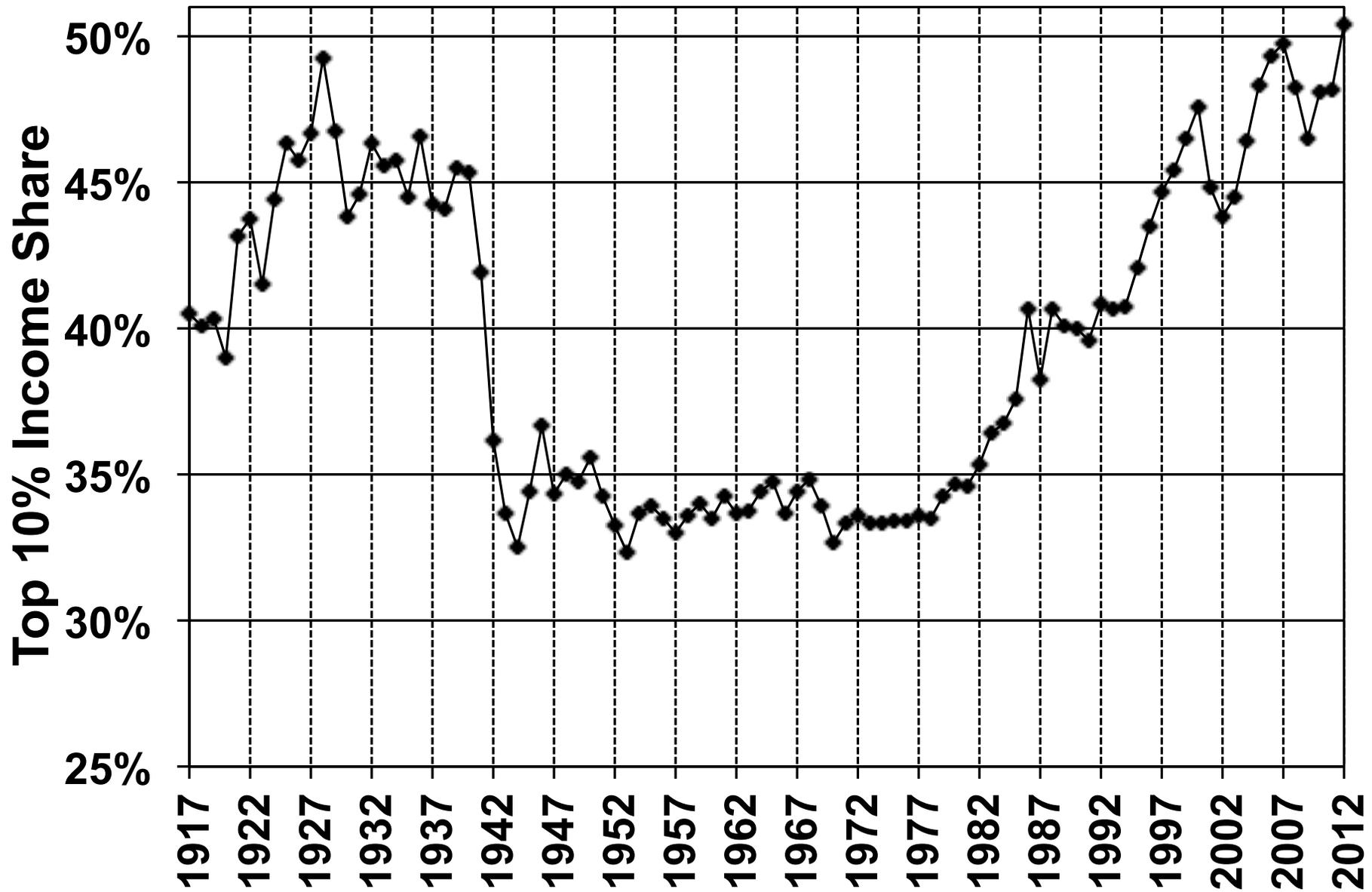
Governments use taxes and transfers to redistribute from rich to poor and reduce inequality

Redistribution through taxes and transfers might reduce incentives to work (**efficiency costs**)

⇒ Redistribution creates an **equity-efficiency trade-off**

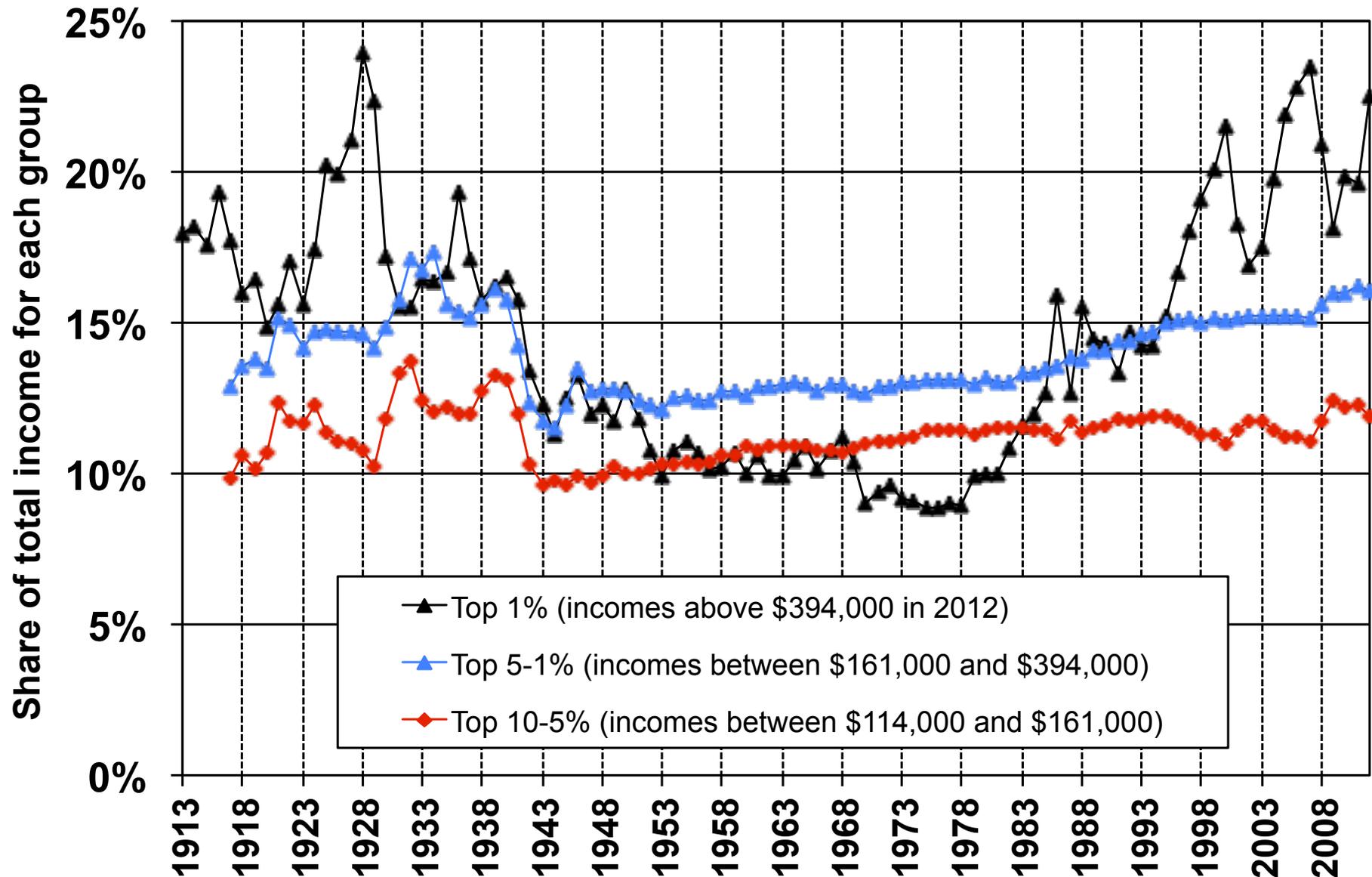
Income inequality has soared in the United States in recent decades, and has moved to the forefront in the public debate (see Piketty's 2014 book success)

Top 10% Pre-tax Income Share in the US, 1917-2012



Source: Piketty and Saez, 2003 updated to 2012. Series based on pre-tax cash market income including realized capital gains and excluding government transfers. 2012 data based on preliminary statistics

Decomposing Top 10% into 3 Groups, 1913-2012



Source: Piketty and Saez, 2003 updated to 2012. Series based on pre-tax cash market income including realized capital gains and excluding government transfers. 2012 data based on preliminary statistics.

How Might the Government Intervene?

- 1) Tax or Subsidize Private Sale or Purchase:** Tax goods that are overproduced (e.g. carbon tax) and subsidized goods underproduced (e.g., flu shots subsidies)
- 2) Restrict or Mandate Private Sale or Purchase:** Restrict the private sale or purchase of overproduced goods (e.g. fuel efficiency requirements), or mandate the private purchase of underproduced goods (e.g., auto insurance)
- 3) Public Provision:** The government can provide the good directly, in order to potentially attain the level of consumption that maximizes social welfare (example is National Defense)
- 4) Public Financing of Private Provision:** Government pays for the good but private sector supplies it (e.g., privately provided health insurance paid for by US government in Medicare-Medicaid)

What Are the Effects of Alternative Interventions?

1) Direct Effects: The effects of government interventions that would be predicted if individuals did not change their behavior in response to the interventions.

Direct effects are relatively easy to compute

2) Indirect Effects: The effects of government interventions that arise only because individuals change their behavior in response to the interventions (sometimes called **unintended effects**)

Empirical public economics analysis tries to estimate indirect effects to inform the policy debate

Example: increasing top income tax rates mechanically raises tax revenue but top earners might work less and earn less, reducing tax revenue relative to mechanical calculation

Why Do Governments Do What They Do?

Political economy: The theory of how the political process produces decisions that affect individuals and the economy

Example: Understanding how the level of taxes and spending is set through voting and voters' preferences

Public choice is a sub-field of political economy from a Libertarian perspective that focuses on **government failures**

government failures = situations where the government does not act in the benefit of society

Normative vs. Positive Public Economics

Normative Public Economics: Analysis of How Things Should be (e.g., should the government intervene in health insurance market? how high should taxes be?, etc.)

Positive Public Economics: Analysis of How Things Really Are (e.g., Does govt provided health care crowd out private health care insurance? Do higher taxes reduce labor supply?)

Positive Public Economics is a required 1st step before we can complete Normative Public Economics

Positive analysis is primarily empirical and Normative analysis is primarily theoretical

Paternalism vs. Individual Failures

In many situations, individuals may not or do not seem to act in their best interests [e.g., many individuals are not able to save for retirement]

Two Polar Views on such situations:

1) **Paternalism [Libertarian View]** Individual failures do not exist and government wants to impose its own preferences against individuals' will

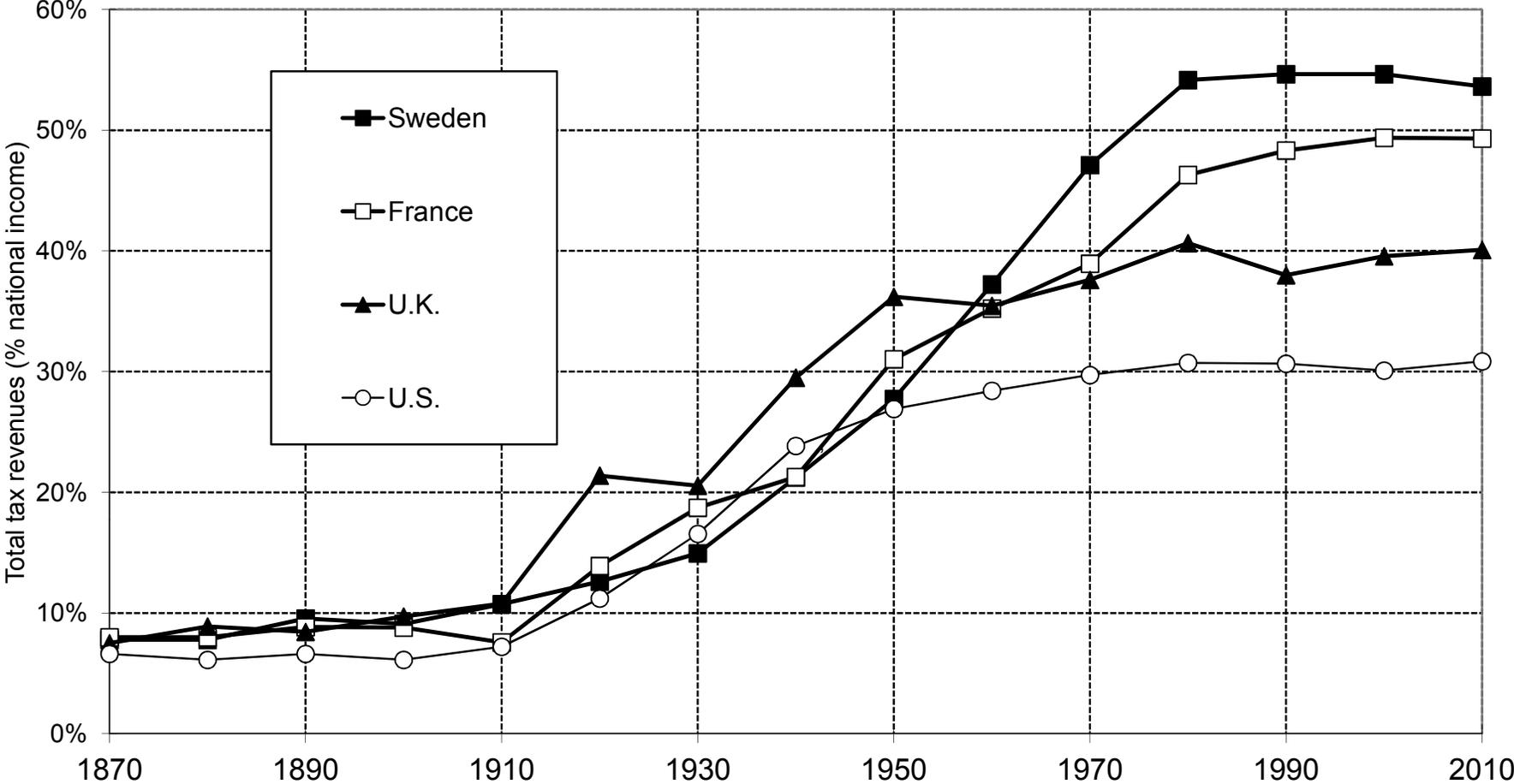
2) **Individual Failures [Behavioral Economics View]** Individual Failures exist: Self-control problems, Cognitive Limitations

Distinguishing the 2 views: Under Paternalism, individuals are opposed to government interventions. If individuals understand they have failures, they will support govt interventions.

Key Facts on Taxes and Spending

- 1) **Government Growth:** Size of government relative to National Income grows dramatically over the process of development from less than 10% in less developed economies to 30-50% in most advanced economies
- 2) **Government Size Stable** in richest countries after 1980
- 3) **Government Growth** is due to the expansion of the **welfare state:** public education, public retirement benefits, public health insurance, income support programs
- 4) **Govt spending > Taxes:** Most rich countries run deficits and have significant public debt (relative to GDP), particularly after Great Recession of 2008

Figure 13.1. Tax revenues in rich countries, 1870-2010

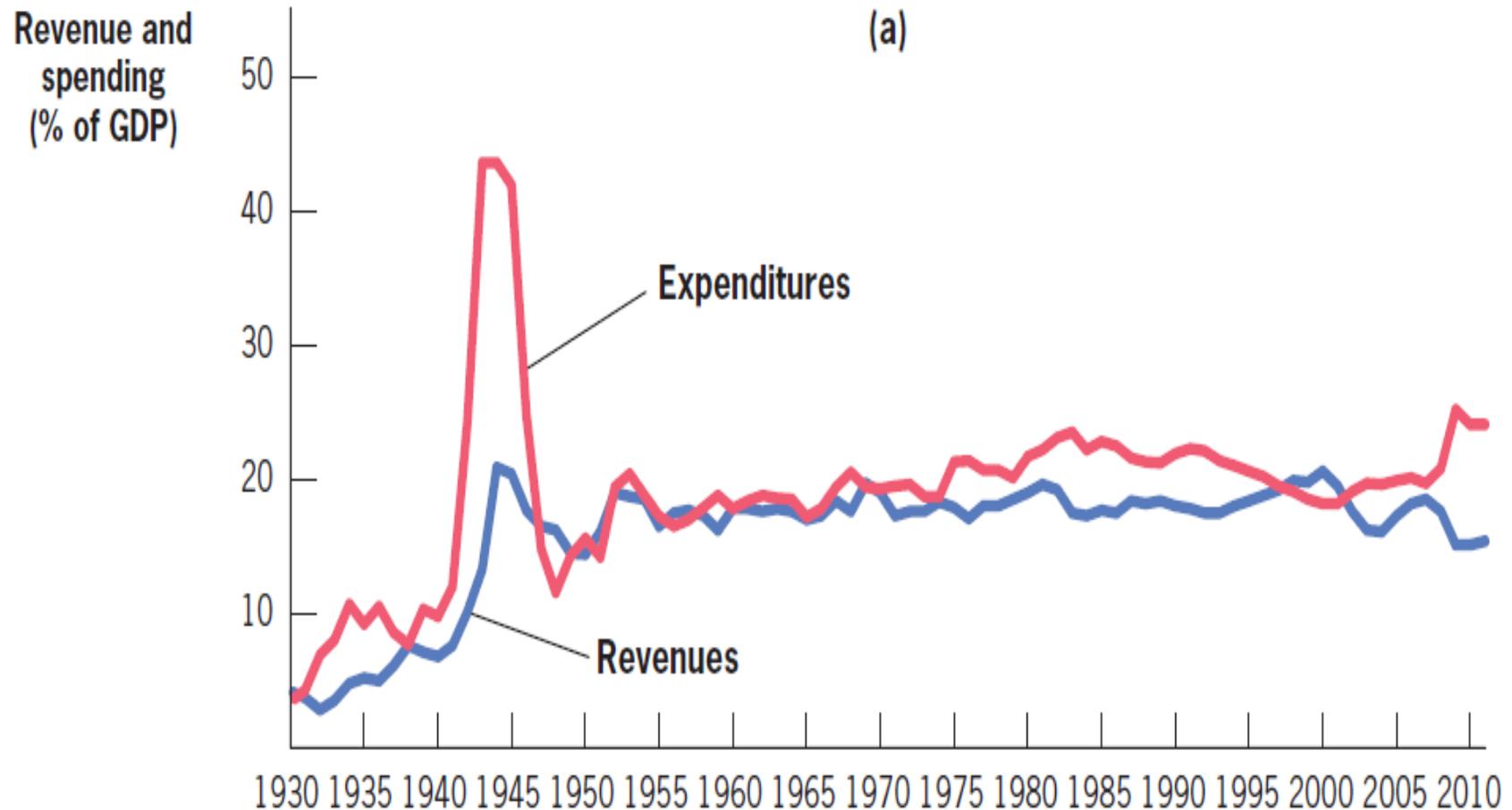


Total tax revenues were less than 10% of national income in rich countries until 1900-1910; they represent between 30% and 55% of national income in 2000-2010. Sources and series: see piketty.pse.ens.fr/capital21c.

Source: Piketty (2014)

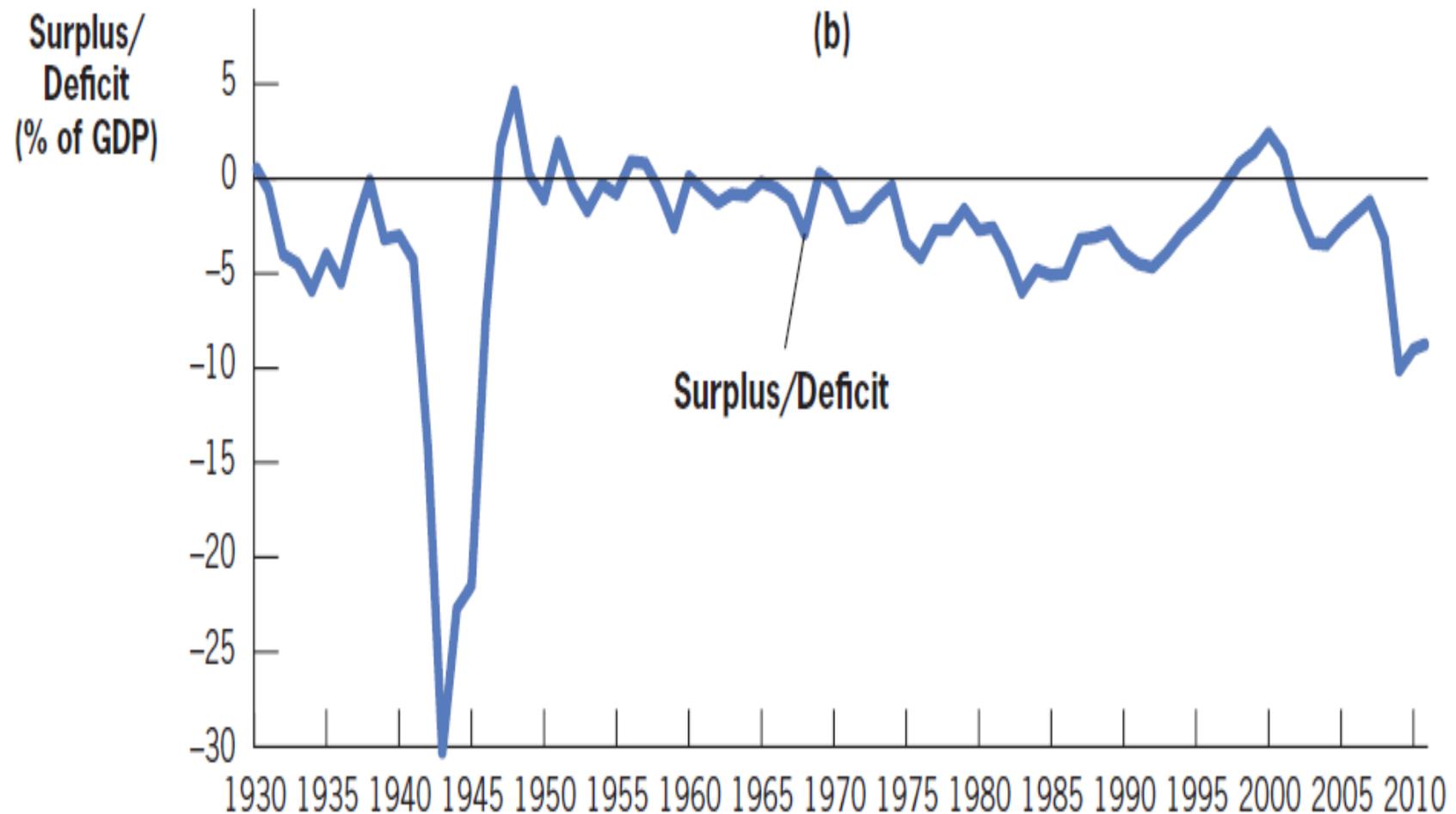
1.2

Federal Revenues and Expenditures, 1930–2011



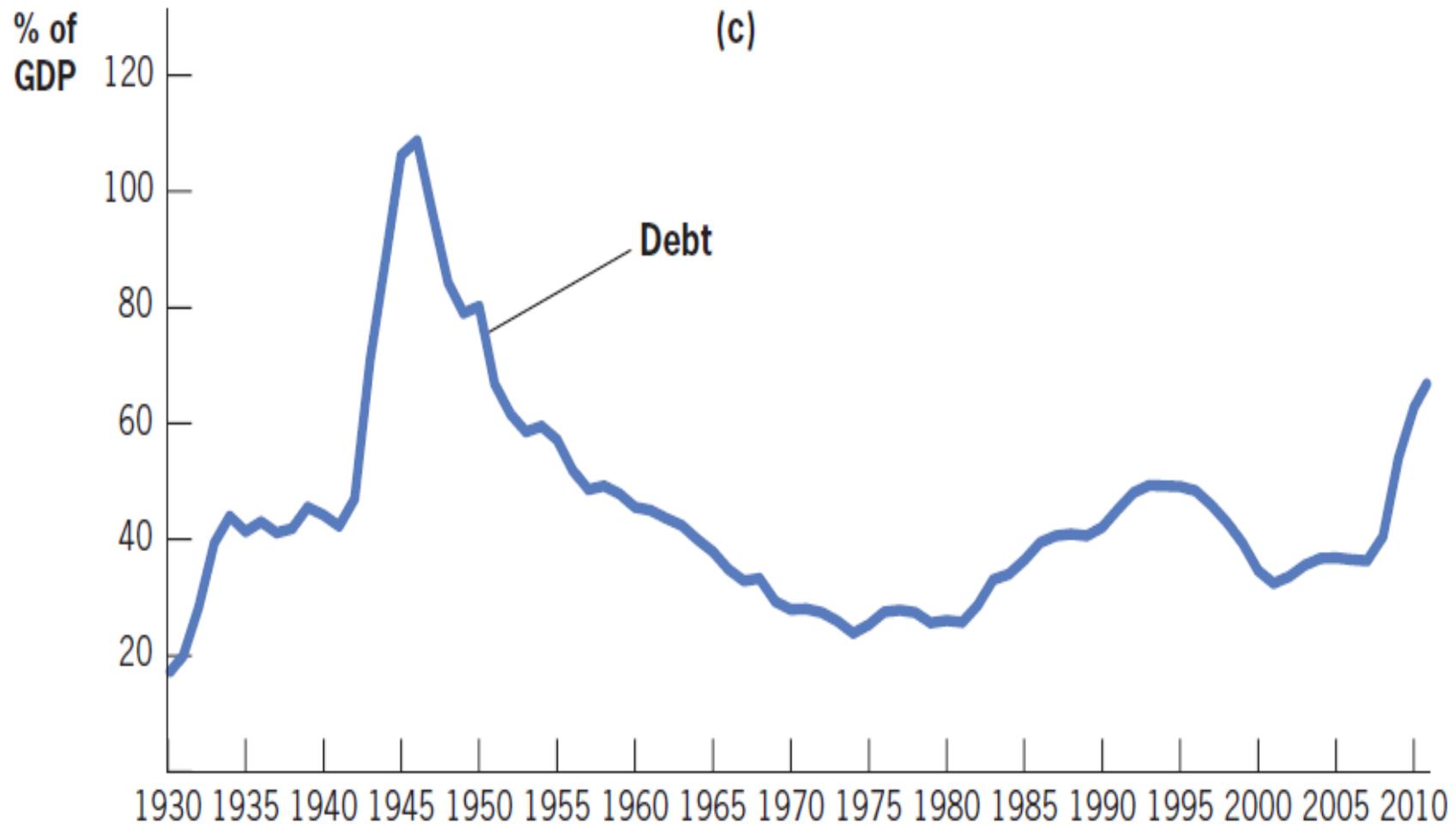
1.2

Federal Surplus/Deficit, 1930–2011



1.2

Federal Debt, 1930–2011



DIFFERENT LEVELS OF GOVERNMENTS

US Federal govt raises about 20% of GDP in taxes

State+Local govts raise about 10% of GDP in taxes

Decentralized states = states where a larger fraction of taxes/spending take place at local level

Decentralized states give additional power to individuals who can also vote with their feet

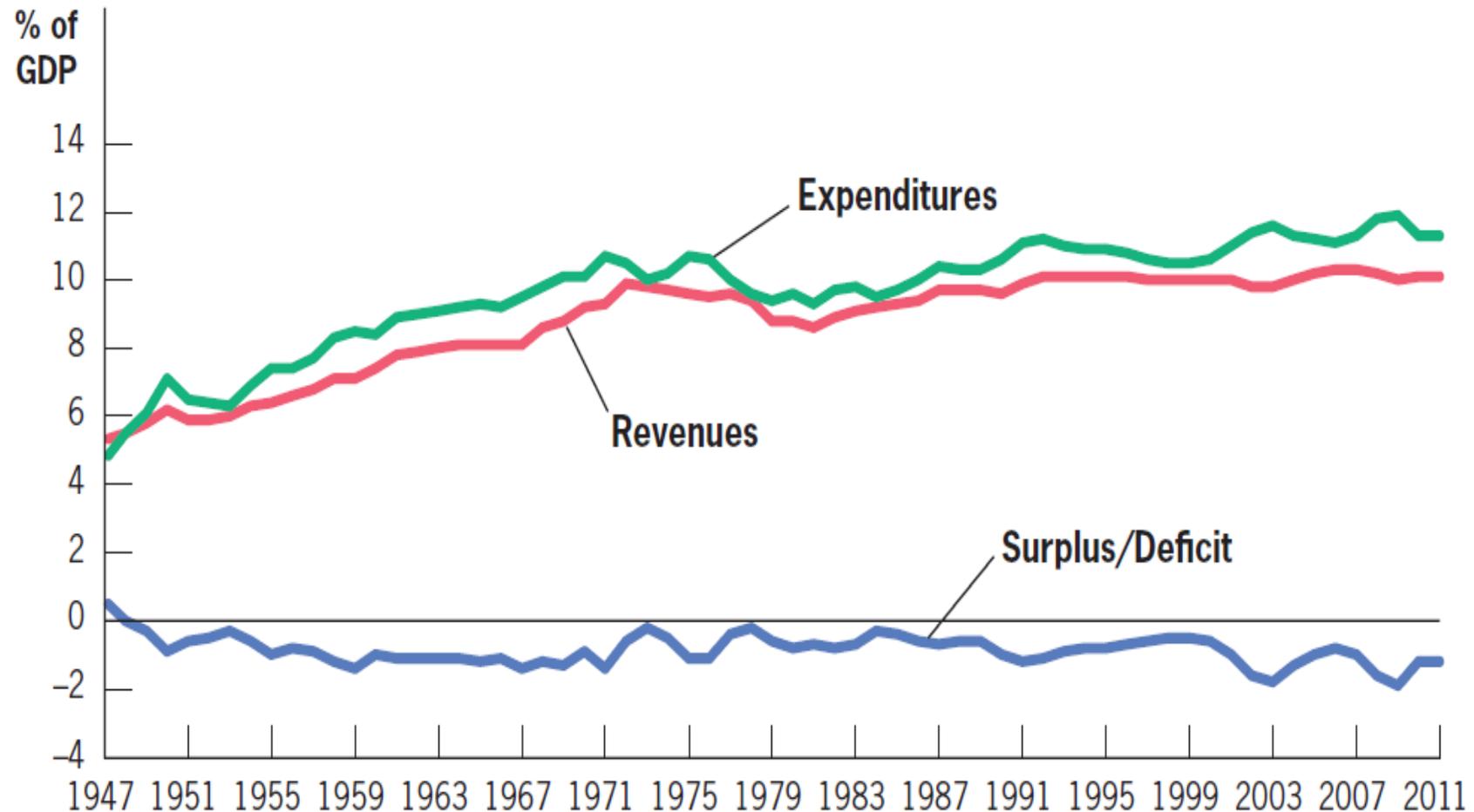
Creates competition between local govts: If local govt is inefficient (high taxes and wasteful spending), residents can leave, putting the local govt out of business

Redistribution through taxes and transfers harder to achieve at local level (rich can leave if local taxes are too high)

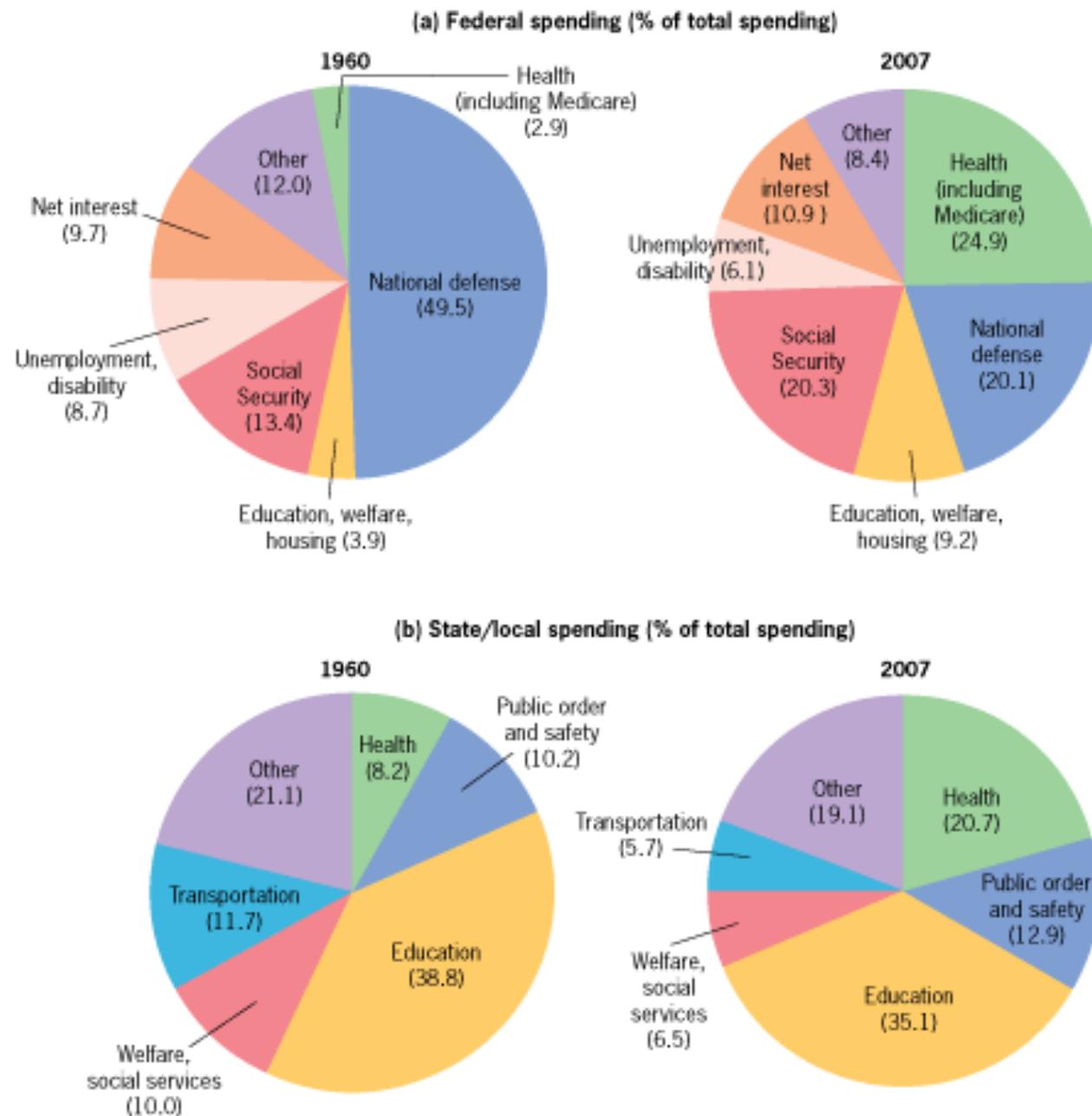
⇒ Conservatives/libertarians tend to prefer decentralized states

1.2

State and Local Government Receipts, Expenditures, and Surplus, 1947–2008



■ FIGURE 1-7



Distribution of Spending

The Distribution of Federal and State Expenditures, 1960 and 2007 • This figure shows the changing composition of federal and state spending over time, as a share of total spending. (a) For the federal government, defense spending has fallen and Social Security and health spending have risen. (b) For the states, the distribution has been more constant, with a small decline in education and welfare spending and a rise in health spending.

DISTRIBUTION OF TAXES

US Federal govt raises about 20% of GDP in taxes, State+Local govt raises about 10% of GDP in taxes.

Main Federal taxes: (1) Individual income tax (40%), (2) payroll taxes on earnings (40%), (3) corporate tax (15%)

Main State taxes: (1) real estate property taxes (30%), (2) sales and excise taxes (30%), (3) individual and corporate state taxes (30%)

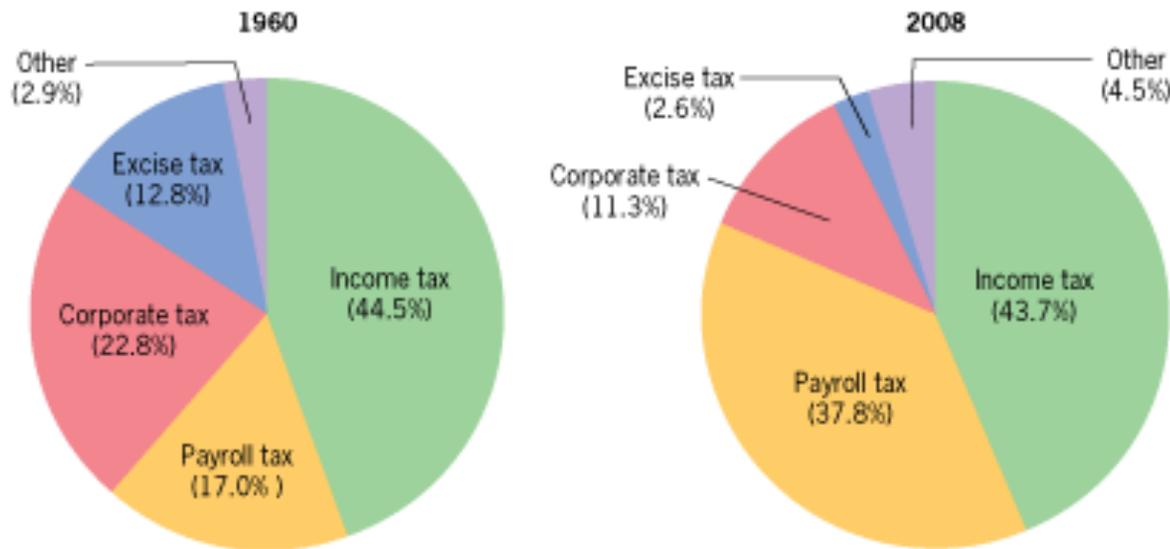
Key questions: who bears the burden of those taxes (tax incidence), what impact do they have on the economy?

■ FIGURE 1-8

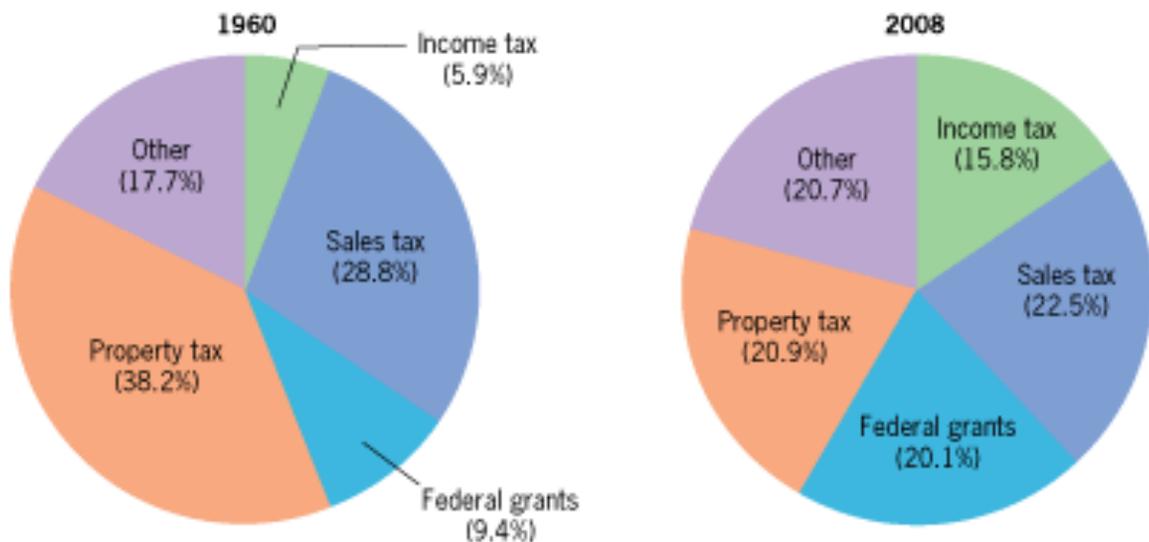
Distribution of Revenue Sources

The Distribution of Federal and State Revenues, 1960 and 2008 • This figure shows the changing composition of federal and state revenue sources over time, as a share of total revenues. (a) At the federal level, there has been a large reduction in corporate and excise tax revenues and a rise in payroll tax revenues. (b) For the states, there has been a decline in property taxes and a rise in income taxes and federal grants.

(a) Federal revenues (% of total revenue)



(b) State/local revenues (% of total revenue)



REGULATORY ROLE OF THE GOVERNMENT

Another critical role the government plays in all nations is that of *regulating economic and social activities*. Examples:

- 1) **Minimum wage** at the Federal level is \$7.25 (States can adopt higher min wages) ⇒ Potential impact on inequality
- 2) The **Food and Drug Administration (FDA)** regulates the labeling and safety of nearly all food products and approves drugs and medical devices to be sold to the public
- 3) The **Occupational Safety and Health Administration (OSHA)** is charged with regulating the workplace safety of American workers
- 4) The **Environmental Protection Agency (EPA)** is charged with minimizing dangerous pollutants in the air, water, and food supplies

PUBLIC DEBATES OVER SOCIAL SECURITY, HEALTH CARE AND EDUCATION

Social Security, health care, and education are each the subject of debate, with both the "liberal" and "conservative" positions holding differing views in their approach to each problem.

Social Security: Social Security is the single largest government expenditure program. The financing structure of this program is basically that today's young workers pay the retirement benefits of today's old.

Health Care: Up to 2013, about 20% of the non-elderly U.S. population did not have health insurance. ObamaCare is reducing this number but US health care costs are very high.

Education: There is an enormous dissatisfaction with current US educational system, highlighted by the poor performance of US K-12 students on international tests.

GOVERNMENT CAN PREVENT PEOPLE FROM MAKING BIG MISTAKES



Source: Bernheim-Rangel (2004)

PUBLIC DEBATES OVER SOCIAL SECURITY, HEALTH CARE AND EDUCATION

Social Security, health care, and education are each the subject of debate, with both the "liberal" and "conservative" positions holding differing views in their approach to each problem.

Social Security: Social Security is the single largest government expenditure program. The financing structure of this program is basically that today's young workers pay the retirement benefits of today's old.

Health Care: Up to 2013, about 20% of the non-elderly U.S. population did not have health insurance. ObamaCare is reducing this number but US health care costs are very high.

Education: There is an enormous dissatisfaction with current US educational system, highlighted by the poor performance of US K-12 students on international tests.

PROFESSOR YAGAN'S RESEARCH

My research (available on my webpage) is in public and labor economics:

- 1) Profits, Taxes, and Inequality
- 2) Labor Market Scars
- 3) Education Effectiveness and Access

I will discuss some of my research in this course when we cover the relevant topics

INSTRUCTION AT A TOP RESEARCH UNIVERSITY

Three things you can get out of a course:

- 1) Information: Facts and tools
- 2) Perspective: What really matters and why
- 3) Inspiration: Belief in the power of ideas and hard work

Research universities are designed to hit #2 and #3 out of the park.

BERKELEY ECONOMICS: AN EMPHASIS ON DATA

BloombergView

HOME

EDITORIAL BOARD

COLUMNISTS

TOPICS

THE LEFT COAST RULES. PHOTOGRAPHER: DAVID MADISON/GETTY IMAGES



ECONOMICS

Econ 101: Chicago? M.I.T.? Nope, Berkeley's on Top

96 JUL 30, 2015 8:00 AM EDT

BERKELEY ECONOMICS: AN EMPHASIS ON DATA

BloombergView

HOME	EDITORIAL BOARD	COLUMNISTS	TOPICS

f
t
in
g+
✉

The point is that Berkeley has done something similar to what the old Chicago School did -- it has changed the entire game. Through intellectual force and creativity, it has turned the ship of the profession in a new direction. We might call it the Berkeley Reformation. The Chicago School made economics about theory; the Berkeley Reformation has made it more about data. The Chicago School made economics about efficiency; the Berkeley Reformation has made it about inequality as well. The Chicago School was Panglossian in its belief that markets work well; the Berkeley Reformation showed deep, fundamental reasons that they break down. The Chicago School described the world in terms of perfectly rational agents; the Berkeley Reformation added the complexity of flawed decision-making.

In the 1980s and 1990s, it could rightly be said that Chicago had conquered the economics world. But in the 2010s, the profession has pointed in Berkeley's direction.

WE WILL FOLLOW THAT TRADITION

This course: Data-driven approach to what government intervention really does to economic efficiency and inequality and what really matters for policy decisions

Along the way: Convey my confidence in you to use your ideas, reason, and hard work to improve lives

REFERENCES

Berheim, B. Douglas and Antonio Rangel “Addiction and Cue-Triggered Decision Processes”, *The American Economic Review*, Vol. 94, No. 5 (Dec., 2004), pp. 1558-1590 (web)

Jonathan Gruber, *Public Finance and Public Policy*, Fourth Edition, 2012 Worth Publishers, Chapter 1

National Center for Education Statistics “Highlights from TIMSS 2007: Mathematics and science achievement of US fourth-and eighth-grade students in an international context.” Institute of Education Sciences, US Department of Education, 2009.(web)

Piketty, Thomas, *Capital in the 21st Century*, Cambridge: Harvard University Press, 2014, Chapter 13, (web)

Piketty, Thomas and Emmanuel Saez “Income Inequality in the United States, 1913-1998”, *Quarterly Journal of Economics*, 118(1), 2003, 1-39, series updated to 2012 in September 2013 (web)

Undergraduate Public Economics

Danny Yagan
UC Berkeley

Lecture 2

Theoretical Tools of Public Finance

THEORETICAL AND EMPIRICAL TOOLS

Theoretical tools: The set of tools designed to understand the mechanics behind economic decision making.

Economists model individuals' choices using the concepts of utility function maximization subject to budget constraint

Empirical tools: The set of tools designed to analyze data and answer questions raised by theoretical analysis.

UTILITY MAPPING OF PREFERENCES

Utility function: A utility function is some mathematical function translating consumption into utility:

$$U = u(X_1, X_2, X_3, \dots)$$

where X_1, X_2, X_3 , and so on are the goods consumed by the individual

Example with two goods: $u(X_1, X_2) = \sqrt{X_1 \cdot X_2}$ with X_1 number of movies, X_2 number of music songs

Individual utility increases with the level of consumption of each good

PREFERENCES AND INDIFFERENCE CURVES

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

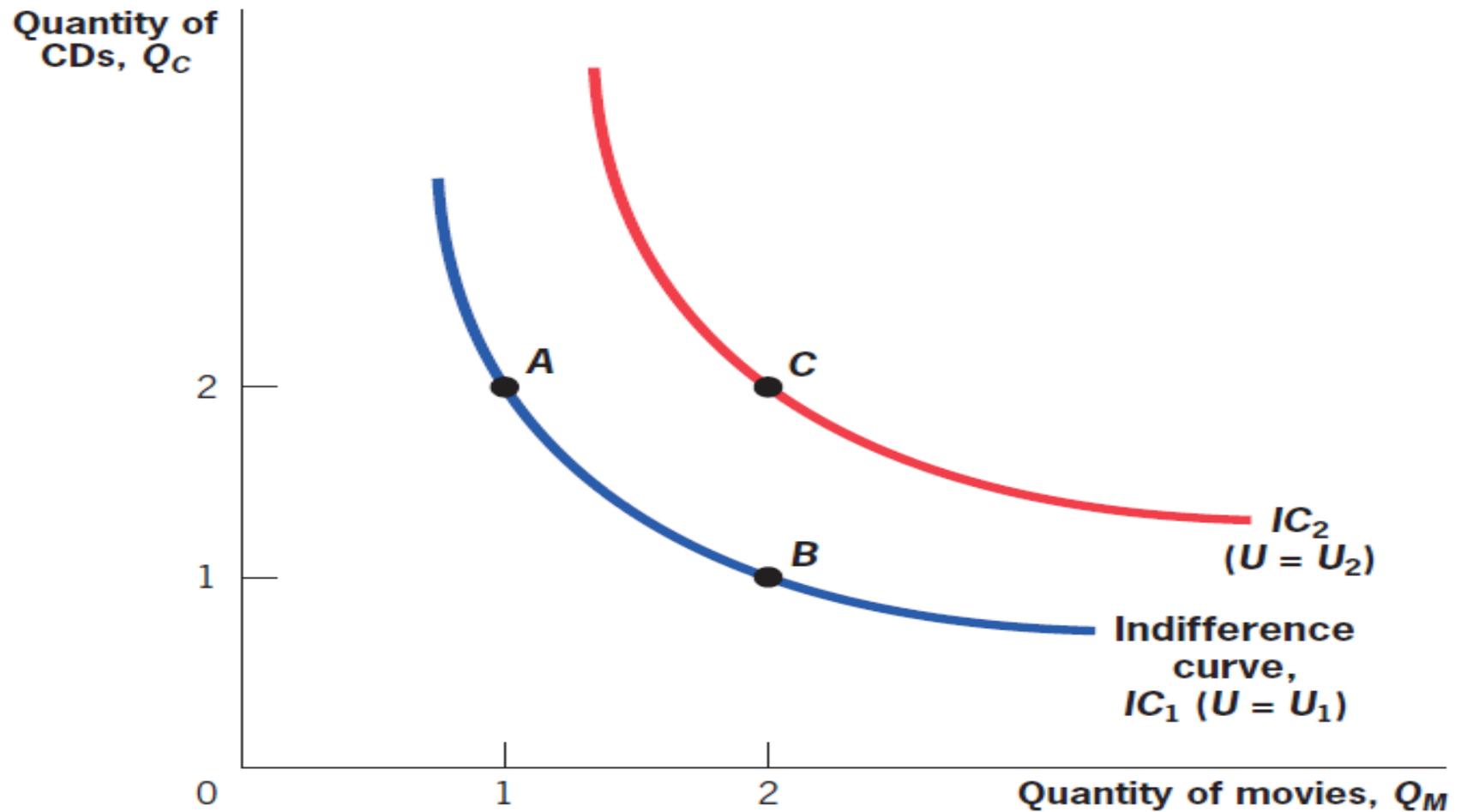
Mathematically, indifference curve giving utility level \bar{U} is given by the set of bundles (X_1, X_2) such that $u(X_1, X_2) = \bar{U}$

Indifference curves have two essential properties, both of which follow naturally from the more-is-better assumption:

1. Consumers prefer higher indifference curves.
2. Indifference curves are always downward sloping.

2.1

Preferences and Indifference Curves



MARGINAL UTILITY

Marginal utility: The additional increment to utility obtained by consuming an additional unit of a good:

Marginal utility of good 1 is defined as:

$$MU_1 = \frac{\partial u}{\partial X_1} \simeq \frac{u(X_1 + dX_1, X_2) - u(X_1, X_2)}{dX_1}$$

It is the derivative of utility with respect to X_1 keeping X_2 constant (called the partial derivative)

Example:

$$u(X_1, X_2) = \sqrt{X_1 \cdot X_2} \Rightarrow \frac{\partial u}{\partial X_1} = \frac{\sqrt{X_2}}{2\sqrt{X_1}}$$

This utility function described exhibits the important principle of **diminishing marginal utility**: $\partial u / \partial X_1$ decreases with X_1 : the consumption of each additional unit of a good gives less extra utility than the consumption of the previous unit

MARGINAL RATE OF SUBSTITUTION

Marginal rate of substitution (MRS): The *MRS* is equal to (minus) the slope of the indifference curve, the rate at which the consumer will trade the good on the vertical axis for the good on the horizontal axis.

Marginal rate of substitution between good 1 and good 2 is:

$$MRS_{1,2} = \frac{MU_1}{MU_2}$$

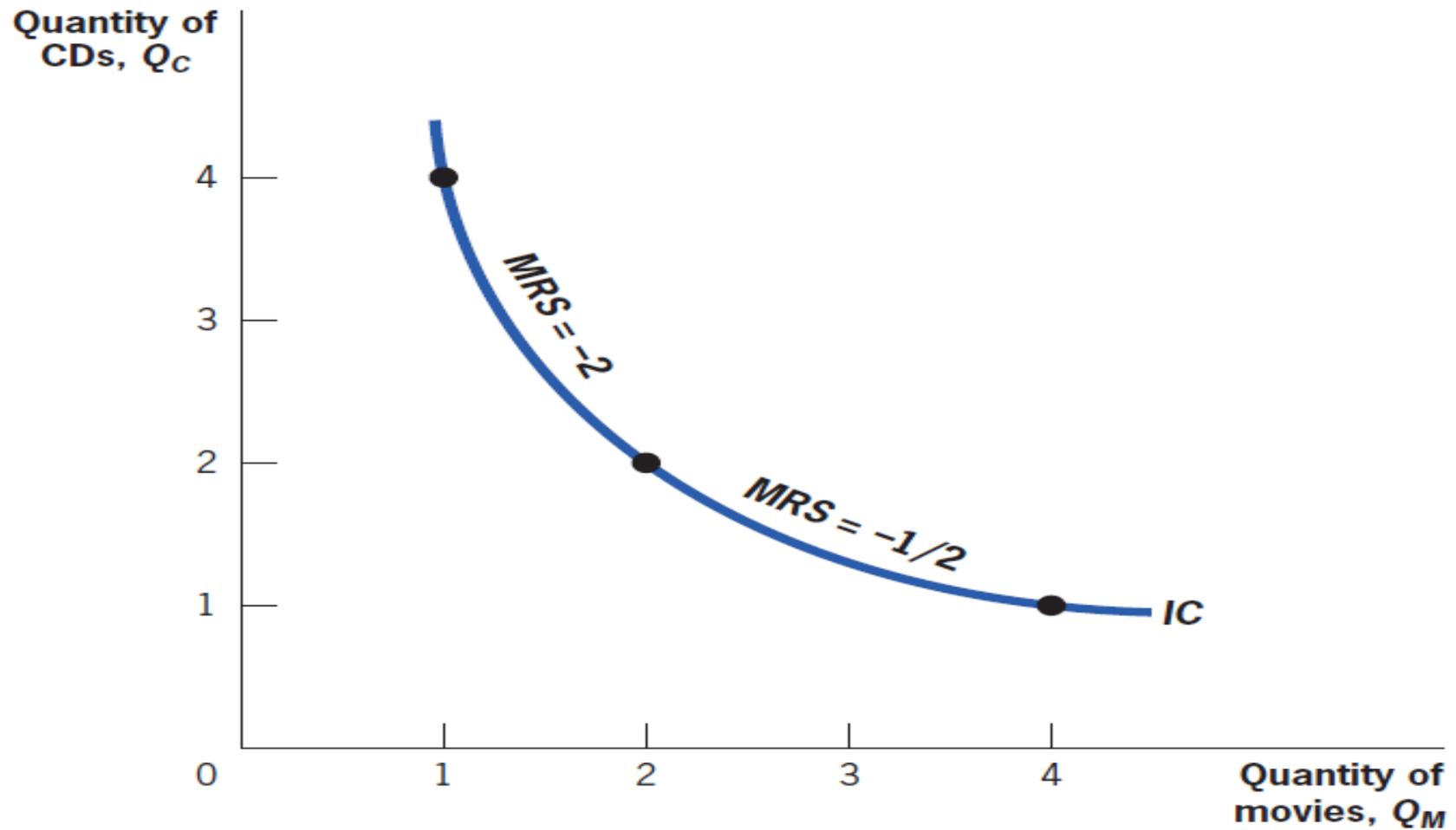
Individual is indifferent between 1 unit of good 1 and $MRS_{1,2}$ units of good 2.

Example:

$$u(X_1, X_2) = \sqrt{X_1 \cdot X_2} \Rightarrow MRS_{1,2} = \frac{X_2}{X_1}$$

2.1

Marginal Rate of Substitution



BUDGET CONSTRAINT

Budget constraint: A mathematical representation of all the combinations of goods an individual can afford to buy if she spends her entire income.

$$p_1X_1 + p_2X_2 = Y$$

with p_i price of good i , and Y disposable income.

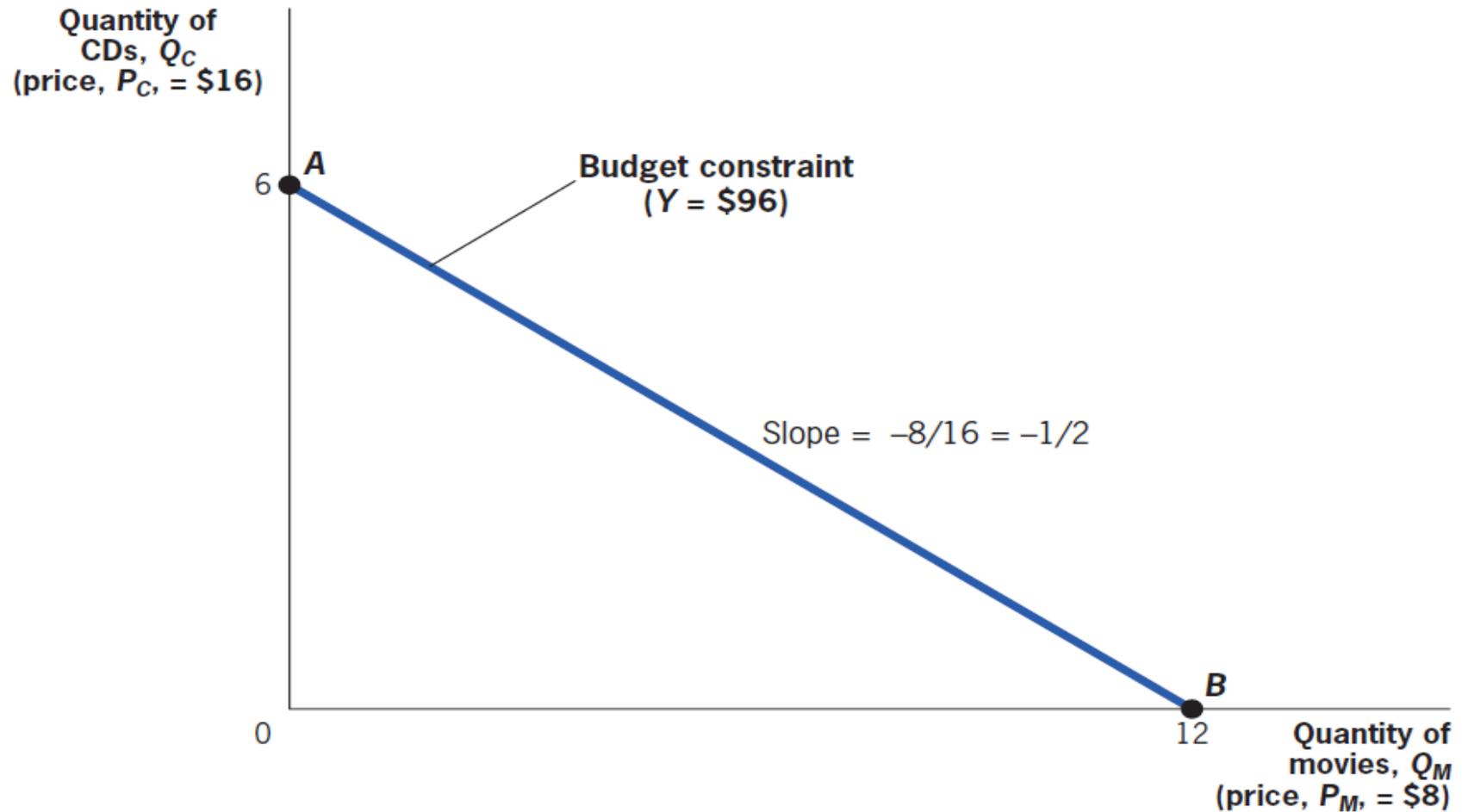
Budget constraint defines a linear set of bundles the consumer can purchase with its disposable income Y

$$X_2 = \frac{Y}{p_2} - \frac{p_1}{p_2}X_1$$

The slope of the budget constraint is $-p_1/p_2$

2.1

Budget Constraints



UTILITY MAXIMIZATION

Individual maximizes utility subject to budget constraint:

$$\max_{X_1, X_2} u(X_1, X_2) \quad \text{subject to} \quad p_1 X_1 + p_2 X_2 = Y$$

$$\text{Solution:} \quad MRS_{1,2} = \frac{p_1}{p_2}$$

Proof: Budget implies that $X_2 = (Y - p_1 X_1)/p_2$

Individual chooses X_1 to maximize $u(X_1, (Y - p_1 X_1)/p_2)$

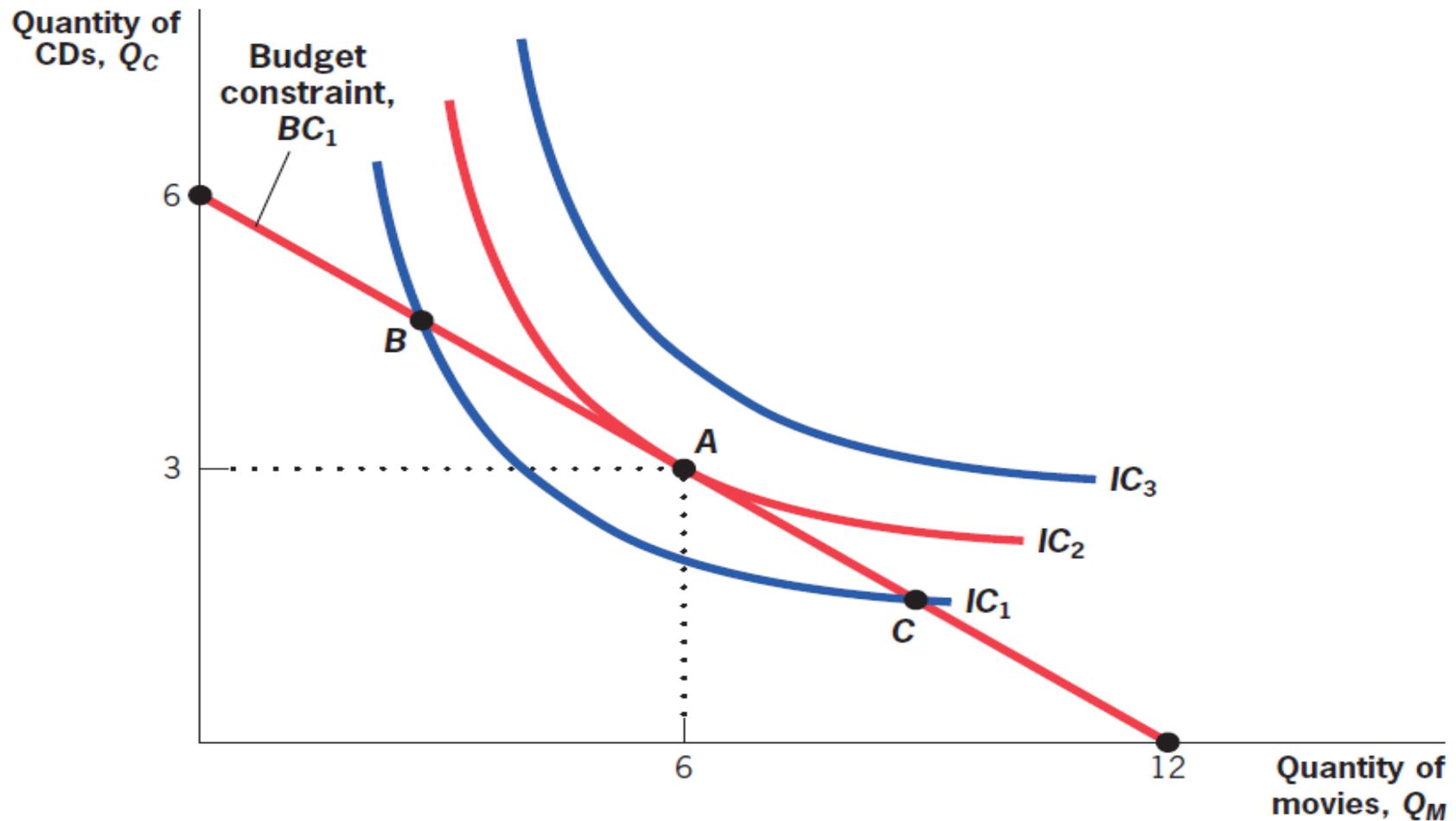
The first order condition (FOC) is:

$$\frac{\partial u}{\partial X_1} - \frac{p_1}{p_2} \cdot \frac{\partial u}{\partial X_2} = 0.$$

At the optimal choice, the individual is indifferent between buying 1 extra unit of good 1 for \$ p_1 and buying p_1/p_2 extra units of good 2 (also for \$ p_1).

2.1

Putting It All Together: Constrained Choice



INCOME AND SUBSTITUTION EFFECTS

Let us denote by $p = (p_1, p_2)$ the price vector

Individual maximization generates demand functions $X_1(p, Y)$ and $X_2(p, Y)$

How does $X_1(p, Y)$ vary with p and Y ?

Those are called price and income effects

Example: $u(X_1, X_2) = \sqrt{X_1 \cdot X_2}$ then $MRS_{1,2} = X_2/X_1$.

Utility maximization implies $X_2/X_1 = p_1/p_2$ and hence $p_1X_1 = p_2X_2$

Budget constraint $p_1X_1 + p_2X_2 = Y$ implies $p_1X_1 = p_2X_2 = Y/2$

Demand functions: $X_1(p, Y) = Y/(2p_1)$ and $X_2(p, Y) = Y/(2p_2)$

INCOME EFFECTS

Income effect is the effect of giving extra income Y on the demand for goods: How does $X_1(p, Y)$ vary with Y ?

Normal goods: Goods for which demand increases as income Y rises: $X_1(p, Y)$ increases with Y (most goods are normal)

Inferior goods: Goods for which demand falls as income Y rises: $X_1(p, Y)$ decreases with Y (example: you use public transportation less when you are rich enough to buy a car)

Example: if leisure is a normal good, you work less (i.e. get more leisure) if you are given a transfer

PRICE CHANGES

How does $X_1(p_1, p_2, Y)$ vary with p_1 ?

Changing p_1 affects the slope of the budget constraint and can be decomposed into 2 effects:

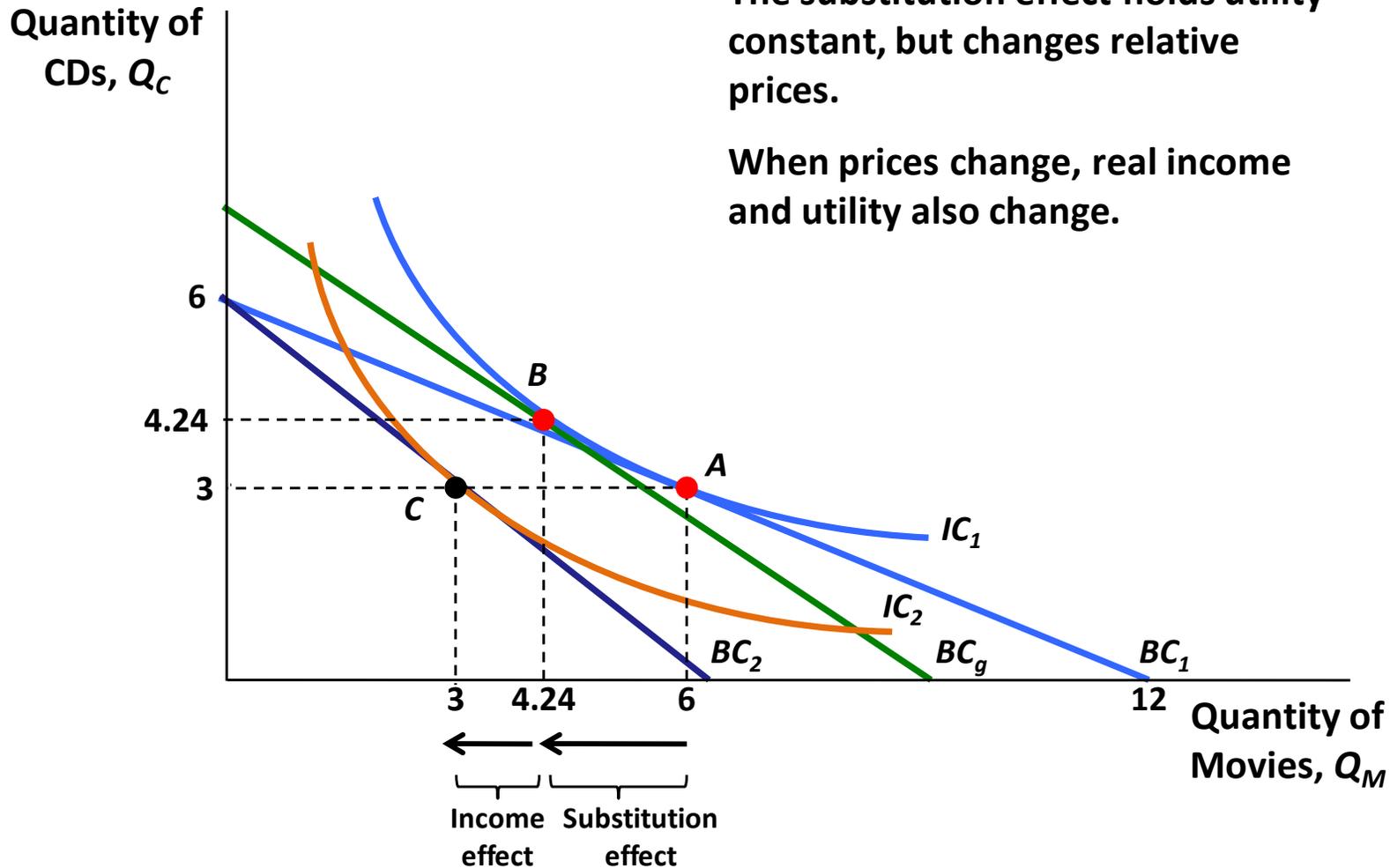
1) Substitution (“price”) effect: Holding utility constant, a relative rise in the price of a good will always cause an individual to choose less of that good

2) Income effect: A rise in the price of a good will typically cause an individual to choose less of all goods because her income can purchase less than before

For normal goods, an increase in p_1 reduces $X_1(p_1, p_2, Y)$ through both substitution and income effects

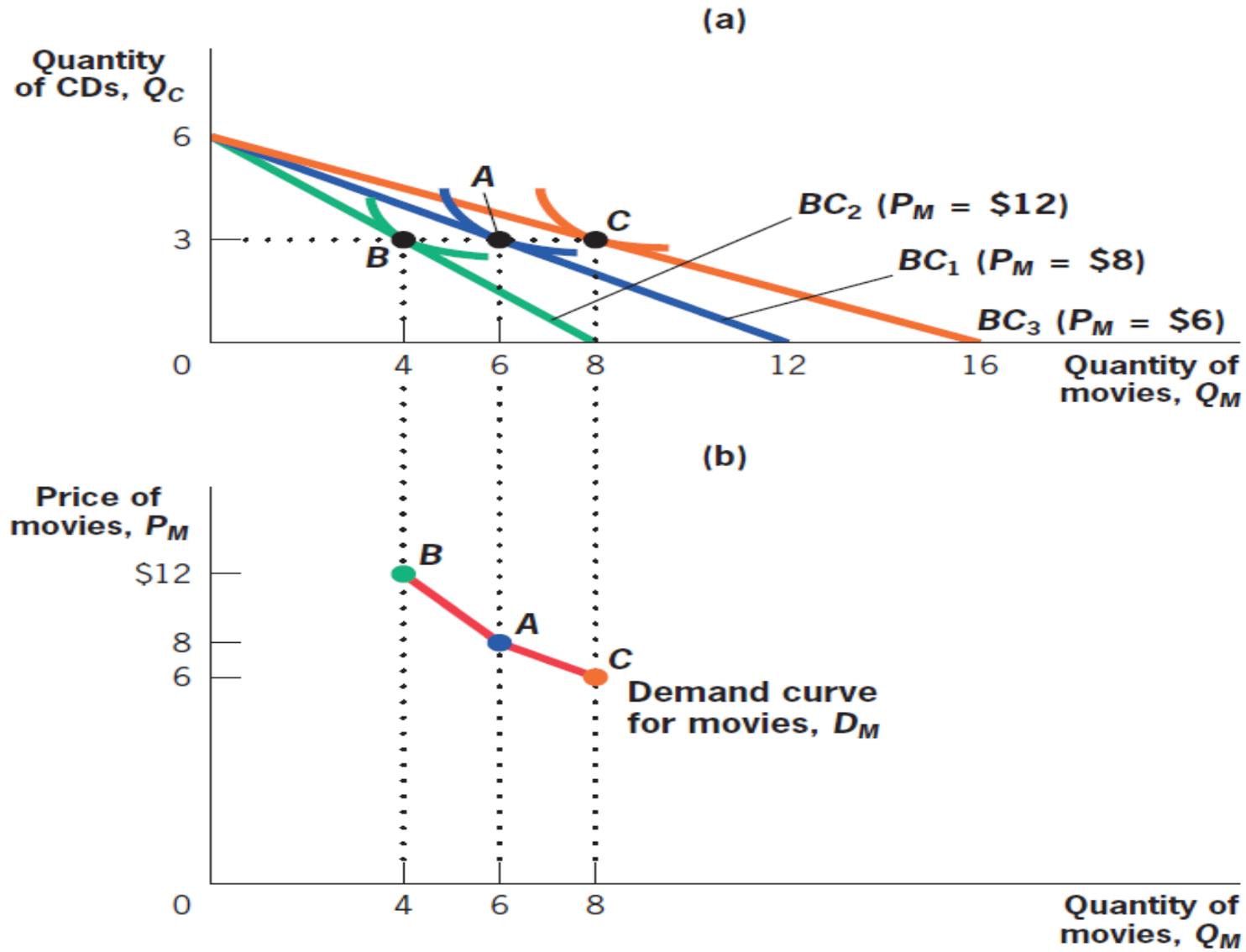
2.1

The Effects of Price Changes: Substitution and Income Effects



2.3

Demand Curves



ELASTICITY OF DEMAND

Each individual has a demand for each good that depends on the price p of the good. Aggregating across all individuals, we get aggregate demand $D(p)$ for the good

At price p , demand is $D(p)$ and p is the \$ value for consumers of the marginal (last) unit consumed

Demand graph: quantity on X-axis, price on Y-axis

Elasticity of demand ε : The % change in demand caused by a 1% change in the price of that good:

$$\varepsilon = \frac{\% \text{ change in quantity demanded}}{\% \text{ change in price}} = \frac{\Delta D / D}{\Delta p / p} = \frac{p}{D} \frac{dD}{dp}$$

Elasticities are unit free

PROPERTIES OF ELASTICITY OF DEMAND

- 1) Typically negative, since quantity demanded typically falls as price rises.
- 2) Typically not constant along a demand curve.
- 3) With vertical demand curve, demand is **perfectly inelastic** ($\varepsilon = 0$).
- 4) With horizontal demand curve, demand is **perfectly elastic** ($\varepsilon = -\infty$).
- 5) The effect of one good's prices on the demand for another good is the **cross-price** elasticity. Typically, not zero.

PRODUCERS

Producers (typically firms) use technology to transform inputs (labor and capital) into outputs (consumption goods)

Goal of producers is to maximize profits = sales of outputs minus costs of inputs

Production decisions (for given prices) define supply functions

SUPPLY CURVES

Marginal cost: The incremental cost to a firm of producing one more unit of a good

Profits: The difference between a firm's revenues and costs, maximized when marginal revenues equal marginal costs

Supply curve $S(p)$ is the quantity that firms in aggregate are willing to supply at each price: typically upward sloping with price due to decreasing returns to scale

At price p , producers produce $S(p)$, and the \$ cost of producing the marginal (last) unit is p

Elasticity of supply ϵ_S is defined as

$$\epsilon_S = \frac{\% \text{ change in quantity supplied}}{\% \text{ change in price}} = \frac{\Delta S/S}{\Delta p/p} = \frac{p dS}{S dp}$$

MARKET EQUILIBRIUM

Demanders and suppliers interact on markets

Market equilibrium: The equilibrium is the price p^* such that $D(p^*) = S(p^*)$

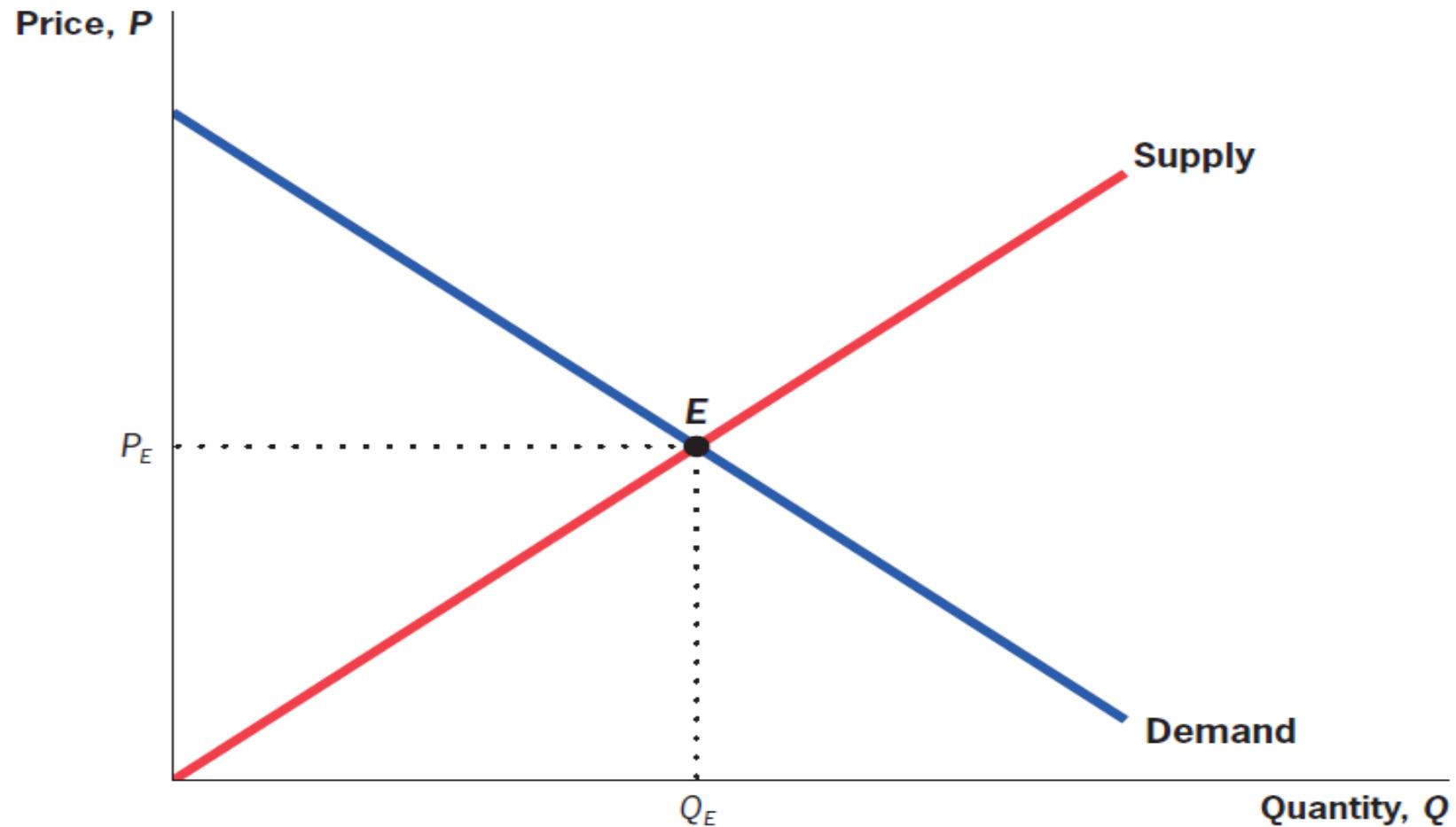
In the simple diagram, p^* is unique if $D(p)$ decreases with p and $S(p)$ increases with p

If $p > p^*$, then supply exceeds demand, and price needs to fall to equilibrate supply and demand

If $p < p^*$, then demand exceeds supply, and price needs to increase to equilibrate supply and demand

2.3

Equilibrium: Graphical Representation



SOCIAL EFFICIENCY

Social efficiency represents the net gains to society from all trades that are made in a particular market, and it consists of two components: consumer and producer surplus.

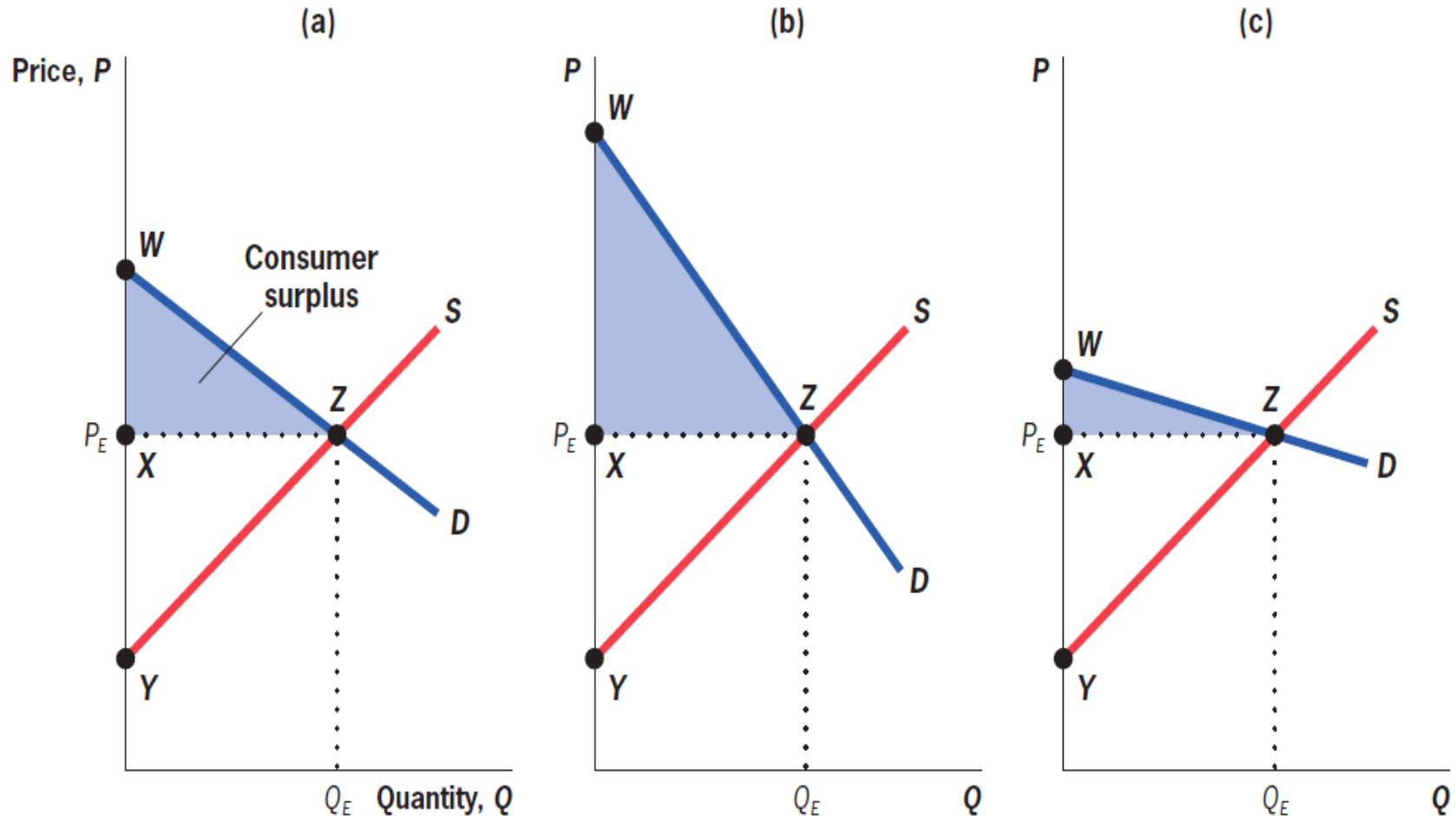
Consumer surplus: The benefit that consumers derive from consuming a good, above and beyond the price they paid for the good. It is the area below demand curve and above market price.

Producer surplus: The benefit producers derive from selling a good, above and beyond the cost of producing that good. It is the area above supply curve and below market price.

Total social surplus (social efficiency): The sum of consumer surplus and producer surplus. It is the area above supply curve and below demand curve.

2.3

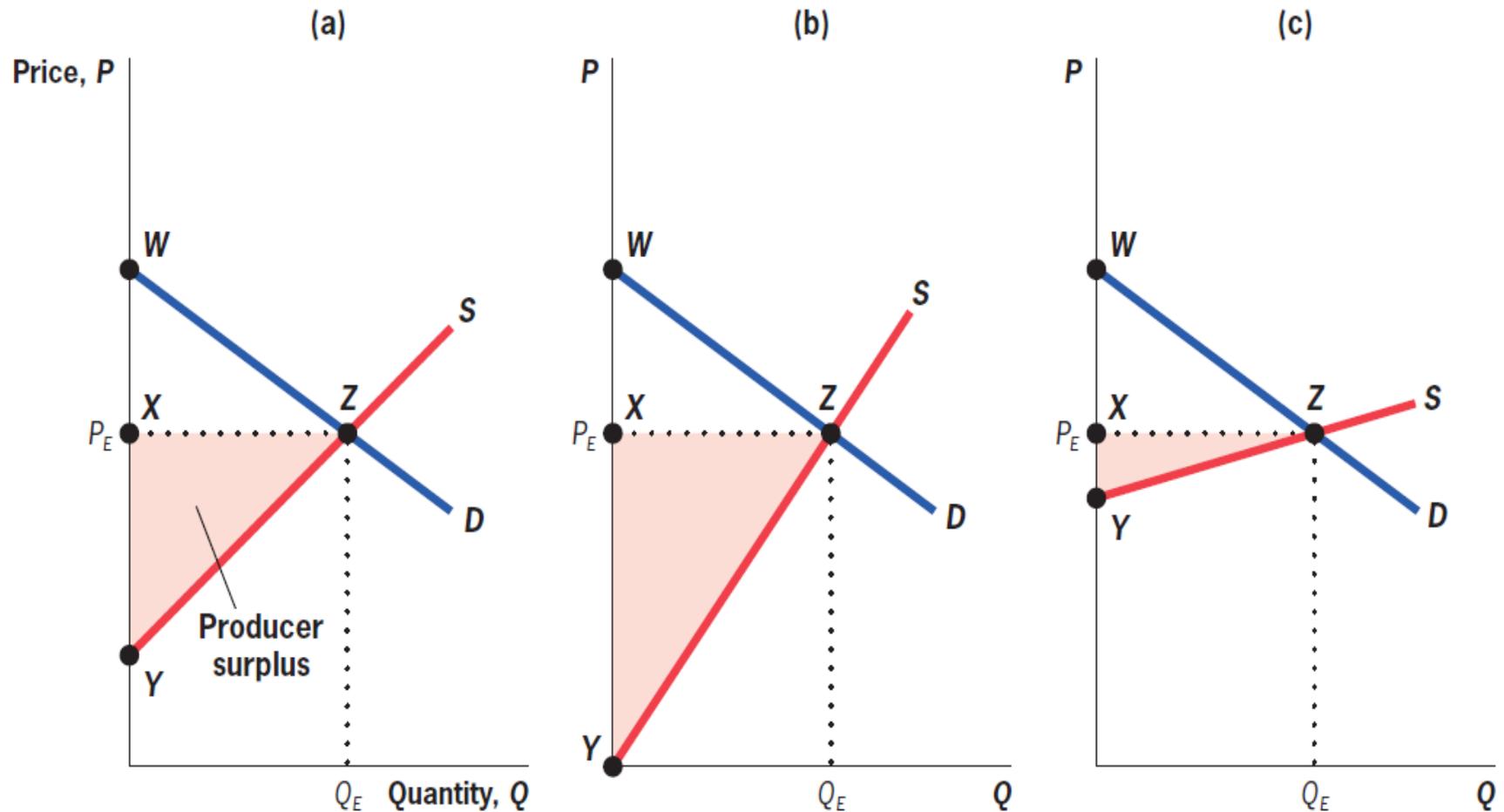
Consumer Surplus: Graphical Representation



- Consumer surplus is the area under the demand curve since demand = willingness to pay.

2.3

Producer Surplus: Graphical Representation



- Producer surplus is the area above the supply curve since supply = marginal cost.

COMPETITIVE EQUILIBRIUM MAXIMIZES SOCIAL EFFICIENCY

First Fundamental Theorem of Welfare Economics:

The competitive equilibrium where supply equals demand, maximizes social efficiency.

Deadweight loss: The reduction in social efficiency from denying trades for which benefits exceed costs when quantity differs from the socially efficient quantity

Key rule: Deadweight loss triangles point to the social optimum, and grow outward from there.

The simple efficiency result from the 1-good diagram can be generalized into the first welfare theorem (Arrow-Debreu, 1940s), most important result in economics

Generalization: 1st Welfare Theorem

1st Welfare Theorem: If (1) no externalities, (2) perfect competition [individuals and firms are price takers], (3) perfect information, (4) agents are rational, then private market equilibrium is **Pareto efficient**

Pareto efficient: Impossible to find a technologically feasible allocation that improves everybody's welfare

Pareto efficiency is desirable but a very weak requirement (a single person consuming everything is Pareto efficient)

Government intervention may be particularly desirable if the assumptions of the 1st welfare theorem fail, i.e., when there are market failures \Rightarrow Govt intervention can potentially improve everybody's welfare

Second part of class considers such market failure situations

2nd Welfare Theorem

Even with no market failures, free market outcome might generate substantial inequality. Inequality is seen as the biggest issue with market economies.

2nd Welfare Theorem: Any Pareto Efficient allocation can be reached by

(1) Suitable redistribution of initial endowments [individualized **lump-sum** taxes based on individual characteristics and not behavior]

(2) Then letting markets work freely

⇒ No conflict between efficiency and equity

2nd Welfare Theorem fallacy

In reality, 2nd welfare theorem does not work because redistribution of initial endowments is not feasible (because initial endowments cannot be observed by the government)

⇒ govt needs to use **distortionary** taxes and transfers based on economic outcomes (such as income or working situation)

⇒ Conflict between efficiency and equity: **Equity-Efficiency trade-off**

First part of class considers policies that trade-off equity and efficiency

Illustration of 2nd Welfare Theorem Fallacy

Suppose economy is populated 50% with disabled people unable to work (hence they earn \$0) and 50% with able people who can work and earn \$100

Free market outcome: disabled have \$0, able have \$100

2nd welfare theorem: govt is able to tell apart the disabled from the able [even if the able do not work]

⇒ can tax the able by \$50 [regardless of whether they work or not] to give \$50 to each disabled person ⇒ the able keep working [otherwise they'd have zero income and still have to pay \$50]

Real world: govt can't tell apart disabled from non working able

⇒ \$50 tax on workers + \$50 transfer on non workers destroys all incentives to work ⇒ govt can no longer do full redistribution ⇒ Trade-off between equity and size of the pie

SOCIAL WELFARE FUNCTIONS

Social welfare function (SWF): A function that combines the utility functions of all individuals into an overall social utility function.

UTILITARIAN SOCIAL WELFARE FUNCTION

With a utilitarian social welfare function, society's goal is to maximize the sum of individual utilities:

$$SWF = U_1 + U_2 + \dots + U_N$$

The utilities of all individuals are given equal weight, and summed to get total social welfare

If marginal utility of money decreases with income (satiation), utilitarian criterion values redistribution from rich to poor

Taking \$1 for a rich person decreases his utility by a small amount, giving the \$1 to a poor person increases his utility by a large amount \Rightarrow Transfers from rich to poor increase total utility

RAWLSIAN SOCIAL WELFARE FUNCTION

Rawls (1971) proposed that society's goal should be to maximize the well-being of its worst-off member. The Rawlsian SWF has the form:

$$SWF = \min(U_1, U_2, \dots, U_N)$$

Since social welfare is determined by the minimum utility in society, social welfare is maximized by maximizing the well-being of the worst-off person in society (=maxi-min)

Rawlsian criterion is even more redistributive than utilitarian criterion: society wants to extract as much tax revenue as possible from the middle and rich to make transfers to the poor as large as possible

OTHER SOCIAL JUSTICE PRINCIPLES

Standard welfarist approach is based on individual utilities. This fails to capture important elements of actual debates on redistribution and fairness

1) Commodity egalitarianism: Society should ensure that individuals meet a set of basic needs (seen as rights), but that beyond that point income distribution is irrelevant

⇒ Rich countries today consider free K-12 education, universal health care, decent retirement/disability benefits as rights

2) Equality of opportunity: Society should ensure that all individuals have equal opportunities for success

⇒ Individuals should be compensated for inequalities they are not responsible for (e.g., family background, inheritance, intrinsic ability) but not for inequalities they are responsible for (being hard working vs. loving leisure)

TESTING PEOPLE SOCIAL PREFERENCES

Saez-Stantcheva '13 survey people online (using Amazon MTurk) by asking hypothetical questions to elicit social preferences.

Key findings:

- 1) People typically do not have “utilitarian” social justice principles (consumption lover not seen as more deserving than frugal person)
- 2) People put weight on whether income has been earned through effort vs. not (hard working vs. leisure lover)
- 3) People put a lot of weight of what people would have done absent the government intervention (deserving poor vs. free loaders)

Which of the following two individuals do you think is most deserving of a \$1,000 tax break?

Individual A earns \$50,000 per year, pays \$10,000 in taxes and hence nets out \$40,000. She greatly enjoys spending money, going out to expensive restaurants, or traveling to fancy destinations. She always feels that she has too little money to spend.

Individual B earns the same amount, \$50,000 per year, also pays \$10,000 in taxes and hence also nets out \$40,000. However, she is a very frugal person who feels that her current income is sufficient to satisfy her needs.

-
- Individual A is most deserving of the \$1,000 tax break
 - Individual B is most deserving of the \$1,000 tax break
 - Both individuals are exactly equally deserving of the tax \$1,000 break

>>

Which of the following two individuals is most deserving of a \$1,000 tax break?

Individual A earns \$30,000 per year, by working in two different jobs, 60 hours per week at \$10/hour. She pays \$6,000 in taxes and nets out \$24,000. She is very hard-working but she does not have high-paying jobs so that her wage is low.

Individual B also earns the same amount, \$30,000 per year, by working part-time for 20 hours per week at \$30/hour. She also pays \$6,000 in taxes and hence nets out \$24,000. She has a good wage rate per hour, but she prefers working less and earning less to enjoy other, non-work activities.

-
- Individual A is most deserving of the \$1,000 tax break
 - Individual B is most deserving of the \$1,000 tax break
 - Both individuals are exactly equally deserving of the \$1,000 tax break

>>

We assume now that the government can increase benefits by \$1,000 for some recipients of government benefits.

Which of the following four individuals is most deserving of the \$1,000 increase in benefits?

Please drag and drop the four individuals into the appropriate boxes on the left. The upper box, marked 1 should contain the individual you think is most deserving. The box labeled "2" should contain the second most-deserving individual, etc.. Please note that you can put two individuals in the same box if you think that they are equally deserving.

Individual A gets \$15,000 per year in Disability Benefits because she cannot work due to a disability and has no other resources.

Individual B gets \$15,000 per year in Unemployment Benefits and has no other resources. She lost her job and has not been able to find a new job even though she has been actively looking for one.

Individual C gets \$15,000 per year in Unemployment Benefits and has no other resources. She lost her job but has not been looking actively for a new job, because she prefers getting less but not having to work.

Individual D gets \$15,000 per year in Welfare Benefits and Food Stamps and has no other resources. She is not looking for a job actively because she can get by living off those government provided benefits.

Source: survey in Saez and Stantcheva (2013)

Table 2: Revealed Social Preferences

	(1)	(2)	(3)	(4)
A. Consumption lover vs. Frugal				
	Consumption lover > Frugal	Consumption lover = Frugal	Consumption lover < Frugal	
# obs. = 1,125	4.1%	74.4%	21.5%	
B. Hardworking vs. leisure lover				
	Hardworking > Leisure lover	Hardworking = Leisure lover	Hardworking < Leisure lover	
# obs. = 1,121	42.7%	54.4%	2.9%	
C. Transfer Recipients and free loaders				
	Disabled person unable to work	Unemployed looking for work	Unemployed not looking for work	Welfare recipient not looking for work
# obs. = 1,098				
Average rank (1-4) assigned	1.4	1.6	3.0	3.5
% assigned first rank	57.5%	37.3%	2.7%	2.5%
% assigned last rank	2.3%	2.9%	25.0%	70.8%

Notes: This table reports preferences for giving a tax break and or a benefit increase across individuals in various scenarios. Panel A considers two individuals with the same earnings, same taxes, and same disposable income but high marginal utility of income (consumption lover) vs. low marginal utility of income (frugal). In contrast to utilitarianism, 74% of people report that consumption loving is irrelevant and 21.5% think the frugal person is most deserving. Panel B considers two individuals with the same earnings, same taxes, and same disposable income but different wage rates and hence different work hours. 54.4% think hours of work is irrelevant and 42.7% think the hardworking low wage person is more deserving. Panel C considers transfer recipients receiving the same benefit levels. Subjects find the disabled person unable to work and the unemployed person looking for work much more deserving than the abled bodied unemployed or welfare recipient not looking for work.

ACTUAL SOCIAL PREFERENCES

General conclusion: People favor redistribution if they feel inequalities are “unfair” but unfair means different things to different people

⇒ Redistribution supported when people don't have control [education for children, health insurance for the sick, retirement/disability benefits for the elderly/disabled unable to work]

⇒ Less support when people have some or full control [unemployment, being low income]

Conservatives tend to frame things: individuals have control

Liberals tend to frame things: individuals do not have control

Conclusion: Two General Rules for Govt Intervention

1) Market Failures: Government intervention can help if there are market failures

2) Redistribution: Free market generates inequality. Govt taxes and spending can reduce inequality

First part of course will analyze 2), second part of course will analyze 1)

[we are inverting the ordering relative to Gruber's textbook so as to cover topics related to Professor Yagan's research first, with richer normative frameworks].

REFERENCES

Gruber, Jonathan, *Public Finance and Public Policy*, Fourth Edition, 2012
Worth Publishers, Chapter 2

Rawls, John, *A Theory of Justice*, 1971, revised in 1999, Cambridge:
Harvard University Press

Saez, Emmanuel and Stefanie Stantcheva “Generalized Social Marginal
Welfare Weights for Optimal Tax Theory,” NBER Working Paper No.
18835, February 2013. (web)

Undergraduate Public Economics

Danny Yagan
UC Berkeley

Lecture 3

Empirical Tools of Public Finance

DEFINITIONS

Empirical public finance: The use of data and statistical methods to measure the impact of government policy on individuals and markets (example: how an increase of taxes affects work behavior)

Correlation: Two economic variables are correlated if they move together (example: height and weight across individuals)

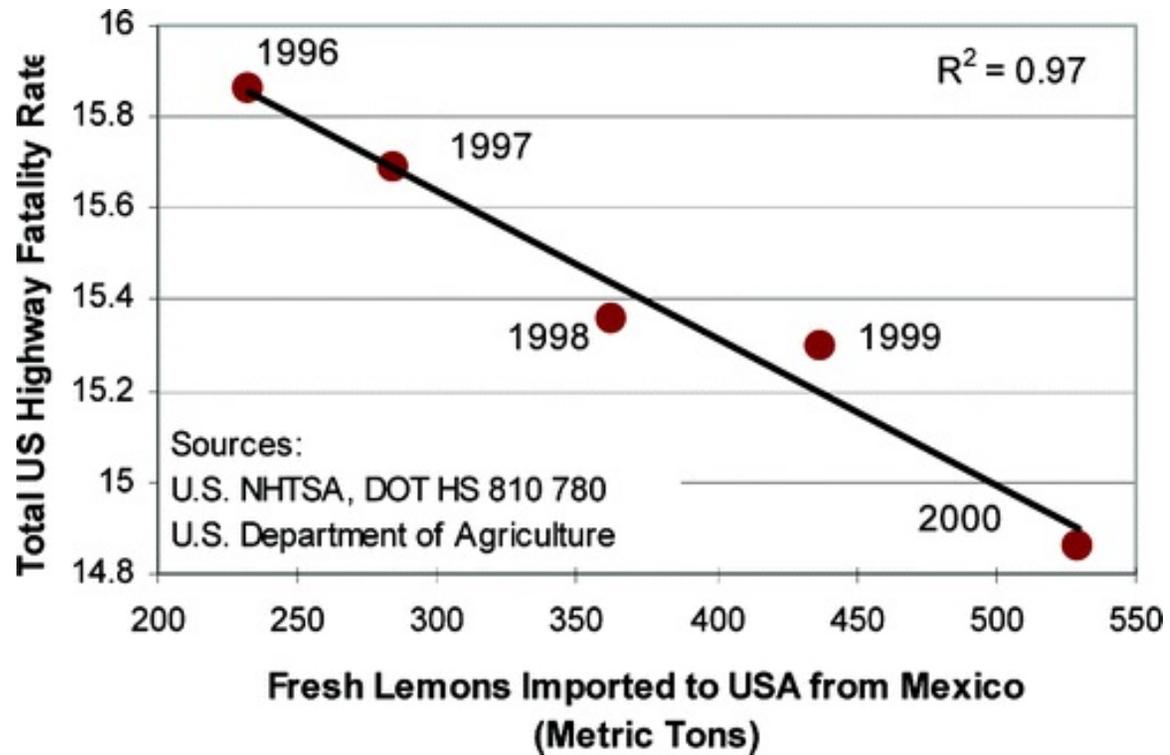
Causality: Two economic variables are causally related if the movement of one causes movement of the other (example: good nutrition as an infant increases adult height)

THE IMPORTANT DISTINCTION BETWEEN CORRELATION AND CAUSALITY

There are many examples where causation and correlation can get confused.

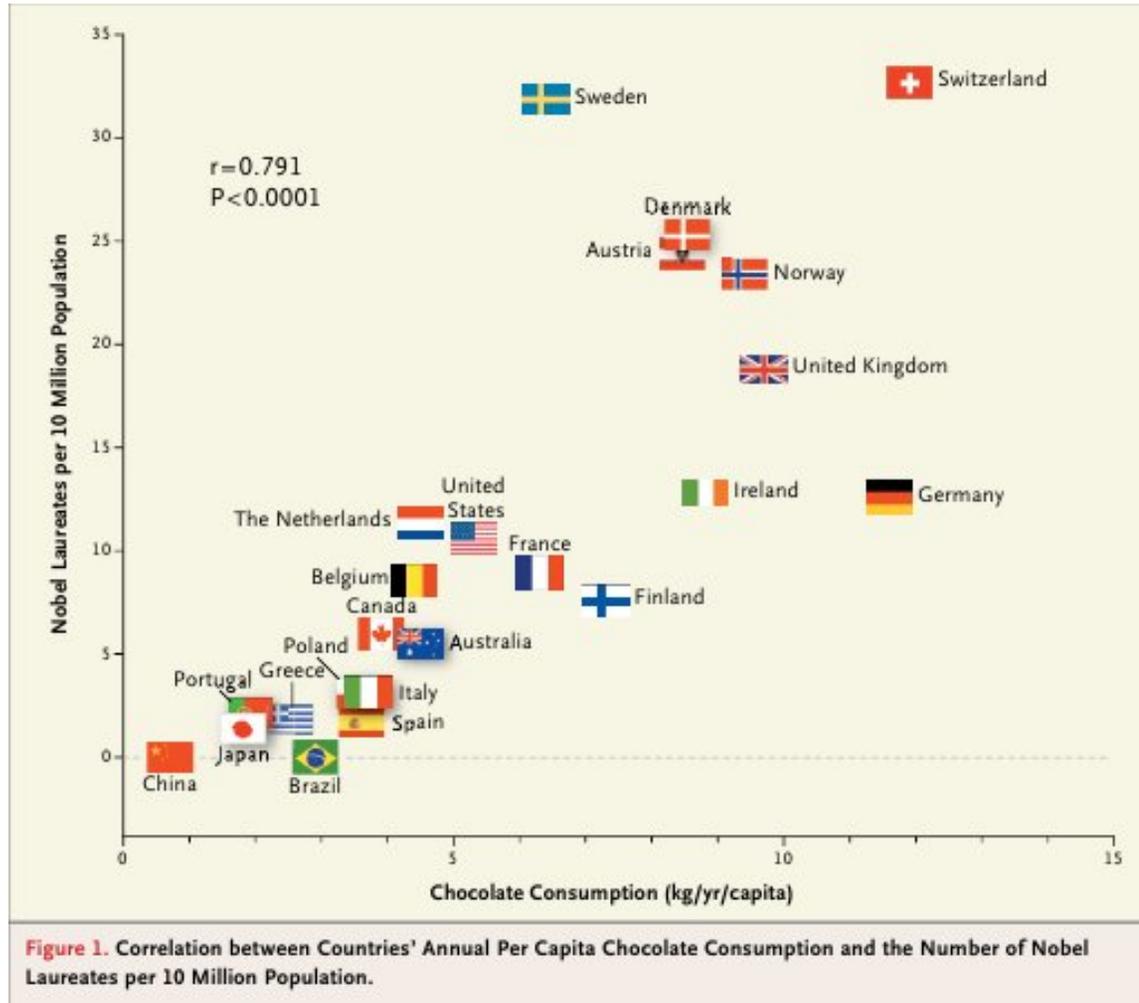
In statistics, this is called the *identification problem*: given that two series are correlated, how do you identify whether one series is causing another?

HAVE LEMON IMPORTS SAVED LIVES?



Source: Johnson (2008) *American Chemical Society*

DOES CHOCOLATE CAUSE NOBEL PRIZES?



Source: Messerli (2012) *New England Journal of Medicine*

THE IDENTIFICATION PROBLEM

The attempt to interpret a correlation as a causal relationship without sufficient thought to the underlying process generating the data is a common problem.

For any correlation between two variables A and B, there are three possible explanations, one or more of which could result in the correlation:

- 1) A is causing B
- 2) B is causing A
- 3) Some third factor is causing both

The general problem that empirical economists face in trying to use existing data to assess the causal influence of one factor on another is that one cannot immediately go from correlation to causation.

RANDOMIZED TRIALS AS A SOLUTION

Randomized trial: The ideal type of experiment designed to test causality, whereby a group of individuals is randomly divided into a treatment group, which receives the treatment of interest, and a control group, which does not.

Treatment group: The set of individuals who are subject to an intervention being studied.

Control group: The set of individuals comparable to the treatment group who are not subject to the intervention being studied.

Randomized trials have been used in medicine for many decades and have become very popular in economics, especially development economics in the last 15 years

THE PROBLEMS OF BIAS

Bias: Any source of difference between treatment and control groups that is correlated with the treatment but is not due to the treatment.

Having large sample sizes allows researchers to eliminate any consistent differences between groups by relying on the statistical principle called *the law of large numbers*: the odds of getting the wrong answer approaches zero as the sample size grows.

Statisticians develop methods to evaluate the precision of estimates and create confidence intervals around estimates

EXAMPLES OF RANDOMIZED TRIALS

Randomized Trials of ERT (estrogen replacement therapy)

The randomized trial of ERT tracked over 16,000 women ages 50–79 who were recruited to participate in the trial by 40 clinical centers in the United States. The study was supposed to last 8.5 years but was stopped after 5.2 years because its conclusion was already clear: ERT did in fact raise the risk of heart disease.

Randomized Trials in the Welfare Reform (TANF) Context

Randomized trials are equally useful in the context of public policy. A number of states implemented randomized trials to test various aspects of proposed welfare reform.

WHY WE NEED TO GO BEYOND RANDOMIZED TRIALS

Even the gold standard of randomized trials has some potential problems.

1) External validity: The results are only valid for the sample of individuals who volunteer to be either treatments or controls, and this sample may be different from the population at large (e.g., randomized experiment in Sweden or US would not necessarily generate the same results)

2) Attrition: Individuals may leave the experiment before it is complete. Reduction in the size of samples over time, which, if not random, can lead to bias estimates.

Outside randomized experiments, bias is a pervasive problem that is not easily remedied. There are, however, methods available that can allow us to approach the gold standard of randomized trials.

OBSERVATIONAL DATA

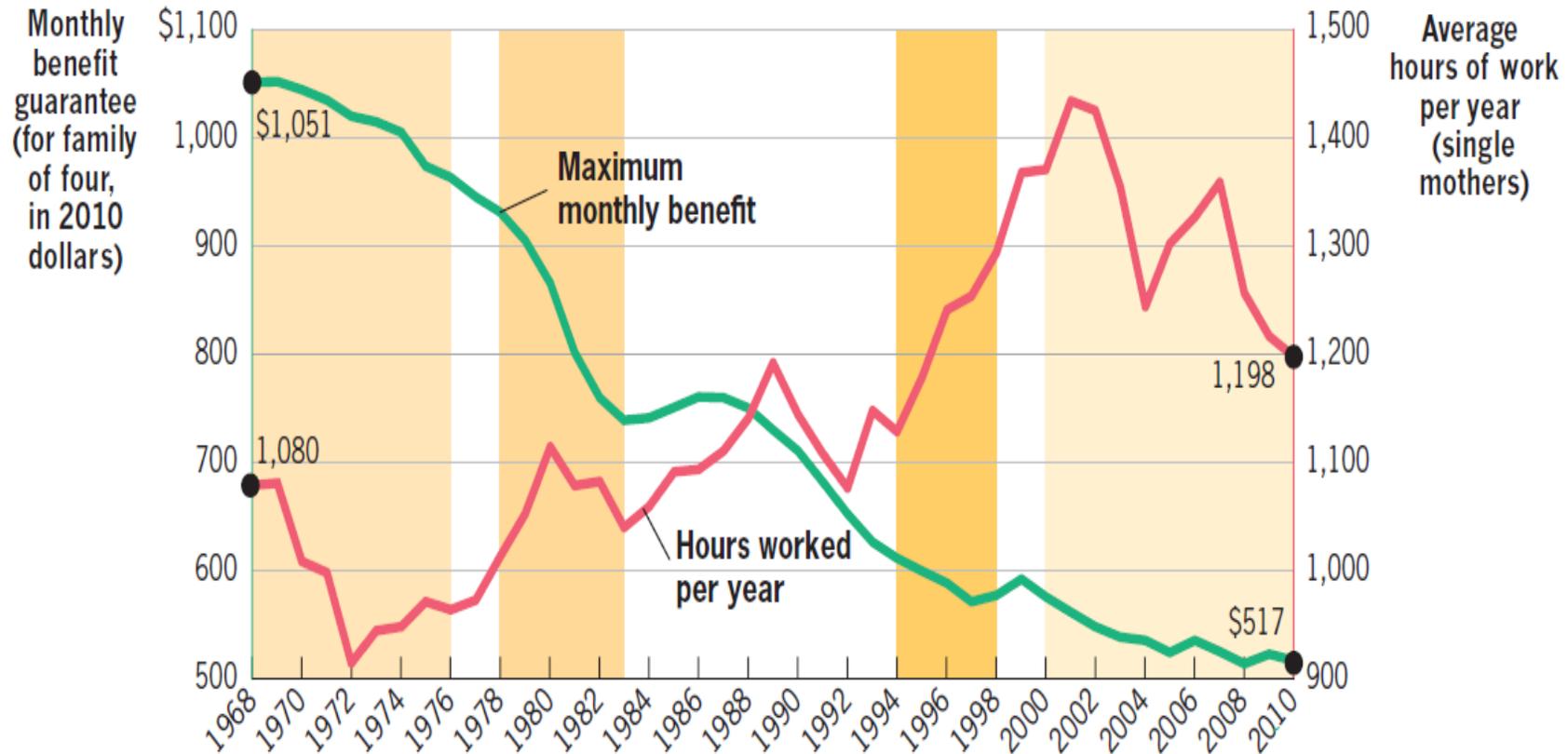
Observational data: Data generated by individual behavior observed in the real world, not in the context of deliberately designed experiments.

Time series analysis: Analysis of the co-movement of two series over time.

Cross-sectional regression analysis: Statistical analysis of the relationship between two or more variables exhibited by many individuals at one point in time.

3.3

Time Series Analysis: Cash Welfare Guarantee and Hours Worked Among Single Mothers



Time series analysis: Analysis of two series over time.

PROBLEMS WITH TIME SERIES ANALYSIS

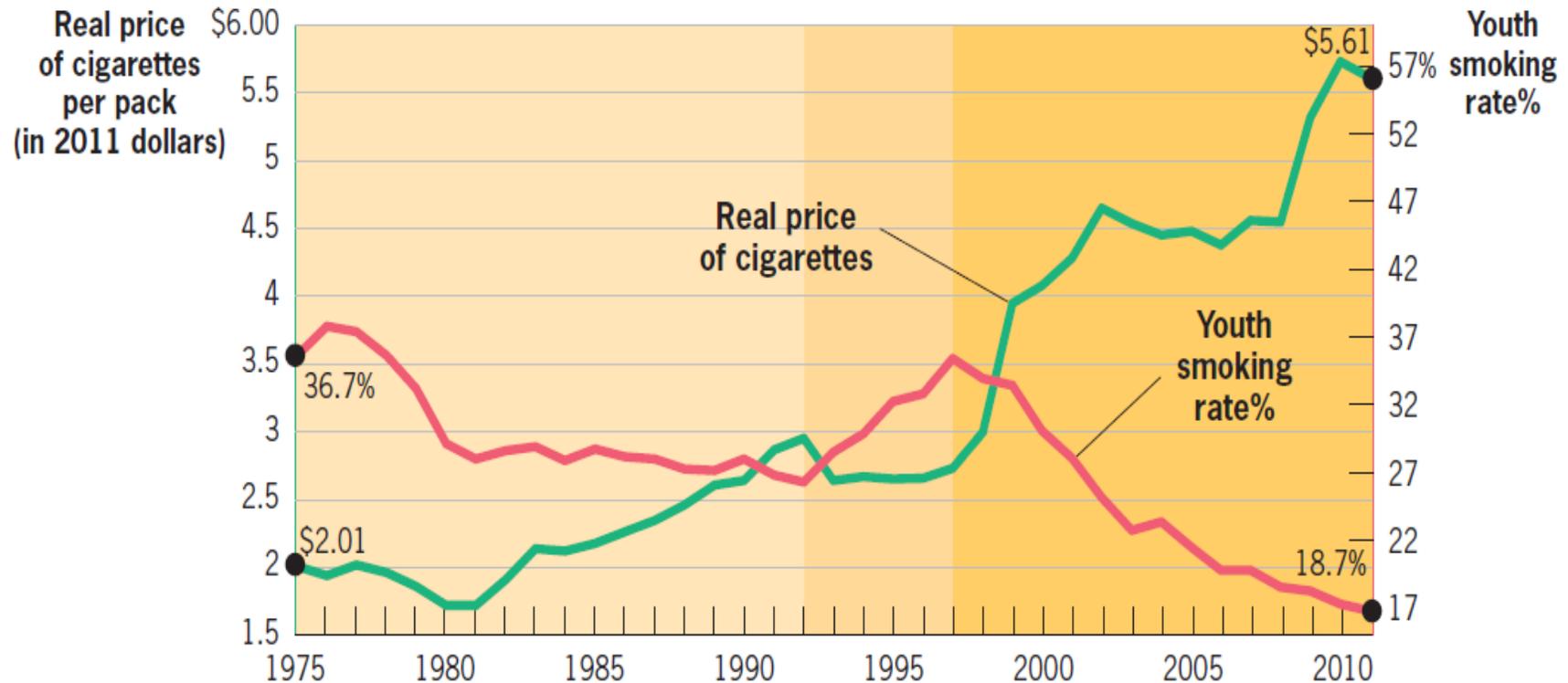
1) Although this time series correlation is striking, it does not necessarily demonstrate a causal effect of TANF benefits on labor supply

When there is a slow-moving trend in one variable through time, as is true for the general decline in income guarantees over this period, it is very difficult to infer its causal effects on another variable.

2) Other factors get in the way of a causal interpretation of this correlation over time; factors such as economic growth and a more generous Earned Income Tax Credit (EITC) can cause bias in this time series analysis because they are also correlated with the outcome of interest.

3.3

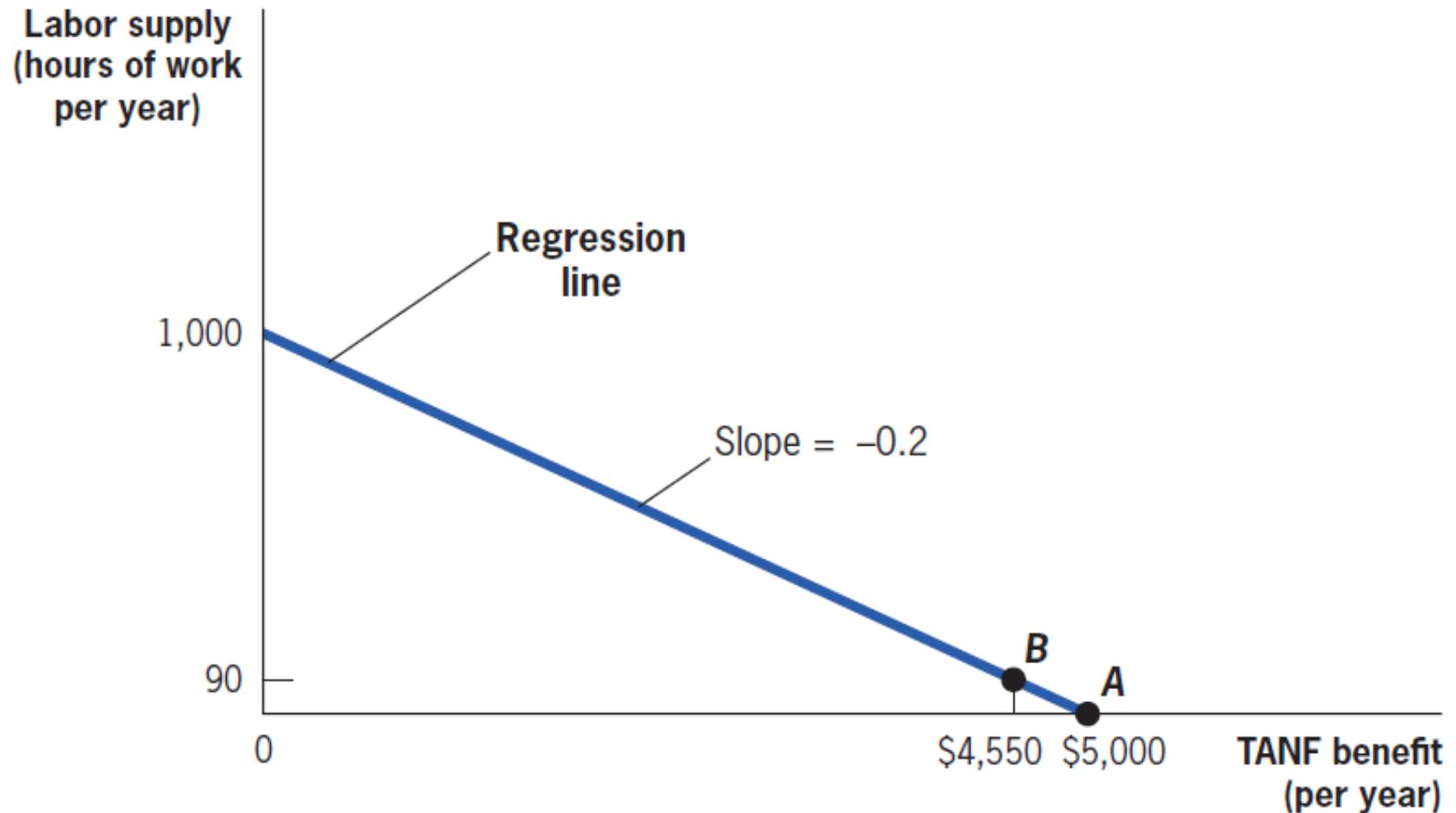
When Is Time Series Analysis Useful? Cigarette Prices and Youth Smoking



- Sharp, simultaneous changes in prices and smoking rates in 1993 and 1998–onward
- Known causes: price war, tobacco settlements

3.3

Cross-Sectional Regression Analysis: Labor Supply and TANF Benefit



REGRESSION

Regression line: The line that measures the best linear approximation to the relationship between any two variables.

$$Y = X\beta + \varepsilon$$

X is the independent variable data (TANF benefit guarantee)

Y is the dependent variable data (labor supply)

β is the coefficient that measures the effect of X on Y

ε is the error term (captures variations in Y not related to X).

Ordinary least square regression (OLS) estimates β without bias if ε is not correlated with X

REGRESSION ESTIMATES

The estimated coefficient $\hat{\beta}$ is reported with standard errors in parentheses

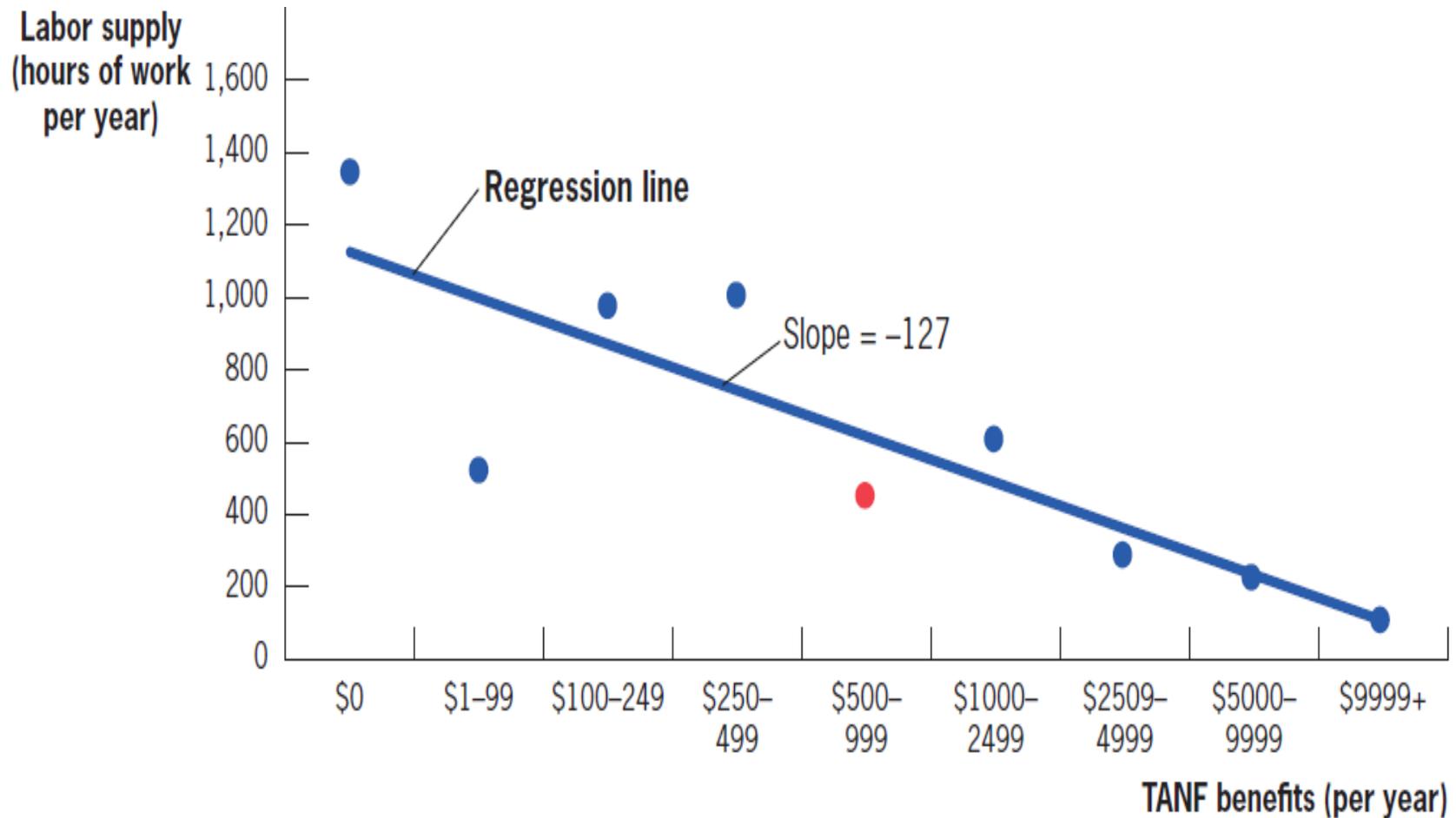
Example: $\hat{\beta} = .5(.1)$ should be understood as β is in confidence interval $(.5 - 2 \cdot .1, .5 + 2 \cdot .1) = (.3, .7)$ with probability 95%.

We have standard errors because we do not know the exact value of β

When estimated coefficient is more than twice the standard error, we can conclude that it is significantly positive (i.e., is above zero with probability 95%).

3.3

Example with Real-World Data: Labor Supply and TANF Benefits



PROBLEMS WITH CROSS-SECTIONAL REGRESSION ANALYSIS

The result summarized in Figure 3-4 seems to indicate strongly that mothers who receive the largest TANF benefits work the fewest hours. Once again, however, there are several possible interpretations of this correlation.

One interpretation is that higher TANF benefits are causing an increase in leisure.

Another possible interpretation is that in places with high TANF benefits, mothers have a high taste for leisure and wouldn't work much even if TANF benefits weren't available (this means exactly that ε is correlated with X)

CONTROL VARIABLES

It is essential in all empirical work to ensure that there are no factors that cause consistent differences in behavior across two groups (ε) and are also correlated with the independent variable X

Control variables: Additional variables Z that are included in cross-sectional regression models to account for differences between treatment and control groups that can lead to bias

$$Y = X\beta + Z\gamma + \varepsilon$$

In TANF case, Z would include race, education, number of children to control for demographic differences across states

Empirically, add Z variables and assess whether they change the estimate β . If estimate β varies a lot, we cannot be confident that identification assumption holds

QUASI-EXPERIMENTS: DEFINITION

Quasi-experiments (also called natural experiments)

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

Example: one state (Arkansas) decreases generosity of welfare benefits while another comparable state (Louisiana) does not. Single mothers in Arkansas are the Treatment (T) group, Single mothers in Louisiana are the control (C) group.

QUASI-EXPERIMENTS: ESTIMATION

We consider a Treatment group (T) and a Control group (C) and outcome Y

Simple difference estimator: $D = Y^{T,After} - Y^{C,After}$ is the difference in average outcomes between treatment and control after the change

In randomized experiment, simple difference $D = Y^{T,After} - Y^{C,After}$ is sufficient because T and C are identical before the treatment

In quasi-experiment, T and C might not be comparable before treatment. You can compute $D^{before} = Y^{T,Before} - Y^{C,Before}$

If $D^{before} = 0$, you can be fairly confident that $D = Y^{T,After} - Y^{C,After}$ estimates the causal effect

Difference-in-Difference estimator

If simple difference $D^{before} = Y^{T,Before} - Y^{C,Before}$ is not zero, you can form the **Difference-in-Difference estimator**

$$DD = [Y^{T,After} - Y^{C,After}] - [Y^{T,Before} - Y^{C,Before}]$$

This measures whether the difference between treatment and control changes after the policy change

Equivalent alternative: simple difference estimator for treatment, minus simple difference estimator for control:

$$DD = [Y^{T,After} - Y^{T,Before}] - [Y^{C,After} - Y^{C,Before}]$$

DD identifies the causal effect of the treatment if, absent the policy change, the difference between T and C would have stayed the same (this is called the “parallel trends” or “common trends” assumption)

3.3

Benefits and Labor Supply in Arkansas and Louisiana

Arkansas			
	1996	1998	Difference
Benefit guarantee (\$)	5,000	4,000	-1,000
Hours worked	1,000	1,200	200
Louisiana			
	1996	1998	Difference
Benefit guarantee (\$)	5,000	5,000	0
Hours worked	1,050	1,100	50

PROBLEMS WITH QUASI-EXPERIMENTS

With quasi-experimental studies, we can never be completely certain that we have purged all bias from the treatment–control comparison.

Quasi-experimental studies present various robustness checks to try to make the argument that they have obtained a causal estimate.

Examples: find alternative control groups, do a placebo comparing treatment and control DD when no policy change took place, etc.

Best way to check validity of DD estimator is to plot times series and assess whether a clear break between the two groups happens at the time of the reform

TWO GRAPHICAL EXAMPLES

1) Effects of lottery winnings on labor supply from Imbens, Rubin, Sacerdote AER'01

Ideal quasi-experiment to measure income effects as lottery generates random assignment conditional on playing \Rightarrow Very compelling graph, DD is convincing

2) Effects of the 1987 EITC expansion (tax credit for low income workers with kids) on labor supply from Eissa and Liebman QJE'96

Compares single mothers (Treatment) to single females with no kids (Control) \Rightarrow No compelling break in graph around 1987, DD is not convincing

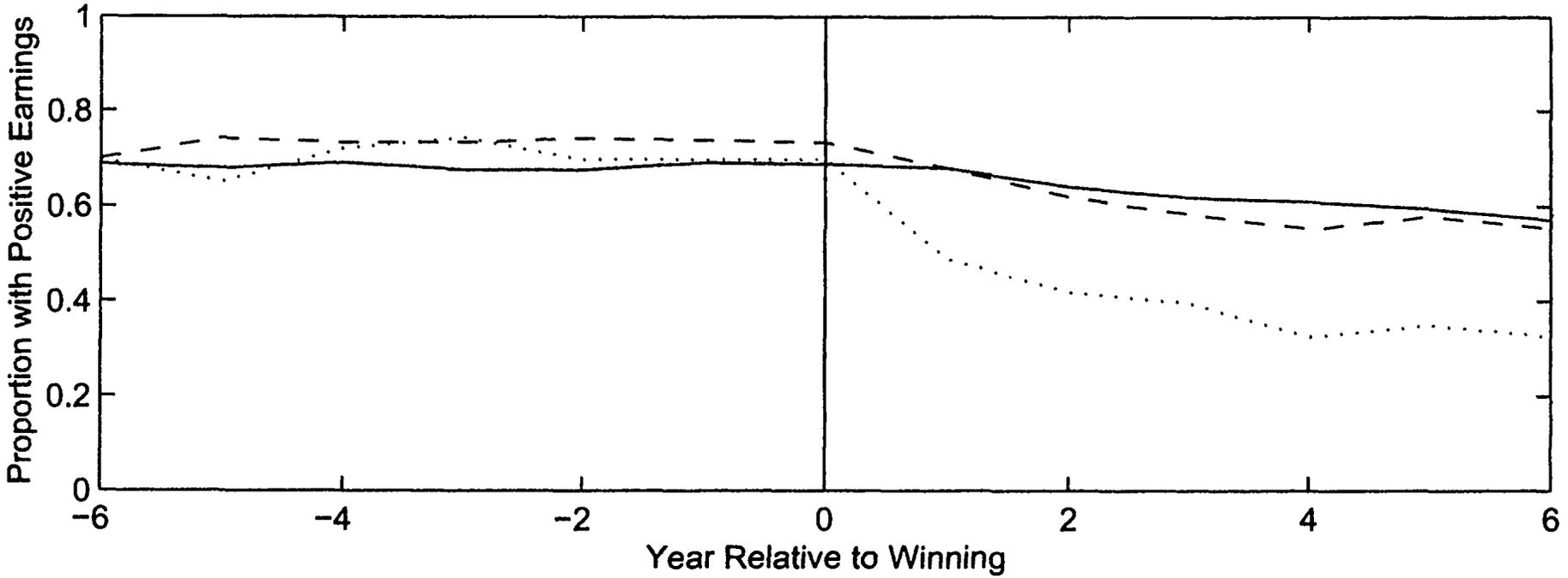
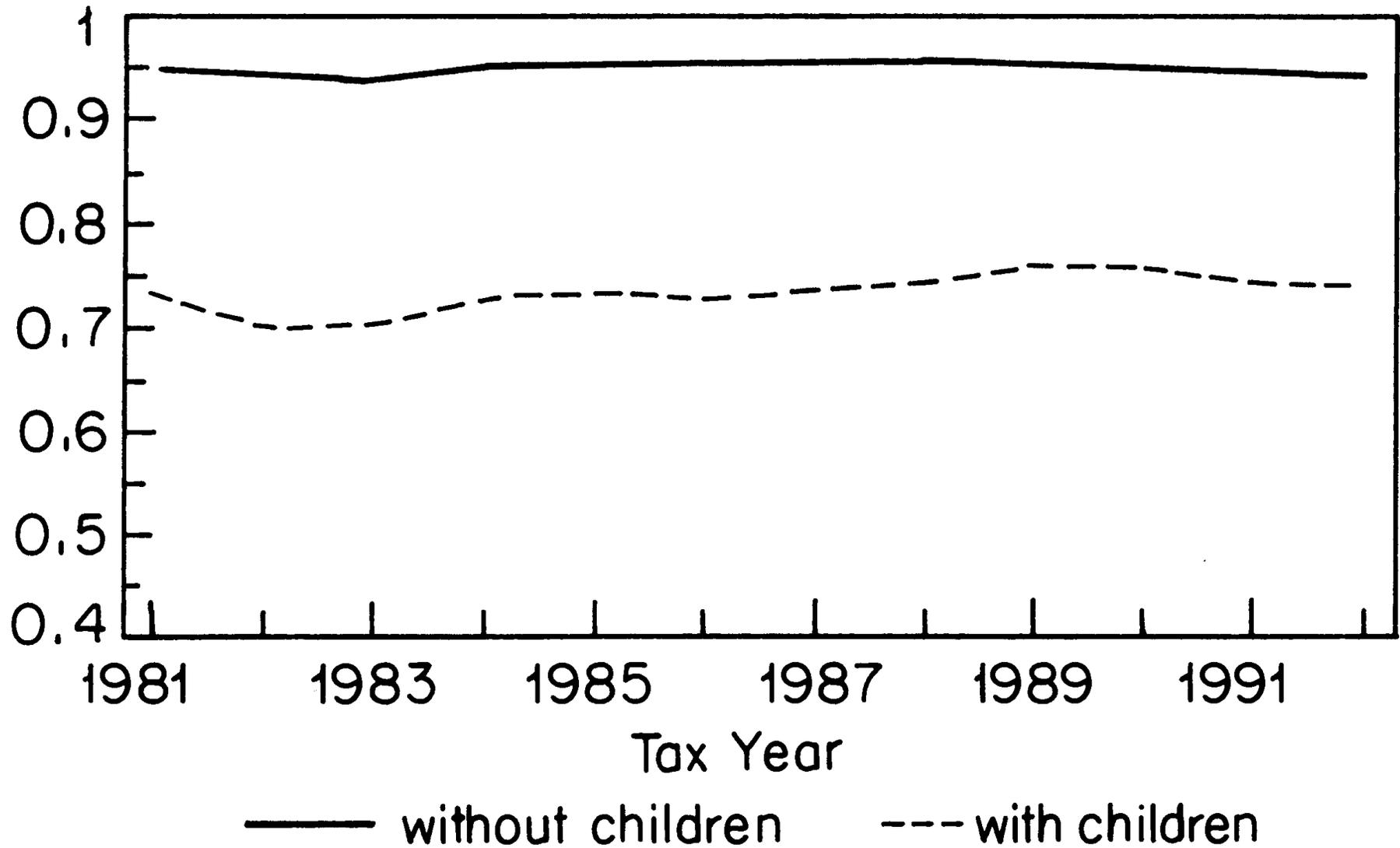


FIGURE 2. PROPORTION WITH POSITIVE EARNINGS FOR NONWINNERS, WINNERS, AND BIG WINNERS

Note: Solid line = nonwinners; dashed line = winners; dotted line = big winners.

Source: Imbens et al (2001), p. 784

All Unmarried Females



VALIDATING A DD ANALYSIS

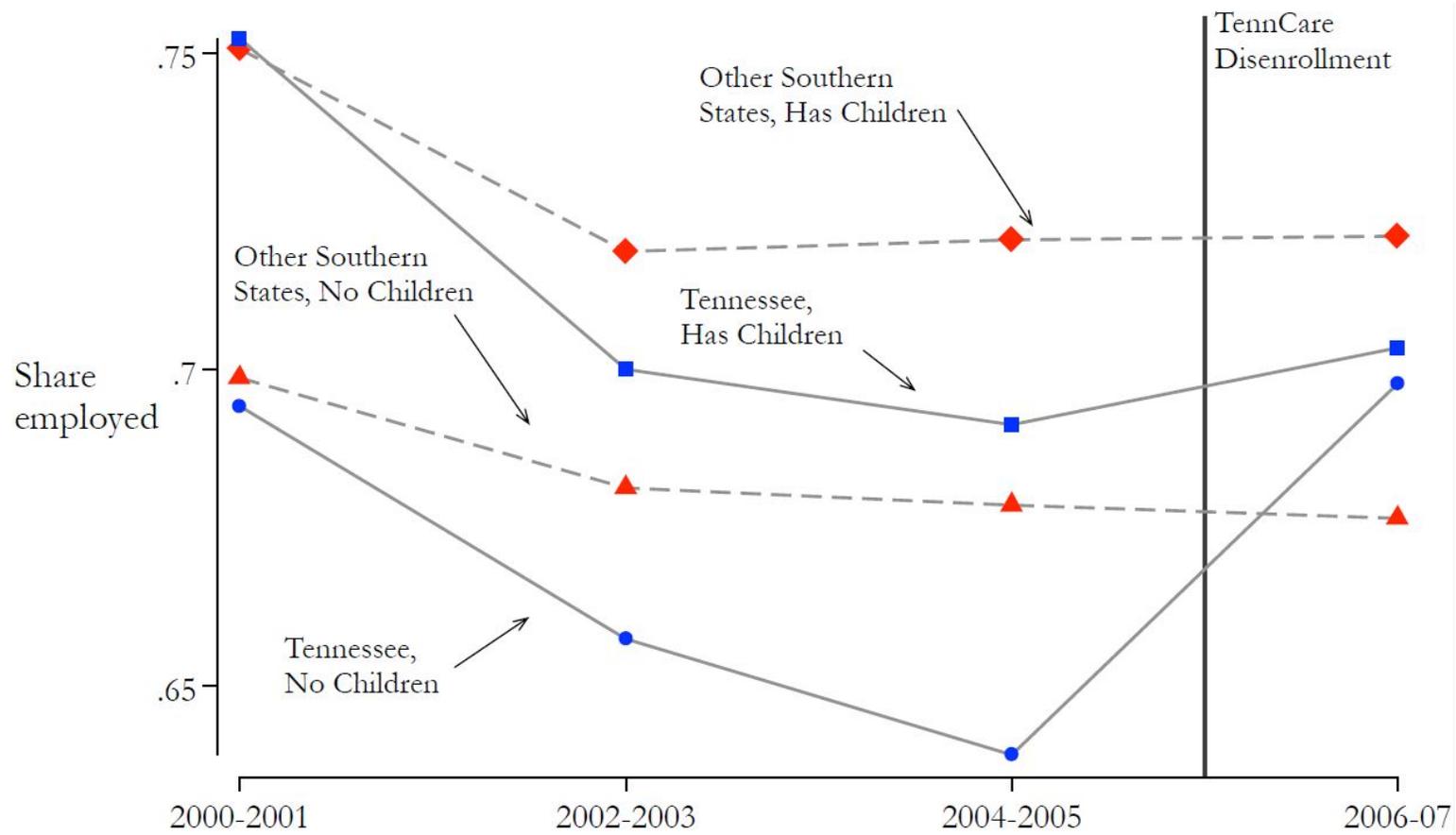
DD “common trends” identifying assumption: In the absence of the treatment, outcomes for the treatment group and control group would have trended similarly.

Typical check: Do outcomes for the treatment group and the control group trend similarly *before* the treatment?

Additional check called a “placebo test”: Are there similar unaffected groups that trend similarly before and after the treatment?

Example: Tennessee in 2005 abruptly removed 170,000 adults from Medicaid (free medical care for the poor), especially affecting poor childless adults. Did those childless adults enter the workforce in order to get health insurance from an employer instead?

PARALLEL TRENDS AND PLACEBO TEST



Source: Garthwaite-Gross-Notowidigdo (QJE 2014)

STRUCTURAL MODELING

Structural estimates: Builds a theoretical model of individual behavior and then estimates the parameters of the model. Estimates of the features that drive individual decisions, such as income and substitution effects or parameters of the utility function.

Reduced form estimates: Measures of the total impact of an independent variable on a dependent variable, without decomposing the source of that behavior response in terms of underlying parameters of the utility functions

Reduced form estimates are more transparent and convincing but structural estimates are more directly useful to make predictions for alternative policies

CONCLUSION

The central issue for any policy question is establishing a causal relationship between the policy in question and the outcome of interest.

We discussed several approaches to distinguish causality from correlation. The gold standard for doing so is the randomized trial, which removes bias through randomly assigning treatment and control groups.

Unfortunately, however, such trials are not available for every question we wish to address in empirical public finance. As a result, we turn to alternative methods such as time series analysis, cross-sectional regression analysis, and quasi-experimental analysis.

Each of these alternatives has weaknesses, but careful consideration of the problem at hand can often lead to a sensible solution to the bias problem that plagues empirical analysis.

REFERENCES

Eissa, Nada, and Jeffrey B. Liebman. "Labor supply response to the earned income tax credit." *The Quarterly Journal of Economics* 111.2 (1996): 605-637.(web)

Garthwaite, Craig, Tal Gross, and Matthew Notowidigdo. "Public Health Insurance, Labor Supply, and Employment Lock" *Quarterly Journal of Economics* 129.2 (2014): 653-696.(web)

Hotz, V. Joseph, Charles H. Mullin, and John Karl Scholz. "Welfare, employment, and income: evidence on the effects of benefit reductions from California." *The American Economic Review* 92.2 (2002): 380-384.(web)

Imbens, Guido W., Donald B. Rubin, and Bruce I. Sacerdote. "Estimating the effect of unearned income on labor earnings, savings, and consumption: Evidence from a survey of lottery players." *American Economic Review* (2001): 778-794.(web)

Jonathan Gruber, *Public Finance and Public Policy*, Fourth Edition, 2012
Worth Publishers, Chapter 3

Marquis, Grace S., et al. "Association of breastfeeding and stunting in Peruvian toddlers: an example of reverse causality." *International Journal of Epidemiology* 26.2 (1997): 349-356.(web)

Undergraduate Public Economics

Danny Yagan
UC Berkeley

Lecture 4

Tools of Budget Analysis

GOVERNMENT BUDGETING

Debt: The amount borrowed by government through bonds to individuals, firms, or foreign governments. Debt is a **stock**

Deficit: government's spending + interest payments on debt minus government revenues in a given year. A negative deficit is called a surplus. Deficit is a **flow**

Evolution of debt from year to year:

$$Debt_{t+1} = Debt_t + Deficit_t = Debt_t \cdot (1 + r_t) + Spending_t - Revenue_t$$

with r_t interest paid on government debt

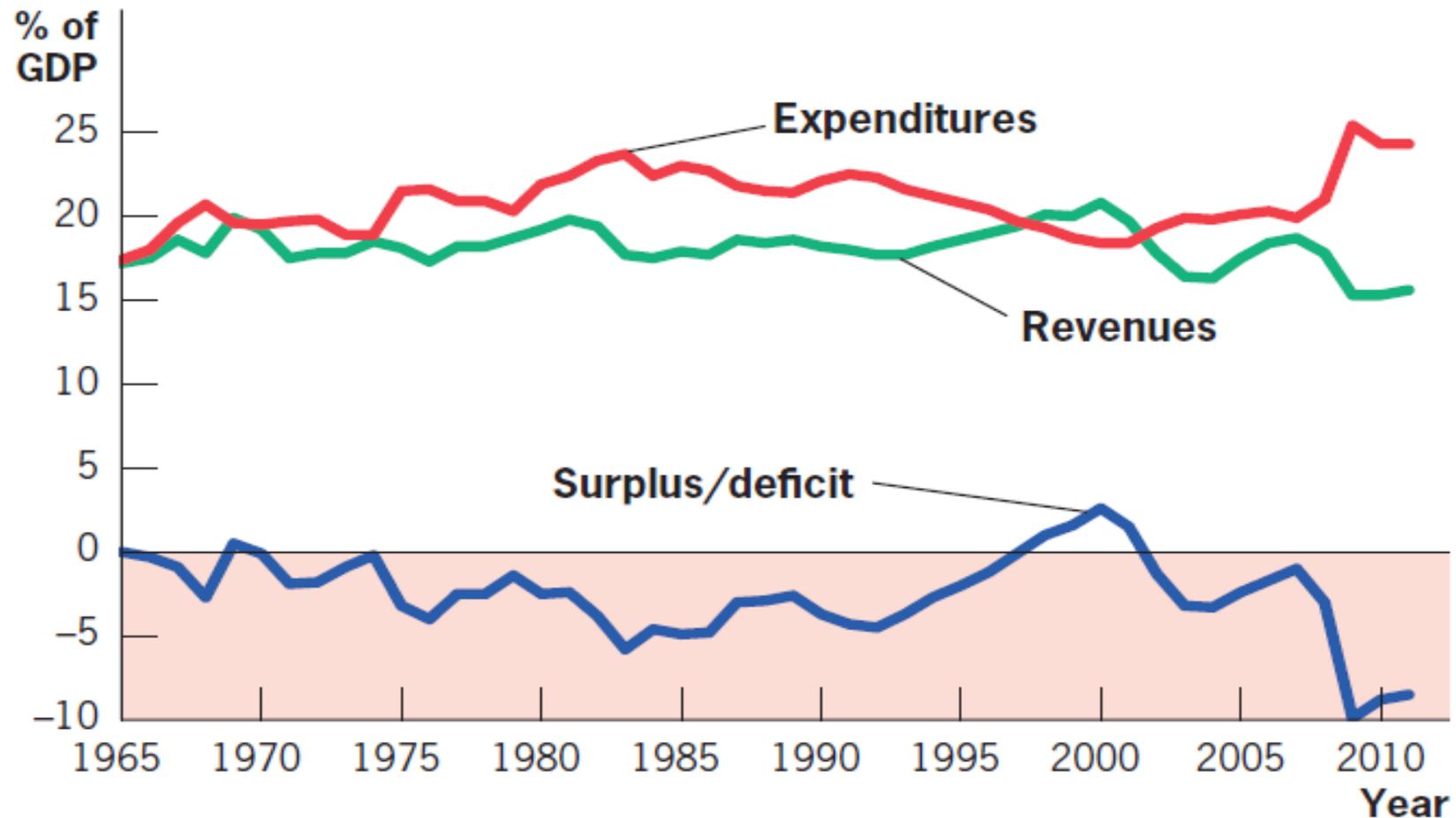
$$\text{Primary Deficit} = \text{Spending} - \text{Revenue}$$

In 2014: US Federal debt is around 75% of GDP, US deficit is 2.8% of GDP

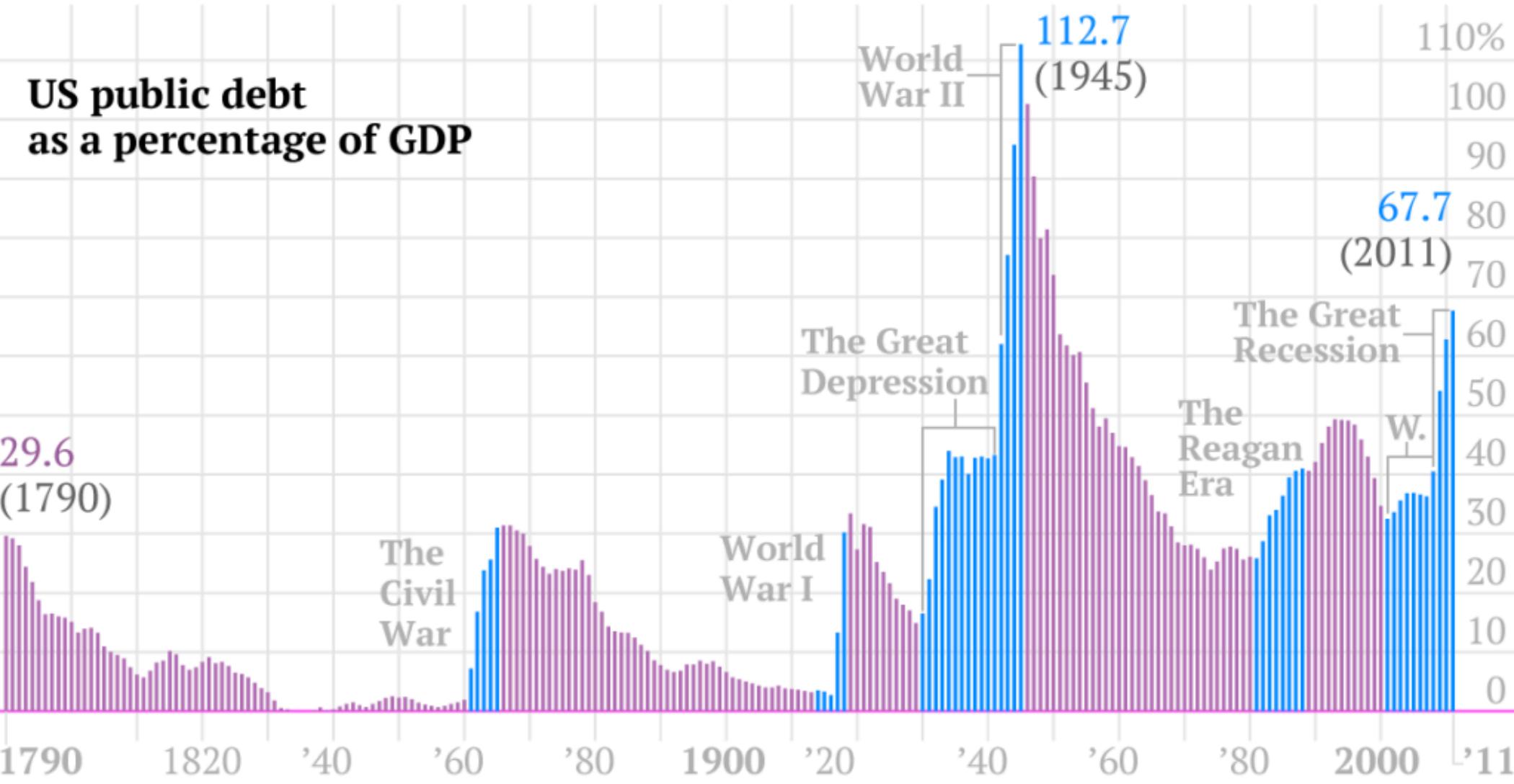
US government owns assets worth about 100% of GDP

4.1

The Budget Deficit in Recent Years



US public debt as a percentage of GDP



Ritchie King | Quartz

Data: Congressional Budget Office

Source: Congressional Budget Office

THE US FEDERAL PROCESS

Taxes, spending, and debt ceiling are decided by Congress and the President

New law requires majority vote both in House and in Senate along with President's signature (veto power)

In recent years, Senate vote requires 60/100 super-majority (due to filibuster)

Two forms of spending:

Entitlement spending: Mandatory funds for programs for which funding levels are automatically set by the number of eligible recipients, not the discretion of Congress (ex: medicare, social security)

Discretionary spending: Optional spending set by appropriation levels each year, at Congress's discretion (ex: defense)

Budget Policies and Deficits at the State Level

Balanced budget requirement (BBR): Law forcing a government to balance its budget each year (spending = revenue).

ex-post BBR: government needs to balance its budget by the end of each fiscal year

ex-ante BBR: government needs to submit/pass a balanced budget at the start of each fiscal year, or both (easier to evade with rosy predictions)

California has ex-ante BBR: recession lowered tax revenue and forced cuts in government spending (plans to have rainy fund)

STATIC VS. DYNAMIC SCORING

Govts have agencies evaluating effects of proposed reforms on govt deficit (Congressional Budget Office in the US)

Static scoring: A method used by budget modelers that assumes that government policy changes only the distribution of total resources, not the amount of total resources.

Dynamic scoring: A method used by budget modelers that attempts to model the effect of government policy on both the distribution of total resources and the amount of total resources.

Example: tax increases on the rich, static scoring assumes no effect on GDP, dynamic scoring incorporates effects on growth

Static scoring is safest in the absence of good empirical estimates of growth effects (dynamic scoring can be manipulated by ideologues).

Intertemporal Government Budget Constraint

Policy debates have traditionally focused on the extent to which this year's governmental spending exceeds this year's governmental revenues.

The existence of implicit obligations in the future, however, suggests that this does not capture the full picture

Intertemporal budget constraint: An equation relating the present discounted value of the government's obligations to the present discounted value of its revenues.

PDV of Tax Payments =
 PDV of All Future Govt Consumption + Current Govt Debt

BACKGROUND: PRESENT DISCOUNTED VALUE

For govt, spending F now has the same cost as spending $F \cdot (1 + r)$ next year with r interest rate on government debt

Present discounted value (PDV): The value of each period's dollar amount in today's terms.

Govt spends F_1, F_2, F_3, \dots in each future year, then the PDV is computed as:

$$PDV = \frac{F_1}{(1+r)} + \frac{F_2}{(1+r)^2} + \frac{F_3}{(1+r)^3} + \dots$$

If $F_1 = F_2 = \dots = F$ then

$$PDV = \frac{F}{1+r} \cdot \left[1 + \frac{1}{(1+r)} + \frac{1}{(1+r)^2} + \dots \right] = \frac{F}{1+r} \cdot \frac{1}{1 - \frac{1}{1+r}} = \frac{F}{r}$$

Paying F in perpetuity is equivalent to paying F/r upfront

LOGIC OF GOVERNMENT ALWAYS HAVING SOME DEBT

Student debt lets you borrow from your future self: If you're going to earn more as an adult than as a college student, it makes sense to borrow as a student and pay off your debt as an adult.

Government debt lets one generation borrow from future generations: future generations will be much richer than us, and government debt lets us borrow from them.

Is this sensible or is it "stealing from our children"? Matter of perspective.

ALTERNATIVE MEASURES OF LONG-RUN GOVERNMENT BUDGETS

Long-run Fiscal Imbalance

If the government continues with today's policies, how much more will the government spend than it will collect in taxes over the entire future?

Example: In 2003 alone, the government added roughly \$20 trillion to the fiscal imbalance (due to tax cuts and medicare prescription drug benefit of Bush administration)

PROBLEMS WITH LONG-RUN FISCAL MEASURES

The fiscal imbalance calculations are fairly tenuous:

1) They depend critically on many assumptions about future growth rates in costs and incomes, and the interest rate used for discounting

⇒ Those assumptions become heroic for long-distance future (example: how will health care costs evolve?)

2) The calculations also assume that government policy remains unchanged (but if big imbalance arises, then government will typically be forced to act and fix it)

⇒ Makes most sense to consider a time window that is longer than 1 year but less than infinity

PROBLEMS WITH LONG-RUN FISCAL MEASURES

Some programs are easier to project than others.

Example: social security retirement benefits are easier to project than medicare benefits

Social security benefits depend on demography and longevity (slow moving variables) \Rightarrow Social security does fairly reliable 75 year projections

Medicare depends on growth of health care costs that have been growing very fast (before the Great recession) \Rightarrow such a rate of growth is not sustainable for ever so making a long-run projection based on those rates is not meaningful

CBO makes budget projections over the next 10 years in its official budget projection

Short-Run Effects of the Govt on the Macroeconomy

Keynesian theory (IS-LM macro model): More government spending or tax cuts stimulates the economy in the short-run [and conversely]

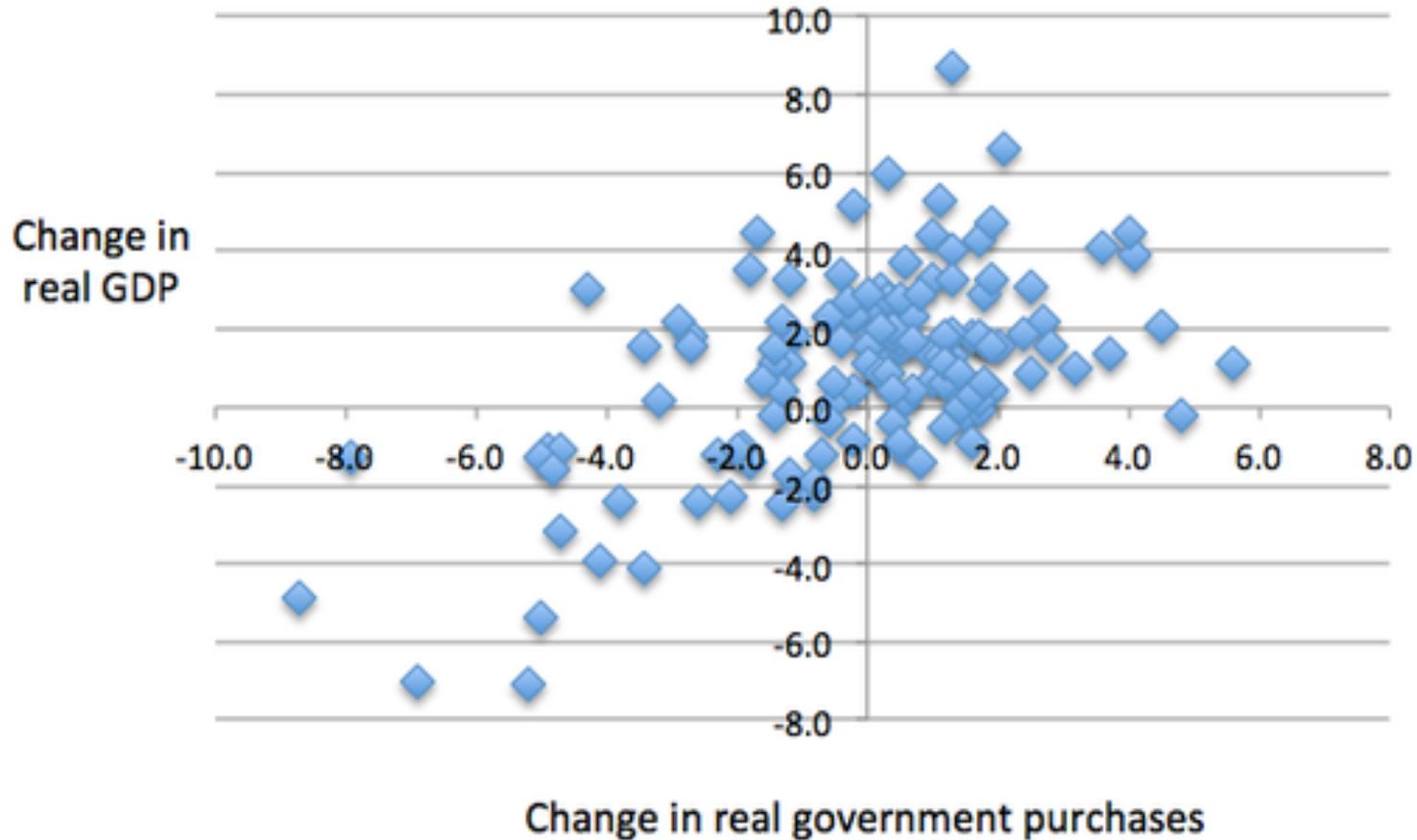
Short-run stabilization: Govt can use taxes and spending policies to smooth the peaks and troughs of the business cycle

Automatic stabilization: Policies that automatically alter taxes or spending in response to economic fluctuations to offset changes in household consumption levels (ex: unemployment insurance, progressive taxation, corporate profits tax)

Discretionary stabilization: Policy actions taken by the government in response to business cycle (ex: Fiscal stimulus with Spring 2008 rebate checks, 2009-12 Obama stimulus, unemployment insurance extensions)

⇒ Ability to run deficits in recessions is a great tool for short-run business cycle stabilization (but need to reduce debt during good times to keep ability to run deficits when needed)

Government spending and growth, 2010-2013



% changes in annual real govt spending and changes in real GDP, 33 EU countries, 2010-11, 2011-2, 2012-3 (=99 dots). Source: Krugman NYtimes blog, January 6, 2015

LOGIC OF GOVERNMENT DEBT RISING AFTER RECESSIONS

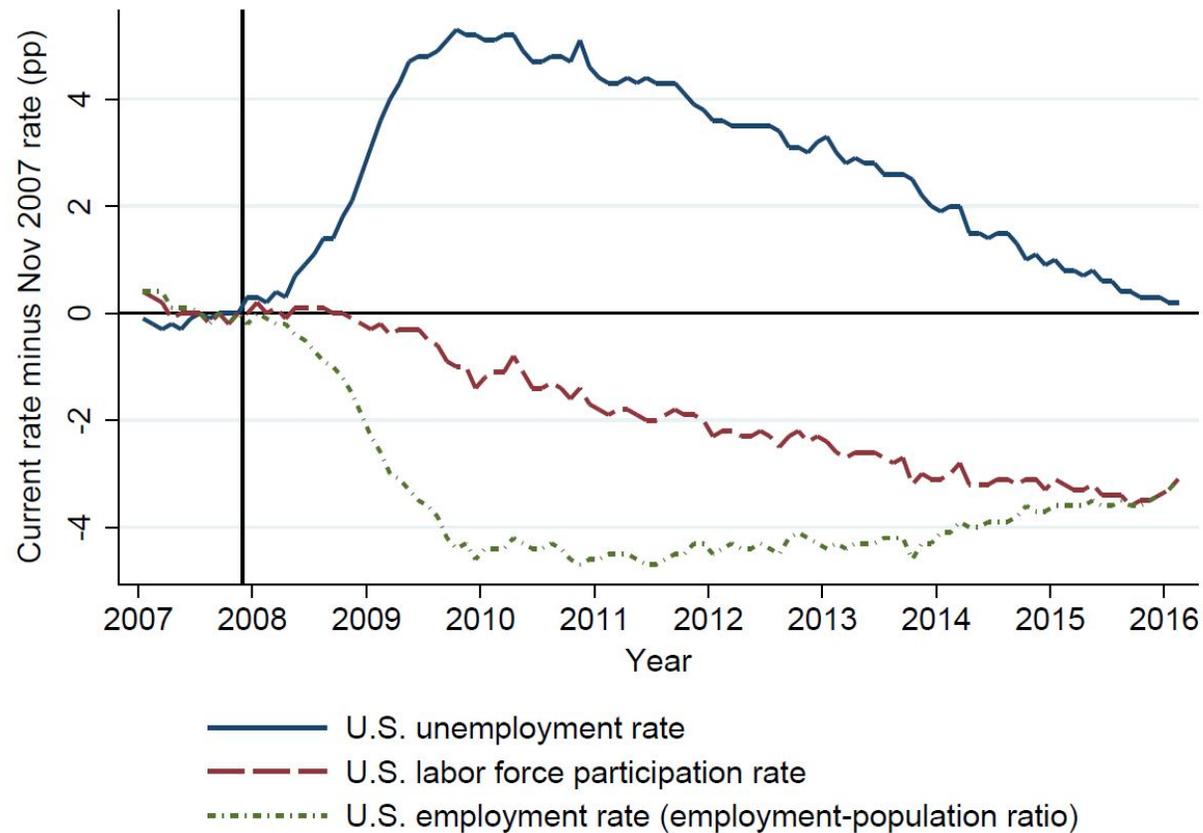
Mechanical: Less economic activity causes less tax revenue which causes higher deficits (unless gov *cuts* spending).

Keynesian: Government steps in when markets fail and spends extra, in order to keep output at capacity. Especially important if recessions causes long-term damage.

Classical: Some of the stuff government buys is more valuable or can be cheaper in recessions. Example #1: Unemployment insurance. Example #2: Building new bridges, airports, and schools when construction unemployment is high and government debt is cheap (interest rates fall during recessions).

IS THE GREAT RECESSION REALLY OVER?

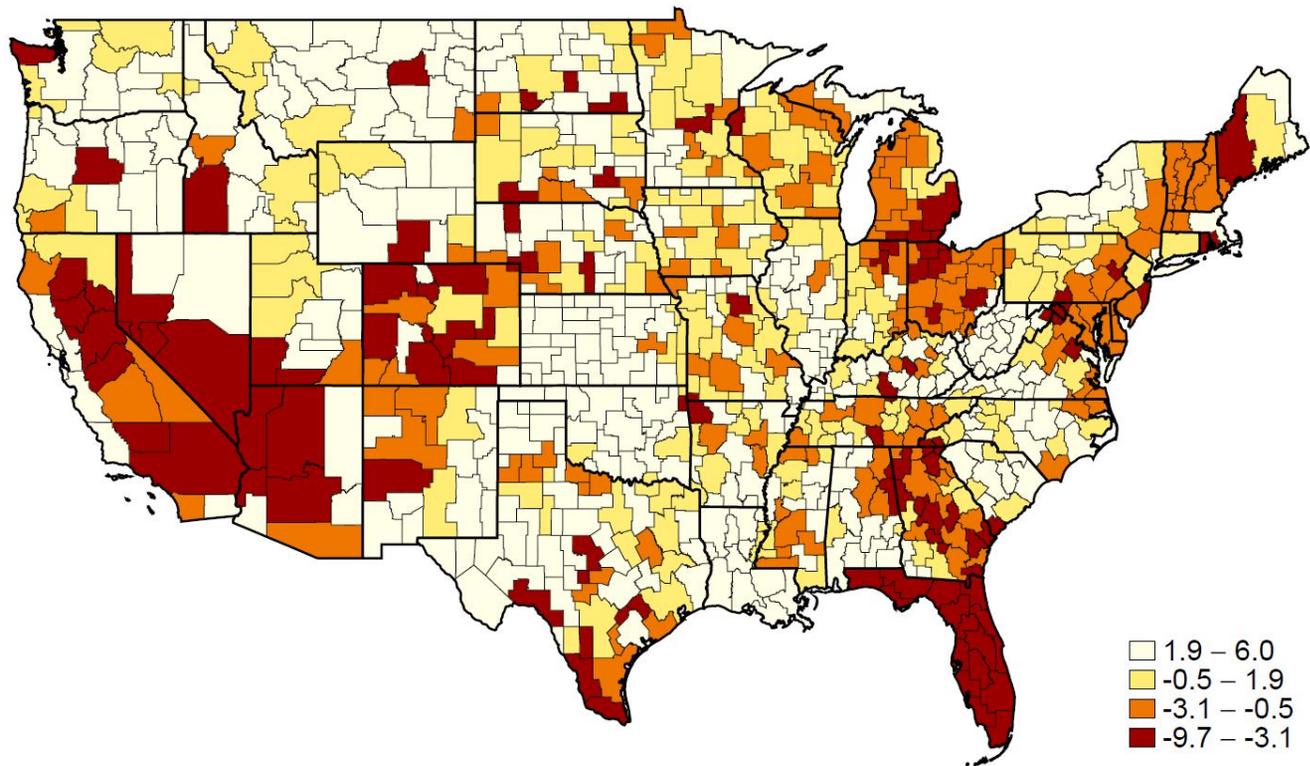
A. Current U.S. Aggregate Minus November 2007 U.S. Aggregate



Source: Yagan (2016)

LOCAL VARIATION IN GREAT RECESSION INTENSITY

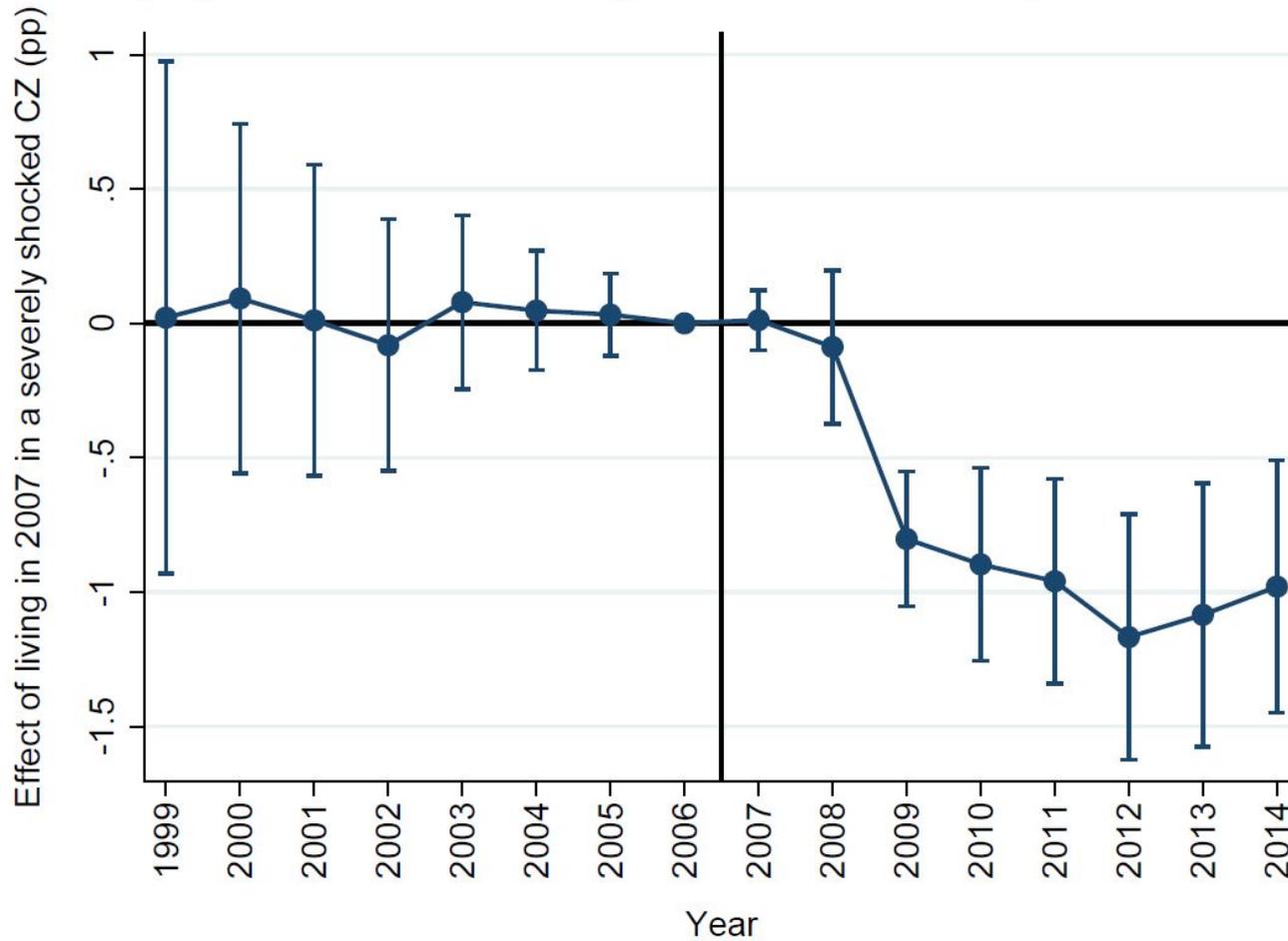
Figure 2: 2007-2009 CZ Shocks



Source: Yagan (2016)

ENDURING IMPACT

A. Employment Effect of Living in 2007 in a Severely Shocked CZ



Source: Yagan (2016)

LONG-RUN EFFECTS OF GOVERNMENT DEBT

In the long-run, government debt affects the capital market where savers meet investors

savings = investment + new govt debt

With more government debt, if savings do not change, less funds available for investment: investment decreases

Two mitigating factors:

1) In an open economy, investment or govt debt can be funded with foreign savings

2) If individuals are forward looking, they understand that higher debt implies high taxes later on and hence they save more to be able to pay higher taxes later on [Ricardian equivalence]

Background: Savings and Economic Growth

The earliest economic growth models emphasized a central role for savings as an engine of growth, and this insight remains important for growth economics today.

More Capital, More Growth: As there is more capital in an economy, each worker is more productive, and total social product rises. A larger capital stock means more total output for any level of labor supply. Thus, the size of the capital stock might be a primary driver of growth.

Neo-classical aggregate production function:

K capital stock, L labor, A technology

$$Y = F(K, L) = A \cdot K^\alpha L^{1-\alpha} \quad \text{with} \quad \alpha \simeq 30$$

HIGH CAPITAL AND HIGH WAGES

Table 1: McDonalds Cashier or Crew Wages and Big Mac Prices, December 1998

<i>Country</i>	Estimated hourly wage rate	Reported Big Mac price	Exchange Rate per \$1	\$ hourly wage rate	\$ Big Mac price	Economist \$ Big Mac 3/99**	Big Macs per hour of work
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Russia	8.00	25.5	19.57	0.41	1.30		0.31
Korea	1700	3000	1210	1.41	2.48		0.57
Brazil	2.87	4.45	1.73*	1.66	2.57	1.71	0.65
Poland	4.12	5.3	3.50	1.18	1.51	1.38	0.78
Czech Rep.	45.00	53	30.30	1.49	1.75		0.85
UK	3.60		0.62*	5.80	3.07	3.07	1.89
USA	6.00		1.00	6.00	2.43	2.43	2.12
Germany	11.28	4.95	1.67	6.76	2.97	2.72	2.28
France	40.22	17.5	5.76	6.99	3.04	2.87	2.30
Italy	10417	4500	1646	6.33	2.73	2.5	2.31
Belgium	280.00	114	34.50	8.12	3.30		2.46
Sweden	64.90	25	8.03	8.09	3.11	2.88	2.60
Japan	844	280	120*	7.03	2.33	2.44	3.01

Savings and Investment market

Interest rate: The rate of return in year $t + 1$ on investments made in year t .

Save S in period t , you get $(1 + r) \cdot S$ in period $t + 1$

Supply and Demand for capital:

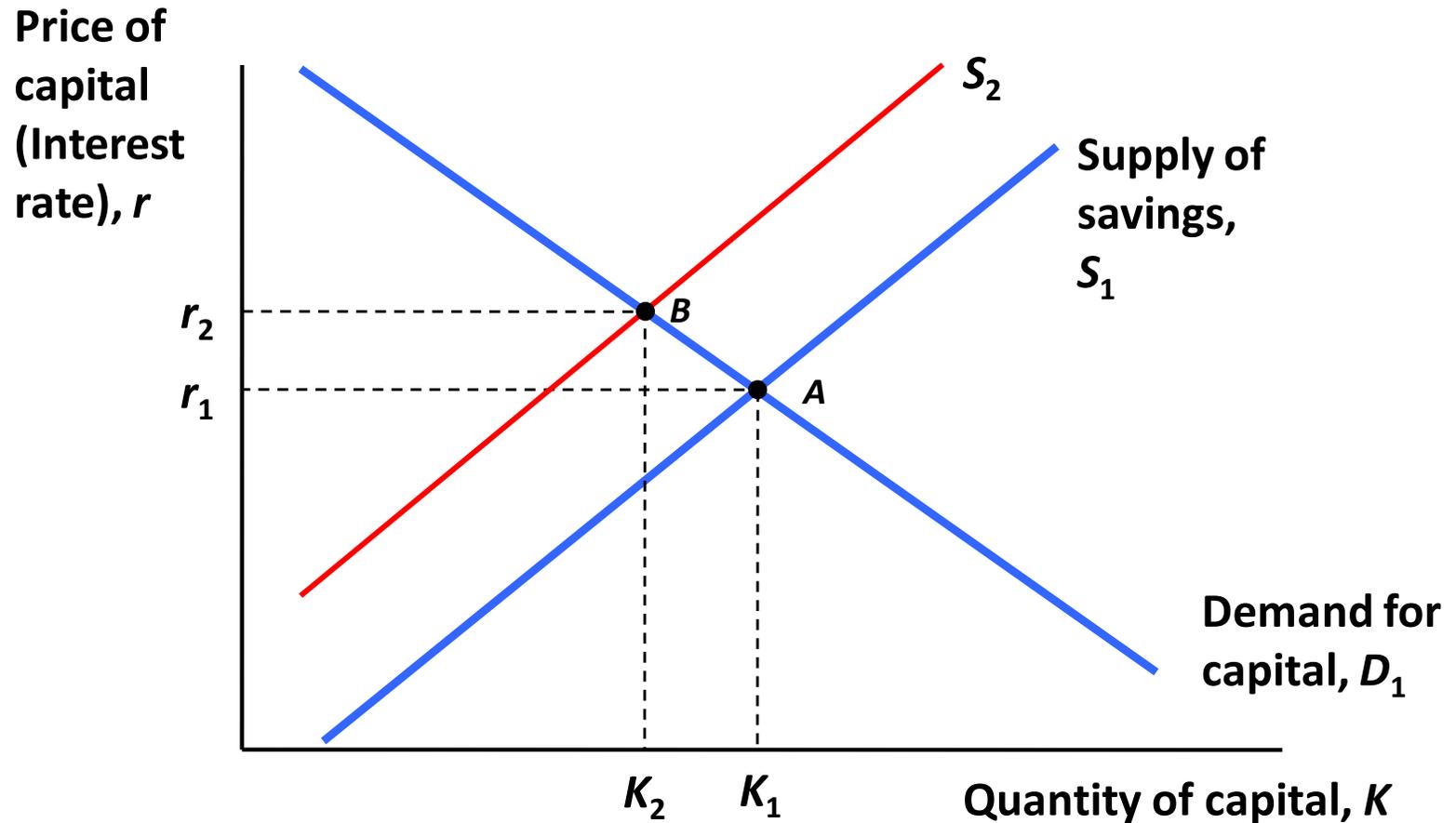
Supply of savings (from households) depends positively on r (higher r means bigger returns to savings)

Demand for investment (from firms) depends negatively on r (firms invest only if return on investment is at least equal to r)

In a competitive market, the equilibrium amount of investment is determined by the intersection of these demand and supply curves: $S(r) = D(r)$

4.4

Capital Market Equilibrium



Effects of Government Debt on Capital Market

If there is a deficit d , the government must borrow to finance the difference between its revenues and its expenditures: individual savings need to cover both d and firms' demand for capital

Equilibrium: $S(r) = d + D(r)$ or $S(r) - d = D(r) \Rightarrow r$ increases and K decreases

The government's borrowing may *crowd out* the borrowing of the private sector and lead to a lower level of capital accumulation

In reality, there are a number of complications of how government financing affects interest rates and growth

INTERNATIONAL CAPITAL MARKETS

With international capital markets, there is a worldwide interest rate r (i.e., supply of savings is perfectly elastic for a small open economy and hence horizontal)

⇒ Government debt has no impact on r and K

There is a large body of economics literature that has investigated the integration of international capital markets

It has generally concluded that while integration is present (and perhaps growing), it is far from perfect.

As a result, the supply of capital to the United States may not be perfectly elastic, and government deficits could crowd out private savings

RICARDIAN EQUIVALENCE

Standard supply and demand savings model assumed that supply of savings $S(r)$ was not affected by government deficit d

In reality, if individuals are rational, they recognize that more debt now means higher taxes (or less spending) in the future

If individuals are fully rational in their savings decisions, individuals understand that \$1 tax reduction today (and \$1 of debt increase)

⇒ \$1 of increases in PDV of taxes so the inter-temporal budget of the individual is unchanged

⇒ Individuals save the extra \$1 of tax reduction and will use it to pay future taxes

⇒ Govt debt increases the supply of individual savings one-for-one so govt debt does not affect r and capital

RICARDIAN EQUIVALENCE

What if increased taxes are on future generations? Will the elderly respond to the tax cut?

Barro (1974) showed that if generations are linked through altruism: parents care about kids and can leave them bequests, then the dynasty behaves as a single individual

\$1 tax cut on elderly leads them to save it and increase their bequests by \$1 so that their kids (or grand kids) can pay the corresponding tax increase down the road (consumption is unaffected)

This model has received only modest empirical support in the economics literature

Example: Individuals are not that rational and strongly respond to tax rebates

THE FEDERAL BUDGET, INTEREST RATES, AND ECONOMIC GROWTH: EVIDENCE

Theory therefore tells us that higher deficits likely lead to higher interest rates and less capital investment, but it does not tell us how much higher and how much less.

Effects of deficits on interest rates depend on circumstances

In normal times, you would expect a positive effect of deficits on interest rates

In recessions (like in the US since 2008), interest rate on govt debt is very low in spite of very large deficits (due to the Fed setting r at zero to stimulate the economy)

The existing empirical literature on this question is somewhat inconclusive, although recent evidence suggests that projected long-term deficits do appear to be reflected to some extent in long-term interest rates

CONCLUSION

The deficit has been a constant source of policy interest and political debate over the last decade

Short-run: should the govt spend more and increase deficit to accelerate recovery from the great recession?

Long-run: should the govt address long-term deficits by reforming retirement and health care benefits?

International evidence suggests that austerity during the Great Recession worsened the recession

Health care cost growth has slowed down sharply since 2008, substantially improving the long-term Federal budget outlook

REFERENCES

Jonathan Gruber, Public Finance and Public Policy, Fourth Edition, 2012
Worth Publishers, Chapter 4

Congressional Budget Office “The Budget and Economic Outlook: Fiscal
Years 2013 to 2023” (web)

Barro, Robert J. “Are government bonds net wealth?.” The Journal of
Political Economy 82.6 (1974): 1095-1117.(web)

Yagan, Danny. “Is the Great Recession Really Over? Longitudinal Evidence
of Enduring Employment Impacts” (web)

Undergraduate Public Economics

Danny Yagan
UC Berkeley

Lecture 5

Income Distribution, Poverty, Taxes, and
Transfers

Recall: Two General Rules for Government Intervention

1) Market Failures: Government intervention can help if there are market failures

2) Redistribution: Free market generates inequality. Public cares about economic disparity. Govt taxes and spending can reduce inequality

Role 2: Redistribution

Even with no market failures, free market outcome might generate substantial inequality

Inequality matters because (a) marginal utility of consumption decreases with wealth, and (b) people evaluate their economic well-being relative to others, not in absolute terms \Rightarrow Public cares about inequality

Govt uses taxes and transfers to redistribute

\Rightarrow Generates an efficiency and equity trade-off (size of economic pie vs. distribution of the economic pie)

Income Inequality: Labor vs. Capital Income

Individuals derive market income (before tax) from **labor** and **capital**: $z = wl + rk$ where w is wage, l is labor supply, k is capital, r is rate of return on capital

1) **Labor income inequality** is due to differences in working abilities (education, talent, physical ability, etc.), work effort (hours of work, effort on the job, etc.), and luck (labor effort might succeed or not)

2) **Capital income inequality** is due to differences in wealth k (due to past saving behavior and inheritances received), and in rates of return r

Capital Income (or wealth) is much more concentrated than Labor Income

Macro-aggregates: Labor vs. Capital Income

Labor income $wl \simeq 75\%$ of market income z

Capital income $rk \simeq 25\%$ of market income z

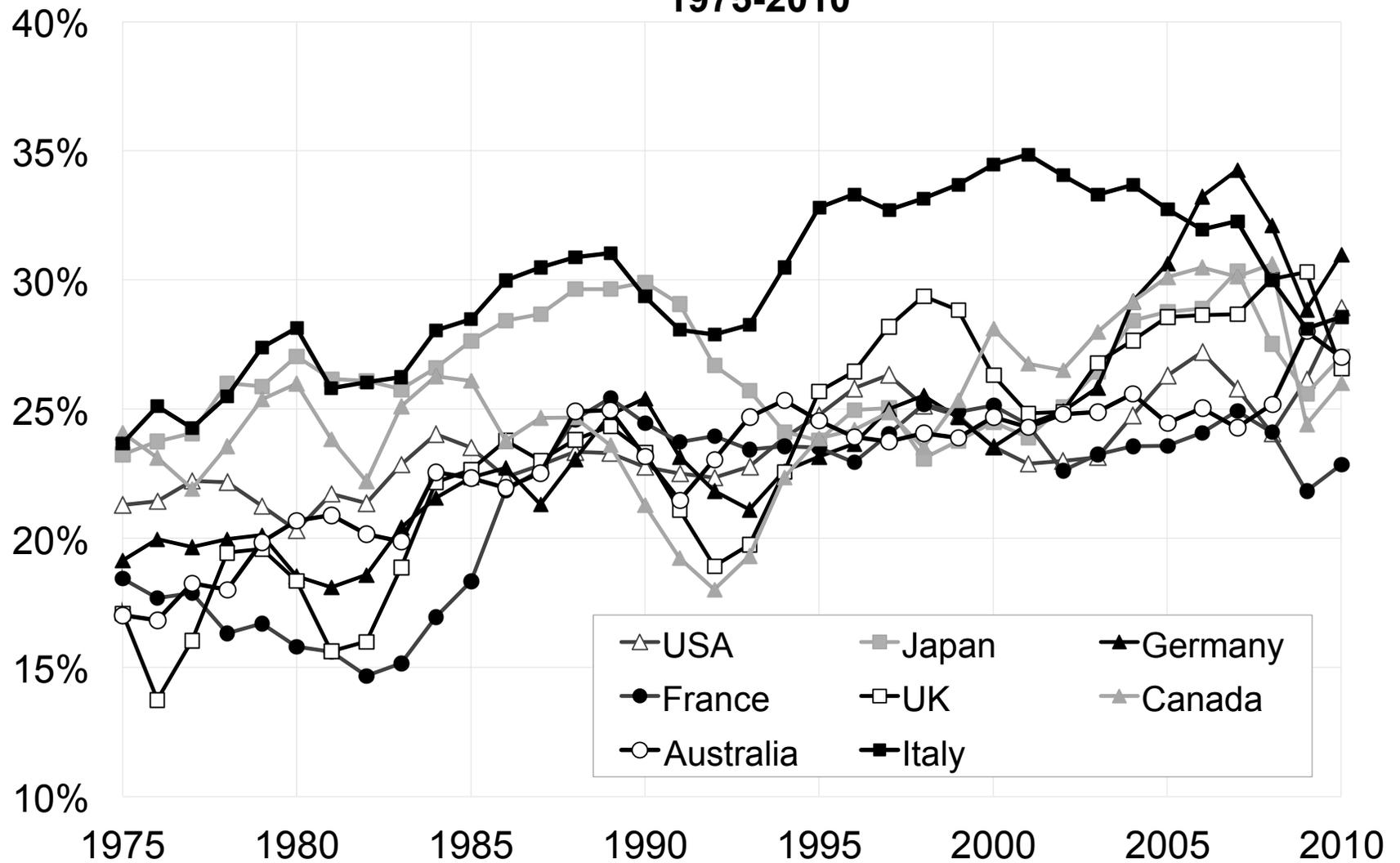
Capital stock $k \simeq 400 - 500\%$ of market income z

Rate of return on capital $r \simeq 5 - 6\%$

In GDP, gross capital share is higher (35%) because it includes depreciation of capital

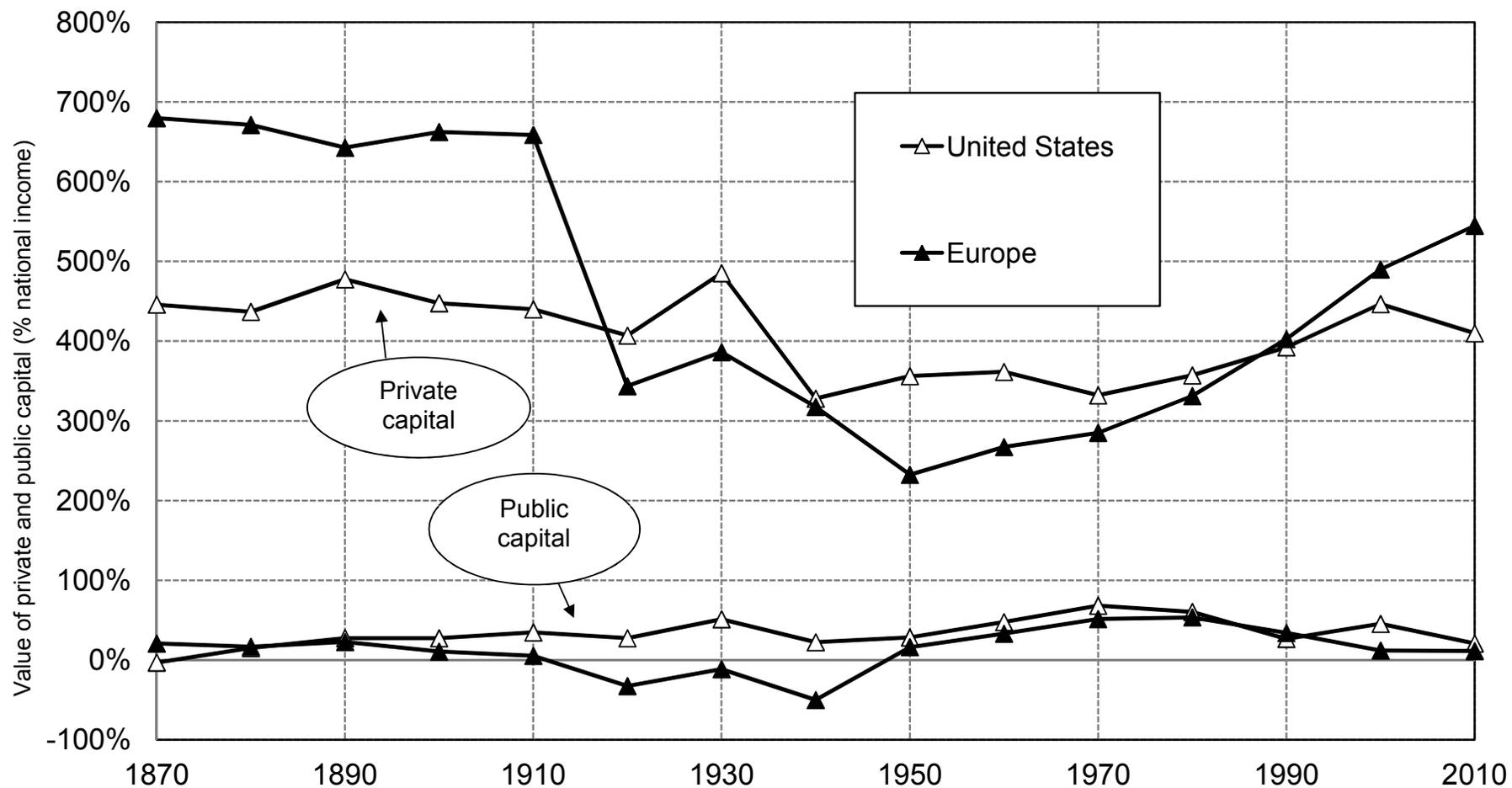
National Income = GDP - depreciation of capital + net foreign income

**Figure 12: Capital shares in factor-price national income
1975-2010**



Source: Piketty and Zucman (2014)
43

Figure 5.1. Private and public capital: Europe and America, 1870-2010



The fluctuations of national capital in the long run correspond mostly to the fluctuations of private capital (both in Europe and in the U.S.). Sources and series: see piketty.pse.ens.fr/capital21c.

Income Inequality Measurement

Inequality can be measured by indexes such as Gini coefficient, quantile income shares which are functions of the income distribution $F(z)$

Most famous inequality index: **Gini coefficient**

Gini = 2 * area between 45 degree line and Lorenz curve

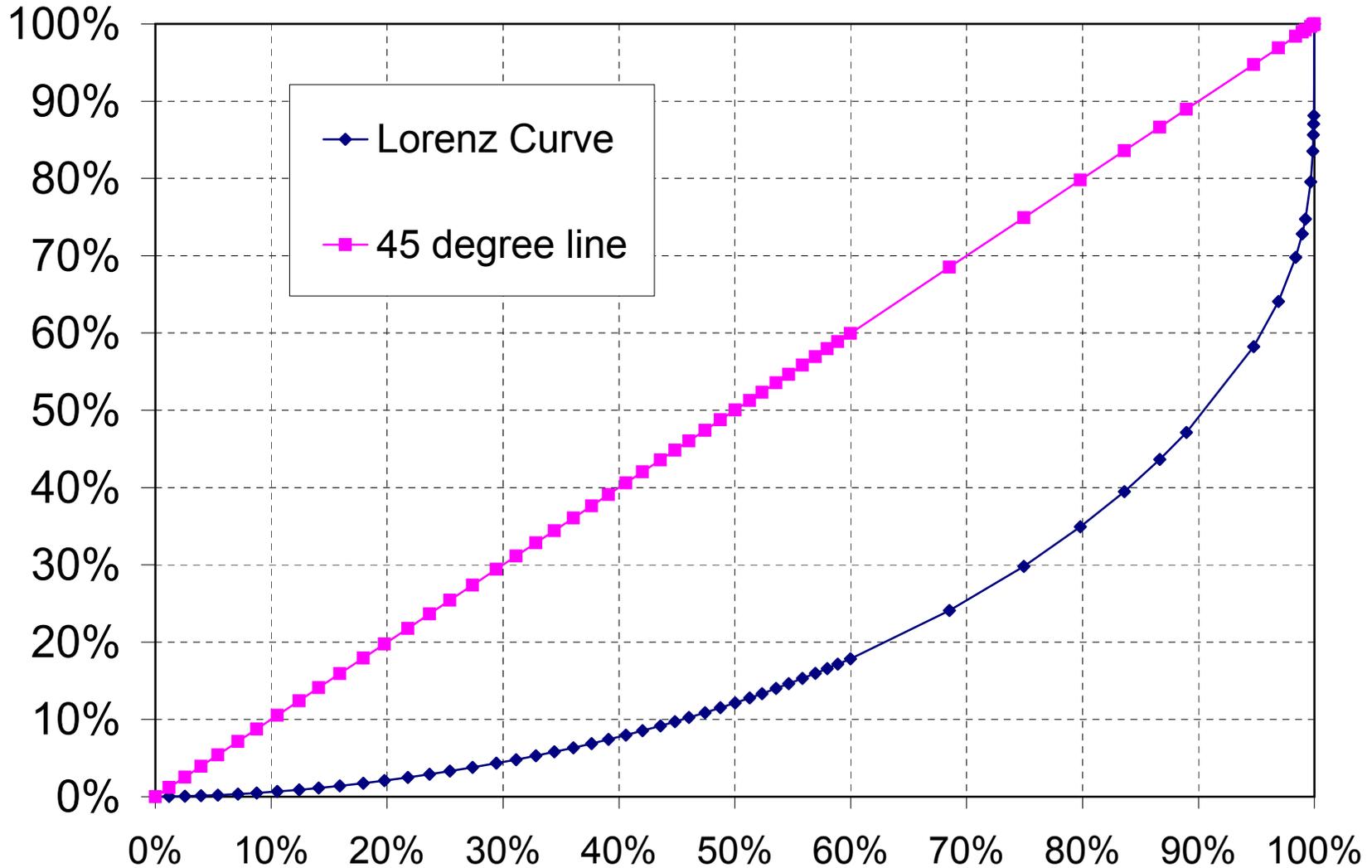
Lorenz curve $L(p)$ at percentile p is fraction of total income earned by individuals below percentile p

$$0 \leq L(p) \leq p$$

Gini=0 means perfect equality

Gini=1 means complete inequality (top person has all the income)

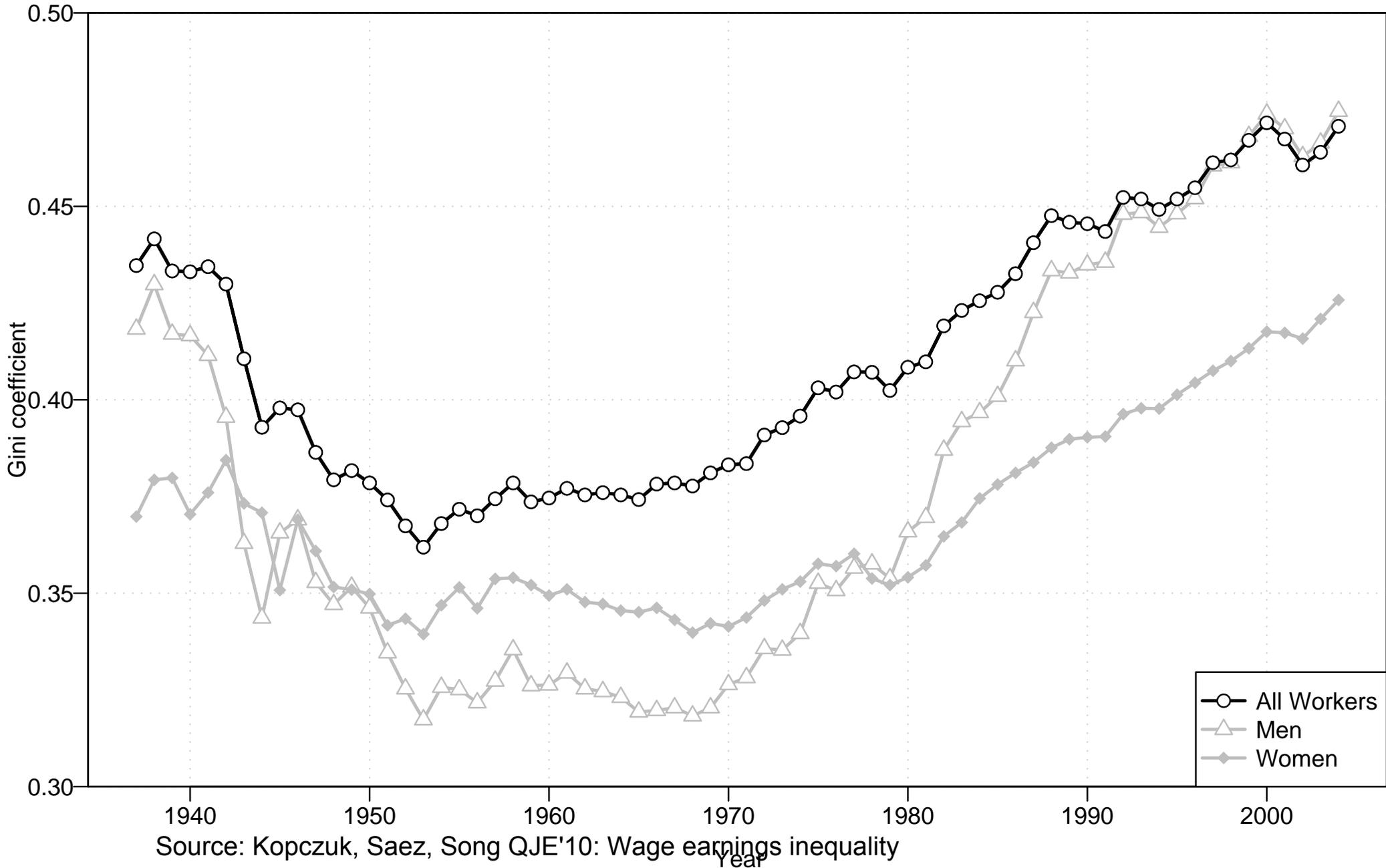
Gini Coefficient California pre-tax income, 2000, Gini=62.1%



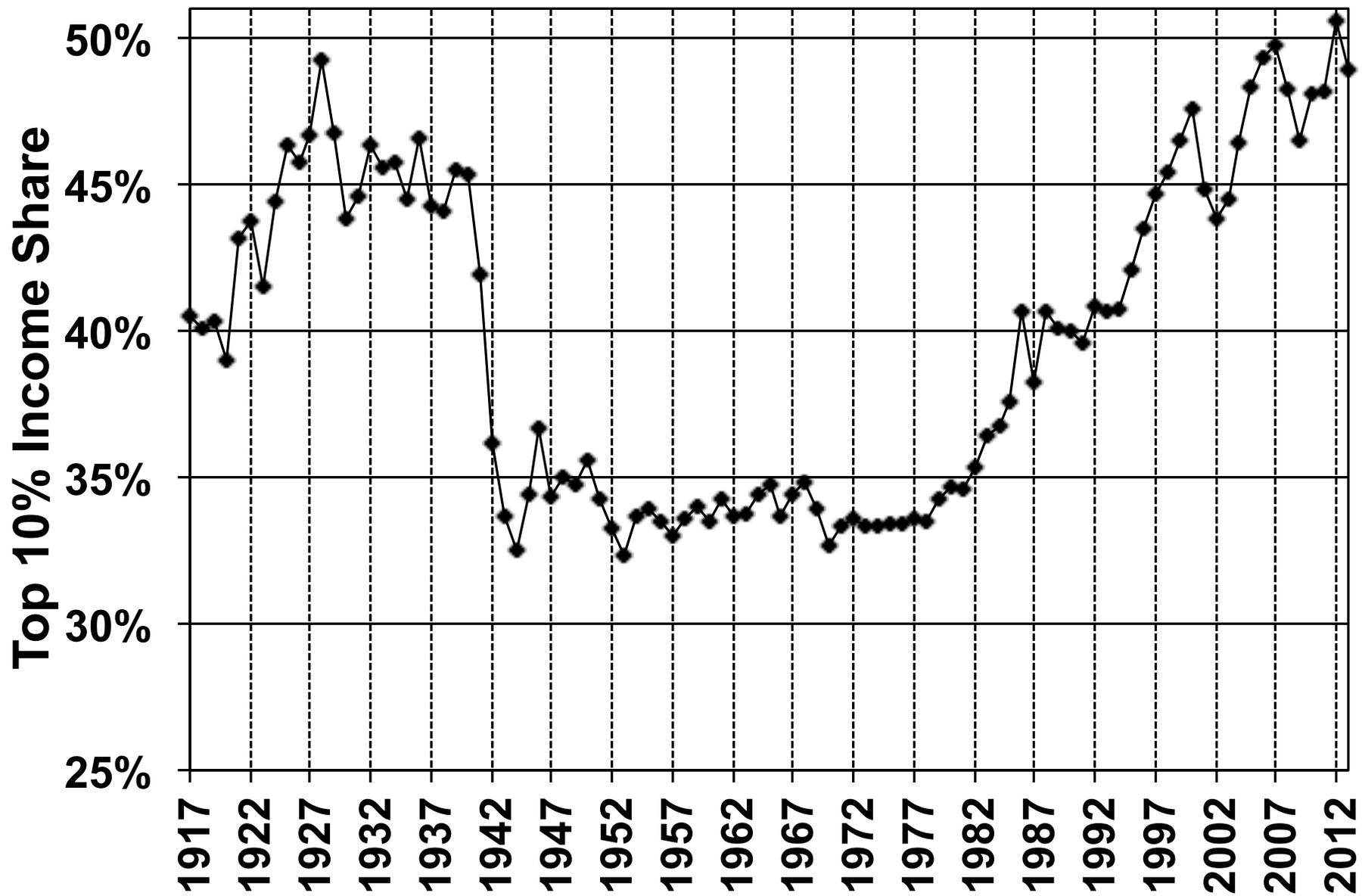
Key Empirical Facts on Income Inequality

- 1) In the US, labor income inequality has increased substantially since 1970: debate between skilled biased technological progress view vs. institution view (min wage and Unions) [Autor-Katz'99]
- 2) In the US, top income shares dropped dramatically from 1929 to 1950 and increased dramatically since 1980 [Piketty and Saez, 2003]
- 3) Top incomes used to be primarily capital income. Now, top incomes are divided 50/50 between labor and capital income (due to explosion of top labor incomes with stock-options, bonuses, etc.)
- 4) Fall in top income shares from 1900-1950 happened in most OECD countries. Surge in top income shares has happened primarily in English speaking countries, not as much in Continental Europe and Japan [Atkinson, Piketty, Saez JEL'11]

Figure 1: Gini coefficient

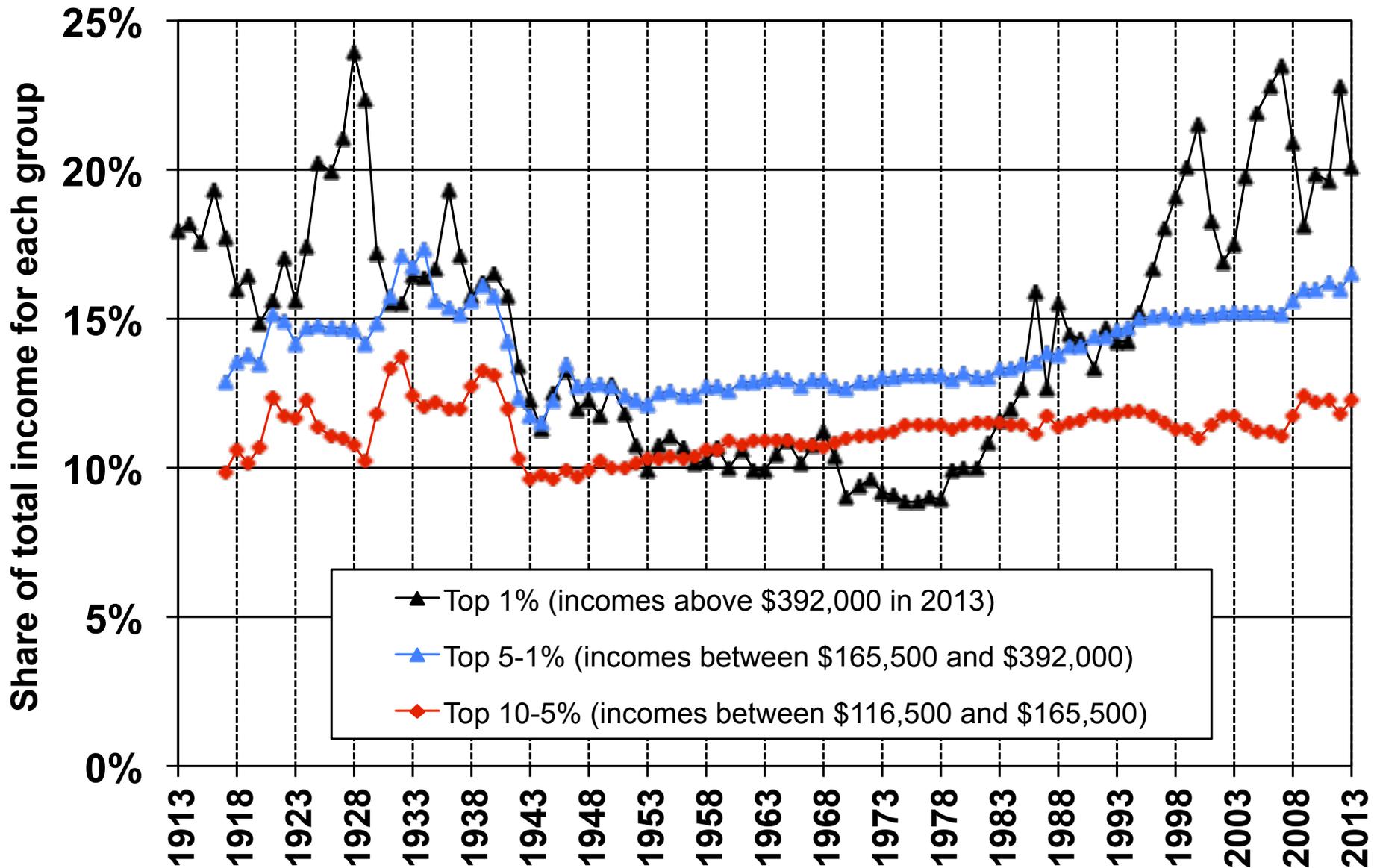


Top 10% Pre-tax Income Share in the US, 1917-2013



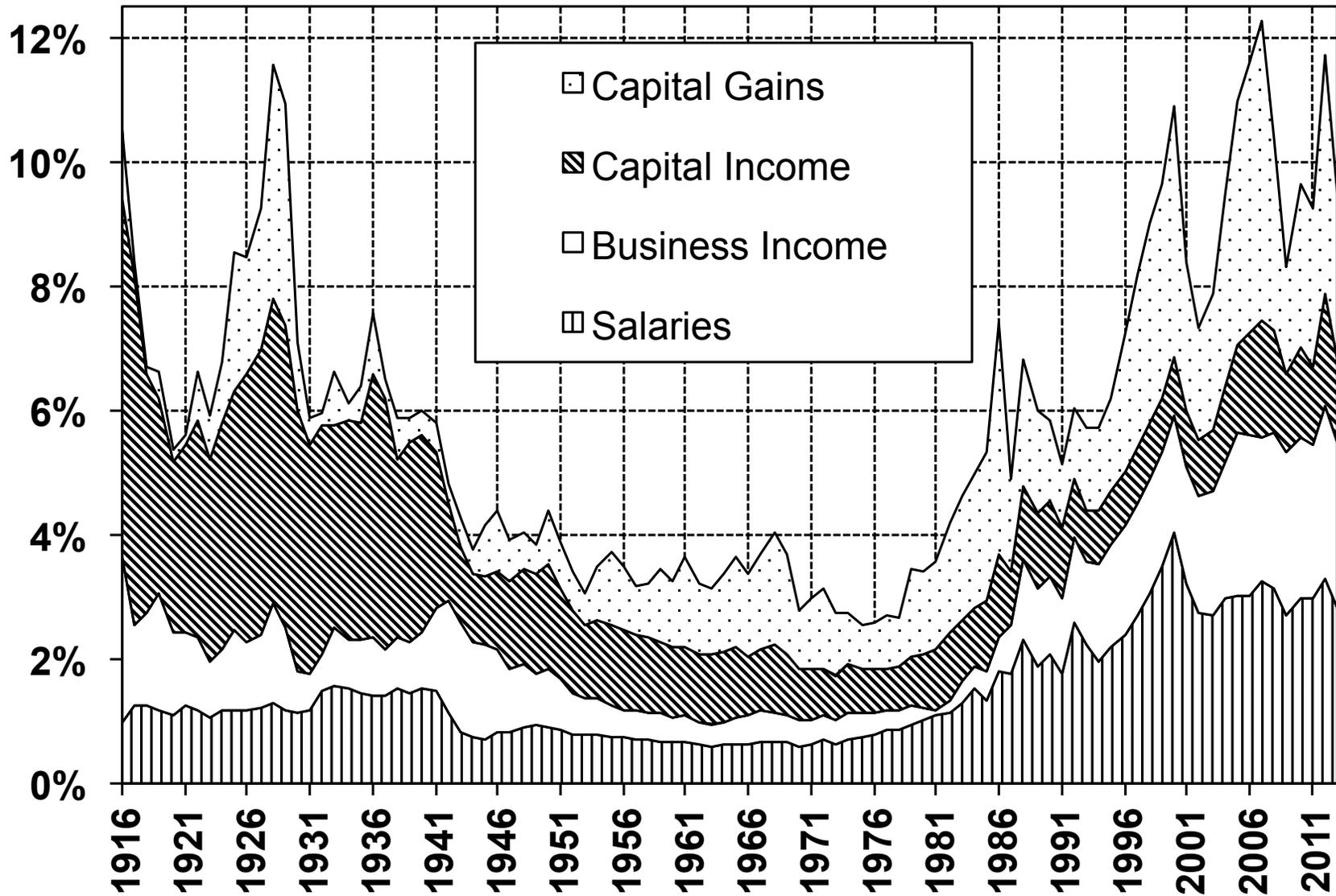
Source: Piketty and Saez, 2003 updated to 2013. Series based on pre-tax cash market income including realized capital gains and excluding government transfers.

Decomposing Top 10% into 3 Groups, 1913-2013



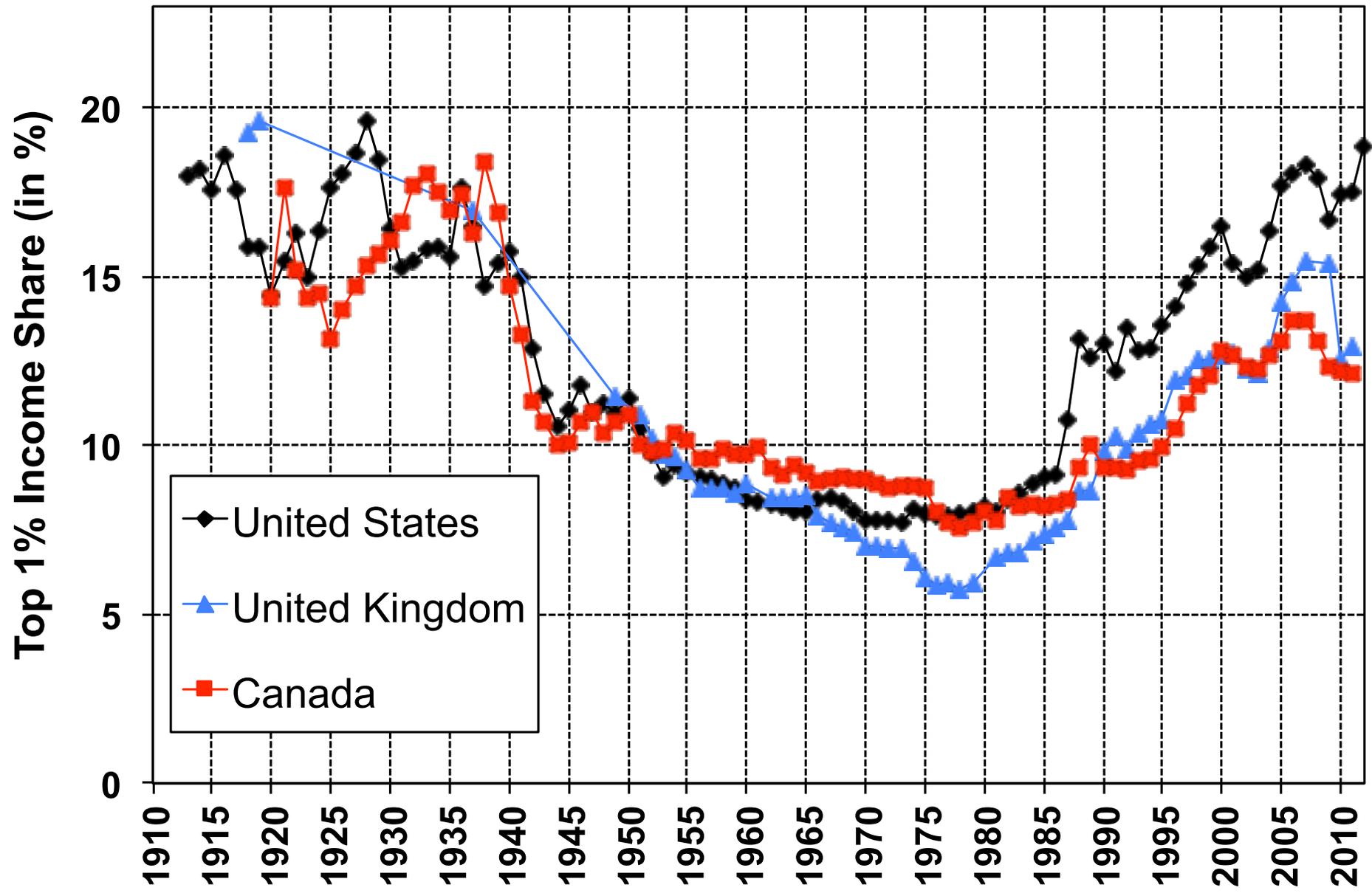
Source: Piketty and Saez, 2003 updated to 2013. Series based on pre-tax cash market income including realized capital gains and excluding government transfers.

US Top 0.1% Pre-Tax Income Share and Composition

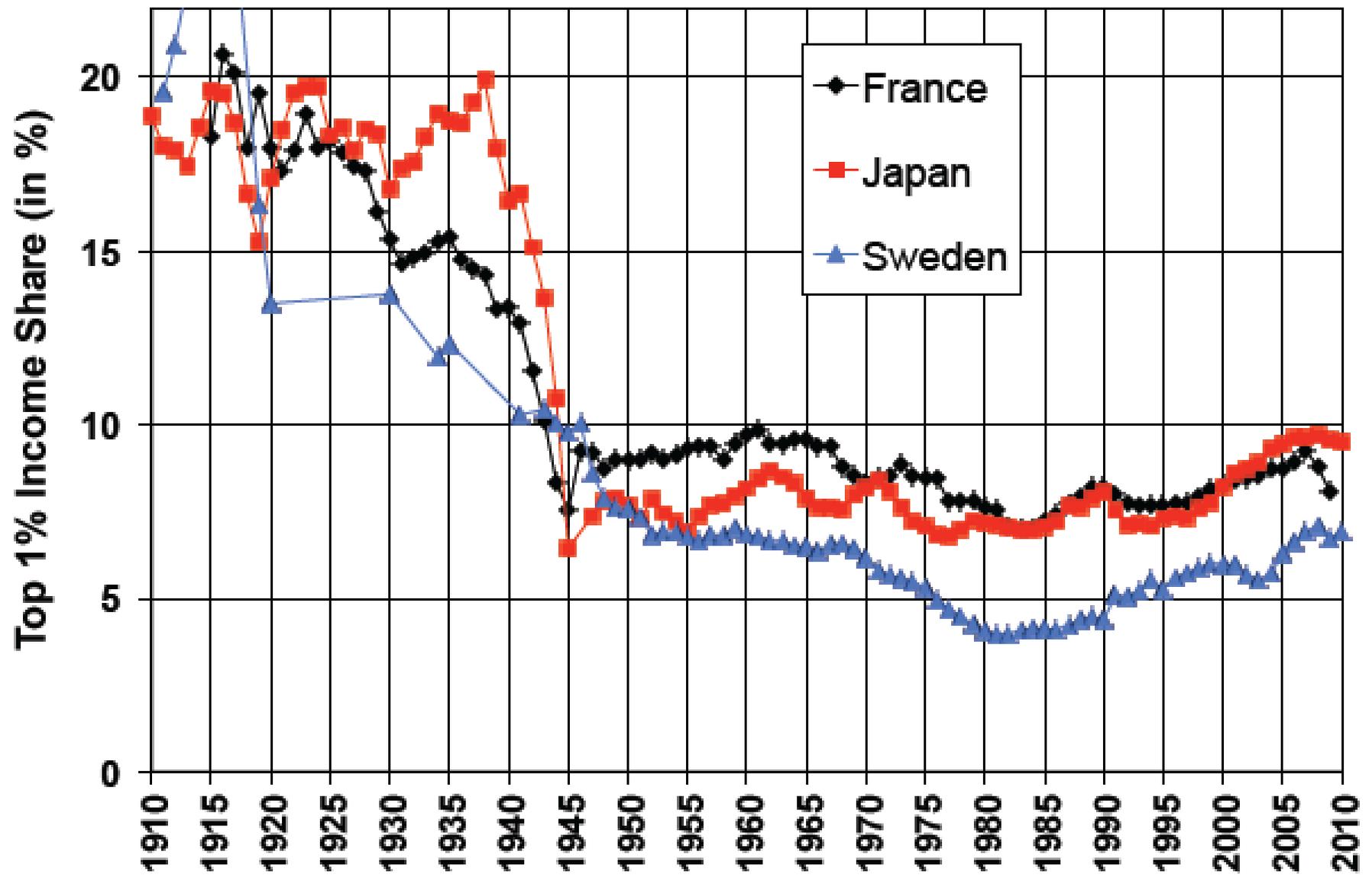


Source: Piketty and Saez, 2003 updated to 2013. Series based on pre-tax cash market income including or excluding realized capital gains, and always excluding government transfers.

Top 1% share: English Speaking countries (U-shaped)



Top 1% share: Continental Europe and Japan (L-shaped)



Source: THE WORLD TOP INCOMES DATABASE

POVERTY RATE DEFINITIONS

1) **Absolute:** Fraction of population with disposable income (normalized by family size) below **poverty threshold** z^* fixed in real terms (e.g., World Bank \$1/day in 1990 dollars)

2) **Relative:** Fraction of population with disposable income (normalized by family size) below **poverty threshold** z^* fixed relative to median (European Union defines poverty threshold as 60% of median)

Absolute poverty falls in the long run with economic growth [nobody in the US is World Bank poor] but relative poverty does not

Absolute poverty captures both growth and inequality effects while relative poverty captures only inequality effects

A recent study by Luttmer (2005) finds that individuals' self-reported well-being rises as their own income rises, but falls as their neighbors' incomes rise, suggesting that it is relative income, and not absolute income, that determines well-being.

Poverty Rate Disposable Income Definition

Most intuitive notion of poverty is based on consumption c
[not pre-tax income z]

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

where $T(z)$ is tax, $B(z)$ govt transfers, E net private transfers
(charity, family, friends), s is net savings (change in assets)

Consumption c is difficult to measure

Disposable Income $z - T(z) + B(z)$ [post-tax income] mea-
sured in traditional Current Population Survey (CPS)

FAMILY SCALE

Ideally, poverty should be defined at the individual level based on individual consumption [e.g., kids better off when mother or grandmother controls income instead of father, Duflo '03]

However, many consumption goods are shared within the family [e.g., housing, joint meals, etc.] and it is difficult to measure consumption at individual level

Measured poverty is therefore based on consumption or disposable income at the family level [or unit sharing resources] and everybody within the family has same poverty status

Bigger families need more resources but economies of scale in consumption: scale disposable income by family size

US POVERTY RATE DEFINITION

Based on **money income** = market income before taxes + cash govt transfers excluding tax credits + cash private transfers

In-kind market income and transfers (employer health insurance, Medicaid, nutrition, public housing) do NOT count

Income and employee payroll taxes are NOT deducted. Income tax credits (EITC, Child Tax Credit) are NOT added.

Threshold: 3x the cost of a minimum food diet in 1963 in today's prices (using the official CPI)

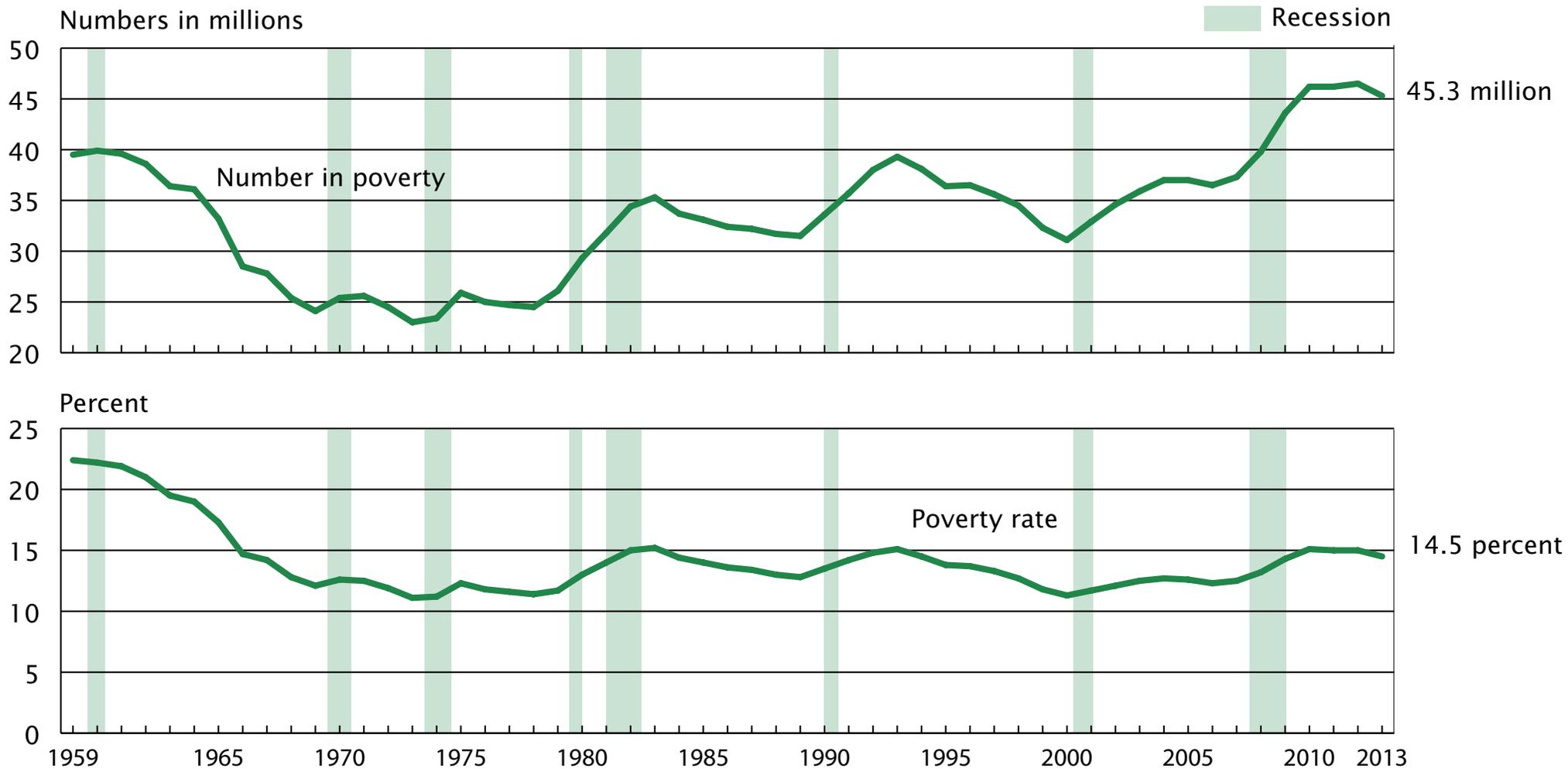
Threshold depends on household size/structure: e.g., \$20K/year for single parent with 2 kids

17.1

Poverty Lines by Family Size (2012)

Size of Family Unit	Poverty Line
1	\$11,170
2	15,130
3	19,090
4	23,050
5	27,010
For each additional person, add	3,960

Figure 4.
Number in Poverty and Poverty Rate: 1959 to 2013

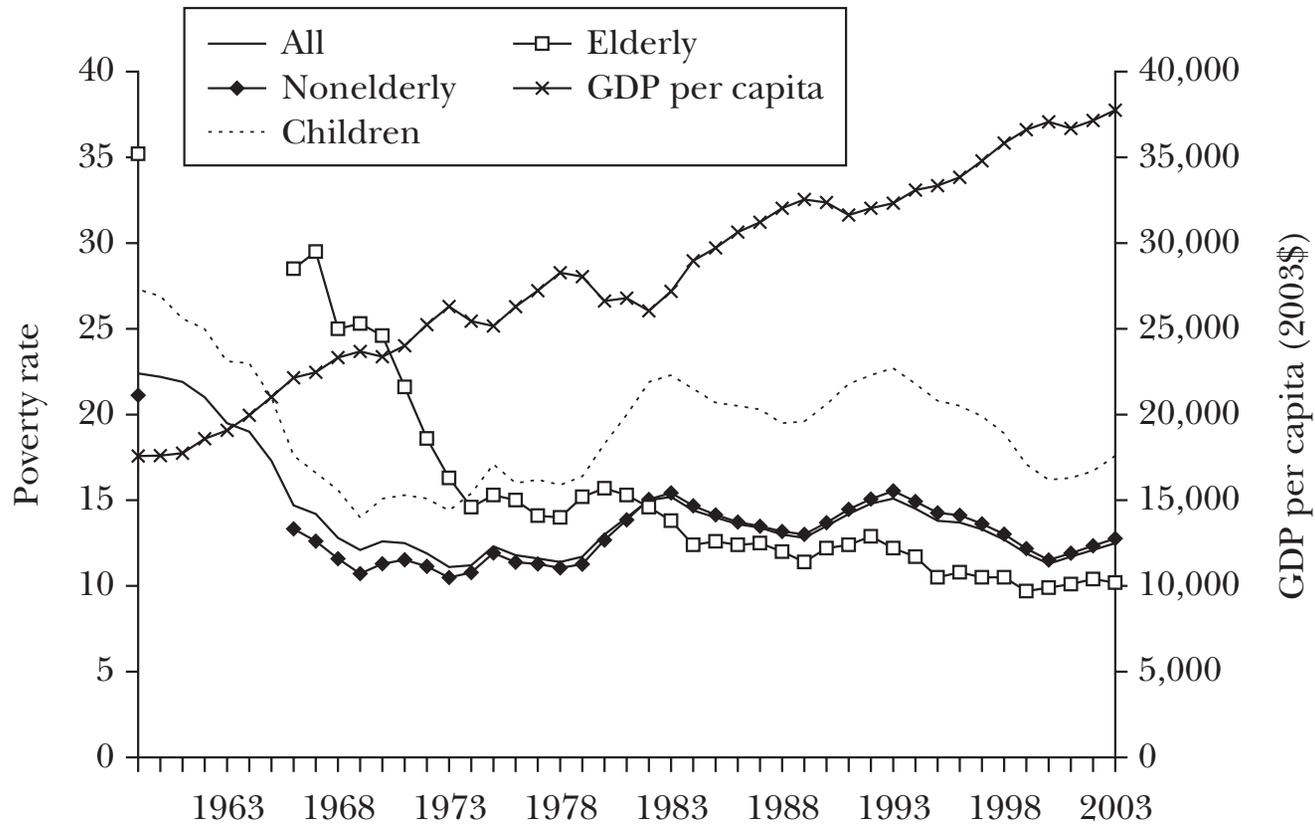


Note: The data points are placed at the midpoints of the respective years. For information on recessions, see Appendix A. For information on confidentiality protection, sampling error, nonsampling error, and definitions, see <ftp://ftp2.census.gov/programs-surveys/cps/techdocs/cpsmar14.pdf>.

Source: U.S. Census Bureau, Current Population Survey, 1960 to 2014 Annual Social and Economic Supplements.

Figure 1

Trends in Individual Poverty Rates and Real GDP per Capita, 1959–2003



Source: Poverty rates are from U.S. Bureau of the Census, Current Population Survey, Annual Social and Economic Supplements. The GDP per capita series is from the Economic Report of the President (2005).

Note: The poverty rate data are unavailable for some subgroups for 1960–1965.

Factors Explaining Evolution of Poverty

Based on Hoynes-Page-Stevens JEP'06

- 1) Stagnant bottom wages (in spite of economic growth per capita [main explanation])
- 2) Changes in family structure: single parent families ↑ from 7% in 1967 to 14.4% in 2003 ⇒ Increases poverty rate by 4 pts [large effect]
- 3) Increase in female labor force participation ⇒ Reduces poverty rate [significant effect only since 1980]
- 4) Immigration: accounts for about 0.7 points in the poverty rate increase from 1969 to 1999 [small effect]
- 5) Means-tested transfers [medium effect because they are concentrated below poverty line]

ISSUES WITH US POVERTY RATE DEFINITION

Definition was close to disposable income when measuring poverty started but no longer:

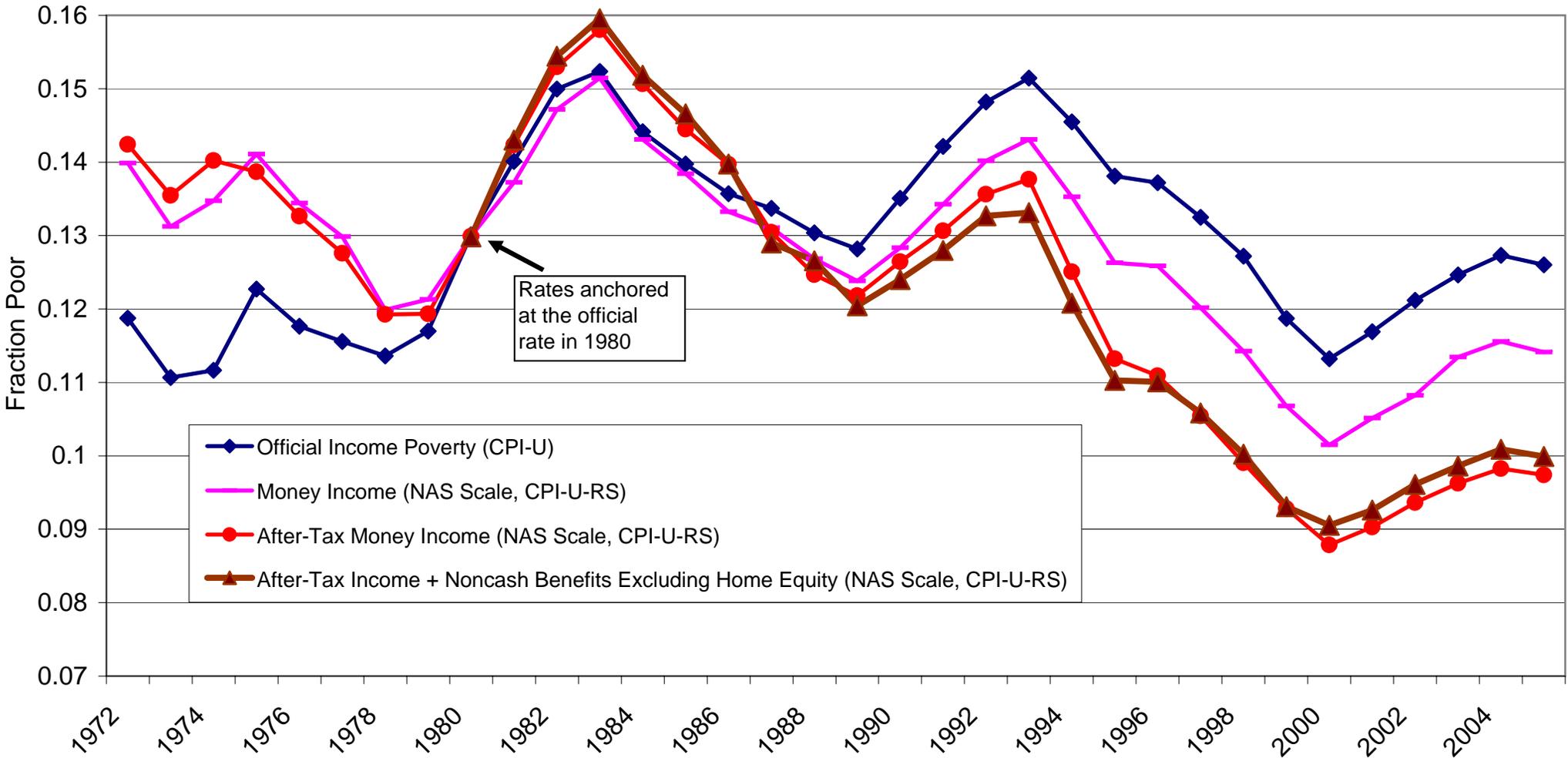
- 1) In-kind transfers have grown substantially [Medicaid]
- 2) Payroll tax and Income tax credits (EITC, Child Tax Credit) have grown substantially for low income families
- 3) Official CPI overstates inflation [and understates real economic growth] because it is not chained [i.e., does not take into account that relative price changes lead to changes in consumption]

Politically difficult to change definition

Recomputing Poverty Rate: Meyer-Sullivan NBER'09

- 1) Change the scaling for family size (no strong effect)
- 2) Change the price index: shift to CPI-U-RS instead of official CPI-U (large legitimate effect, CPI-U-RS better index)
- 3) Shift to households [people living in same unit] instead of family [people in same unit related by blood/adoption]: not clear which is best, depends on sharing [some effect]
- 4) Shift to after-tax income [deduct income/payroll taxes, add tax credits]: large legitimate effect
- 5) Add non-cash benefits [nutrition, housing, health insurance]: tiny net effect [medicaid \uparrow , other programs \downarrow]
- 6) Shift to consumption [modest effect on poverty rate, huge effect on deep poverty]

Figure 1: Official and Alternative Income Poverty Rates, 1972-2005



Notes: The rates are anchored at the official rate in 1980. Data are from the CPS-ASEC/ADF. Official Income Poverty follows the U.S. Census definition of income poverty using official thresholds. For measures other than the official one, the threshold in 1980 is equal to the value that yields a poverty rate equal to the official poverty rate in 1980 (13.0 percent). The thresholds in 1980 are then adjusted overtime using the CPI-U-RS. Poverty status is determined at the family level and then person weighted. After-Tax Money Income includes taxes and credits (calculated using TAXSIM). After-Tax Money Income + Noncash Benefits Excluding Home Equity also includes food stamps and CPS-imputed measures of housing and school lunch subsidies, and the fungible value of Medicaid and Medicare. This last series is only available starting with the 1980 CPS-ASEC/ADF. See Data Appendix for more details.

Source: Meyer, Bruce D., and James X. Sullivan (2009)

Measuring Intergenerational Income Mobility

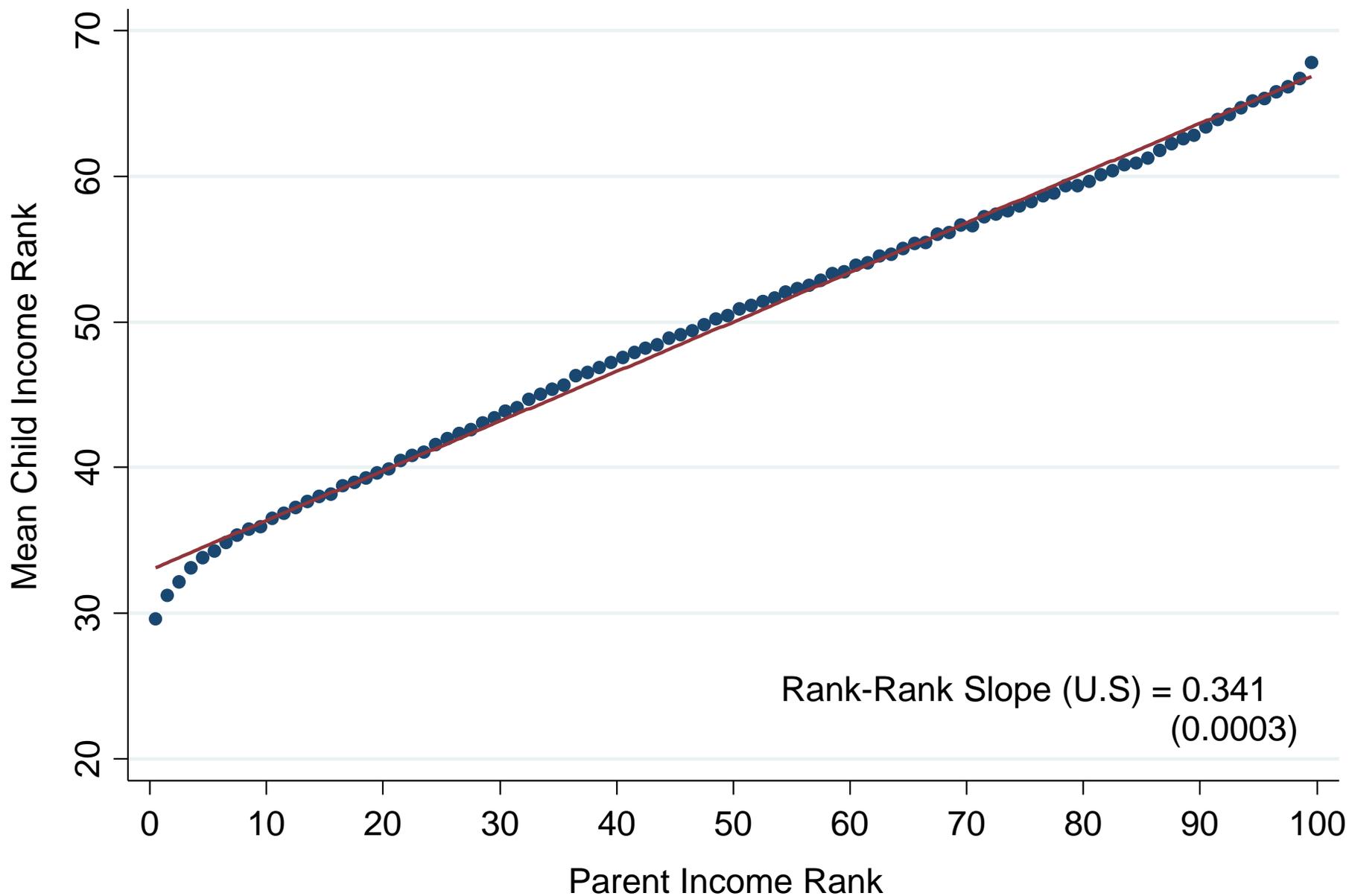
Most believe that children's success should not depend too much on parental income

Studies linking adult children to their parents can measure link between children and parents income

Simple measure: average income rank of children by income rank of parents (Chetty et al. '14)

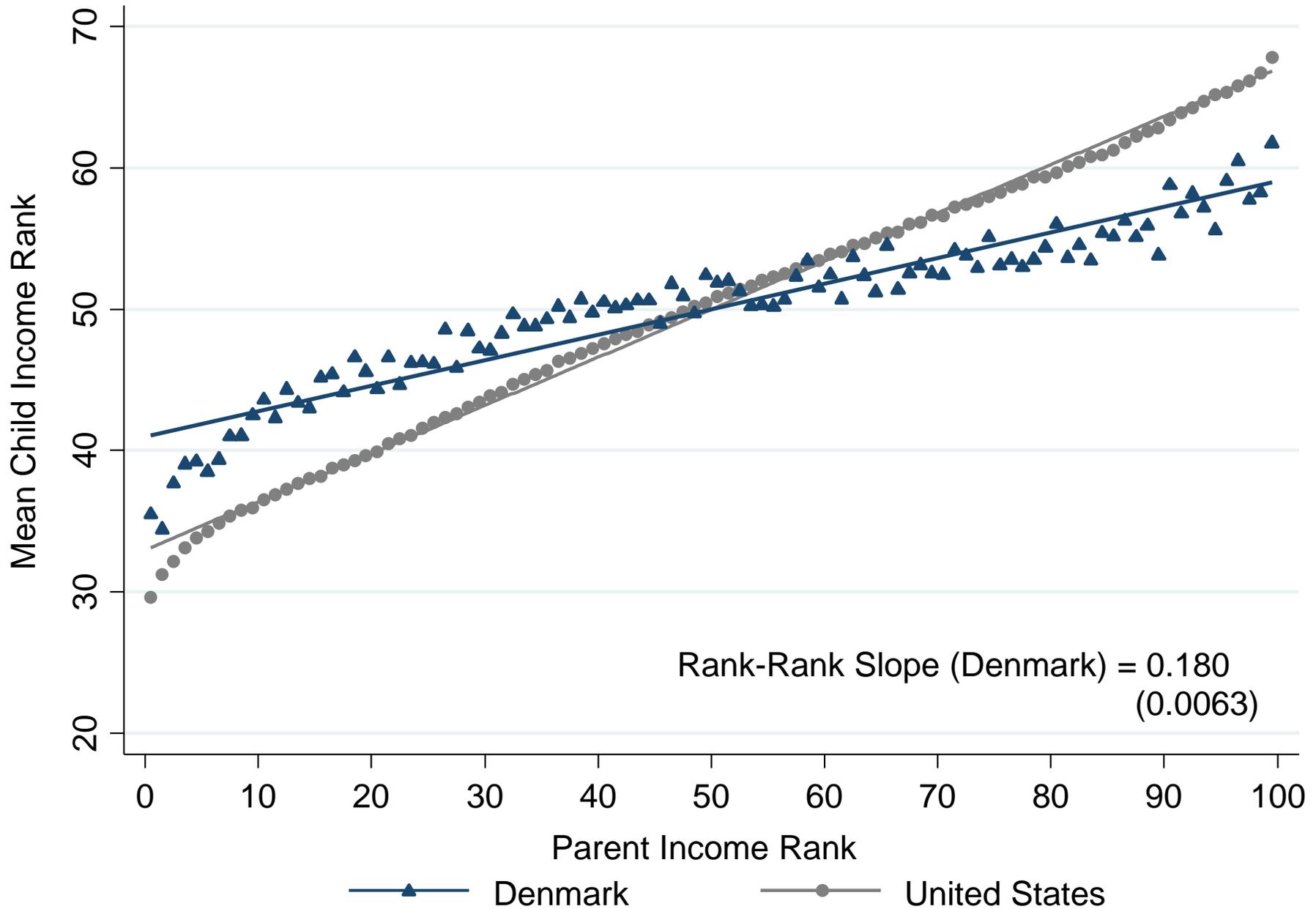
- 1) US has less mobility than European countries (especially Scandinavian countries such as Denmark)
- 2) Substantial heterogeneity in mobility across cities in the US
- 3) Places with low segregation, low income inequality, good K-12 schools, high social capital, high family stability tend to have high mobility [this is a correlation and not necessarily causal]

A. Mean Child Income Rank vs. Parent Income Rank in the U.S.



Source: Chetty, Hendren, Kline, Saez (2014)

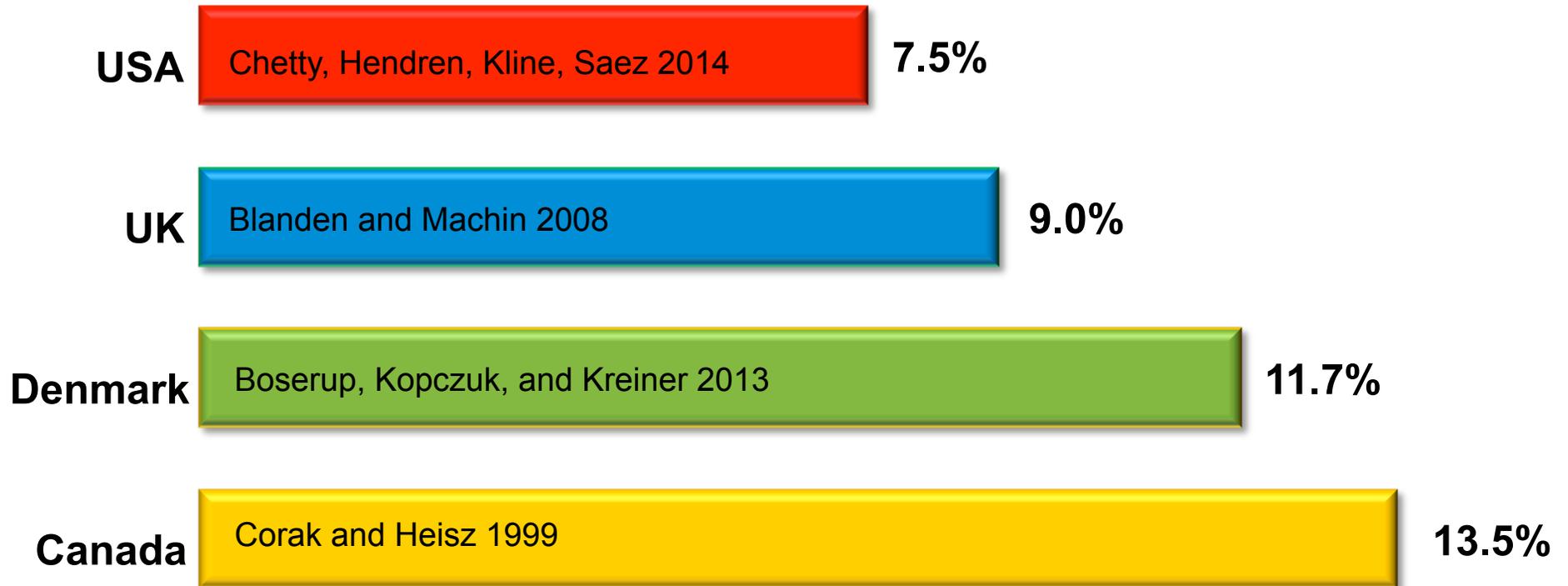
B. United States vs. Denmark



Source: Chetty, Hendren, Kline, Saez (2014)

The American Dream?

- Probability that a child born to parents in the bottom fifth of the income distribution reaches the top fifth:

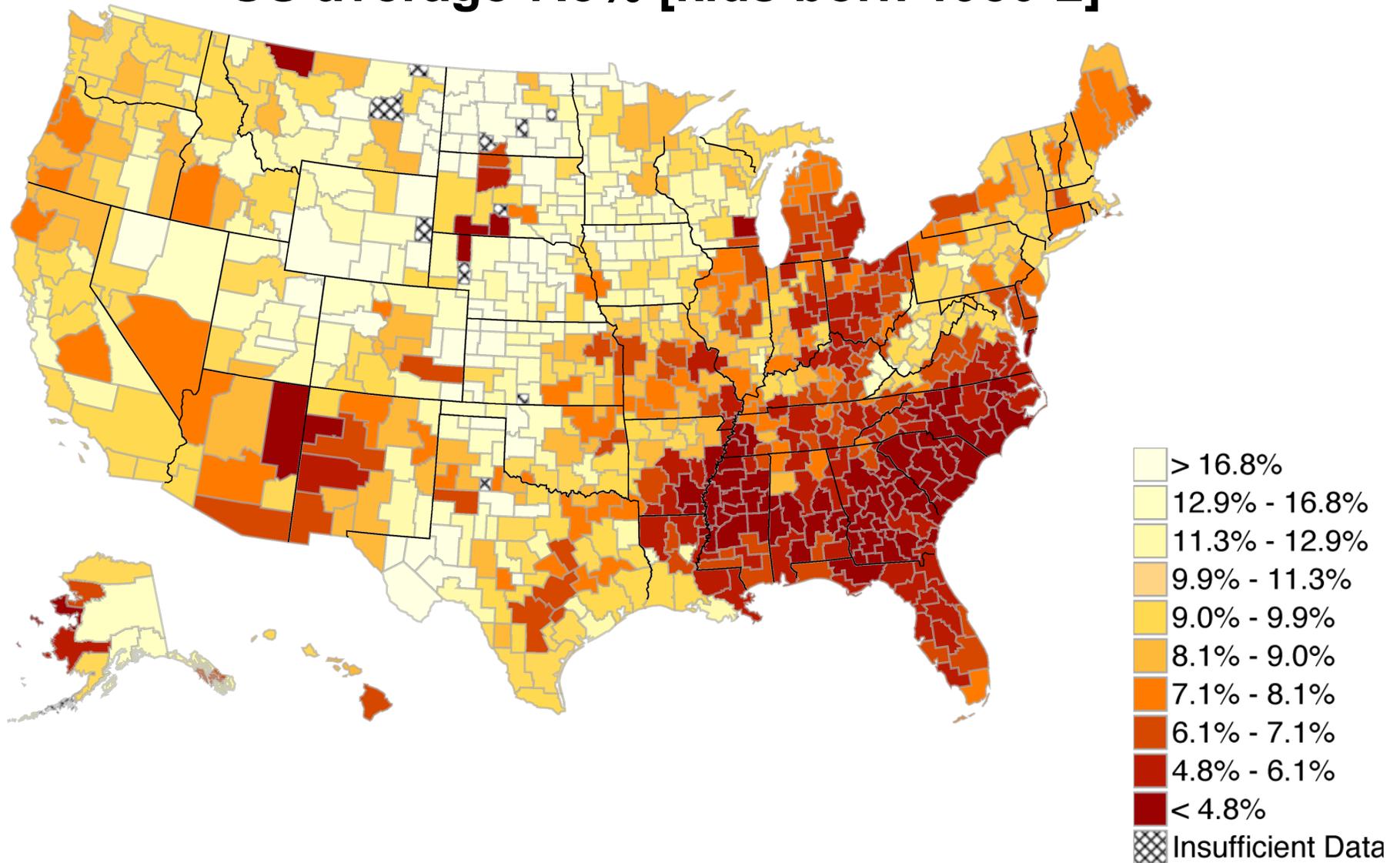


→ Chances of achieving the “American Dream” are almost two times higher in Canada than in the U.S.

The Geography of Upward Mobility in the United States

Probability of Reaching the Top Fifth Starting from the Bottom Fifth

US average 7.5% [kids born 1980-2]



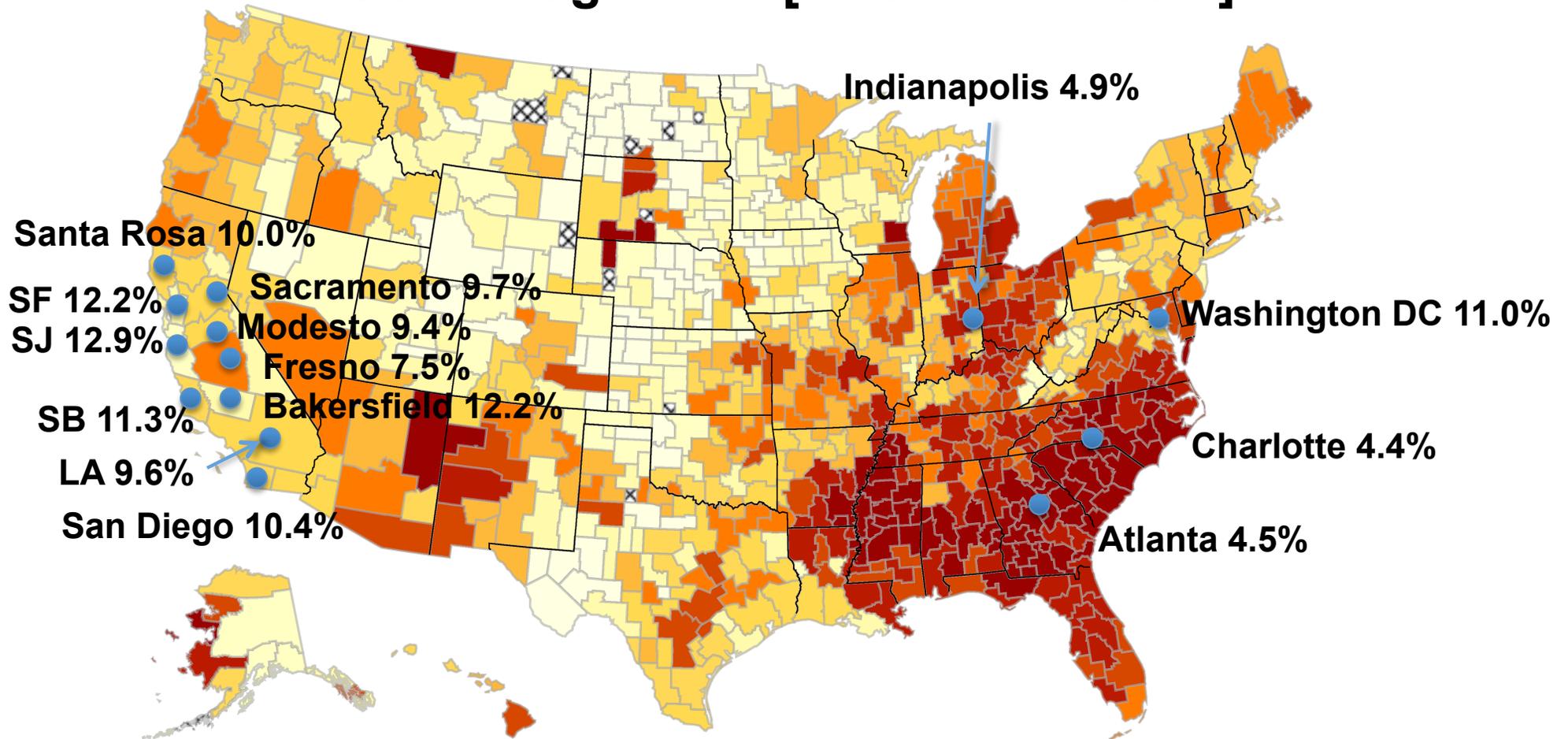
Note: Lighter Color = More Upward Mobility

Download Statistics for Your Area at www.equality-of-opportunity.org

The Geography of Upward Mobility in the United States

Odds of Reaching the Top Fifth Starting from the Bottom Fifth

US average 7.5% [kids born 1980-2]



Note: Lighter Color = More Upward Mobility

Download Statistics for Your Area at www.equality-of-opportunity.org

TABLE 1. Upward Mobility in the 50 Largest Metro Areas: The Top 10 and Bottom 10

Rank	Commuting Zone	Odds of Reaching Top Fifth from Bottom Fifth	Rank	Commuting Zone	Odds of Reaching Top Fifth from Bottom Fifth
1	San Jose, CA	12.9%	41	Cleveland, OH	5.1%
2	San Francisco, CA	12.2%	42	St. Louis, MO	5.1%
3	Washington, D.C.	11.0%	43	Raleigh, NC	5.0%
4	Seattle, WA	10.9%	44	Jacksonville, FL	4.9%
5	Salt Lake City, UT	10.8%	45	Columbus, OH	4.9%
6	New York, NY	10.5%	46	Indianapolis, IN	4.9%
7	Boston, MA	10.5%	47	Dayton, OH	4.9%
8	San Diego, CA	10.4%	48	Atlanta, GA	4.5%
9	Newark, NJ	10.2%	49	Milwaukee, WI	4.5%
10	Manchester, NH	10.0%	50	Charlotte, NC	4.4%

Note: This table reports selected statistics from a sample of the 50 largest commuting zones (CZs) according to their populations in the 2000 Census. The columns report the percentage of children whose family income is in the top quintile of the national distribution of child family income conditional on having parent family income in the bottom quintile of the parental national income distribution—these probabilities are taken from Online Data Table VI of Chetty et al., 2014a.

Source: Chetty et al., 2014a.

Govt Redistribution with Taxes and Transfers

Govt taxes individuals based on income and consumption and provides transfers: z is pre-tax income, $y = z - T(z) + B(z)$ is post-tax income

1) If inequality in y is less than inequality in $z \Leftrightarrow$ tax and transfer system is redistributive (or progressive)

2) If inequality in y is more than inequality in $z \Leftrightarrow$ tax and transfer system is regressive

a) If $y = z \cdot (1 - t)$ with constant t , tax/transfer system is neutral

b) If $y = z \cdot (1 - t) + G$ where G is a universal transfer, then tax/transfer system is progressive

Actual tax/transfer systems in rich countries roughly like b) with G welfare state transfers [education, health, retirement]

Federal US Tax System: Overview

- 1) Individual income tax (on both labor+capital income) [progressive](40% of fed tax revenue)
- 2) Payroll taxes (on labor income) financing social security programs [about neutral] (40% of revenue)
- 3) Corporate income tax (on capital income) [progressive if incidence on capital income] (15% of revenue)
- 4) Estate taxes (on capital income) [very progressive] (2% of revenue)
- 5) Minor excise taxes (mostly labor income) [regressive] (3% of revenue)

State+Local Tax System: Overview

- 1) Individual+Corporate income taxes [progressive] (30% of state+local tax revenue)
- 2) Sales + Excise taxes (tax on consumption = income - savings) [slightly regressive] (30% of revenue)
- 3) Real estate property taxes (on capital income) [slightly progressive] (30% of revenue)

<http://www.census.gov/govs/www/qtax.html>

18.1

Tax Revenue by Type of Tax in the United States (2010, % of Total Tax Revenue)

	Federal	State and Local	Total
Individual income taxes	42%	20%	34%
Social insurance contributions (payroll tax)	35	0	24
Corporate taxes	13	4	10
Consumption tax	3	34	14
Property tax	0	33	11
Other	7	9	7

18.1

Taxation Around the World

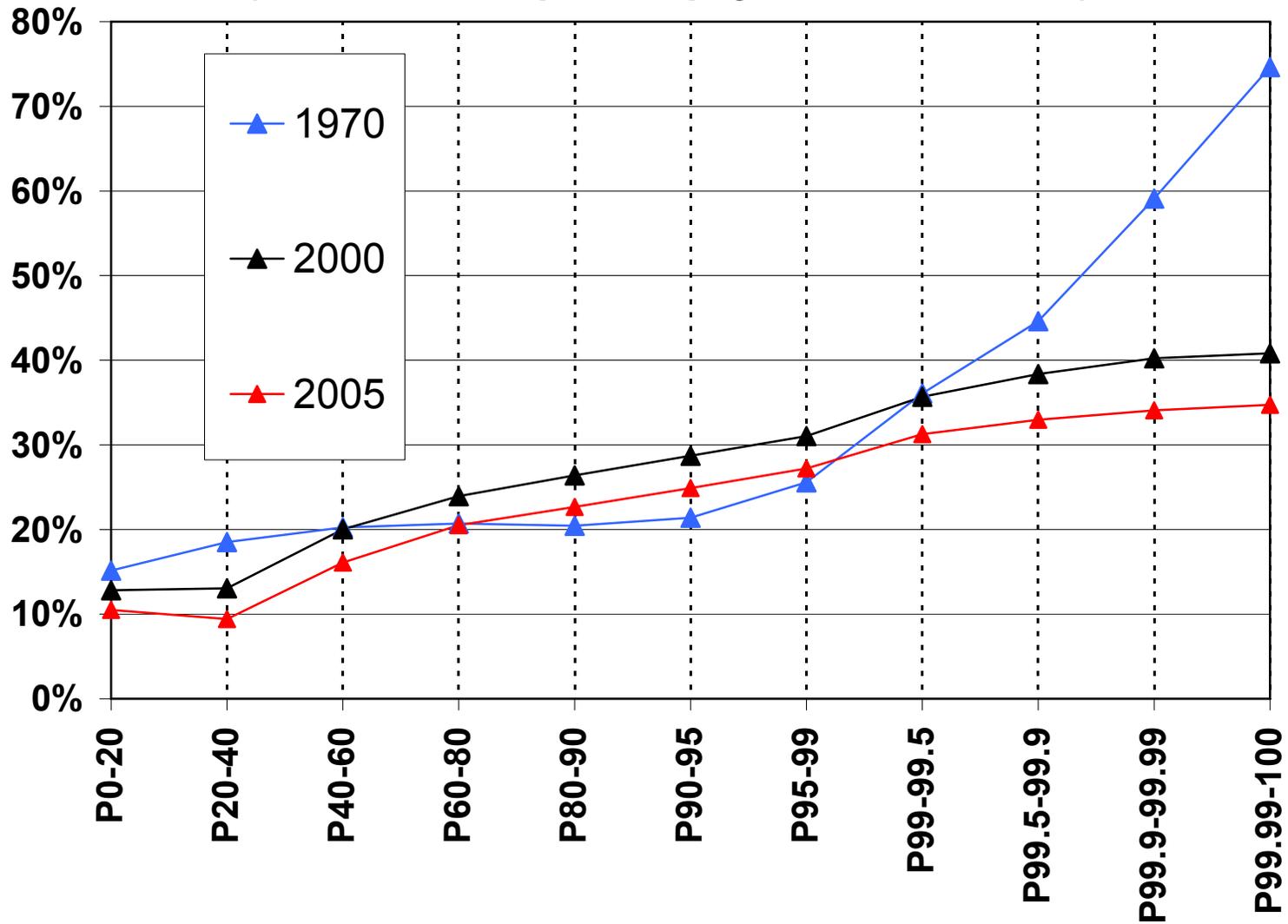
	Norway	Denmark	OECD Average
Individual income taxes	24%	55%	25%
Social insurance contributions (payroll tax)	23	2	27
Corporate taxes	22	5	8
Consumption tax	26	30	31
Property tax	3	4	5
Other	2	4	4

US Tax System: Progressivity and Evolution

1) Medium Term Changes: Federal Tax Progressivity has declined since 1970 but govt redistribution remains substantial especially when including transfers (Medicaid, Social Security, UI, DI, various income support programs)

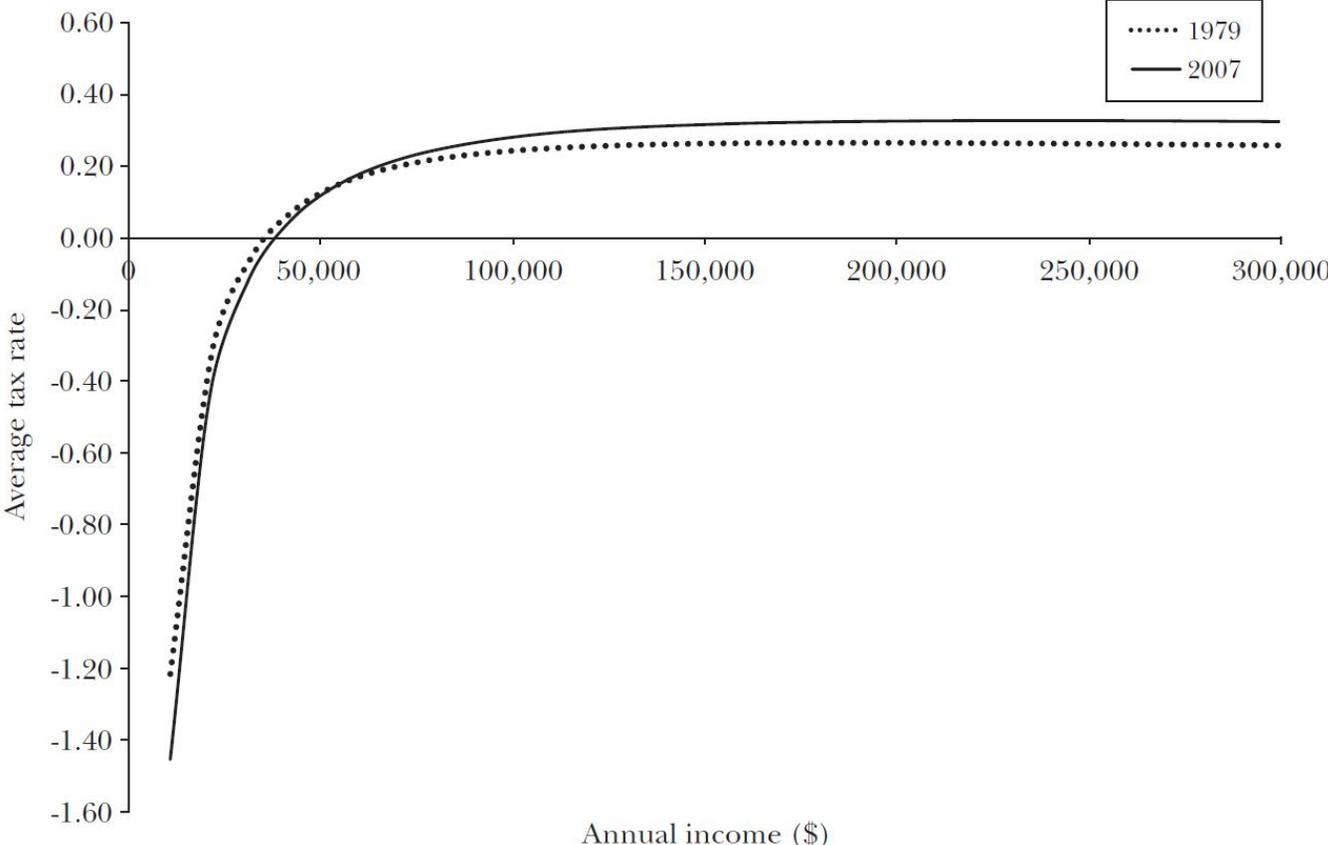
2) Long Term Changes: Before 1913, US taxes were primarily tariffs, excises, and real estate property taxes [slightly regressive], no transfer programs (and hence small govt)

2. Federal Average Tax Rates by Income Groups (individual+corporate+payroll+estate taxes)



POLICY ANALYSIS PREVIEW: PROGRESSIVITY SHOULD RISE WHEN INEQUALITY RISES

Optimal Average Tax Rates for Different Wage Distributions, 1979 and 2007



Source: Mankiw, Weinzierl, Yagan (2009)

Plan for Lectures on Taxation/Redistribution

1) Tax incidence (who bears the burden of taxation), efficiency costs of taxation, optimal commodity taxation

2) Taxation of labor income:

Optimal design of labor income taxation and means-tested transfers

Empirical analysis of tax and transfer programs on labor supply and earnings

3) Taxation of capital income (savings, wealth, and corporate profits)

REFERENCES

Jonathan Gruber, *Public Finance and Public Policy*, Fourth Edition, 2012
Worth Publishers, Chapter 17 and Chapter 18

Atkinson, A., F. Alvaredo, T. Piketty and E. Saez *The World Top Incomes Database*,(web)

Atkinson, Anthony B., Thomas Piketty, and Emmanuel Saez. "Top Incomes in the Long Run of History." *Journal of Economic Literature* 49.1 (2011): 3-71.(web)

Blanden, J and Machin, S (2008) "Up and down the generational income ladder in Britain: Past changes and future prospects" *National Institute Economic Review* 205 (1). 101–116.

Boserup, Simon, Wojciech Kopczuk, and Claus Kreiner "Stability and Persistence of Intergenerational Wealth Formation: Evidence from Danish Wealth Records of Three Generations", October 2014 (web)

Chetty, Raj, Nathan Hendren, Lawrence Katz, "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment," *American Economic Review* forthcoming (web)

Chetty, Raj, Nathan Hendren, Patrick Kline, and Emmanuel Saez, "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States," *Quarterly Journal of Economics* 2014. (web)

Chetty, Raj, Nathan Hendren, Patrick Kline, Emmanuel Saez, and Nicholas Turner “Is the United States Still a Land of Opportunity? Trends in Intergenerational Mobility Over 25 Years,” NBER Working Paper No. 19844, January 2014. (web)

Corak, Miles, and Andrew Heisz, “The Intergenerational Earnings and Income Mobility of Canadian Men: Evidence from Longitudinal Income Tax Data,” *Journal of Human Resources*, 34, no. 3 (1999), 504–533. (web)

Duflo, Esther. “Grandmothers and Granddaughters: Old-Age Pensions and Intrahousehold Allocation in South Africa”, *The World Bank Economic Review* Vol. 17, 2003, 1-25 (web)

Hoynes, Hilary W., Marianne E. Page, and Ann Huff Stevens. “Poverty in America: Trends and explanations.” *The Journal of Economic Perspectives* 20.1 (2006): 47-68.(web)

Katz, Lawrence F., and David H. Autor. “Changes in the wage structure and earnings inequality.” *Handbook of Labor Economics* 3 (1999): 1463-1555. (web)

Kopczuk, Wojciech, Emmanuel Saez, and Jae Song. “Earnings inequality and mobility in the United States: evidence from social security data since 1937.” *The Quarterly Journal of Economics* 125.1 (2010): 91-128.(web)

Luttmer, Erzo FP. “Neighbors as negatives: Relative earnings and well-being.” *Quarterly Journal of Economics* 120.3 (2005): 963-1002.(web)

Mankiw, N.G., Weinzierl, M. and Yagan, D., 2009. Optimal taxation in theory and practice. *The Journal of Economic Perspectives*, 23(4), pp.147-174. (web)

Meyer, Bruce D., and James X. Sullivan. "Consumption and Income Inequality in the US: 1960-2008." (2010).(web)

Piketty, Thomas, and Emmanuel Saez. "Income inequality in the United States, 1913-1998." *The Quarterly Journal of Economics* 118.1 (2003): 1-41.(web)

Piketty, Thomas, and Emmanuel Saez. "How Progressive is the US Federal Tax System? A Historical and International Perspective." *The Journal of Economic Perspectives* 21.1 (2007): 3-24.(web)

US Census Bureau, 2014. "Income and Poverty in the United States: 2013", report P60-249. (web)

Undergraduate Public Economics

Danny Yagan
UC Berkeley

Lecture 6

Tax Incidence and Efficiency Costs of Taxation

19.1

Tax Incidence

Sources of federal government revenue, 1960 and 2008:

Category:	1960	2008
Income taxes	44.5%	43.7%
Corporate taxes	22.8	11.3
Payroll tax	17.0	37.8
Excise taxes	12.8	2.6
Other	2.9	4.5

- **Tax incidence:** Assessing which party (consumers or producers) bears the true burden of a tax.

TAX INCIDENCE

Tax incidence is the study of the effects of tax policies on prices and the welfare of individuals

What happens to market prices when a tax is introduced or changed?

Example: what happens when impose \$1 per pack tax on cigarettes?

Effect on price \Rightarrow distributional effects on smokers, profits of producers, shareholders, farmers, etc.

This is positive analysis: typically the first step in policy evaluation; it is an input to later thinking about what policy maximizes social welfare.

TAX INCIDENCE

Tax incidence is not an accounting exercise but an analytical characterization of changes in economic equilibria when taxes are changed.

Key point: Taxes can be shifted: taxes affect directly the prices of goods, which affect quantities because of behavioral responses, which affect indirectly the price of other goods.

If prices are constant economic incidence would be the same as legislative incidence.

Example: Liberals favor capital income taxation because capital income is concentrated at the high end of the income distribution. Taxing capital means taxing disproportionately the rich.

Argument neglects implicitly general equilibrium price effects: if people save less because of capital taxes, capital stock may go down driving also the wages down and hurting workers. The capital tax might be shifted partly on workers.

Partial Equilibrium Tax Incidence

Partial Equilibrium Model:

Simple model goes a long way to showing main results.

Government levies an excise tax on good x

Excise means it is levied on a quantity (gallon, pack, ton, ...).
Typically fixed in nominal terms (e.g, \$1 per pack)

[ad-valorem tax is a fraction of prices (e.g. 5% sales tax)]

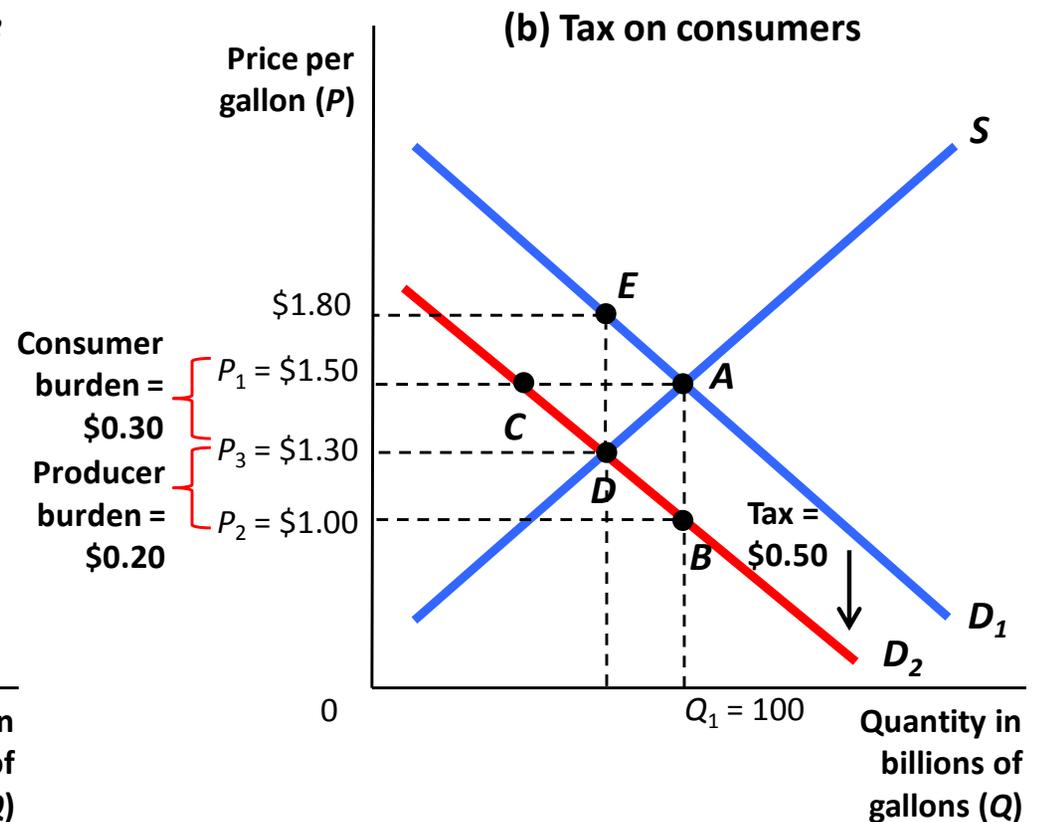
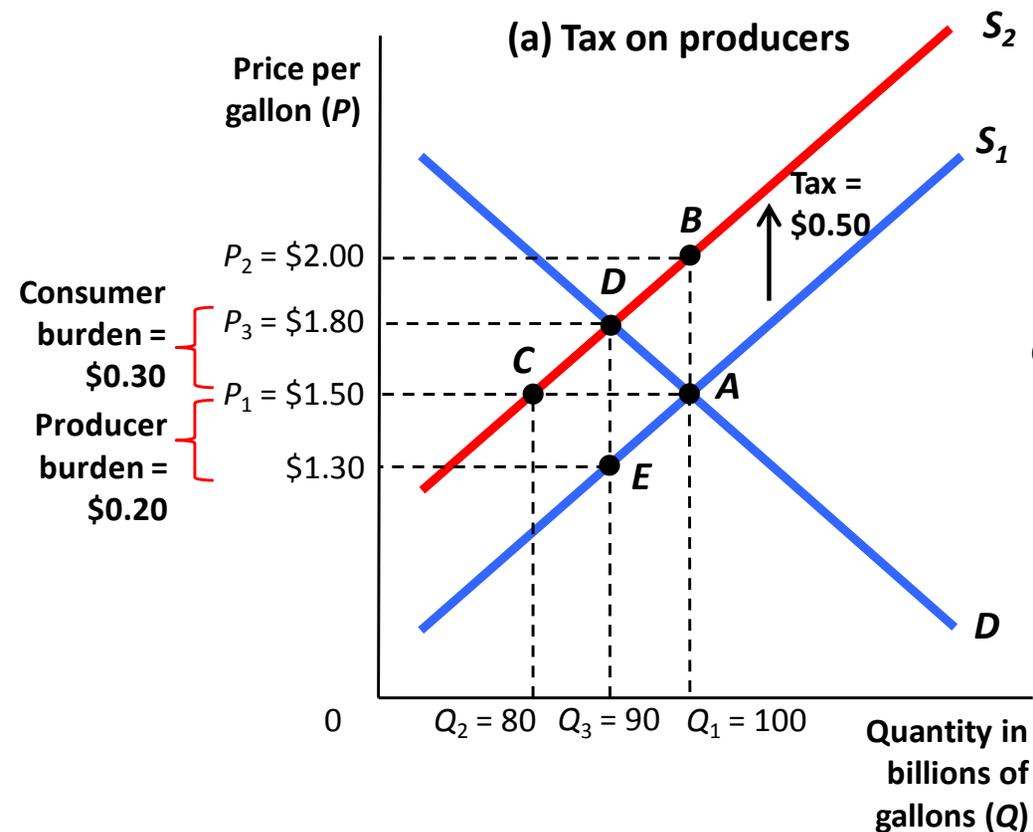
Let p denote the pretax price of x (producer price)

Let $q = p + t$ denote the tax inclusive price of x (consumer price)

Draw graph on blackboard

19.1

The Statutory Burden of a Tax Does Not Describe Who Really Bears the Tax, and Is Irrelevant to the Tax Burden



Computing Burden Shares

For either side of the market: [burden share] = [dollars lost per unit] / [increase in tax per unit]

Consumers: [consumers' dollars lost per unit] = [increase in tax per unit paid by consumers] + [increase in price paid per unit by consumers to producers]

Producers: [producers' dollars lost per unit] = [increase in tax per unit paid by producers] – [increase in price paid per unit by consumers to producers]

Ex. from part (b) in previous slide:

Consumer burden share = $[\$0.50 + (\$1.30 - \$1.50)] / [\$0.50] = 60\%$

Producer burden share = $[\$0.00 - (\$1.30 - \$1.50)] / [\$0.50] = 40\%$

TAX INCIDENCE

Demand for good x is $D(q)$ decreases with $q = p + t$

Supply for good x is $S(p)$ increases with p

Equilibrium condition: $Q = S(p) = D(p + t)$

Start from $t = 0$ and $S(p) = D(p)$. We want to characterize dp/dt : effect of a small tax increase on price, which determines who bears effective burden of tax:

Change dt generates change dp so that equilibrium holds:

$$S(p + dp) = D(p + dp + dt) \Rightarrow$$

$$S(p) + S'(p)dp = D(p) + D'(p)(dp + dt) \Rightarrow$$

$$S'(p)dp = D'(p)(dp + dt) \Rightarrow$$

$$\frac{dp}{dt} = \frac{D'(p)}{S'(p) - D'(p)}$$

TAX INCIDENCE

Useful to use elasticities in economics because elasticities are unit free

Elasticity: percentage change in quantity when price changes by one percent

$\varepsilon_D = \frac{q}{D} \frac{dD}{dq} = \frac{qD'(q)}{D(q)} < 0$ denotes the price elasticity of demand

$\varepsilon_S = \frac{p}{S} \frac{dS}{dp} = \frac{pS'(p)}{S(p)} > 0$ denotes the price elasticity of supply

$$\frac{dp}{dt} = \frac{D'(p)}{S'(p) - D'(p)} = \frac{\varepsilon_D}{\varepsilon_S - \varepsilon_D}$$

$$-1 \leq \frac{dp}{dt} \leq 0 \quad \text{and} \quad 0 \leq \frac{dq}{dt} = 1 + \frac{dp}{dt} \leq 1$$

TAX INCIDENCE

$$\frac{dp}{dt} = \frac{\varepsilon_D}{\varepsilon_S - \varepsilon_D}$$

When do consumers bear the entire burden of the tax? ($dp/dt = 0$ and $dq/dt = 1$)

1) $\varepsilon_D = 0$ [inelastic demand]

example: short-run demand for gasoline inelastic (need to drive to work)

2) $\varepsilon_S = \infty$ [perfectly elastic supply]

example: perfectly competitive industry

When do producers bear the entire burden of the tax? ($dp/dt = -1$ and $dq/dt = 0$)

1) $\varepsilon_S = 0$ [inelastic supply]

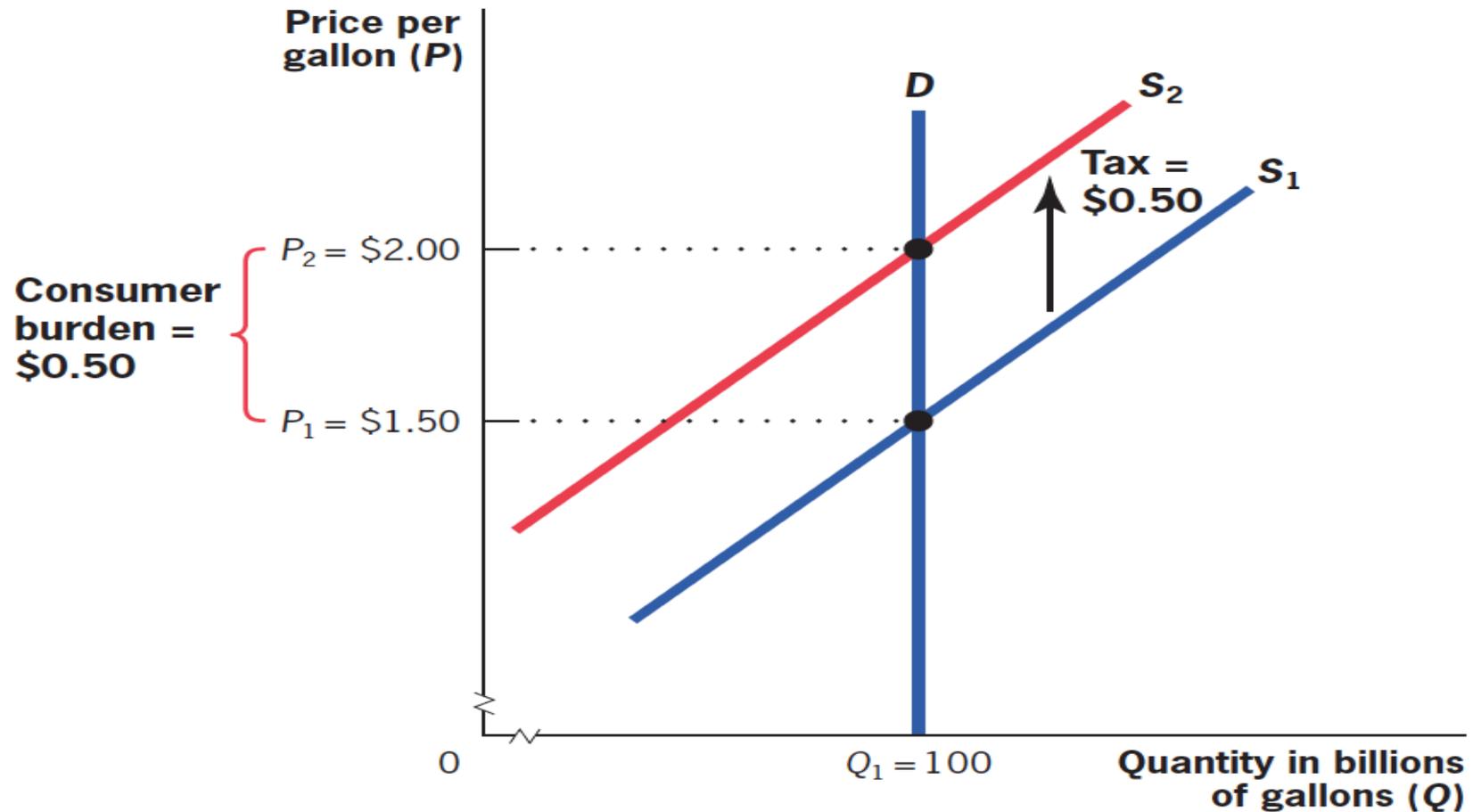
example: fixed quantity supplied

2) $\varepsilon_D = -\infty$ [perfectly elastic demand]

example: there is a close substitute, and demand shifts to this substitute if price changes.

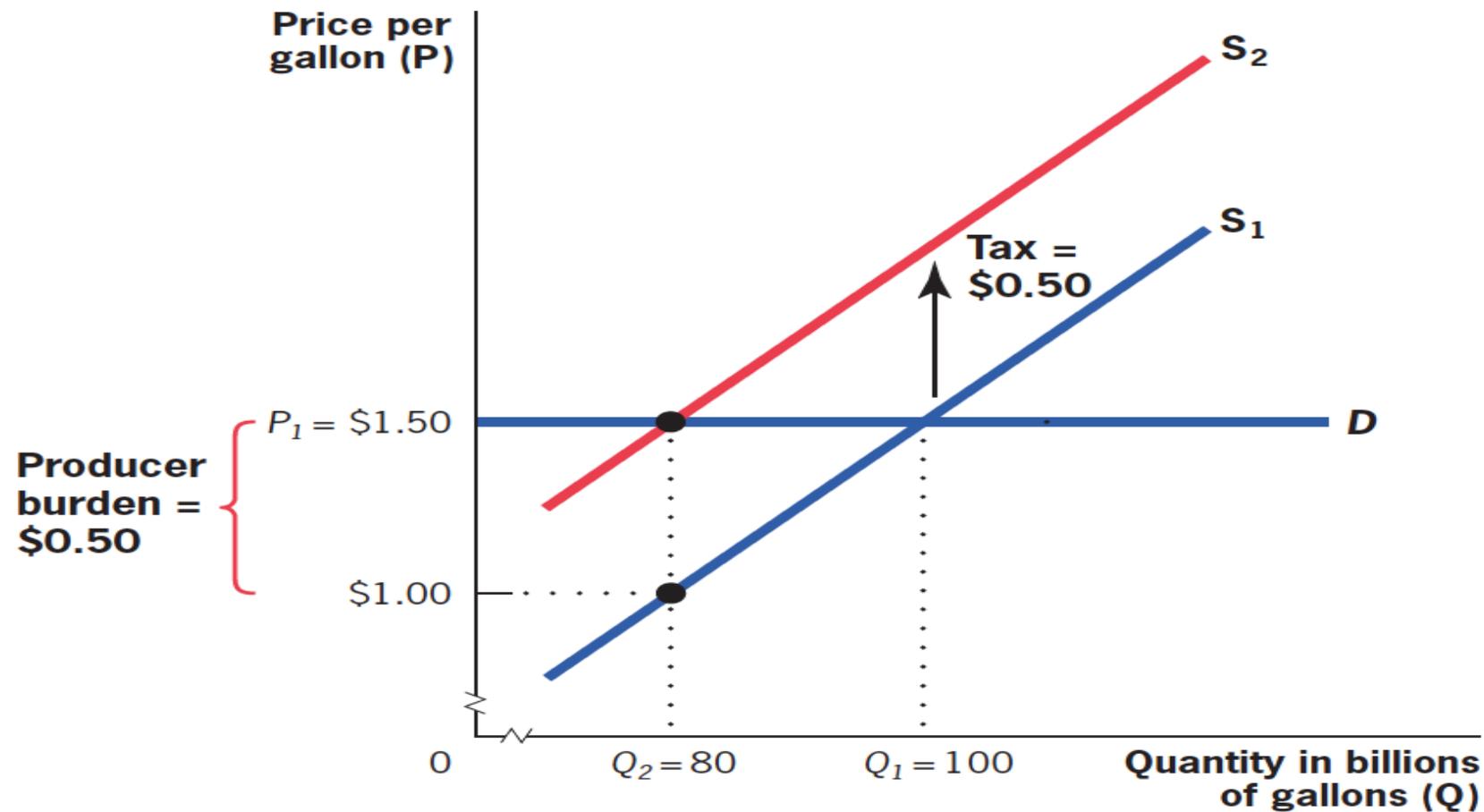
19.1

Perfectly Inelastic Demand



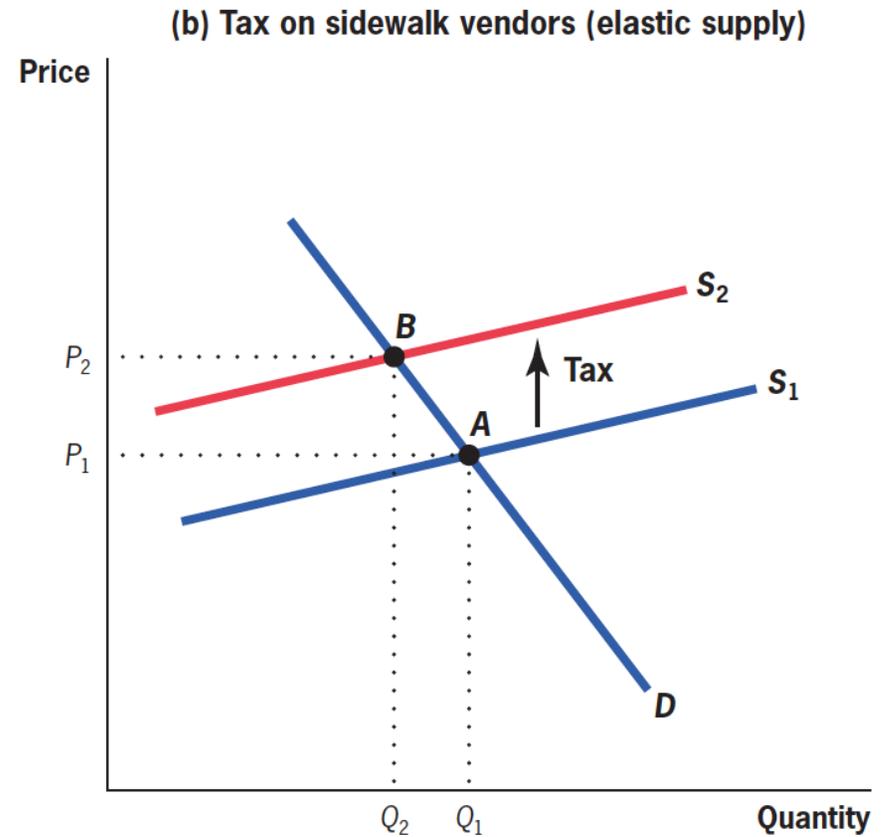
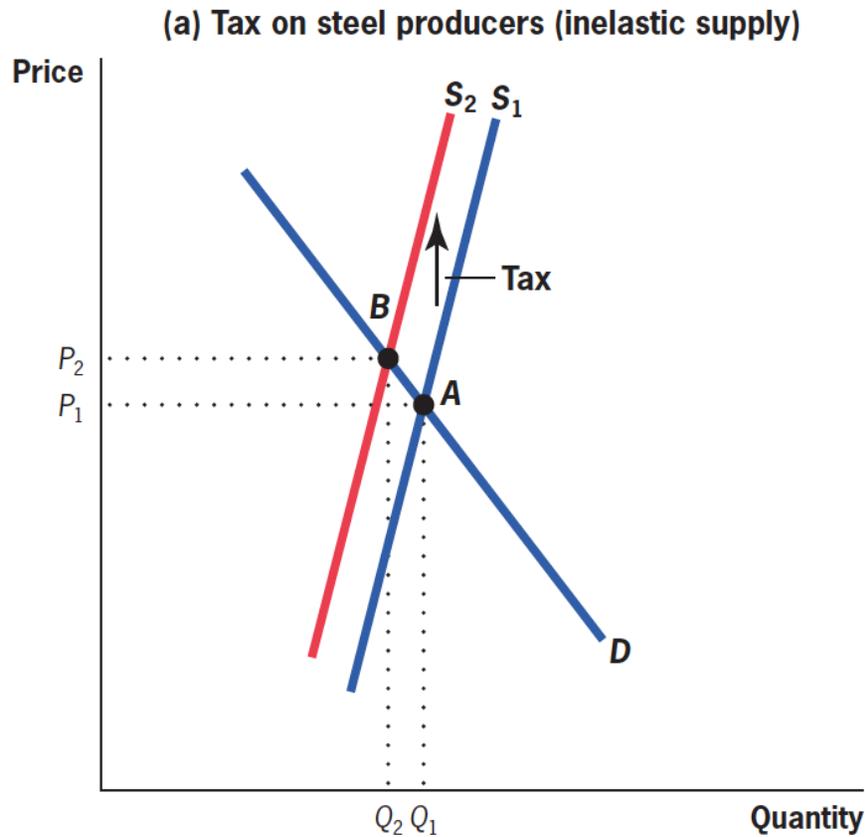
19.1

Perfectly Elastic Demand



19.1

Supply Elasticities



TAX INCIDENCE: KEY RESULTS

- 1) statutory incidence not equal to economic incidence
- 2) equilibrium is independent of who nominally pays the tax
- 3) more inelastic factor bears more of the tax

These are robust conclusions that hold with more complicated models

Efficiency Costs of Taxation

Deadweight burden (also called excess burden) of taxation is defined as the welfare loss (measured in dollars) created by a tax over and above the tax revenue generated by the tax

In the simple supply and demand diagram, welfare is measured by the sum of the consumer surplus and producer surplus

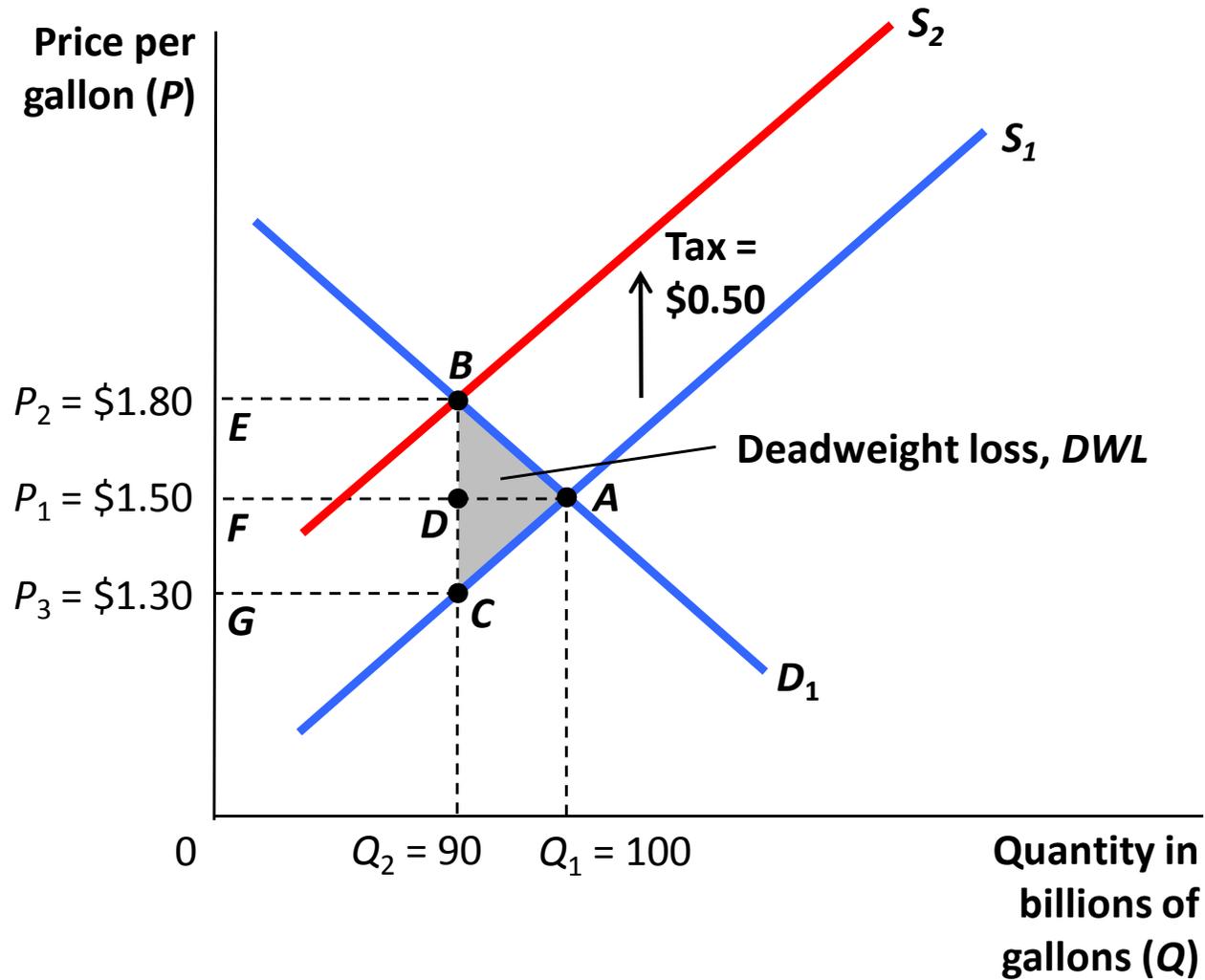
The welfare loss of taxation is measured as change in consumer+producer surplus minus tax collected: it is the triangle on the figure

The inefficiency of any tax is determined by the extent to which consumers and producers change their behavior to avoid the tax; deadweight loss is caused by individuals and firms making inefficient consumption and production choices in order to avoid taxation.

If there is no change in quantities consumed, the tax has no efficiency costs

20.1

Taxation and Economic Efficiency: Graphical Approach



Efficiency Costs of Taxation

Deadweight burden (or deadweight loss) of small tax dt (starting from zero tax) is measured by the **Harberger Triangle**:

$$DWB = \frac{1}{2} (-dQ) \cdot dt = -\frac{1}{2} S'(p) \cdot dp \cdot dt = -\frac{1}{2} \frac{p S'(p)}{S(p)} \cdot \frac{Q}{p} \cdot dp \cdot dt$$

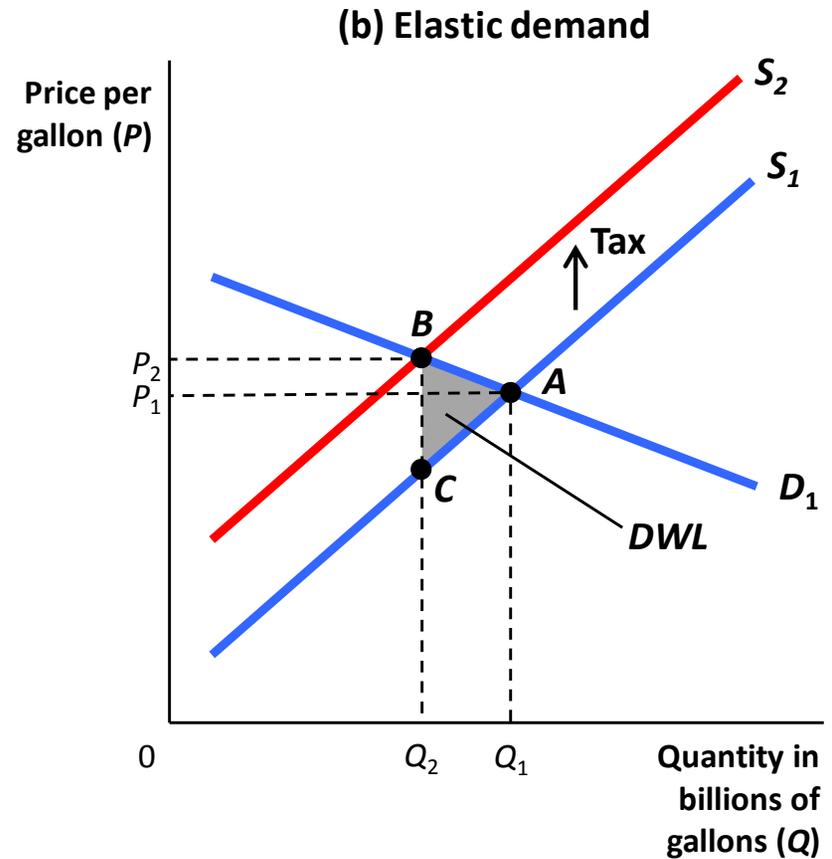
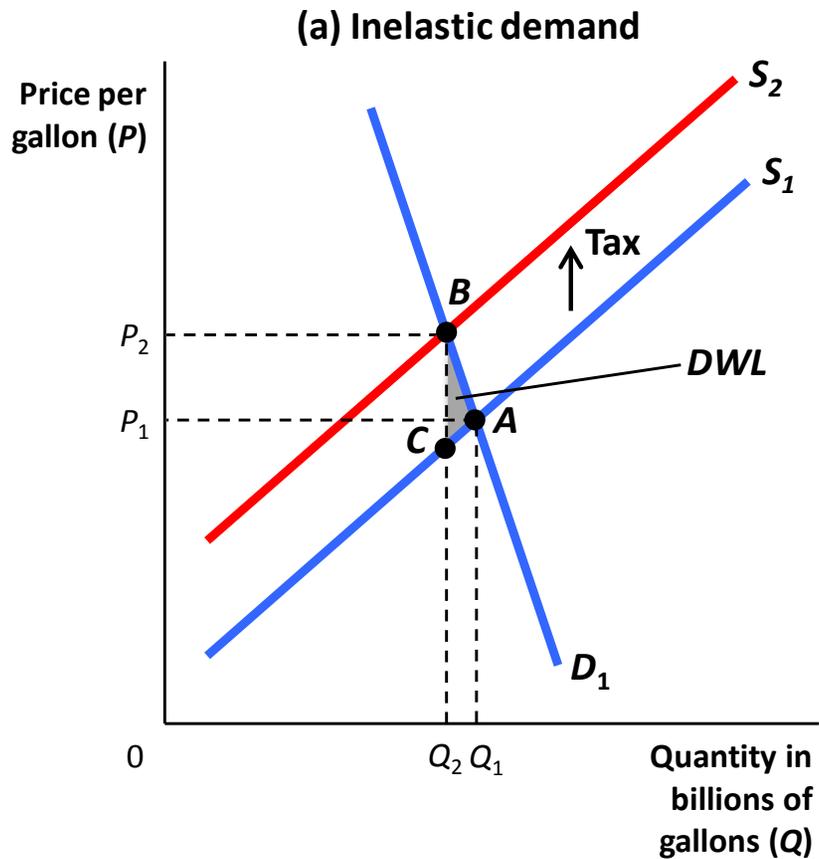
[recall that $Q = S(p)$ and hence $dQ = S'(p)dp$]

Recall that $dp/dt = \varepsilon_D / (\varepsilon_S - \varepsilon_D)$, hence:

$$DWB = -\frac{1}{2} \cdot \frac{\varepsilon_S \cdot \varepsilon_D}{\varepsilon_S - \varepsilon_D} \cdot \frac{Q}{p} (dt)^2$$

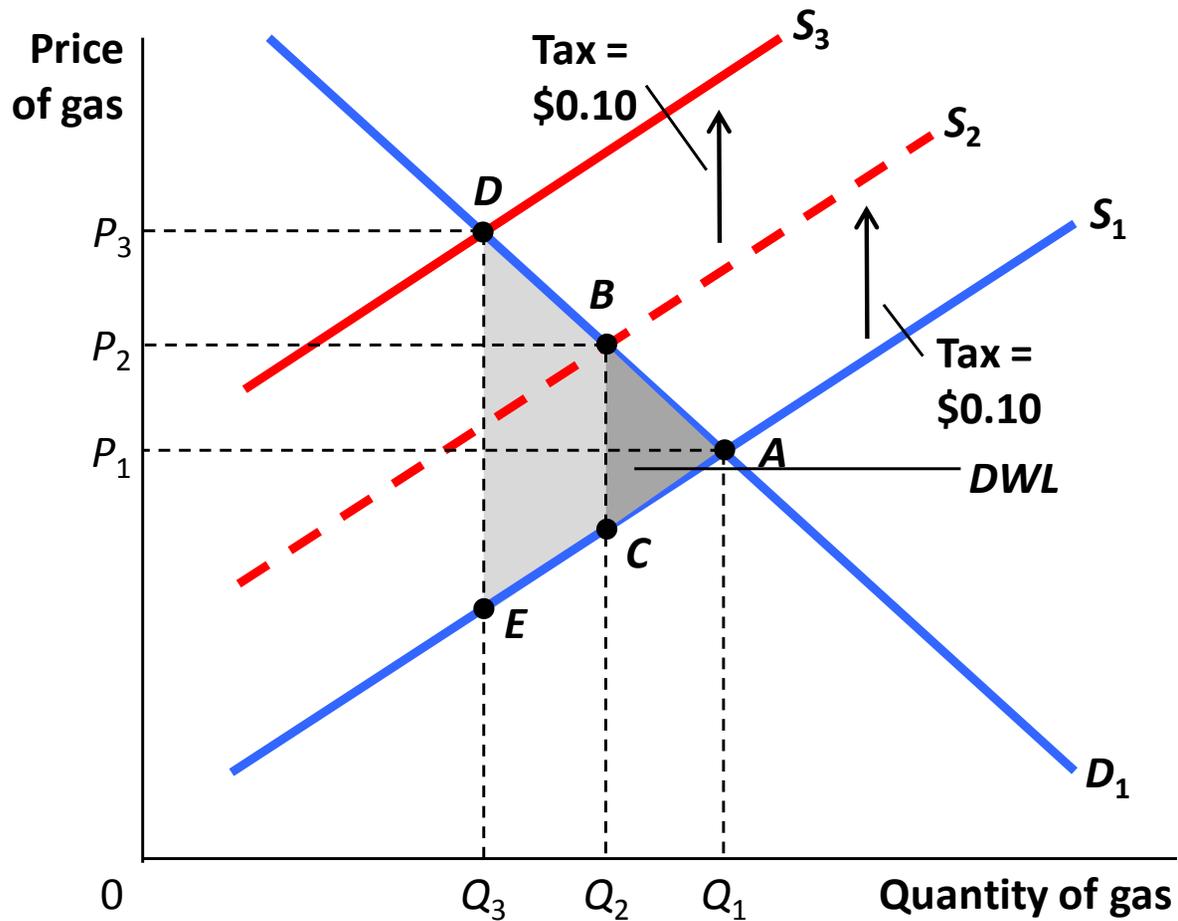
20.1

Elasticities Determine Tax Inefficiency



20.1

Marginal DWL Rises with Tax Rate



Efficiency Costs of Taxation

$$DWB = -\frac{1}{2} \cdot \frac{\varepsilon_S \cdot \varepsilon_D}{\varepsilon_S - \varepsilon_D} \cdot \frac{Q}{p} (dt)^2$$

1) DWB increases with the absolute size of elasticities $\varepsilon_S > 0$ and $-\varepsilon_D > 0$

⇒ More efficient to tax relatively inelastic goods

2) DWB increases with the square of the tax rate t : small taxes have relatively small efficiency costs, large taxes have relatively large efficiency costs

⇒ More efficient to spread taxes across all goods to keep each tax rate low

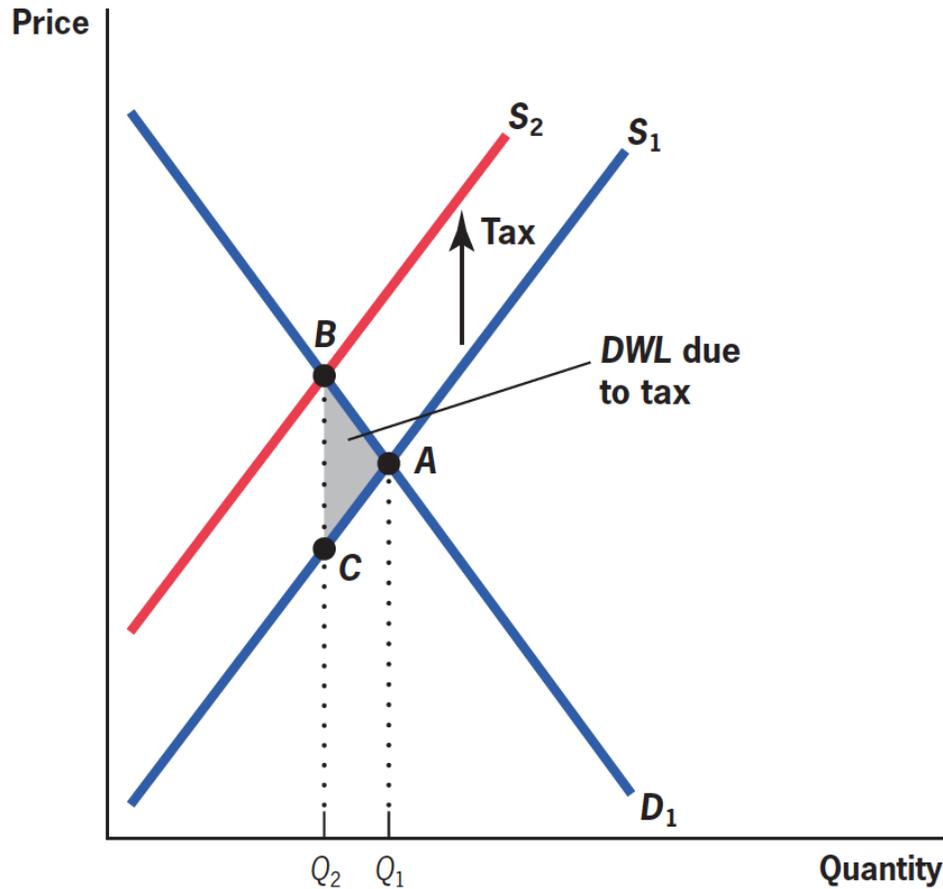
⇒ Better to fund large one time govt expense (such as a war) with debt and repay slowly afterwards than have very high taxes only during war

3) Pre-existing distortions (such as an existing tax or externality) makes the cost of taxation higher: move from the triangle to trapezoid

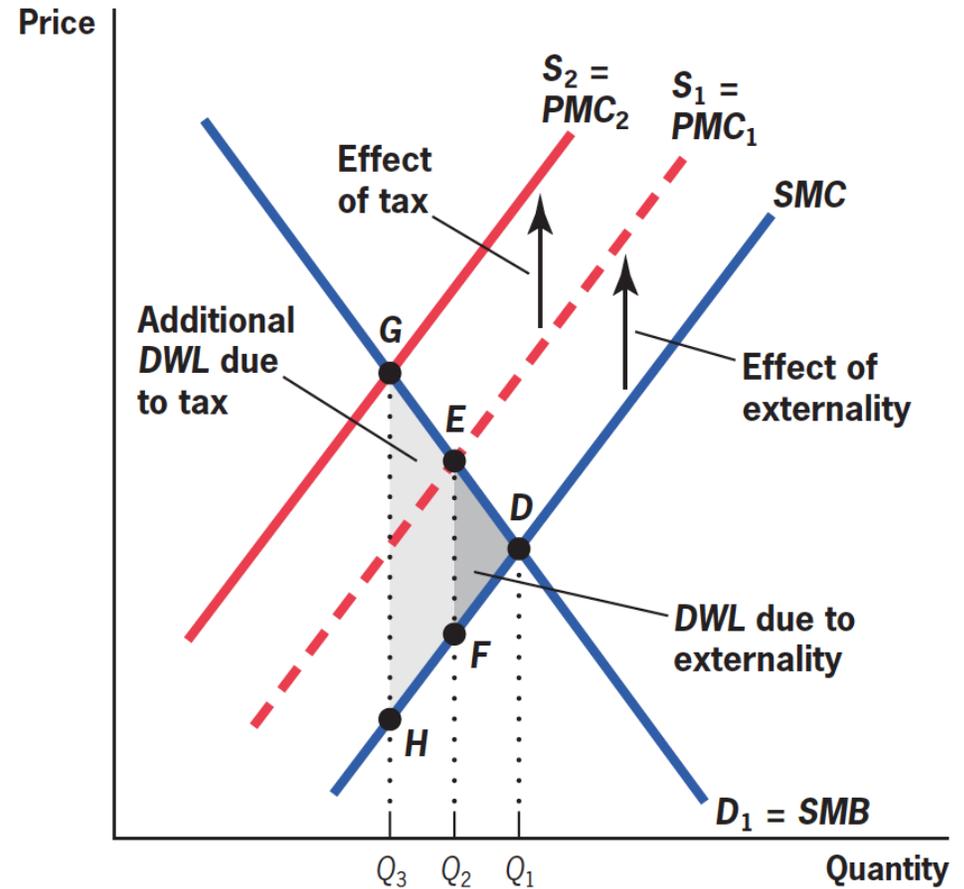
20.1

A Tax System's Efficiency Is Affected by a Market's Preexisting Distortions

(a) No externality



(b) Positive production externality



Application: Optimal Commodity Taxation

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

Consider one consumer who consumes K different goods

What are the tax rates t_1, \dots, t_K of each good that raise a given amount of revenue while minimizing the welfare loss to the individual? (Assume perfectly elastic supply of each good.)

Uniform tax rates $t = t_1 = \dots = t_K$ is not optimal if the individual has more elastic demand for some goods than for others

Optimum is called the **Ramsey tax rule**: optimal tax rates are such that the marginal DWB for last dollar of tax collected is the same across all goods

⇒ Tax more the goods that have inelastic demands [and tax less the goods that have elastic demands]

Note: this abstracts from redistribution and focuses solely on efficiency

Tax Incidence: Empirical Application

Doyle and Sampatharank (2008) study the Gas Tax Holidays in Indiana (IN) and Illinois (IL).

These gas taxes are taxes on producers (statutory incidence is on producers since gas stations pay the government).

Are gas tax cuts passed through to consumers? Or do producers pocket the tax cut and leave consumer price unchanged?

Study this question using state-level gas tax reforms:

Gas prices spike above \$2.00 in 2000

IN suspends 5% gas tax on July 1. Reinstated on Oct 30.

IL suspends 5% gas tax on July 1. Reinstated on Dec 31.

Tax Incidence: Empirical Application

Empirical approach in paper: difference-in-difference (DD), compare treated states with neighboring states (MI, OH, MO, IA, WI) before and after tax change

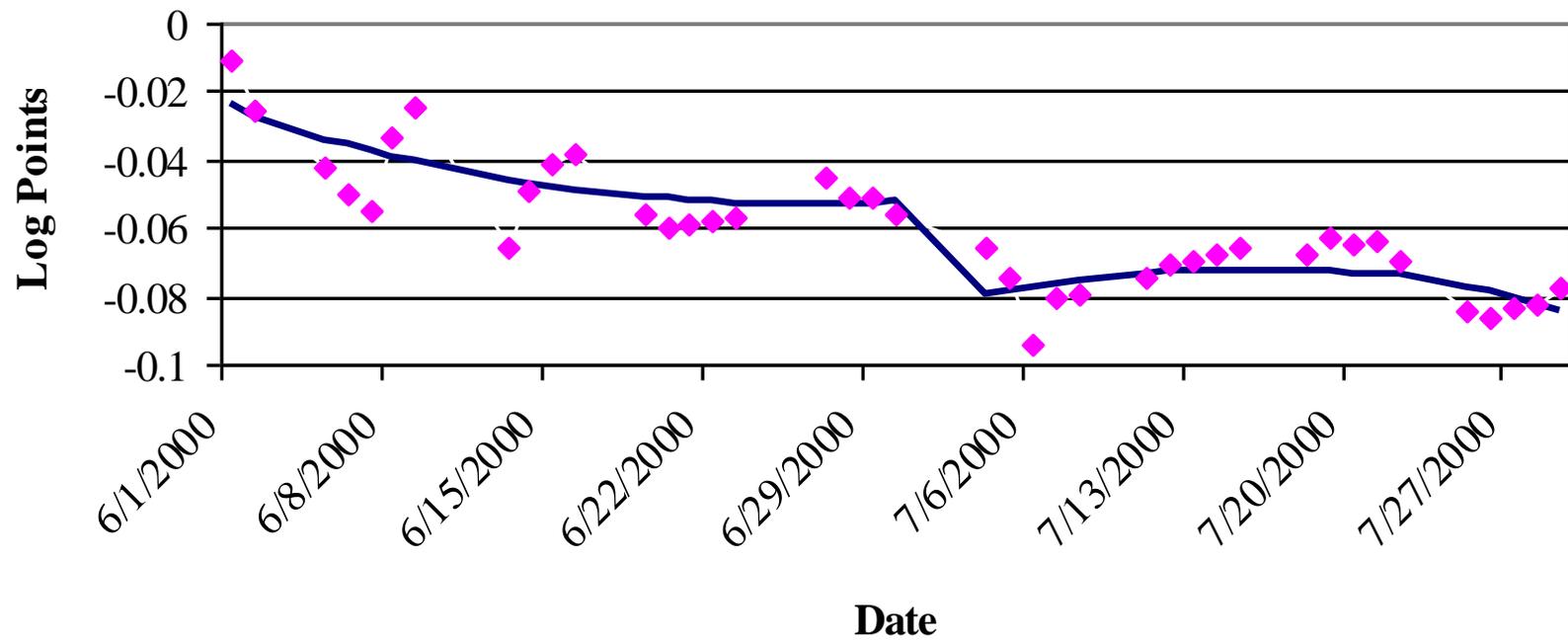
Graphical evidence is most transparent. Findings:

1) 10 cent decrease in gas tax \Rightarrow 7 cent decrease in price paid by consumers

2) So consumers bear 70% of the burden of the gas tax: $([\$0.00] + [-\$0.07]) / [-\$0.10] = 70\%$ (applying formula from “Computing Burden Share” slide)

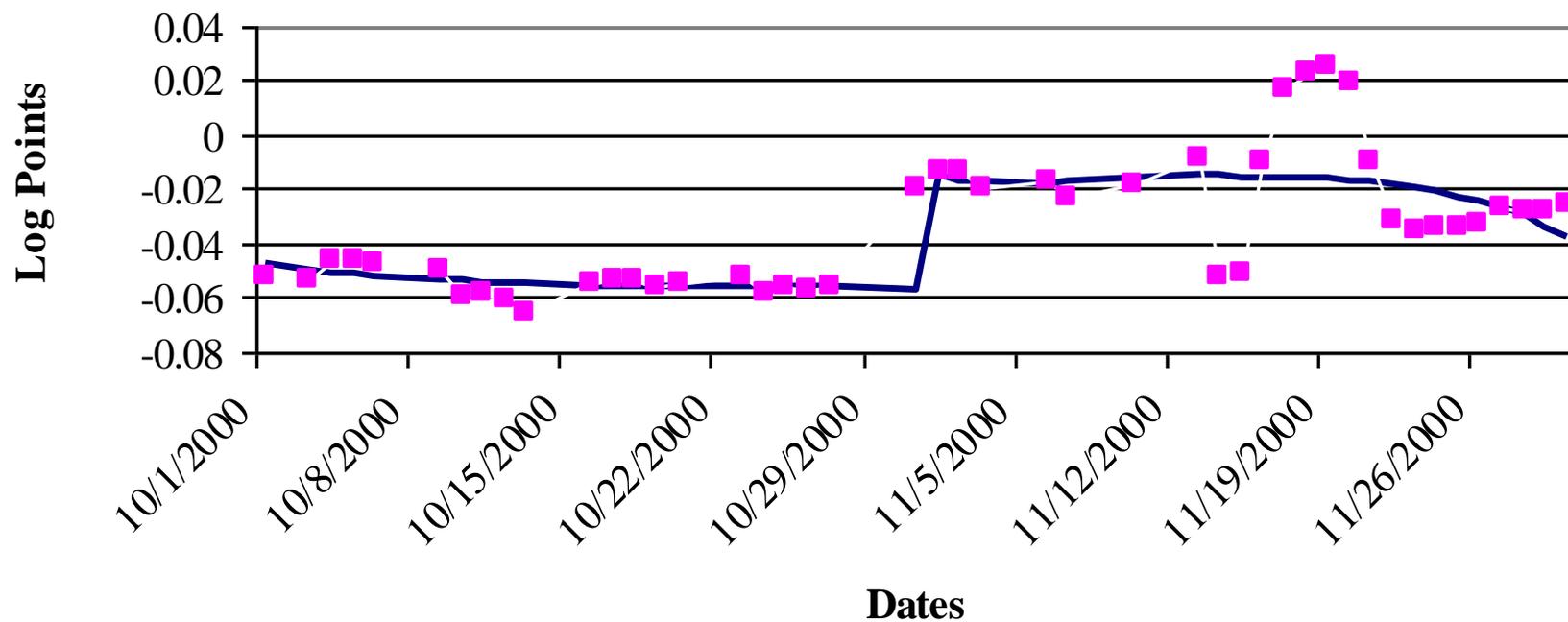
(Note: Consumers bearing 70% of gas tax burden means that consumers get 70% of the benefit of a gas tax cut.)

Figure 2A: Summer 2000 Difference in Log Gas Prices
IL/IN vs. Neighboring States: MI, OH, MO, IA, WI



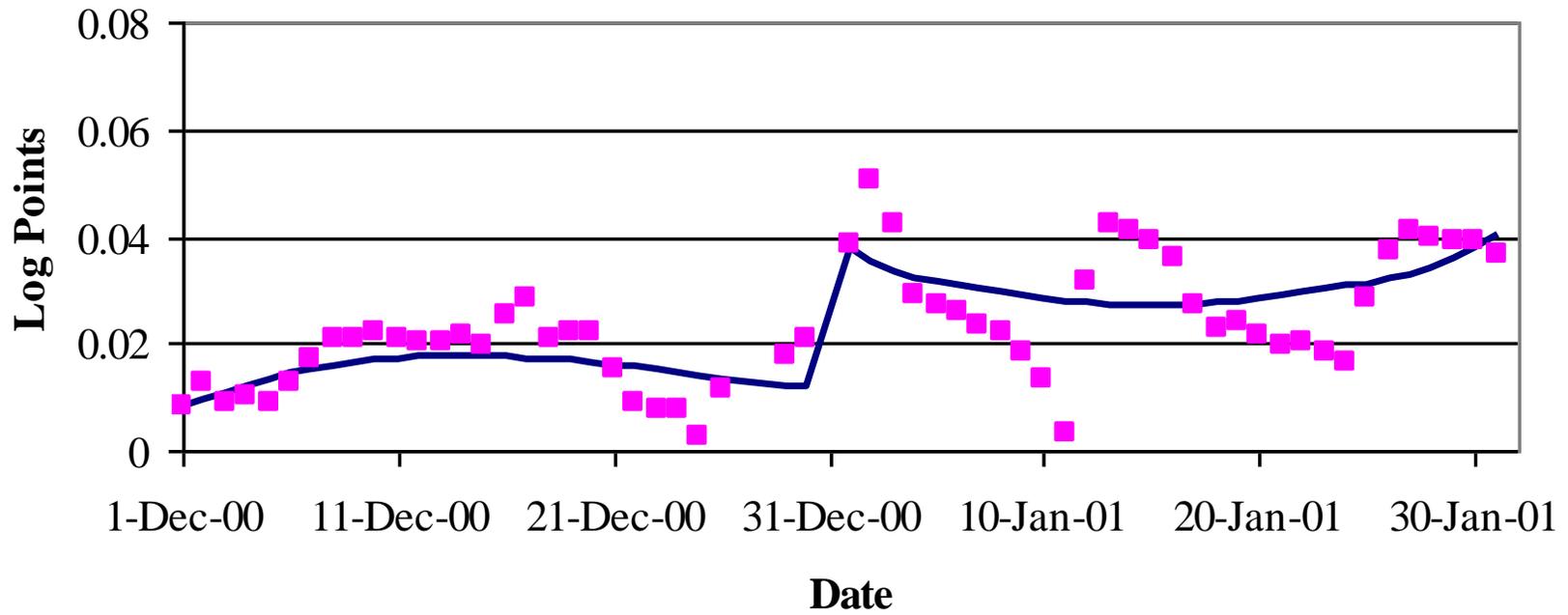
Source: Doyle and Samphantharak 2008.

Figure 2B: Fall 2000 Difference in Log Gas Prices
IN vs. Neighboring States: MI, OH, IL



Source: Doyle and Samphantharak 2008.

Figure 2C: Winter 2000/2001 Difference in Log Gas Prices
IL vs. Neighboring States: MO, IA, WI, IN



Source: Doyle and Samphantharak 2008.

EVIDENCE: The Incidence of Excise Taxation

- Excises tax on cigarettes varies widely across the United States.
 - Low of \$0.025/pack per pack in VA.
 - High of \$1.51/pack in CT and MA.
 - Since 1990, NJ increased its tax rate nearly sixfold.
 - Arizona has increased its tax nearly eightfold.
- Many studies examine how taxes affect prices.
- These studies uniformly conclude that the price of cigarettes rises by the full amount of the excise tax.

General Equilibrium Tax Incidence

Examples so far have focused on **partial equilibrium** incidence which considers impact of a tax on one market in isolation

General equilibrium models consider the effects on related markets of a tax imposed on one market

E.g. imposition of a tax on cars may reduce demand for steel
⇒ additional effects on prices in equilibrium beyond car market.

General Equilibrium Tax Incidence: Example: Soda Tax in Berkeley

Consider the market for Soda in Berkeley

Berkeley imposes a Soda tax (voted in 2014)

Who bears the incidence?

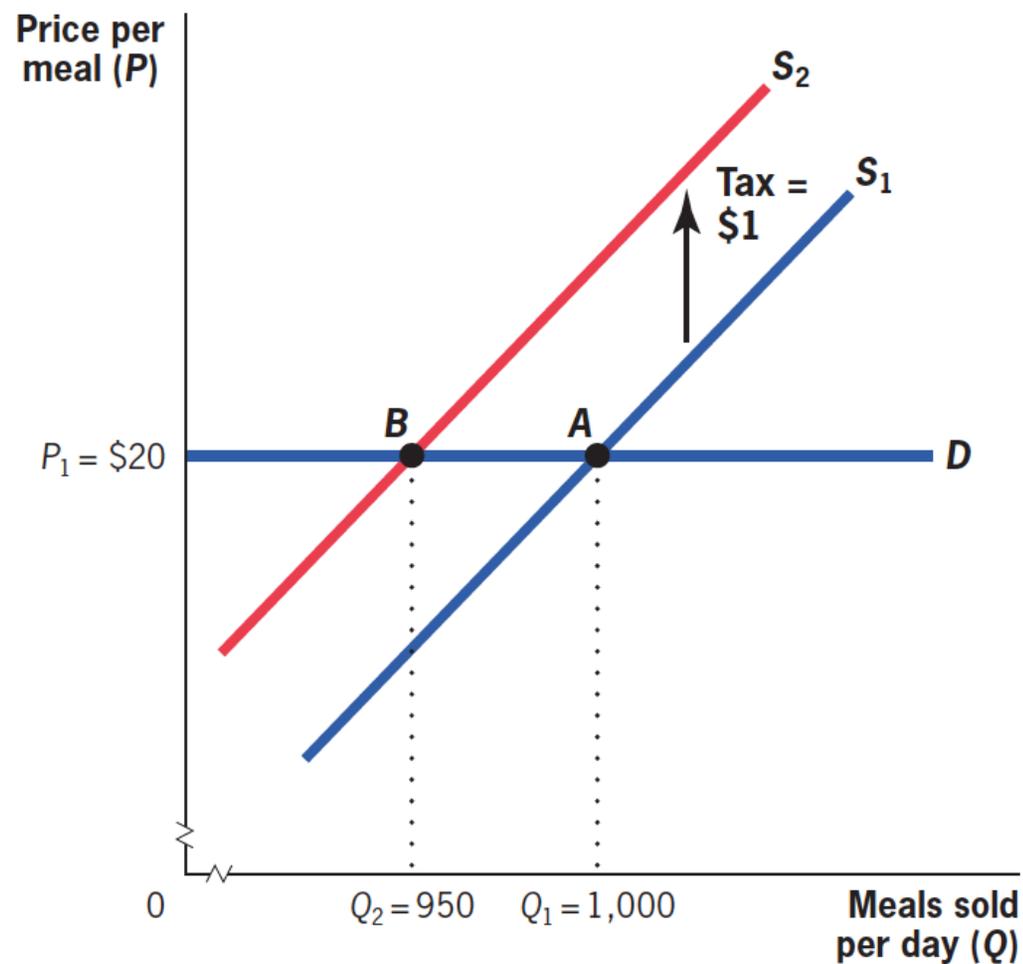
If soda demand is inelastic, then consumers bear burden

Demand for Soda in Berkeley is likely to be elastic: if price of Soda in Berkeley goes up, you consume less Soda [intention of the tax] or buy Soda in Oakland

Consider extreme case of perfectly elastic demand

19.3

Effects of a Restaurant Tax: A General Equilibrium Example



General Equilibrium Tax Incidence: Example: Soda Tax

If Soda demand perfectly elastic then:

1) Soda sellers (supermarkets, restaurants) bear the full burden of the tax.

2) But Soda sellers are not self-contained entities

Companies are just a technology for combining capital and labor to produce an output.

Capital: land, physical inputs like building, kitchen equipment, etc.

Labor: cashier staff, cooks, waitstaff, etc.

3) Ultimately, these two factors (capital or labor) must bear the loss in profits due to the tax [if consumer demand is perfectly elastic]

General Equilibrium Tax Incidence: Example: Soda Tax

Incidence is “shifted backward” to capital and labor.

Assume that labor supply is perfectly elastic because cashiers can always go and work in Oakland if they get paid less in Berkeley

Capital, in contrast, is perfectly inelastic in short-run: you cannot pick up the shop and move it in the short run.

In short run, capital bears tax because it is completely inelastic
⇒ Soda tax owners lose (not consumers or workers)

In the longer-run, the supply of capital is also likely to be highly elastic: Investors can close or sell the shop, take their money, and invest it elsewhere.

General Equilibrium Tax Incidence: Long-run effects

If both labor and capital are highly elastic in the long run, who bears the tax?

The one additional inelastic factor is land.

The supply is clearly fixed.

When both labor and capital can avoid the tax, the only way Soda sellers will remain in Berkeley is if they pay a lower rent on their land.

⇒ Soda tax ends up hurting Berkeley landowners in general equilibrium [if Soda demand, labor and capital are fully elastic]

This is of course an idealized example, in practice, demand, labor, and capital are not fully elastic

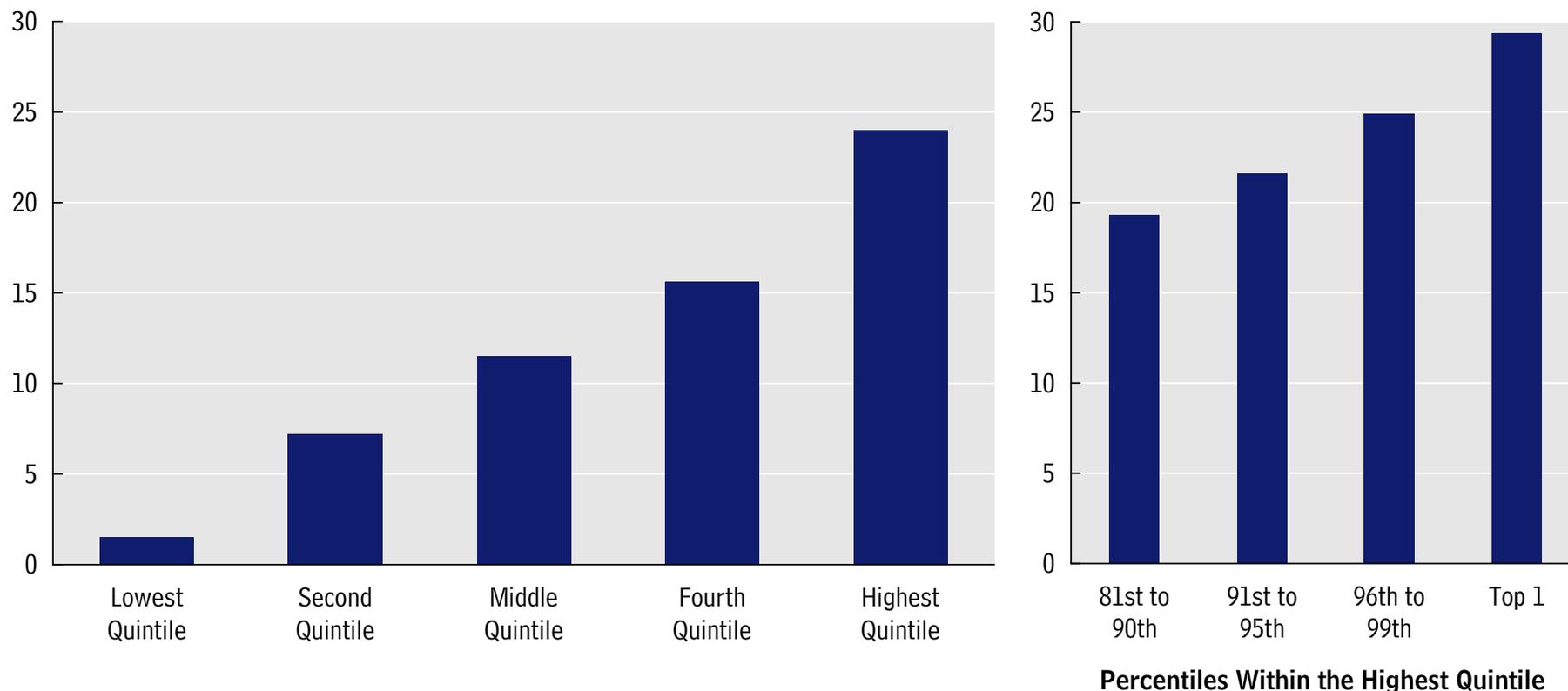
CBO INCIDENCE ASSUMPTIONS

The Congressional Budget Office (CBO) analysis considers the incidence of the full set of taxes levied by the federal government. Their key assumptions follow:

1. **Individual Income taxes** are borne fully by the households that pay them.
2. **Payroll taxes** are borne fully by workers, regardless of whether these taxes are paid by the workers or by the firm.
3. **Excise taxes** are fully shifted to prices and so are borne by individuals in proportion to their consumption of the taxed item.
4. **Corporate taxes** are allocated 75% to owners of capital (not only shareholders but owners of capital in general) in proportion to capital income and 25% to labor in proportion to labor income [controversial]

Figure 1.**Average Federal Tax Rates, by Income Group, 2010**

(Percent)



Source: Congressional Budget Office.

Notes: Average federal tax rates are calculated by dividing federal tax liabilities by before-tax income.

Before-tax income is the sum of market income and government transfers. Market income is composed of labor income, business income, capital gains, capital income (excluding capital gains), income received in retirement for past services, and other sources of income. Government transfers are cash payments and in-kind benefits from social insurance and other government assistance programs.

Federal taxes include individual and corporate income taxes, social insurance (or payroll) taxes, and excise taxes.

Income groups are created by ranking households by before-tax income. Quintiles (fifths) contain equal numbers of people; percentiles (hundredths) contain equal numbers of people as well.

Source: Congressional Budget Office (CBO) 2013

DECEMBER 2013

THE DISTRIBUTION OF HOUSEHOLD INCOME AND FEDERAL TAXES, 2010

Table 2.

Average Federal Tax Rates, by Income Group, 2009 and 2010

(Percentage of before-tax income)

Income Group	All Federal Taxes	Individual Income Taxes	Social Insurance Taxes	Corporate Income Taxes	Excise Taxes
			2010		
Lowest Quintile	1.5	-9.2	8.4	0.7	1.6
Second Quintile	7.2	-2.3	7.8	0.7	1.0
Middle Quintile	11.5	1.6	8.3	0.8	0.8
Fourth Quintile	15.6	5.0	9.0	1.0	0.6
Highest Quintile	24.0	13.8	6.7	3.1	0.4
All Quintiles	18.1	7.7	7.7	2.1	0.6
81st to 90th Percentiles	19.3	8.1	9.4	1.2	0.5
91st to 95th Percentiles	21.6	10.7	8.9	1.5	0.4
96th to 99th Percentiles	24.9	15.1	7.1	2.3	0.4
Top 1 Percent	29.4	20.1	2.2	6.9	0.2

INCIDENCE OF FEDERAL TAXES

1. **Individual Income taxes** is progressive due to tax credits for low earners and progressive tax brackets
2. **Payroll taxes** are a constant tax rate of 15% but only up to \$120K of earnings \Rightarrow Regressive at the top
3. **Excise taxes** are regressive because share of income devoted to consumption of goods with excise tax (alcohol, tobacco, gas) falls with income
4. **Corporate taxes** are progressive because capital income is highly concentrated

State+local taxes are less progressive than Federal taxes

Mandated Benefits

Now consider incidence of a mandated benefit instead of a tax

Examples: (a) requirement that employers pay for healthcare (employers with 50+ employees will be required to do that with new Obamacare law or pay a fine), (b) workers compensation benefits [for injuries on the job]

Affects firms like a tax

But effect of mandated benefits on equilibrium wages and employment differ from a tax (Summers 1989) because workers value the mandated benefit

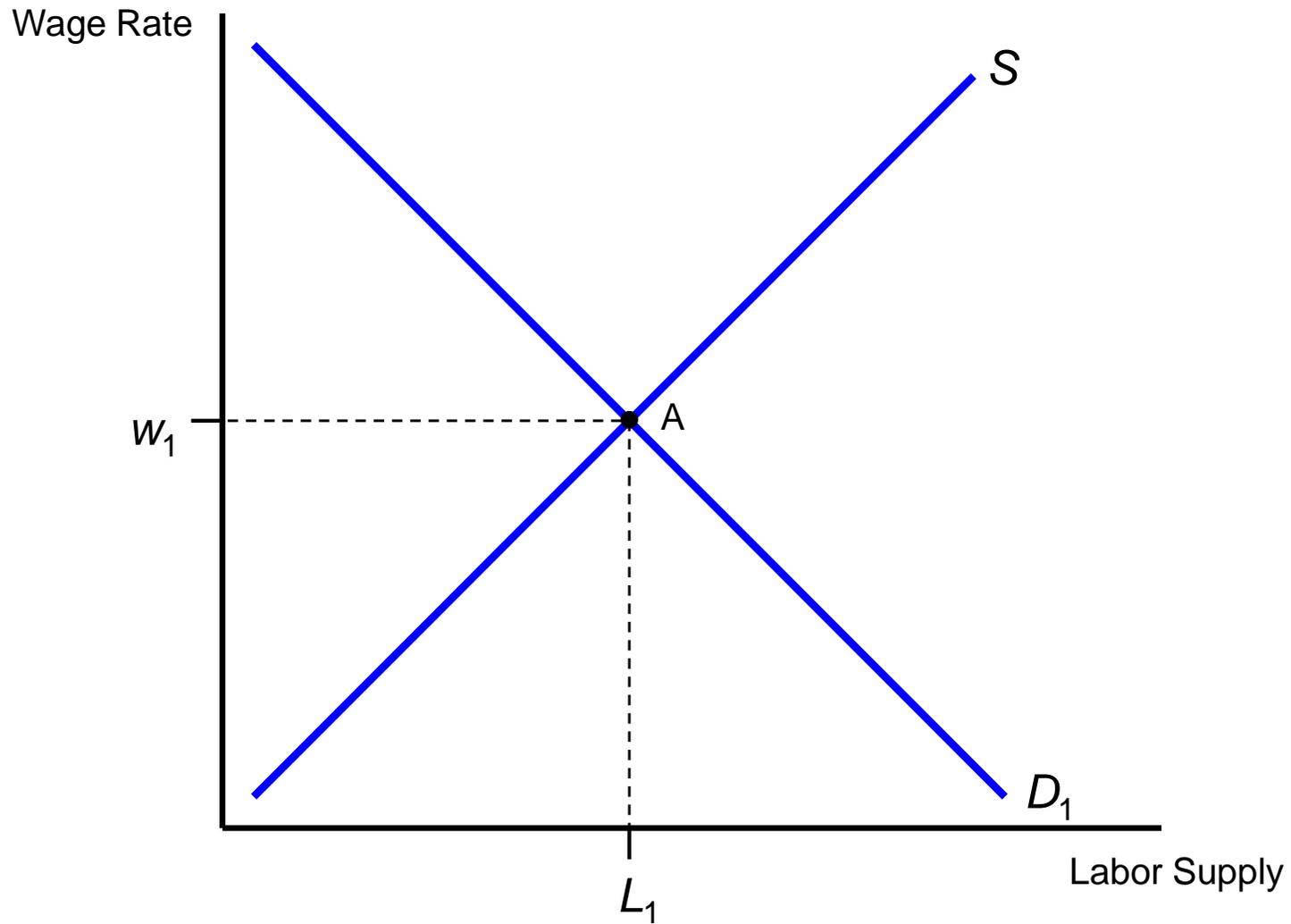
Suppose workers value \$1 of mandated benefit at \$ $\alpha \geq 0$

Could have $\alpha < 1$ if benefit not as valuable as cash

Could have $\alpha > 1$ if benefit more valuable than cash (e.g., can't buy health insurance on individual market)

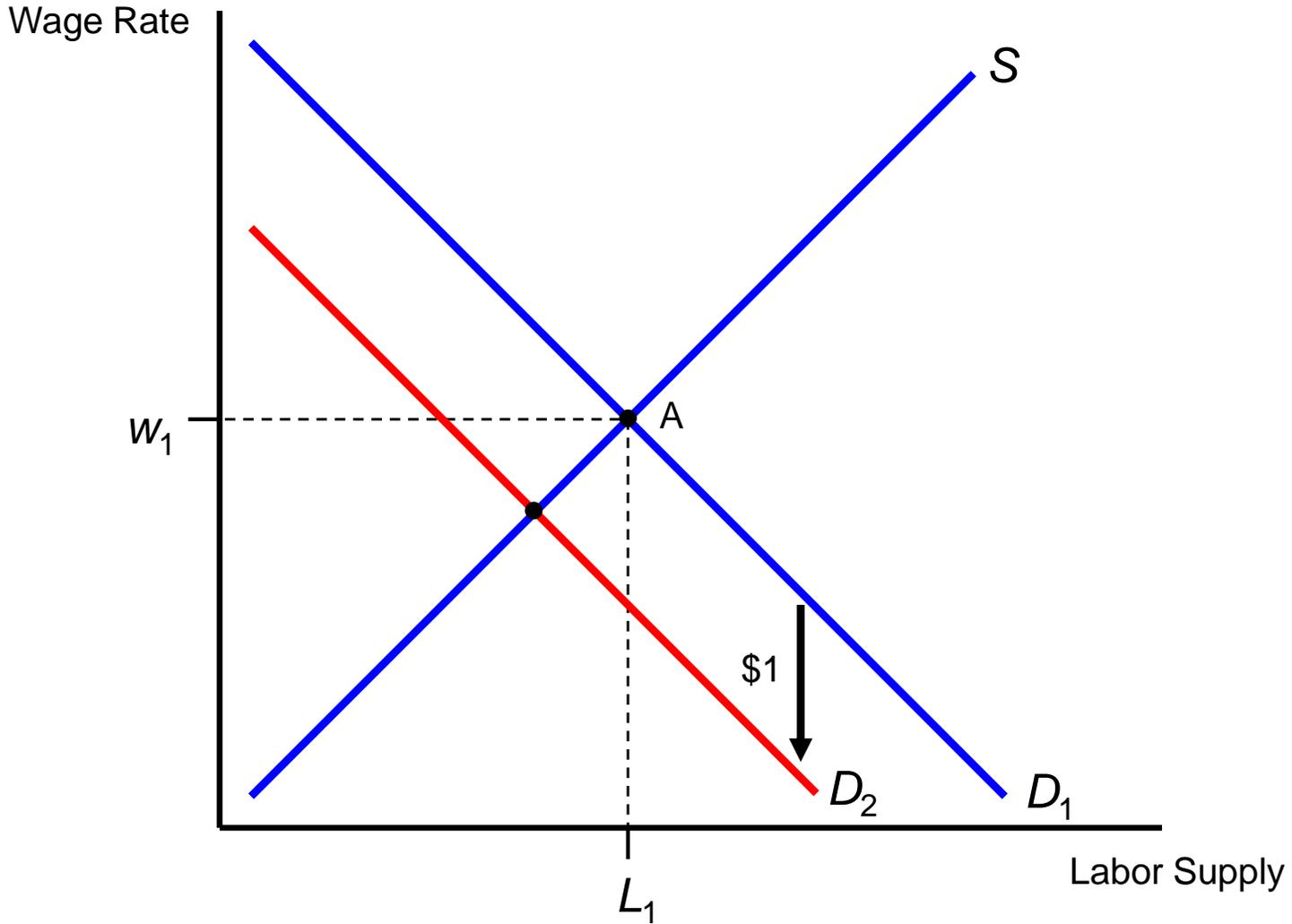
If $\alpha = 1$ then no change in employment

Figure 1: Mandated Benefit



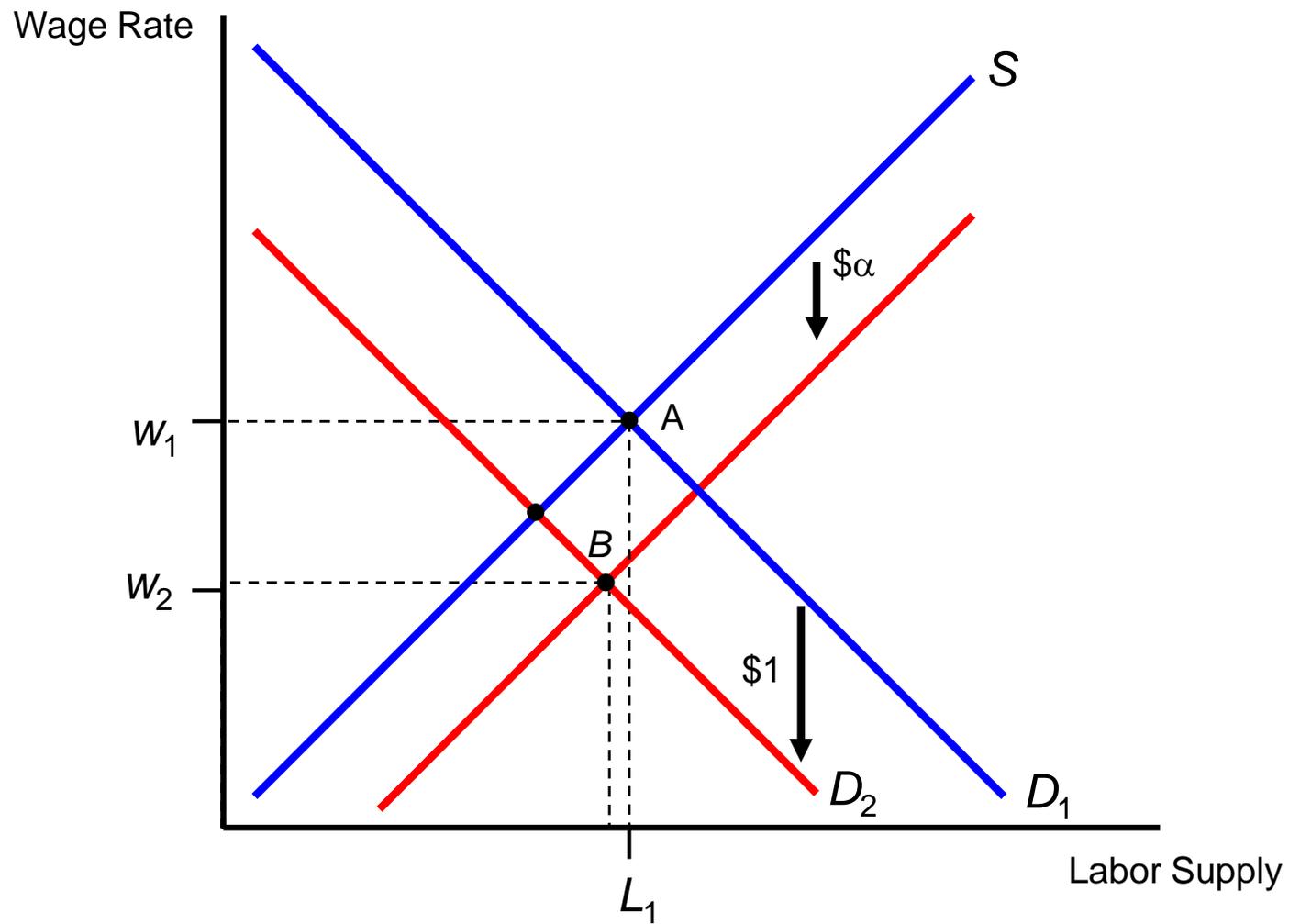
Source: Raj Chetty's Undergraduate Lecture Note

Figure 1: Mandated Benefit



Source: Raj Chetty's Undergraduate Lecture Note

Figure 1: Mandated Benefit



Source: Raj Chetty's Undergraduate Lecture Note

Tax Salience: A New Theory

Traditional model assumes that all individuals are fully aware of taxes that they pay

Is this true in practice? May not be because (unlike gas tax) many taxes are not fully salient.

Do you know your exact marginal income tax rate? Do you think about it when choosing a job?

Do you know the sales tax you have to pay in addition to posted prices at cash register?

Chetty, Looney, Kroft AER '09: test this assumption in the context of commodity taxes and develop a theory of taxation with inattentive consumers

Tax Salience: A New Theory

Chetty, Looney, Kroft AER'09 develop two empirical strategies to test whether salience matters for sales tax incidence

Sales tax is paid at the cash register and not displayed on price tags in stores

1) **Randomized field experiment** with supermarket stores

In one treatment store: they display new price tags showing the level of sales tax and total price was displayed on a subset of products

Compare shopping behavior in for treated products vs control products in treated store, before and after new tags are implemented (this is called difference-in-difference [DD] strategy)

Repeat the analysis in control stores as a placebo DD strategy

2) **Policy experiment** using variation in beer excise and sales taxes across states

Excise tax is salient because built into posted price while sales tax is not salient because it is not included in posted price



Orig. Tag

Exp. Tag

Effect of Posting Tax-Inclusive Prices: Mean Quantity Sold

TREATMENT STORE			
Period	<u>Control Categories</u>	<u>Treated Categories</u>	<u>Difference</u>
Baseline	26.48 (0.22)	25.17 (0.37)	-1.31 (0.43)
Experiment	27.32 (0.87)	23.87 (1.02)	-3.45 (0.64)
Difference over time	0.84 (0.75)	-1.30 (0.92)	DD_{TS} = -2.14 (0.64)

CONTROL STORES			
Period	<u>Control Categories</u>	<u>Treated Categories</u>	<u>Difference</u>
Baseline	30.57 (0.24)	27.94 (0.30)	-2.63 (0.32)
Experiment	30.76 (0.72)	28.19 (1.06)	-2.57 (1.09)
Difference over time	0.19 (0.64)	0.25 (0.92)	DD_{CS} = 0.06 (0.90)

DDD Estimate **-2.20**
(0.58)

Figure 2a

Per Capita Beer Consumption and State Beer Excise Taxes

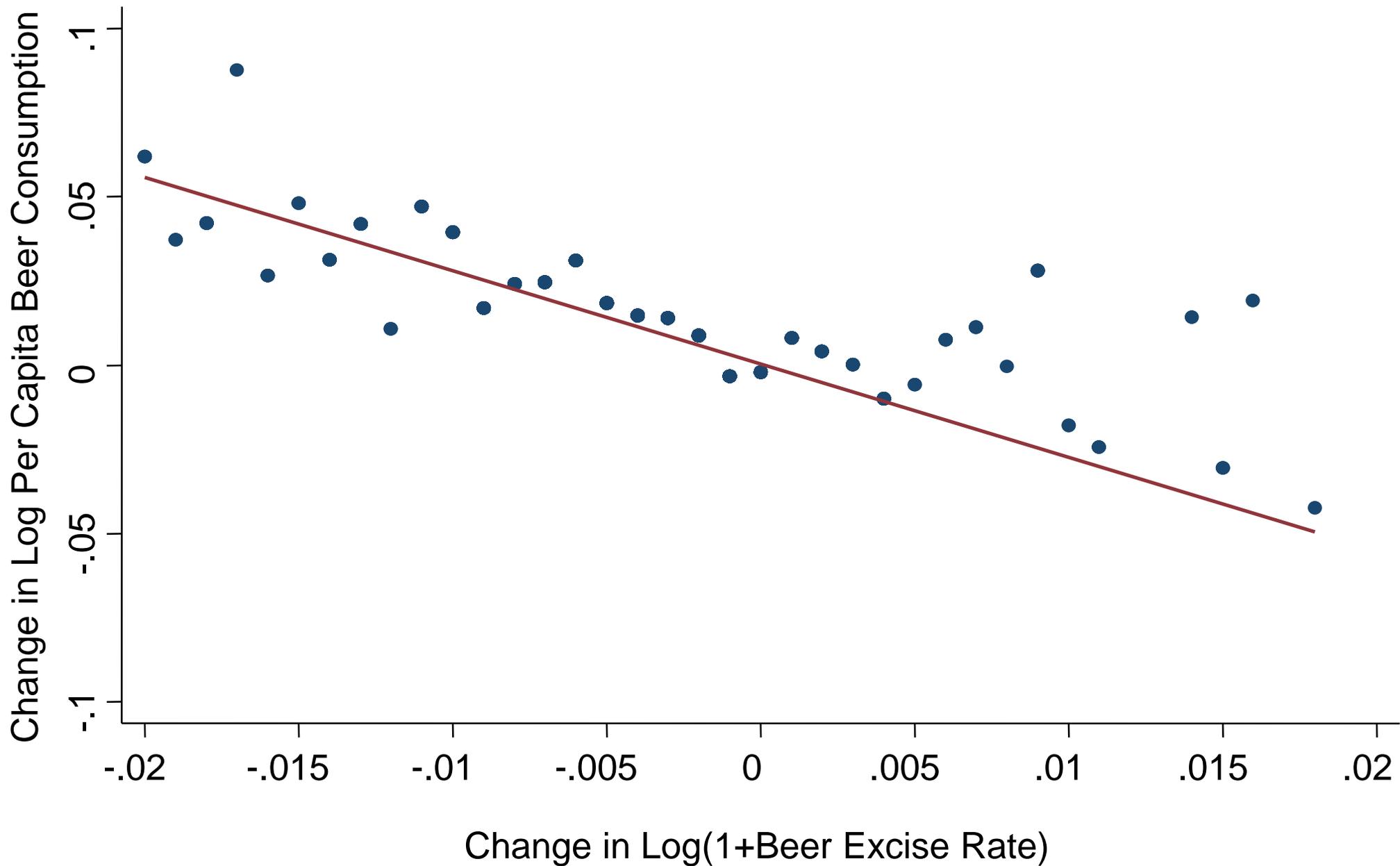
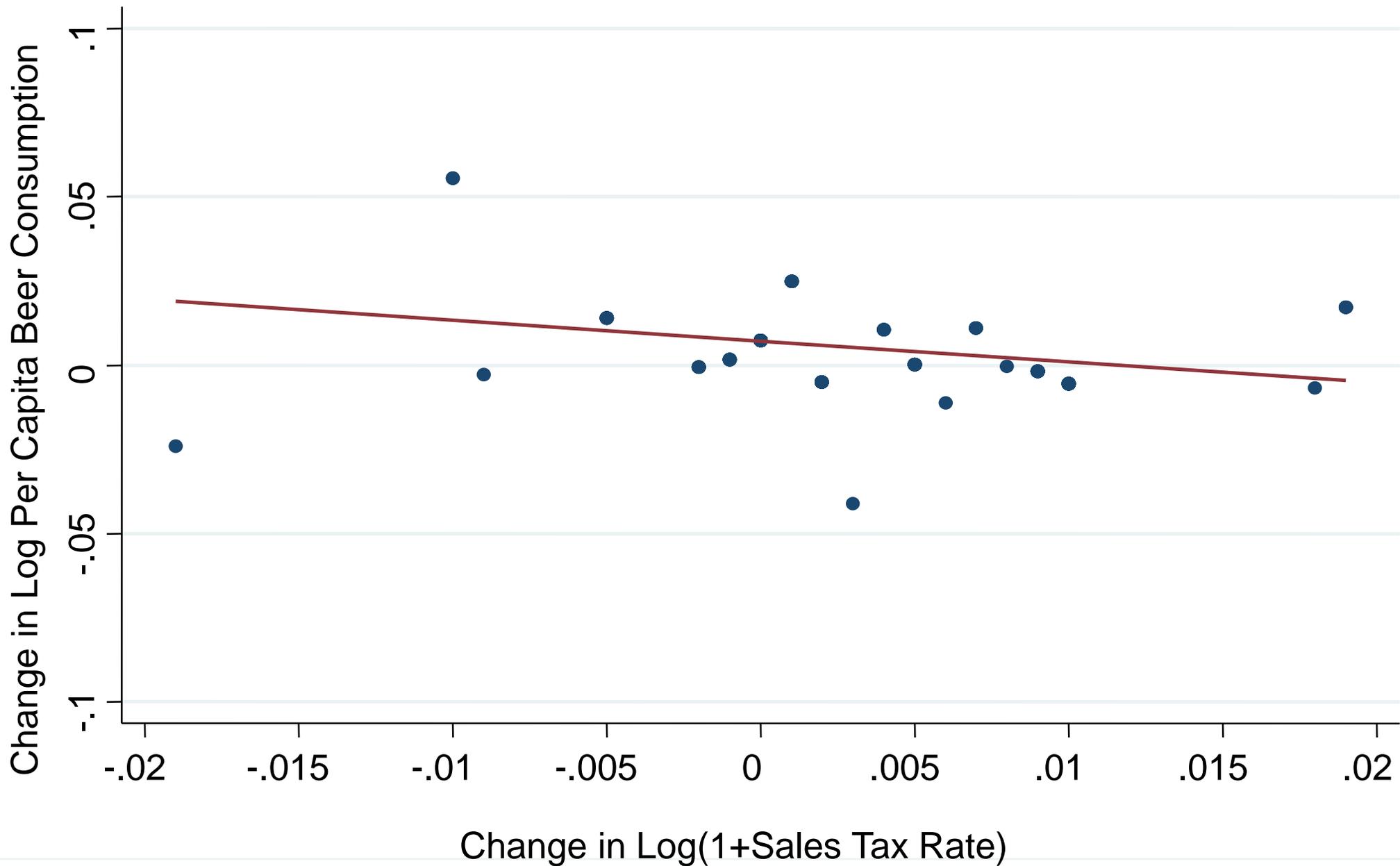


Figure 2b

Per Capita Beer Consumption and State Sales Taxes



Effect of Excise and Sales Taxes on Beer Consumption

Dependent Variable: Change in Log(per capita beer consumption)

	Baseline (1)	Bus Cyc, Alc Regs. (2)	3-Year Diffs (3)	Food Exempt (4)
$\Delta\text{Log}(1+\text{Excise Tax Rate})$	-0.87 (0.17) ^{***}	-0.89 (0.17) ^{***}	-1.11 (0.46) ^{**}	-0.91 (0.22) ^{***}
$\Delta\text{Log}(1+\text{Sales Tax Rate})$	-0.20 (0.30)	-0.02 (0.30)	-0.00 (0.32)	-0.14 (0.30)
Business Cycle Controls		x	x	x
Alcohol Regulation Controls		x	x	x
Year Fixed Effects	x	x	x	x
F-Test for Equality of Coeffs.	0.05	0.01	0.05	0.04
Sample Size	1,607	1,487	1,389	937

Note: Estimates imply $\theta_\tau \approx 0.06$

Key Empirical Result: Salience matters

1) Posting sales taxes reduces demand for those goods

⇒ Breaks tax incidence result #2: equilibrium depends on who nominally pays the tax

2) Beer consumption is elastic to excise tax rate (built in posted price) but not to the sales tax rate (not built in the posted price)

⇒ If tax is not salient to consumers, they are less elastic, and hence more likely to bear the tax burden

A number of recent empirical studies show that individuals are not fully informed and fully rational and this has large consequences for policy

REFERENCES

Jonathan Gruber, Public Finance and Public Policy, Third Edition, 2010
Worth Publishers, Chapter 19

Chetty, Raj, Adam Looney, and Kory Kroft. "Salience and Taxation: Theory and Evidence." The American Economic Review 99.4 (2009): 1145-1177.(web)

Doyle Jr, Joseph J., and Krislert Samphantharak "\$2.00 Gas! Studying the effects of a gas tax moratorium." Journal of Public Economics 92.3 (2008): 869-884.(web)

Ramsey, Frank P. "A Contribution to the Theory of Taxation." The Economic Journal 37.145 (1927): 47-61.(web)

Summers, Lawrence H. "Some simple economics of mandated benefits." The American Economic Review 79.2 (1989): 177-183.(web)

Undergraduate Public Economics

Danny Yagan
UC Berkeley

Lecture 7

Labor Income Taxes and Transfers:
Optimal Policy Theory

TAXATION AND REDISTRIBUTION

Key question: Should government reduce inequality using taxes and transfers?

1) Governments use **taxes** to raise revenue

2) This revenue funds **transfer** programs:

a) Universal Transfers: Public Education, Health Care Benefits (only 65+ in the US), Retirement and Disability Benefits, Unemployment benefits

b) Means-tested Transfers: In-kind (e.g., public housing or Medicaid in the US) and cash benefits

Modern governments raise large fraction of GDP in taxes (30-45%) and spend significant fraction of GDP on transfers

FACTS ON US TAXES AND TRANSFERS

References: Comprehensive description in:

<http://www.taxpolicycenter.org/taxfacts/>

A) Taxes: (1) individual income tax (fed+state), (2) payroll taxes on earnings (fed, funds Social Security+Medicare), (3) corporate income tax (fed+state), (4) sales taxes (state)+excise taxes (state+fed), (5) property taxes (state)

B) Means-tested Transfers: (1) refundable tax credits (fed), (2) in-kind transfers (fed+state): Medicaid, public housing, nutrition (SNAP), education, (3) cash welfare: TANF for single parents (fed+state), SSI for old/disabled (fed)

FEDERAL US INCOME TAX

US income tax assessed on **annual family** income (not individual) [most other OECD countries have shifted to individual assessment]

Sum all cash income sources from family members (both from labor and capital income sources) = called **Adjusted Gross Income (AGI)**

Main exclusions: fringe benefits (health insurance, pension contributions), imputed rent of homeowners, unrealized capital gains

FEDERAL US INCOME TAX

Taxable income = AGI - personal exemptions - deduction

personal exemption = \$4K × # family members (in 2014)

deduction is max of standard deduction or itemized deductions

Standard deduction is a fixed amount depending on family structure (\$12.4K for couple, \$6.2K for single in 2014)

Itemized deductions: mortgage interest payments, charitable giving, state and local taxes paid, various other small items

[about 10% of AGI lost through itemized deductions, called tax expenditures]

FEDERAL US INCOME TAX: TAX BRACKETS

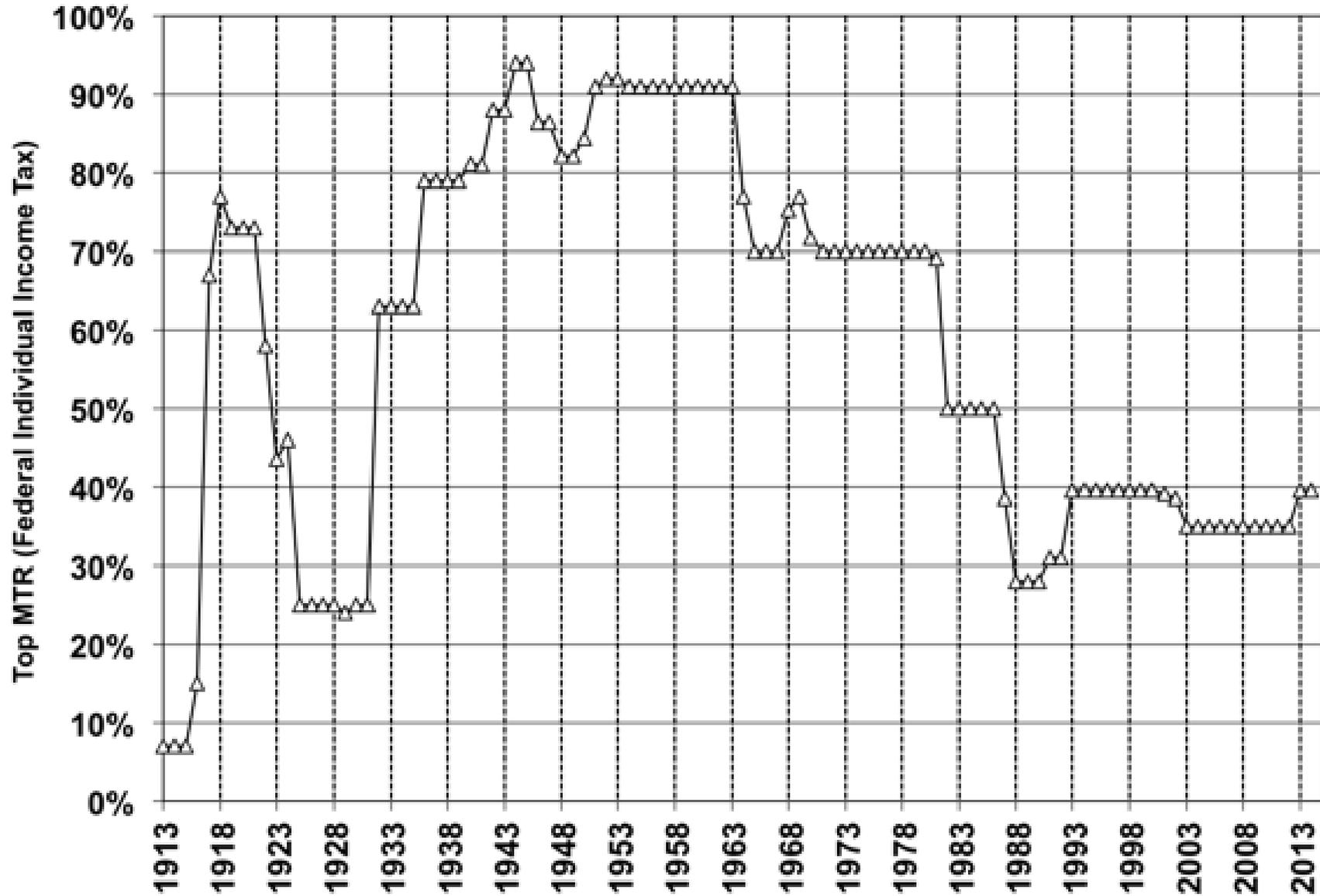
Tax $T(z)$ is **piecewise linear** and **continuous** function of taxable income z with constant marginal tax rates (MTR) $T'(z)$ by brackets [draw graph]

In 2016, 7 brackets with MTR 10%, 15%, 25%, 28%, 33%, 35%, 39.6% (top bracket for z above \$458K), indexed on price inflation

Lower preferential rates (up to a max of 20%) apply to dividends (since 2003) and realized capital gains [in part to offset double taxation of corporate profits]

Tax rates change frequently over time. Top MTRs have declined drastically since 1960s (as in most OECD countries)

US Top Marginal Tax Rate (Federal Individual Income Tax)



Source: IRS. Statistics of Income Division. Historical Table 23

FEDERAL US INCOME TAX: TAX CREDITS

Tax credits: Additional reduction in taxes

(1) **Non refundable** (cannot reduce taxes below zero): foreign tax credit, child care expenses, education credits, energy credits, and many others

(2) **Refundable** (can reduce taxes below zero, i.e., be net transfers): EITC (earned income tax credit, up to \$3.3K, \$5.5K, \$6.1K for working families with 1, 2, 3+ kids), Child Tax Credit (\$1000 per kid, partly refundable)

Refundable tax credits are now the largest means-tested cash transfer for low income families

EITC Amount as a Function of Earnings

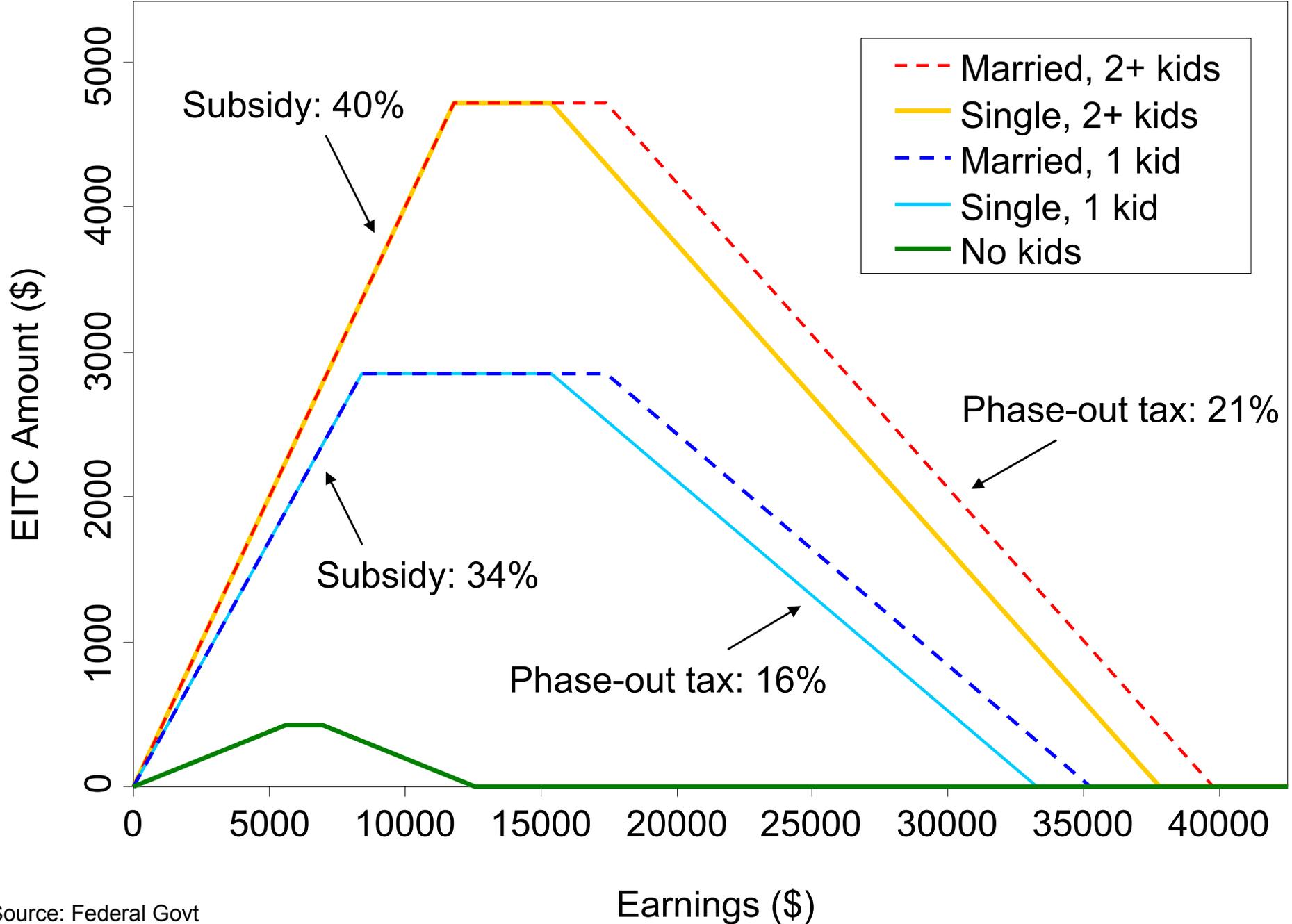
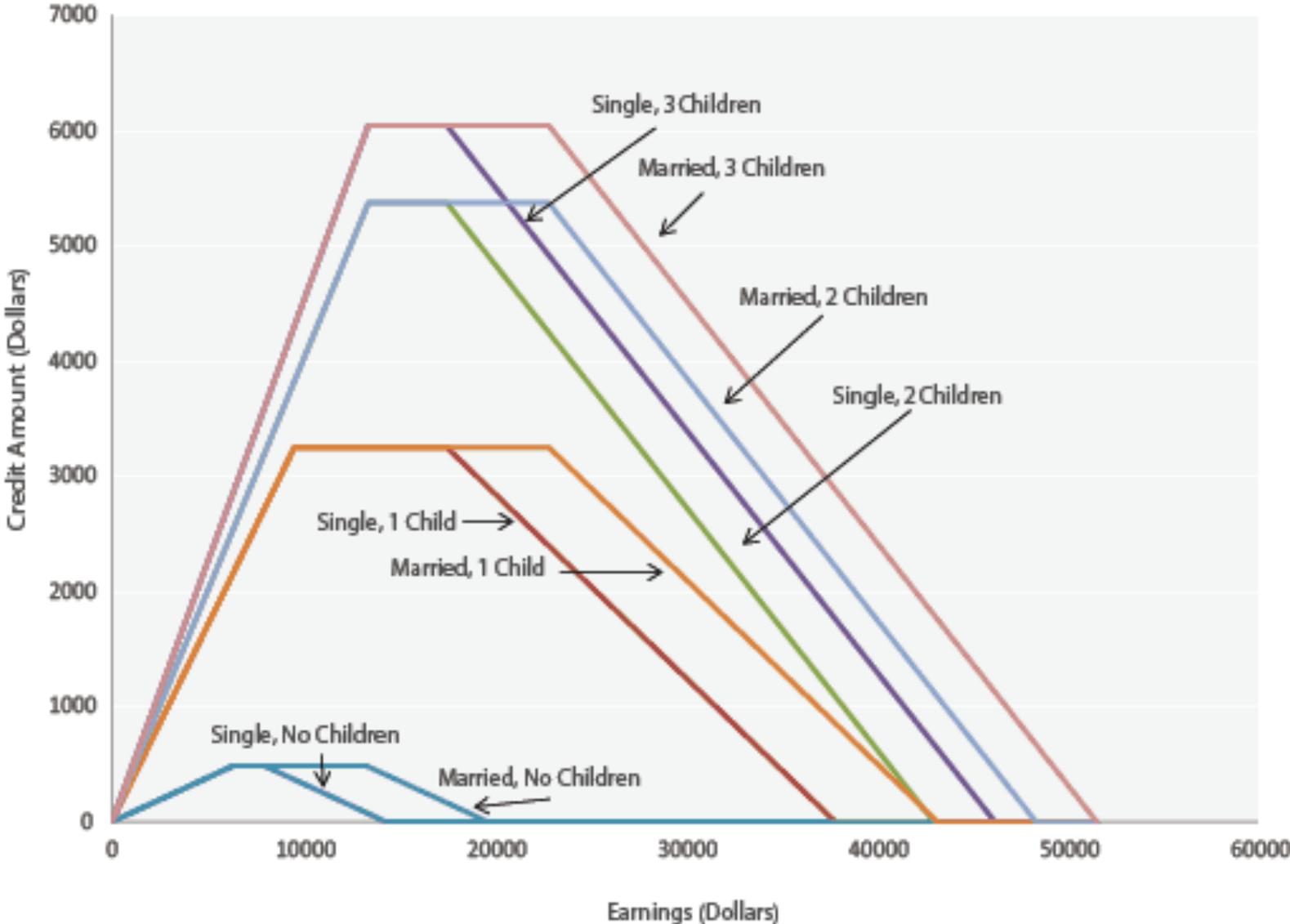


Figure 1: Earned Income Tax Credit by Number of Children and Filing Status, 2013



Source: 2013 EITC parameters taken from <http://www.taxpolicycenter.org/taxfacts/displayafact.cfm?Docid=36>

FEDERAL US INCOME TAX: TAX FILING

Taxes on year t earnings are withheld on paychecks during year t (pay-as-you-earn)

Income tax return filed in Feb-April 15, year $t + 1$ [filers use either software or tax preparers, huge private industry, most OECD countries provide pre-populated returns]

Most tax filers get a tax refund as withholdings larger than taxes owed in general

Payers (employers, banks, etc.) send income information to govt (3rd party reporting)

3rd party reporting + withholding at source is key for successful enforcement

MAIN MEANS-TESTED TRANSFER PROGRAMS

1) **Traditional transfers:** managed by welfare agencies, paid on monthly basis, high stigma and take-up costs \Rightarrow low take-up rates (often only around 50%)

Main programs: Medicaid (health insurance for low incomes), SNAP (former food stamps), public housing, TANF (welfare), SSI (aged+disabled)

2) **Refundable income tax credits:** managed by tax administration, paid as an annual lumpsum in year $t + 1$, low stigma and take-up cost \Rightarrow high take-up rates

Main programs: EITC and Child Tax Credit [large expansion since the 1990s] for low income working families with children

KEY CONCEPTS FOR TAXES/TRANSFERS

Draw budget $(z, z - T(z))$ which integrates taxes and transfers

1) Transfer benefit with zero earnings $-T(0)$ [sometimes called demogrant or lumpsum grant]

2) Marginal tax rate (or phasing-out rate) $T'(z)$: individual keeps $1 - T'(z)$ for an additional \$1 of earnings (intensive labor supply response)

3) Participation tax rate $\tau_p = [T(z) - T(0)]/z$: individual keeps fraction $1 - \tau_p$ of earnings when moving from zero earnings to earnings z (extensive labor supply response):

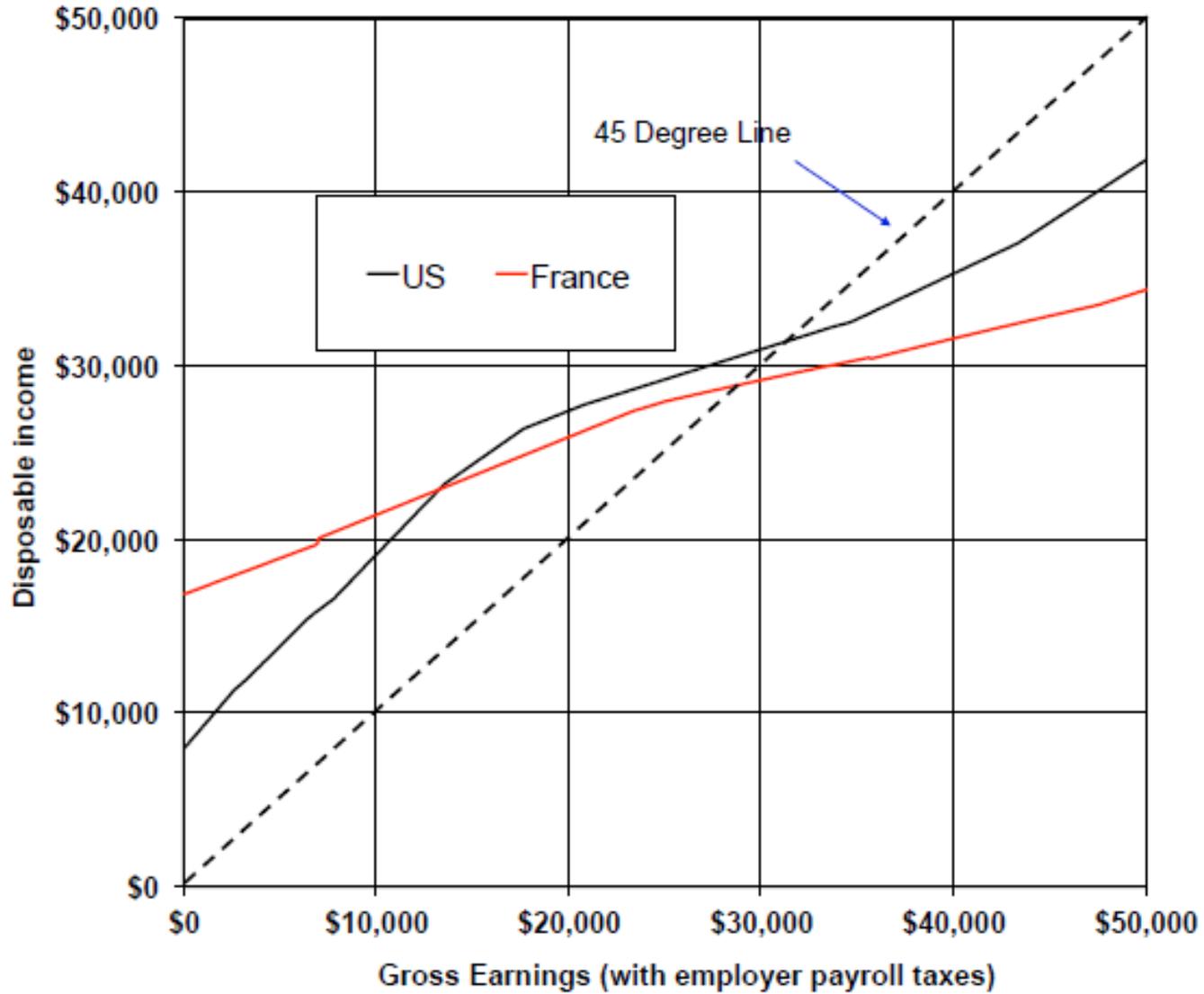
$$z - T(z) = -T(0) + z \cdot (1 - \tau_p)$$

4) Break-even earnings point z^* : point at which $T(z^*) = 0$

US Tax/Transfer System, single parent with 2 children, 2009



Source: Computations made by Emmanuel Saez using tax and transfer system parameters



Source: Piketty, Thomas, and Emmanuel Saez (2012)

Profile of Current Means-tested Transfers

Traditional means-tested programs reduce incentives to work for low income workers

Refundable tax credits have significantly increased incentive to work for low income workers

However, refundable tax credits cannot benefit those with zero earnings

Trade-off: US chooses to reward work more than most European countries (such as France) but therefore provides smaller benefits to those with no earnings

Optimal Taxation: Case with No Behavioral Responses

Utility $u(c)$ strictly increasing and concave

Same for everybody where c is after tax income.

Income z is fixed for each individual, $c = z - T(z)$ where $T(z)$ is tax/transfer on z .

N individuals with fixed incomes $z_1 < \dots < z_N$

Government maximizes **Utilitarian** objective:

$$SWF = \sum_{i=1}^N u(z_i - T(z_i))$$

subject to **budget constraint** $\sum_{i=1}^N T(z_i) = 0$ (taxes need to fund transfers)

Simple Model With No Behavioral Responses

Replace $T(z_1) = -\sum_{i=2}^N T(z_i)$ from budget constraint:

$$SWF = u\left(z_1 + \sum_{i=2}^N T(z_i)\right) + \sum_{i=2}^N u(z_i - T(z_i))$$

First order condition (FOC) in $T(z_i)$ for a given $i = 2, \dots, N$:

$$0 = \frac{\partial SWF}{\partial T(z_i)} = u'\left(z_1 + \sum_{j=2}^N T(z_j)\right) - u'(z_i - T(z_i)) = 0 \Rightarrow$$

$u'(z_i - T(z_i)) = u'(z_1 - T(z_1)) \Rightarrow z_i - T(z_i) = \text{constant across } i = 1, \dots, N$

Perfect equalization of after-tax income = 100% tax rate and redistribution [draw graph]

Utilitarianism with decreasing marginal utility leads to perfect egalitarianism [Edgeworth, 1897]

ISSUES WITH SIMPLE MODEL

1) **No behavioral responses:** Obvious missing piece: 100% redistribution would destroy incentives to work and thus the assumption that z is exogenous is unrealistic

⇒ Optimal income tax theory incorporates behavioral responses

2) **Issue with Utilitarianism:** Even absent behavioral responses, many people would object to 100% redistribution [perceived as confiscatory]

⇒ Citizens' views on fairness impose **bounds** on redistribution
govt can do [political economy / public choice theory]

EQUITY-EFFICIENCY TRADE-OFF

Taxes can be used to raise revenue for transfer programs which can reduce inequality in disposable income \Rightarrow Desirable if society feels that inequality is too large

Taxes (and transfers) reduce incentives to work \Rightarrow High tax rates create economic inefficiency if individual respond to taxes

Size of behavioral response limits the ability of govt to redistribute with taxes/transfers

\Rightarrow Generates an equity-efficiency trade-off

Empirical tax literature estimates the size of behavioral responses to taxation

Labor Supply Theory

Individual has utility over labor supply l and consumption c : $u(c, l)$ increasing in c and decreasing in l [i.e., increasing in leisure]

$$\max u(c, l) \quad \text{subject to} \quad c = w \cdot l + R$$

with $w = \bar{w} \cdot (1 - \tau)$ the net-of-tax wage (\bar{w} is before tax wage rate and τ is tax rate), and R non-labor income

FOC $w \frac{\partial u}{\partial c} + \frac{\partial u}{\partial l} = 0$ defines Marshallian labor supply $l = l(w, R)$

$$\text{Uncompensated labor supply elasticity : } \varepsilon^u = \frac{w}{l} \cdot \frac{\partial l}{\partial w}$$

$$\text{Income effects : } \eta = l \frac{\partial l}{\partial R} \leq 0$$

Labor Supply Theory

Substitution effects: Hicksian labor supply: $l^c(w, u)$ minimizes cost needed to reach u given slope $w \Rightarrow$

$$\text{Compensated elasticity } \varepsilon^c = \frac{w}{l} \cdot \frac{\partial l^c}{\partial w} > 0$$

$$\text{Slutsky equation } \frac{\partial l}{\partial w} = \frac{\partial l^c}{\partial w} + l \frac{\partial l}{\partial R} \Rightarrow \varepsilon^u = \varepsilon^c + \eta$$

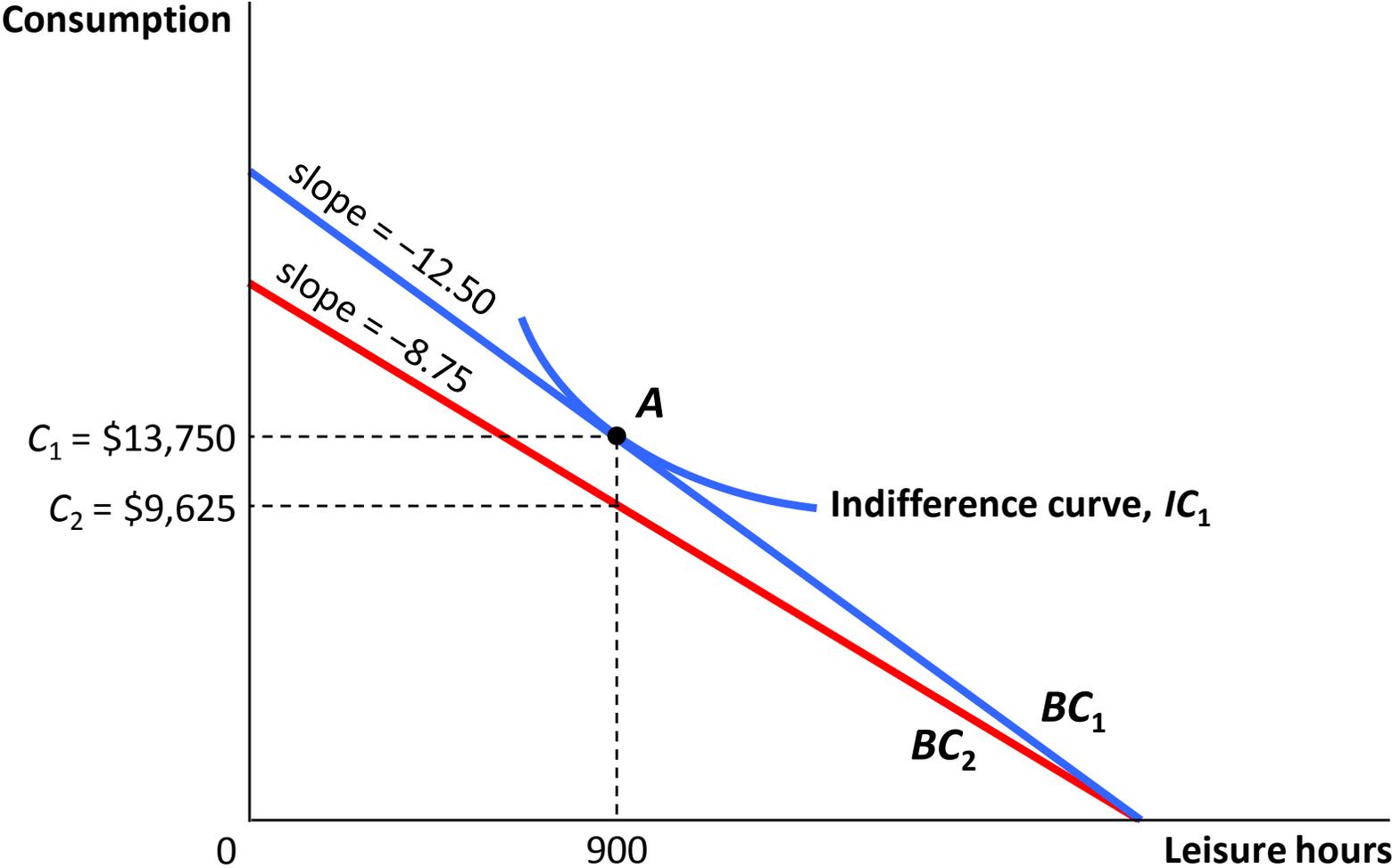
Tax rate τ discourages work through substitution effects (work pays less at the margin)

Tax rate τ encourages work through income effects (taxes make you poorer and hence in more need of income)

Net effect ambiguous (captured by sign of ε^u)

21.1

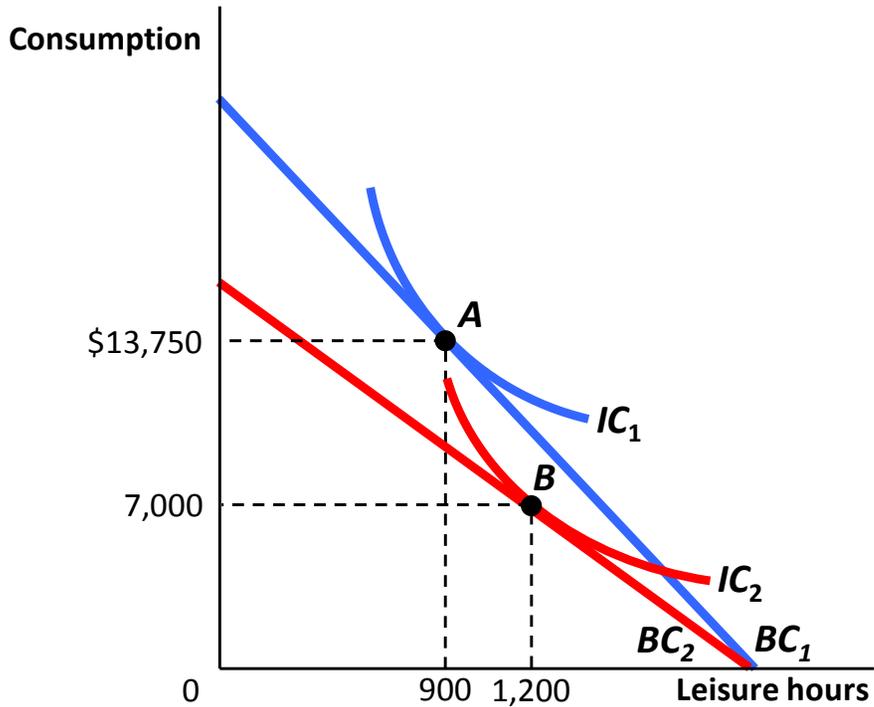
Basic Theory



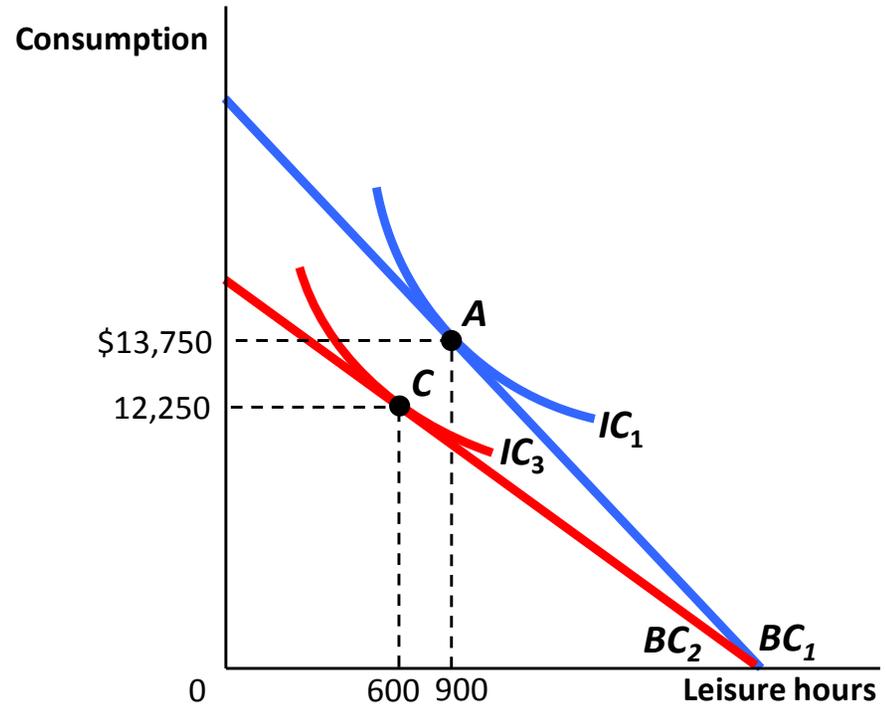
21.1

Substitution versus Income Effect

(a) Substitution effect is larger



(b) Income effect is larger



General nonlinear income tax [draw graph]

With no taxes: $c = z$ (consumption = earnings)

With taxes $c = z - T(z)$ (consumption = earnings - net taxes)

$T(z) \geq 0$ if individual pays taxes on net, $T(z) \leq 0$ if individual receives transfers on net

$T'(z) > 0$ reduces net wage rate and reduces labor supply through substitution effects

$T(z) > 0$ reduces disposable income and increases labor supply through income effects

$T(z) < 0$ increases disposable income and decreases labor supply through income effects

Transfer program such that $T(z) < 0$ and $T'(z) \geq 0$ always discourages labor supply

OPTIMAL LINEAR TAX RATE: LAFFER CURVE

$c = (1 - \tau) \cdot z + R$ with τ linear tax rate and R fixed universal transfer funded by taxes $R = \tau Z$ with Z average earnings

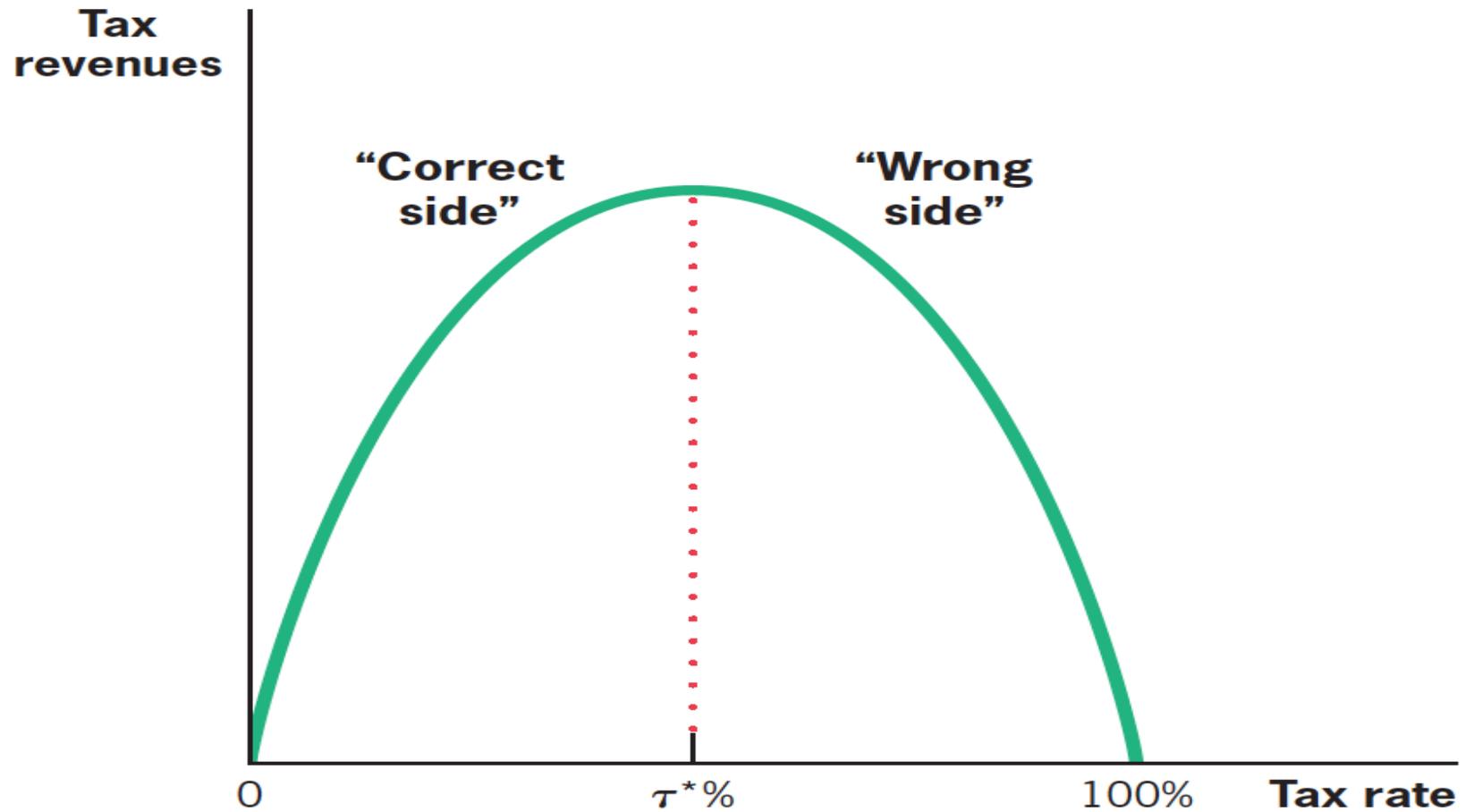
Individual i choose l_i to maximize $u^i((1 - \tau) \cdot w_i l_i + R, l_i)$

Labor supply choices l_i determine individual earnings $z_i = w_i l_i$. Assume small income effects \Rightarrow average earnings $Z = \sum_i z_i / N$ depend positively on net-of-tax rate $1 - \tau$.

Tax Revenue per person $R(\tau) = \tau \cdot Z(1 - \tau)$ is inversely U-shaped with τ : $R(\tau = 0) = 0$ (no taxes) and $R(\tau = 1) = 0$ (nobody works): called the Laffer Curve

20.3

The Laffer Curve



OPTIMAL LINEAR TAX RATE: LAFFER CURVE

Top of the Laffer Curve is at τ^* maximizing tax revenue:

$$0 = R'(\tau^*) = Z - \tau^* \frac{dZ}{d(1-\tau)} \Rightarrow \frac{\tau^*}{1-\tau^*} \cdot \frac{1-\tau^*}{Z} \frac{dZ}{d(1-\tau)} = 1$$

$$\text{Revenuemaximizingtaxrate} : \tau^* = \frac{1}{1+e} \text{ with } e = \frac{1-\tau}{Z} \frac{dZ}{d(1-\tau)}$$

e is the elasticity of average income Z with respect to the net-of-tax rate $1 - \tau$ [empirically estimable]

Inefficient to have $\tau > \tau^*$ because decreasing τ would make taxpayers better off (they pay less taxes) and would increase tax revenue for the government

If government is **Rawlsian** (maximizes welfare of the worst-off person with no earnings) then $\tau^* = 1/(1+e)$ is optimal to make transfer $R(\tau)$ as large as possible

OPTIMAL LINEAR TAX RATE: FORMULA

Government chooses τ to maximize **utilitarian** social welfare

$$SWF = \sum_i u^i((1 - \tau)w_i l_i + \tau \cdot Z(1 - \tau), l_i)$$

taking into account that labor supply l_i responds to taxation and hence that this affects the tax revenue per person $\tau \cdot Z(1 - \tau)$ that is redistributed back as transfer to everybody

Government first order condition: (using the envelope theorem as l_i maximizes u^i):

$$0 = \frac{dSWF}{d\tau} = \sum_i \frac{\partial u^i}{\partial c} \cdot \left[-z_i + Z - \tau \frac{dZ}{d(1 - \tau)} \right],$$

OPTIMAL LINEAR TAX RATE: FORMULA

Hence, we have the following optimal linear income tax formula

$$\tau = \frac{1 - \bar{g}}{1 - \bar{g} + e} \quad \text{with} \quad \bar{g} = \frac{\sum_i z_i \cdot \frac{\partial u^i}{\partial c}}{Z \cdot \sum_i \frac{\partial u^i}{\partial c}}$$

$0 \leq \bar{g} < 1$ as $\frac{\partial u^i}{\partial c}$ is decreasing with z_i (marginal utility falls with consumption)

τ decreases with elasticity e [efficiency] and with parameter \bar{g} [equity]

Formula captures the **equity-efficiency trade-off**

\bar{g} is low and τ close to Laffer rate $\tau^* = 1/(1 + e)$ when

(a) inequality is high

(b) marginal utility decreases fast with income

OPTIMAL TOP INCOME TAX RATE (Diamond and Saez JEP'11)

In practice, individual income tax is progressive with brackets with increasing marginal tax rates. What is the optimal top tax rate?

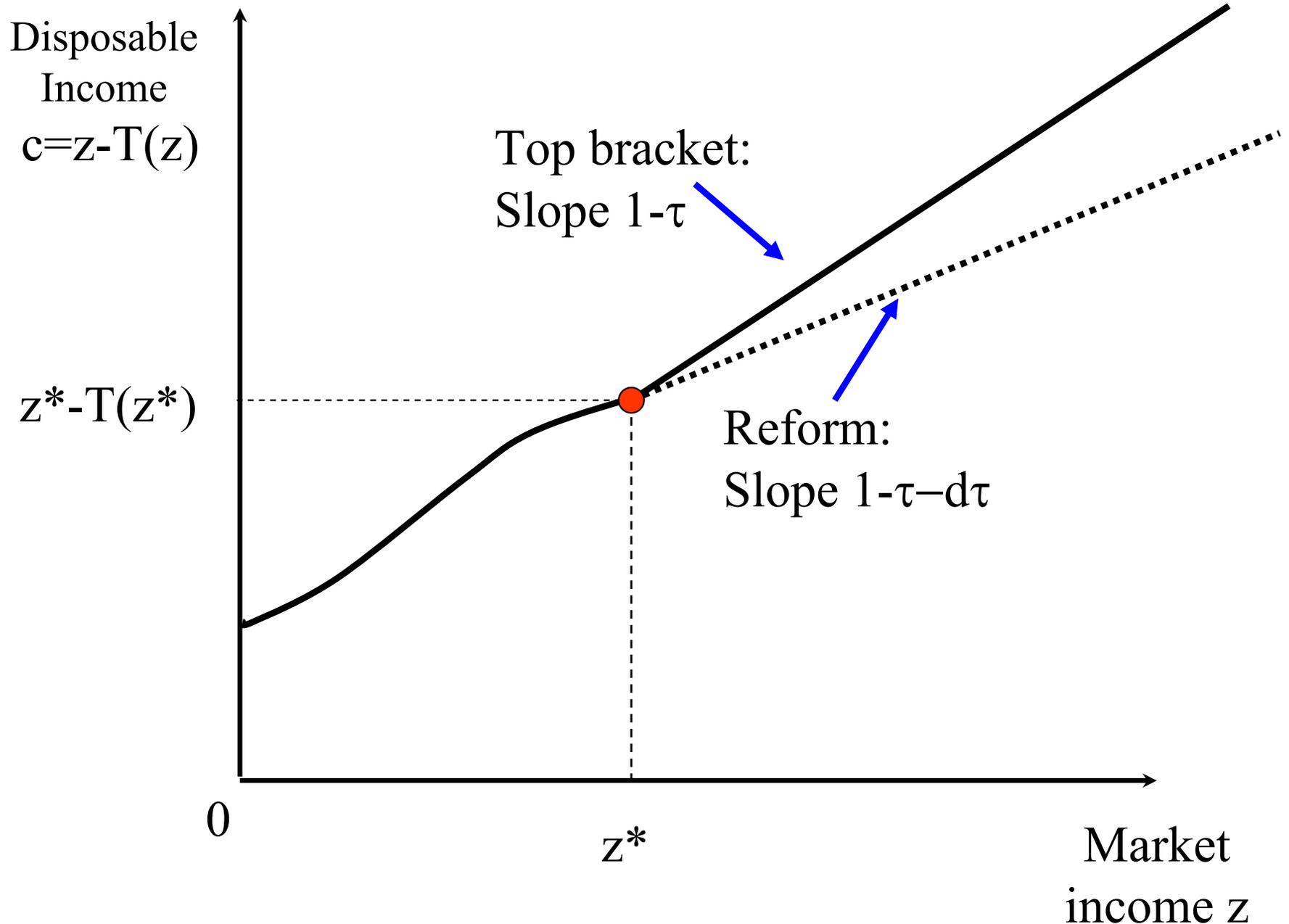
Consider constant MTR τ above fixed z^* . Goal is to derive optimal τ

In the US in 2014, $\tau = 39.6\%$ and $z^* = \$458,000$ (\simeq top 1%).

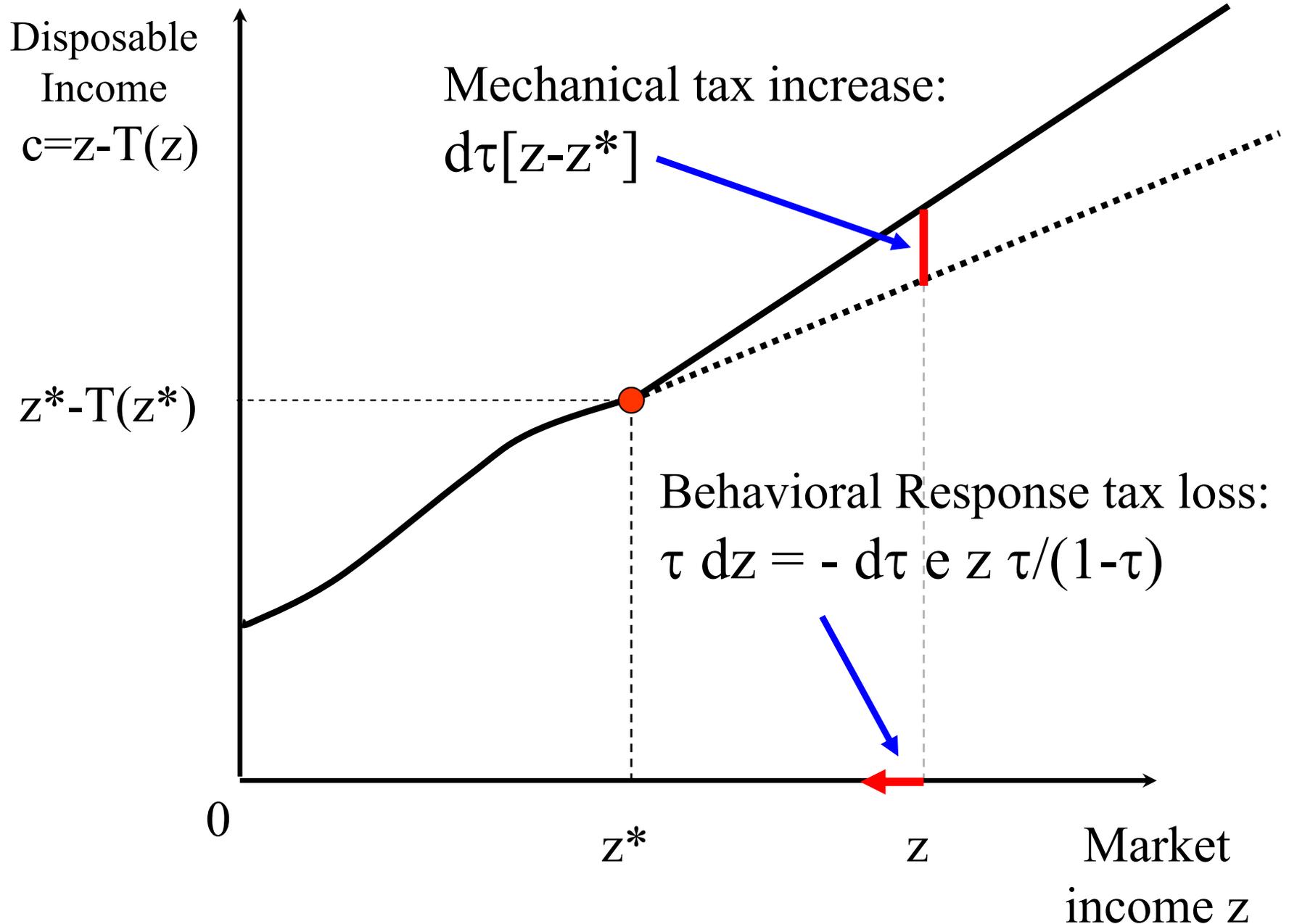
Denote by z average income of top bracket earners [depends on net-of-tax rate $1 - \tau$], with elasticity $e = [(1 - \tau)/z] \cdot dz/d(1 - \tau)$

Suppose the government wants to maximize tax revenue collected from top bracket taxpayers (marginal utility of consumption of top 1% earners is small)

Optimal Top Income Tax Rate (Mirrlees '71 model)



Optimal Top Income Tax Rate (Mirrlees '71 model)



OPTIMAL TOP INCOME TAX RATE

Consider small $d\tau > 0$ reform above z^* . Denote mean income above z^* as z_m . Assume constant elasticity e above z^* .

1) **Mechanical increase** in tax revenue:

$$dM = [z_m - z^*]d\tau$$

2) **Behavioral response** reduces tax revenue:

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

$$dM + dB = d\tau \left\{ [z_m - z^*] - e \frac{\tau}{1-\tau} z_m \right\}$$

Optimal τ such that $dM + dB = 0$

$$\Rightarrow \frac{\tau}{1-\tau} = \frac{1}{e} \cdot \frac{z_m - z^*}{z_m} \Rightarrow \tau = \frac{1}{1 + a \cdot e} \quad \text{with} \quad a = \frac{z_m}{z_m - z^*}$$

OPTIMAL TOP INCOME TAX RATE

$$\text{Optimal tax rate } \tau = \frac{1}{1 + a \cdot e} \quad \text{with} \quad a = \frac{z_m}{z_m - z^*}$$

Optimal τ decreases with e [efficiency]

Optimal τ decrease with a [thinness of top tail]

Empirically $a \simeq 1.5$, easy to estimate using distributional data

Empirically e is harder to estimate [controversial]

Example: If $e = .25$ then $\tau = 1/(1 + 1.5 \cdot 0.25) = 1/1.75 = 73\%$

Example: If $e = 1$ then $\tau = 1/(1 + 1.5 \cdot 1) = 1/2.5 = 40\%$

REAL VS. TAX AVOIDANCE RESPONSES

Behavioral response to income tax comes not only from reduced labor supply but from tax avoidance or tax evasion

Tax avoidance: legal means to reduce tax liability (exploiting tax loopholes)

Tax evasion: illegal under-reporting of income

Labor supply vs. tax avoidance/evasion distinction matters because:

1) If people work less when tax rates increase, there is not much the government can do about it

2) If people avoid/evade more when tax rates increase, then the govt can reduce tax avoidance/evasion opportunities [closing tax loopholes, broadening the tax base, increasing tax enforcement, etc.]

REAL VS. AVOIDANCE RESPONSES

Key policy question: Is it possible to eliminate avoidance responses using base broadening, etc.? or would new avoidance schemes keep popping up?

a) Some forms of tax avoidance are due to **poorly designed tax codes** (preferential treatment for some income forms or some deductions)

b) Some forms of tax avoidance/evasion can only be addressed with **international cooperation** (off-shore tax evasion in tax heavens)

c) Some forms of tax avoidance/evasion are due to technological limitations of tax collection (impossible to tax informal cash businesses)

OPTIMAL PROFILE OF TRANSFERS

If individuals respond to taxes only through intensive margin (how much they work at the margin and not whether they work), optimal transfer at bottom takes the form of a “Negative Income Tax”:

- 1) Lumpsum grant $-T(0) > 0$ for those with no earnings
- 2) High MTRs $T'(z)$ at the bottom to phase-out the lumpsum grant quickly

Intuition: high MTRs at bottom are efficient because:

- (a) they target transfers to the most needy
- (b) earnings at the bottom are low to start with \Rightarrow intensive labor supply response does not generate large output losses

Optimal Transfers: Participation Responses

Empirical literature shows that participation labor supply responses [whether to work or not] are large at the bottom [much larger and clearer than intensive responses]

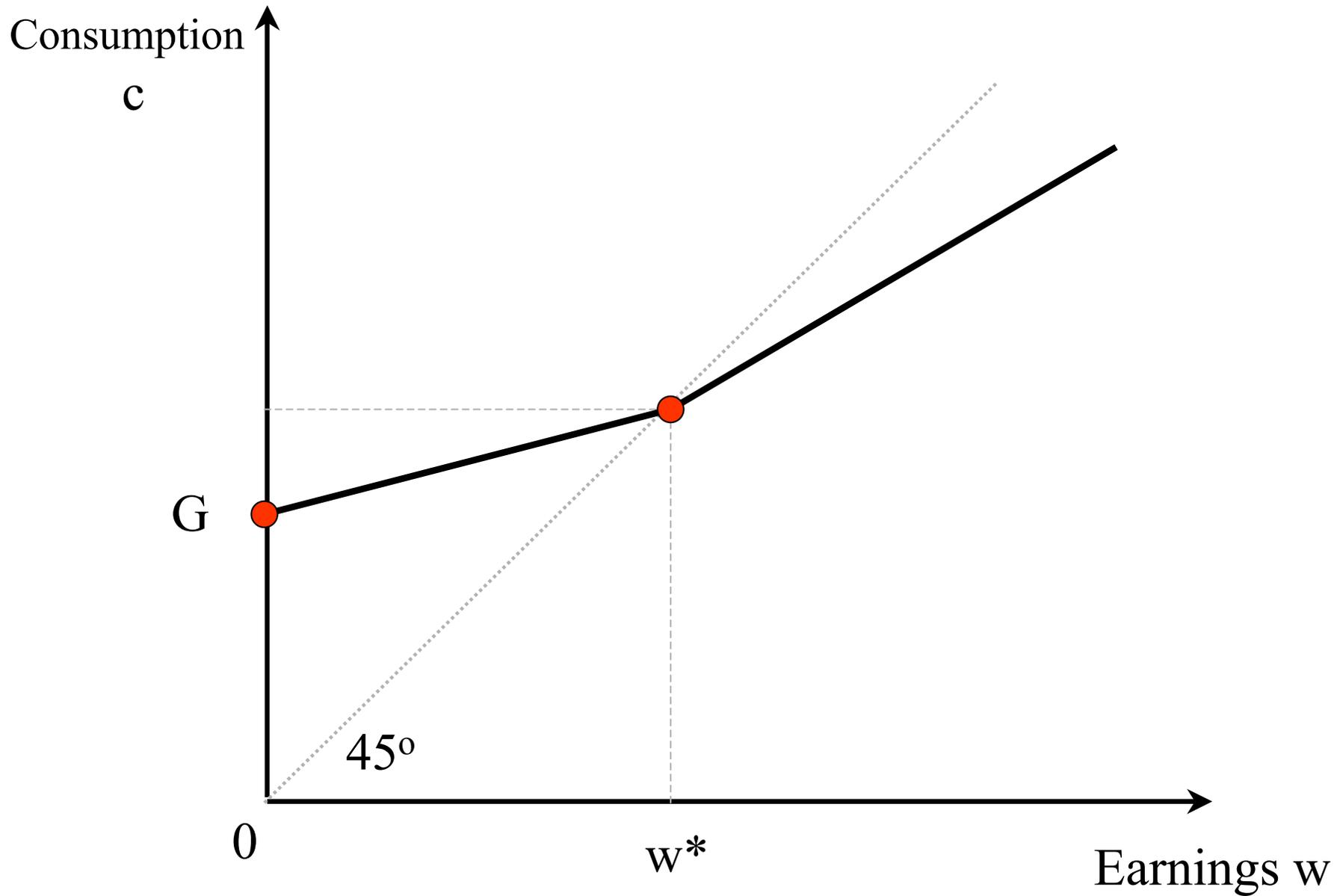
Participation depends on participation tax rate:

$$\tau_p = [T(z) - T(0)]/z$$

Individual keeps fraction $1 - \tau_p$ of earnings when moving from zero earnings to earnings z : $z - T(z) = -T(0) + z \cdot (1 - \tau_p)$

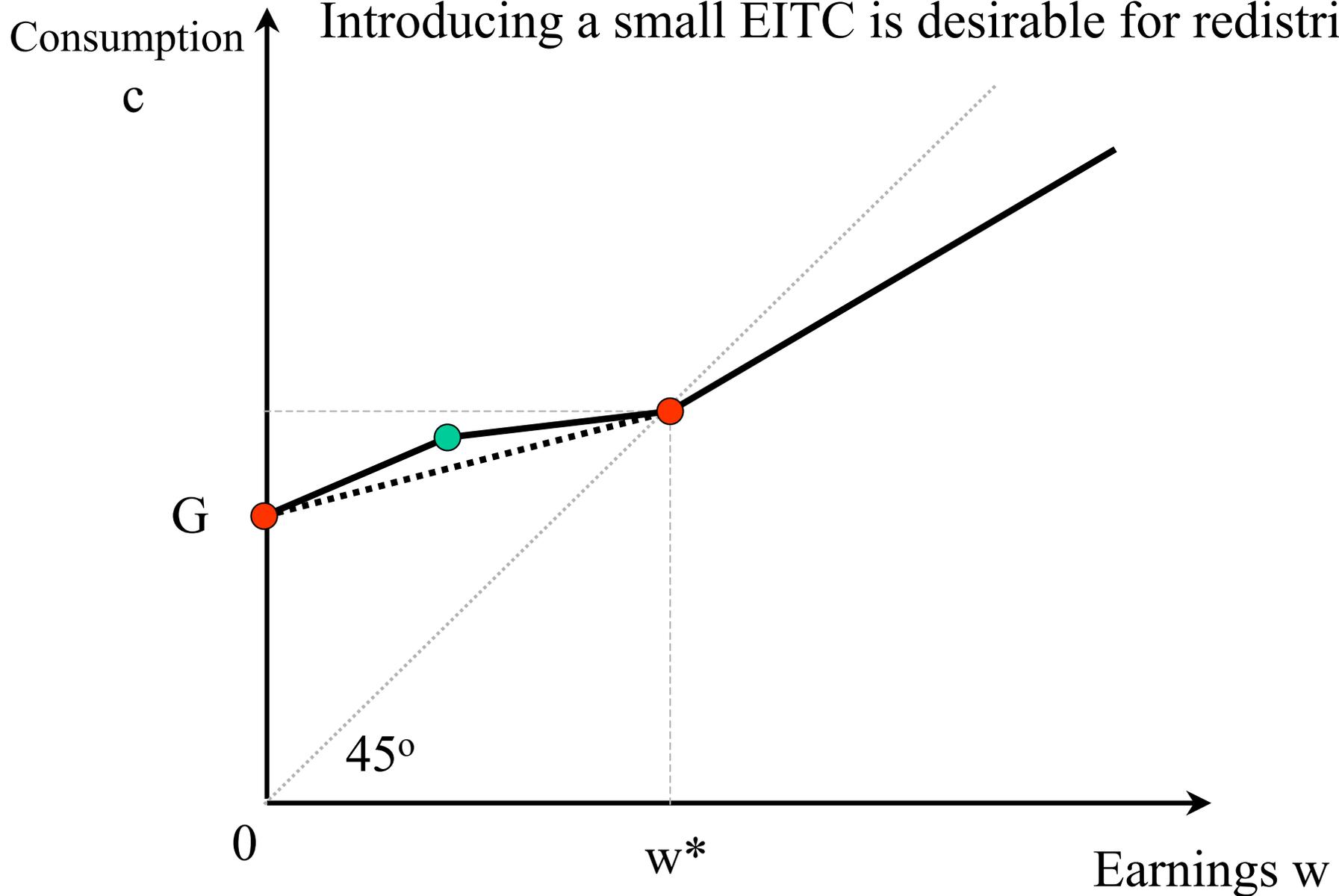
Key result: in-work subsidies with $T'(z) < 0$ are optimal when labor supply responses are concentrated along extensive margin and govt cares about low income workers.

Starting from a Means-Tested Program



Starting from a Means-Tested Program

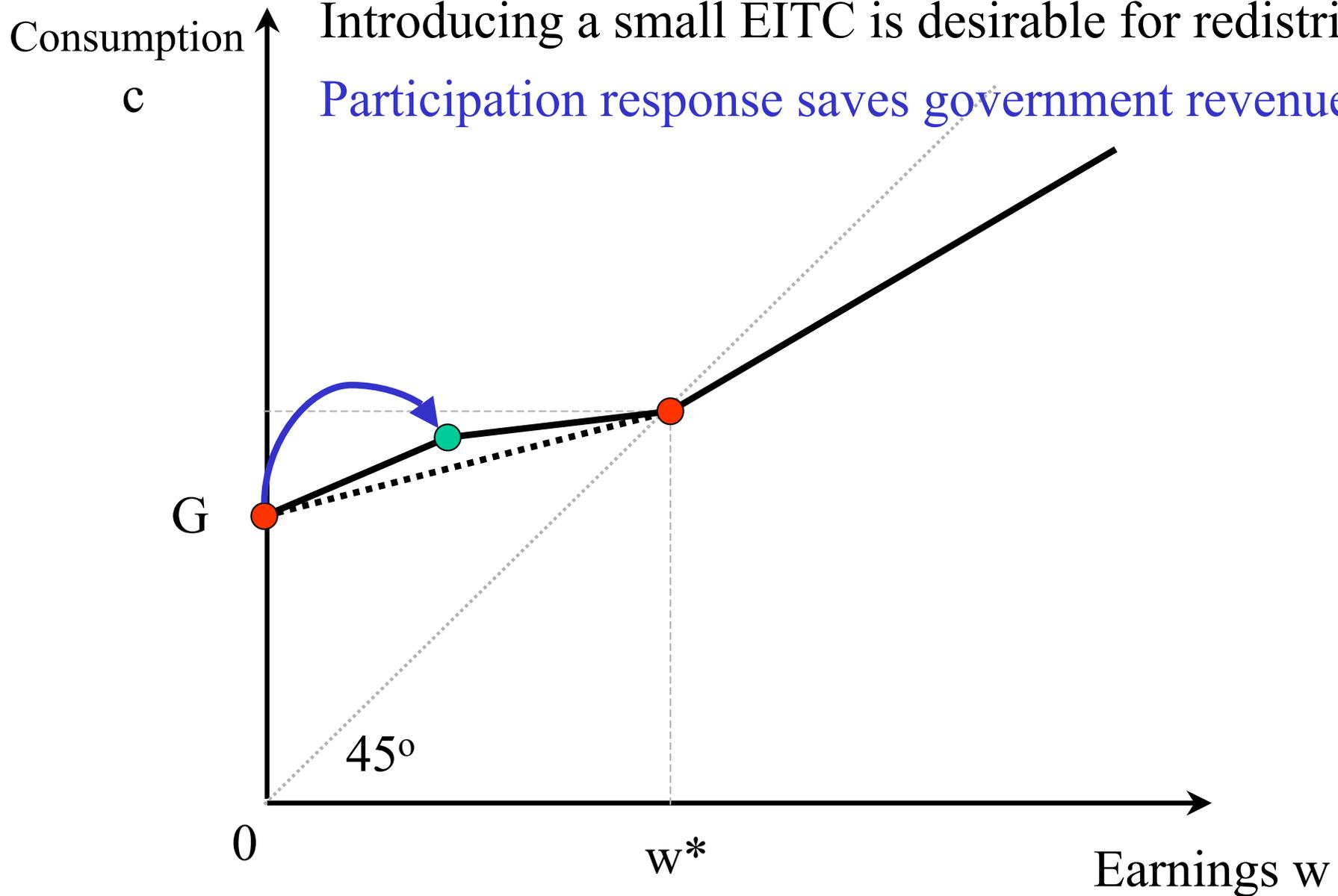
Introducing a small EITC is desirable for redistribution



Starting from a Means-Tested Program

Introducing a small EITC is desirable for redistribution

Participation response saves government revenue



FAMILY TAXATION: MARRIAGE AND CHILDREN

Two important issues in policy debate:

- 1) Marriage: What is the optimal taxation of couples vs. singles?
- 2) Children: What should be the net transfer (transfer or tax reduction) for family with children (as a function of family income and structure)?

TAXATION OF COUPLES

Three potentially desirable properties:

(1) income tax should be based on resources (i.e., family income if families fully share their income)

(2) income tax should be marriage neutral: no higher/lower tax when two single individuals marry

(3) income tax should be progressive (i.e., higher incomes pay a larger fraction of their income in taxes)

It is **impossible** to have a tax system that satisfies all 3 conditions simultaneously:

Income tax that is based on family income and marriage neutral has to satisfy: $T(z^h + z^w) = T(z^h) + T(z^w)$ and hence be linear i.e. $T(z) = \tau \cdot z$

TAXATION OF COUPLES

(1) If couples share their incomes, then family taxation is better. If couples don't share their incomes, then individualized tax is better

(2) If marriage responds to tax/transfer differential \Rightarrow better to reduce marriage penalty, i.e., move toward individualized system

Particularly important when cohabitation is close substitute for marriage (as in Scandinavian countries)

(3) If labor supply of secondary earners more elastic than labor supply of primary earner \Rightarrow Secondary earnings should be taxed less (Boskin-Sheshinski JpubE'83)

Labor supply elasticity differential between primary and secondary earners is decreasing over time as earnings gender gap decreases

TRANSFERS OR TAX CREDITS FOR CHILDREN

1) Children reduce **normalized family income** \Rightarrow Children increase marginal utility of consumption \Rightarrow Transfer for children T_{kid} should be positive

In practice, transfers for children are always positive

2) Should $T_{kid}(z)$ increase with income z ?

Pro: rich spend more on their kids than lower income families

Cons: Lower income families need child transfers most

In practice, $T_{kid}(z)$ is fairly constant with z

Europe has much more generous pre-kindergarten child care benefits, US has more generous cash tax credits for families with children

Best Additional Reading

For both this lecture and the next lecture:

Diamond, P. and E. Saez "From Basic Research to Policy Recommendations: The Case for a Progressive Tax", *Journal of Economic Perspectives*, 25.4, (2011): 165-190. (web)

Mankiw, N.G., Weinzierl, M. and Yagan, D., 2009. Optimal taxation in theory and practice. *The Journal of Economic Perspectives*, 23(4), pp.147-174. (web)

Saez, Emmanuel, Joel Slemrod, and Seth H. Giertz. "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review." *Journal of Economic Literature* 50(1) (2012): 3-50. (web)

REFERENCES

Jonathan Gruber, Public Finance and Public Policy, Third Edition, 2010
Worth Publishers, Chapter 20 and Chapter 21

Boskin, Michael J., and Eytan Sheshinski. "Optimal tax treatment of the family: Married couples." *Journal of Public Economics* 20.3 (1983): 281-297.(web)

Diamond, P. and E. Saez "From Basic Research to Policy Recommendations: The Case for a Progressive Tax", *Journal of Economic Perspectives*, 25.4, (2011): 165-190. (web)

IRS, Statistics of Income Division "U.S. Individual Income Tax: Personal Exemptions and Lowest and Highest Tax Bracket"(2013) (web)

Piketty, Thomas, and Emmanuel Saez. "Optimal labor income taxation." No. w18521. National Bureau of Economic Research, 2012.(web)

Saez, Emmanuel. "Optimal income transfer programs: intensive versus extensive labor supply responses." *The Quarterly Journal of Economics* 117.3 (2002): 1039-1073.(web)

Saez, Emmanuel, Joel Slemrod, and Seth H. Giertz. "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review." *Journal of Economic Literature* 50(1) (2012): 3-50. (web)

Undergraduate Public Economics

Danny Yagan
UC Berkeley

Lecture 8

Labor Income Taxes and Transfers:
Evidence

This Lecture

Last lecture: Theoretical groundwork to understand effect of taxes and identify two elasticities needed for optimal policy at top and bottom incomes

- 1) EITC: Is the extensive (participation) margin elasticity large?
- 2) Top marginal rate: How large is the top income elasticity?

Will also discuss additional outside-the-model considerations

Responses to Low-Income Transfer Programs

- 1) Particular interest in treatment of low incomes in a progressive tax/transfer system: are they responsive to incentives?
- 2) Complicated set of transfer programs in US
 - a) In-kind: food stamps, Medicaid, public housing, job training, education subsidies
 - b) Cash: TANF, EITC, SSI

Overall Costs of Anti Poverty Programs

1) US government (fed+state and local) spent \$800bn in 2012 on income-tested programs

a) About 5% of GDP but 15% of \$5 Trillion govt budget (fed+state+local).

b) About 50% is health care (Medicaid)

2) Only \$200 billion in cash (1.3% of GDP, or 25% of transfer spending)

EITC Expansions and Welfare Reform

Nationwide expansions in the EITC 1987-1994, largest in 1994.

Also state-wide expansions (understudied b/c CPS data too thin)

1996 Welfare Reform: Turned AFDC (Aid to Families with Dependent Children) into TANF (Temporary Assistance to Needy Families) which required recipients to work or get training, limited benefit duration, and reduced phase-out rate

General shift from welfare to “workfare”

Federal Income Tax and EITC Parameters, 1984–2005

Year	Federal Income Tax Parameters					
	[Lowest Marginal Tax Rate; Highest Marginal Tax Rate]	Personal Exemption, Standard Deduction ^{a,b}	EITC Parameters (Family with One Child; Family Two or More Children)			
			Phase-In Rate	Maximum Credit	Phase-Out Rate	Maximum Earnings
1984	[0.000; 0.500]	\$1,000; \$0	0.100	\$ 500	0.125	\$10,000
1985	[0.000; 0.500]	\$1,040; \$0	0.110	\$ 550	0.122	\$11,000
1986	[0.000; 0.500]	\$1,080; \$0	0.110	\$ 550	0.122	\$11,000
<i>TRA86</i>						
1987	[0.110; 0.390]	\$1,900; \$2,540	0.140	\$ 851	0.100	\$15,432
1988	[0.150; 0.330]	\$1,950; \$4,400	0.140	\$ 874	0.100	\$18,576
1989	[0.150; 0.330]	\$2,000; \$4,550	0.140	\$ 910	0.100	\$19,340
1990	[0.150; 0.330]	\$2,050; \$4,750	0.140	\$ 953	0.100	\$20,264
<i>OBRA90^c</i>						
1991	[0.150; 0.310]	\$2,150; \$5,000	0.167; 0.173	\$1,192; \$1,235	0.119; 0.124	\$21,250
1992	[0.150; 0.310]	\$2,300; \$5,250	0.176; 0.184	\$1,324; \$1,384	0.126; 0.131	\$22,370
1993	[0.150; 0.396]	\$2,350; \$5,450	0.185; 0.195	\$1,434; \$1,511	0.132; 0.139	\$23,050
<i>OBRA1993^d</i>						
1994	[0.150; 0.396]	\$2,450; \$5,600	0.263; 0.300	\$2,038; \$2,528	0.160; 0.177	\$23,755; \$25,296
1995	[0.150; 0.396]	\$2,500; \$5,750	0.340; 0.360	\$2,094; \$3,110	0.160; 0.202	\$24,396; \$26,673
1996	[0.150; 0.396]	\$2,550; \$5,900	0.340; 0.400	\$2,152; \$3,556	0.160; 0.211	\$25,078; \$28,495

Source: Eissa Hoynes (2006)

Year	Federal Income Tax Parameters		EITC Parameters (Family with One Child; Family Two or More Children)			
	[Lowest Marginal Tax Rate; Highest Marginal Tax Rate]	Personal Exemption, Standard Deduction ^{a,b}	Phase-In Rate	Maximum Credit	Phase-Out Rate	Maximum Earnings
	1997	[0.150; 0.396]	\$2,650; \$6,050	0.340; 0.400	\$2,210; \$3,656	0.160; 0.211
2000	[0.150; 0.396]	\$2,900; \$6,650	0.340; 0.400	\$2,353; \$3,888	0.160; 0.211	\$27,413; \$31,152
<i>EGTRRA2001^e</i>						
2002	[0.100; 0.386]	\$3,050; \$7,000	0.340; 0.400	\$2,547; \$4,204	0.160; 0.211	\$29,201; \$33,178
2004	[0.100; 0.35]	\$3,100; \$7,150	0.340; 0.400	\$2,604; \$4,300	0.160; 0.211	\$30,338; \$34,458
2005	[0.100; 0.35]	\$3,200; \$7,300	0.340; 0.400	\$2,662; \$4,400	0.160; 0.211	\$31,030; \$35,263

Source: Eissa Hoynes (2006)

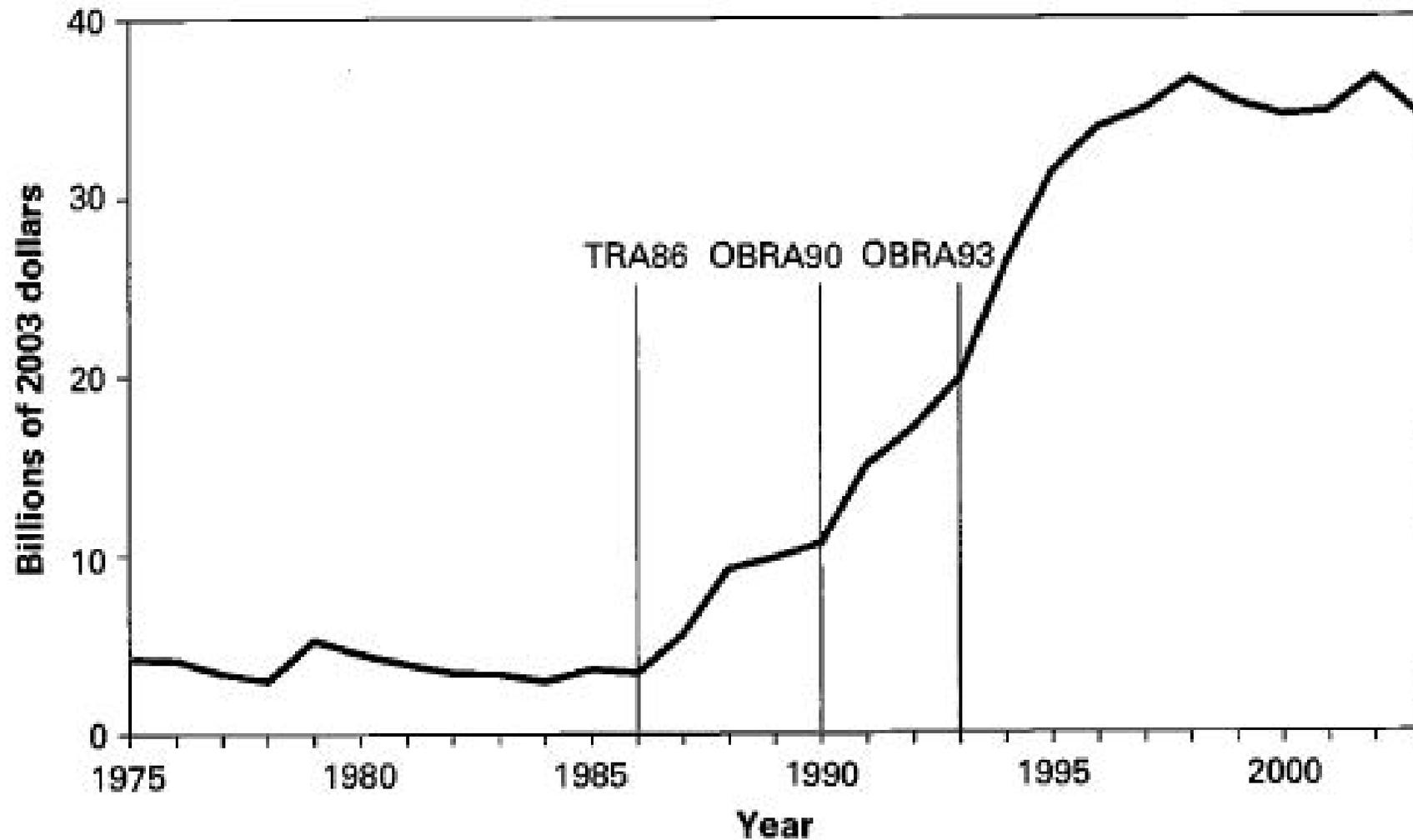


Figure 3.2
 Real Spending on the EITC (Billions of 2003 Dollars).
 Source: Eissa Hoynes (2006)

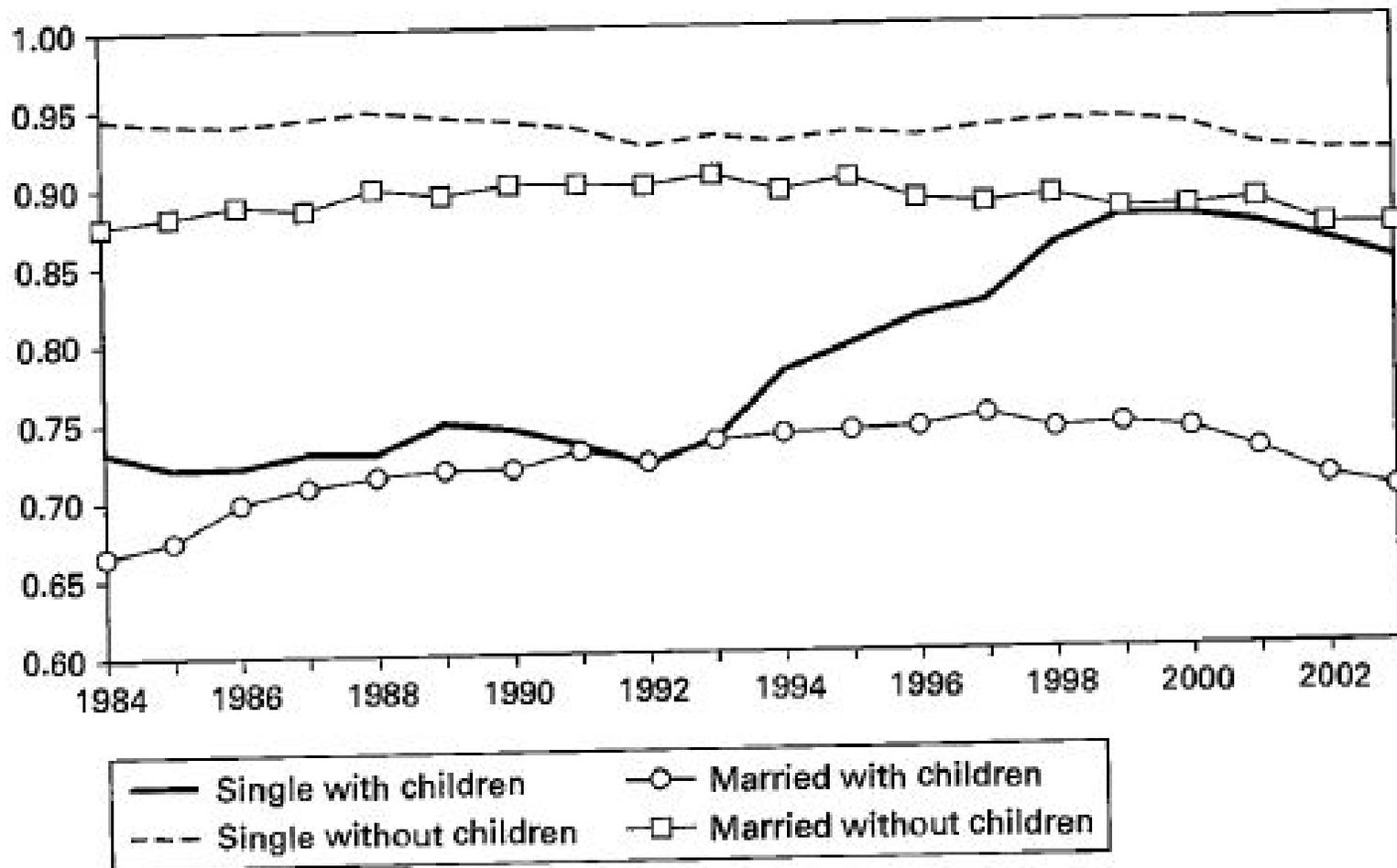


Figure 3.5
Annual Employment Rates for Women by Marital Status and Presence of Children, 1984–2003.

Source: Eissa Hoynes (2006)

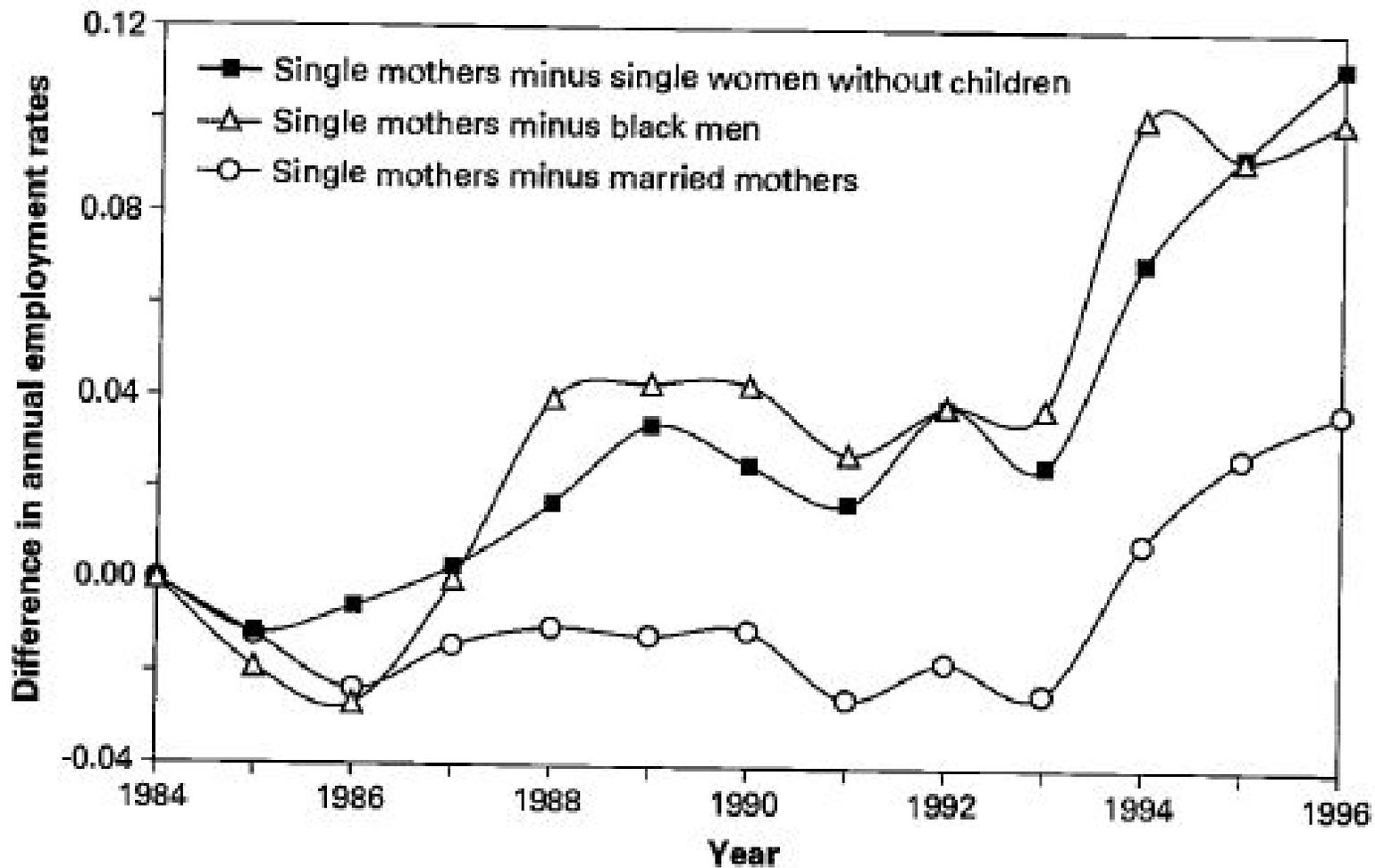


Figure 3.6
 Difference-in-Difference Estimates of Annual Employment Rates, Single Mothers Versus Control Groups, 1984–1996.

Source: Eissa Hoynes (2006)

EITC Extensive Margin Elasticity

Large extensive margin elasticity, ranging between 0.69 and 1.16 (Hotz Scholz 2003)

EITC responsible for most of relative increase in single mothers' employment 1984-1996 (Meyer Rosenbaum 2001), welfare reform (AFDC→TANF) responsible for less

Implies that an EITC is likely optimal (Saez 2002)

EITC Follow-up #1: Net Poverty Reduction?

EITC increases participation [reduces poverty]

EITC phase-out should reduce work [can increase poverty]

Need estimates of effects across full EITC distribution

Chetty, Friedman, Saez AER'13 EITC information

Use US population wide tax return data since 1996

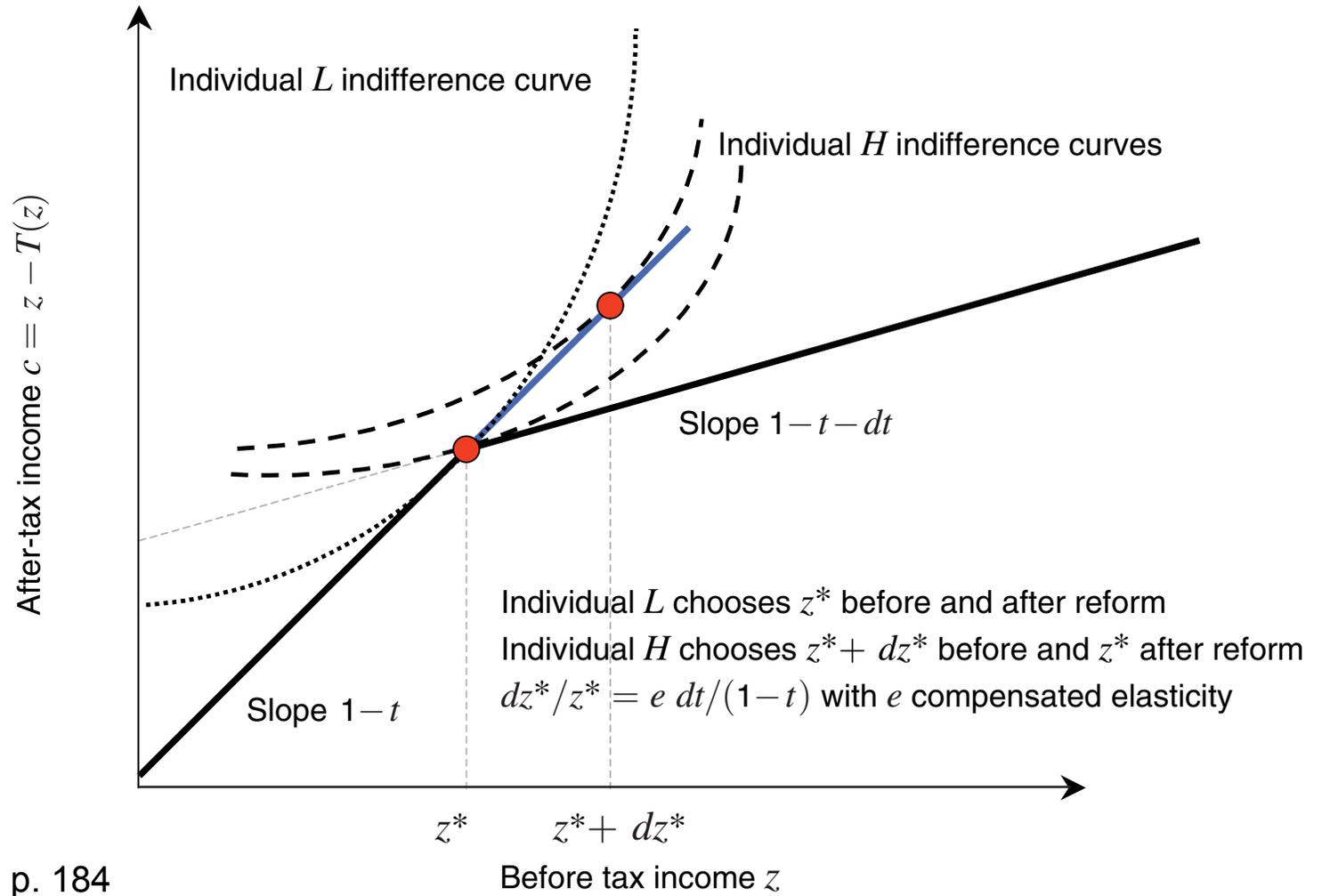
1) Substantial heterogeneity fraction of EITC recipients bunching (using self-employment) across geographical areas \Rightarrow Information about EITC varies across areas

2) Places with high self-employment EITC bunching display **wage earnings** distribution more concentrated around plateau

\Rightarrow Evidence of wage earnings response to EITC along intensive margin

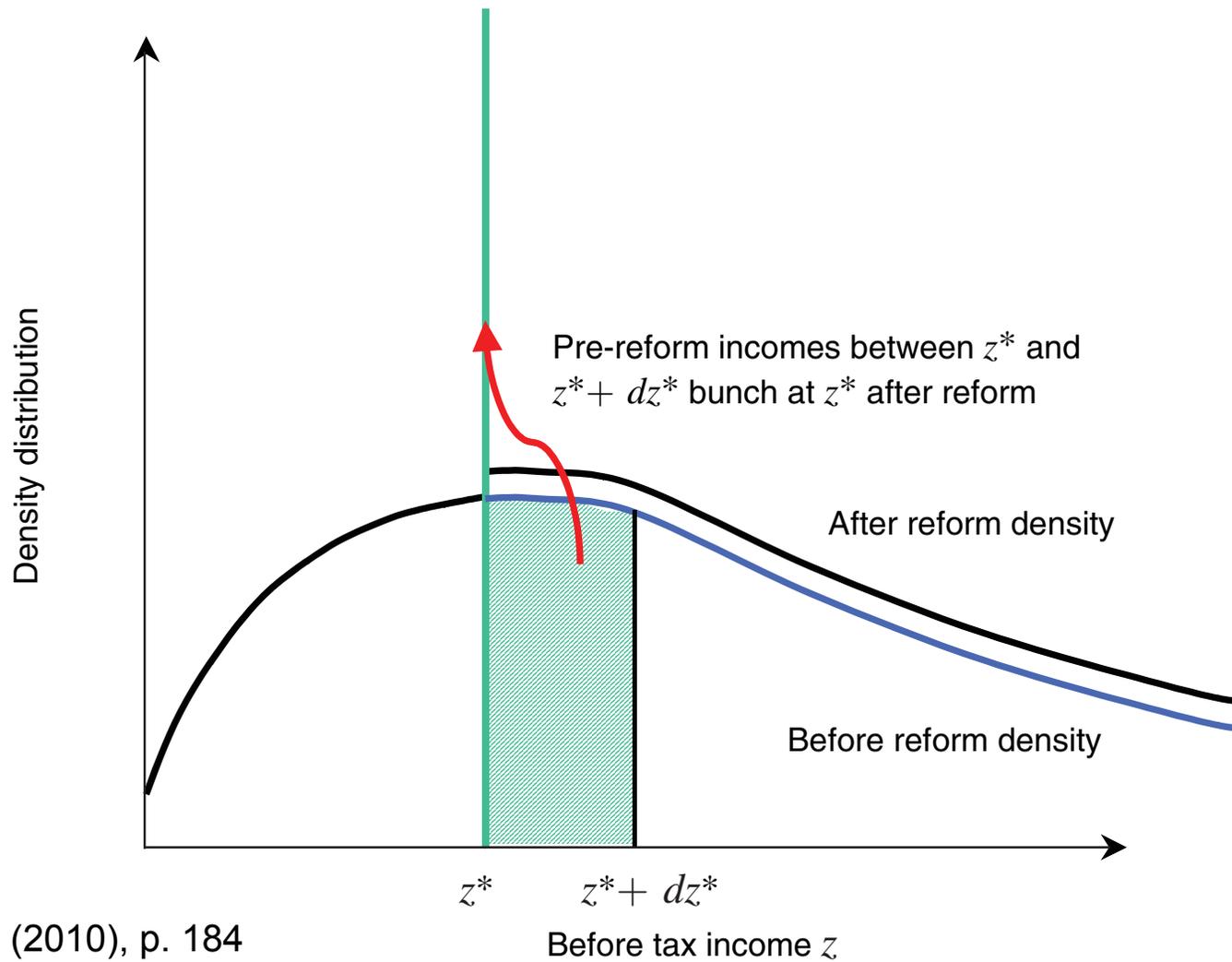
3) Key strategy: use birth of first child to estimate causal effect of EITC on wage earnings

Panel A. Indifference curves and bunching



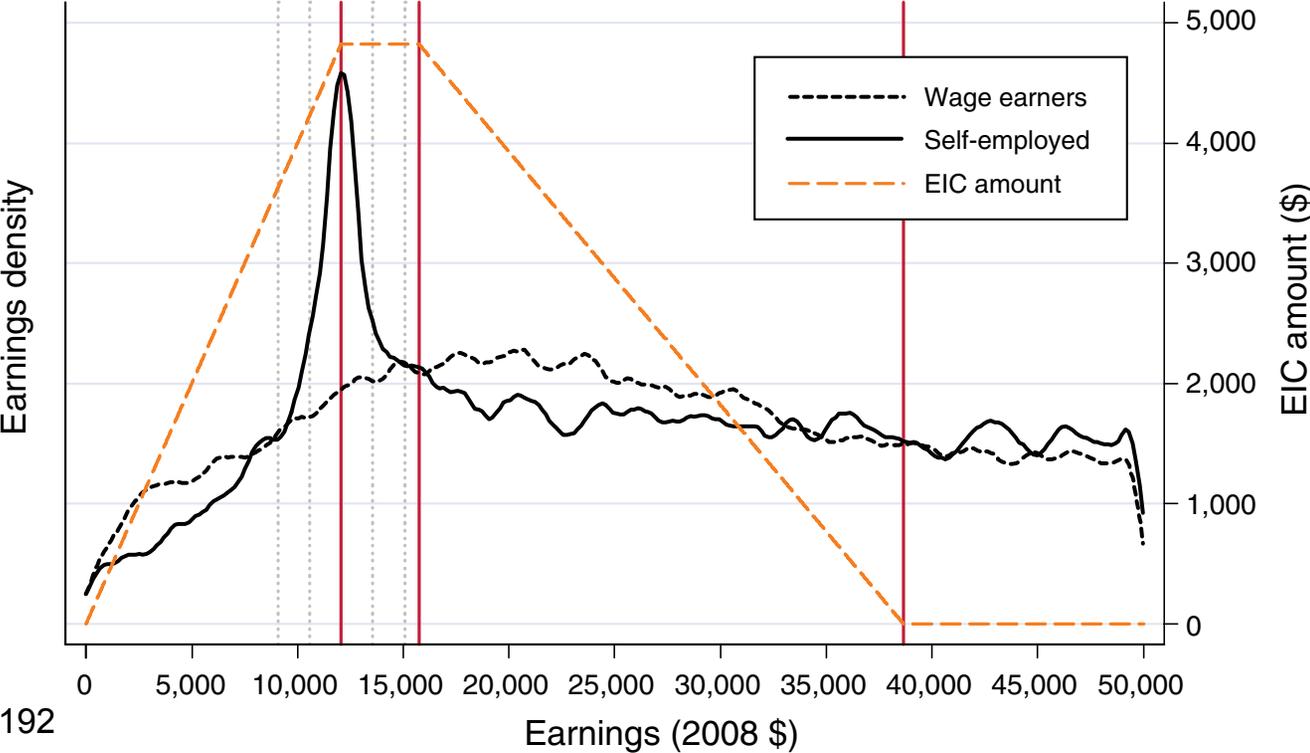
Source: Saez (2010), p. 184

Panel B. Density distributions and bunching



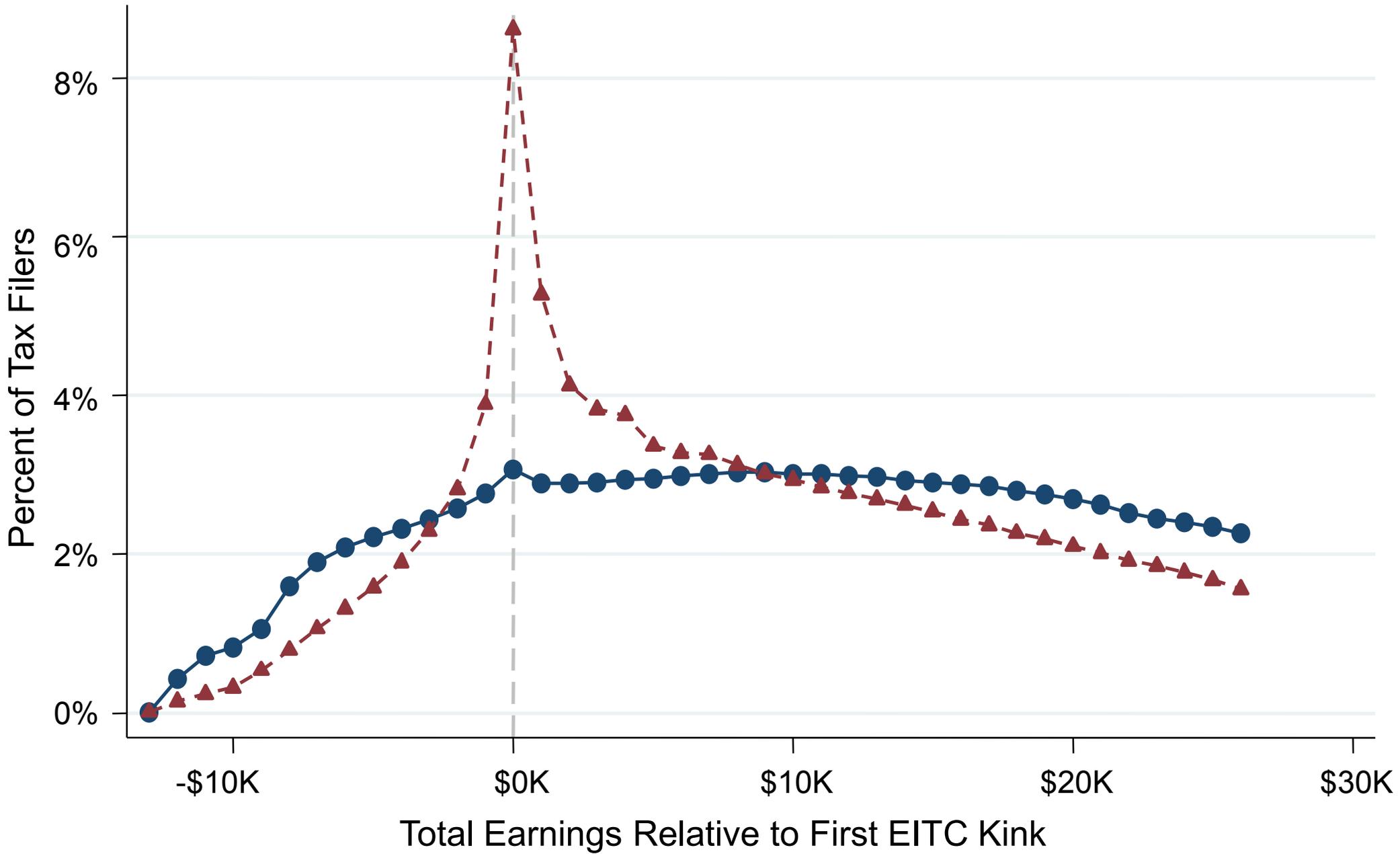
Source: Saez (2010), p. 184

Panel B. Two or more children



Source: Saez (2010), p. 192

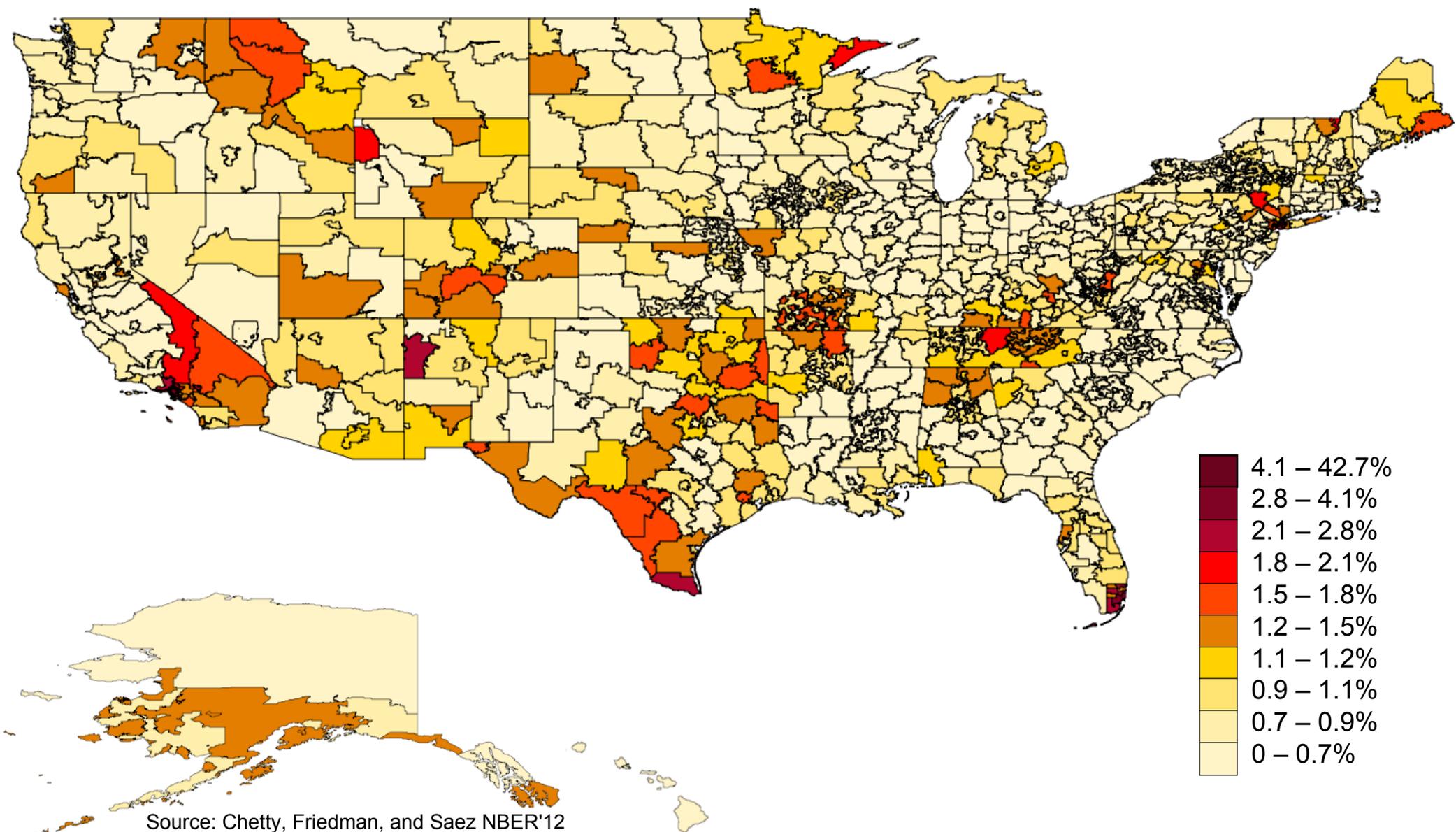
Earnings Distributions in Lowest and Highest Bunching Deciles



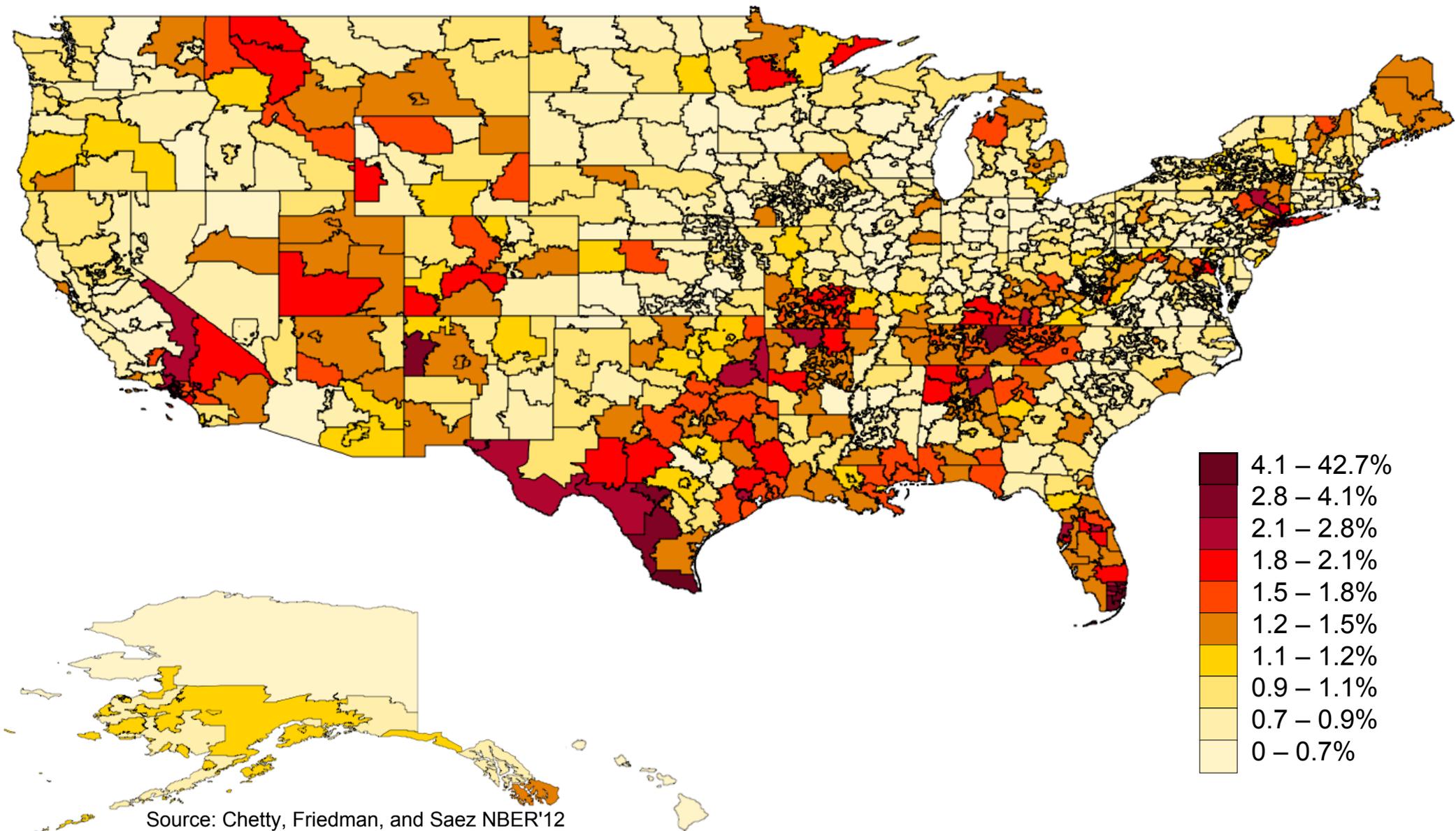
Source: Chetty, Friedman, and Saez NBER'12

Lowest Bunching Decile — Highest Bunching Decile

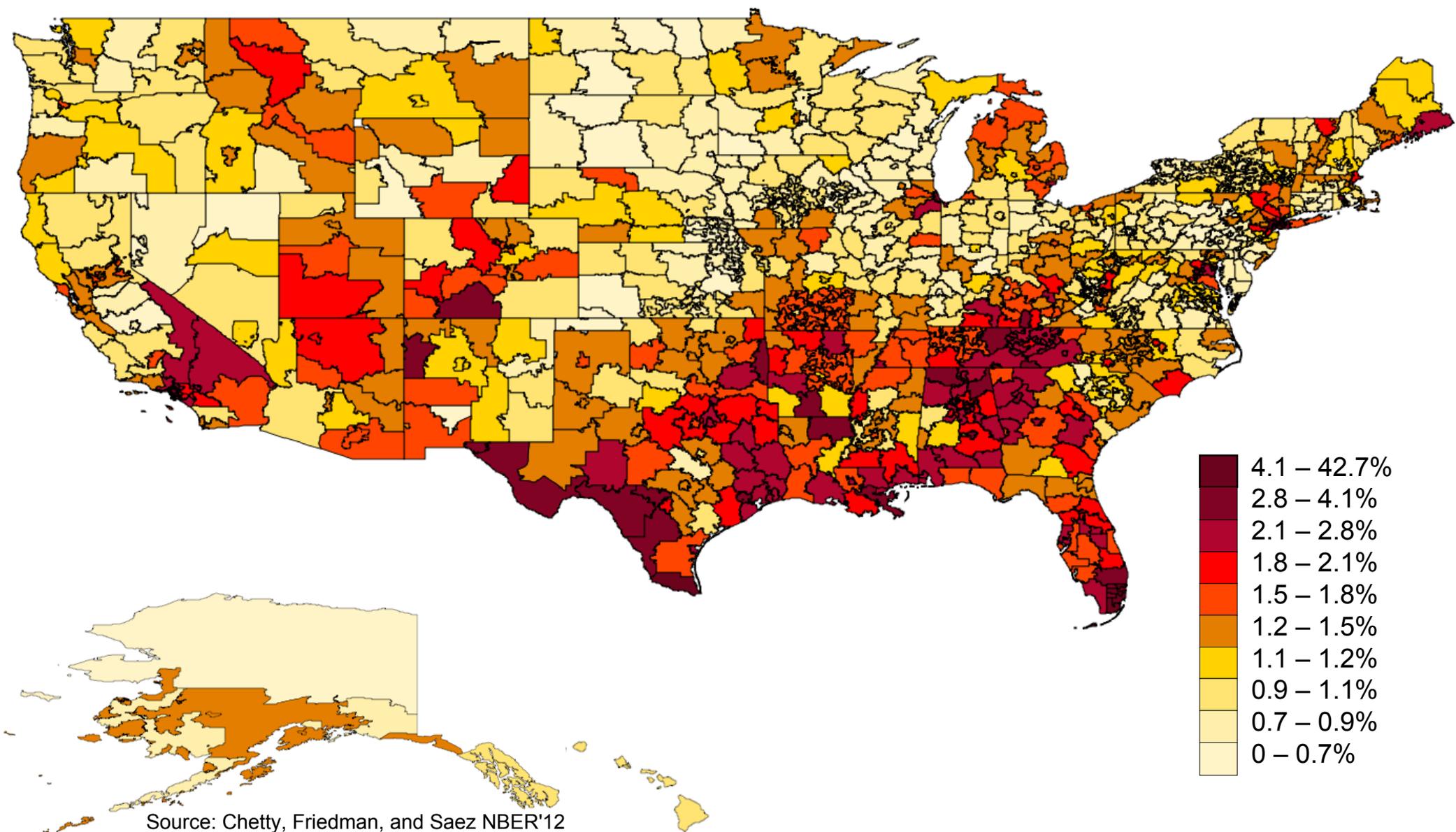
Fraction of Tax Filers Who Report SE Income that Maximizes EITC Refund in 1996



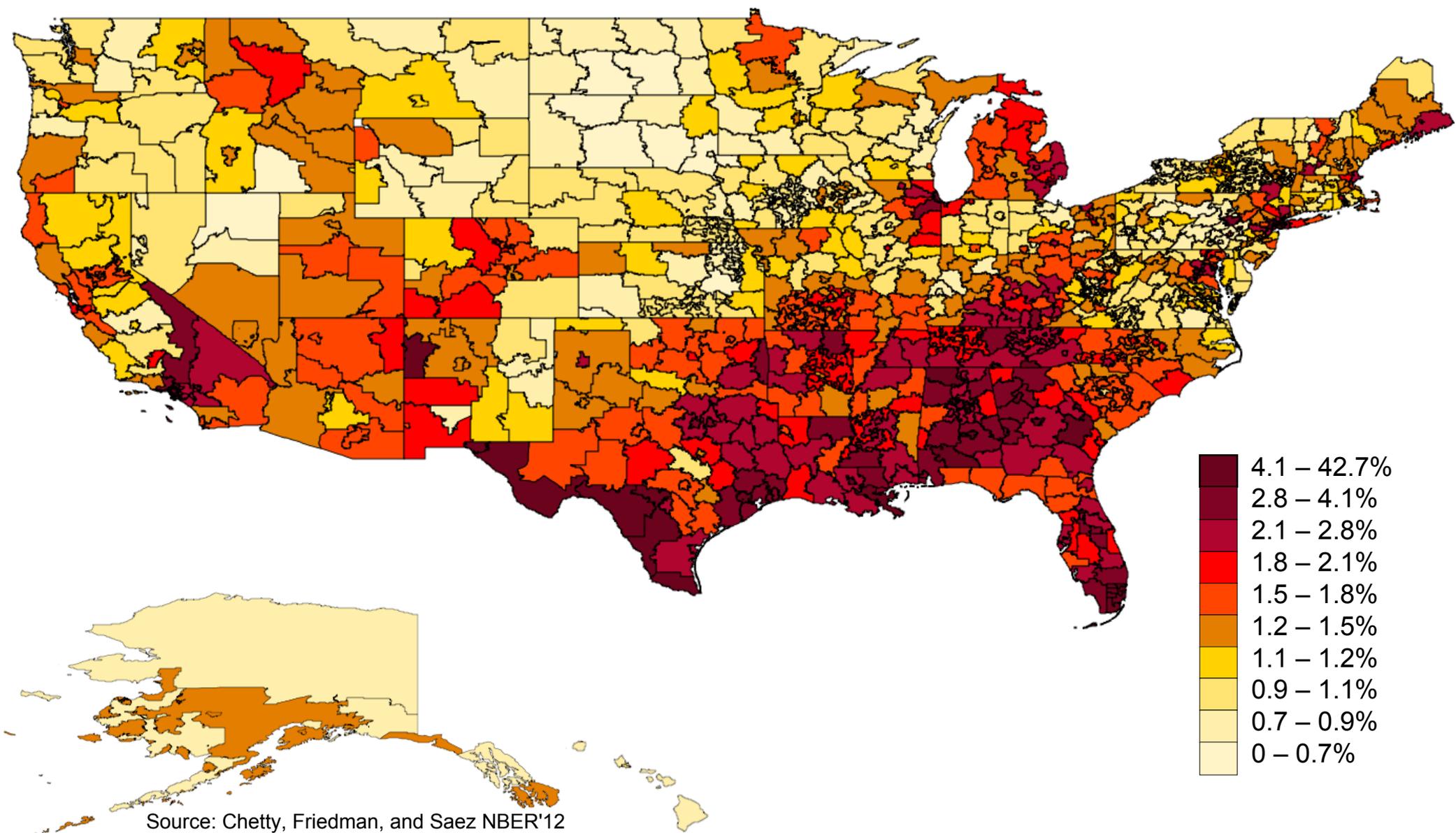
Fraction of Tax Filers Who Report SE Income that Maximizes EITC Refund in 1999



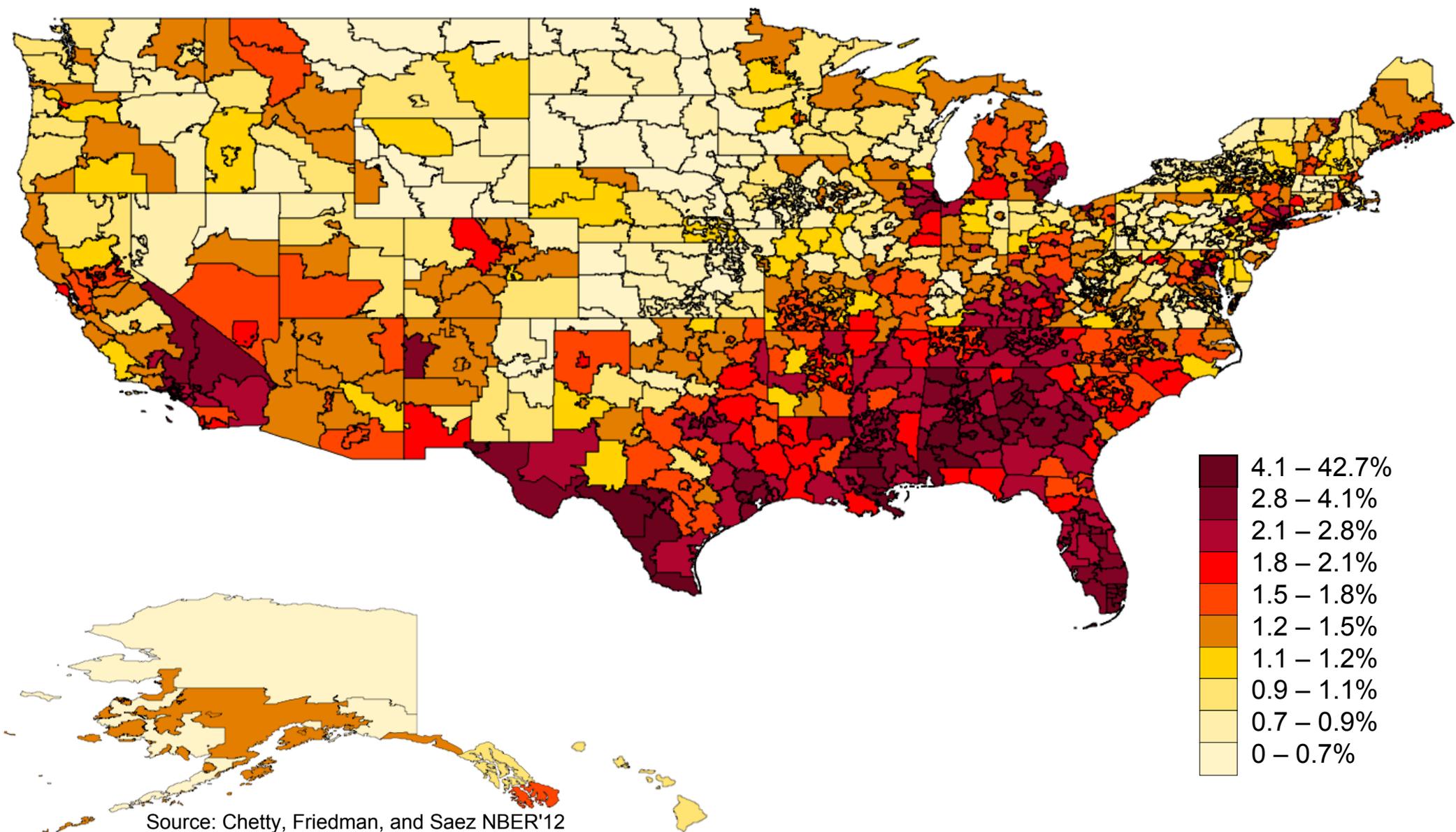
Fraction of Tax Filers Who Report SE Income that Maximizes EITC Refund in 2002



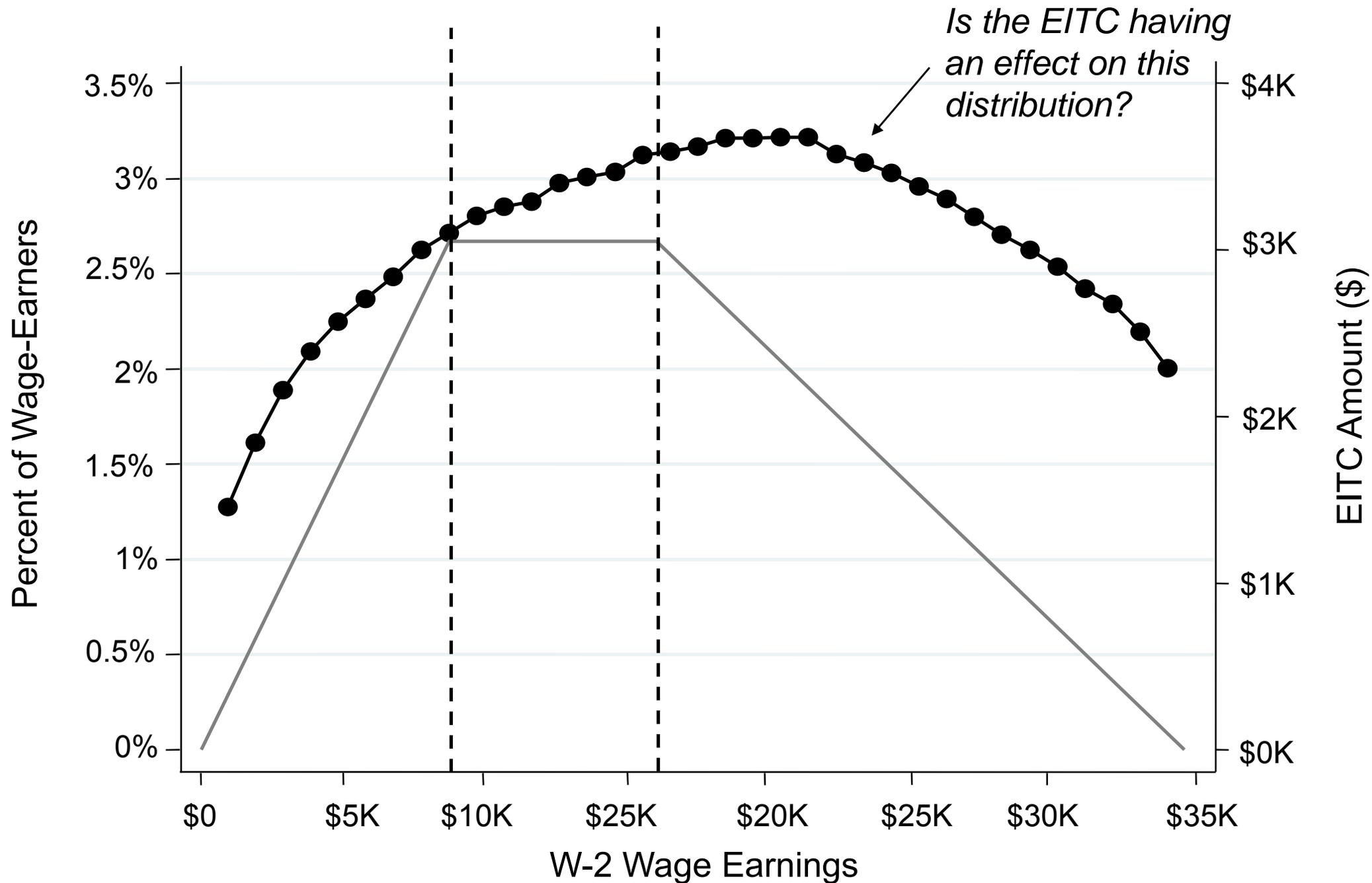
Fraction of Tax Filers Who Report SE Income that Maximizes EITC Refund in 2005



Fraction of Tax Filers Who Report SE Income that Maximizes EITC Refund in 2008

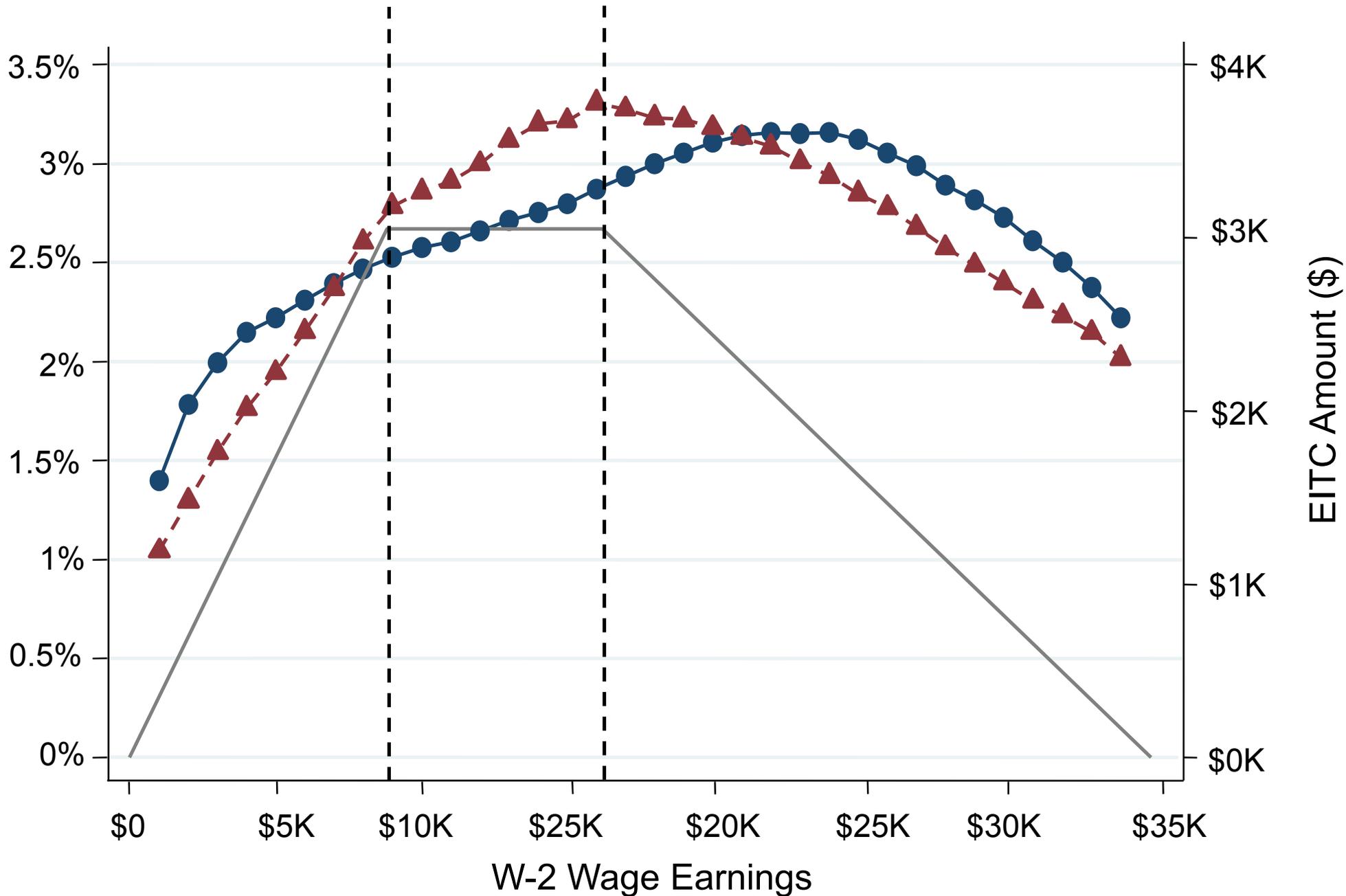


Income Distribution For Single Wage Earners with One Child



Source: Chetty, Friedman, and Saez NBER'12

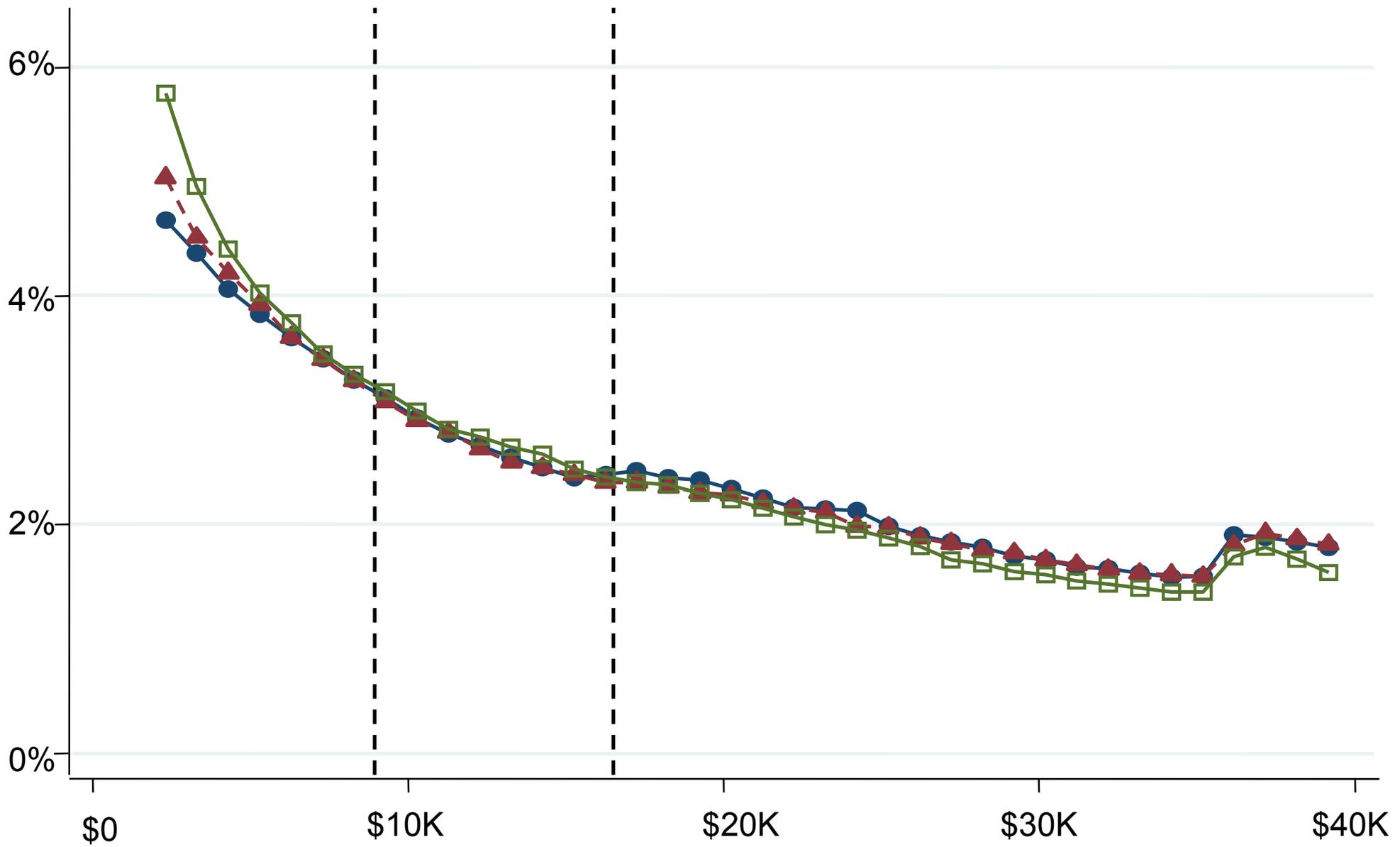
Income Distribution For Single Wage Earners with One Child High vs. Low Bunching Areas



Source: Chetty, Friedman, and Saez NBER'12

Lowest Bunching Decile
Highest Bunching Decile

Earnings Distribution in the Year Before First Child Birth for Wage Earners



Wage Earnings

Source: Chetty, Friedman, and Saez NBER'12

—●— Lowest Sharp Bunching Decile —▲— Middle Sharp Bunching Decile —□— Highest Sharp Bunching Decile

Earnings Distribution in the Year of First Child Birth for Wage Earners

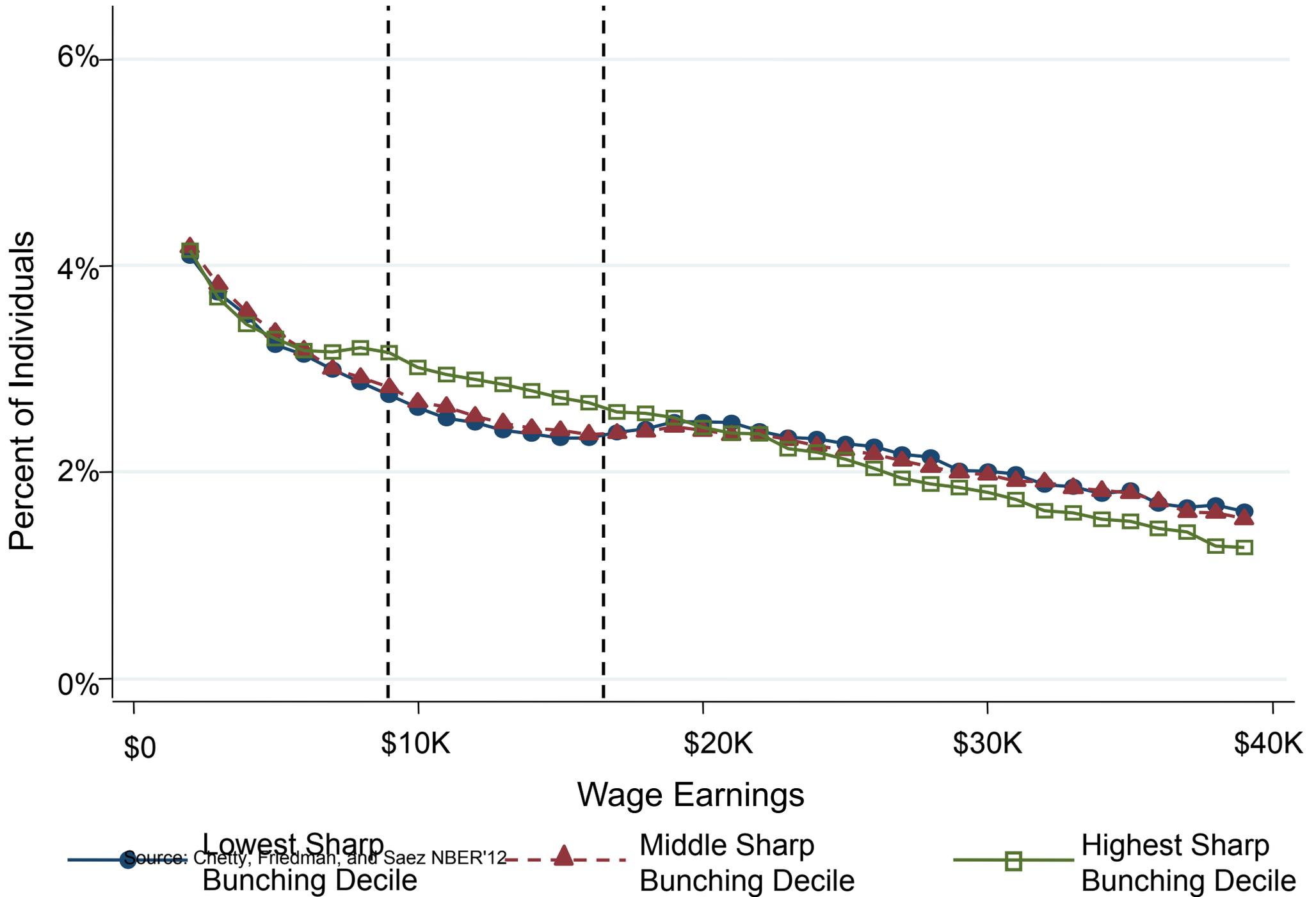


TABLE 4—IMPACT OF EITC ON WAGE EARNINGS DISTRIBUTION OF EITC-ELIGIBLE HOUSEHOLDS

	Percent of EITC-eligible households below threshold			
	50 percent of poverty line	100 percent of poverty line	150 percent of poverty line	200 percent of poverty line
	(1)	(2)	(3)	(4)
No EITC counterfactual	13.15	31.31	53.81	77.06
EITC with no behavioral response	8.92	21.37	41.56	70.82
EITC with avg. behavioral response in United States	8.16	21.00	41.97	71.29
EITC with top decile behavioral response	6.73	20.24	42.56	72.08

Source: Chetty-Friedman-Saez (2013)

Chetty-Friedman-Saez: Results

- 1) Phase-in elasticity (increases earnings of very poor): 0.31.
Phase-out elasticity (decreases earnings of moderately poor): 0.14
- 2) Reduces poverty on net

Welfare reform and consumption: Meyer and Sullivan 2004

U.S. poverty rate defined by money income, not consumption

1) Examine the consumption patterns of single mothers and their families from 1984–2000 using CEX data

2) Question: Did combination of EITC and welfare reform raise, lower, or have no effect on consumption of single mothers?

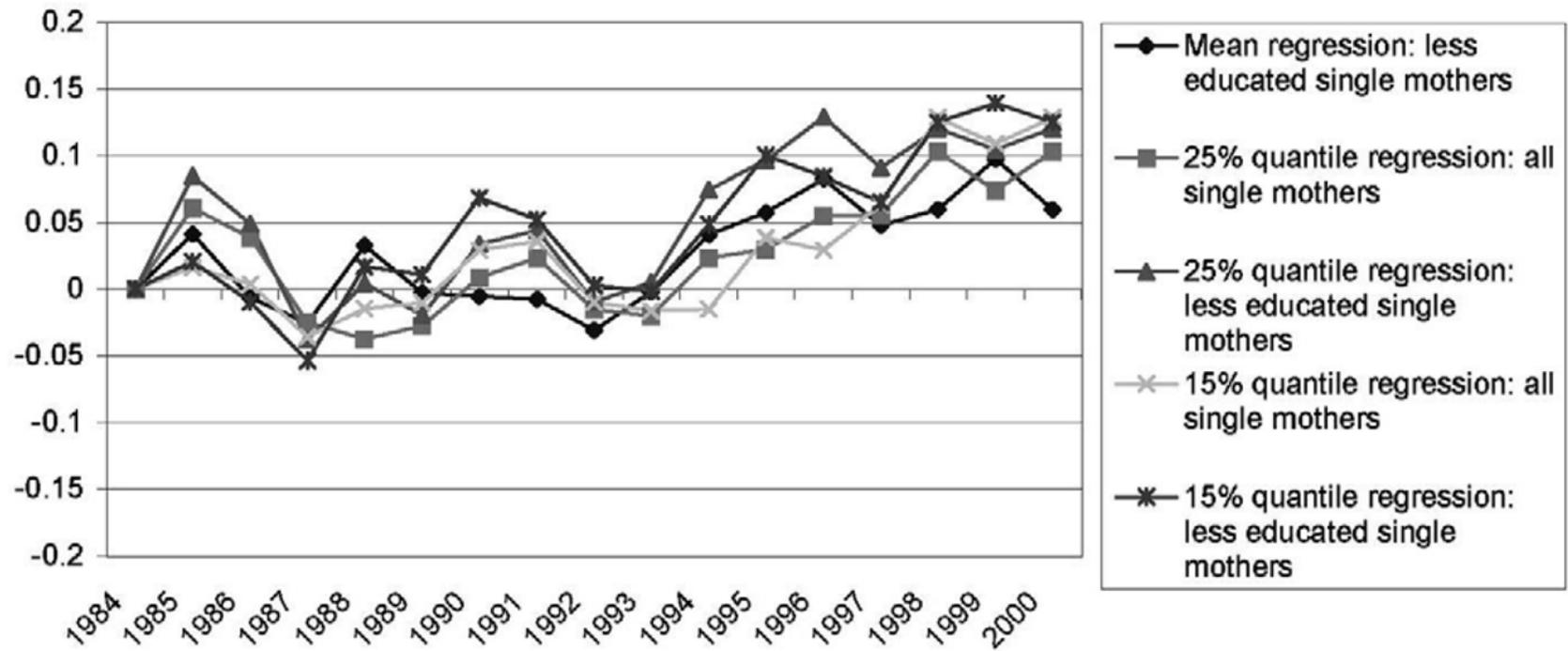


Fig. 2. Total consumption: single mothers, 1984–2000.

Source: Meyer and Sullivan (2004), p. 1407

Meyer and Sullivan: Results

1) Material conditions of single mothers did not decline in 1990s, either in absolute terms or relative to single childless women

2) In most cases, evidence suggests that the material conditions of single mothers have improved slightly

Q: Is this because economy was booming in 1990s?

Q: If no change in consumption, do we really want single mothers spending less time raising their kids and more time working?

Q: Is workfare approach still suitable? [SNAP households surged from 12M in '07 to 20M in '10 while TANF households increased slightly from 1.7M in '07 to 1.85M in '10]

EITC Follow-up #2: Cash vs. In-Kind Redistribution

Most means-tested transfers are in-kind and often rationed (health care, child care, education, public housing, nutrition subsidies) [care not cash San Francisco reform]

1) Rational Individual perspective:

(a) If in-kind transfer is **tradeable** at market price \Rightarrow in-kind equivalent to cash

(b) If in-kind transfer **non-tradeable** \Rightarrow in-kind inferior to cash

Cash transfer preferable to in-kind transfer from individual perspective

EITC Follow-up #2: Cash vs. In-Kind Redistribution

2) **Social perspective:** 4 justifications:

a) Commodity Egalitarianism: some goods (education, health, shelter, food) seen as **rights** and ought to be provided to all

b) Paternalism: society imposes its preferences on recipients [recipients prefer cash]

c) Behavioral: Recipients do not make choices in their best interests (self-control, myopia) [recipients understand that in-kind is better for them]

d) Efficiency: It could be efficient to give in-kind benefits if it can prevent those who don't really need them from getting them (i.e., force people to queue to get free soup kitchen)

EITC Follow-up #3: Who Enjoys the Incidence?

General Equilibrium: Influx of low-wage workers \Rightarrow Drives down all low-wage workers' wages \Rightarrow Firms capture part of the EITC

Argument: For every \$1 of EITC, single mothers keep \$0.70 and EITC-ineligible workers lose \$0.42 \Rightarrow Firms capture \$0.72 (Rothstein 2008, 2010)!

Evidence questionable but principle still relevant: Future EITC expansions that bring (or keep) more workers into the labor force may generate less benefits to poor

EITC Follow-up #4: Intergenerational Welfare Culture

Conservative concern that welfare promotes a culture of dependency: kids growing up in welfare supported families are more likely to use welfare

Correlation in welfare use across generations is obviously not necessarily causal

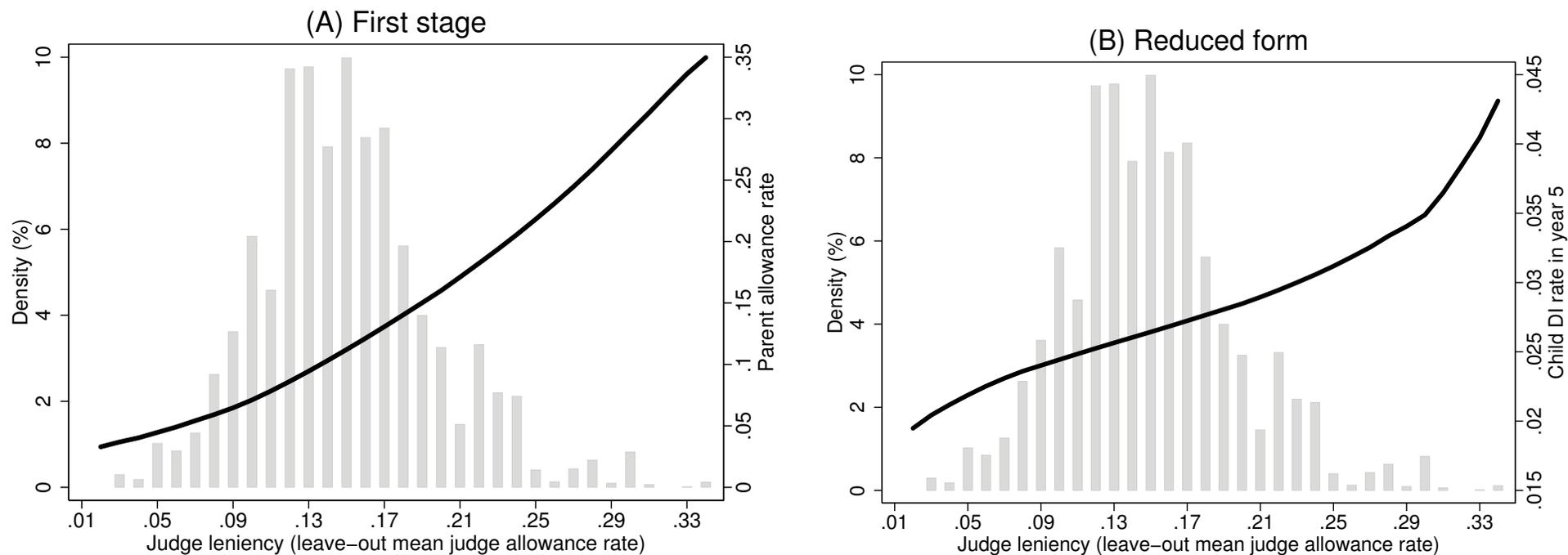
Dahl, Kostol, Mogstad (2013) analyze causal effect of parental use of Disability Insurance (DI) on children use (as adults) of DI in Norway

Identification uses random assignment of judges to denied DI applicants who appeal [some judges severe, others lenient]

Find evidence of causality: parents on DI increases odds of kids on DI over next 5 years by 6 percentage points

Mechanism seems to be learning about DI availability rather than reduced stigma from using DI [because no effect on other welfare programs use]

Figure 3: Effect of Judge Leniency on Parents (First Stage) and Children (Reduced Form).



Notes: Baseline sample, consisting of parents who appeal an initially denied DI claim during the period 1989-2005 (see Section 3 for further details). There are 14,893 individual observations and 79 different judges. Panel (A): Solid line is a local linear regression of parental DI allowance on judge leniency. Panel (B): Solid line is a local linear regression of child DI receipt on their parent's judge leniency measure. All regressions include fully interacted year and department dummies. The histogram of judge leniency is shown in the background of both figures (top and bottom 0.5% excluded from the graph).

Source: Dahl, Kostol, Mogstad (2013)

Top Taxable Income Elasticity

Modern public finance literature focuses on taxable income elasticities instead of hours/participation elasticities, most relevant at top

Two main reasons:

1) What matters for policy is the total behavioral response of taxable income to tax rates (not only hours of work but also occupational choices, avoidance, etc.)

2) Data availability: taxable income is precisely measured in tax return data

Recent overview of this literature: Saez-Slemrod-Giertz JEL'12

Tax Reforms

Tax rates change frequently over time

Biggest tax rate changes have happened at the top

Key recent reforms:

Reagan I: ERTA'81: top rate ↓ 70% to 50% (1981-1982)

Reagan II: TRA'86: top rate ↓ 50% to 28% (1986-1988)

Clinton: OBRA'93: top rate ↑ 31% to 39.6% (1992-1993)

Bush: EGTRRA '01: top rate ↓ 39.6% to 35% (2001-2003)

Obama: top rate ↑ 35% to 39.6% (2012-2013)

Taxable Income = Ordinary Income + Realized Capital Gains
- Deductions ⇒ Each component can respond to *MTRs*

Basic DD Doesn't Deliver What You Want

Goolsbee JPE'00 analyzes CEO pay around the 1993 Clinton top tax rate increase [from 31% in 1992 to 39.6% in 1993 announced in late 1992]

Find a strong taxable income effect via re-timing: exercising stock options (executives can choose the timing of their stock-option exercises)

⇒ Large short-term response due to re-timing, much smaller long-term response

Economists will study whether 2013 Obama top tax rate has generated income shifting from 2013 to 2012

TABLE 2
 AVERAGE COMPENSATION BY TYPE FOR HIGH-INCOME EXECUTIVES
 (in Thousands)

	1991	1992	1993	1994	1995
Taxable income	911	1,153	974	965	1,173
Salary	347	336	336	351	373
Bonus	198	207	241	284	330
LTIP payout	57	72	57	64	89
Options exercised	268	496	293	235	381
Other income (nontaxed)	36	37	66	54	78

SOURCE.—Author's calculations for executives with permanent income greater than \$275,000 per year.

Inter-Temporal Substitution: Stock Options

Major form of compensation of US top executives. Theoretical goal is to motivate executives to increase the value of the company (stock price $P(t)$)

Stock-option is granted at date t_0 allow executives to buy N company shares at price $P(t_0)$ on or after t_1 (in general $t_1 - t_0 \simeq 3 - 5$ years = vesting period)

Executive exercise option at (chosen) time $t_2 \geq t_1$: pays $N \cdot P(t_0)$ to get shares valued $N \cdot P(t_2)$. Exercise profit $N[P(t_2) - P(t_0)]$ (considered and taxed as wage income in the US)

After t_2 , executive owns N shares, eventually sold at time $t_3 \geq t_2$: realized capital gain $N[P(t_3) - P(t_2)]$ (taxed as capital gains)

Inter-Temporal Substitution: Realized Capital Gains

Realized capital gains occur when individual sells an asset [e.g., stock] at a higher price than buying price

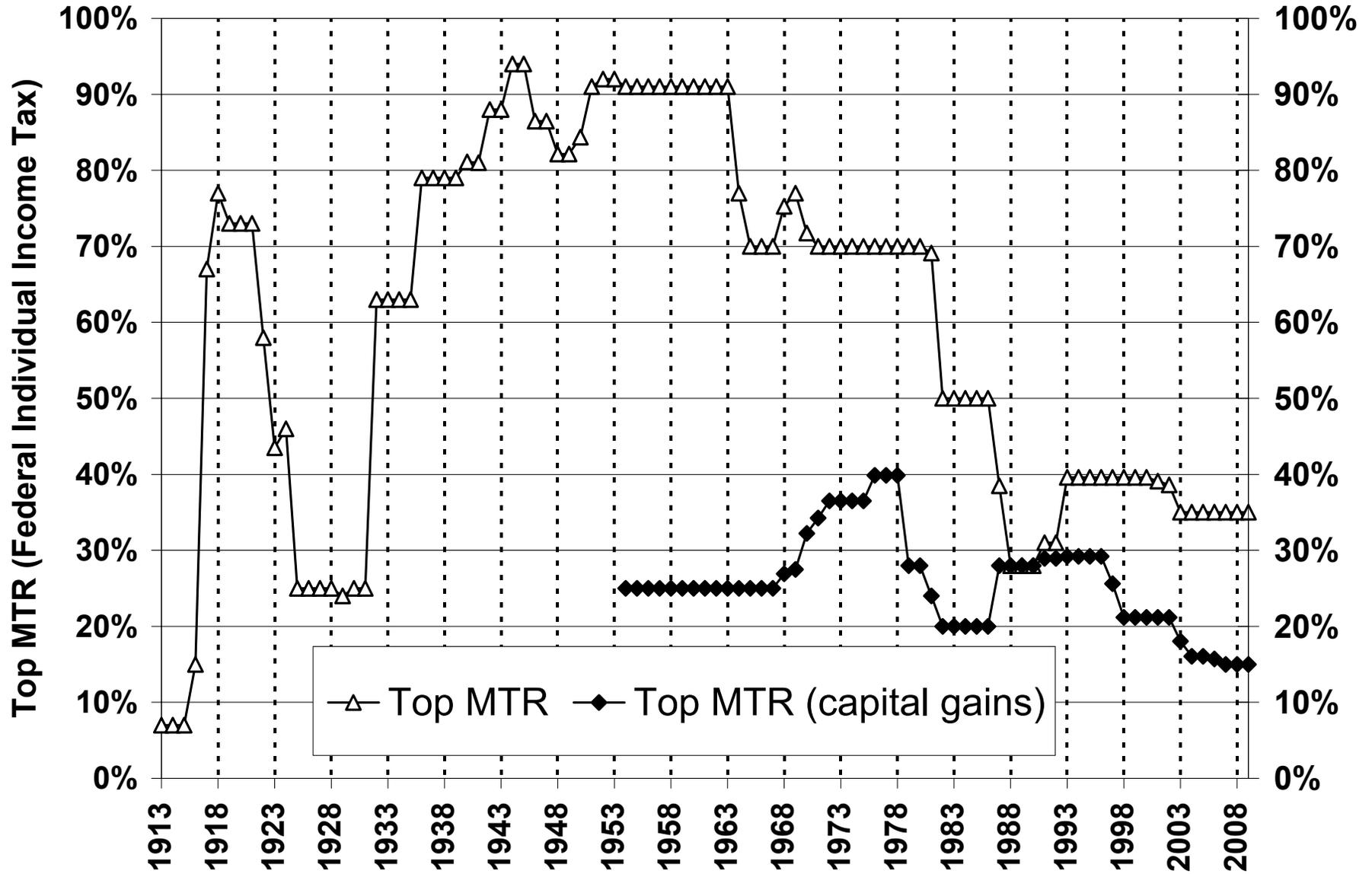
Individuals have flexibility in the timing of asset sales and capital gains realizations (any asset, not just stock options)

Tax Reform Act of 1986 lowered the top tax rate on ordinary income from 50% to 28% but increased the top tax rate on realized capital gains from 20% to 28%

⇒ Surge in capital gains realizations in 1986 [and depressed capital gains in 1987] to take advantage of low 20% rate before 28% tax rate applies

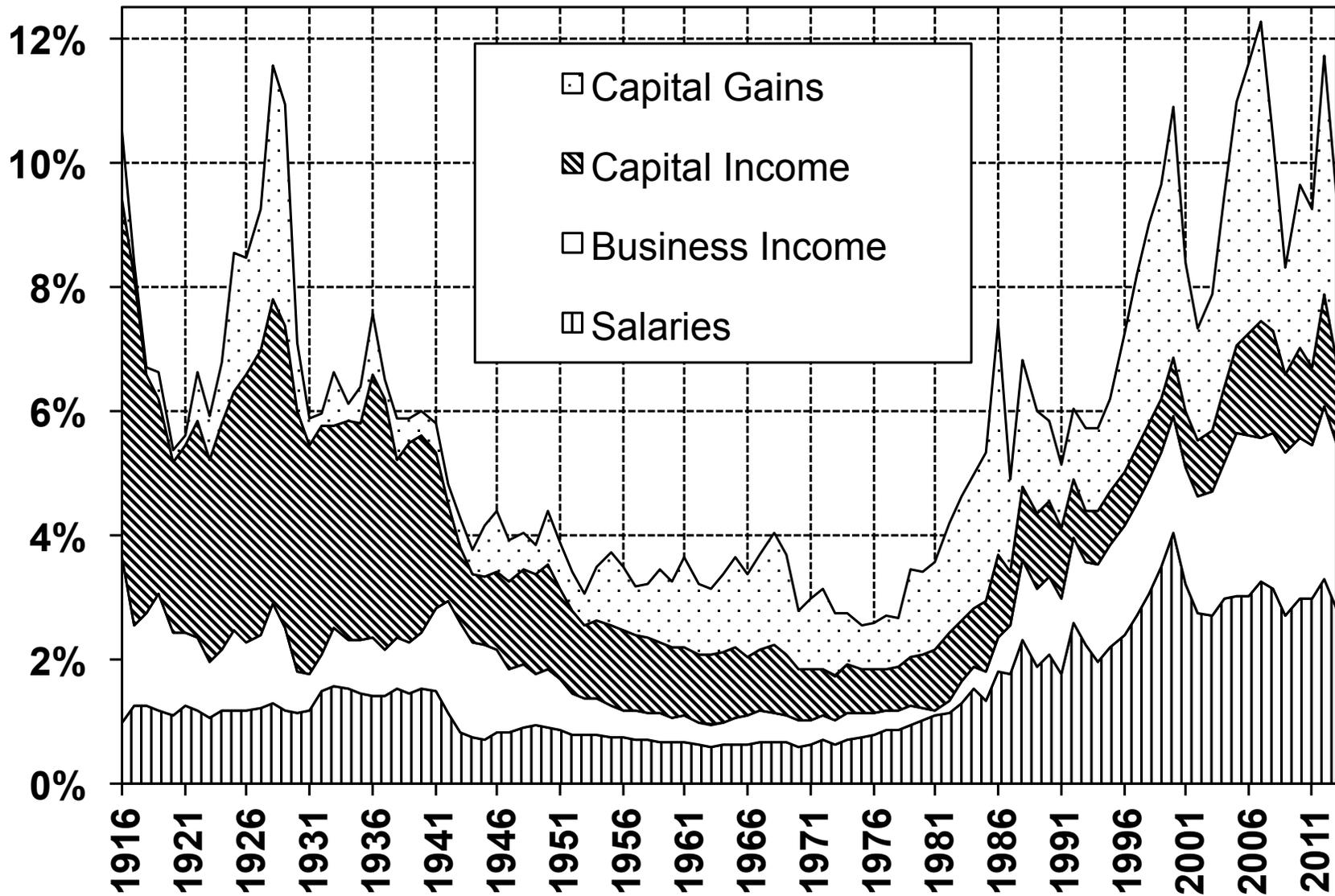
⇒ Short-term elasticity is very large but long-term elasticity is certainly much smaller

US Top MTR ordinary income vs. capital gains



Source: Department of the Treasury, Office of Tax Analysis

US Top 0.1% Pre-Tax Income Share and Composition



Source: Piketty and Saez, 2003 updated to 2013. Series based on pre-tax cash market income including or excluding realized capital gains, and always excluding government transfers.

Long-Run Evidence

Goal: evaluate whether top **pre-tax** incomes respond to changes in one minus the marginal tax rate (=net-of-tax rate)

Focus is on **pre-tax** income before deductions and excluding realized capital gains (taxed at lower separate rate, timing flexibility)

Strategy #1: U.S top rate changes

Strategy #2: International top rate changes

Income Share Based Elasticity Estimation

1) **Tax Reform Episode:** Compare top **pre-tax** income shares at t_0 (before reform) and t_1 (after reform)

$$e = \frac{\log sh_{t_1} - \log sh_{t_0}}{\log(1 - \tau_{t_1}) - \log(1 - \tau_{t_0})}$$

where sh_t is top income share and τ_t is the average MTR for top group in year t

Identification assumption: absent tax change, $sh_{t_0} = sh_{t_1}$

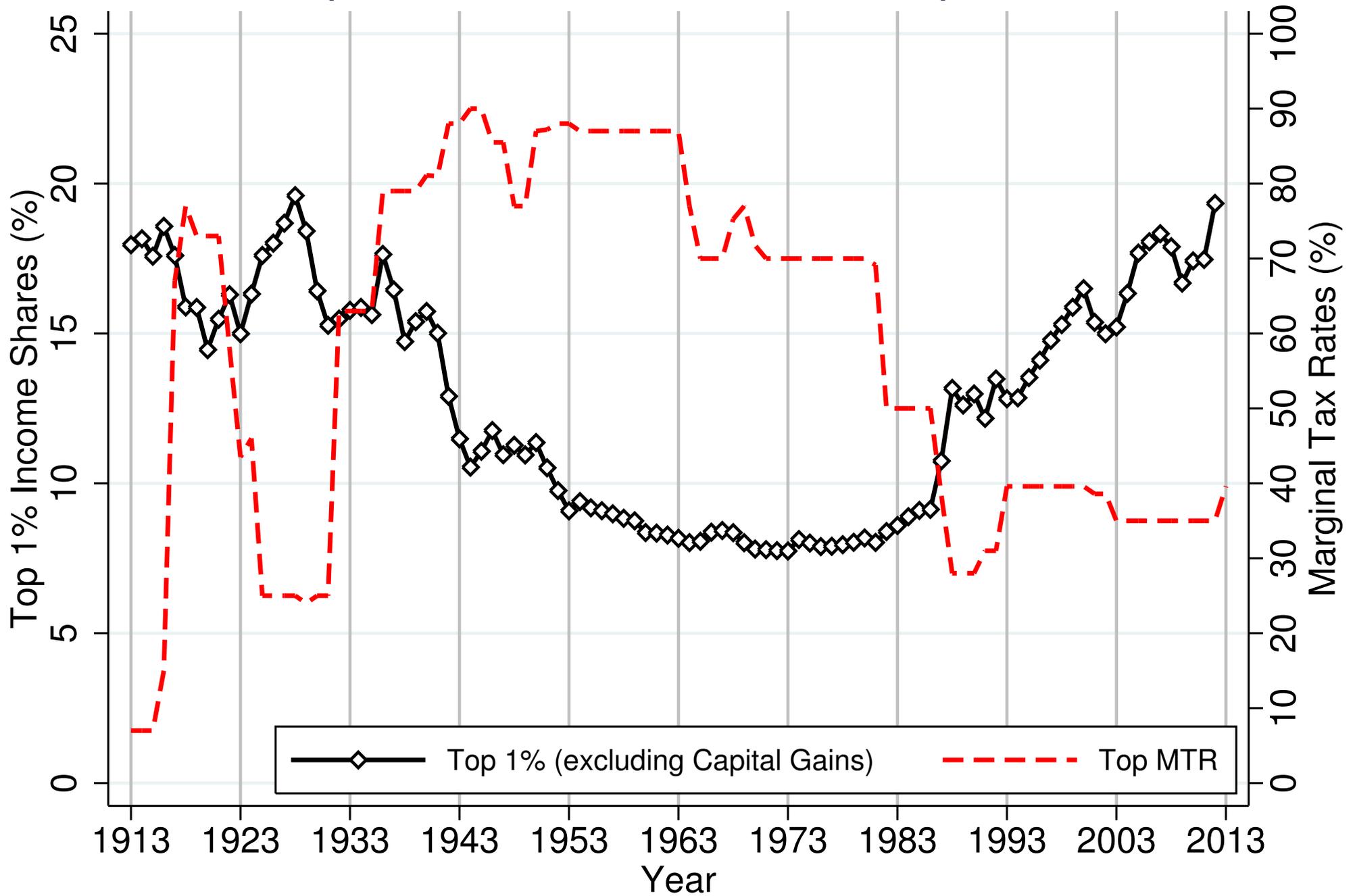
2) **Full Time Series:** Run regression:

$$\log sh_t = \alpha + e \cdot \log(1 - \tau_t) + \varepsilon_t$$

and adding time controls to capture non-tax related top income share trends

Identification assumption: non-tax related changes in $sh_t \perp \tau_t$

Top 1% Income Share and Top MTR



Long-Run U.S. Evidence

Top-1% income share elasticity of 1.8, but 0.5 when controlling for time trends

Red flags in the time trends: (a) top share keeps rising after 1986, and (b) top share doesn't move in lock-step (little change after 1981, wrong sign after 1993)

Key issue: Skill-biased technical change (e.g. computers) may have been increasing top-incomes at the same time \Rightarrow Counterfactual (what would have happened in absence of tax cut) is unknown

Separate issue: Shifting from corporate tax base to personal tax base (talk about later)

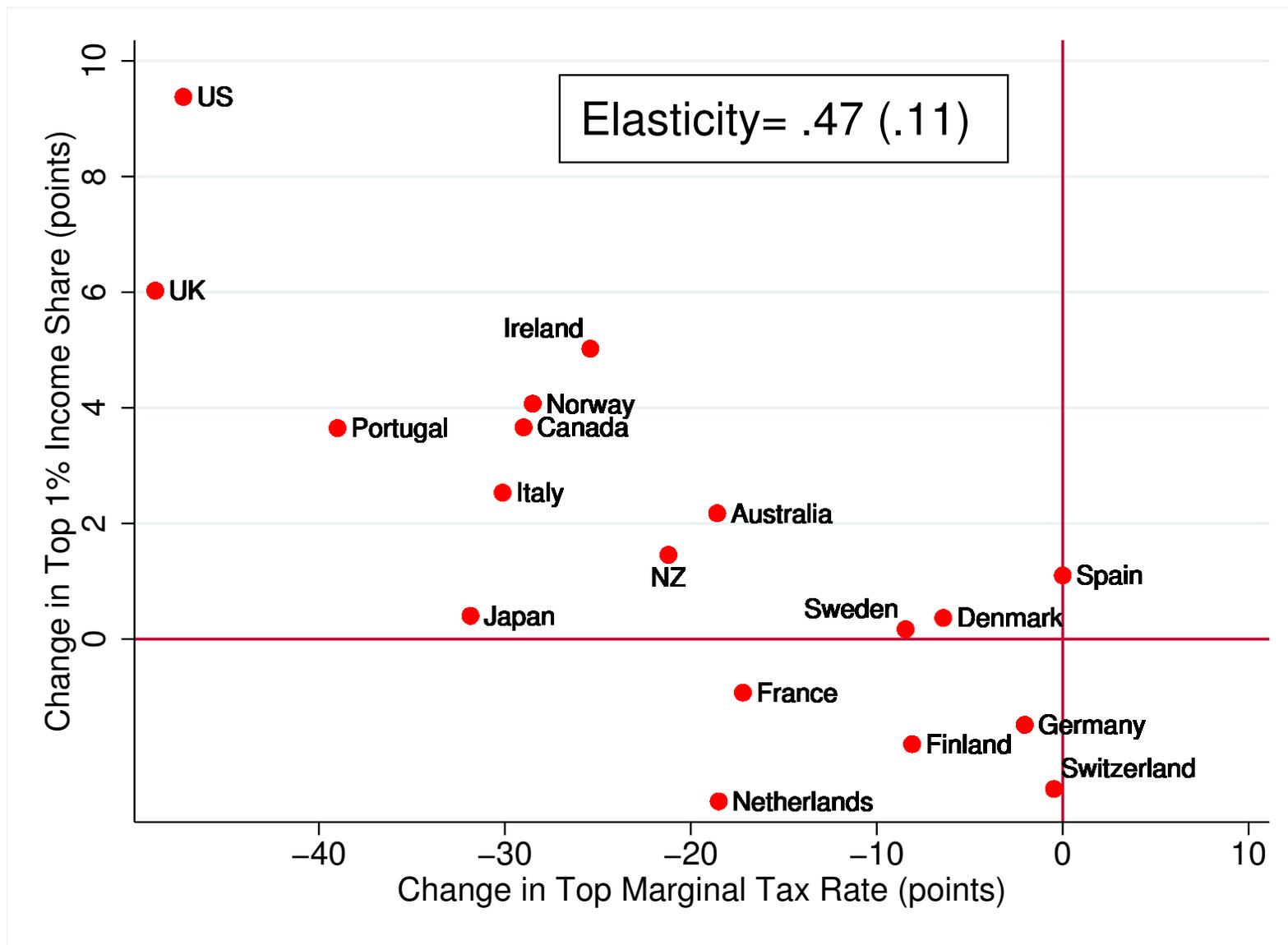
Long-Run International Evidence

1) Use pre-tax top 1% income share data from 18 OECD countries since 1960 using the **World Top Incomes Database**

2) Compute top (statutory) individual income tax rates using OECD data [including both central and local income taxes].

Plot change in top 1% pre-tax income share (between 1960-4 and 2005-9) against change in top MTR (between 1960-4 and 2005-9)

Key advantage: Controls flexibly for time trends (every country got computers but only U.S. and U.K. had huge reductions in the top MTR)



Change in Top Tax Rate and Top 1% Share, 1960-4 to 2005-9

Top Taxable Income Elasticity

Best current estimates: around 0.5

Still considerable uncertainty due to lack of counterfactual

Implies optimal top tax rate of $1/(1 + 1.5 * .5) = 57\%$ if gov does not value their consumption

Current top tax rate: 40% Federal + 4% Medicare + about 8% state/local = 52%

Top Tax Rate Follow-Up #1: Externalities?

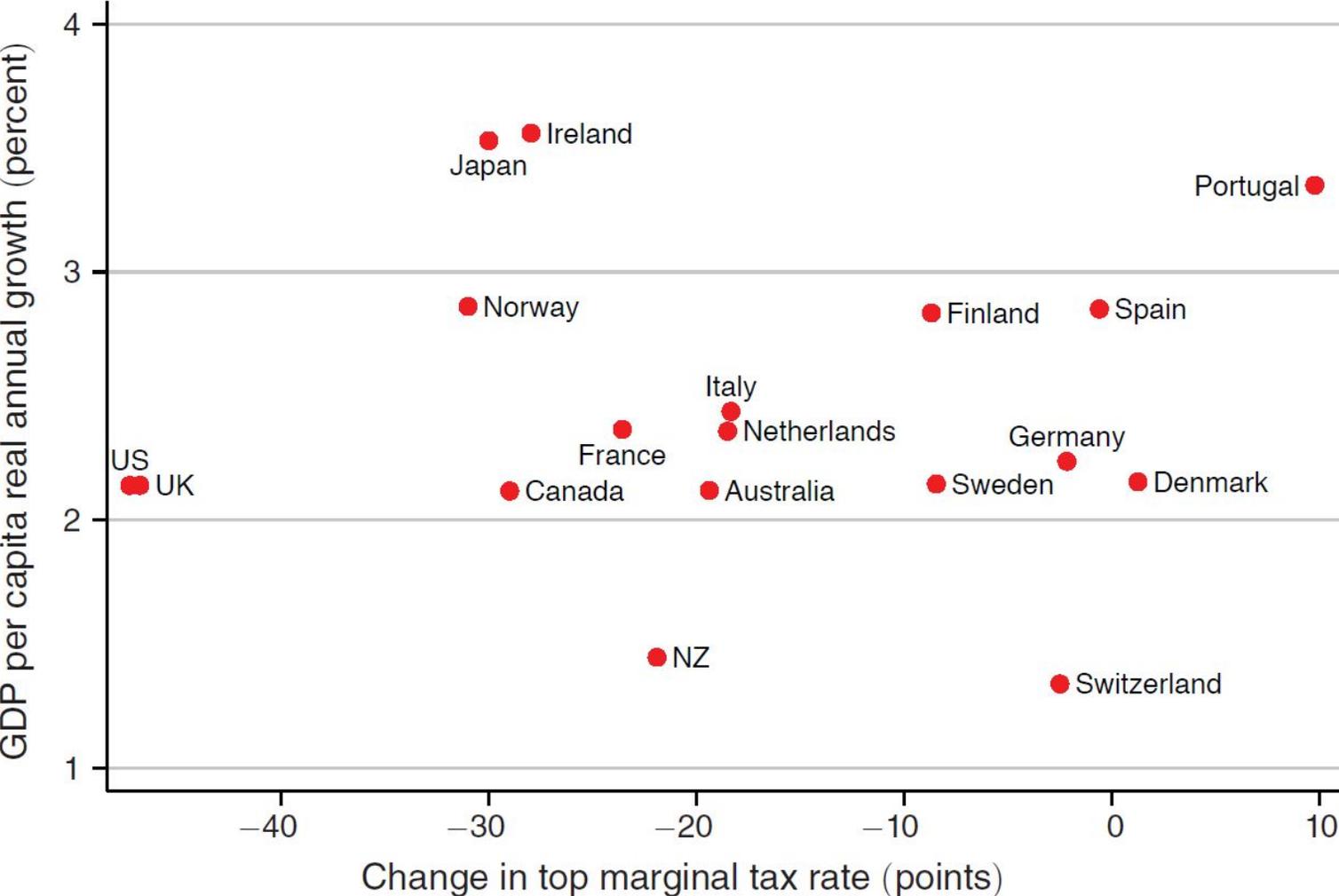
1) If top-1% give a lot in charity or are “job creators” \Rightarrow Their income/consumption have positive externalities \Rightarrow Lower optimal rate

Top 1% do give some to charity, noisy job creation evidence but even small effect would be very important

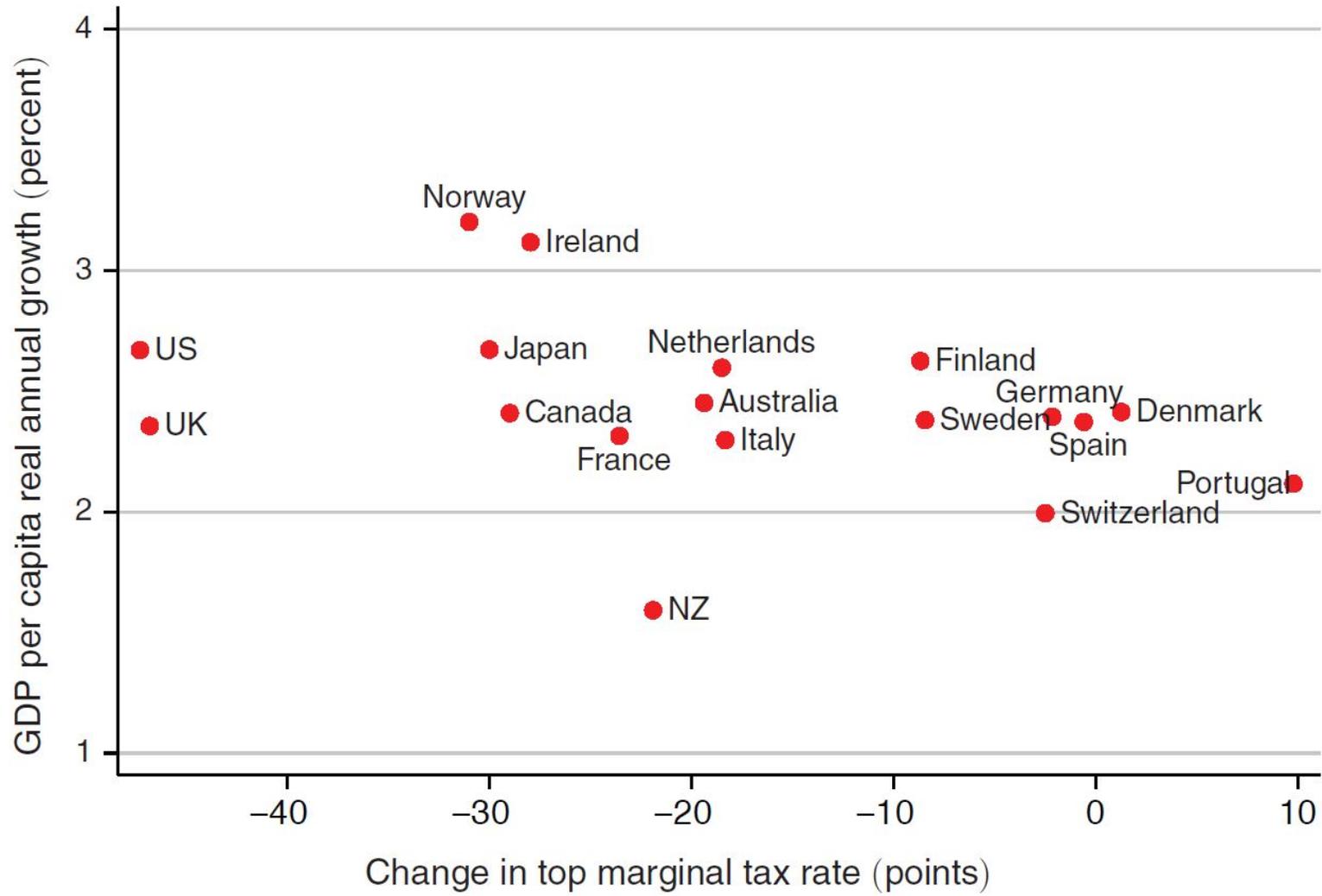
2) If top-1% extract pay at expense of the 99% when tax rates are low \Rightarrow Their income has a negative externality \Rightarrow Higher optimal rate

Evidence consistent with CEOs exploiting poor governance to extract more when tax rates are low (Piketty-Saez-Stantcheva-2014)

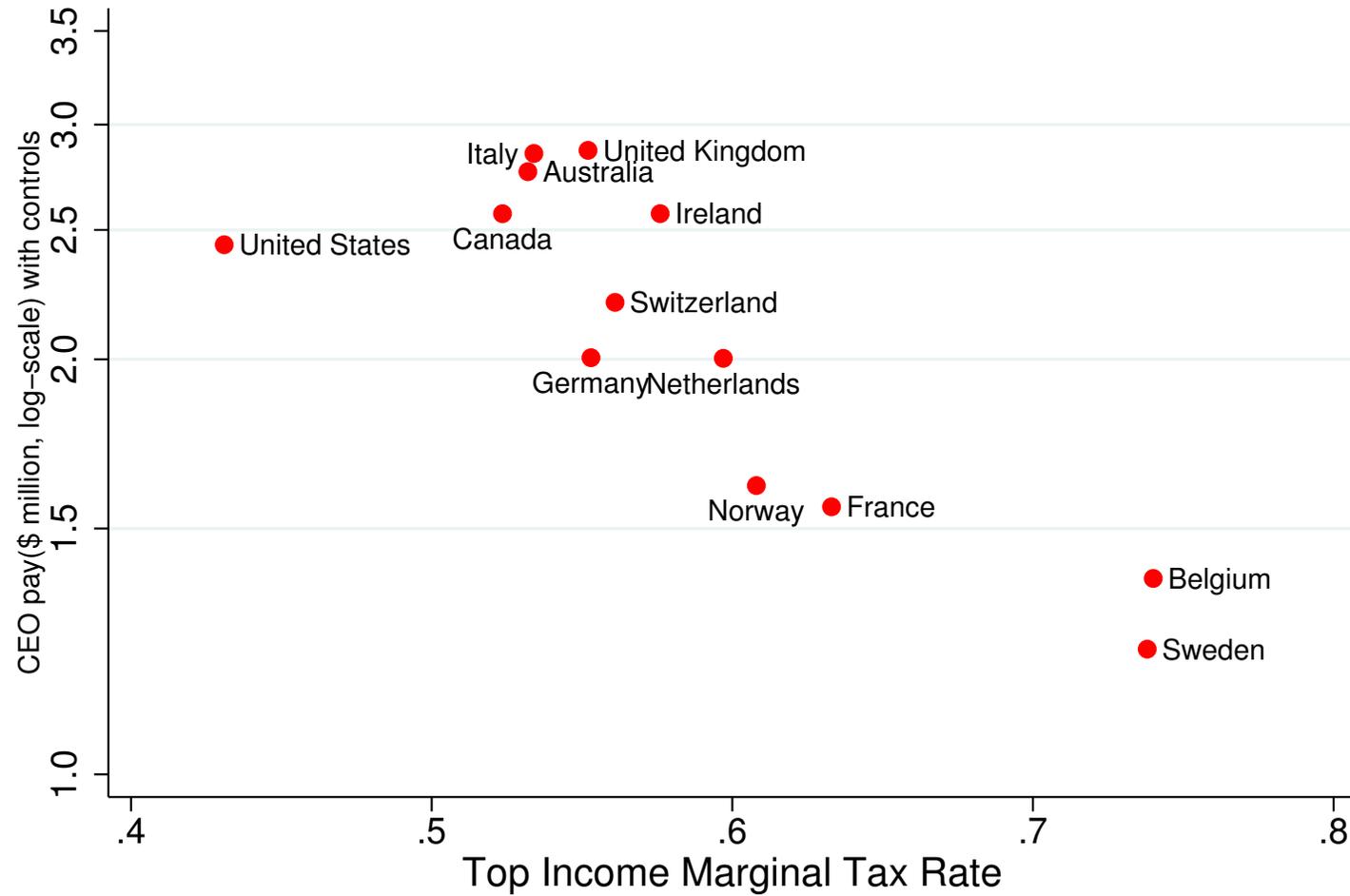
Panel A. Growth and change in top marginal tax rate



Panel B. Growth (adjusted for initial 1960 GDP)



B. Average CEO compensation with controls



Controlling for firm profitability, governance, size, and industry

Top Tax Rate Follow-Up #2: International Migration

Public debate concern that top skilled individuals move to low tax countries (e.g., within the EU) or low tax States (within US)

Migration concern bigger in public debate than concern about supply-side within a country

Little work on tax induced international migration of top skilled workers (see Young et al. 14 within US states)

Hard to get data but interesting variation due to proliferation of special low tax schemes for highly paid foreigners in Europe

Kleven-Landais-Saez AER'13 look at football players in Europe (highly mobile group, many tax reforms) \Rightarrow Find significant migration responses to taxes after football market was de-regulated in '95

KLEVEN-LANDAIS-SAEZ-SCHULTZ QJE'14

Exploit the 1991 tax scheme in Denmark: high earnings immigrants ($\geq 103,000$ Euros/year) taxed at flat 25% rate (instead of regular tax with top 59% rate) for 3 years

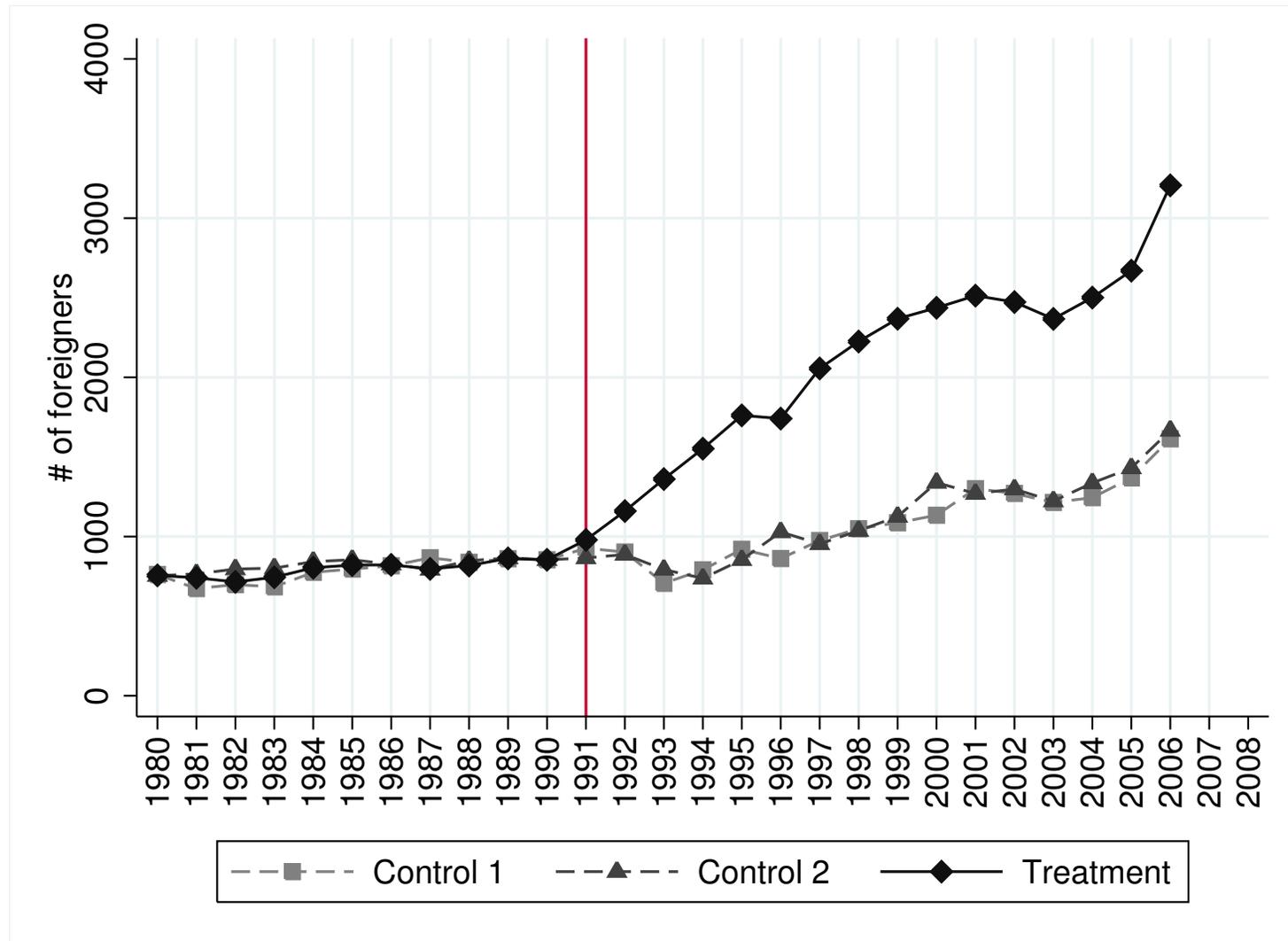
Use population wide Danish tax data and DD strategy: compare immigrants above eligibility earnings threshold (treatment) to immigrants slightly below threshold (control)

Key Finding: Scheme doubles the number of highly paid foreigners in Denmark relative to controls

⇒ Elasticity of migration with respect to the net-of-tax rate above one (much larger than the within country elasticity of earnings)

⇒ Tax coordination will be key to preserve progressive taxation in the European Union (like the U.S. states coordinating to redistribute federally)

Figure 3: Total number of foreigners in different income groups



Control 1= annualized income between .8 and .9 of threshold

Control 2= annualized income between .9 and .995 of threshold.

REFERENCES

Atkinson, A., F. Alvaredo, T. Piketty and E. Saez *The World Top Incomes Database*, (web)

Chetty, R., J. Friedman and E. Saez “Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings”, *American Economic Review*, 2013, 103(7), 2683-2721 (web)

Dahl, Gordon B., Andreas Ravndal Kostol, Magne Mogstad “Family Welfare Cultures” NBER Working Paper No. 19237, 2013, forthcoming *Quarterly Journal of Economics* (web) Department of the Treasury(2012) “Capital Gains and Taxes Paid on Capital Gains” (web)

Eissa, Nada and Hilary Hoynes. “Behavioral Responses to Taxes: Lessons from the EITC and Labor Supply”, *Tax Policy and the Economy* Volume 20, 2006. (web)

Goolsbee, A. “What Happens When You Tax the Rich? Evidence from Executive Compensation”, *Journal of Political Economy*, Vol. 108, 2000, 352-378. (web)

IRS, Statistics of Income Division(2013) “U.S. Individual Income Tax: Personal Exemptions and Lowest and Highest Tax Bracket”(web)

Kleven, Henrik, Camille Landais, and Emmanuel Saez “Taxation and International Mobility of Superstars: Evidence from the European Football Market,” *American Economic Review*, 103(5), 2013. (web)

Kleven, Henrik Jacobsen, Camille Landais, Emmanuel Saez, and Esben Anton Schultz. “Migration and Wage Effects of Taxing Top Earners: Evidence from the Foreigners’ Tax Scheme in Denmark.” *Quarterly Journal of Economics* 127(1), (2014).(web)

Meyer, B. and D. Rosenbaum “Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers”, *Quarterly Journal of Economics*, Vol. 116 (2001), 1063-1114. (web)

Meyer, B. and X. Sullivan “The effects of welfare and tax reform: the material well-being of single mothers in the 1980s and 1990s”, *Journal of Public Economics*, Vol. 88 (2004), 1387-1420. (web)

Piketty, T. and E. Saez “Income Inequality in the United States, 1913-1998”, *Quarterly Journal of Economics*, Vol. 116, (2003): 1-39. (web)

Piketty, Thomas, Emmanuel Saez, and Stefanie Stantcheva “Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities,” *American Economic Journal: Economic Policy*, 6(1), 2014. (web)

Rothstein, Jesse. “The Unintended Consequences of Encouraging Work: Tax Incidence and the EITC.” Working paper 2008. (web)

Rothstein, Jesse. "Is the EITC as Good as an NIT? Conditional Cash Transfers and Tax Incidence." *American Economic Journal: Economic Policy* 2010. (web)

Saez, E. "Do Taxpayers Bunch at Kink Points?", *American Economic Journal: Economic Policy*, Vol. 2, 2010, 180-212. (web)

Saez, Emmanuel, Joel Slemrod, and Seth H. Giertz. "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review." *Journal of Economic Literature* 50(1) (2012): 3-50. (web)

Undergraduate Public Economics

Danny Yagan
UC Berkeley

Lecture 9

Taxes on Capital and Savings

MOTIVATION

1) Capital income is about 25-30% of national income (labor income is 70-75%) but distribution of capital income is much more unequal than labor income

Capital income inequality is due to differences in savings behavior but also inheritances received

⇒ Equity suggests it should be taxed more than labor

2) Capital Accumulation correlated strongly with growth [although causality link is not obvious] and capital accumulation might be sensitive to the net-of-tax return.

⇒ Efficiency cost of capital taxation might be high.

MOTIVATION

3) Capital more mobile internationally than labor

Key distinction is **residence** vs. **source** base capital taxation:

Residence: Capital income tax based on residence of owner of capital.

Most individual income tax systems are residence based (with credits for taxes paid abroad)

Incidence falls on the owner \Rightarrow can only escape tax through tax evasion (tax heavens) or changing residence (mobility)

Tax evasion through tax heavens is a very serious concern (Zucman QJE'13)

Source: Capital income tax based on location of capital (corporate income tax except US which taxes worldwide income of US corporations)

Incidence is then partly shifted to labor if capital is mobile

Mechanism: tax on capital, capital flees the country, hurts the wage of workers (as workers are less productive with less capital) \Rightarrow Workers bear part of the burden

4) Capital taxation is extremely complex and provides many tax avoidance opportunities (Gravelle, 94)

FACTS ABOUT WEALTH AND CAPITAL INCOME

Definition: Capital Income = Returns from Wealth Holdings

Aggregate US **Private** Wealth $\simeq 4 \times$ Annual National Income

Housing: residential real estate (land+buildings) [income = rents] net of mortgage debt

Unincorporated business assets: value of sole proprietorships and partnerships [income = individual business profits]

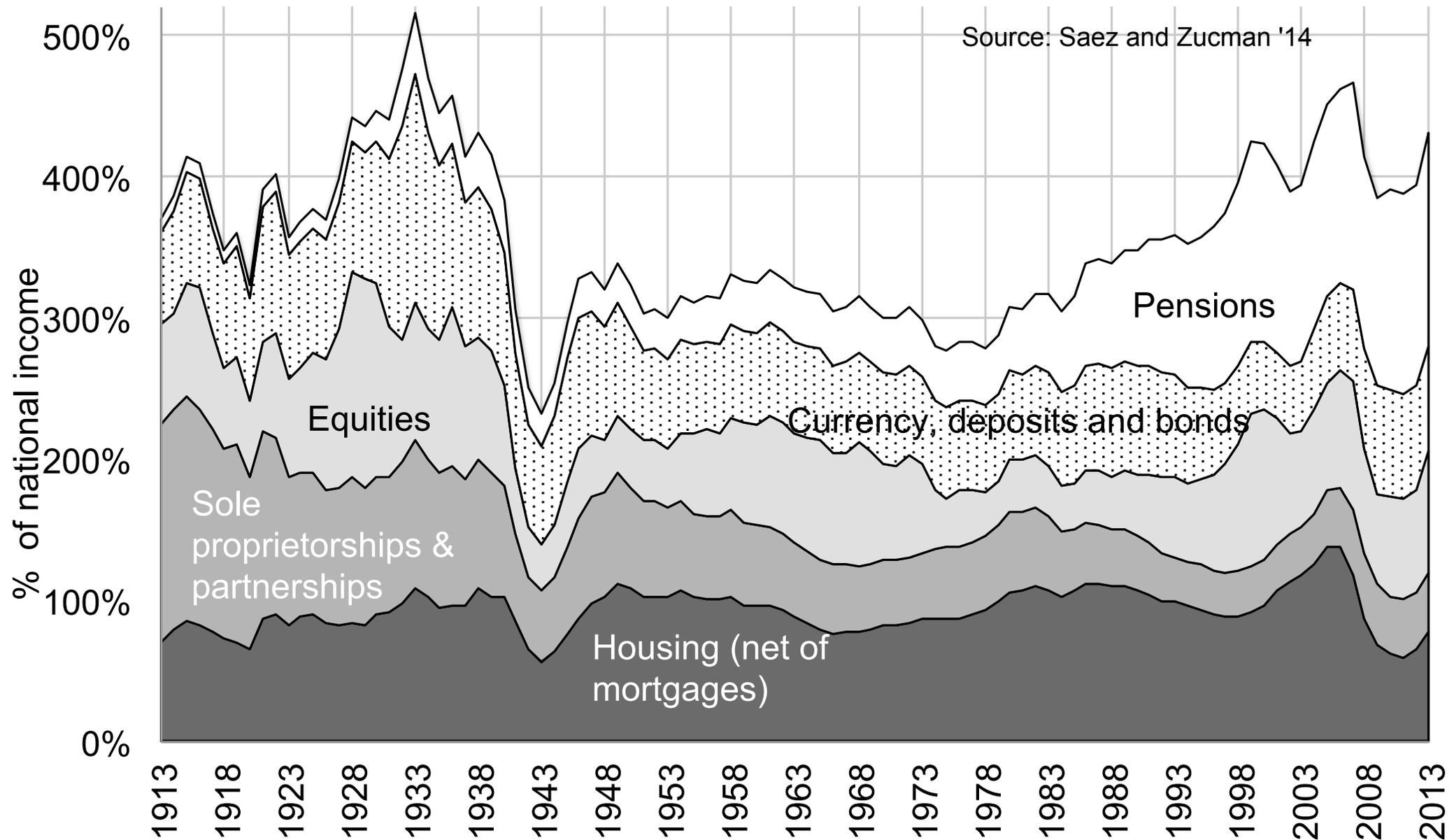
Corporate equities: Value of corporate stock [income = dividends + retained earnings]

Fixed claim assets: Currency, deposits, bonds [income = interest income] minus debts [credit card, student loans]

Pension funds: Substantial amount of equities and fixed claim assets held indirectly through pension funds

The composition of household wealth in the U.S., 1913-2013

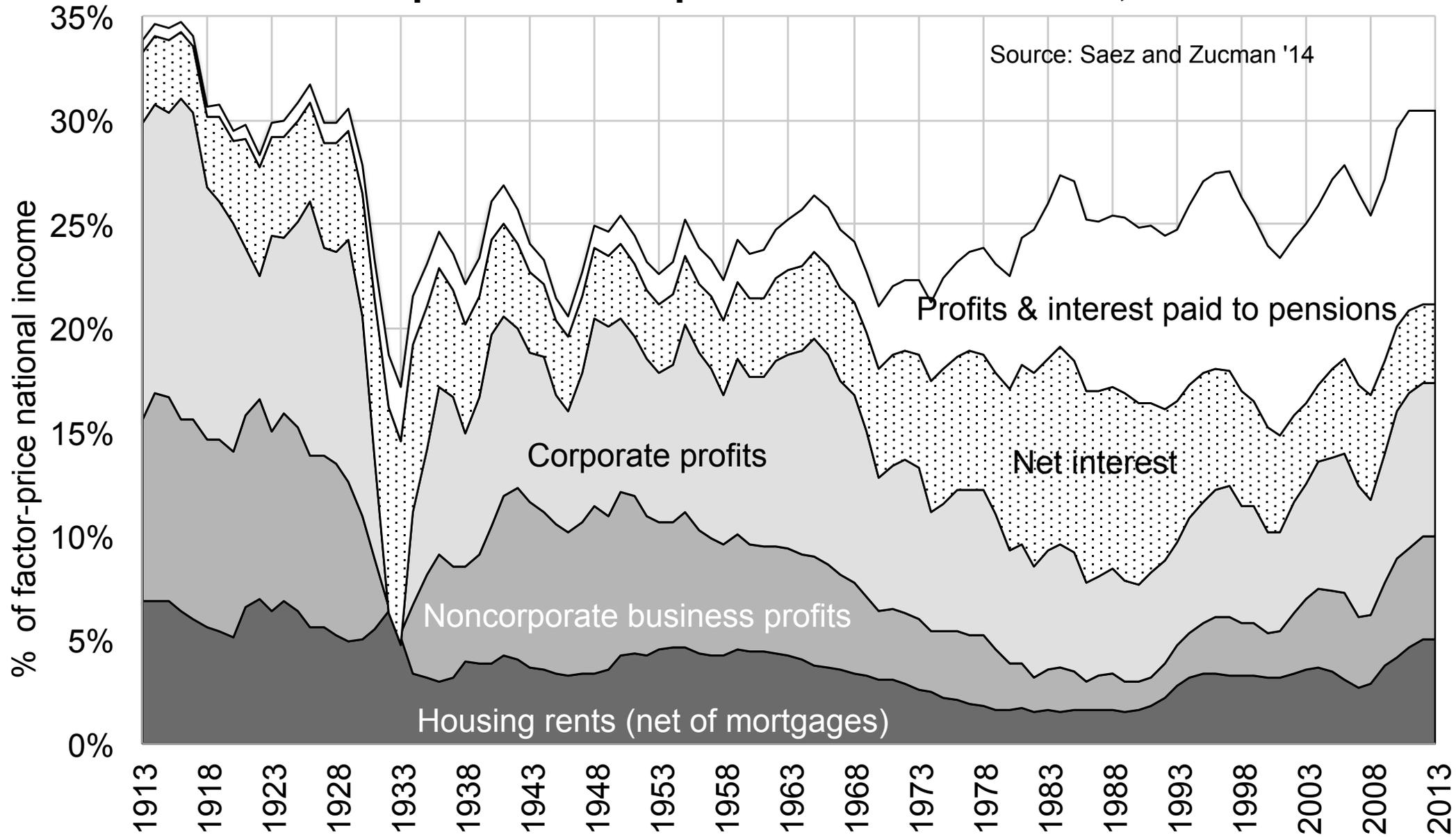
Source: Saez and Zucman '14



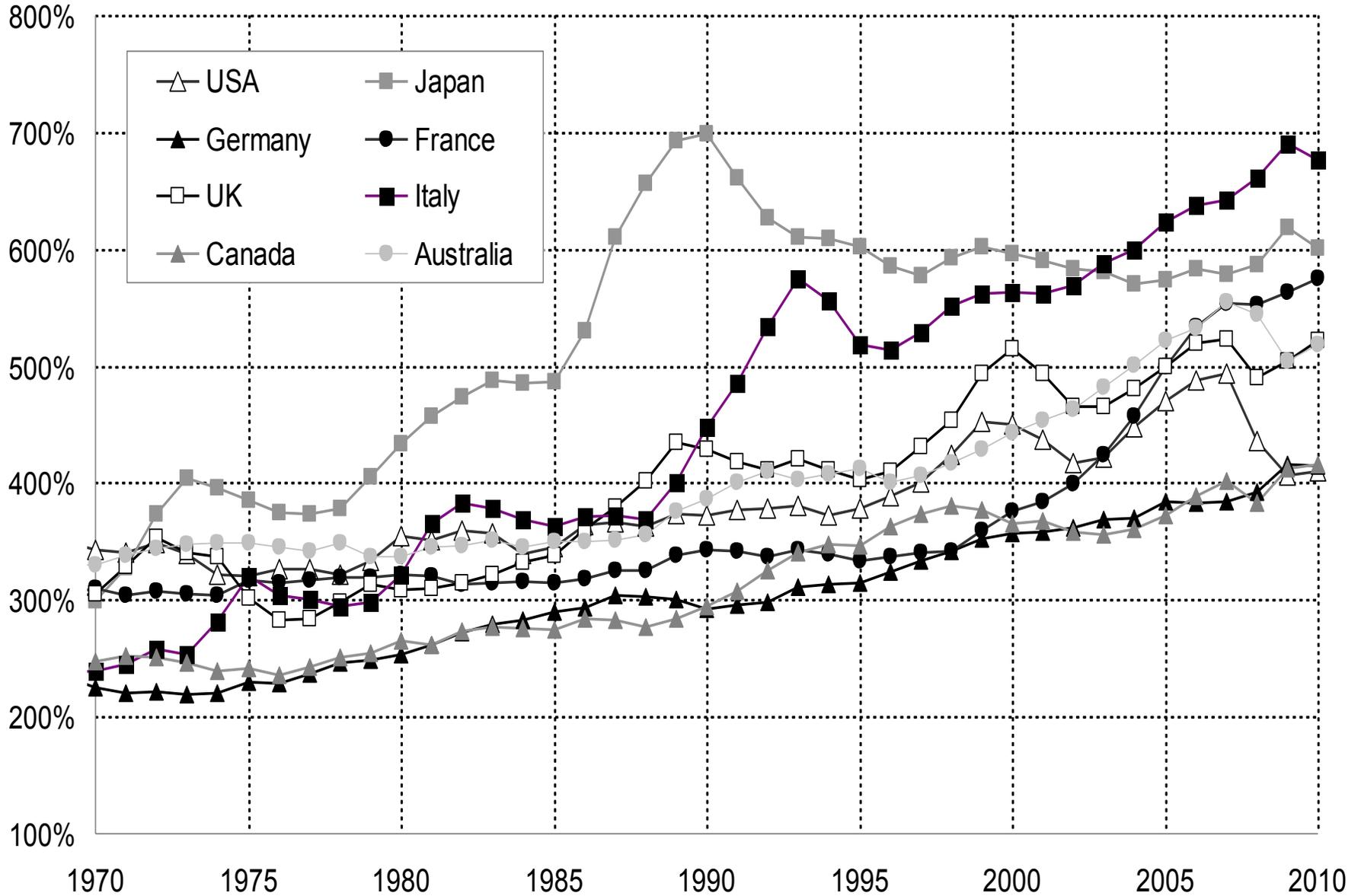
This figure depicts the evolution of the ratio of total household wealth to national income. This ratio has followed a U-shaped evolution and the composition of wealth has changed markedly since 1913. Source: Appendix Table A1.

The composition of capital income in the U.S., 1913-2013

Source: Saez and Zucman '14

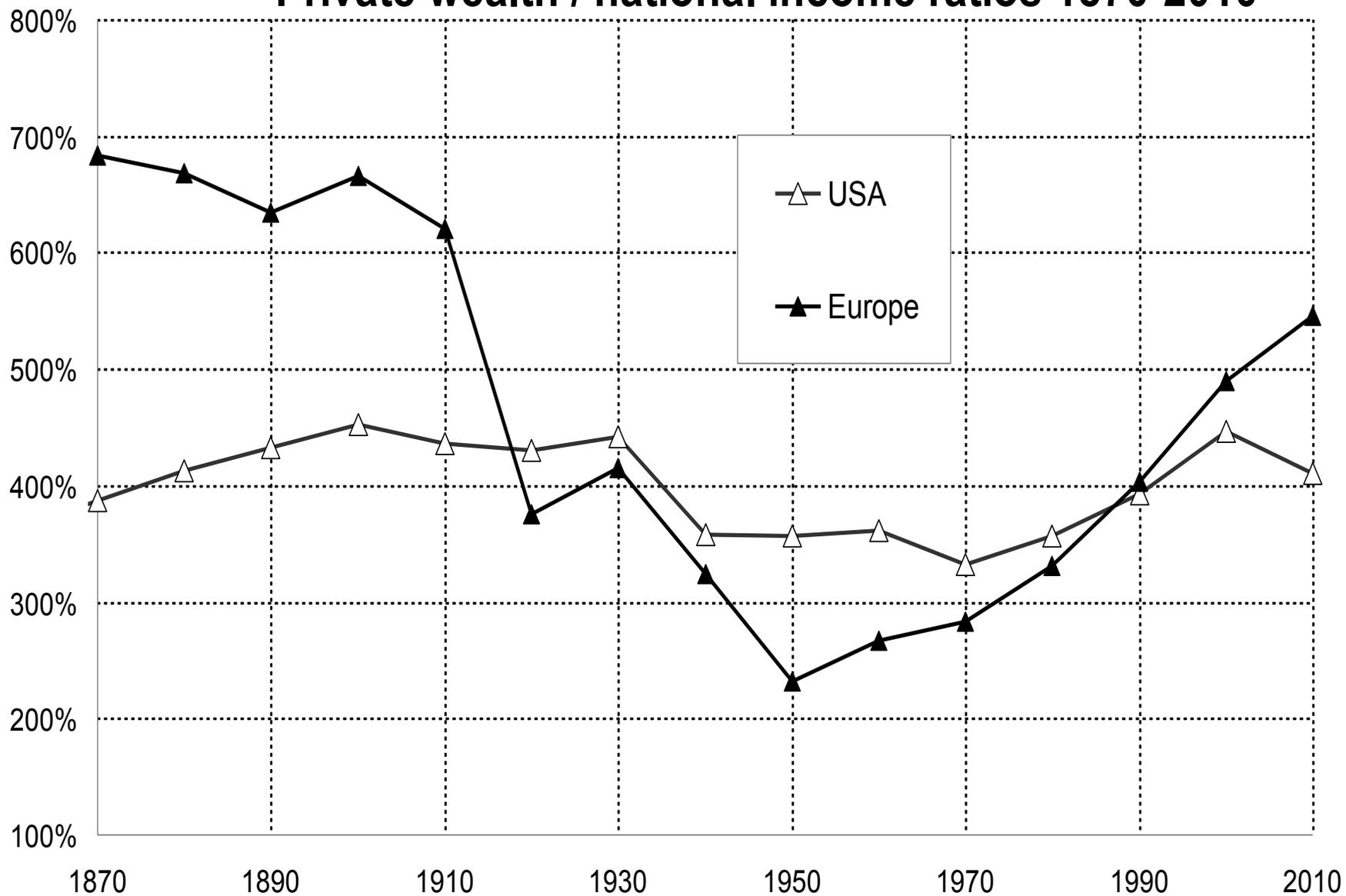


Private wealth / national income ratios, 1970-2010



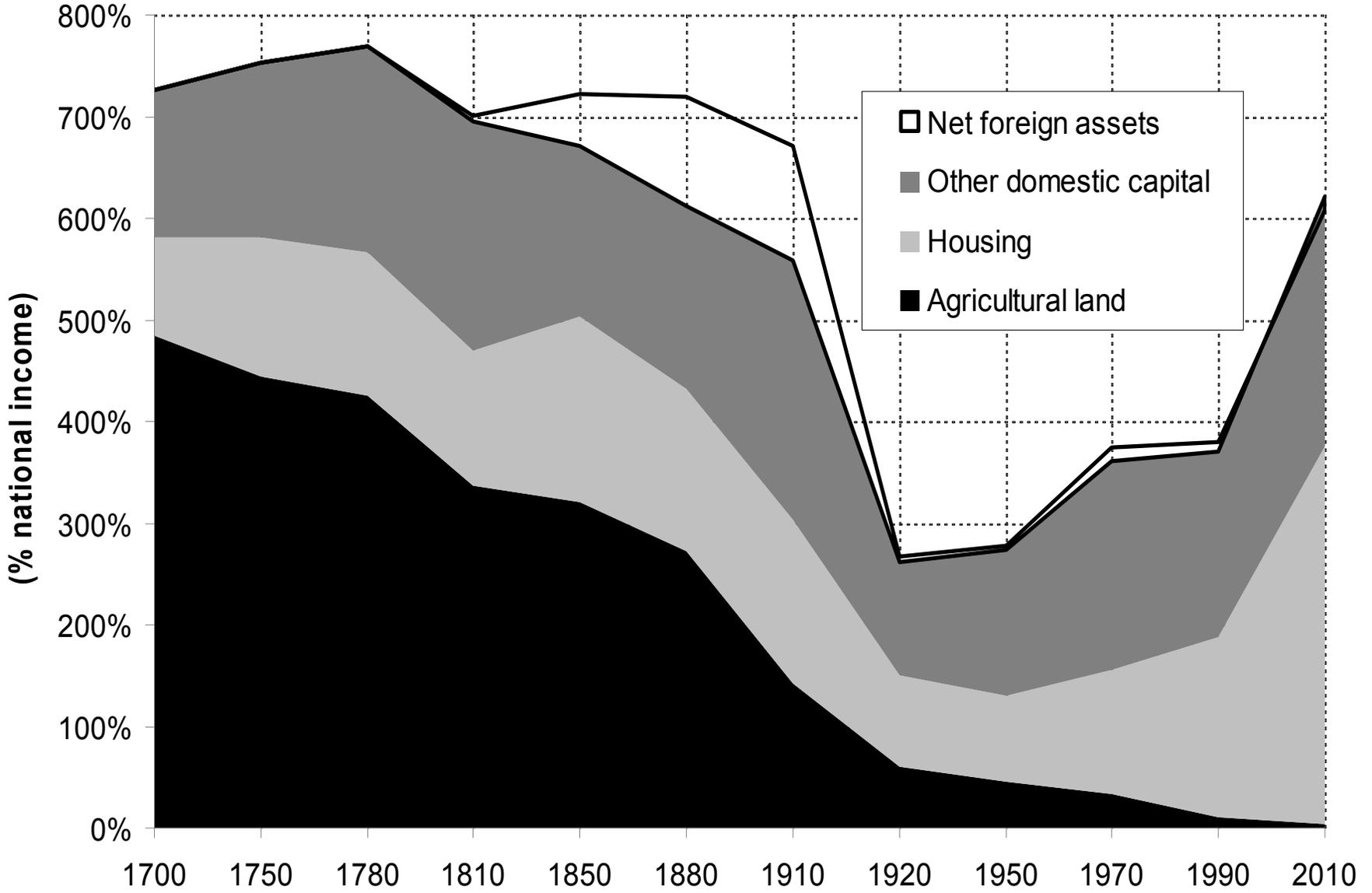
Authors' computations using country national accounts. Private wealth = non-financial assets + financial assets - financial liabilities (household & non-profit sectors)

Private wealth / national income ratios 1870-2010



Authors' computations using country national accounts. Private wealth = non-financial assets + financial assets - financial liabilities (household & non-profit sectors)

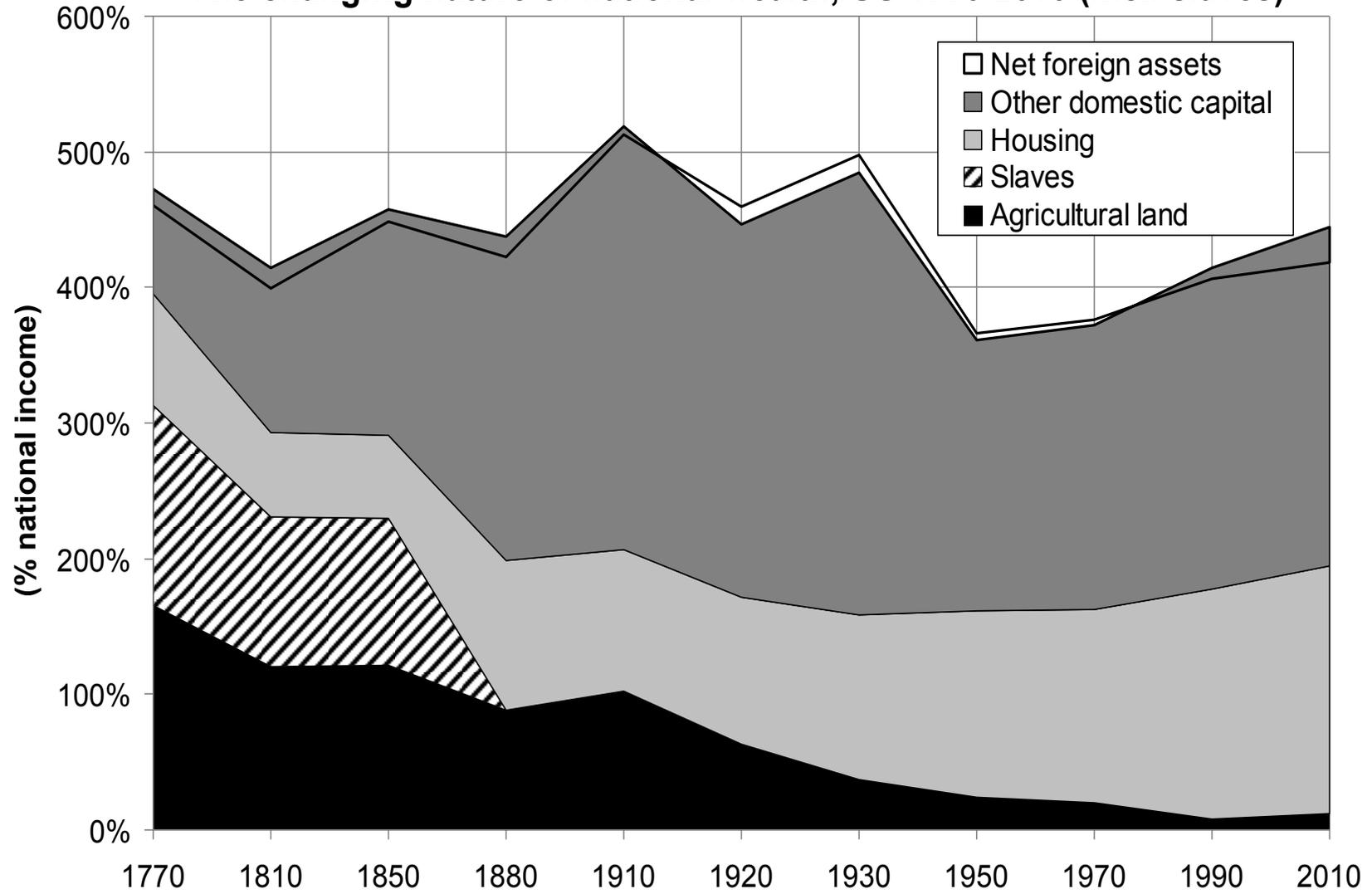
The changing nature of national wealth, France 1700-2010



National wealth = agricultural land + housing + other domestic capital goods + net foreign assets

Source: Piketty, Thomas [book]Capital in the 21st Century

The changing nature of national wealth, US 1770-2010 (incl. slaves)



National wealth = agricultural land + housing + other domestic capital goods + net foreign assets

FACTS ABOUT WEALTH AND CAPITAL INCOME

Wealth = W , Return = r , Capital Income = rW

$$W_t = W_{t-1} + r_t W_{t-1} + E_t + I_t - C_t$$

where W_t is wealth at age t , C_t is consumption, E_t labor income earnings (net of taxes), r_t is the average (net) rate of return on investments and I_t net inheritances (gifts received and bequests - gifts given).

Differences in Wealth and Capital income due to:

- 1) Age
- 2) past earnings, and past saving behavior $E_t - C_t$ [life cycle wealth]
- 3) Net Inheritances received I_t [transfer wealth]
- 4) Rates of return r_t

Wealth Inequality (Saez and Zucman '14)

Wealth inequality is very large (always much higher than income inequality)

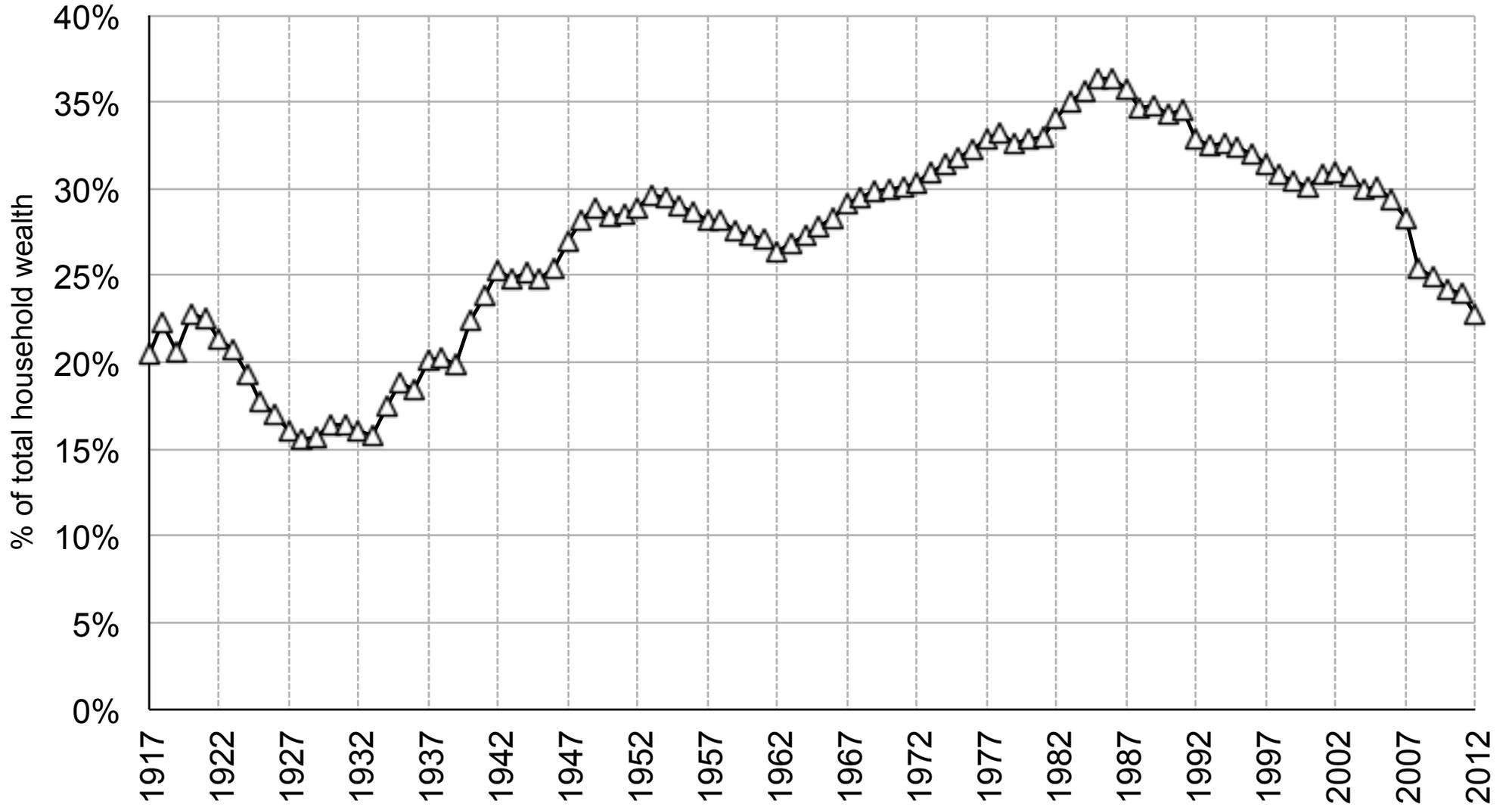
In the US 2012: Top 1% wealthiest families get 40% of total wealth, Next 9% get about 35%, next 40% get 20%, bottom 50% get about 0%

Wealth inequality decreases from 1929 to 1980: wealth democratization due to rise in homeownership and pensions

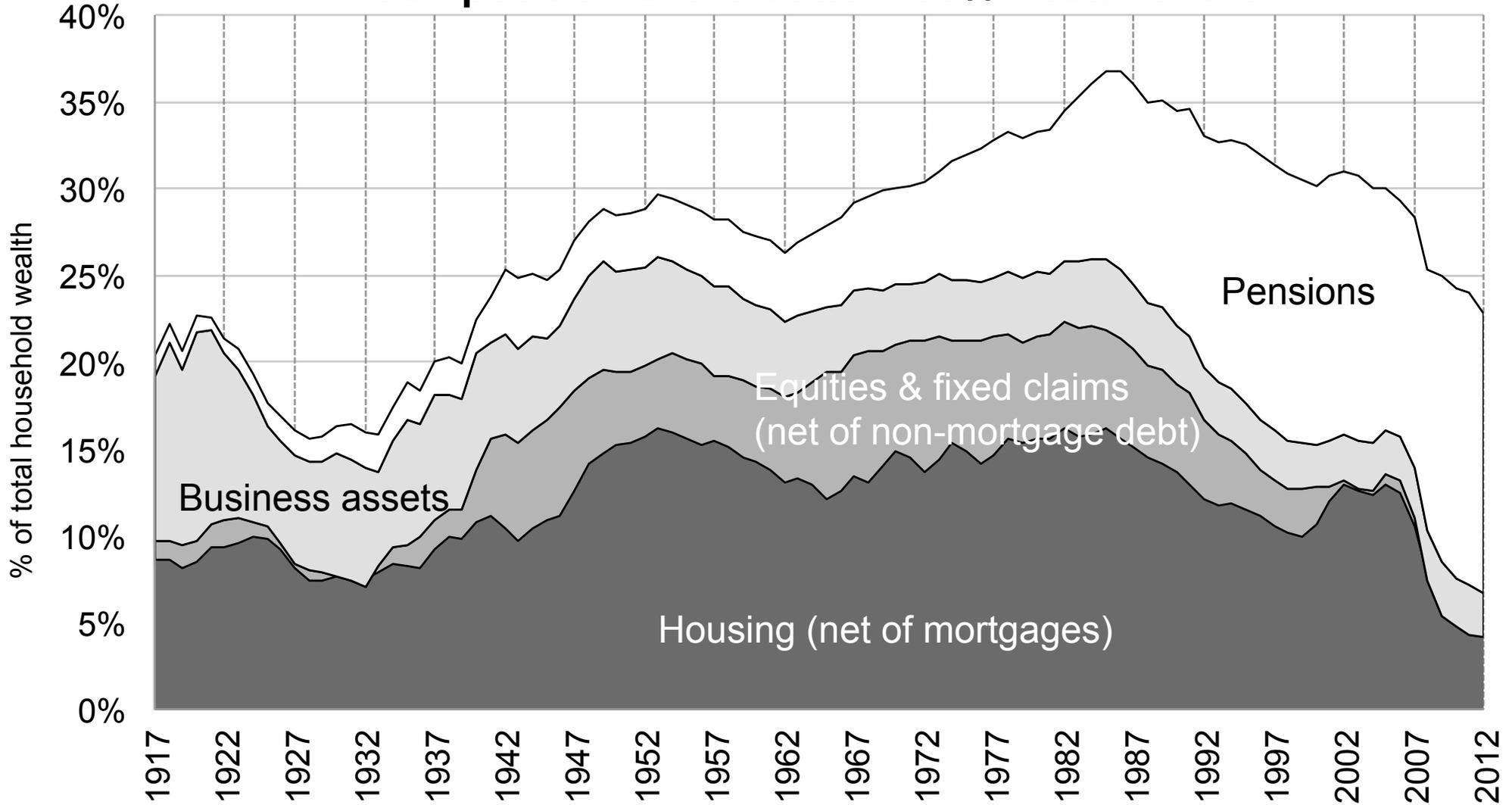
Wealth inequality increases sharply since 1980 fueled by increases in top incomes and a sharp drop in savings rate for bottom 90%

US public underestimates extent of wealth inequality and thinks the ideal wealth distribution should be a lot less unequal [Norton-Ariely '11]

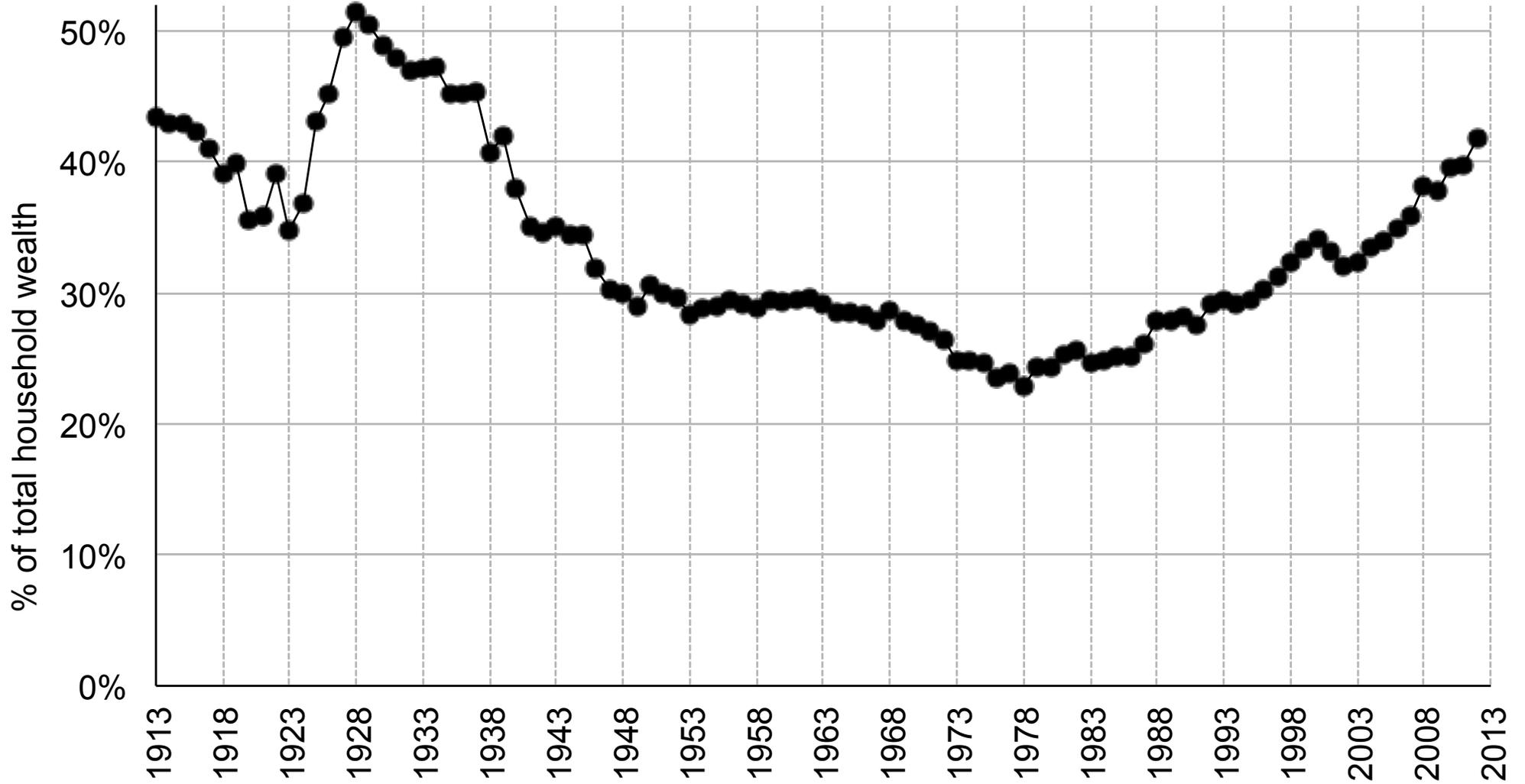
Bottom 90% wealth share in the United States, 1917-2012



Composition of the bottom 90% wealth share

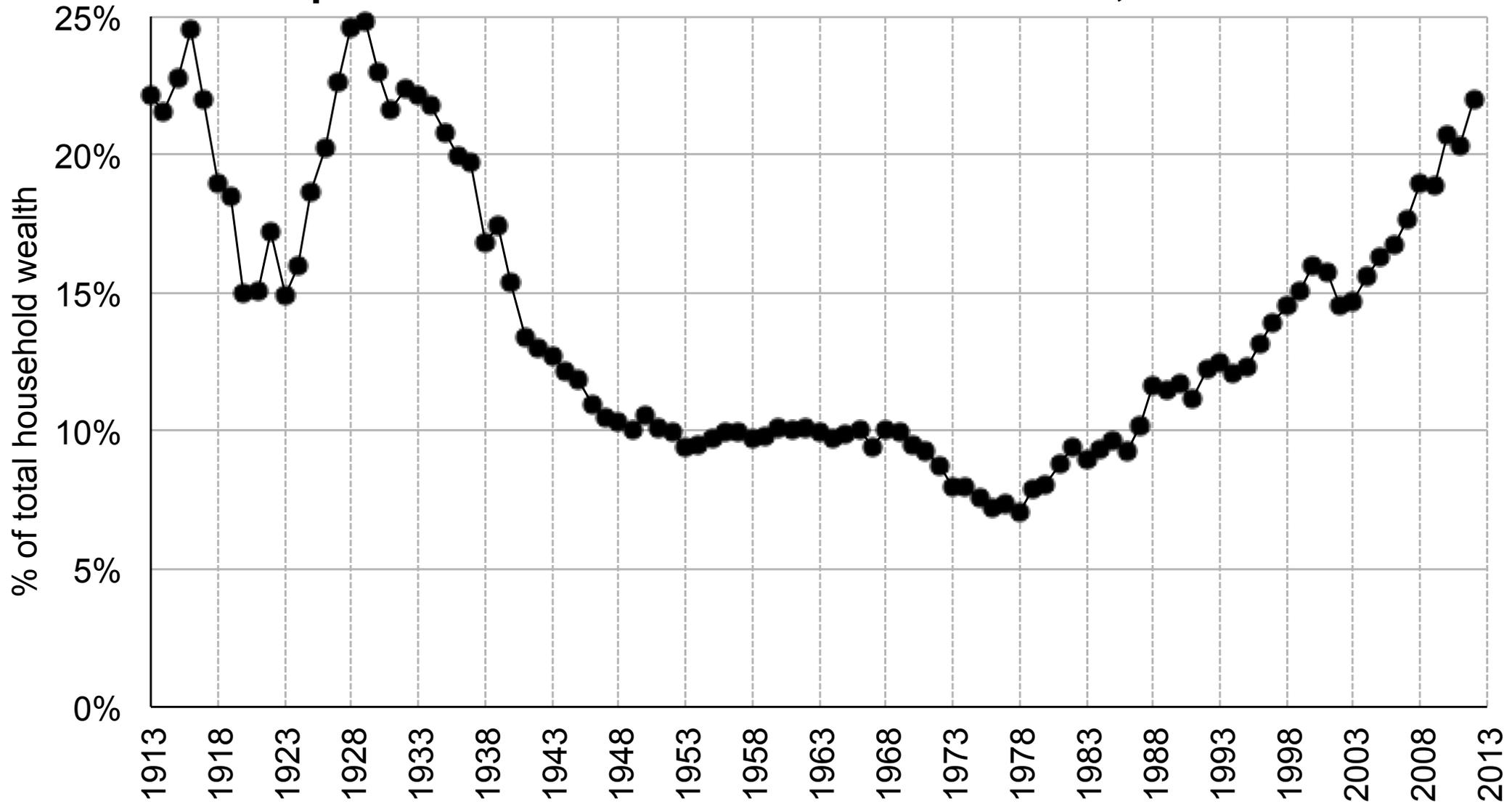


Top 1% wealth share in the United States, 1913-2012



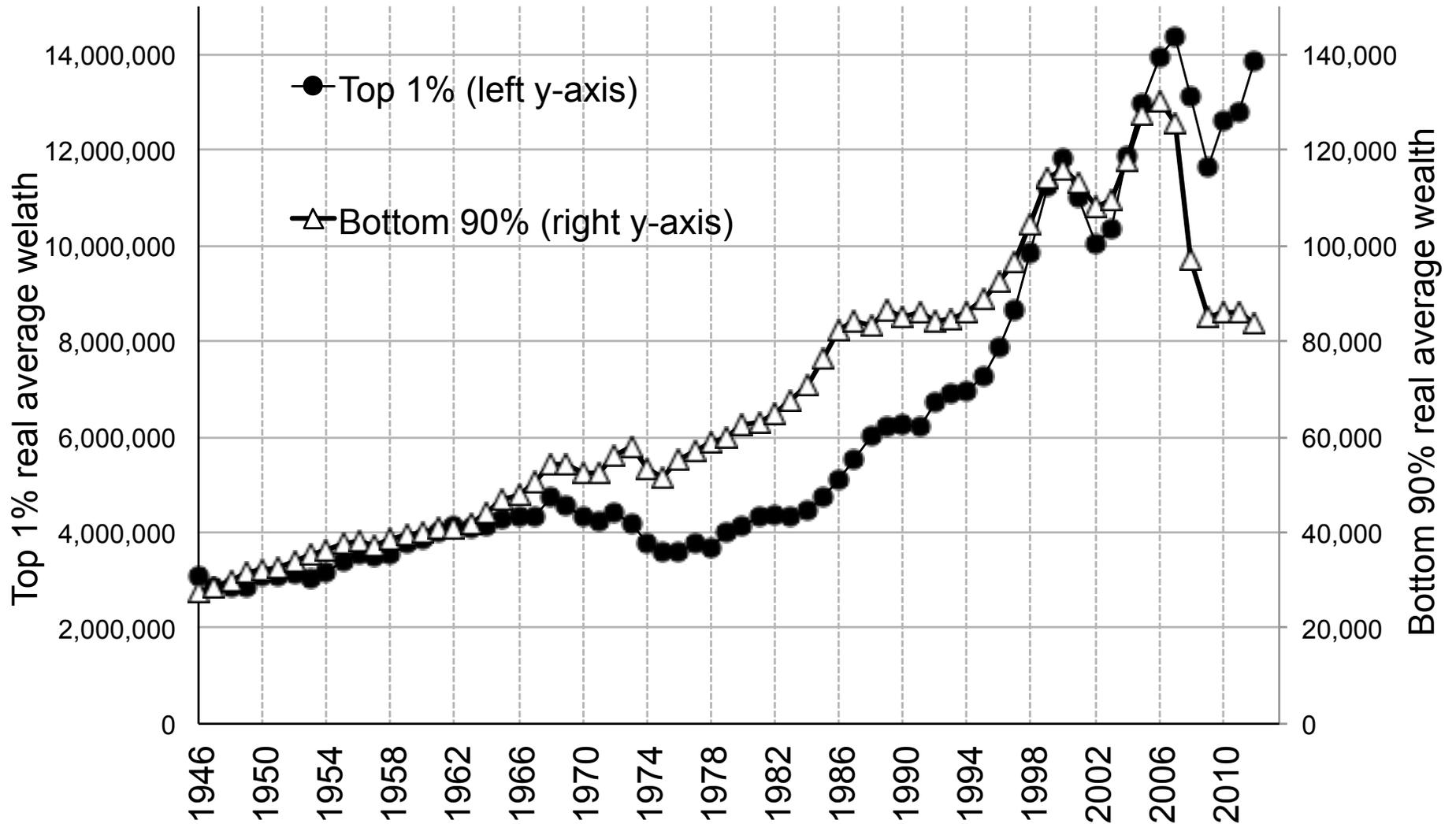
This figure depicts the share of total household wealth held by the 1% richest families, as estimated by capitalizing income tax returns. Source: Saez and Zucman (2014).

Top 0.1% wealth share in the United States, 1913-2012



This figure depicts the share of total household wealth held by the 0.1% richest families, as estimated by capitalizing income tax returns. In 2012, the top 0.1% includes about 160,000 families with net wealth above \$20.6 million. Source: Appendix Table B1.

Real average wealth of bottom 90% and top 1% families



Real values are obtained by using the GDP deflator, 2010 dollars. Source: Appendix Tables B3.

Source: Norton and Ariely 2011

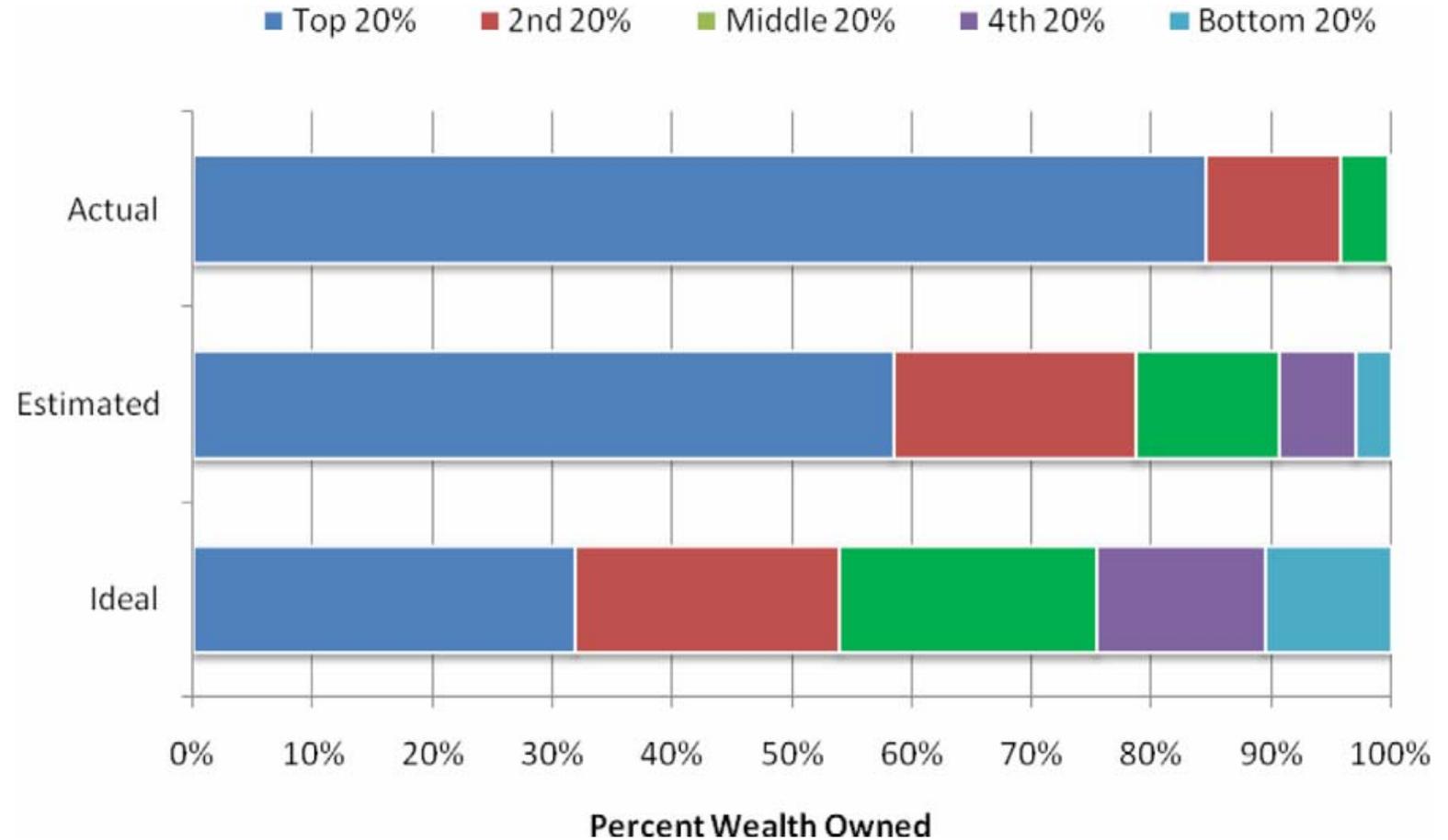


Fig. 2. The actual United States wealth distribution plotted against the estimated and ideal distributions across all respondents. Because of their small percentage share of total wealth, both the “4th 20%” value (0.2%) and the “Bottom 20%” value (0.1%) are not visible in the “Actual” distribution.

FACTS OF US CAPITAL INCOME TAXATION

1) **Corporate Income Tax** (fed+state): 35% Federal tax rate on profits of corporations [complex rules with many industry specific provisions]: effective tax rate much lower

2) **Individual Income Tax** (fed+state): taxes many forms of capital income

Realized capital gains and dividends receive preferential treatment (to lower double taxation of corporate profits)

Imputed rent of home owners, returns on pension funds, state+local bonds interest are exempt

3) **Estate tax**: tax on very large estates (above \$5m) bequeathed to heirs

4) **Property taxes** (local) on real estate (old tax):

Tax varies across jurisdictions. About 0.5% of market value on average

LIFE CYCLE VS. INHERITED WEALTH

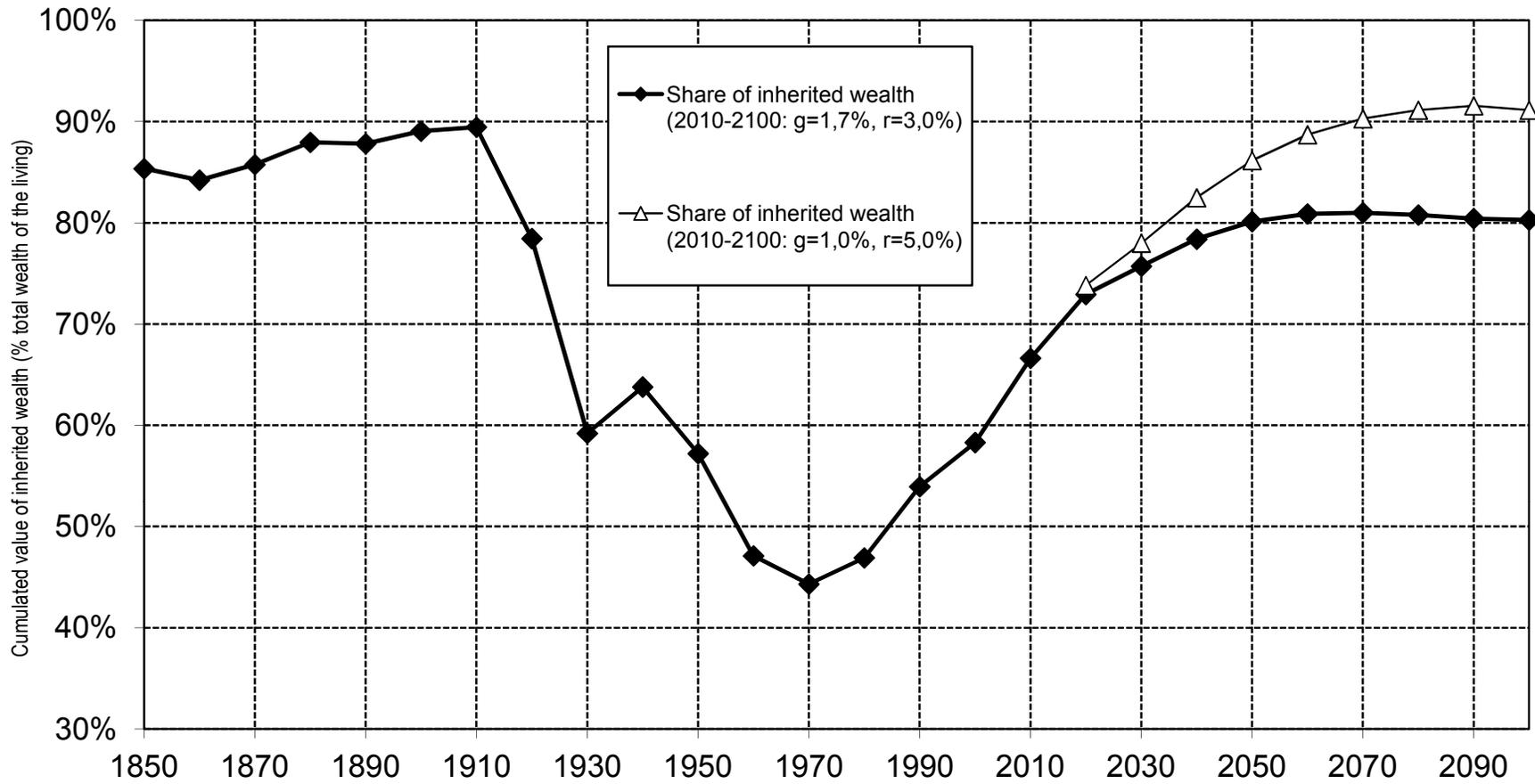
Economists divide existing wealth into 2 categories:

- 1) **Life-cycle wealth** is wealth from savings earlier in your life
- 2) **Inherited wealth** is wealth from inheritances received

Distinction matters for taxation because individuals are responsible for life-cycle wealth but not inherited wealth

Inherited wealth used to be very large in Europe (before World-War I), became small in post-World War II period, but is growing in recent decades (especially in Europe)

Figure 11.7. The share of inherited wealth in total wealth, France 1850-2100



Inherited wealth represents 80-90% of total wealth in France in the 19th century; this share fell to 40%-50% during the 20th century, and might return to 80%-90% during the 21st century. Sources and series: see piketty.pse.ens.fr/capital21c

Source: Piketty (2014)

Piketty (2014) book: Capital in the 21st Century

Analyzes income, wealth, inheritance data over the long-run:

1) Growth rate $g = \text{population growth} + \text{growth per capita}$. Population growth will converge to zero, growth per capita for frontier economies is modest (1-1.5%) \Rightarrow long-run $g \simeq 1-1.5\%$

2) Long-run Wealth to income ratio (β) = savings rate (s) / annual growth (g):

$$\text{Proof: } W_{t+1} = (1 + g) \cdot W_t = W_t + s \cdot Y_t \Rightarrow W_t/Y_t = s/g$$

With $s = 8\%$ and $g = 2\%$, $\beta = 400\%$ but with $s = 8\%$ and $g = 1\%$, $\beta = 800\% \Rightarrow$ Wealth will become important

Piketty (2014) book: Capital in the 21st Century

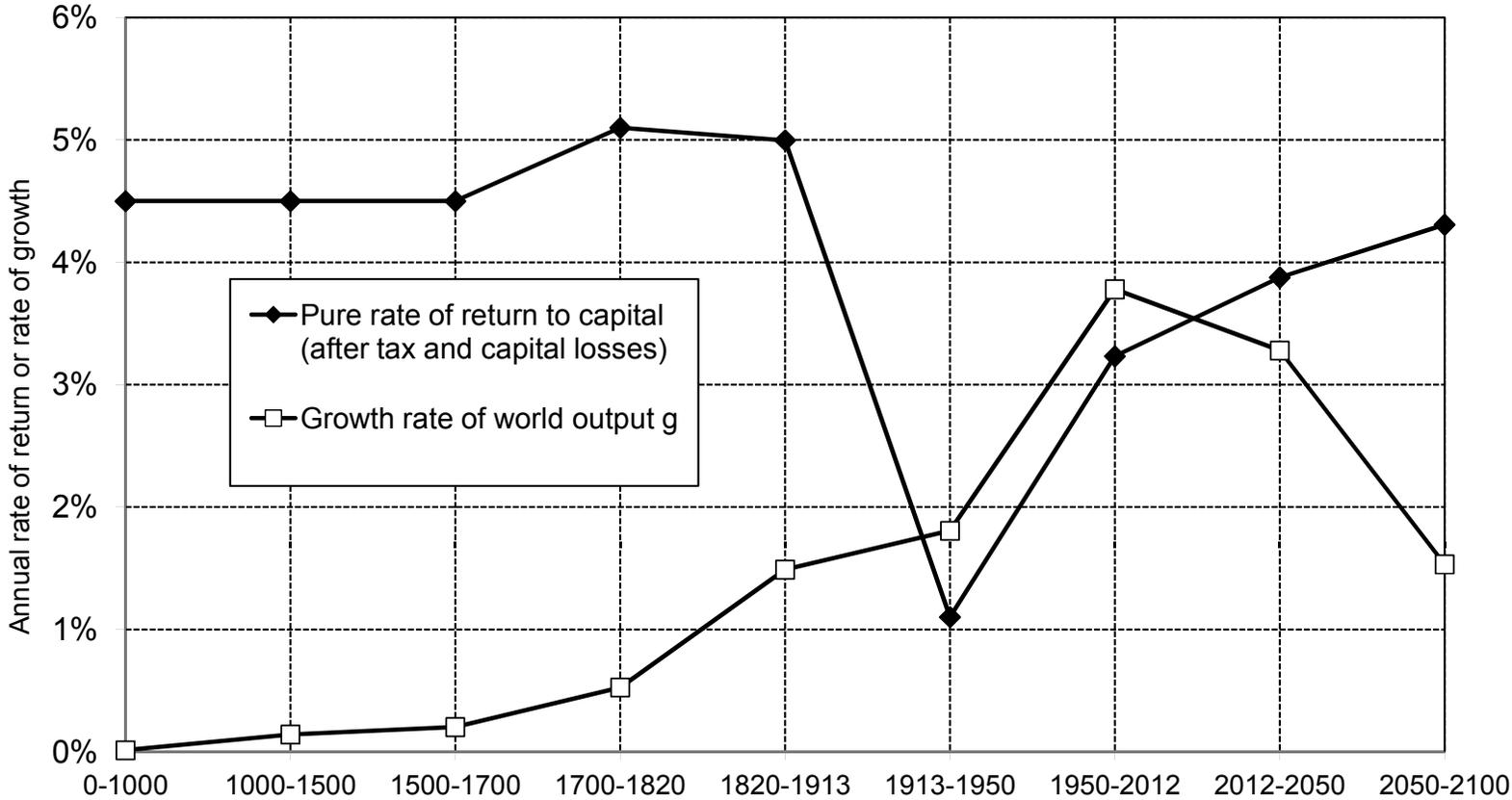
3) Rate of return on wealth $r \simeq 5\%$ significantly larger than g [except exceptional period of 1940s-1960s]

With $r \gg g$, role of inheritance in wealth grows and wealth inequality increases [past swallows the future]

Explanation: Rentier who saves all his return on wealth accumulates wealth at rate r bigger than g and hence his wealth grows relative to the size of the economy. The bigger $r - g$, the easier it is for wealth to “snowball”: fortunes are created faster and last longer

\Rightarrow Capital income taxation reduces r to $r \cdot (1 - \tau_K)$ \Rightarrow This reduces wealth concentration and the relative weight of inherited wealth

Figure 10.10. After tax rate of return vs. growth rate at the world level, from Antiquity until 2100



The rate of return to capital (after tax and capital losses) fell below the growth rate during the 20th century, and may again surpass it in the 21st century. Sources and series : see piketty.pse.ens.fr/capital21c

Source: Piketty (2014)

LIFE-CYCLE MODEL

Individual lives for 2 periods, works l , earns wl , consumes c_1 in period 1, consumes c_2 in period 2:

$$U = u(c_1, l) + \delta v(c_2)$$

Start with case with no taxes

Savings $s = wl - c_1$, $c_2 = (1 + r)s$. Capital income rs

$$\text{Intertemporal budget : } c_1 + \frac{c_2}{1 + r} \leq wl$$

$$\max_{l, c_2} u \left(wl - \frac{c_2}{1 + r}, l \right) + \delta v(c_2)$$

$$\text{First order condition labor Supply : } w \frac{\partial u}{\partial c_1} + \frac{\partial u}{\partial l} = 0$$

$$\text{First order conditions savings : } \frac{\partial u}{\partial c_1} = \delta \cdot (1 + r) \frac{\partial v}{\partial c_2}$$

TAXES IN LIFE-CYCLE MODEL

1) Budget with consumption tax t_c :

$$(1 + t_c)[c_1 + c_2/(1 + r)] \leq wl$$

Budget with labor income tax τ_L :

$$c_1 + c_2/(1 + r) \leq (1 - \tau_L)wl$$

2) Consumption and labor income tax are equivalent if

$$1 + t_c = 1/(1 - \tau_L)$$

Both taxes distort only labor supply and not savings

TAXES IN LIFE-CYCLE MODEL

3) Budget with capital income tax τ_K :

$$c_1 + c_2/(1 + r(1 - \tau_K)) \leq wl$$

τ_K distorts only savings choice (and not labor supply)

4) Budget with comprehensive income tax τ on both labor and capital income:

$$c_1 + c_2/(1 + r(1 - \tau)) \leq (1 - \tau)wl$$

τ distorts both labor supply and savings

τ imposes “double” tax: (1) tax on earnings, (2) tax on savings

EFFECT OF CAPITAL TAX ON SAVINGS

Consider simpler model (fixed earnings w in period 1)

$$\max_{c_1, c_2} u(c_1) + \delta u(c_2) \quad \text{subject to} \quad c_1 + \frac{c_2}{1 + r(1 - \tau_K)} \leq w$$

Recall that $c_1 = w - s$ and $c_2 = [1 + r(1 - \tau_K)] \cdot s$ [draw graph]

Suppose τ_K increases and hence $1/[1 + r(1 - \tau_K)] \uparrow$

1) Substitution effect: price of $c_2 \uparrow \Rightarrow c_2 \downarrow$, $c_1 \uparrow \Rightarrow$ savings $s = w - c_1$ decrease

2) Income effect: Price of $c_2 \uparrow \Rightarrow$ consumer is poorer and both c_1 and $c_2 \downarrow \Rightarrow$ savings s increase

Total net effect is theoretically ambiguous $\Rightarrow \tau_K$ has ambiguous effects on s

Fundamental tax reform: Shift to consumption taxation

Current US tax system is an income tax taxing both earnings and capital income

Some conservatives advocate shifting to consumption tax

Consumption tax is equivalent to taxing only earnings

Shift from labor tax to consumption tax generates double taxation of transitional generation (who have paid labor tax when working and need to pay consumption tax when old)

Actual consumption taxes (such as value-added taxes) tend to be flat while actual income taxes are generally progressive

OPTIMAL CAPITAL INCOME TAXATION

Two broad types of models:

- 1) Life-cycle models: wealth is due solely to life-cycle savings
- 2) Models with bequests: wealth is due solely to inheritances

Optimal Tax in Life-Cycle model

Government can use both a progressive labor income tax $T(wl)$ and a linear capital income tax τ_K

Individuals live 2 periods, earn in period 1, retired in period 2

$$\max_{c_1, c_2, l} u(c_1) - h(l) + \delta u(c_2) \quad s.t. \quad c_1 + \frac{c_2}{1 + r(1 - \tau_K)} \leq wl - T(wl)$$

Individuals differ only according to their earning ability w

Government maximizes social welfare function based on individual utilities

Atkinson-Stiglitz JpubE'76 theorem: The optimal tax τ_K on capital income should be zero. Using a labor tax on earnings $T(wl)$ is sufficient.

Optimal Tax in Life-Cycle model

Atkinson-Stiglitz' theorem shows that life-time savings should not be taxed, tax only labor income

Key intuition: in basic life-cycle model, inequality in life-time resources is due solely to differences in earnings ability. This inequality can be addressed with labor income taxation. Capital income taxation just distorts saving behavior.

From justice view: seems fair to not discriminate against savers if labor earnings is the only source of inequality.

LIMITS OF LIFE-CYCLE MODEL

In reality, capital income inequality also due

- (1) difference in rates of returns across individuals
- (2) shifting of labor income into capital income
- (3) inheritances

Difference in Rates of Returns Across Individuals

Rate of return on wealth varies significantly over time and across individuals

In general, richer and higher-skilled individuals are able to invest in higher return assets due to ability to take risks and scale effects in financial advice [e.g., large University endowments get a larger return than smaller ones, Piketty 2014, Chapter 12]

⇒ Capital income is correlated with underlying skill so taxing it improves upon Atkinson-Stiglitz outcome [two dimensions of inequality (labor, savings technology) require two dimensions of tax policy (labor tax, capital tax)]

SHIFTING OF LABOR / CAPITAL INCOME

In practice, difficult to distinguish between capital and labor income [e.g., small business profits, professional traders]

Differential tax treatment can induce shifting

(1) Carried interest in the US: hedge fund and private equity fund managers receive fraction of profits of assets they manage for clients. Those profits are really labor income but are taxed as realized capital gains

(2) Finnish Dual income tax system: taxes separately capital income at preferred rates since 1993: Pirtila and Selin SJE'11 show that it induced shifting from labor to capital income especially among self-employed

With income shifting, taxing capital income becomes desirable to curb this tax avoidance opportunity

Inheritance: Estate Taxation in the United States

Estate federal tax imposes a tax on estates above \$5.5M exemption (only about .1% of deceased liable), tax rate is 40% above exemption (in 2013 and after)

Charitable and spousal giving are fully exempt from the tax

E.g.: if Bill Gates / Warren Buffet give all their wealth to charity, they won't pay estate tax

Popular support for estate tax is pretty weak ("death tax") but public does not know that estate tax affects only richest

Support for estate tax increase shoots up from 17% to 53% when survey respondents are informed that only richest pay it (Kuziemko-Norton-Saez-Stantcheva AER'15 do an online Mturk survey experiment)

Treatment example: Information about the Estate Tax

Besides the income tax, the government can also level the playing field with **the federal estate tax**.

The **Federal Estate Tax** (also known as the **Death Tax**) applies when a deceased person leaves **more than \$5 million** in wealth to his or her heirs. Wealth left to a spouse or charitable organizations is exempt from estate tax.



Only 1 person out of 1000 is wealthy enough to face the estate tax.

Average Americans do not have anything close to \$5 million in wealth, so the estate tax does not affect them and they can pass on their property to their children tax-free.

Eliminating the estate tax would allow the very richest families to pass down all of their wealth to their children tax-free. Hence, children of rich people would also start off very rich themselves.

Increasing the estate tax is a way to level the playing field between the children of wealthy parents and children of middle-class parents.

Taxation of Inheritances: Welfare Effects

Inheritances (or gifts from living parents) raise difficult issues of social justice [see Kaplow 2001]:

(1) Inequality in inheritances contributes to economic inequality and individuals not responsible for inheritances they receive:

⇒ seems fair to redistribute from those who received inheritances to those who did not

(2) However, it seems unfair to tax the parents who worked hard (and already paid tax on income) to pass on wealth to children

Taxation of Inheritances: Behavioral Responses

Potential behavioral response effects of inheritance tax:

(1) reduces wealth accumulation of altruistic parents (and hence tax base) [not very good estimates, Kopczuk-Slemrod 2001 suggest small effects]

(2) reduces labor supply of altruistic parents (less motivated to work if cannot pass wealth to kids) [no good estimates]

(3) induces inheritors to work more through income effects because they receive smaller inheritances (Carnegie effect, decent evidence from Holtz-Eakin, Joulfaian, Rosen QJE'93)

Critical to understand why there are inheritances for optimal inheritance tax policy. 3 models of bequests: (a) accidental, (b) altruistic bequests, (c) manipulative bequest motive

ACCIDENTAL BEQUESTS

People die with a stock of wealth they intended to spend on themselves (or that they accumulated out of love for wealth, Carroll '98):

Bequest taxation has no distortionary effect on behavior of parent and can only increase labor supply of inheritors (through income effects) \Rightarrow strong case for taxing bequests heavily

Surveys show that bequest motives are not the main driver of wealth accumulation (Kopczuk-Lupton '07)

Only 1/3 of people surveyed say that the main reason they accumulate wealth is for bequests to their children

Altruistic Bequests (Piketty and Saez ECMA'13)

Utility $u(c) - h(l) + \delta v(b^{left})$ where c is own consumption, l is labor supply, and b^{left} is net-of-tax bequests left to next generation and $v(b^{left})$ is utility of leaving bequests for donor

Individual receives $b^{received}$, works and earns $wl - T(wl)$, consumes c , saves $s = wl - T(wl) + b^{received} - c$, which translates into $b^{left} = s(1 + r)(1 - \tau_B)$ for heir (τ_B is bequest tax rate)

Bequests provide an additional source of life-income:

$$c + \frac{b^{left}}{(1 - \tau_B)(1 + r)} = wl - T(wl) + b^{received}$$

In this model, Atkinson-Stiglitz breaks down and using bequest taxation is desirable to supplement labor income taxation

⇒ Two-dimensional inequality (labor, bequests) requires two-dimensional tax policy tool (labor tax, bequest tax)

MANIPULATIVE BEQUESTS

Parents use potential bequest to extract favors from children

Empirical Evidence: Bernheim-Shleifer-Summers JPE '85 show that number of visits of children to parents is correlated with bequeathable wealth but not annuitized wealth of parents

[Annuitized wealth is wealth that disappears at death such as a pension or an annuity]

$$Visits_i = \alpha + \beta BequeathableWealth_i + \gamma Annuitizedwealth_i + \varepsilon_i$$

In regression, they find $\beta > 0$ and $\gamma = 0$ (but causality not clear)

⇒ Bequest becomes one additional form of labor income for inheritor and one consumption good for parent

⇒ Inheritances should be taxed as labor income for donees

SOCIAL-FAMILY PRESSURE BEQUESTS

Parents may not want to leave bequests but feel compelled to by pressure of heirs or society: bargaining between parents and children

With estate tax, parents do not feel like they need to give as much \Rightarrow parents are made better-off by the estate tax \Rightarrow Case for estate taxation stronger

Empirical evidence:

Aura JpubE'05: reform of private pension annuities in the US in 1984 requiring both spouses signatures when worker decides to get a single annuity or couple annuity: reform increases sharply couple annuities choice

Equal division of estates [Wilhelm AER'96, Light-McGarry '04]: estates are very often divided equally probably to avoid conflicts [gifts before death are not as equally split]

WEALTH IN TAX HAVENS

Official statistics substantially underestimate the net foreign asset positions of rich countries bc they do not capture most of the assets held by households in off-shore tax havens

Example: Wealthy US individual opens a Cayman Islands account and buys mutual fund shares (composed of US corporate stock): Cayman Islands record a liability but US do not record an asset (because this is not reported in the US)

⇒ Total world liabilities are larger than world total assets

Zucman QJE'13 compiles international financial stats and estimates that around 8% of the global financial wealth of households is held in tax havens (3/4 of which is unrecorded = 6%)

If top 1% hold about 50% of total financial wealth, then about 12% of financial wealth of the rich is hidden in tax heavens

CURBING OFF-SHORE TAX EVASION

Off-shore tax evasion possible because of bank secrecy: US cannot get a list of US individuals owning Swiss bank accounts from Switzerland

⇒ No 3rd party reporting makes tax enforcement very difficult

In principle, problem could be solved with exchange of information across countries BUT need all countries to cooperate

Johannesen-Zucman AEJ-EP'14 analyze tax haven crackdown: G20 countries forced number of tax havens to sign bilateral treaties on bank information sharing

Key result: Instead of repatriating funds, tax evaders shifted deposits to havens not covered by treaty with home country.

CURBING OFF-SHORE TAX EVASION

FATCA'13 US regulations try to impose info exchange for all entities dealing with US:

If foreign bank B does not provide list of all its US account holders, any financial transaction between B and US will carry 30% tax withholding ⇒ Interesting to see what it will do

Long-term solution will require:

a) Systematic registration of assets to ultimate owners [already exists within countries for domestic tax enforcement]

b) Systematic information exchange between tax countries with no exceptions for tax heavens

⇒ Could be enforced with tariffs threats on tax heavens [Zucman JEP'14 and book '15]

REFERENCES

Atkinson, A.B. and J. Stiglitz “The Design of Tax Structure: Direct Versus Indirect Taxation”, *Journal of Public Economics*, Vol. 6, 1976, 55-75. (web)

Aura, S. “Does the Balance of Power Within a Family Matter? The Case of the Retirement Equity Act”, *Journal of Public Economics*, Vol. 89, 2005, 1699-1717. (web)

Bernheim, B. D., A. Shleifer, and L. Summers “The Strategic Bequest Motive”, *Journal of Political Economy*, Vol. 93, 1985, 1045-76. (web)

Carroll, C. “Why Do the Rich Save So Much?”, NBER Working Paper No. 6549, 1998. (web)

Gravelle, Jane. “The Economic Effects of Taxing Capital Income.” [Book] MIT Press, 1994.

Holtz-Eakin, D., D. Joulfaian and H.S. Rosen “The Carnegie Conjecture: Some Empirical Evidence”, *Quarterly Journal of Economics* Vol. 108, May 1993, pp.288-307. (web)

Johannesen, Niels and Gabriel Zucman “The End of Bank Secrecy? An Evaluation of the G20 Tax Haven Crackdown,” forthcoming in *American Economic Journal: Economic Policy*, 2013. (web)

Kaplow, L. "A Framework for Assessing Estate and Gift Taxation", in Gale, William G., James R. Hines Jr., and Joel Slemrod (eds.) *Rethinking estate and gift taxation* Washington, D.C.: Brookings Institution Press, 2001. (web)

Kopczuk, Wojciech and Joseph Lupton 2007. "To Leave or Not to Leave: The Distribution of Bequest Motives," *Review of Economic Studies*, 74(1), 207-235. (web)

Kopczuk, Wojciech and Joel Slemrod, "The Impact of the Estate Tax on the Wealth Accumulation and Avoidance Behavior of Donors", in William G. Gale, James R. Hines Jr., and Joel B. Slemrod (eds.), *Rethinking Estate and Gift Taxation*, Washington, DC: Brookings Institution Press, 2001, 299-343. (web)

Kuziemko, Ilyana, Michael I. Norton, Emmanuel Saez, and Stefanie Stantcheva "How Elastic are Preferences for Redistribution? Evidence from Randomized Survey Experiments," NBER Working Paper No. 18865, 2013, forthcoming *American Economic Review*, 2015 (web)

Light, Audrey and Kathleen McGarry. "Why Parents Play Favorites: Explanations For Unequal Bequests," *American Economic Review*, 2004, v94(5,Dec), 1669-1681. (web)

Norton, M. and D. Ariely "Building a Better America—One Wealth Quintile at a Time", *Perspectives on Psychological Science* 2011 6(9). (web)

Piketty, Thomas, *Capital in the 21st Century*, Cambridge: Harvard University Press, 2014, (web)

Piketty, T. and G. Zucman “Capital is Back: Wealth-Income Ratios in Rich Countries, 1700-2010”, forthcoming *Quarterly Journal of Economics*, 2014 (web)

Piketty, T. and E. Saez “A Theory of Optimal Inheritance Taxation”, *Econometrica*, 81(5), 2013, 1851-1886. (web)

Pirttila, Jukka and Hakan Selin, “Income Shifting Within a Dual Income Tax System: Evidence from the Finnish Tax Reform,” *Scandinavian Journal of Economics*, 113(1), 120-144, 2011. (web)

Saez, Emmanuel and Gabriel Zucman, “Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data”, NBER Working Paper, No. 20625, October 2014 (web)

Wilhelm, Mark O. “Bequest Behavior and the Effect of Heirs’ Earnings: Testing the Altruistic Model of Bequests,” *American Economic Review*, 86(4), 1996, 874-892. (web)

Zucman, G. “The Missing Wealth of Nations: Are Europe and the US Net Debtors or Net Creditors”, *Quarterly Journal of Economics*, 2013, 1321-1364. (web)

Zucman, G. *La Richesse Cachée des Nations*, Paris: Seuil, 2012 [translation in english in 2015], summary in Journal of Economic Perspectives, 2014 (web)

Zucman, G. "Taxing across Borders: Tracking Personal Wealth and Corporate Profits" Journal of Economic Perspectives, 28(4), 2014, 121-148. (web)

Undergraduate Public Economics

Danny Yagan
UC Berkeley

Lecture 10 Business Taxation

LAST LECTURE VS. THIS LECTURE

Last lecture was from the perspective of the **saver**: How do taxes distort savings decisions and should we tax savings?

This lecture is from the perspective of the **investor**: How do taxes affect investment / should we tax investment?

Motivations:

- 1) Might care only about total investment/capital, not savers (e.g. if savers are all rich and government doesn't value their consumption, then Atkinson-Stiglitz result isn't main focus)
- 2) Savings doesn't have to equal investment in open economy

THIS LECTURE'S BIG QUESTIONS

- 1) If the gov. doesn't care about savers, should we tax capital?
- 2) What are the theoretical investment effects of a business income tax?
- 3) What are the theoretical and empirical investment effects of a dividend income tax?
- 4) Is capital accumulation actually good for workers?
- 5) Are governments fighting a losing business tax battle?

Basic Definitions

Business (firm) is a for-profit legal entity owned by shareholders typically with limited liability (if business goes bankrupt, share price drops to zero but shareholders not liable for unpaid bills/debt)

Shareholders: Individuals who have purchased ownership stakes in a business.

Ownership vs. control: owners are shareholders. Managers (CEO and top executives) in general do not own the company but run the corporation on behalf of shareholders

Firm Financing

Firms can finance themselves through debt or through equity

Equity finance: The raising of funds by sale of ownership shares in a firm. Shareholders receive dividends from corporation and capital gain if the share price increases

Debt finance: The raising of funds by borrowing from lenders such as banks, or by selling corporate bonds. Corporate bonds are promises by a business to make periodic interest payments, as well as ultimate repayment of principal, to the bondholders (the lenders)

Bond holders have priority on shareholders for repayment in case of bankruptcy

Firm Taxation

Businesses use capital (land, buildings, machines, equipment) and labor (workers) to transform inputs (raw materials) into outputs (goods produced and sold).

Profits = revenues from goods sold - expenses (labor costs, inputs, capital depreciation, interest payments on debt)

Profits are **double taxed**:

Business income tax (τ^{INC}): Tax on annual profits (e.g. the corporate income tax of 35%)

Payout tax: Tax on after-business-income-tax profits once they are distributed (paid out) to shareholders

Dividend: Standard payout form. A specific amount is paid per share owned, often in regular amounts. Owners pay **dividend tax** (τ^{DIV})

Capital gain: If business keeps after-business-income-tax profits, share value rises. Owners realize cap. gain when selling those shares, pay cap. gains tax

Why Not Tax Payouts Only?

Businesses are not people but are ultimately owned by people. In principle, we want to tax people based on their economic resources but:

1) Taxing Pure Profits: Some firms have market power (e.g., Microsoft) and hence earn pure profits. Taxing pure profits does not distort behavior because firms maximize profits anyway

2) Back-up for individual taxes: If corporations were not taxed on their earnings, then individuals who owned shares in corporations could postpone taxes indefinitely by having the corporations never pay out their earnings

3) Taxing foreign owners: Corporations often have foreign owners. Countries want to tax economic activity on their territory. E.g., consider developing country with foreign owned mineral/oil extraction companies

4) Tax collection convenience: Historically, corporations are more convenient to tax than individuals because they are large, visible, and have detailed accounts (for transparency for their shareholders). So taxing corporate income (profits) was attractive

(1) Should We Tax Capital if Don't Care about Savers?

Mankiw (2000) and Judd (1985): Suppose you only want to maximize the welfare of “spenders” (workers who consume all their income every year and save nothing) and don't care about “savers”.

In standard model where government has to fund some base spending level (e.g. an army):

(a) Substitution effect of capital tax on savings dominates income effect

(b) Savings = Investment

(c) Firms do not earn pure profits

(d) Capital augments labor

(1) Should We Tax Capital if Don't Care about Savers?

⇒ Spenders (in the long run) would want a zero capital tax and a positive labor income tax!

Mechanism: Capital augments labor (makes labor so much more productive and thus increases workers' wages) that any reduction in the capital stock via a capital tax is bad for workers

Key Questions

Do capital taxes really reduce investment?

Does capital really agument labor?

Business Income Tax Theory

Classic cost-of-capital model (Hall-Jorgenson 1967):

Firm has concave gross-profit function $F(K)$ (equal to revenue minus input and labor costs), has to pay capital cost $r \cdot K$ (cost of buying or leasing machines/buildings).

Firm maximizes:

$$\max_K F(K) - r \cdot K$$

$$FOC : F'(K) = r$$

If $F(K^*) - rK^* > 0$, we say firm earns pure profit.

If $F(K^*) - rK^* = 0$, firm does not earn pure profit (and thus barely breaks even).

Business Income Tax Theory

Now firm has to pay a business income tax on its gross profit $F(K)$.

Firm maximizes:

$$\max_K (1 - \tau^{INC}) \cdot F(K) - r \cdot K$$

$$FOC : F'(K) = \frac{r}{(1 - \tau^{INC})}$$

⇒ Business income tax reduces investment

Business Income Tax Theory

But what if the business income tax is assessed on profits after capital costs?

Firm maximizes:

$$\max_K (1 - \tau^{INC}) \cdot [F(K) - r \cdot K]$$

$$FOC : F'(K) = r$$

⇒ Business income tax has no effect on investment! (Robertson 1927)

What Just Happened Here?

When all costs are deducted from taxable income, the business income tax is a tax on pure profit

The K that maximizes pure profit $F(K) - r \cdot K$ also maximizes 65% of pure profit

Firms in perfectly competitive markets do not earn pure profit. Best they can do is break even (after capital costs).

How competitive are markets?

Introspection test: Would you care if you were banned from buying from your favorite grocery store, airline, car company, restaurant, etc.?

Empirical test: Many markets are quite competitive, many are not

Are Business Costs Fully Deductible?

Most costs are fully and immediately deductible.

“Deductible” means “you can deduct the amount from your taxable income (in this case, from your gross profits)”

Capital costs (purchases of machines/buildings) cannot be deducted immediately and instead must be made in pre-defined sequence of annual “depreciation deductions”

Deductions add up to original nominal cost, but total effective deductibility is only partial because value of future tax deductions must be discounted

Longer depreciation schedules \Rightarrow less effective deductibility

Ex. Cars depreciated over 5 years, railroads over 20 years

Evidence on Business Income Taxes and Investment

Large literature trying to estimate elasticity of business investment with respect to $(1 - \tau^{INC})$ and other changes in **cost of capital**: the pre-tax rate of return required on marginal investments

Hassett-Hubbard (2002) consensus range of cost-of-capital elasticity of investment: $[-0.5, -1.0]$ (very elastic)

Some of that evidence is based on (questionable) analysis of changes in τ^{INC}

Would imply that business income tax is very far from pure profits tax

Business Income Tax Theory vs. Practice

Theory recommends: *narrow* the business income tax base (allow for more capital cost deductions) and *increase* the business income tax rate

Emerging policy consensus: broaden the base and lower the rate

“President’s Framework for Business Tax Reform” (Obama 2012): Lengthen depreciation schedules and reduce corporate income tax rate from 35% to 28%

Why the Disconnect?

Possibility 1: Interest deductibility means we actually subsidize investment, so narrowing the base makes the tax code neutral (but then why lower the rate?)

Possibility 2: International competition (discussed later)

Possibility 3: Ideal solution cannot be enforced.

Ex.: Acme Inc. has \$10m in pure profit, suddenly buys a \$10m airplane in order to reduce taxable income to zero and save \$3.5m in taxes. Acme shuts down the next day, plane becomes owners' personal property, and owner sells the plane for \$10m. Owner should have to declare \$10m in ordinary income (and thus pay \$3.96m in taxes if in top bracket), but hard to enforce. \Rightarrow Full deductibility bankrupts government.

Possibility 4: Policymakers are wrong (many think so)

(3) Dividend Tax Theory

Dividend taxes: second layer of double-taxation of business profit. Applies to after-annual-business-income-tax profits, once paid to shareholders (owners).

Two views of dividend taxation: “traditional view” and “new view”

“Traditional view”: Dividend tax is just like the annual business income tax

“New view”: Dividend tax is different

Key: Source of investment funds

Traditional View of Dividend Taxes

Leading treatment: Poterba-Summers (1985). Same as Hall-Jorgenson, but will write in slightly different way.

Think of an entrepreneur (owner of new firm) maximizing her after-tax wealth

Entrepreneur starts with wealth W and decides how much of that wealth to use for her firm's investment: i.e. how much equity to contribute so the firm can buy capital K

Firm shuts down after one year and profits paid as dividends

Whatever she doesn't use for investment is kept in a bank and earns r

Traditional View of Dividend Taxes

Entrepreneur chooses K to maximize her after-tax return on wealth:

$$\max_K [(1 - \tau^{DIV}) \cdot (1 - \tau^{INC}) \cdot F(K) + r \cdot (W - K)]$$

$$FOC : F'(K) = \frac{r}{(1 - \tau^{DIV}) \cdot (1 - \tau^{INC})}$$

⇒ Dividend tax reduces investment, in exactly the same way as the annual business income tax

New View of Dividend Taxes

Developers: King (1977), Auerbach (1979), Bradford (1981)

Think of owner of an already profitable firm (e.g. Microsoft)

Firm starts with profits from preexisting operations *PPO* that are enough to pay for any profitable amount of investment

Owner decides how much of profits from preexisting operations to use for her firm's investment

Firm shuts down after one year and profits paid as dividends

Whatever she doesn't use for investment is paid out as a dividend and earns r in the bank

New View of Dividend Taxes

Owner chooses K to maximize her after-tax return on wealth:

$$\max_K [(1 - \tau^{DIV}) \cdot (1 - \tau^{INC}) \cdot F(K) + r \cdot (1 - \tau^{DIV}) \cdot (PPO - K)]$$

$$FOC : F'(K) = \frac{r}{(1 - \tau^{INC})}$$

⇒ Dividend tax does not affect investment!

To see what's going on:

$$FOC : (1 - \tau^{DIV}) \cdot (1 - \tau^{INC}) \cdot F'(K) = r \cdot (1 - \tau^{DIV})$$

Traditional View vs. New View

Key: What is the **opportunity cost** of investment (i.e. what is the cost of capital)?

Traditional view: Owner forgoes $\$r$ for every \$1 of investment

New view: Owner forgoes only $\$r \cdot (1 - \tau^{DIV})$ for every \$1 of investment

Why? Because there are available funds (profits from preexisting operations) that will be subject to dividend taxes regardless of whether they're used for investment

Traditional View vs. New View

Thus in traditional view: dividend tax does not affect the opportunity cost of investment (r) but reduces the post-tax return on investment by $(1 - \tau^{DIV})$

But in new view: dividend tax reduces the opportunity cost of investment by the exact same factor $(1 - \tau^{DIV})$ that it reduces the post-tax return on investment

In new view, dividend tax is effectively a tax on past investment, not new investment

Testing between the Two Views

Facebook in 2005: Clearly a traditional-view firm (no preexisting operations)

Microsoft now: Clearly a new-view firm (huge profits from preexisting operations)

What about the average firm? Hard to say from just looking at investment data.

Need variation in dividend tax rates

The 2003 Dividend Tax Cut

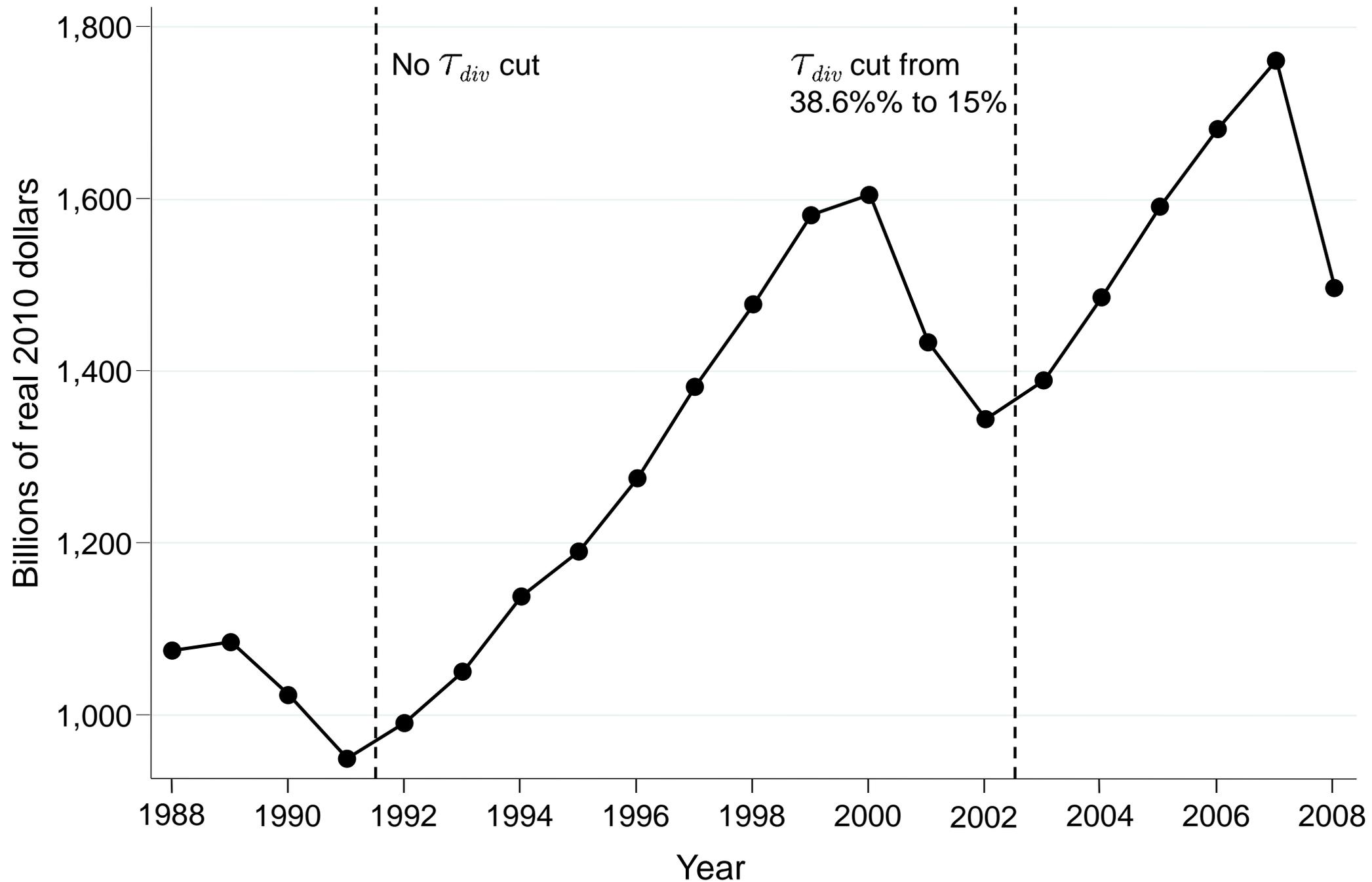
Reduced top federal dividend tax rate from 38.6% to 15%

Traditional-view argument from President: provide “near-term support to investment” and “capital to build factories, to buy equipment, hire more people”

Want to test whether business investment rose after 2003

Problem: Must control for business cycle

U.S. Corporate Investment NIPA Private Fixed Non-residential Investment



DD Design around 2003 Dividend Tax Cut

Yagan (2015): Use unaffected corporations to control for the business cycle

C-corporations: Traditional corporation type, subject to double taxation (annual business income tax and dividend tax)

S-corporations: “Pass-through” business type subject only to annual business income tax (assessed at owner level), never subject to dividend tax

DD between C-corporations (directly affected by 2003 dividend tax cut) and S-corporations (not directly affected)

Strategy: C-corporations vs. S-corporations

- After incorporating, a corporation elects either C or S tax status

	Tax rate on annual income	Tax rate on dividends
C-corporations (treatment)	35%	15%
S-corporations (control)	35%	0%

- S-corporations: < 100 non-institutional investors, one stock class
- Operate in same narrow industries and at the same scale throughout United States → common trends

Example: Retail hardware chains



- Largest hardware chain
- C-corporation



- Third-largest hardware chain
- S-corporation

Note: Only public sources were used to populate this slide

Example: Retail Hardware Chains in Suburban Chicago



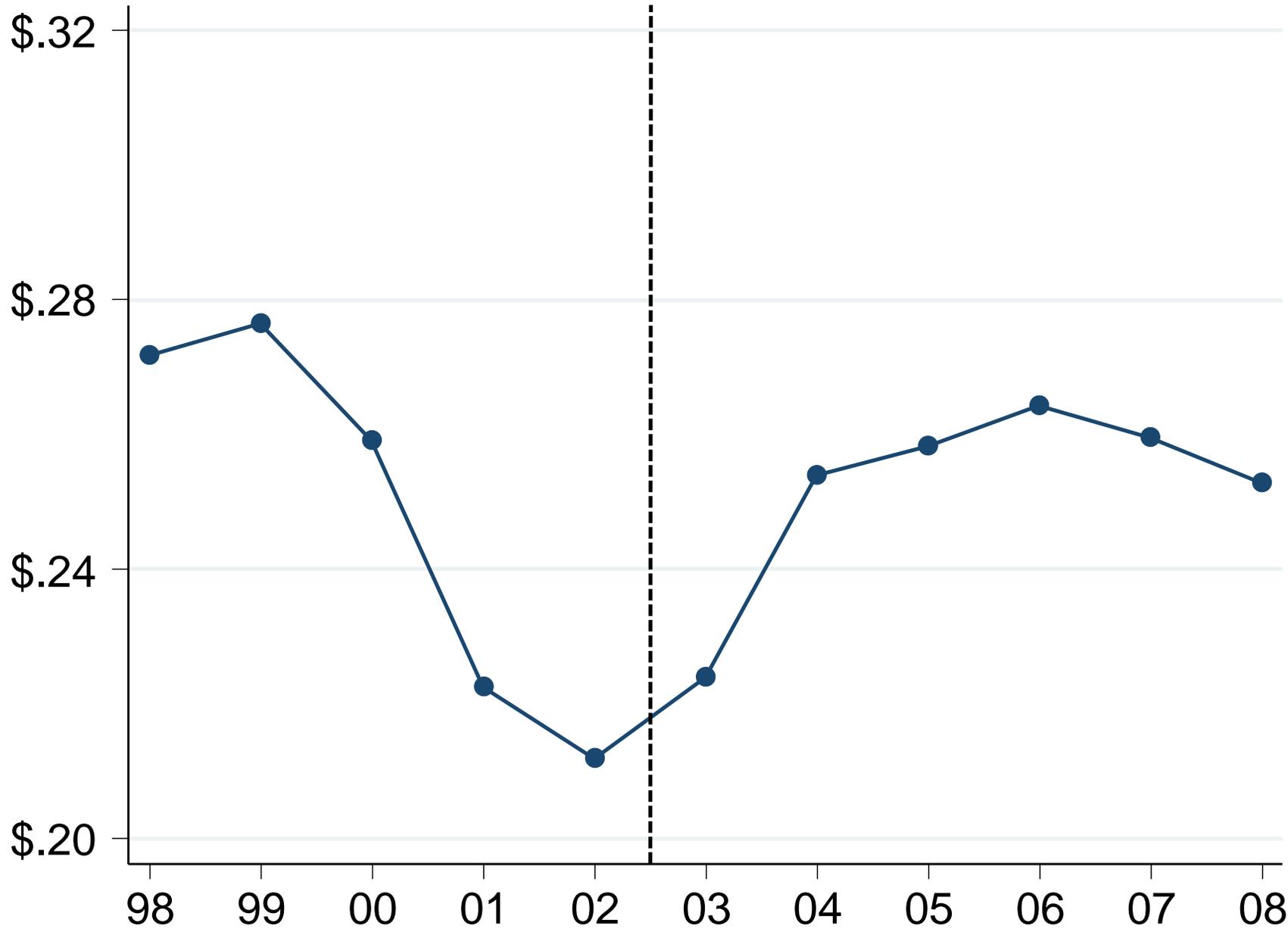
📍 Home Depot (C-corporation)

📍 Menard Inc. (S-corporation)

Note: Only public sources were used to populate this slide.

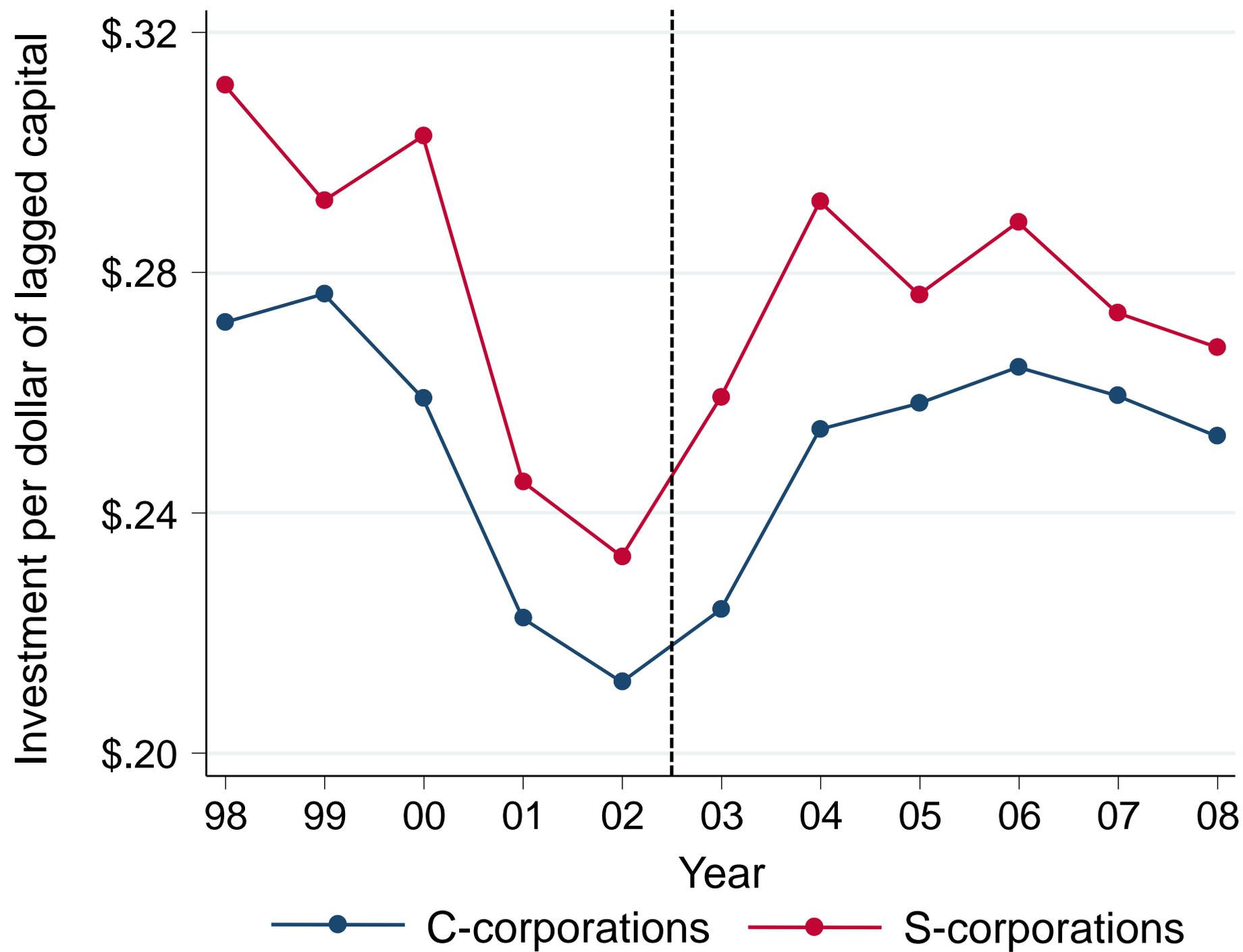
Investment

Investment per dollar of lagged capital

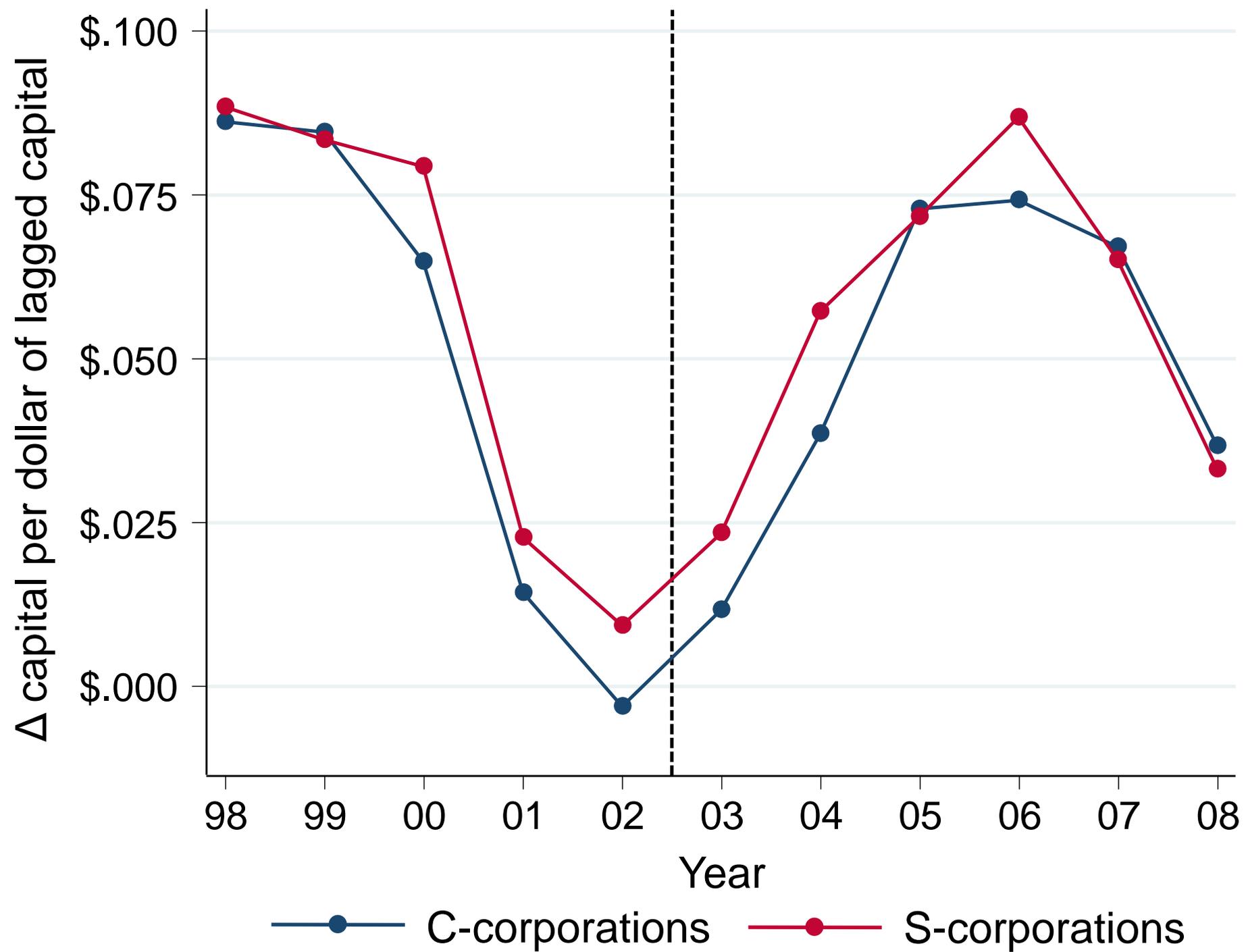


—●— C-corporations —●— S-corporations

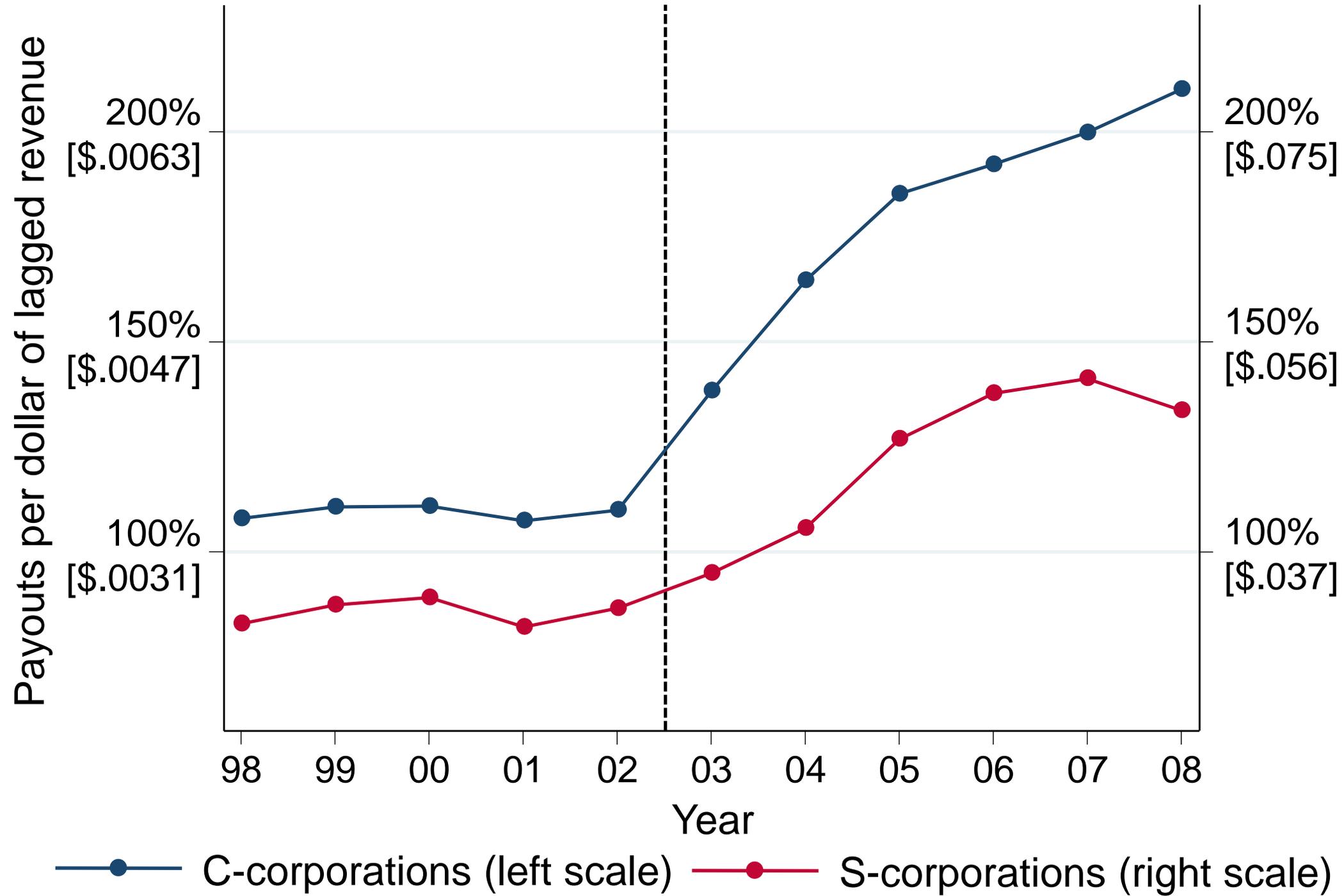
Investment



Net Investment



Payouts to Shareholders



Yagan (2015) Results

Elasticity of investment wrt $(1 - \tau_{DIV}) = 0.00$

95% confidence upper bound: 0.08

Traditional view prediction: $[0.21, 0.41]$, depending on cost-of-capital elasticity of investment (Hassett-Hubbard consensus range $[-0.5, -1.0]$)

Yagan (2015) Results

Natural interpretation: New view is mostly right

Caveat #1: Traditional view could be right, but Hasset-Hubbard are wrong and investment barely responds to cost-of-capital changes (i.e. $F(K)$ very concave)

Caveat #2: Traditional view could be right if tax/political system didn't get in the way (e.g. gov cannot commit to permanently low dividend rate)

Regardless: Future dividend tax changes unlikely to affect investment

(4) Is Capital Accumulation Good for Workers?

Central assumption in all these models: Capital complements labor \Rightarrow investment is good for workers

Strong reason to think this has been true in long arc of history

Now/future: Capital may complement high-skilled (e.g. computers) but not low-skilled

Intuition: You don't need more than one car.

Formally: Products produced by low-skilled workers may have hit inelastic demand \Rightarrow More efficient production no longer increases quantity sold \Rightarrow Need fewer and fewer low-skilled workers making cars

Big Mac Index

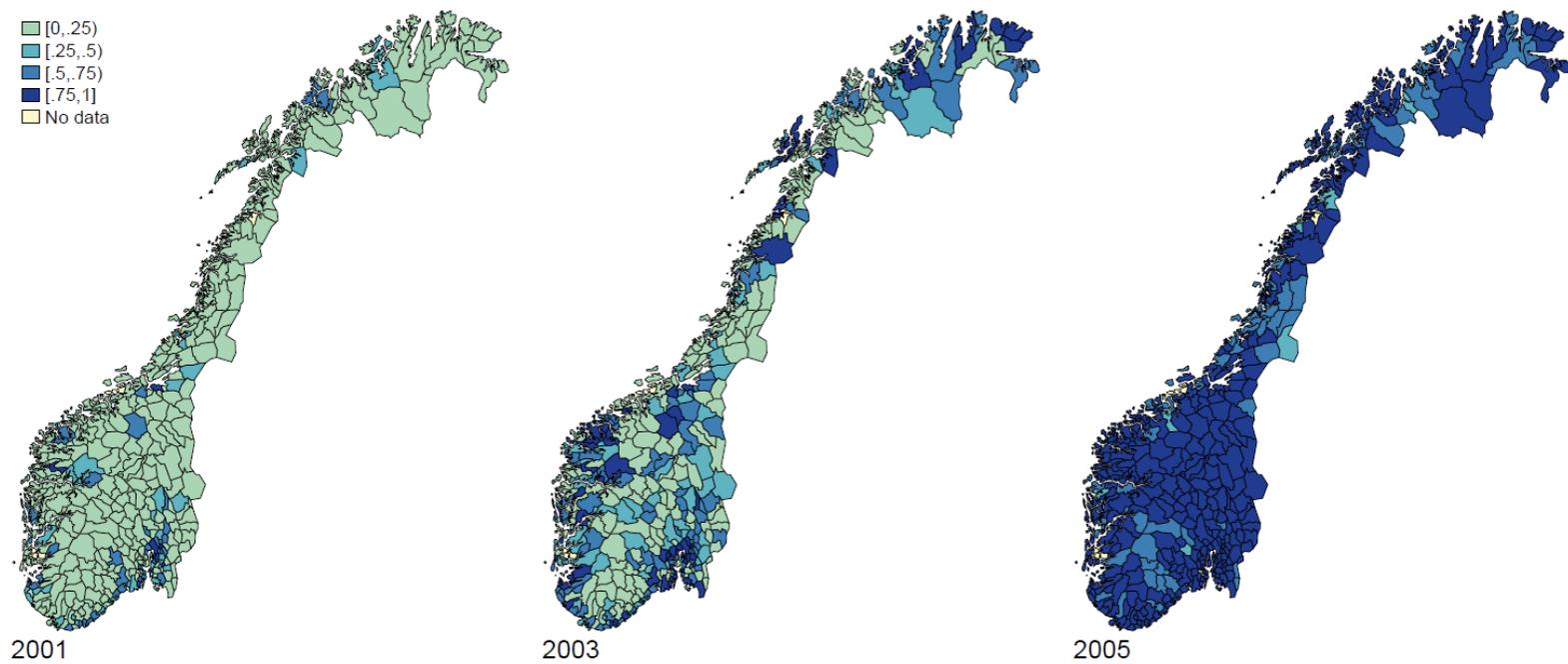
Table 1: McDonalds Cashier or Crew Wages and Big Mac Prices, December 1998

<i>Country</i>	Estimated hourly wage rate	Reported Big Mac price	Exchange Rate per \$1	\$ hourly wage rate	\$ Big Mac price	Economist \$ Big Mac 3/99**	Big Macs per hour of work
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Russia	8.00	25.5	19.57	0.41	1.30		0.31
Korea	1700	3000	1210	1.41	2.48		0.57
Brazil	2.87	4.45	1.73*	1.66	2.57	1.71	0.65
Poland	4.12	5.3	3.50	1.18	1.51	1.38	0.78
Czech Rep.	45.00	53	30.30	1.49	1.75		0.85
UK	3.60		0.62*	5.80	3.07	3.07	1.89
USA	6.00		1.00	6.00	2.43	2.43	2.12
Germany	11.28	4.95	1.67	6.76	2.97	2.72	2.28
France	40.22	17.5	5.76	6.99	3.04	2.87	2.30
Italy	10417	4500	1646	6.33	2.73	2.5	2.31
Belgium	280.00	114	34.50	8.12	3.30		2.46
Sweden	64.90	25	8.03	8.09	3.11	2.88	2.60
Japan	844	280	120*	7.03	2.33	2.44	3.01

Source: Ashenfelter-Jurajda (2001)

Case study of complementarity: Rollout of broadband

Figure A4. Geographical distribution of broadband coverage rates.

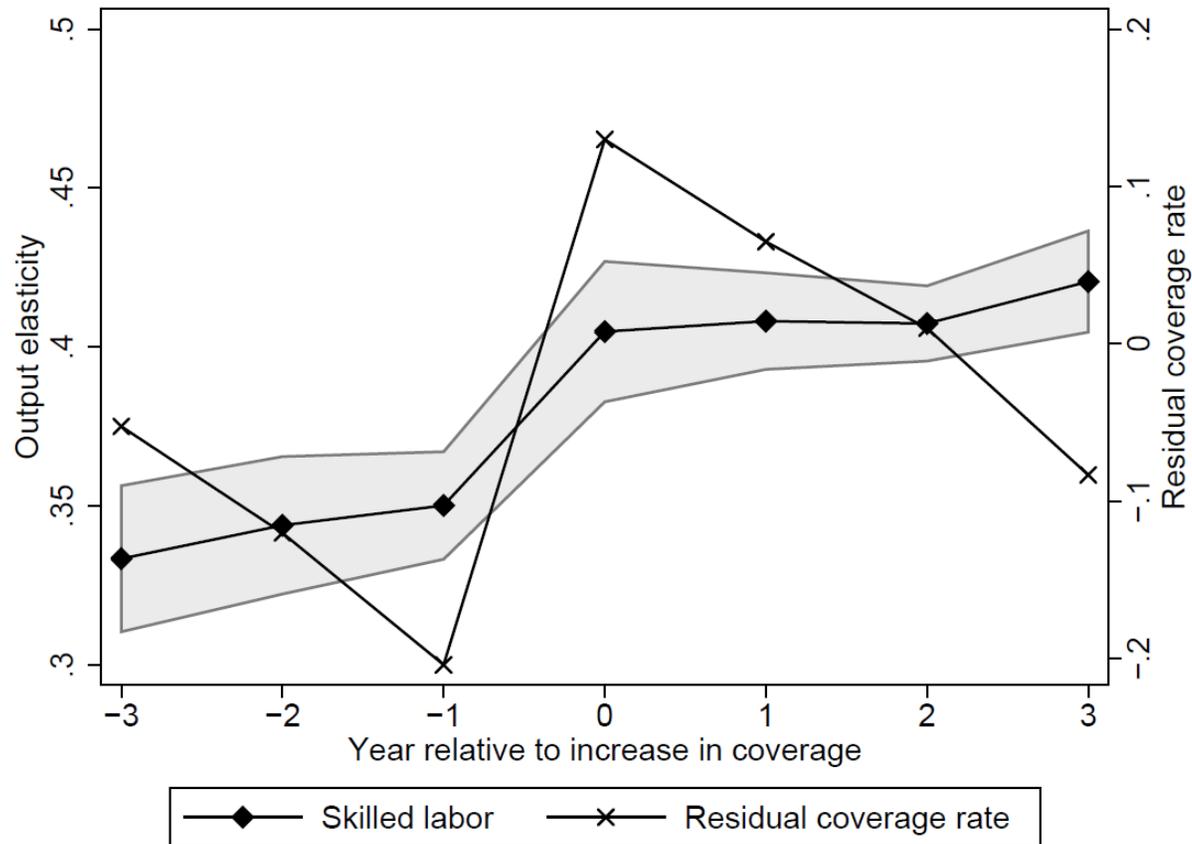


Note: The graphs show the geographical distribution of broadband coverage rates of households in 2001, 2003 and 2005.

Source: Akerman-Gaarder-Mogstad (2013)

Case study of complementarity: Rollout of broadband

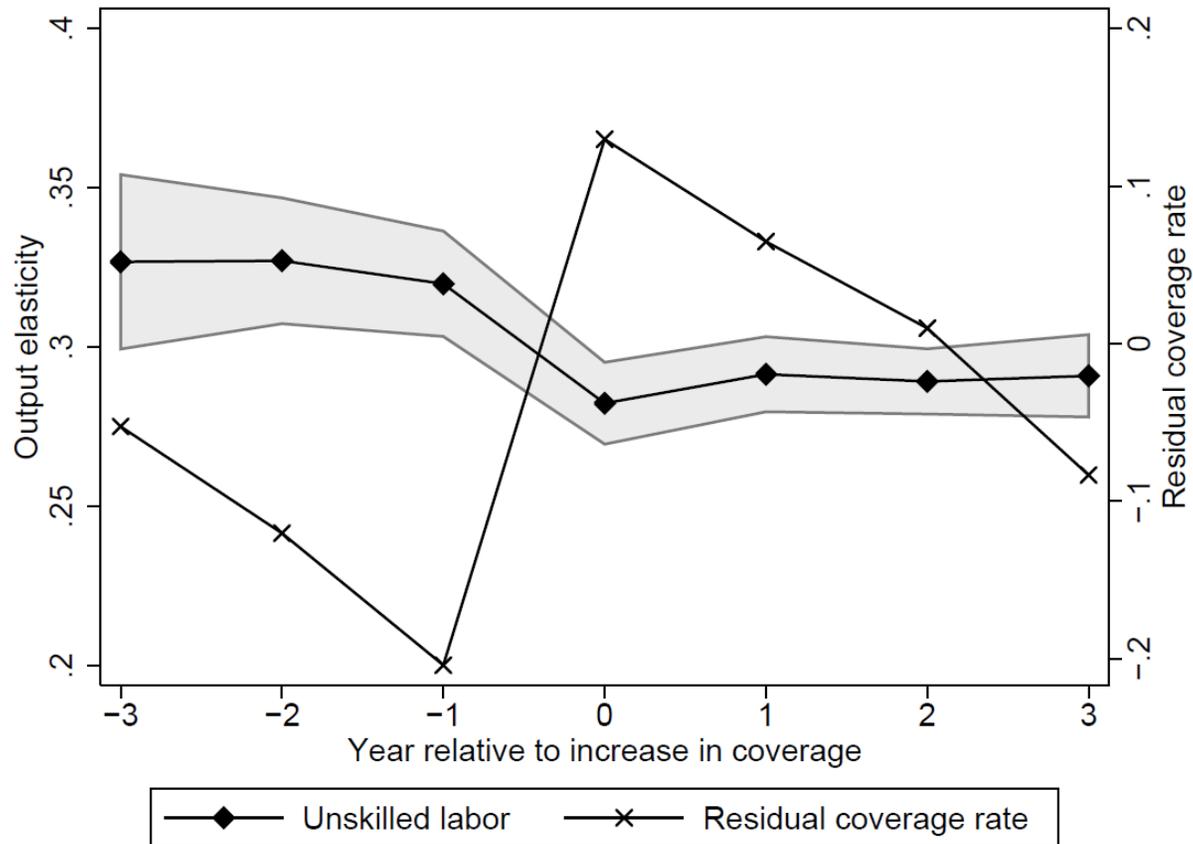
(a) Output elasticity: Skilled labor



Source: Akerman-Gaarder-Mogstad (2013)

Case study of complementarity: Rollout of broadband

(b) Output elasticity: Unskilled labor



Source: Akerman-Gaarder-Mogstad (2013)

(5) Are Govt's Fighting a Losing Business Tax Battle?

Firms are very good at minimizing taxes

Multinational firms avoid taxes through complex international relationships

Domestic firms avoid taxes by strategic use of entity classification

Taxation of Multinational Firms

Multinational firms: Firms that operate in multiple countries.

Subsidiaries: The production arms of a corporation that are located in other nations.

Territorial tax system: Corporations earning income abroad pay tax only to the government of the country in which the income is earned (most countries use this system)

Global tax system: Corporations are taxed by their home countries on their income regardless of where it is earned (US system, but foreign profits taxed only when “repatriated”)

Foreign tax credit: U.S.-based multinational corporations may claim a credit against their U.S. taxes for any tax payments made to foreign governments when funds are repatriated to the US parent.

Tax Avoidance of Multinationals (Zucman '14)

Share of profits made abroad by US corporations is about 1/3 today (was less than 5% in the 1930s)

50% of foreign profits of multinationals are reported in tax havens (such as Ireland)

Multinational companies are particularly savvy to avoid corporate income tax by reporting most of their profits in low tax countries using **transfer pricing**: one subsidiary buys/sells to another at manipulated prices to transfer profits

⇒ 20% of profits of US corporations is retained in tax havens

⇒ Effective corporate tax rate is lower than nominal US Federal tax rate

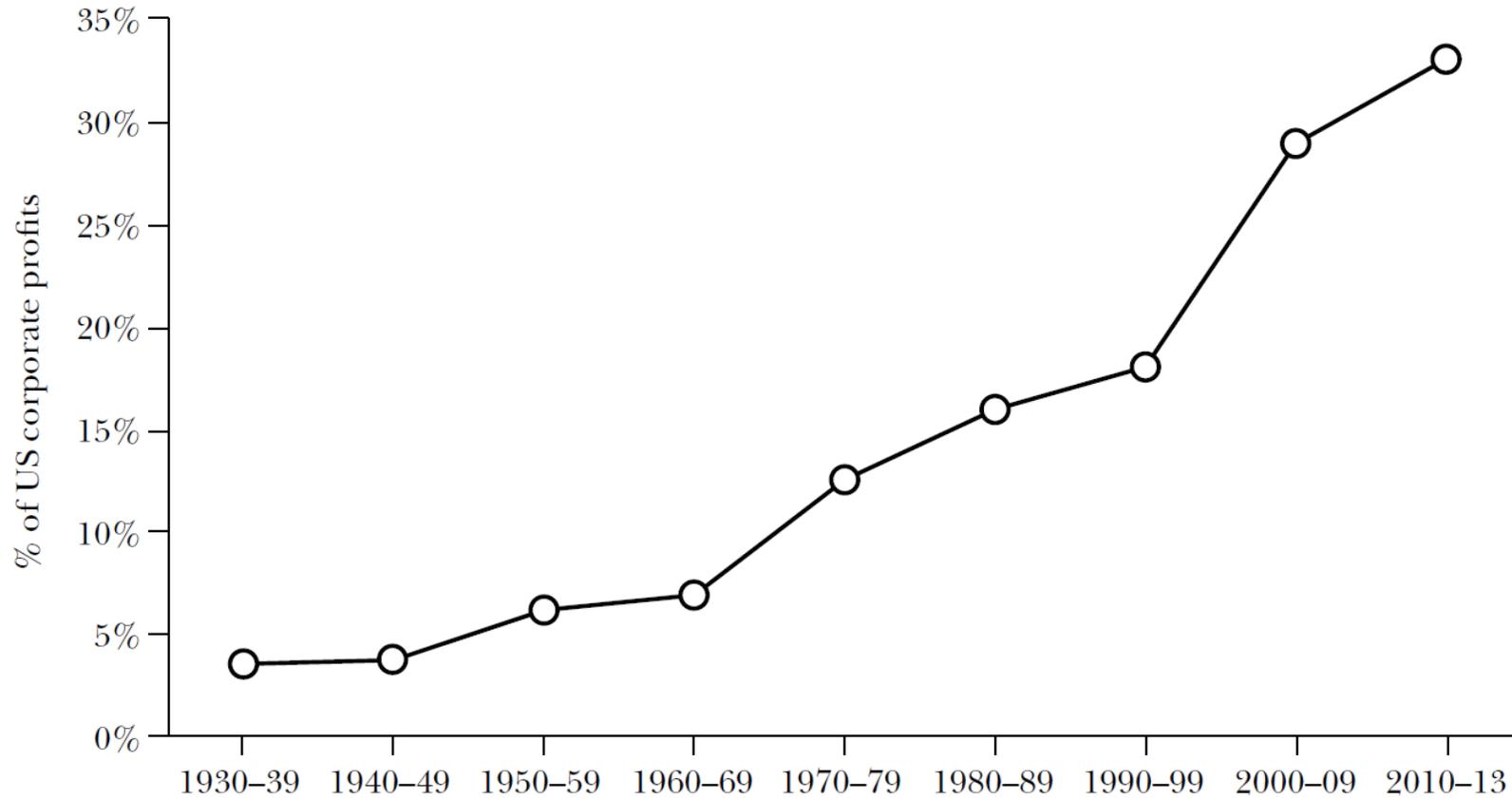
Ex: Microsoft's Irish sub Round Island Inc.



- In 2005: 1.5% of employees, 23% of assets
- Microsoft 2004 average tax rate: 33%
- Microsoft 2005 average tax rate: 26% due to “foreign earnings taxed at lower rates”

Rising importance of earnings booked abroad

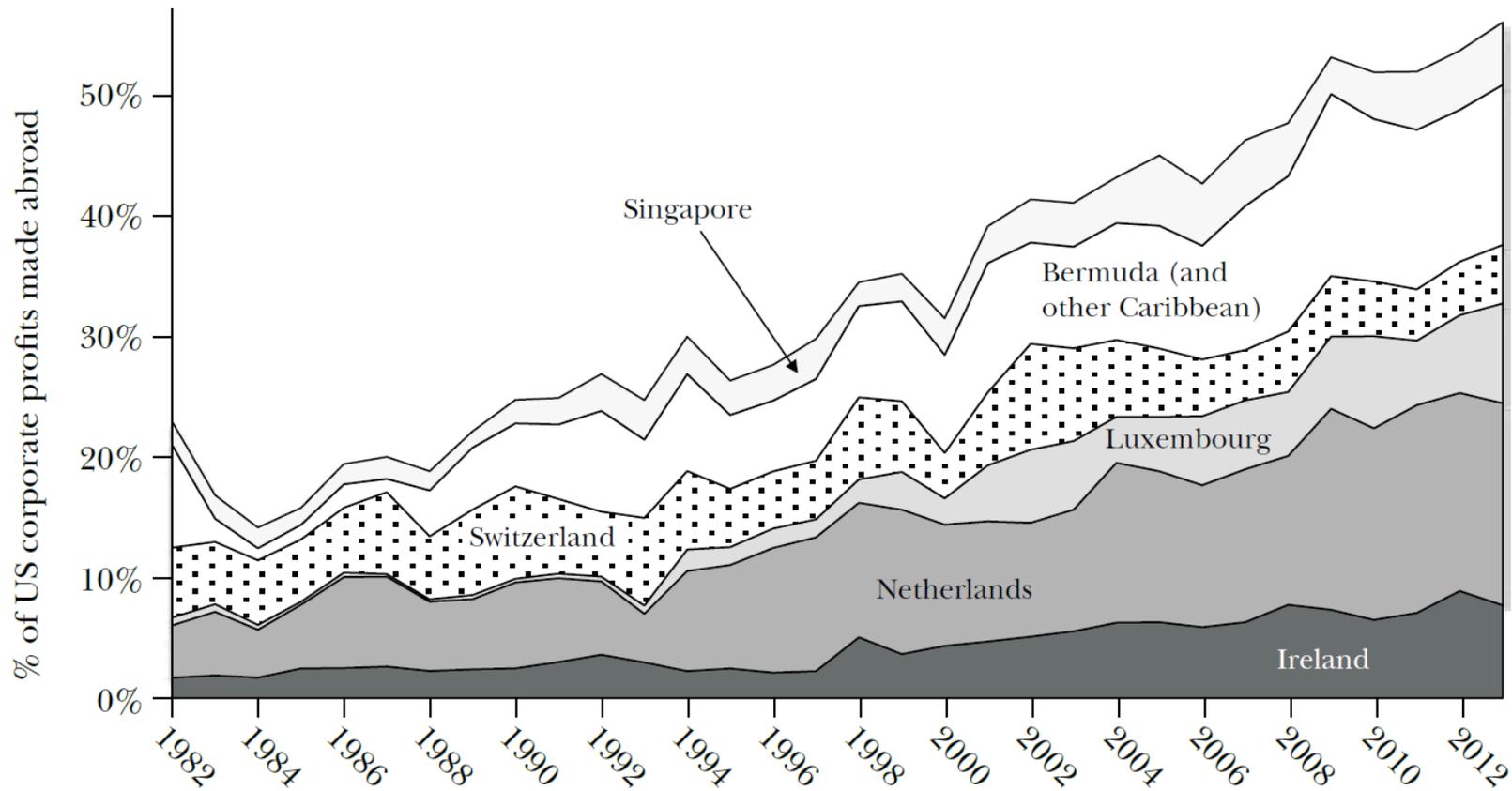
The Share of Profits Made Abroad in US Corporate Profits



Source: Zucman (2014)

Rising importance of earnings booked abroad

The Share of Tax Havens in US Corporate Profits Made Abroad



Source: Zucman (2014)

Repatriation Holidays

Owners eventually want the income repatriated from abroad and paid out to them as dividends

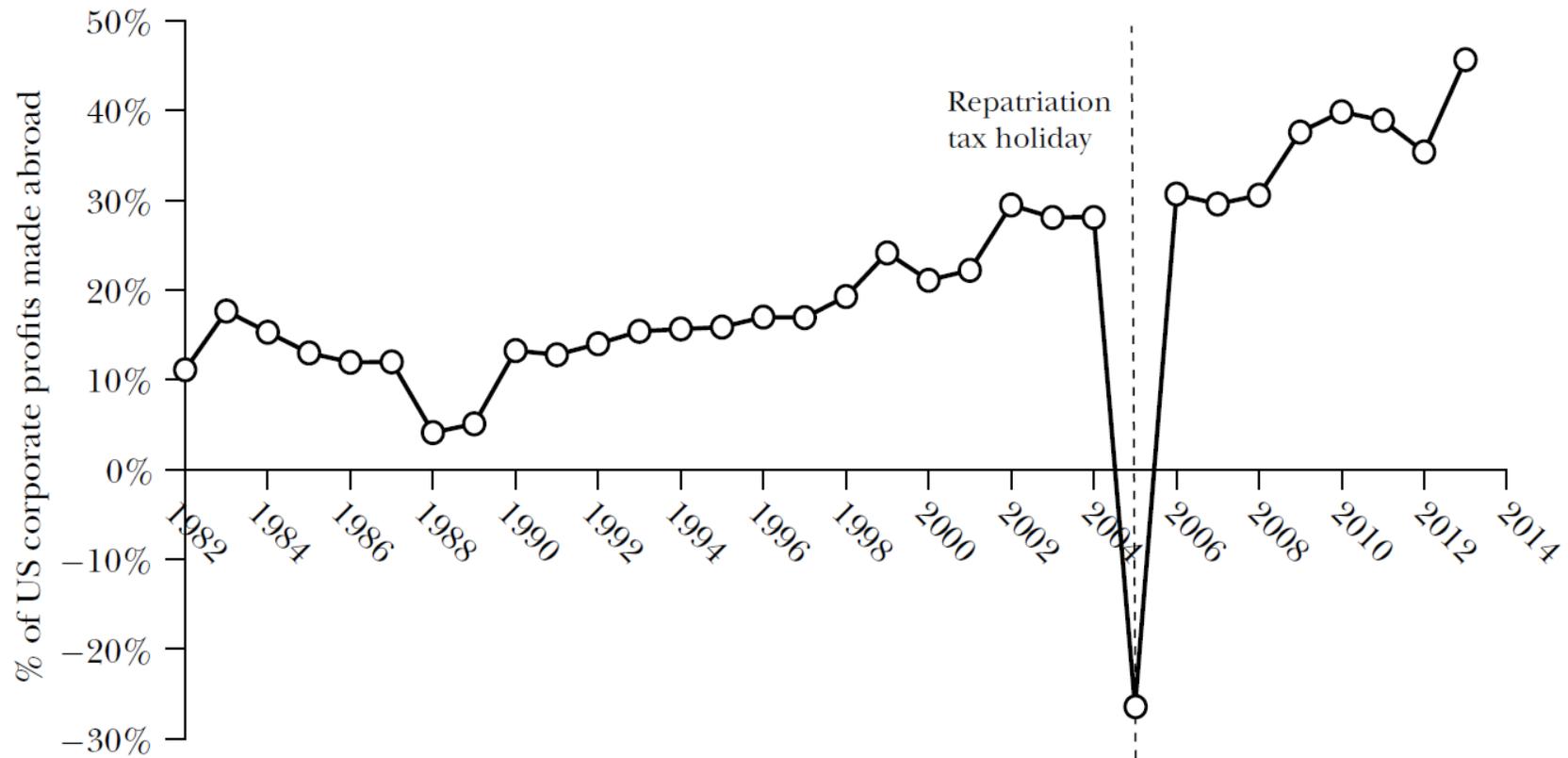
Corporations typically pay 35% tax on foreign profits once repatriated

Massive amount of profits kept abroad \Rightarrow Temptation for politicians to offer repatriation tax holiday

American Jobs Creation Act of 2004: Reduced tax rate on repatriated profits from 35% to 5.25% for 2005 only

Sensitivity of repatriations to tax rate on repatriations

US Corporate Profits Retained in Tax Havens



Source: Zucman (2014)

Inversions

Other way for U.S. corporations to dodge U.S. corporate tax: change country of incorporation to a tax haven

Cannot just say “I’m an Irish corporation now.” Must merge with an Irish corporation first, called “corporate inversion”

Ex. Medtronic (maker of heart pacemakers) merged with Irish Covidien in 2014 → Declared legal headquarters in Ireland → Avoided U.S. tax on \$14bn held overseas

Potential rationale for low U.S. corporate tax rate: Corporations will move headquarters/jobs overseas

No evidence though that many actual jobs move (e.g. Medtronic kept operational HQ in Minnesota)

Domestic Firms

Businesses taxed as either C-corporations or pass-throughs

C-corporate taxation: Annual corporate income tax (35%)
+ dividend tax (now 20%)

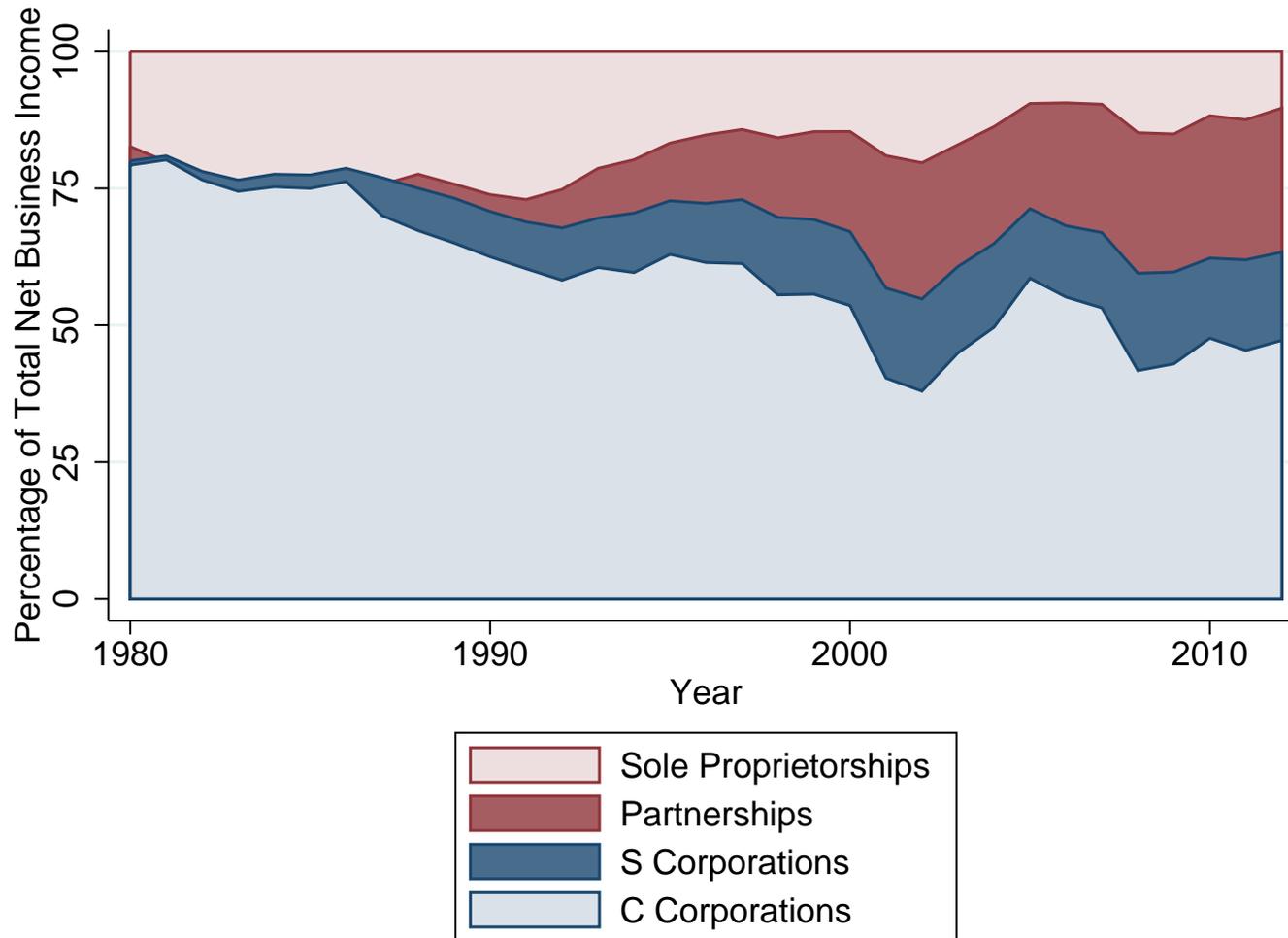
Pass-through business taxation: Tax annual income at owner level (max 39.6%) + no dividend tax

Major pass-through types: S-corporations and Partnerships

Issues: Does not tax foreign owners. And allows firms to label profits as capital gains (e.g. hedge fund carried interest) to enjoy especially low rates (20%)

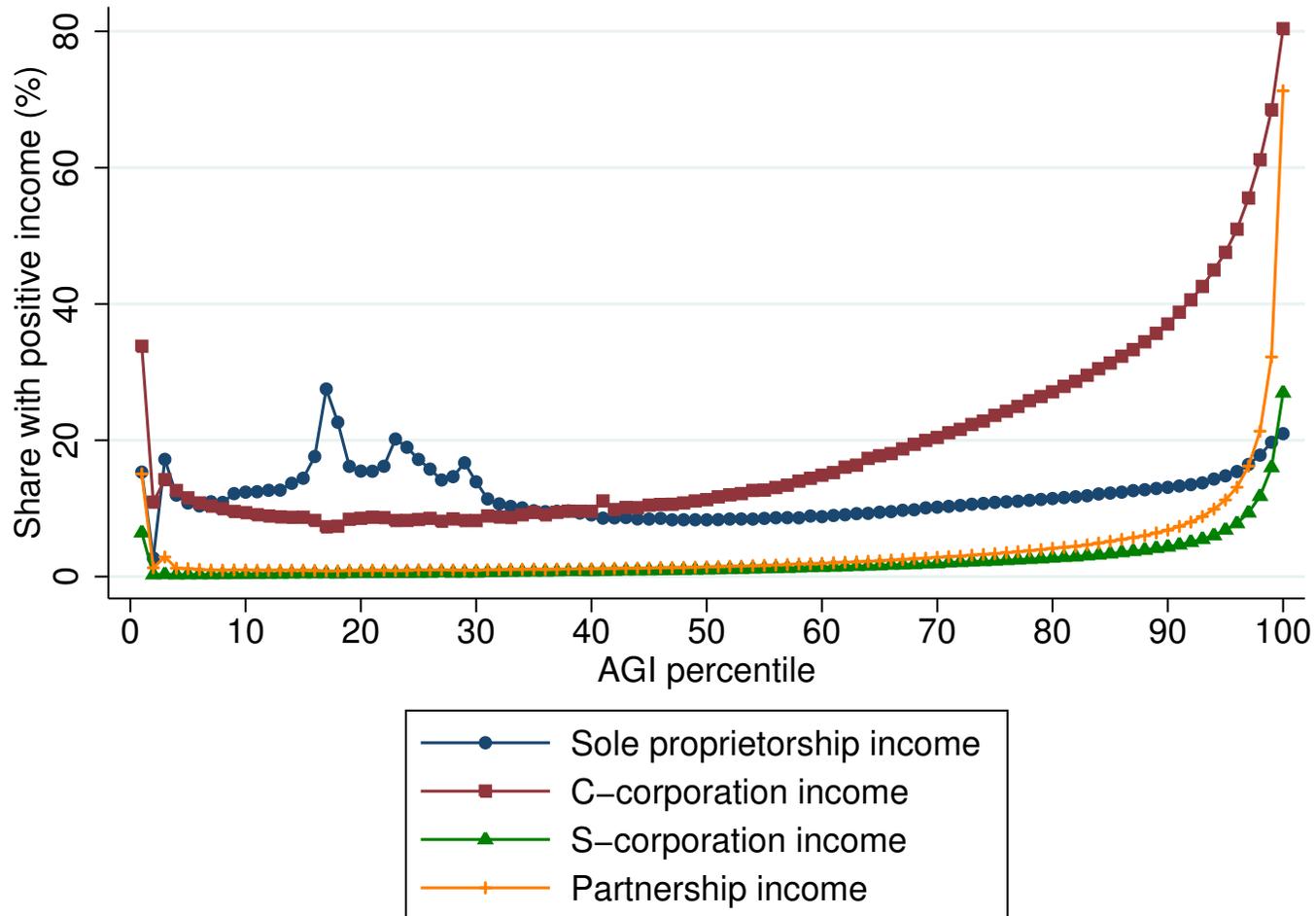
Evidence: Cooper-McClelland-Pearce-Prisinzano-Sullivan-Yagan-Zidar-Zwick (2016)

The Rise of Pass-Throughs



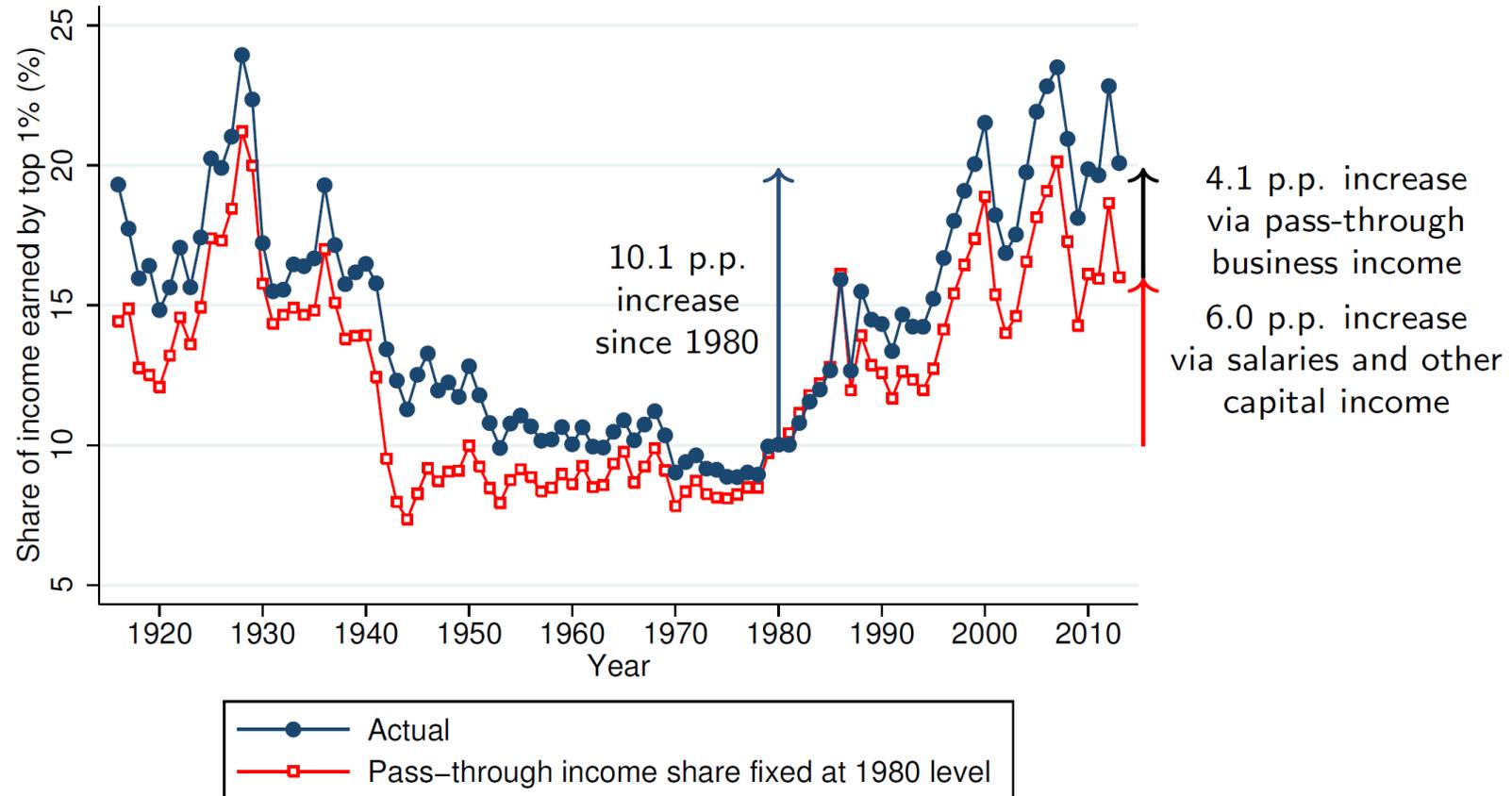
Source: Cooper-McClelland-Pearce-Prisinzano-Sullivan-Yagan-Zidar-Zwick (2016)

Business Participation Rates by AGI Pctile



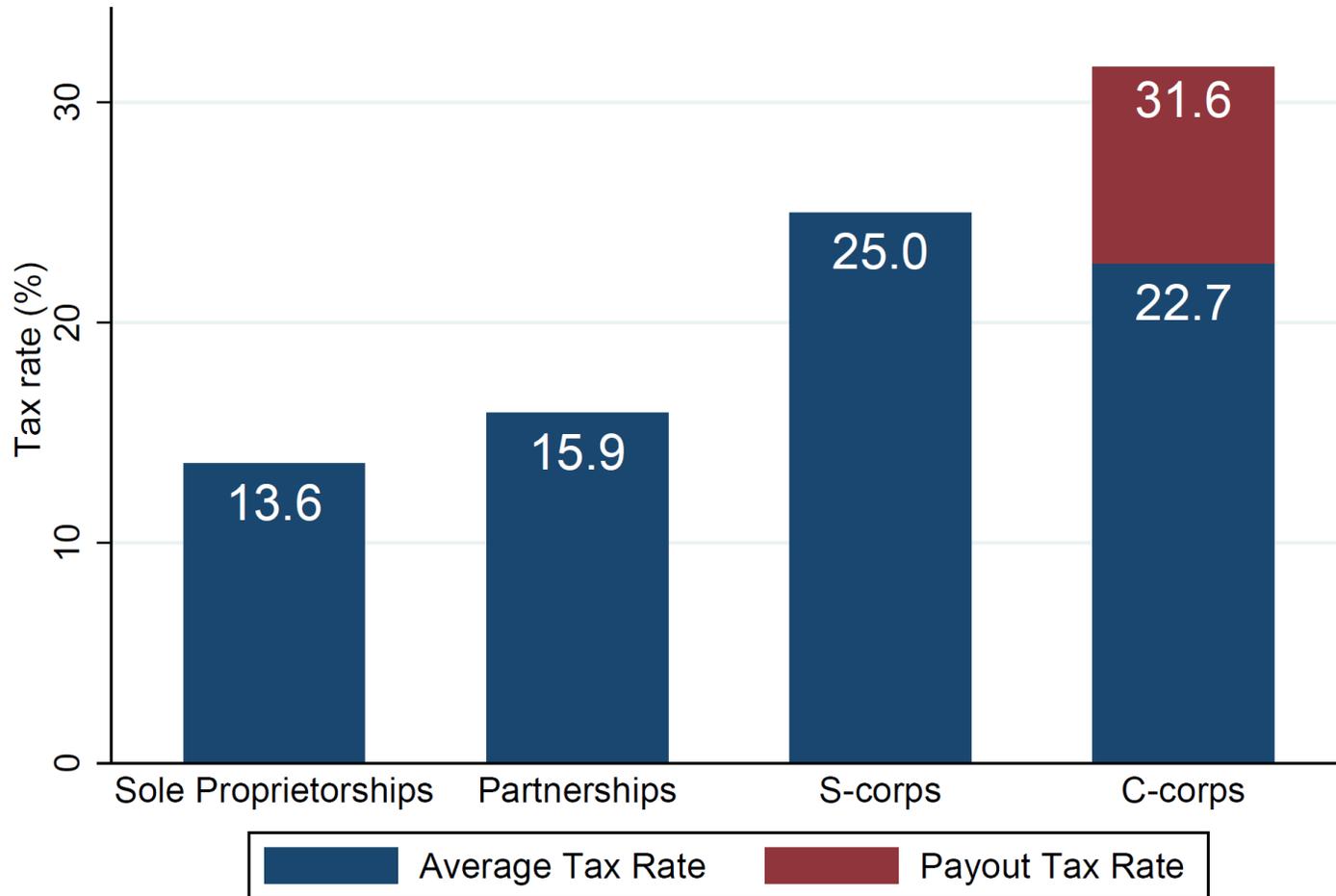
Source: Cooper-McClelland-Pearce-Prisinzano-Sullivan-Yagan-Zidar-Zwick (2016)

Pass-Throughs and Top-1% Income Shares



Source: Cooper-McClelland-Pearce-Prisinzano-Sullivan-Yagan-Zidar-Zwick (2016)

Tax Rate by Entity Type



Source: Cooper-McClelland-Pearce-Prisinzano-Sullivan-Yagan-Zidar-Zwick (2016)

Proposals to Tax Business More Effectively

Multinationals: Apportion global profits based on sales location

Ex. Microsoft 2005 profit = \$12bn and 67% of sales were to U.S. customers \Rightarrow \$8bn would be subject to U.S. tax

Makes Irish subsidiary/inversion irrelevant

Customers will not all move to Ireland

This is exactly what U.S. states do

Proposals to Tax Business More Effectively

Domestic businesses: Put size limit on pass-throughs

Difference between labor and capital income is fuzzy for small businesses \Rightarrow Good reason to tax both identically at owner level, even if small businesses taxed less than big businesses

Little good reason to allow big businesses to be taxed at different rates (e.g. Home Depot vs. Menards)

REFERENCES

- Akerman, Anders, Ingvil Gaarder, and Magne Mogstad. 2015. "The Skill Complementarity of Broadband Internet." *QJE*. (web)
- Ashenfelter and Jurajda. "Cross-country Comparisons of Wage Rates: The Big Mac Index." Working Paper. 2001. (web)
- Auerbach, Alan J. 1979. "Wealth Maximization and the Cost of Capital." *Quarterly Journal of Economics*, 93(3): 433-46. (web)
- Bradford, David F. 1981. "The Incidence and Allocation Effects of a Tax on Corporate Distributions." *Journal of Public Economics*, 15(1): 1–22. (web)
- Cooper, Michael, John McClelland, James Pearce, Richard Prisinzano, Joseph Sullivan, Danny Yagan, Owen Zidar, and Eric Zwick. 2016. *Tax Policy and the Economy* (web)
- Hall, Robert E. and Dale W. Jorgenson. 1967. "Tax Policy and Investment Behavior." *American Economic Review*, 57(3): 391-414. (web)
- Hassett, Kevin A. and R. Glenn Hubbard. 2002. "Tax Policy and Business Investment." In Alan J. Auerbach and Martin Feldstein, editors, *Handbook of Public Economics*, 1293–1343.

Judd, K. "Redistributive Taxation in a Simple Perfect Foresight Model", *Journal of Public Economics*, Vol. 28, 1985, 59-83. (web)

King, Mervyn A. 1977. *Public Policy and the Corporation*. London: Chapman and Hall.

Mankiw, N. Gregory. "The Savers-Spenders Theory of Fiscal Policy." *AER Papers and Proceedings*. 2000. (web)

Obama, Barack. "The President's Framework for Business Tax Reform." (web)

Poterba, James M., and Lawrence H. Summers. 1985. "The Economic Effects of Dividend Taxation." In *Recent Advances in Corporate Finance*, ed. Edward I. Altman and Marti G. Subrahmanyam, 227-84. Homewood, IL: Dow Jones-Irwin Publishing. (web)

Robertson, D. H. "The Colwyn Committee, the Income Tax, and the Price Level." *Economic Journal*, 1927.(web)

Yagan, Danny. "Capital Tax Reform and the Real Economy: The Effects of the 2003 Dividend Tax Cut." *American Economic Review*, 2015.(web)

Zucman, Gabriel "Taxing Across Borders: Tracking Personal Wealth and Corporate Profits," *Journal of Economic Perspectives*, 2014, 28(4): 121-148.(web)

Undergraduate Public Economics

Danny Yagan
UC Berkeley

Lecture 11 Externalities

OUTLINE

Second part of course is going to cover market failures and show how government interventions can help

- 1) Externalities and public goods
- 2) Asymmetric information (social insurance)
- 3) Individual failures (savings for retirement)

EXTERNALITIES

Market failure: A problem that violates one of the assumptions of the 1st welfare theorem and causes the market economy to deliver an outcome that does not maximize efficiency

Externality: Externalities arise whenever the actions of one economic agent **directly** affect another economic agent outside the market mechanism

Externality example: a steel plant that pollutes a river used for recreation

Not an externality example: a steel plant uses more electricity and bids up the price of electricity for other electricity customers

Externalities are one important case of market failure

EXTERNALITY THEORY: ECONOMICS OF NEGATIVE PRODUCTION EXTERNALITIES

Negative production externality: When a firm's production reduces the well-being of others who are not compensated by the firm.

Private marginal cost (PMC): The direct cost to producers of producing an additional unit of a good

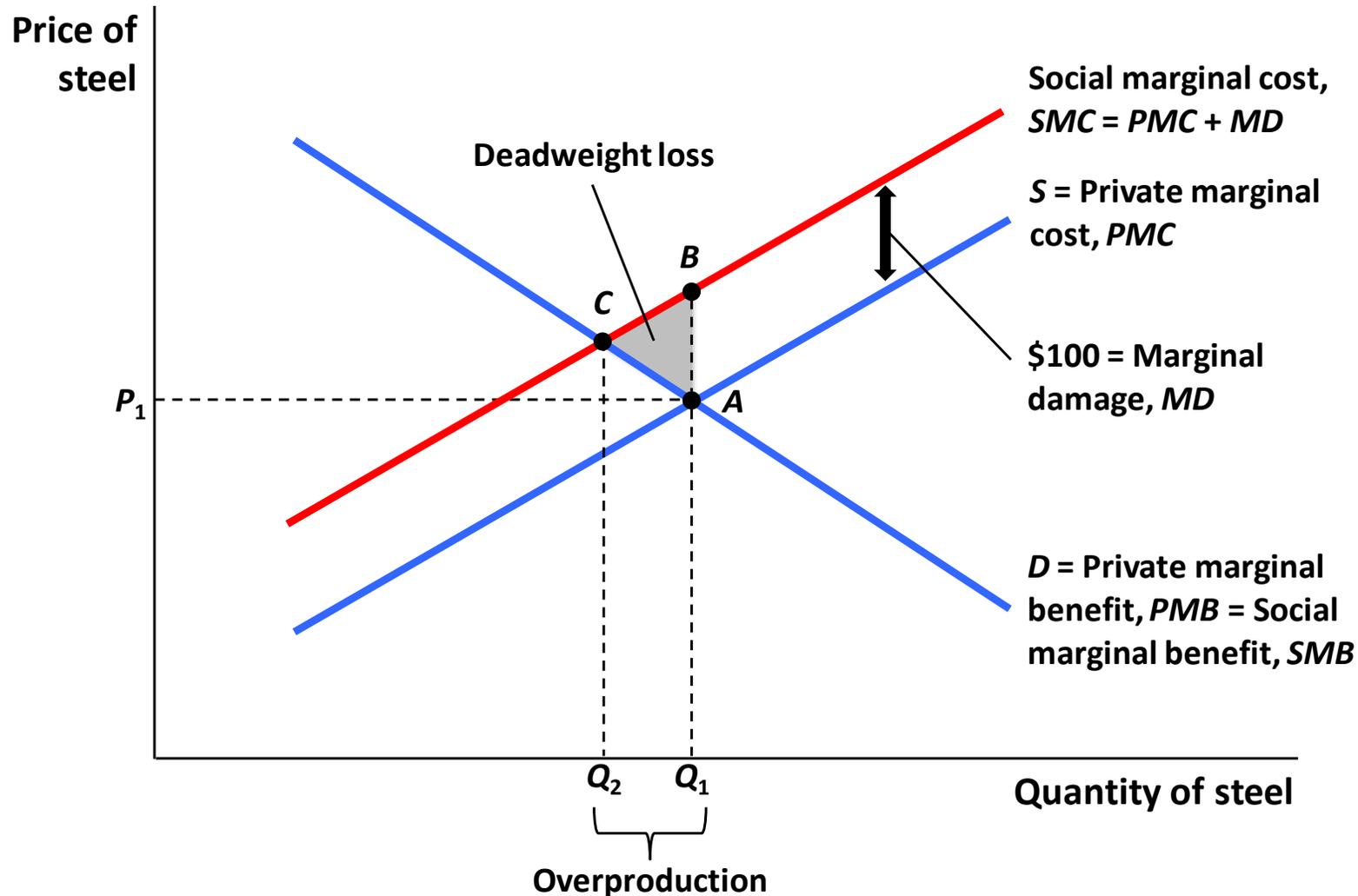
Marginal Damage (MD): Any additional costs associated with the production of the good that are imposed on others but that producers do not pay

Social marginal cost (SMC = PMC + MD): The private marginal cost to producers plus marginal damage

Example: steel plant pollutes a river but plant does not face any pollution regulation (and hence ignores pollution when deciding how much to produce)

5.1

Economics of Negative Production Externalities: Steel Production



EXTERNALITY THEORY: ECONOMICS OF NEGATIVE CONSUMPTION EXTERNALITIES

Negative consumption externality: When an individual's consumption reduces the well-being of others who are not compensated by the individual.

Private marginal benefit (PMB): The direct benefit to consumers of consuming an additional unit of a good by the consumer.

Social marginal benefit (SMB): The private marginal benefit to consumers plus any costs associated with the consumption of the good that are imposed on others

Example: Using a car and emitting carbon contributing to global warming

5.1

APPLICATION: The Externality of SUVs

The consumption of large cars such as SUVs produces three types of negative externalities:

1. Environmental externalities: Compact cars get 25 miles/gallon, but SUVs get only 20.
2. Wear and tear on roads: Larger cars wear down the roads more.
3. Safety externalities: The odds of having a fatal accident quadruple if the accident is with a typical SUV and not with a car of the same size.

Externality Theory: Positive Externalities

Positive production externality: When a firm's production increases the well-being of others but the firm is not compensated by those others. ($SMC < PMC$)

Example: Beehives of honey producers have a positive impact on pollination and agricultural output

Positive consumption externality: When an individual's consumption increases the well-being of others but the individual is not compensated by those others. ($SMB > PMB$)

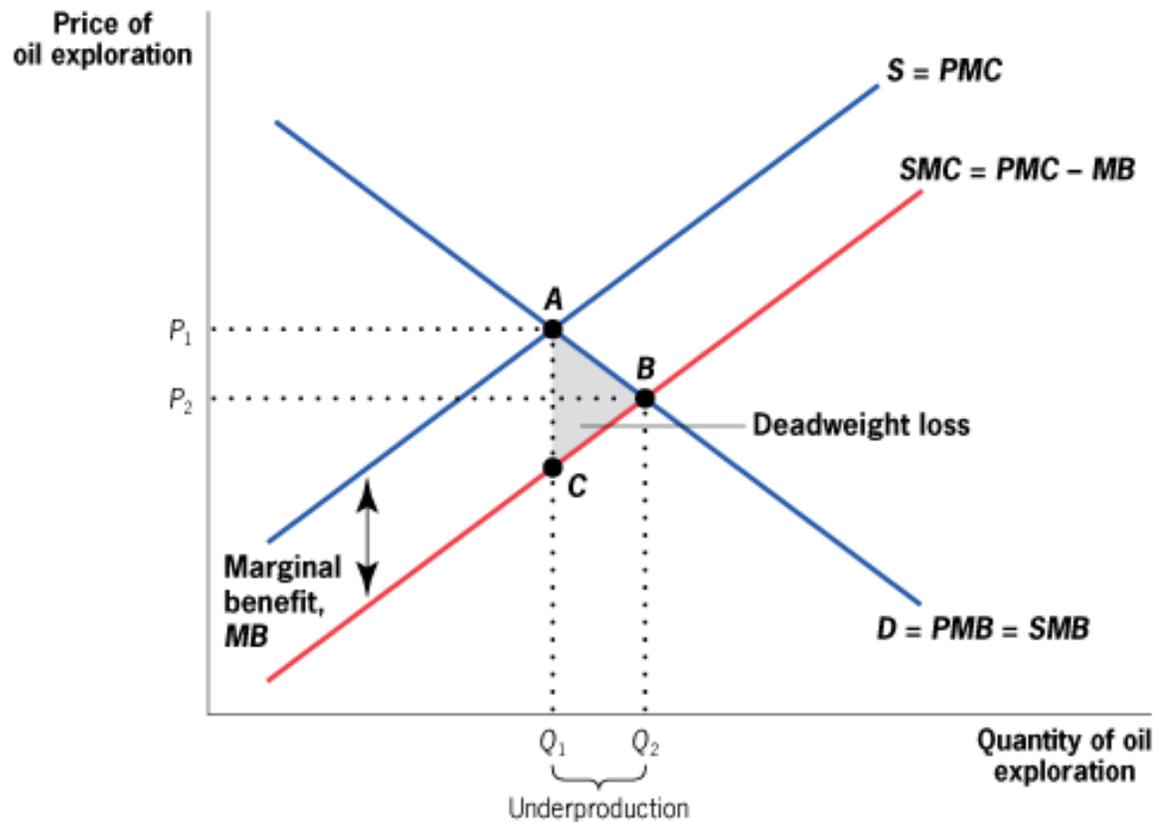
Example: Beautiful private garden that passers-by enjoy seeing

5.1

Externality Theory

Positive Externalities

■ FIGURE 5-4



Market Failure Due to Positive Production Externality in the Oil Exploration Market • Expenditures on oil exploration by any company have a positive externality because they offer more profitable opportunities for other companies. This leads to a social marginal cost that is below the private marginal cost, and a social optimum quantity (Q_2) that is greater than the competitive market equilibrium quantity (Q_1). There is underproduction of $Q_2 - Q_1$, with an associated deadweight loss of area ABC.

Externality Theory: Market Outcome is Inefficient

With a free market, quantity and price are such that $PMB = PMC$

Social optimum is such that $SMB = SMC$

⇒ Private market leads to an inefficient outcome (1st welfare theorem does not work)

Negative production externalities lead to over production

Positive production externalities lead to under production

Negative consumption externalities lead to over consumption

Positive consumption externalities lead to under consumption

Private-Sector Solutions to Negative Externalities

Key question raised by Ronald Coase (famous Nobel Prize winner Chicago libertarian economist):

Are externalities really outside the market mechanism?

Internalizing the externality: When either private negotiations or government action lead the price to the party to fully reflect the external costs or benefits of that party's actions.

PRIVATE-SECTOR SOLUTIONS TO NEGATIVE EXTERNALITIES: COASE THEOREM

Coase Theorem (Part I): When (a) there are well-defined property rights, (b) bargaining is costless, and (c) markets are competitive, then negotiations between the party creating the externality and the party affected by the externality can bring about the socially optimal market quantity.

Coase Theorem (Part II): The efficient quantity for a good producing an externality does not depend on which party is assigned the property rights, as long as someone is assigned those rights.

COASE THEOREM EXAMPLE

Firms pollute a river enjoyed by individuals. If firms ignore individuals, there is too much pollution

1) Individuals own river: If river is owned by individuals then individuals can charge firms for polluting the river. They will charge firms the marginal damage (MD) per unit of pollution.

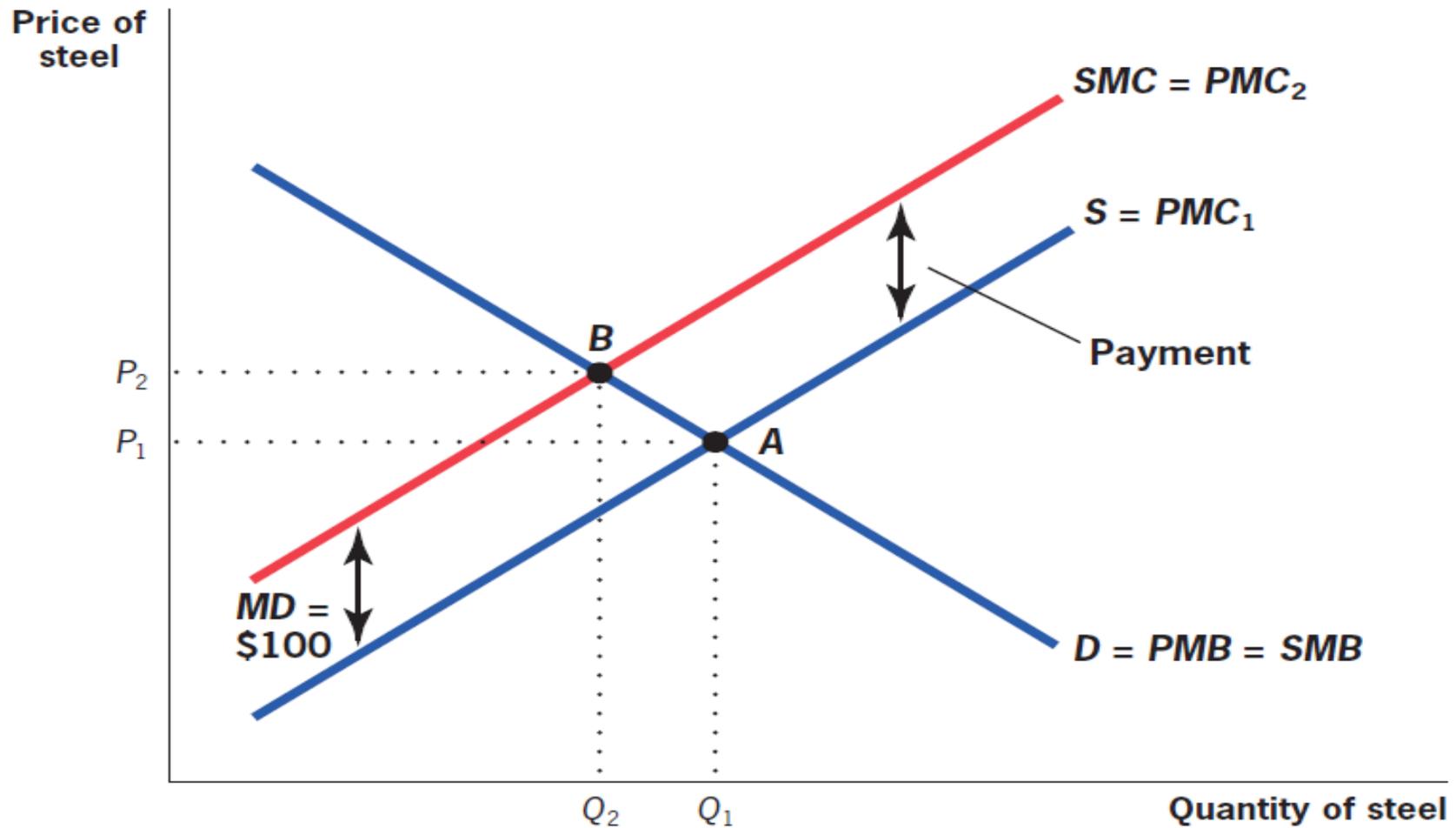
Why price pollution at MD? If charge more than MD, individuals would want to sell an extra unit of pollution, so price must rise under competition. MD is the equilibrium efficient price in the newly created pollution market.

2) Firms own river: If river is owned by firms then firm can charge individuals for polluting less. They will also charge individuals the MD per unit of pollution reduction.

Final level of pollution will be the same in 1) and 2)

5.2

The Solution: Coasian Payments



PROBLEMS WITH COASIAN SOLUTION

In practice, the Coase theorem is unlikely to solve many of the types of externalities that cause market failures.

1) The assignment problem: In cases where externalities affect many agents (e.g. global warming), assigning property rights is difficult

⇒ Coasian solutions are likely to be more effective for small, localized externalities than for larger, more global externalities involving large number of people and firms

2) The holdout problem: Shared ownership of property rights gives each owner power over all the others (because joint owners have to all agree to the Coasian solution)

As with the assignment problem, the holdout problem would be amplified with an externality involving many parties.

PROBLEMS WITH COASIAN SOLUTION

3) The Free Rider Problem: When an investment has a personal cost but a common benefit, individuals will underinvest (example: a single country is better off walking out of Kyoto protocol for carbon emission controls)

4) Transaction Costs and Negotiating Problems: The Coasian approach ignores the fundamental problem that it is hard to negotiate when there are large numbers of individuals on one or both sides of the negotiation.

This problem is amplified for an externality such as global warming, where the potentially divergent interests of billions of parties on one side must be somehow aggregated for a negotiation.

PROBLEMS WITH COASIAN SOLUTION: BOTTOM LINE

Ronald Coase's insight that externalities can sometimes be internalized was useful.

It provides the competitive market model with a defense against the onslaught of market failures.

It is also an excellent reason to suspect that the market may be able to internalize some small-scale, localized externalities.

It won't help with large-scale, global externalities, where only a "government" can successfully aggregate the interests of all individuals suffering from externality

Public Sector Remedies For Externalities

The Environmental Protection Agency (EPA) was formed in 1970 to provide public-sector solutions to the problems of externalities in the environment.

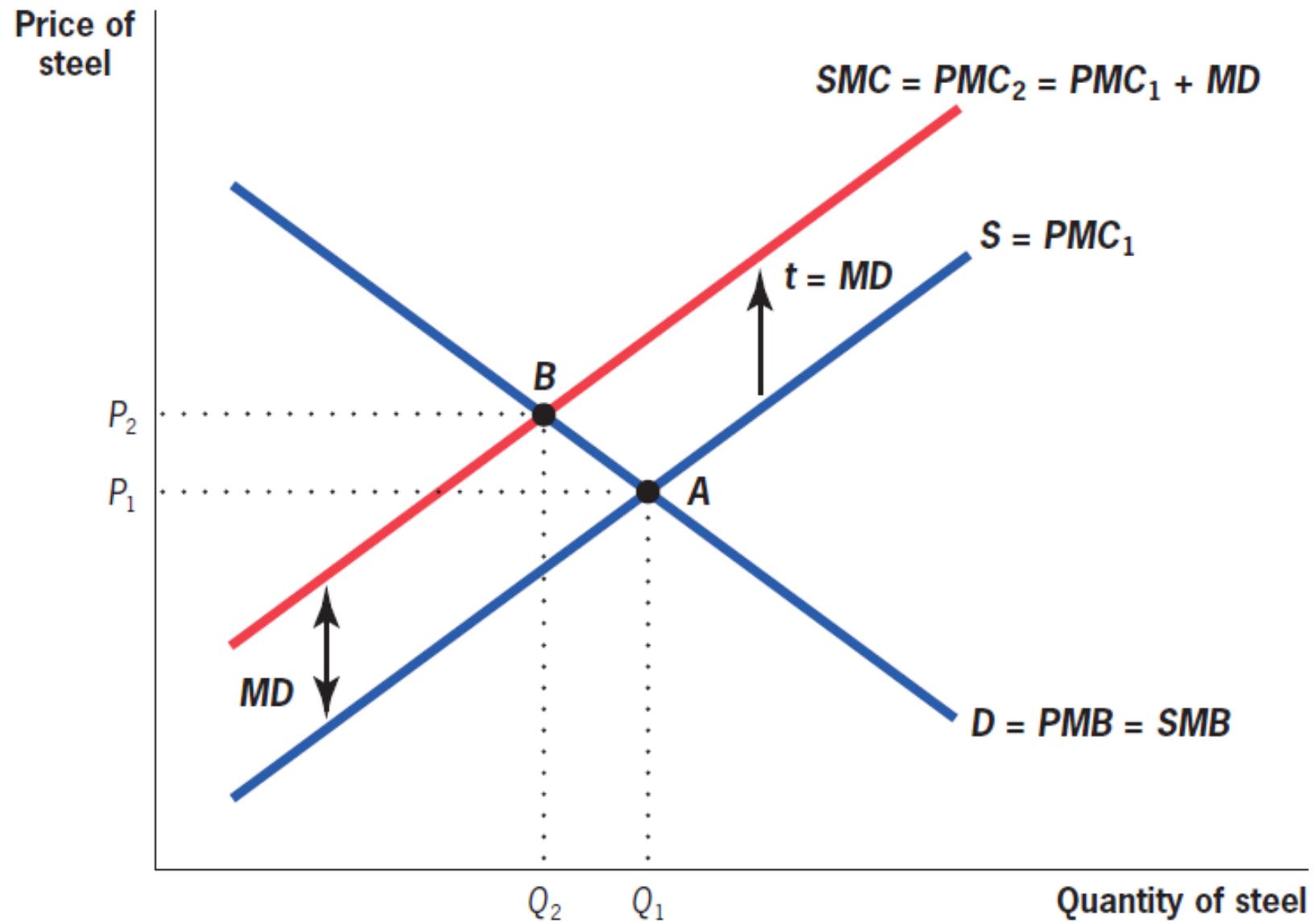
Public policy makers employ two types of remedies to resolve the problems associated with negative externalities:

1) price policy: corrective tax or subsidy equal to marginal damage per unit

2) quantity regulation: government forces firms to produce the socially efficient quantity

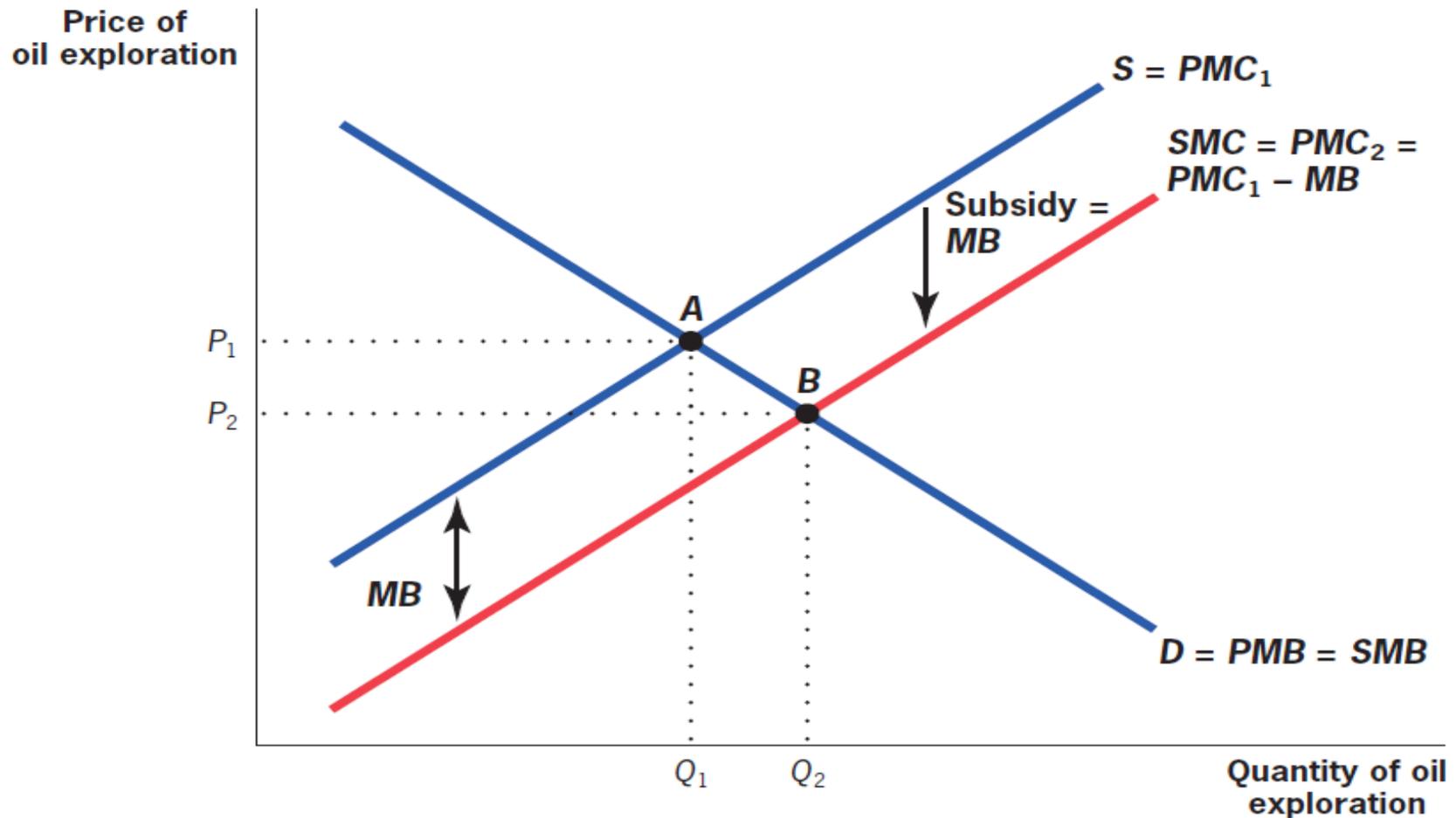
5.3

Corrective Taxation



5.3

Corrective Subsidies



PUBLIC SECTOR REMEDIES FOR EXTERNALITIES: REGULATION

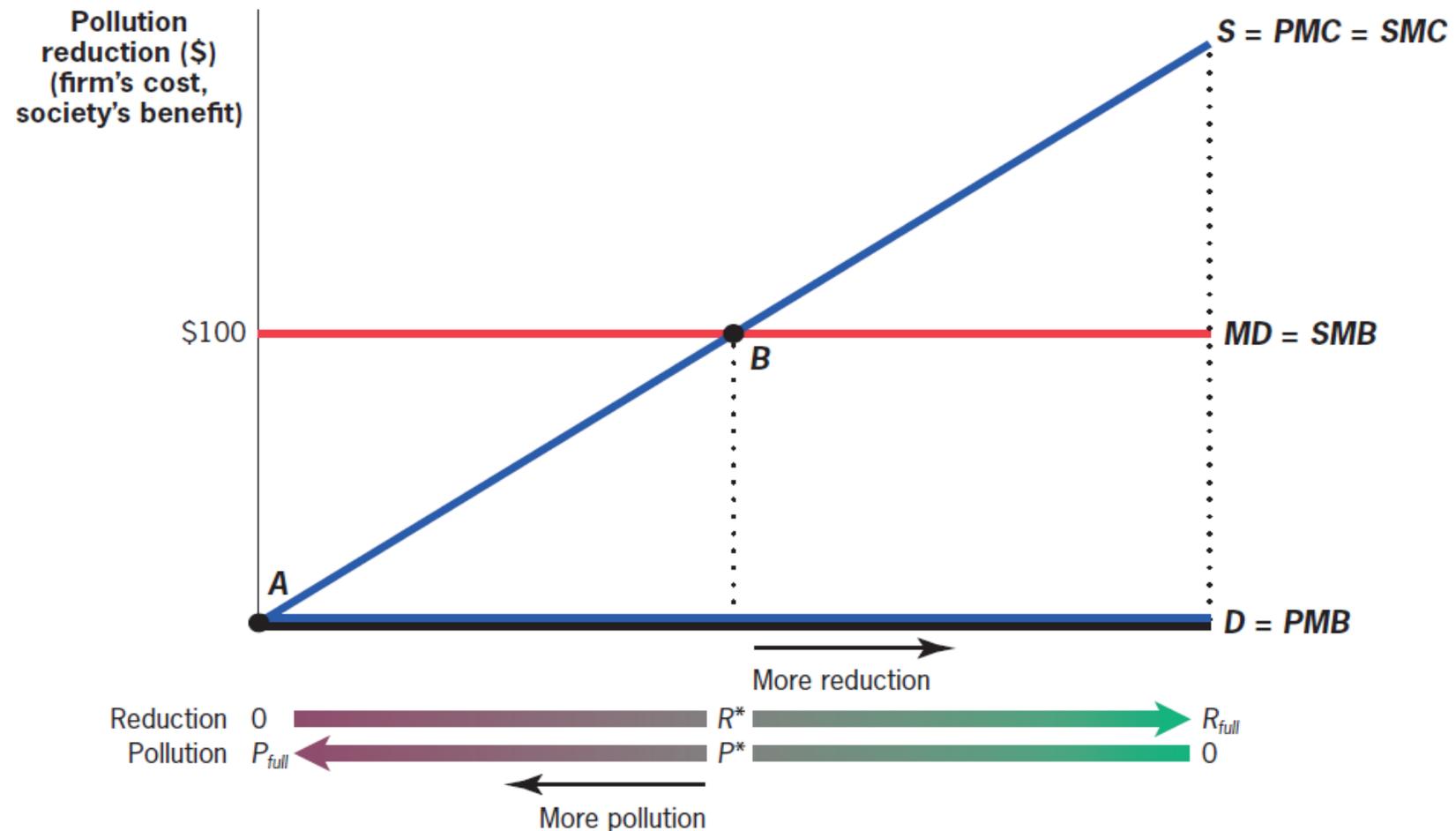
In an ideal world, Pigouvian taxation and quantity regulation would generate equal efficiency (exact same quantity: the optimal quantity)

Quantity regulation seems more straightforward, hence, it has been the traditional choice for addressing environmental externalities

In practice, there are complications that may make taxes a more effective means of addressing externalities.

5.4

Distinctions Between Price and Quantity Approaches to Addressing Externalities: Basic Model



MODEL WITH HETEROGENEOUS COSTS

Assume MD of pollution is \$1 per unit of pollution

2 firms with low (L) or high (H) cost of pollution reduction q :

$$c_H(q) = 1.5q^2 \Rightarrow MC_H(q) = c'_H(q) = 3q$$

$$c_L(q) = .75q^2 \Rightarrow MC_L(q) = c'_L(q) = 1.5q$$

With no taxes, no regulations, firms do $q_L = q_H = 0$

Social welfare maximization:

$$V = \max_{q_H, q_L} q^H + q^L - c_H(q^H) - c_L(q^L) \Rightarrow$$

$$MC_H = 1, MC_L = 1 \Rightarrow q^H = 1/3, q^L = 2/3$$

Optimum outcome is to have the low cost firm do more pollution reduction than the high cost firm

TAX VERSUS REGULATION SOLUTION

Socially optimal outcome can be achieved by \$1 tax per unit of pollution (same tax across firms):

Firm H chooses q_H to maximize $q^H - c_H(q^H) \Rightarrow MC_H = 1$

Firm L chooses q_L to maximize $q^L - c_L(q^L) \Rightarrow MC_L = 1$

Uniform quantity regulation $q^H = q^L = 1/2$ is not efficient because firm H has higher MC of polluting than firm L :

Proof: Firm H would be happy to pay firm L to reduce q^L and increase q^H to keep $q^L + q^H = 1$, firm L is happier and society has same level of pollution

Quantity Regulation with Trading Permits

Suppose start with quantity regulation $q_0^H = q_0^L = 1/2$ and allow firms to trade pollution reductions as long as $q^H + q^L = 1$

Generates a market for pollution reduction at price p

Firm H maximizes $pq^H - c_H(q^H) \Rightarrow MC_H = p$ and $q^H = p/3$

Firm L maximizes $pq^L - c_L(q^L) \Rightarrow MC_L = p$ and $q^L = 2p/3$

$\Rightarrow q^H + q^L = p$. As $1 = q_0^L + q_0^H = q^H + q^L$, in equilibrium $p = 1$ and hence $q_H = 1/3$ and $q_L = 2/3$

Final outcome q_H, q_L does not depend on initial regulation q_0^H, q_0^L

Quantity regulation with tradable permits is efficient as long as total quantity $q_0^L + q_0^H = 1$

MULTIPLE PLANTS WITH DIFFERENT REDUCTION COSTS

Policy Option 1: Price Regulation Through a Uniform Corrective Tax (efficient)

Policy Option 2: Uniform Quantity Regulation (not efficient unless quantity can be based on each firm's reduction costs)

Policy Option 3: Quantity Regulation with Tradable Permits (efficient)

CORRECTIVE TAXES VS. TRADABLE PERMITS

Two differences between corrective taxes and tradable permits (carbon tax vs. cap-and-trade in the case of CO₂ emissions)

1) Initial allocation of permits: If the government sells them to firms, this is equivalent to the tax

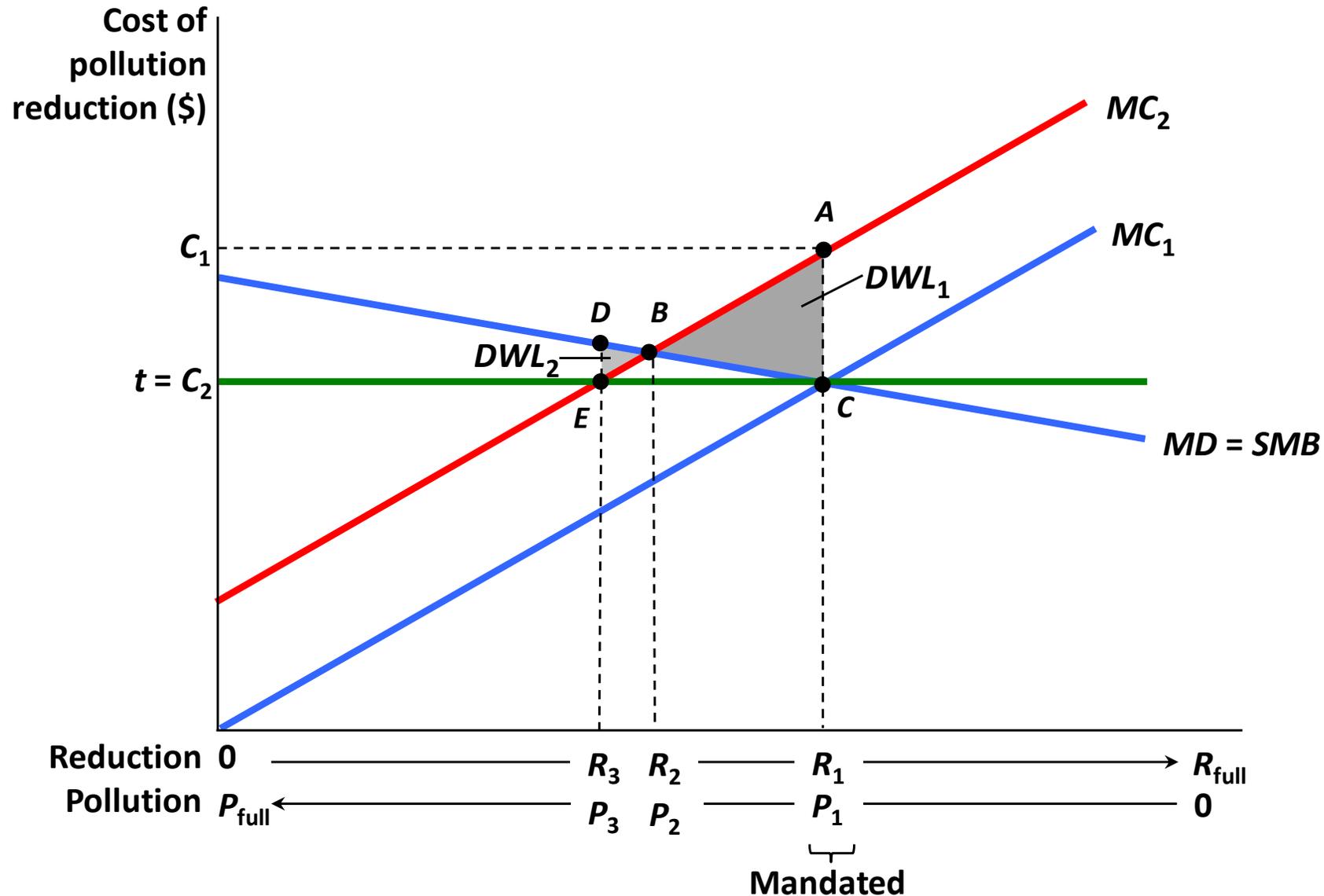
If the government gives them to current firms for free, this is like the tax + large transfer to initial polluting firms.

2) Uncertainty in marginal costs: With uncertainty in costs of reducing pollution, tax cannot target a specific quantity while tradable permits can \Rightarrow two policies no longer equivalent.

Taxes preferable when MD curve is flat. Tradable permits are preferable when MD curve is steep.

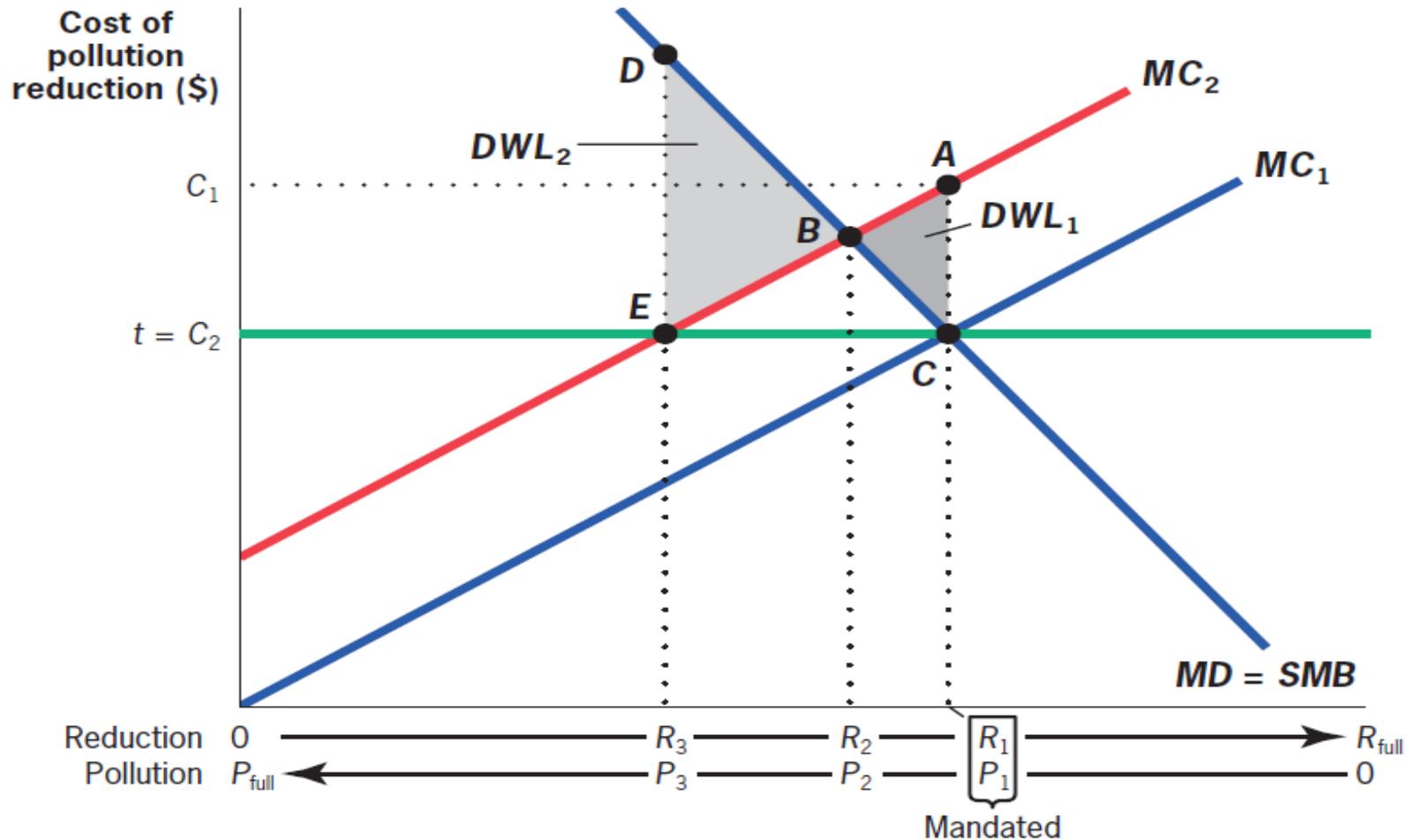
5.4

Uncertainty About Costs of Reduction: Case 1: Flat MD Curve (Global Warming)



5.4

Uncertainty About Costs of Reduction: Case 2: Steep MD Curve (Nuclear leakage)



Empirical Example: Acid Rain and Health

Acid rain due to contamination by emissions of sulfur dioxide (SO_2) and nitrogen oxide (NO_x).

1970 Clean Air Act: Landmark federal legislation that first regulated acid rain-causing emissions by setting maximum standards for atmospheric concentrations of various substances, including SO_2 .

The 1990 Amendments and Emissions Trading:

SO_2 allowance system: The feature of the 1990 amendments to the Clean Air Act that granted plants permits to emit SO_2 in limited quantities and allowed them to trade those permits.

Empirical Example: Effects of Clean Air Act of 1970

How does acid rain (or SO₂) affect health?

Observational approach: relate mortality in a geographical area to the level of particulates (such as SO₂) in the air

Problem: Areas with more particulates may differ from areas with fewer particulates in many other ways, not just in the amount of particulates in the air

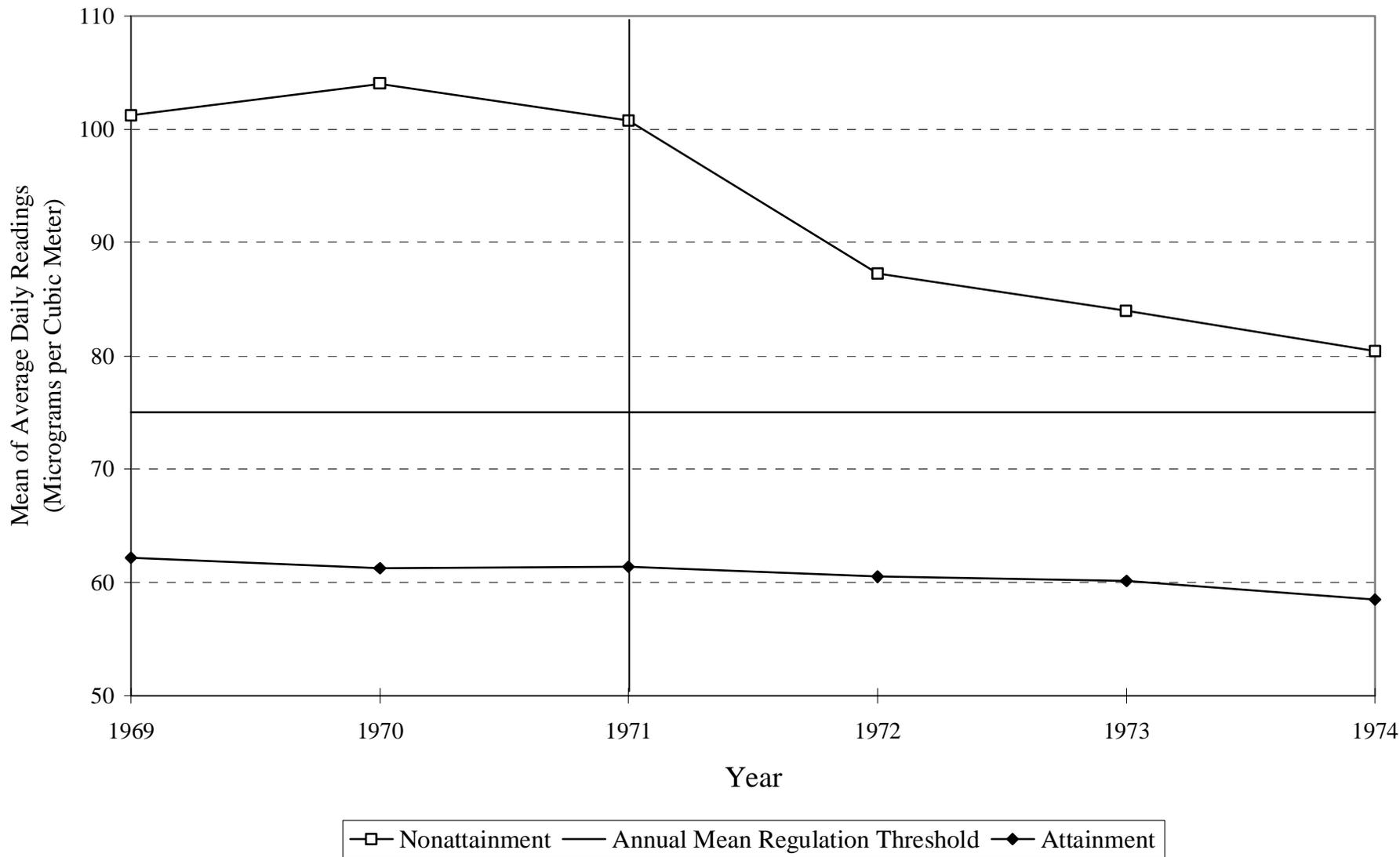
Chay and Greenstone (2003) use clean air act of 1970 to resolve the causality problem:

Areas with more particulates than threshold required to clean up air [treatment group]. Areas with less particulates than threshold are control group.

Compares infant mortality across 2 types of places before and after (DD approach)

Figure 2: Trends in TSPs Pollution and Infant Mortality, by 1972 Nonattainment Status

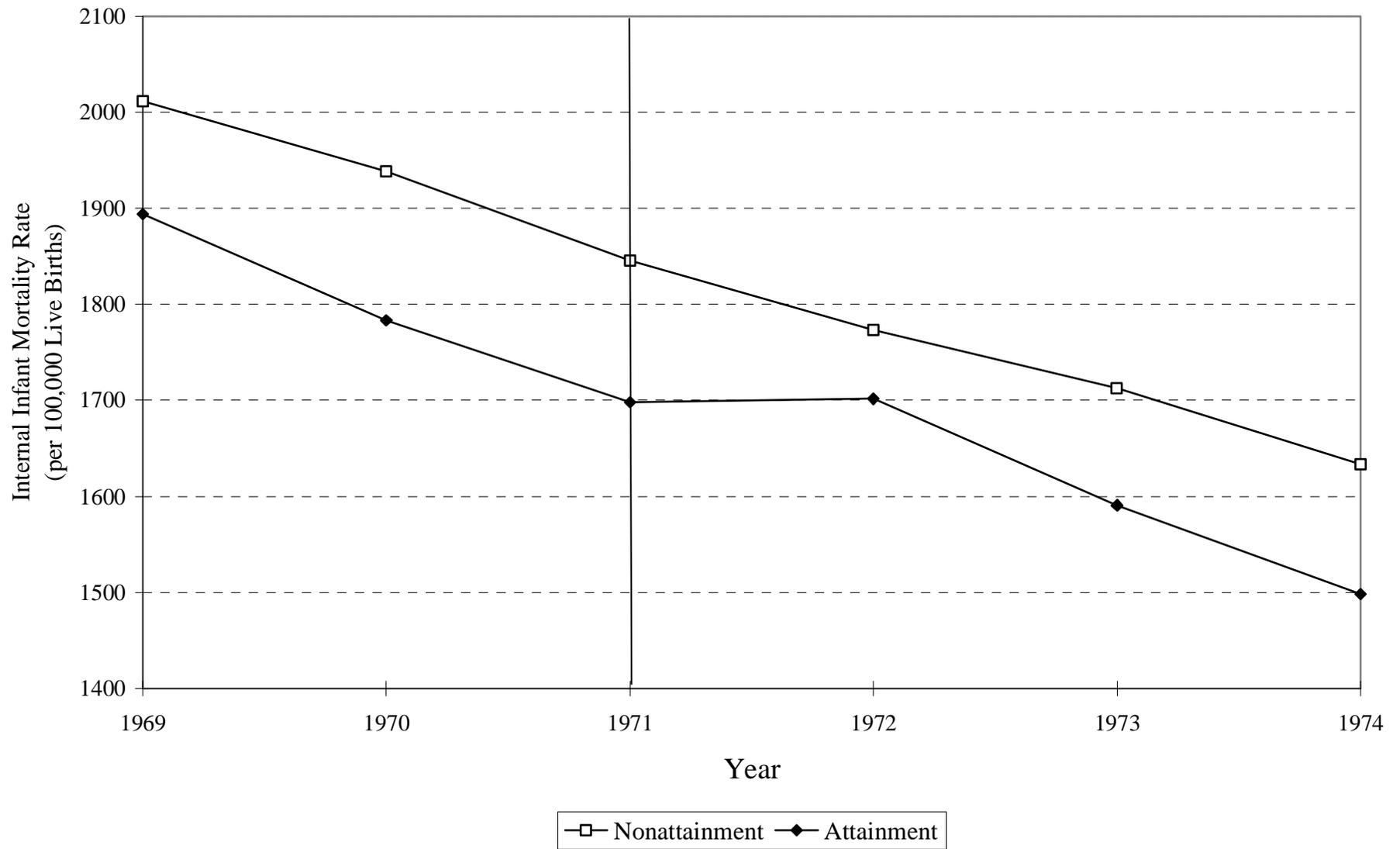
A. Trends in Mean TSPs Concentrations, by 1972 Nonattainment Status



Source: Authors' tabulations from EPA's "Quick Look Reports" data file.

Source: Chay and Greenstone (2003)

B. Trends in Internal Infant Mortality Rate, by 1972 Nonattainment Status



Source: Chay and Greenstone (2003)

Climate Change and CO2 Emissions

- 1) Industrialization has dramatically increased CO2 emissions and atmospheric CO2 generates global warming
- 2) Atmospheric CO2 has long life (35% remains after 100 years) [absent any carbon capture tech breakthrough]
- 3) Great uncertainty in costs of global warming [mitigation or amplifying feedback loops] and thus benefit of reducing CO2
- 4) Great uncertainty in costs of reducing CO2 [easy to innovate?]
- 5) Costs of global warming are decades/centuries away [how should this be discounted?]

How fast should we start reducing emissions? [Stern-Weitzman want a fast reduction, Nordhaus advocates a slower path]

Empirical Example: Costs of Global Warming

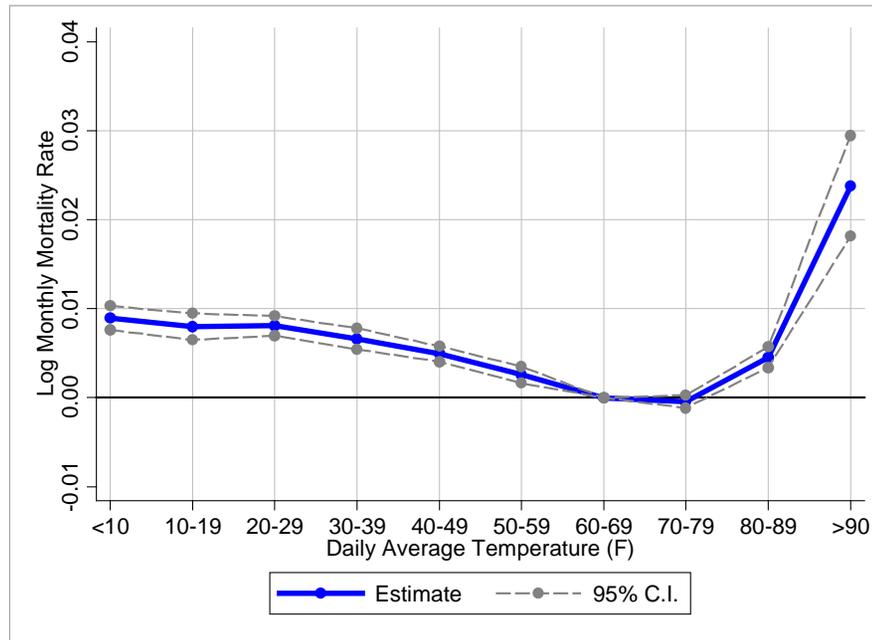
Will we really need to reduce global warming, since we may innovate adaptations as we go that will reduce its costs?

Example: heat waves and mortality analysis of Barreca-Clay-Deschenes-Greenstone-Shapiro

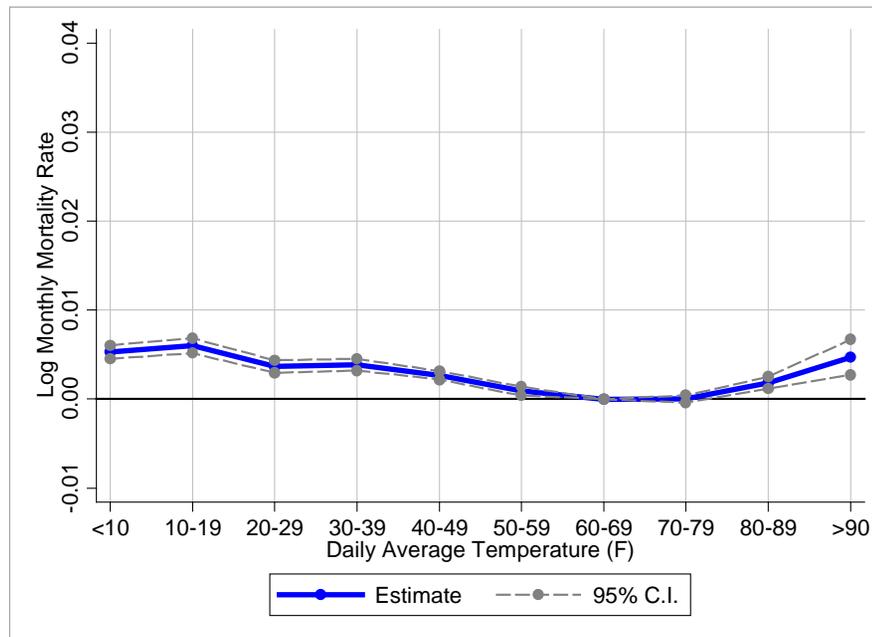
- 1) The mortality effect of an extremely hot day (90°F+) declined by about 80% between 1900-1959 and 1960-2004.
- 2) Adoption of residential air conditioning (AC) explains the entire decline
- 3) Worldwide adoption of AC will speed up the rate of climate change (if fossil fuel powered)

Figure 2: Estimated Temperature-Mortality Relationship (Continued)

(c) 1929-1959



(d) 1960-2004



Notes: Figure 2 plots the response function between log monthly mortality rate and average daily temperatures, obtained by fitting Equation (1). The response function is normalized with the 60°F – 69°F category set equal to zero so each estimate corresponds to the estimated impact of an additional day in bin j on the log monthly

CURBING GLOBAL WARMING: KYOTO TREATY NAD PARIS ACCORD

Kyoto 1997: 35 industrialized nations (but not US) agreed to reduce their emissions of greenhouse gases to 5% below (depends on country) 1990 levels by the year 2012

Industrialized countries are allowed to trade emissions rights among themselves, as long as the total emissions goals are met [=quantity regulation with trading permits]. Developing countries not in treaty.

In the US, Obama has directed EPA to regulate CO2 emission [carbon tax or cap-and-trade requires congress law, 27 states now suing]

Paris 2016: US China India and others agreed to set binding emissions targets to reduce a pre-set share of global emissions. May not be enforceable without Senate approval.

REFERENCES

Jonathan Gruber, *Public Finance and Public Policy*, Fourth Edition, 2012
Worth Publishers, Chapters 5 and 6

Barreca, Alan, et al. "Adapting to Climate Change: The Remarkable Decline in the US Temperature-Mortality Relationship over the 20th Century." No. w18692. National Bureau of Economic Research, 2013.(web)

Chay, K. and M. Greenstone "Air Quality, Infant Mortality, and the Clean Air Act of 1970," NBER Working Paper No. 10053, 2003.(web)

Ellerman, A. Denny, ed. "Markets for clean air: The US acid rain program." Cambridge University Press, 2000.(web)

Gruber, Jonathan. "Tobacco at the crossroads: the past and future of smoking regulation in the United States." *The Journal of Economic Perspectives* 15.2 (2001): 193-212.(web)

Nordhaus, William D., and Joseph Boyer. "Warning the World: Economic Models of Global Warming." MIT Press (MA), 2000.(web)

Nordhaus, William D. "After Kyoto: Alternative mechanisms to control global warming." *The American Economic Review* 96.2 (2006): 31-34.(web)

Nordhaus, William D. *The Climate Casino: Risk, Uncertainty, and Economics for a Warming World*, Yale University Press, 2013.

Undergraduate Public Economics

Danny Yagan
UC Berkeley

Lecture 12 Public Goods

PUBLIC GOODS: DEFINITIONS

Pure public goods: Goods that are perfectly **non-rival in consumption** and are **non-excludable**

Non-rival in consumption: One individual's consumption of a good does not affect another's opportunity to consume the good.

Non-excludable: Individuals cannot deny each other the opportunity to consume a good.

Impure public goods: Goods that satisfy the two public good conditions (non-rival in consumption and non-excludable) to some extent, but not fully.

7.1

Defining Pure and Impure Public Goods

		Is the good rival in consumption?	
		Yes	No
Is the good excludable?	Yes	Private good (ice cream)	Impure public good (Cable TV)
	No	Impure public good (crowded sidewalk)	Public good (defense)

OPTIMAL PROVISION OF PRIVATE GOODS

Two goods: ic (ice-cream) and c (cookies) with prices P_{ic}, P_c

$P_c = 1$ is normalized to one (numéraire good):

Two individuals B and J demand different quantities of the good at the same market price.

$MRS_{ic,c} = MU_{ic}/MU_c = \#$ cookies the consumer is willing to give up for 1 ice-cream

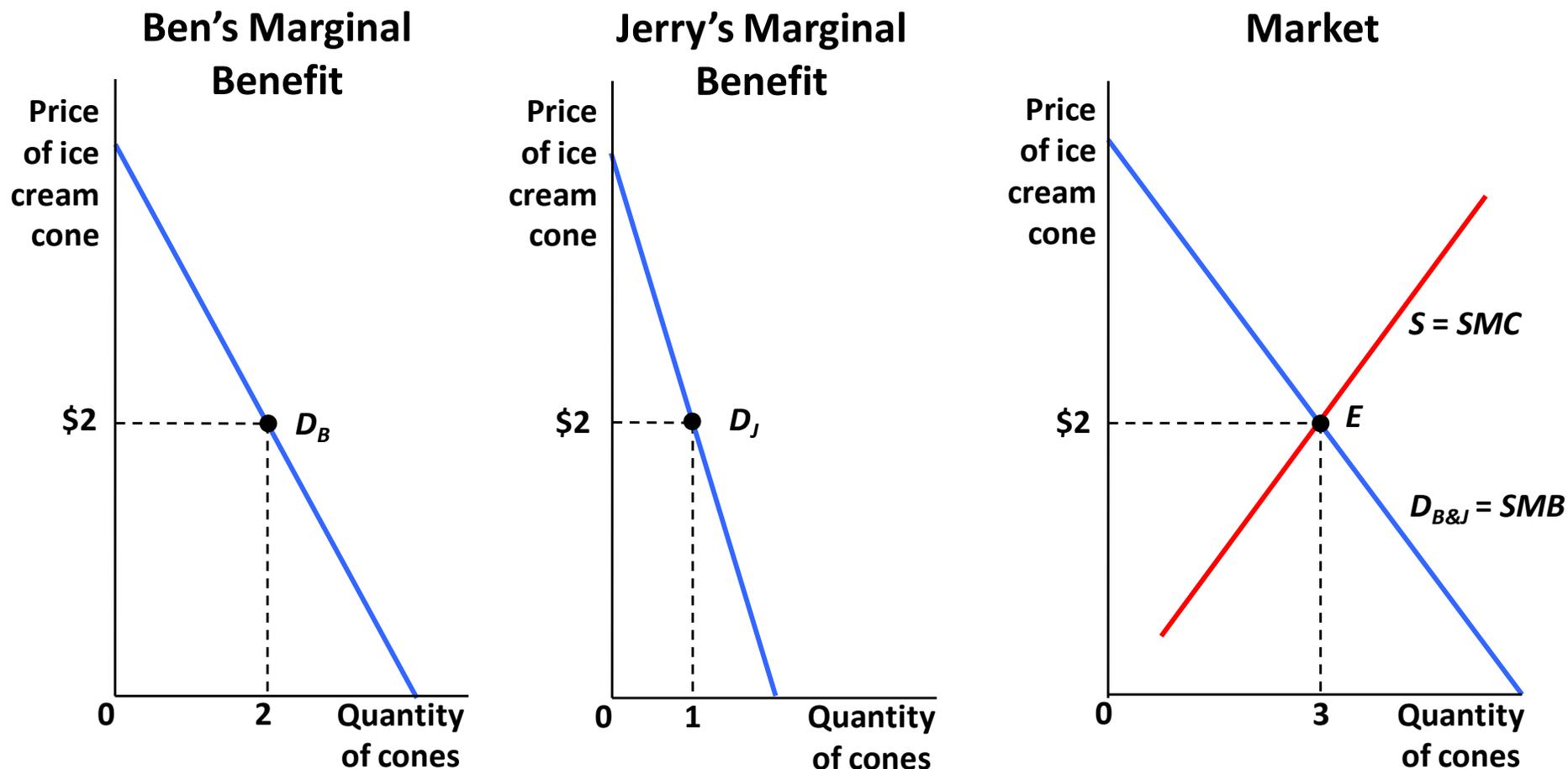
The optimality condition for the consumption of private goods is written as: $MRS_{ic,c}^B = MRS_{ic,c}^J = P_{ic}/P_c = P_{ic}$

Equilibrium on the supply side requires: $MC_{ic} = P_{ic}$

In equilibrium, therefore: $MRS_{ic,c}^B = MRS_{ic,c}^J = MC$

7.1

Horizontal Summation in the Private Goods Market



- To find social demand curve, add quantity at each price—sum horizontally.

OPTIMAL PROVISION OF PUBLIC GOODS

Replace private good ice-cream ic by a public good missiles m

$MRS_{m,c}^B = \#$ cookies B is willing to give up for 1 missile

$MRS_{m,c}^J = \#$ cookies J is willing to give up for 1 missile

In net, society is willing to give up $MRS_{m,c}^B + MRS_{m,c}^J$ cookies for 1 missile

Social-efficiency-maximizing condition for the public good is:

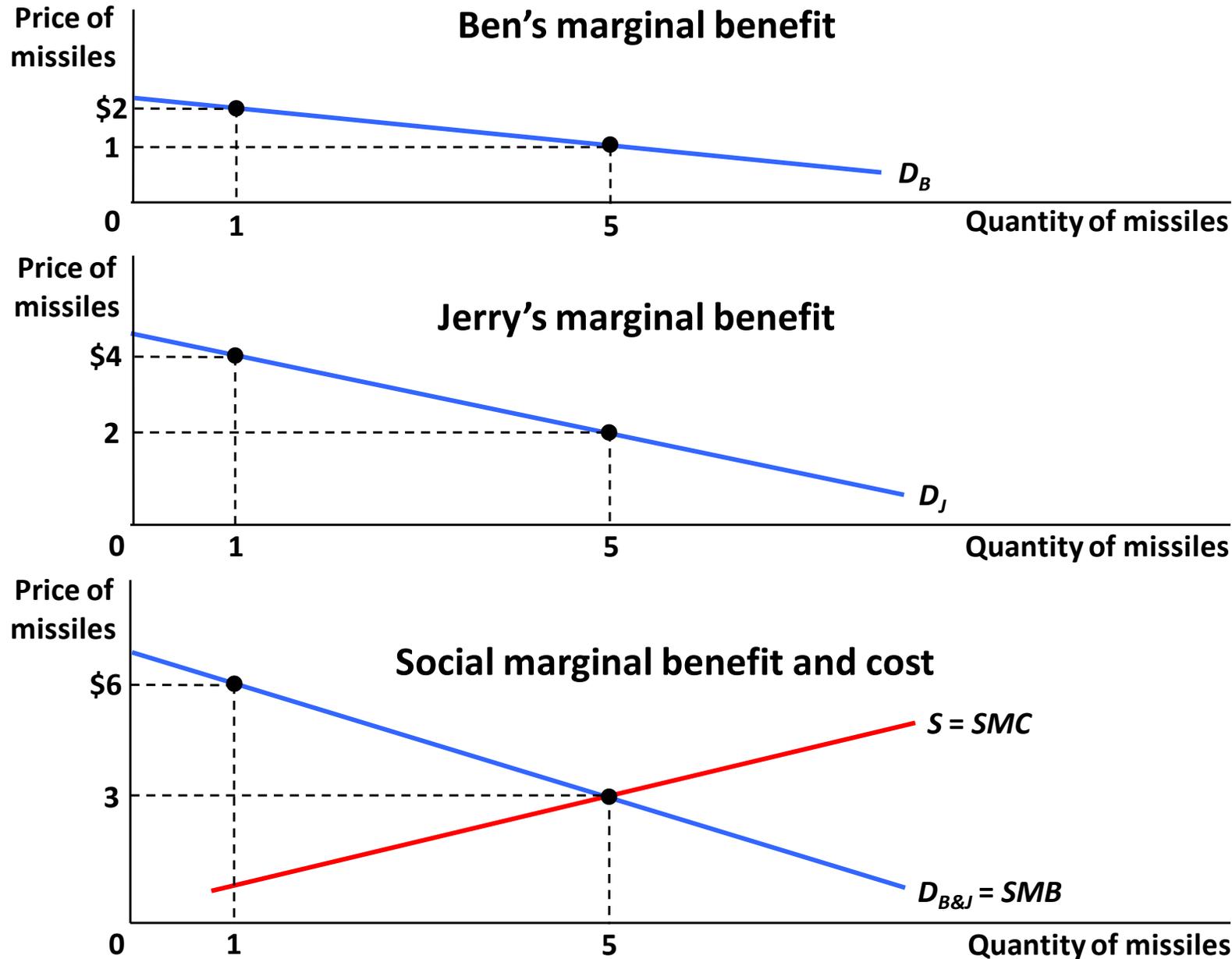
$$MRS_{m,c}^B + MRS_{m,c}^J = MC$$

Social efficiency is maximized when the marginal cost is set equal to the *sum of the MRSs, rather than being set equal to each individual MRS.*

This is called the **Samuelson rule** (Samuelson, 1954)

7.1

Vertical Summation in the Public Goods Market



PRIVATE-SECTOR UNDERPROVISION

Private sector provision such that $MRS_{mc} = MC_m$ for each individual so that $\sum MRS_{mc} > MC_m \Rightarrow$ Outcome is not efficient, could improve the welfare of everybody by having more missiles (and less cookies)

Free rider problem: When an investment has a personal cost but a common benefit, individuals will underinvest.

Because of the **free rider** problem, the private market under-supplies public goods

Another way to see it: private provision of a public good creates a positive externality (as everybody else benefits) \Rightarrow Goods with positive externalities are under-supplied by the market

PRIVATE PROVISION OF PUBLIC GOOD

2 individuals with identical utility functions defined on X private good (cookies) and F public good (fireworks)

$F = F_1 + F_2$ where F_i is contribution of individual i

Utility of individual i is $U_i = 2 \log(X_i) + \log(F_1 + F_2)$ with budget $X_i + F_i = 100$

Individual 1 chooses F_1 to maximize $2 \log(100 - F_1) + \log(F_1 + F_2)$ taking F_2 as given

First order condition: $-2/(100 - F_1) + 1/(F_1 + F_2) = 0 \Rightarrow F_1 = (100 - 2F_2)/3$

Note that F_1 goes down with F_2 due to the free rider problem (called the reaction curve, show graph)

Symmetrically, we have $F_2 = (100 - 2F_1)/3$

PRIVATE PROVISION OF PUBLIC GOOD

Nash equilibrium definition: Each agent maximizes his objective taking as given the actions of the other agents

At the Nash equilibrium, the two reaction curves intersect:

$$F_1 = (100 - 2F_2)/3 \text{ and } F_2 = (100 - 2F_1)/3$$

$$\Rightarrow F_1 + F_2 = (200 - 2(F_1 + F_2))/3 \Rightarrow F = F_1 + F_2 = 200/5 = 40 \Rightarrow F_1 = F_2 = 20$$

What is the Social Optimum? $\sum MRS = MC = 1$

$$MRS_{FX}^i = MU_F^i / MU_X^i = (1/(F_1 + F_2)) / (2/X_i) = X_i / (2F)$$

$$\Rightarrow \sum MRS^i = (X_1 + X_2) / (2F) = (200 - F) / (2F)$$

$$\Rightarrow \sum MRS^i = 1 \Rightarrow 200 - F = 2F \Rightarrow F = 200/3 = 66.6 > 40$$

Public good is under-provided by the market

Can Private Provision Overcome Free Rider Problem?

The free rider problem does not lead to a complete absence of private provision of public goods. Private provision works better when:

1) Some Individuals Care More than Others:

Private provision is particularly likely to surmount the free rider problem when individuals are not identical, and when some individuals have an especially high demand for the public good.

2) Altruism:

When individuals value the benefits and costs to others in making their consumption choices.

3) Warm Glow:

Model of public goods provision in which individuals care about both the total amount of the public good and their particular contributions as well.

Experimental evidence on free riding

Laboratory experiments are a great device to test economic theories

Subjects (often students) are brought to the lab where they sit through a computer team game and get paid based on the game outcomes

Many public good lab experiments. Example (Marwell and Ames 1981):

- 10 repetitions for each game
- In each game, group of 5 people, each with 10 tokens to allocate between cash and public good.
- If take token in cash, get \$1 in cash for yourself. If contribute to common good, get \$.5 to each of all five players.

Nash equilibrium: get everything in cash

Socially optimal equilibrium: contribute everything to public good

In the lab, subjects contribute about 50% to public good, but public good contributions fall as game is repeated (Isaac, McCue, and Plott, 1985)

Explanations: people are willing to cooperate at first but get upset and retaliate if others take advantage of them

Crowding out of private contributions by govt provision

Suppose government forces each individual to provide 5 so that now $F = F_1 + F_2 + 10$ where F_i is voluntary contribution of individual i

Utility of individual i is $U_i = 2 \log(X_i) + \log(F_1 + F_2 + 10)$ with budget $X_i + F_i = 95$

You will find that the private optimum is such that $F_1 = F_2 = 15$ so that government forced contribution crowds out one-to-one private contributions

Why? Rename $F'_i = F_i + 5$. Choosing F'_i is equivalent to choosing F_i : $U_i = 2 \log(X_i) + \log(F'_1 + F'_2)$ with budget $X_i + F'_i = 100$

⇒ Equivalent to our initial problem with no government provision hence the solution in F'_i must be the same

However, government forced contributions will have an effect as soon as private contributions fall to zero (as individuals cannot contribute negative amounts and undo government provision)

EMPIRICAL EVIDENCE ON CROWD-OUT

Two strands of empirical literature

- 1) Field evidence (observational studies)
- 2) Lab experiments (Andreoni, AER'93)

Traditionally, lab experiments have been more influential but recent field studies may change this

Lab experiments show imperfect crowd-out in public good games (where you compare situation with no forced public goods contributions and with forced public good contributions)

Lab experiment may not capture important motives for giving: warm glow, prestige, solicitations from fund raisers

CHARITABLE GIVING

Charitable giving is one form of private provision of public good (big in the US, 2% of National Income given to charities).

Funds (1) religious activities, (2) education, (3) human services, (4) health, (5) arts, (6) various causes (environment, animal protection, etc.)

Encouraged by government: giving can be deducted from income for income tax purposes

People give out of (1) warm-glow (religion, name on building), (2) reciprocity (alumni), (3) social pressure (religion), (4) altruism (poverty relief)

Those effects are not captured in basic economic model

Charities have big fund-raising operations to induce people to give based on those psychological effects

Empirical Evidence on Crowd-Out: Hungerman 2005

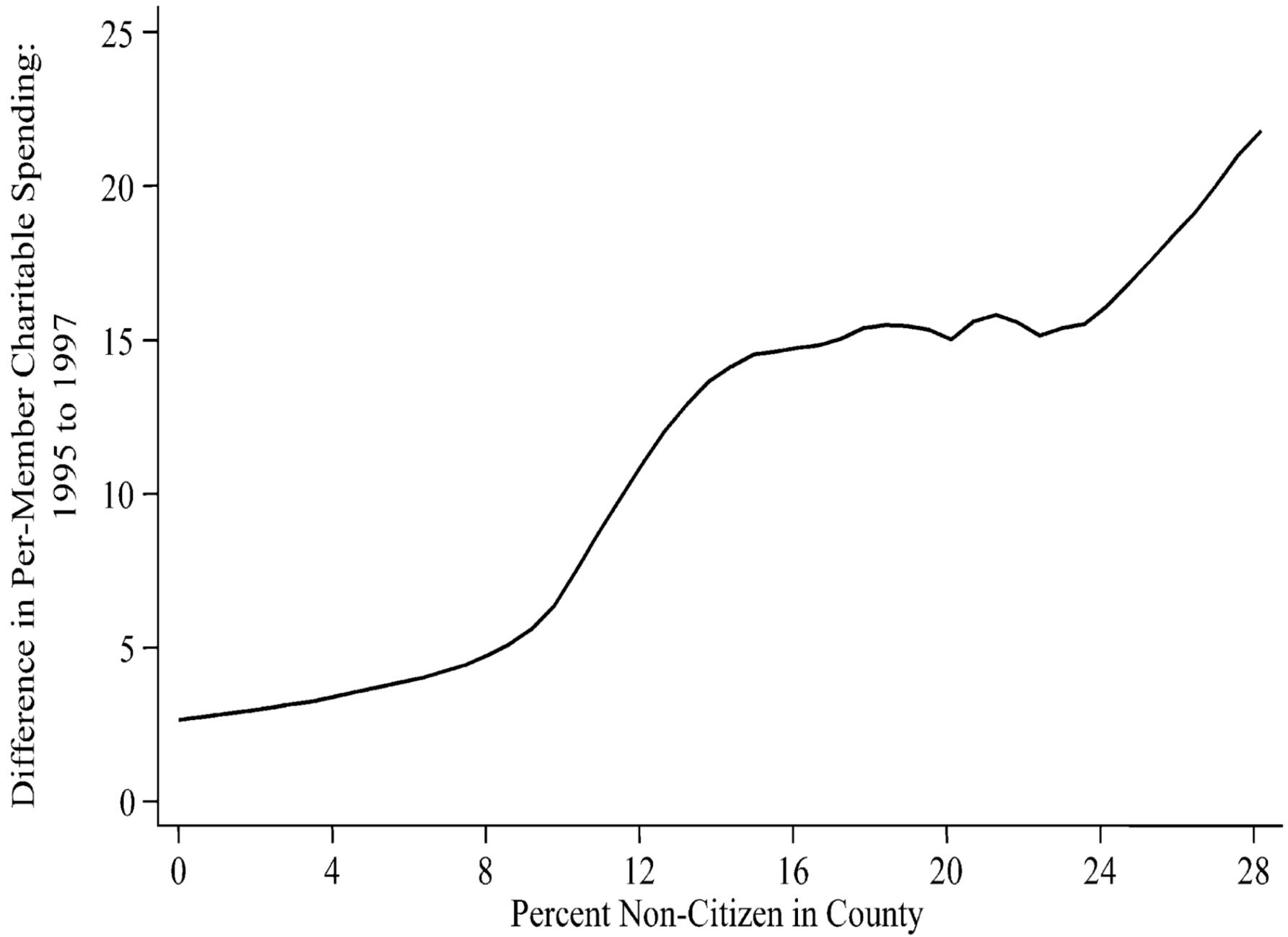
Studies crowdout of church-provided welfare (soup kitchens, etc.) by government welfare.

Uses 1996 Clinton welfare reform act as an instrument for welfare spending cuts.

One aspect of reform: reduced/eliminated welfare for non-citizens

Motivates a diff-in-diff strategy: compare churches in high non-citizen areas with churches in low non-citizen areas before/after 1996 reform

Estimates imply that total church expenditures in a state increase by 40 cents when welfare spending is cut by \$1



Source: Hungerman 2005

Empirical Evidence on Crowd-Out: Andreoni-Payne '03

Government spending crowds out private donations through two channels: willingness to donate + fundraising

Use tax return data on arts and social service organizations

Panel study: follows the same organizations overtime

Results: \$1000 increase in government grant leads to \$250 reduction in private fundraising

Suggests that crowdoout could be non-trivial if fundraising is a powerful source of generating private contributions

Subsequent study by Andreoni and Payne confirms this

Find that \$1 more of government grant to a charity leads to 56 cents less private contributions

70 percent (\$0.40) due to the fundraising channel

Suggests that individuals are relatively passive actors

Randomized field experiment to test reciprocity

Falk (2007) conducted a field experiment to investigate the relevance of reciprocity in charitable giving

In collaboration with a charitable organization, sent 10,000 Christmas solicitation letters for funding schools for street children in Bangladesh to potential donors (in Switzerland) randomized into 3 groups

- 1) 1/3 of letters contained no gift (control group)
- 2) 1/3 contained a small gift: one post-card (children drawings)+one-envelope (small-gift treatment)
- 3) 1/3 contained a larger gift: 4 post-cards (children drawings)+4-envelopes (large-gift treatment)

Likelihood of giving: 12% in control, 14% in small-gift treatment, 21% in large-gift treatment

“large gift” was very effective (even relative to cost)

Empirical Evidence on Social Pressure

Dellavigna-List-Malmendier '12 design a door-to-door fundraiser randomized experiment:

Control: no advance warning of fund-raiser visit

Treatment group 1: flyer at doorknob informs about the exact time of solicitation (hence can seek/avoid fund-raiser)

Treatment group 2: same as treatment 1 but flyer has a check box "Do not disturb"

Results (relative to control):

Treatment group 1: 9-25% less likely to open door for fund-raiser, same (unconditional giving)

Treatment group 2: a number of people opt out and (unconditional giving) is 28-42% lower

⇒ Social pressure is an important determinant of door-to-door giving and door-to-door fund-raising campaigns lower utility of potential donors

Social Prices as a Policy Instrument

Traditional focus in economics is on changing prices of economic goods

Different set of policy instruments: “social prices”

Suppose people care about social norms and policy maker can manipulate social norms

Should make status good one that generates positive externalities.

E.g. large SUVs are frowned upon as gas guzzlers contributing to global warming while electric cars are admired

Creates another set of policy instruments to explore

Recent examples from psychology and political science suggest that social price elasticities can be large

Example: Gerber, Green, Larimer '08: randomized experiment using social pressure via letters to increase voter turnout

Civic duty mailing

Dear Registered Voter:

DO YOUR CIVIC DUTY AND VOTE!

Why do so many people fail to vote? We've been talking about this problem for years, but it only seems to get worse.

The whole point of democracy is that citizens are active participants in government; that we have a voice in government. Your voice starts with your vote. On August 8, remember your rights and responsibilities as a citizen. Remember to vote.

DO YOUR CIVIC DUTY – VOTE!

Hawthorne mailing

Dear Registered Voter:

YOU ARE BEING STUDIED!

Why do so many people fail to vote? We've been talking about this problem for years, but it only seems to get worse.

This year, we're trying to figure out why people do or do not vote. We'll be studying voter turnout in the August 8 primary election.

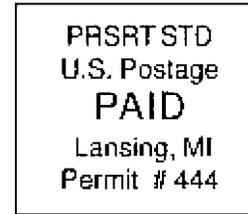
Our analysis will be based on public records, so you will not be contacted again or disturbed in anyway. Anything we learn about your voting or not voting will remain confidential and will not be disclosed to anyone else.

DO YOUR CIVIC DUTY – VOTE!

Self mailing

3 0 4 2 2 - 4 ||| ||| ||| |||

For more information: (517) 351-1975
email: etov@grebner.com
Practical Political Consulting
P. O. Box 6249
East Lansing, MI 48826



ECRLOT **C050
THE WAYNE FAMILY
9999 OAK ST
FLINT MI 48507

Dear Registered Voter:

WHO VOTES IS PUBLIC INFORMATION!

Why do so many people fail to vote? We've been talking about the problem for years, but it only seems to get worse.

This year, we're taking a different approach. We are reminding people that who votes is a matter of public record.

The chart shows your name from the list of registered voters, showing past votes, as well as an empty box which we will fill in to show whether you vote in the August 8 primary election. We intend to mail you an updated chart when we have that information.

We will leave the box blank if you do not vote.

DO YOUR CIVIC DUTY—VOTE!

OAK ST	Aug 04	Nov 04	Aug 06
9999 ROBERT WAYNE		Voted	_____
9999 LAURA WAYNE	Voted	Voted	_____

Neighbors mailing

Dear Registered Voter:

WHAT IF YOUR NEIGHBORS KNEW WHETHER YOU VOTED?

Why do so many people fail to vote? We've been talking about this problem for years, but it only seems to get worse. This year, we're taking a new approach. We're sending this mailing to you and your neighbors to publicize who does and does not vote.

The chart shows the names of some of your neighbors, showing which have votes in the past. After the August 8 election, we intend to mail an updated chart. You and your neighbors will all know who voted and who did not

DO YOUR CIVIC DUTY – VOTE!

MAPLE DR	Aug 04	Nov 04	Aug 06
9995 JOSEPH JAMES SMITH	VOTED	VOTED	_____
9995 JENNIFER KAY SMITH	VOTED		_____
9997 RICHARD B JACKSON	VOTED		_____
9999 KATHY MARIE JACKSON		VOTED	_____
9987 MARIA S. JOHNSON	VOTED	VOTED	_____
9987 TOM JACK JOHNSON	VOTED	VOTED	_____

Source: Gerber, Green, and Larimer (2008)

TABLE 2. Effects of Four Mail Treatments on Voter Turnout in the August 2006 Primary Election

	Experimental Group				
	Control	Civic Duty	Hawthorne	Self	Neighbors
Percentage Voting	29.7%	31.5%	32.2%	34.5%	37.8%
N of Individuals	191,243	38,218	38,204	38,218	38,201

Source: Gerber, Green, and Larimer (2008)

Welfare Analysis of Social Pricing

Should social pricing be used on top of standard pricing through corrective taxes (or tradable permits)?

1) Making people feel bad about driving an SUV is inefficient relative to gas tax: destroys welfare without bringing tax revenue

Could still be desirable if imposing a gas tax is impossible. Some negative actions (such as littering) are hard to enforce with fines so social norm on feeling bad about littering is desirable.

2) Making people feel good about driving an energy efficient car is efficient relative to gas tax: adds to welfare as driving an energy efficient car becomes more enjoyable

REFERENCES

Jonathan Gruber, Public Finance and Public Policy, Fourth Edition, 2012
Worth Publishers, Chapter 7

Andreoni, James. "An experimental test of the public-goods crowding-out hypothesis." The American Economic Review (1993): 1317-1327.(web)

Andreoni, James, and A. Abigail Payne. "Do government grants to private charities crowd out giving or fund-raising?." American Economic Review (2003): 792-812.(web)

Dellavigna, Stefano, John A. List and Ulrike Malmendier, "Testing for Altruism and Social Pressure in Charitable Giving," Quarterly Journal of Economics, 2012, 127(1), 1-56. (web)

Falk, Armin. "Gift exchange in the field." Econometrica 75.5 (2007): 1501-1511.(web)

Gerber, Alan S., Donald P. Green, and Christopher W. Larimer. "Social pressure and vote turnout: Evidence from a large-scale field experiment." American Political Science Review 102.1 (2008): 33. (web)

Hungerman, Daniel M. "Are church and state substitutes? Evidence from the 1996 welfare reform." Journal of Public Economics 89.11 (2005): 2245-2267.(web)

Isaac, Mark R., Kenneth F. McCue, and Charles R. Plott. "Public goods provision in an experimental environment." *Journal of Public Economics* 26.1 (1985): 51-74.(web)

Marwell, Gerald, and Ruth E. Ames. "Economists free ride, does anyone else?: Experiments on the provision of public goods." *Journal of Public Economics* 15.3 (1981): 295-310. (web)

Undergraduate Public Economics

Danny Yagan
UC Berkeley

Lecture 13 Political Economy

Political Economy

Political Economy is the positive analysis of government: why do governments do what they do?

In democracies, citizens vote to elect politicians to run the government

In principle, government decisions should reflect the will of citizens

Even non-democratic rulers are in part subject to people's preferences

MAJORITY VOTING: WHEN IT WORKS

Majority voting: Mechanism used to aggregate individual votes into a social decision: individual policy options are put to a vote and the option that receives the majority of votes is chosen

Majority voting can produce a consistent aggregation of individual preferences only if preferences are restricted to take a certain form

Example: funding for local public schools using property taxes could be chosen as high (H), medium (M), or low (L)

9.2

Majority Voting: When It Works

- There are three types of voters in a town: *parents*, *elders*, and *young couples without children*.
- They have different preferences over the level of school spending (high, medium, or low).

	Parents (33.3%)	Elders (33.3%)	Young Couples (33.3%)
First choice	<i>H</i>	<i>L</i>	<i>M</i>
Second choice	<i>M</i>	<i>M</i>	<i>L</i>
Third choice	<i>L</i>	<i>H</i>	<i>H</i>

MAJORITY VOTING: WHEN IT WORKS

The town could proceed as follows:

- Vote on funding level H versus funding level L: L wins H
- Vote on funding level H versus funding level M: M wins H
- Vote on funding level L versus funding level M: M wins L

M has beaten both H and L so M is the overall winner.

Majority voting has aggregated individual preferences to produce a preferred social outcome: medium school spending and taxes.

9.2

Majority Voting: When It Doesn't Work

- **Cycling:** When majority voting does not deliver a consistent aggregation of individual preferences.

	Public school parents (33.3%)	Private school parents (33.3%)	Young Couples (33.3%)
First choice	<i>H</i>	<i>L</i>	<i>M</i>
Second choice	<i>M</i>	<i>H</i>	<i>L</i>
Third choice	<i>L</i>	<i>M</i>	<i>H</i>

MAJORITY VOTING: WHEN IT DOES NOT WORK

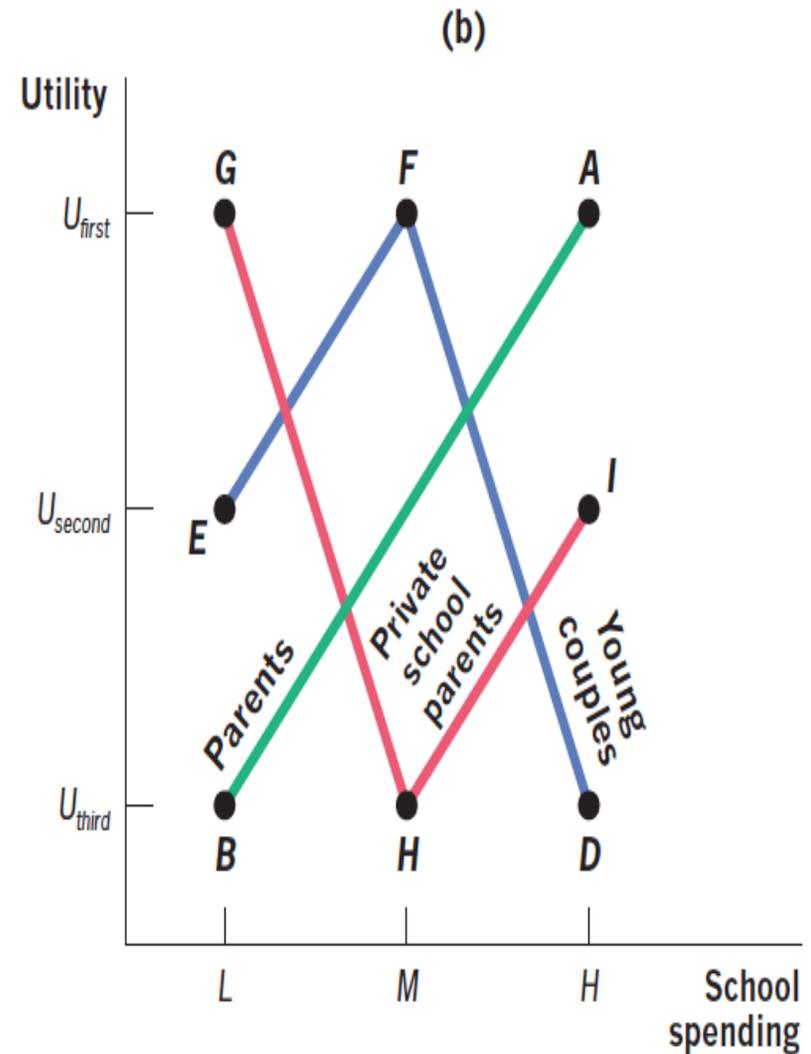
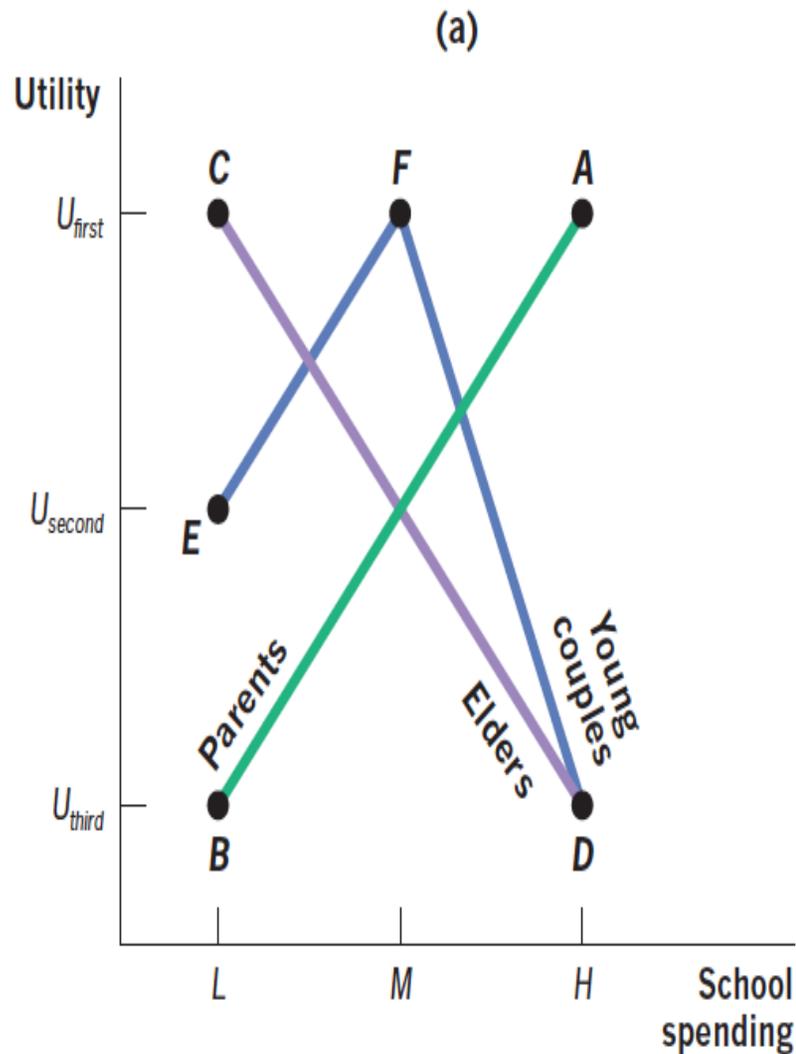
- Vote on funding level H versus funding level L: L wins H
- Vote on funding level H versus funding level M: H wins M
- Vote on funding level L versus funding level M: M wins L

Cycle with no clear winner.

Majority voting is unable to aggregate preferences in a meaningful way in that case

9.2

Single-Peaked versus Non-Single-Peaked Preferences



MEDIAN VOTER THEOREM

Consider choice along a single dimension (e.g., level of funding)

Single peaked preferences: The preferences for funding increase and then decrease (always increasing, or always decreasing also fine). Peak is preferred funding level for the individual.

Median voter is the voter whose peak is at the median (half have lower peaks, half have higher peaks)

Voting Equilibrium (or Condorcet winner) is an outcome that wins in majority voting against any other alternative

Median Voter Theorem: Peak of median voter is a voting equilibrium

PROOF OF MEDIAN VOTER THEOREM

Let $a_1 < \dots < a_{median} < \dots < a_I$ be the peaks of individuals $1, \dots, I$

Suppose vote between a_{median} and a with $a_{median} < a^*$

a_{median} wins because $i = 1, \dots, median$ all prefer a_{median} to a^* (because they all have decreasing preferences for a beyond a_{median})

Symmetrically a_{median} wins against $a^* < a_{median}$ because $i = median, \dots, I$ prefer a_{median} to a^*

Median voter outcome from majority voting is very useful and a hugely influential result in the political economy literature

ABSTRACT SOCIAL CHOICE PROBLEM

$n = 1, \dots, N$ possible choices society can make

$i = 1, \dots, I$ individuals have preferences \prec_i over the N choices

Social decision rule: It aggregates individuals preferences $(\prec_i)_{i=1, \dots, I}$ into a social preference \prec_S over N choices that satisfies 3 key properties:

1) Pareto Dominance: if $a \prec_i b$ for all i then $a \prec_S b$

2) Transitivity: if $a \prec_S b$ and $b \prec_S c$ then $a \prec_S c$

3) Independence of irrelevant alternatives: whether $a \prec_S b$ or $a \succ_S b$ depends only on how individuals rank a vs. b (and not any other alternative).

Importantly, 3) rules out “intensity of preferences effects” to focus

ABSTRACT SOCIAL CHOICE PROBLEM

ARROW'S IMPOSSIBILITY THEOREM: There is no social decision rule that converts individual preferences into a consistent aggregate decision without either

(a) restricting preferences or

(b) imposing a dictatorship (i.e. $\prec_S = \prec_i$ for some i)

Geanakoplos (2005) provides simple proofs

This result was very influential and shows that the abstract social choice problem cannot have a general solution

Most common solutions are to:

(1) restrict preferences to single peaked preferences (median voter theorem)

(2) let intensity of preferences play a role (social welfare function and Samuelson rule for efficiency)

MEDIAN VOTER AND EFFICIENCY

Efficiency requires

\sum social marginal benefits = social marginal costs

\Rightarrow Public good is worth providing if \sum benefits $>$ costs

What matters for efficiency is the **average** marginal benefit across individuals and not the **median** marginal benefit

\Rightarrow Median outcome is not efficient unless Median = Average (not true in general)

Example: bridge project would serve 10 people. 6 people value bridge at \$50, 4 people value bridge at \$100. Total social value of bridge is \$700 = $6 \cdot 50 + 4 \cdot 100$

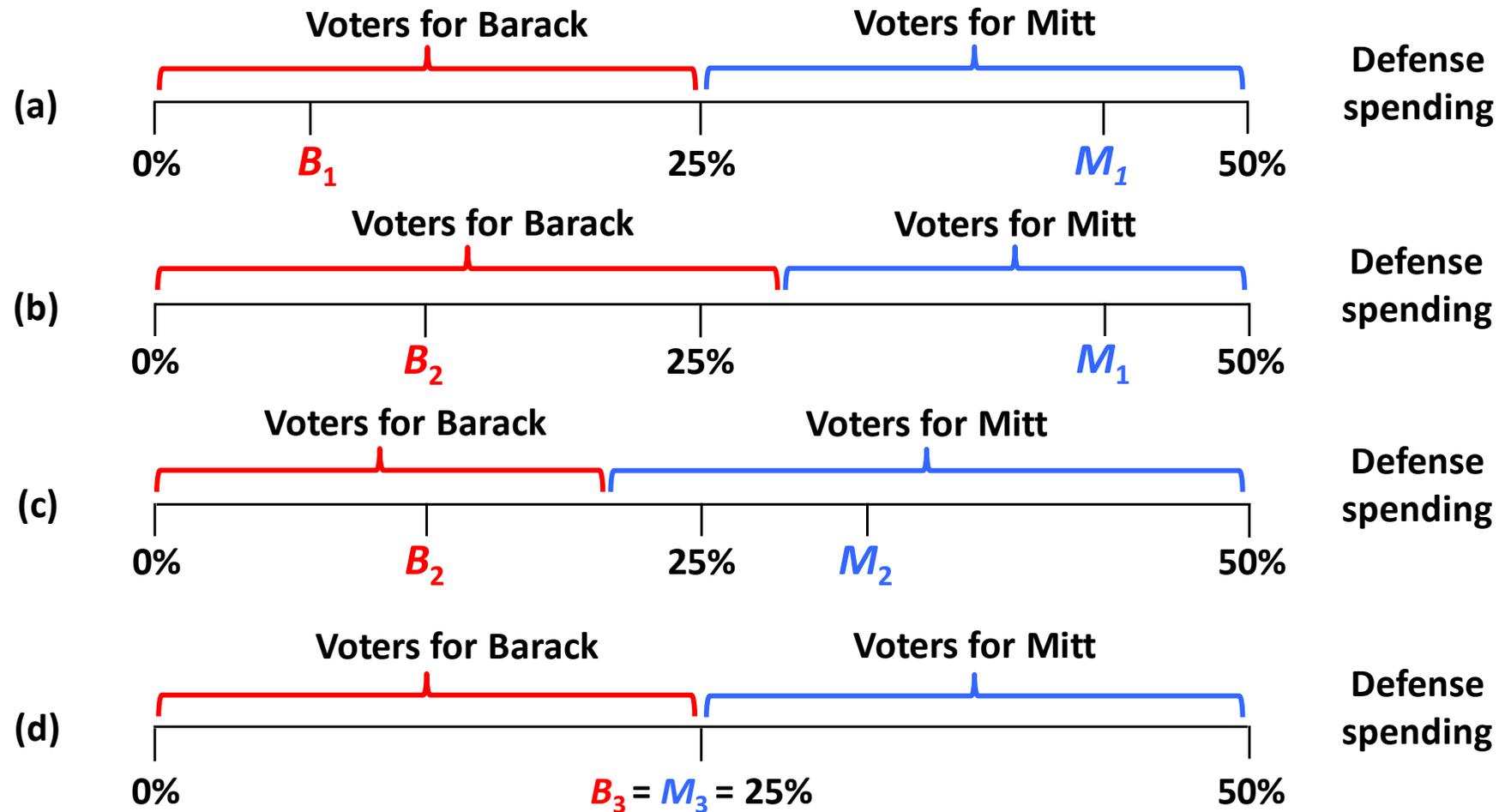
Suppose cost is \$60 per person so total cost = \$600 = $\$60 \cdot 10$.

Mean net benefit is $70 - 60 = \$10$, median net benefit is $50 - 60 = -\$10$

Project is socially desirable but is opposed by 6 people to 4 in majority voting \Rightarrow Median voter leads to an inefficient outcome

9.3

Vote-Maximizing Politicians Represent the Median Voter



ASSUMPTIONS OF THE MEDIAN VOTER MODEL

Although the median voter model is a convenient way to predict outcomes of representative democracy, it does so by making a number of assumptions.

1) Single-dimensional Voting

The median voter model assumes that voters are basing their votes on a single issue.

In reality, representatives are elected not based on a single issue but on a bundle of issues.

Individuals may lie at different points of the voting spectrum on different issues, so appealing to one end of the spectrum or another on some issues may be vote-maximizing.

ASSUMPTIONS OF THE MEDIAN VOTER MODEL

2) Only Two Candidates

The median voter model assumes that there are only two candidates for office.

If there are more than two candidates, the simple predictions of the median voter model break down.

Indeed, there is no stable equilibrium in the model with three or more candidates because there is always an incentive to move in response to your opponents' positions.

In many nations, the possibility of three or more valid candidates for office is a real one.

ASSUMPTIONS OF THE MEDIAN VOTER MODEL

3) No Ideology or Influence

The median voter theory assumes that politicians care only about maximizing votes.

Ideological convictions could lead politicians to position themselves away from the center of the spectrum and the median voter.

4) No Selective Voting

The median voter theory assumes that all people affected by public goods vote, but in fact, only a fraction of citizens vote in the United States. Appealing to the base (by moving away from median voter) is a way to increase turnout.

ASSUMPTIONS OF THE MEDIAN VOTER MODEL

5) No Money

The median voter theory ignores the role of money as a tool of influence in elections.

If taking an extreme position on a given topic maximizes fundraising, even if it does not directly maximize votes on that topic, it may serve the long-run interests of overall vote maximization by allowing the candidate to advertise more strongly.

6) Full Information

The median voter model assumes perfect information along three dimensions: voter knowledge of the issues; politician knowledge of the issues; and politician knowledge of voter preferences.

All three of these assumptions are unrealistic.

LOBBYING

Lobbying: The expending of resources by certain individuals or groups in an attempt to influence a politician

In principle, lobbying could correct inefficiencies due to median voter theorem: those who really want the bridge pay politicians who can provide transfers to those who don't want the bridge as much and get it built

However, lobbying can also lead to inefficiencies if public does not have perfect information and hence does not care to pay attention

Example: 5 people value bridge net of cost at \$100, 100 people value bridge net of cost at -\$6. Median voter does not produce the bridge (the socially desirable outcome)

However, 5 people have strong incentives to lobby and may get the project approved (if the 100 do not pay attention)

TESTING THE MEDIAN VOTER MODEL

Evidence from US congress representatives:

1) Senate: 2 senators for each state in US senate: represent the same constituency and hence should vote in the same way in the senate if median voter model is right (Poole and Rosenthal, '96)

Yet, in the US, when a state has 1 republican senator and 1 democratic senator, those 2 senators vote very differently in the senate (contradicts the median voter model)

TESTING THE MEDIAN VOTER MODEL

2) House of Representatives: Using close elections for US representatives (Lee, Moretti, Butler QJE'04):

When a candidate crosses 50%, he/she gets elected. However, the constituency is virtually the same whether a candidate gets 49.9% or 50.1% of the vote.

Therefore, median voter implies that a Democratic representative elected with 50.1% should vote similarly in congress to a Republican representative elected with 50.1% of the votes.

Yet, in reality, closely elected representatives vote very differently (measured by Americans for Democratic Action ADA scores) if they are Democratic vs. Republican

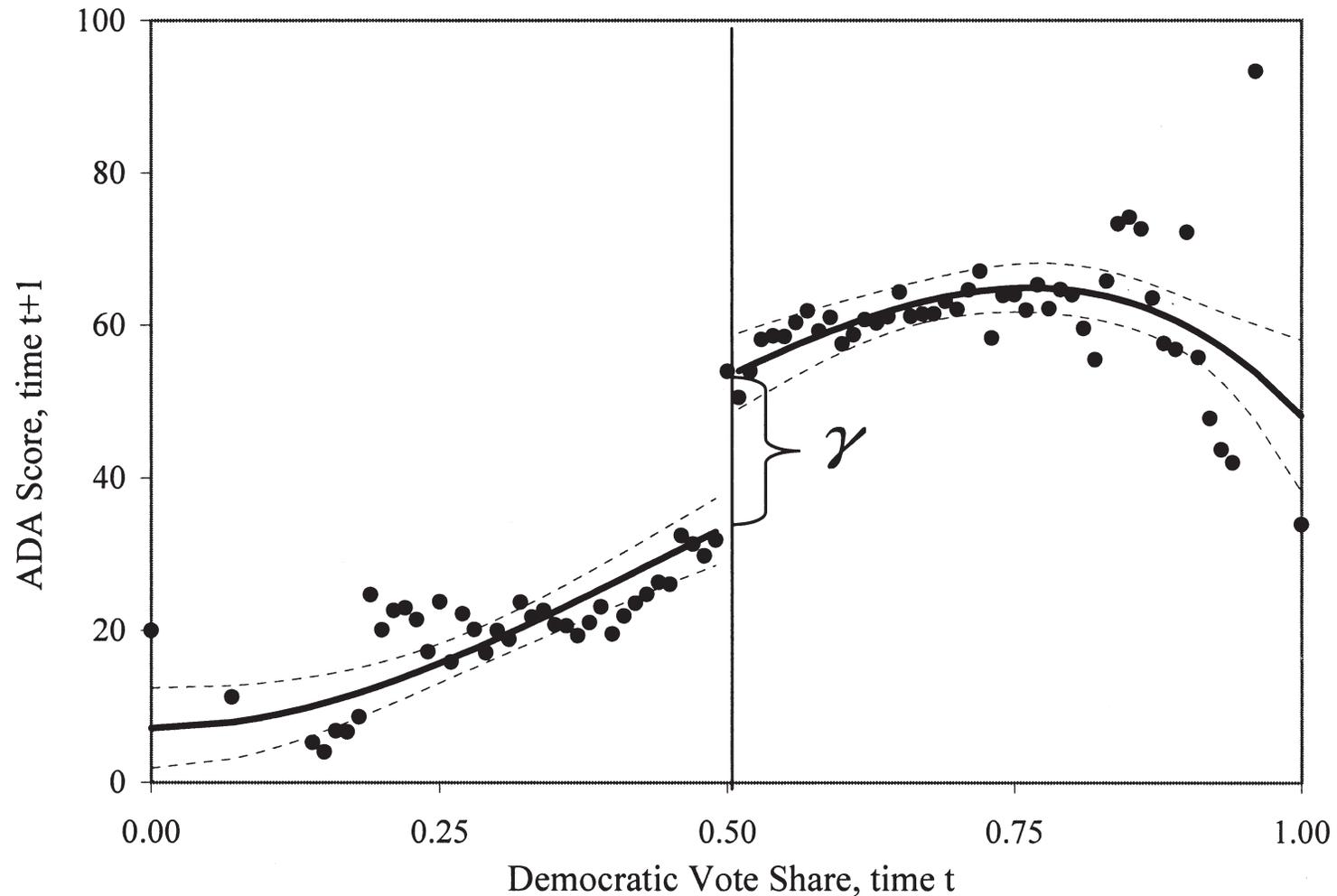


FIGURE I

Total Effect of Initial Win on Future ADA Scores: γ

This figure plots ADA scores after the election at time $t + 1$ against the Democrat vote share, time t . Each circle is the average ADA score within 0.01 intervals of the Democrat vote share. Solid lines are fitted values from fourth-order polynomial regressions on either side of the discontinuity. Dotted lines are pointwise 95 percent confidence intervals. The discontinuity gap estimates

$$\gamma = \underbrace{\pi_0(P_{t+1}^{*D} - P_{t+1}^{*R})}_{\text{"Affect"}} + \underbrace{\pi_1(P_{t+1}^{*D} - P_{t+1}^{*R})}_{\text{"Elect"}}.$$

Source: Lee, Moretti, Butler

PUBLIC CHOICE THEORY: THE FOUNDATIONS OF GOVERNMENT FAILURE

Public choice theory: Government may not act to maximize the well-being of its citizens.

Government failure: The inability or unwillingness of the government to act primarily in the interest of its citizens.

Bureaucracies: Organizations of civil servants that are in charge of carrying out the services of government but follow their self-interest

Leviathan Theory: Under this theory, voters cannot trust the government to spend their tax dollars efficiently and must design ways to combat government greed.

This view of government can explain the many rules in place in the United States and elsewhere that explicitly tie the government's hands in terms of taxes and spending.

PUBLIC VS. PRIVATE PROVISION

The key question raised by this discussion is whether goods and services are provided more efficiently by the public or the private sector

For the production of most goods and services [such as steel, telecommunications, or banking] evidence suggests that private production is more efficient

For goods the public does not understand well (pension funds, health insurance, education), private competition can lead to wasteful advertising (private firms compete on attractive mailings rather than underlying product quality) \Rightarrow higher costs than public provision

PROBLEMS WITH PRIVATIZATION

Natural monopoly: A market in which, because of the uniformly decreasing marginal cost of production, there is a cost advantage to have only one firm provide the good to all consumers in a market [e.g. Microsoft operating system, Google search, Facebook, high speed optical fiber]

With *economies of scale*, the average cost of production falls as the quantity of the output increases.

Private monopoly maximizes profits and under-produces and over-prices relative to efficient outcome: if the government runs or regulates the monopoly, it can restore efficient quantity

Contracting out: Government retains responsibility for providing a good or service, but hires private sector firms to actually provide the good or service. Raises potential for corruption.

CORRUPTION

Corruption: The abuse of power by government officials in order to maximize their own personal wealth or that of their associates.

Electoral accountability is the ability of voters to throw out corrupt regimes.

Corruption also appears more rampant in political systems that feature more *red tape*, bureaucratic barriers that make it costly to do business in a country

Bribes can be seen as “informal taxes” that need to be paid to access government services [corruption often arises when public servants are underpaid]

Do Government Failures Affect Economic Growth?

Studies that suggest that poor government structure can have long-lasting negative impacts on economic growth

1) Acemoglu-Robinson (2012) *Why Nations Fail*: countries with “inclusive governments” (extending political and property rights broadly) grow faster than countries with “extractive governments” (power held by small self-serving elite)

Striking example demonstrating role of political structure: North and South Korea had similar economies when they split in 1948 and have had dramatically different economic development (10 to 1 per capita income ratio today) ⇒ Government policies/failures can have a huge impact

2) Mauro (1995) uses data rating the quality of government along various dimensions (red tape, corruption, etc.): finds that countries with low quality government have lower growth

The difficulty is that the nations with high-quality governments (the treatment group) may differ from those with low-quality governments (the control group) for other reasons as well, biasing the estimates of the effect of government quality.

REFERENCES

Jonathan Gruber, Public Finance and Public Policy, Fourth Edition, 2012
Worth Publishers, Chapter 9

Acemoglu, Daron, and James Robinson. "Why nations fail: the origins of power, prosperity, and poverty." [Book] Random House Digital, Inc., 2012.

Geanakoplos, John. "Three brief proofs of Arrow's impossibility theorem." Economic Theory 26.1 (2005): 211-215. (web)

Lee, David S., Enrico Moretti, and Matthew J. Butler. "Do voters affect or elect policies? Evidence from the US House." The Quarterly Journal of Economics 119.3 (2004):807-859. (web)

Mauro, Paolo. "Corruption and growth." The Quarterly Journal of Economics 110.3 (1995): 681-712.(web)

Poole, Keith T. and Howard Rosenthal, "Are legislators ideologues or the agents of constituents?" European Economic Review,40(3-5), 1996, 707-717.(web)

Washington, Ebonya L. "Female Socialization: How Daughters Affect Their Legislator Fathers." Voting on Women's Issues." American Economic Review 98.1 (2008): 311-332.(web)

Undergraduate Public Economics

Danny Yagan
UC Berkeley

Lecture 14

State and Local Government Expenditures

FISCAL FEDERALISM

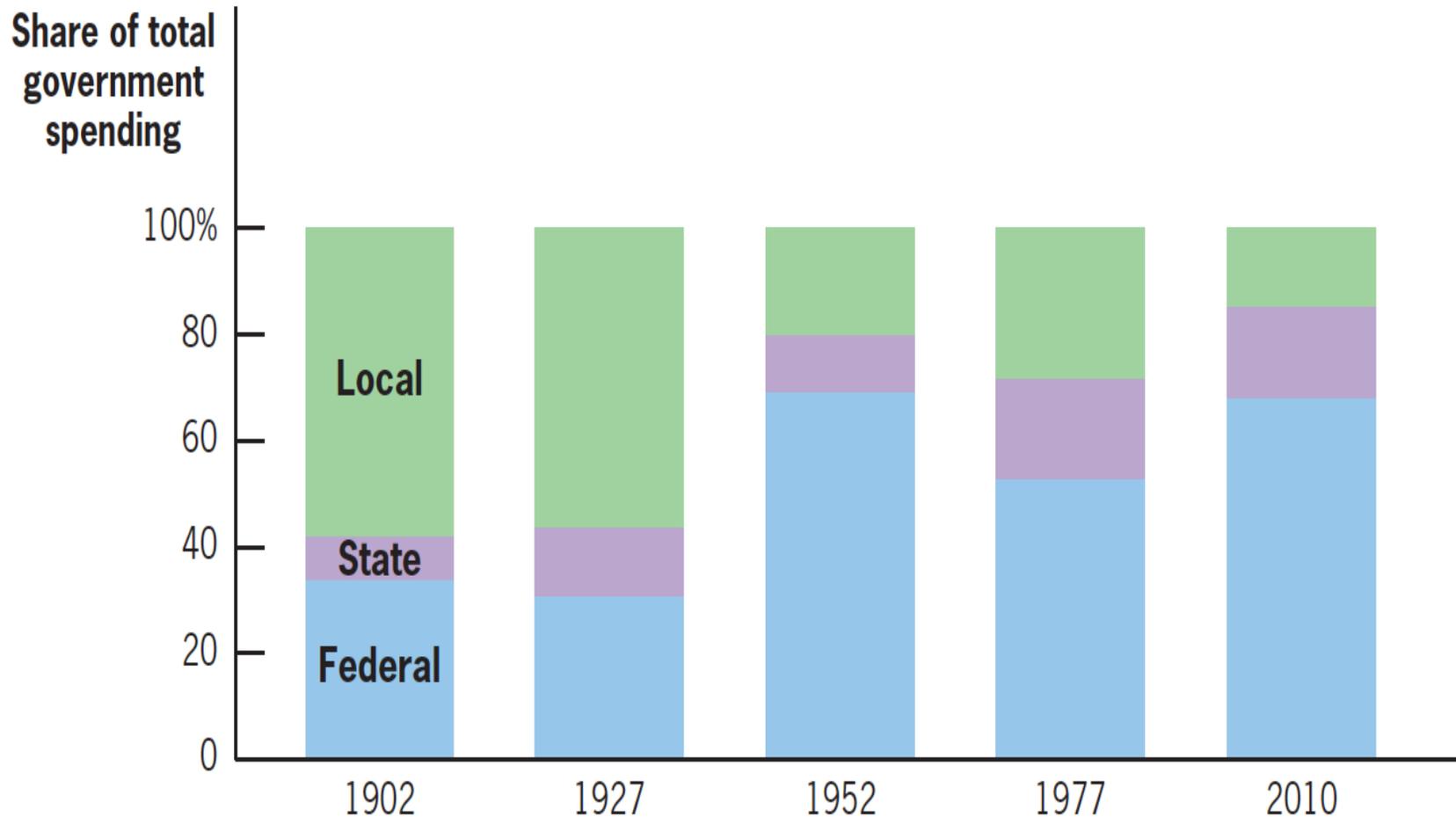
optimal fiscal federalism: The question of which activities should take place at which level of government

The distribution of government spending has changed dramatically over time in the United States

- Local and state spending have declined considerably.
- Much state and local spending now supported by intergovernmental grants [transfers from the federal government]

10.1

State and Local Spending in the United States, 1902–2010



SPENDING AND REVENUE OF STATE AND LOCAL GOVERNMENTS

Property tax: The tax on land and any buildings on it, such as commercial businesses or residential homes.

Main source of revenue from local governments due to:

1) History: real estate property is visible and hence taxable even in archaic economies with informal businesses

2) Immobile tax base: the real estate tax base cannot flee to another jurisdiction (mobility of the tax base is an issue for local governments)

Today, property tax is about 1/3 of revenue raised by state+local government (rest is 1/3 income tax, 1/3 sales taxes)

10.1

Spending and Revenue of State and Local Governments

	Spending			Revenue	
	State	\$/PC		State	\$/PC
Education spending	AK	3,010	Income taxes	NY	2,311
	MA	2,643		MT	854
	TN	1,50		Many	0
Health care spending	DC	10,349	Sales taxes	DC	1,847
	LA	6,759		Iowa	698
	UT	5,031		Many	0

THE TIEBOUT MODEL

What is it about the private market that guarantees optimal provision of private goods that is missing in the case of public goods?

Tiebout's insight was that the factors missing from the market for public goods were shopping and competition

The situation is different when public goods are provided at the local level by cities and towns:

Competition will naturally arise because individuals can *vote with their feet*: if they don't like the level or quality of public goods provision in one town, they can move to the next town

This threat of exit can induce efficiency in local public goods production

THE TIEBOUT FORMAL MODEL

We consider a very simple model to illustrate Tiebout's insight and theorem

Suppose there are $2 \cdot N$ families with identical income Y and 2 towns with N homes each

Towns supply level G_1, G_2 of local public schools

There are 2 types of families:

1) N families with kids, with utility $U^K(C, G)$, value private consumption C and schools G

2) N elderly families, with utility $U^E(C)$, value only private consumption C

THE TIEBOUT EQUILIBRIUM DEFINITION

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

1) In each town, G is decided by median voter and financed equally by town residents with budget $Y = G/N + C$

⇒ If majority in town is elderly then $G = 0$ as this maximizes $U^E(Y - G/N)$

⇒ If majority in town is families with kids then $G = G^*$ that maximizes $U^K(Y - G/N, G)$

2) No 2 families want to exchange locations across towns

THE TIEBOUT THEOREM

Tiebout Theorem Part I: In equilibrium, families will sort themselves in towns according to their taste for public good (1 town with elderly only, 1 town with families with kids only)

Proof: Suppose elderly dominate in town 1 and $G_1 = 0$, then families with kids dominate in town 2 and $G_2 = G^*$. If there is a family with kids in town 1, then there is an elderly family in town 2 and they are willing to switch \Rightarrow not an equilibrium.

Tiebout Theorem Part II: In each town, the level of local public good is efficient

Proof: In elderly town, $G = 0$ which is efficient as nobody values G .

In kids town, G^* maximizes $U^K(Y - G/N, G)$ so that $-U_C^K/N + U_G^K = 0 \Rightarrow U_G^K/U_C^K = 1/N \Rightarrow \sum U_G^K/U_C^K = \sum MRS_{GC} = N/N = 1 = MC$ which is the Samuelson rule

CAVEAT

Tiebout efficiency result holds when there is no cost advantage to everyone living in one big city

Pure public goods always have a cost advantage (economies of scale). Ex. need only one missile defense shield for country

Tiebout shows that efficiency can be restored when public good is rival (i.e. an impure public good). Ex. my kid's desk at school displaces another kid from using the desk

Nevertheless proceed calling the locally publicly provided item a "public good" even though it is not a pure public good

THE TIEBOUT MODEL

People can vote with their feet by choosing the locality that best fits their tastes and provides the best public goods given the tax

The main message of the model is that competition across local jurisdictions puts competitive pressure on the provision of local public goods:

- 1) Public goods need to reflect tastes of local residents
- 2) Public goods need to be efficiently provided (without waste)

PROBLEMS WITH THE TIEBOUT MODEL

The Tiebout model is an idealized model that requires a number of assumptions that may not hold perfectly in reality:

- 1) Individuals can move costlessly across towns (low mobility costs)
- 2) Individuals have perfect information on the benefits and taxes paid in each town
- 3) There must be enough towns so that individuals can sort themselves into groups with similar preferences for public goods
- 4) No externalities/spillovers of public goods across towns [with spillovers across towns, public goods will be under provided in Tiebout model, e.g. parks, police]

PROBLEMS WITH THE TIEBOUT MODEL

The Tiebout model requires equal financing of the public good among all residents.

Lump-sum tax: A fixed taxation amount independent of a person's income, consumption of goods and services, or wealth. Sometimes called a poll tax.

Towns typically finance their public goods instead through a property tax where rich pay more than poor (because they live in nicer houses). The problem that property taxation causes is that the *poor chase the rich* (rich also want to be with rich)

Two mechanisms prevent poor from chasing the rich:

- 1) **Housing prices:** places with rich people have high housing prices
- 2) **Zoning:** Restrictions that towns place on the use of real estate (e.g., each house must sit on a parcel of at least 6000 sq feet)

Zoning regulations protect the tax base of wealthy towns by pricing lower-income people out of the housing market.

1990 Poll Tax Riots: 200,000 People in Central London

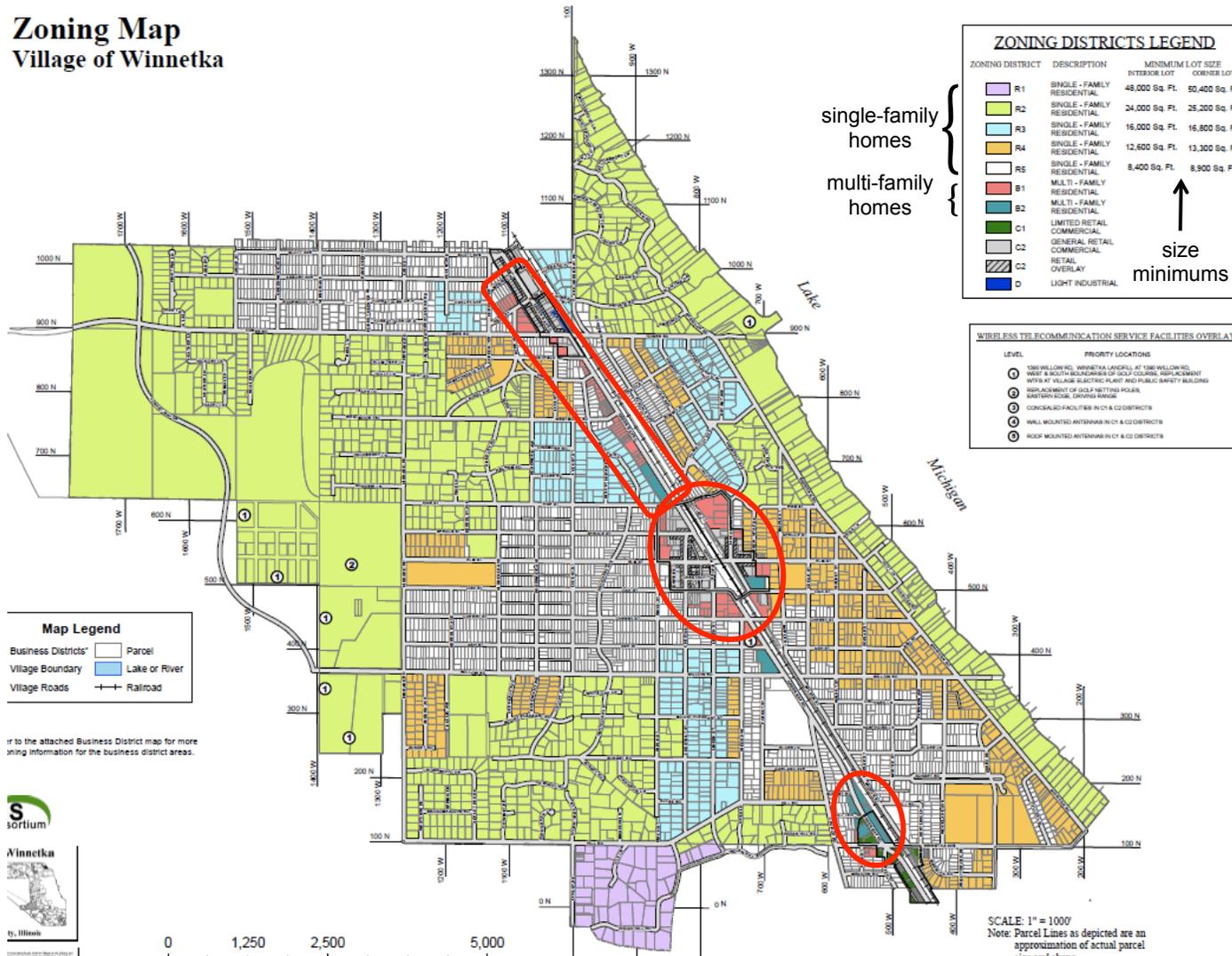


1990 Poll Tax Riots: 200,000 People in Central London



Zoning Example: Winnetka IL

Zoning Map Village of Winnetka



single-family homes
multi-family homes

EVIDENCE ON THE TIEBOUT MODEL

Tiebout Sorting: Resident Similarity Across Areas

A testable implication of the Tiebout model is that when people have more choice of local community, the tastes for public goods will be more similar among residents than when people do not have many choices

This fact is indeed pretty well established

More Efficiency when there is more Tiebout sorting

This fact is controversial

Evidence on the Tiebout Model: Hoxby (2000)

Hoxby (2000) considers public school districts in the US. She compares cities where:

A) There are few large school districts and hence little choice for residents (such as Miami)

B) There are many small school districts and hence a lot of choice for residents (such as Boston)

2 key findings:

I) Cities with few districts have less sorting across neighborhood (in terms of school quality) than cities with many districts (this result is well established)

II) Cities with many districts have **higher** test scores on average: this result is controversial (see Rothstein, 2007 critique)

EVIDENCE ON THE TIEBOUT MODEL

Capitalization of Fiscal Differences into House Prices

House price capitalization: Incorporation into the price of a house of the costs (including local property taxes) and benefits (including local public goods) of living in the house.

⇒ High property taxes (relative to public goods quality) depresses housing prices

⇒ Low property taxes (relative to public goods quality) increases housing prices

Example: Suppose \$1 cut in property tax (in perpetuity) with no change in value of local public good

Capitalized value: $\Delta V = 1 + 1/(1 + r) + 1/(1 + r)^2 + \dots = 1/[1 - 1/(1 + r)] = (1 + r)/r = \21 if $r = 5\%$

10.2

EVIDENCE: Evidence for Capitalization from California's Proposition 13

- California's Proposition 13 became law in 1978.
 - Set the maximum amount of any tax on property at 1% of the "full cash value."
 - Full cash value: Value as of 1976, with annual increases of 2% at most.
- Reduced property taxes immensely in some areas, little change in others.

10.2

EVIDENCE: Evidence for Capitalization from California's Proposition 13

- Each \$1 of property tax reduction increased house values by about \$7, about equal to the PDV of a permanent \$1 tax cut.
- In principle, the fall in property taxes would result in a future reduction in public goods and services, which would lower home values.
- The fact that house prices rose by almost the present discounted value of the taxes suggests that Californians did not think that they would lose many valuable public goods and services when taxes fell.

KEY CONSEQUENCE OF TIEBOUT MODEL

It is very hard for a local government to redistribute from rich to poor:

If local redistribution is high \Rightarrow

- 1) Poor flock to the city which provides welfare benefits
- 2) Rich flee to other cities to avoid paying for redistribution
 \Rightarrow Local redistribution program will break down

Redistribution programs work better if implemented at higher level: state or federal (harder to leave the state or country). At local level, need to have tax-benefit linkage to avoid migration

tax-benefit linkages: The relationship between the taxes people pay and the government goods and services they get in return.

REDISTRIBUTION ACROSS COMMUNITIES

There is currently enormous inequality in both the ability of local communities to finance public goods and the extent to which they do so.

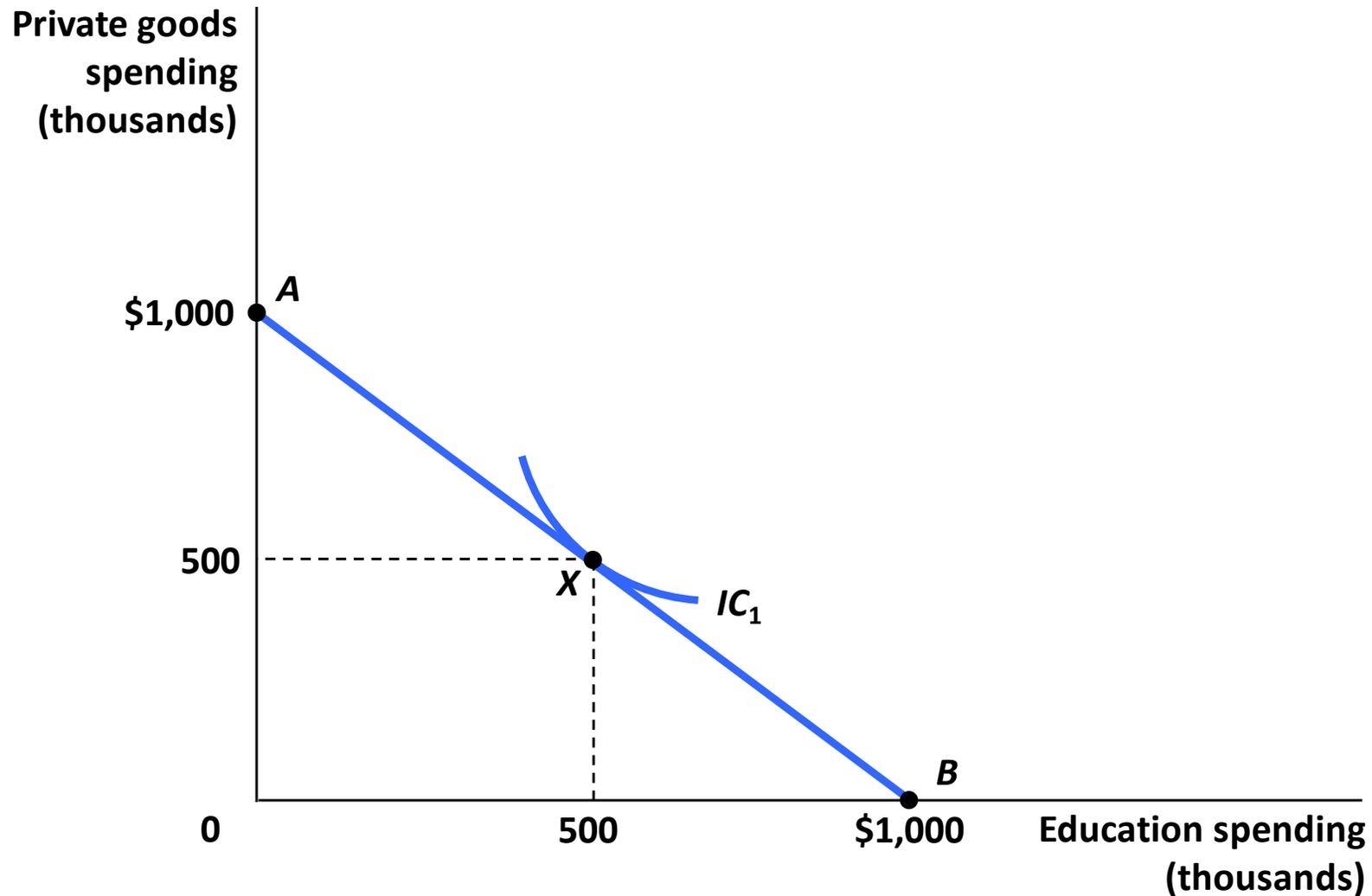
Central government can redistribute across communities **directly** using taxes and spending but also **indirectly** by giving grants to lower levels of government

Higher levels of government can redistribute across lower levels of government through *intergovernmental grants*.

We assume in graphical analysis that local community chooses public spending and private spending according the preferences of Median voter in the community

10.3

Tools of Redistribution: Grants



Intergovernmental Grants

Higher level government can provide grants to redistribute across communities and incentivize communities to spend on public goods

Three main forms of grants:

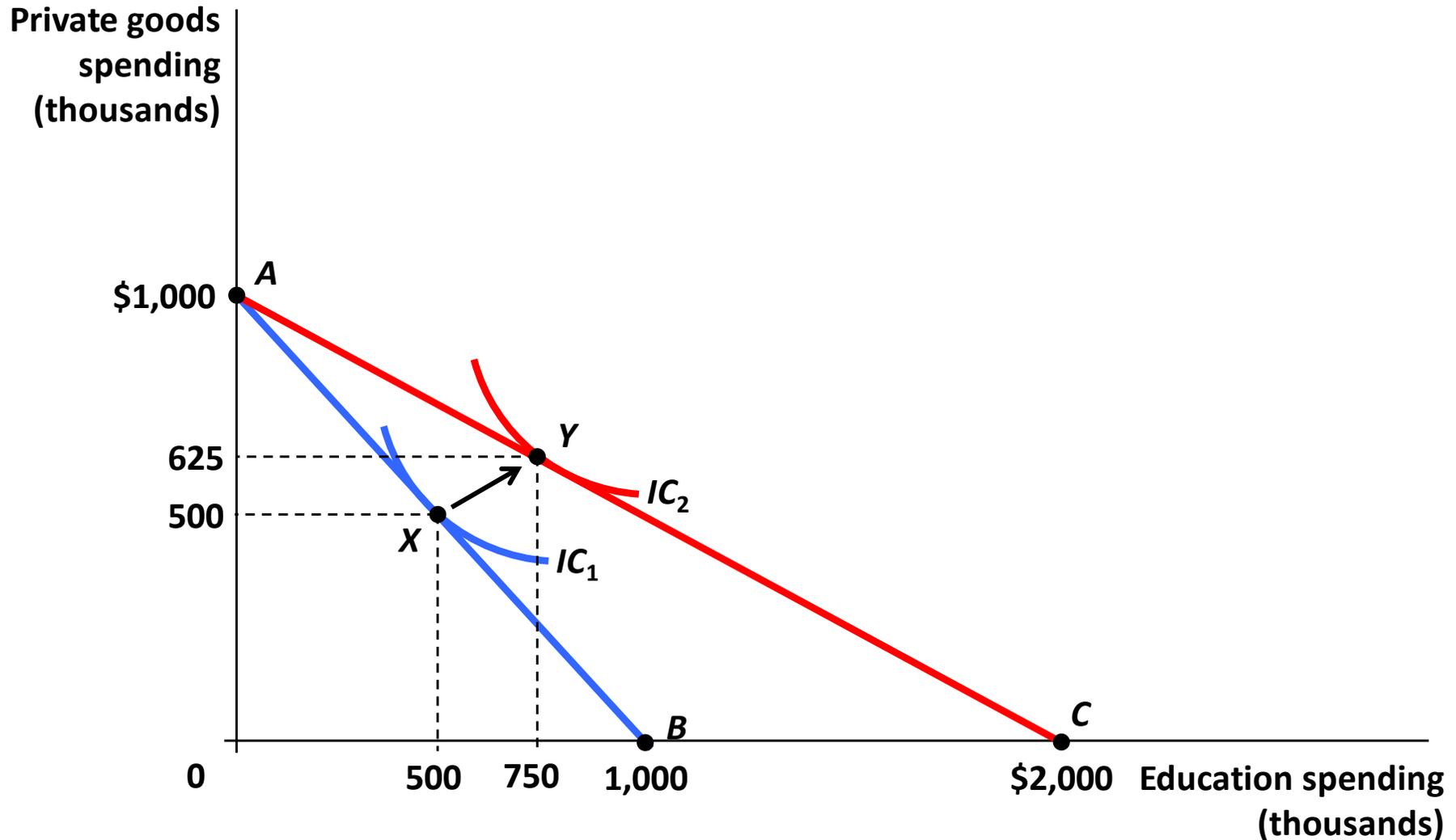
1) Matching grant: A grant, the amount of which is tied to the amount of public good spending by the local community.

2) Block grant: A grant of some fixed amount with no mandate on how it is to be spent.

3) Conditional block grant: A grant of some fixed amount with a mandate that the money be spent in a particular way.

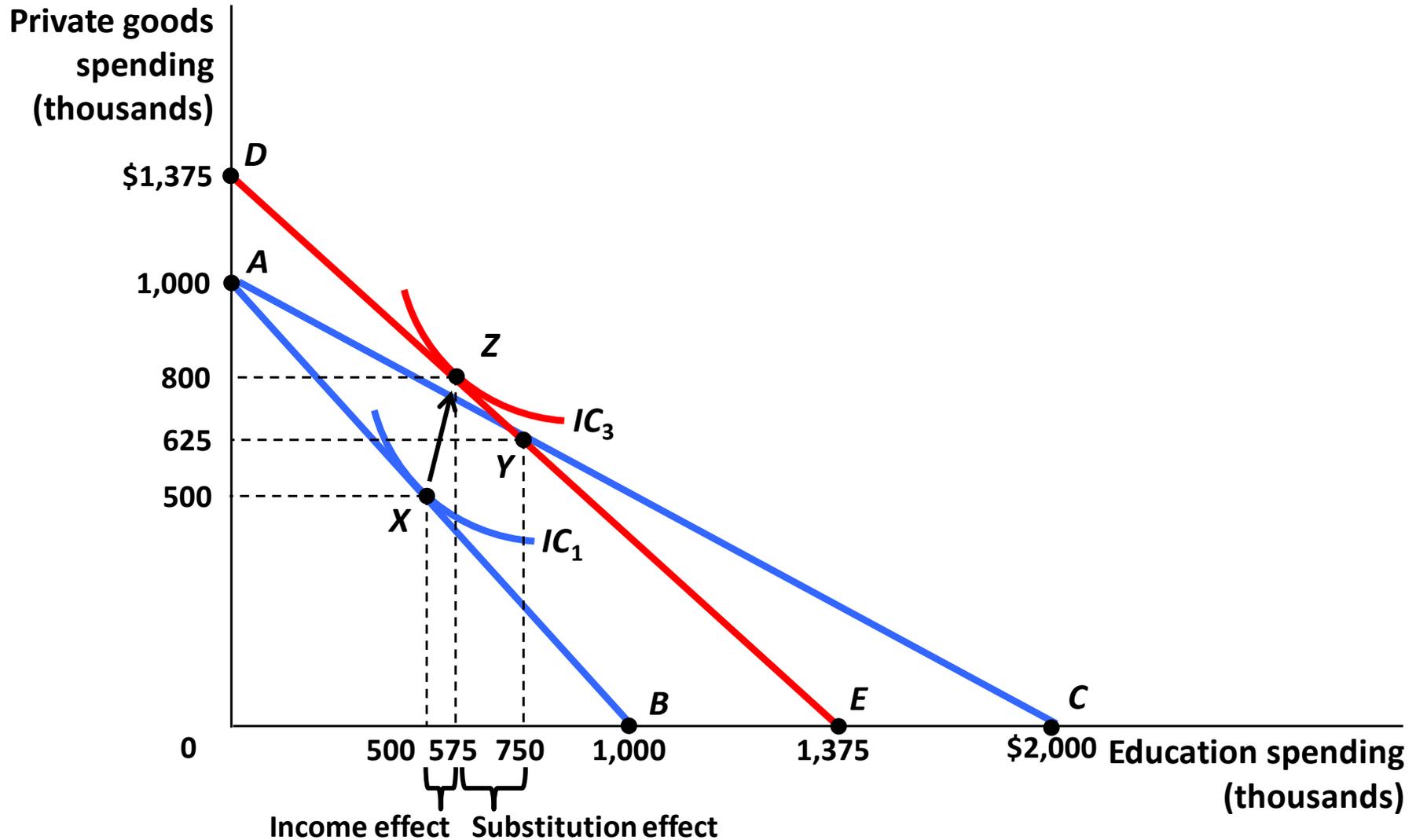
10.3

Matching Grants



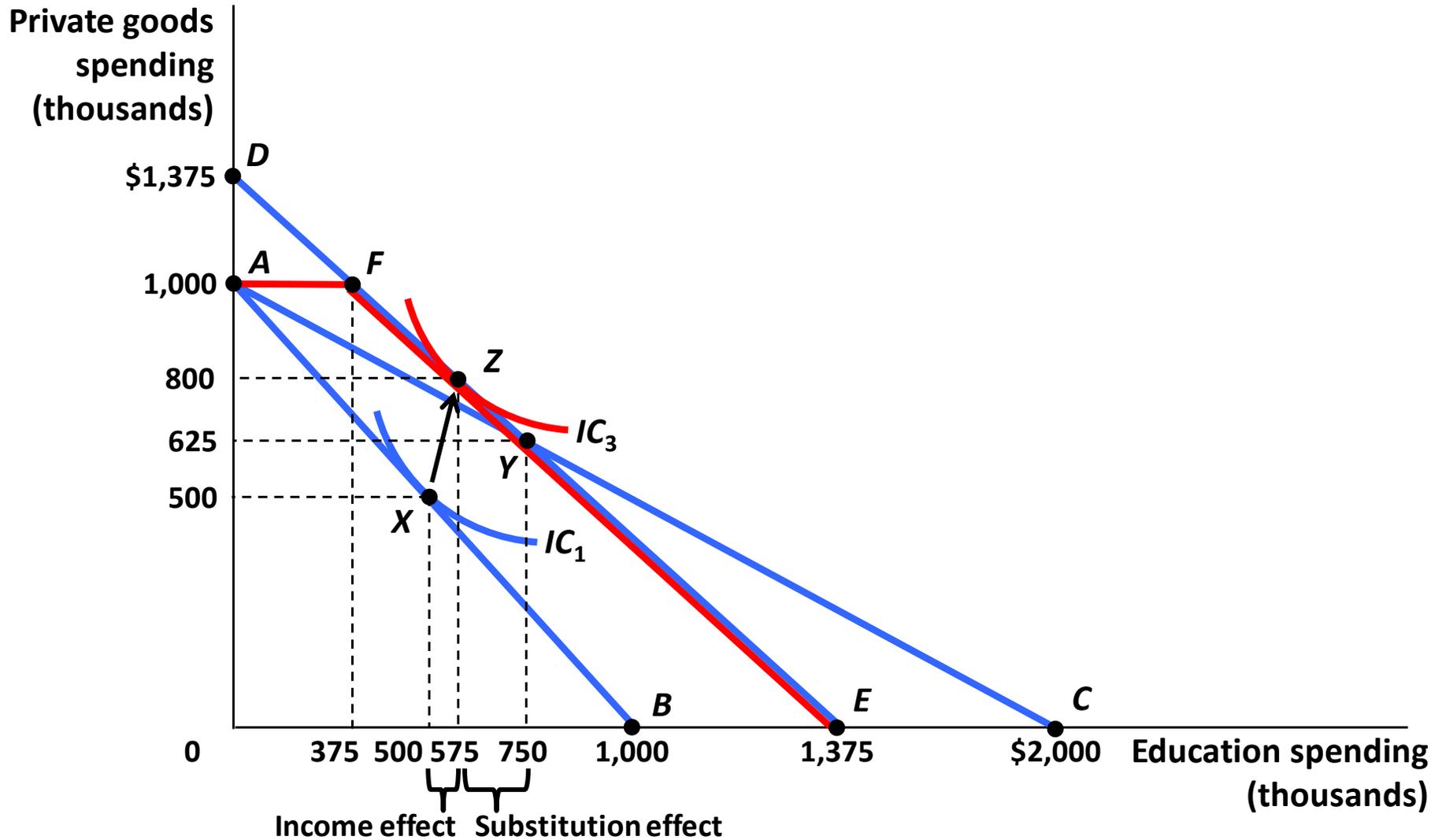
10.3

Block Grant



10.3

Conditional Block Grant



KEY PREDICTION OF THEORY: CROWD-OUT

In the theory presented, a \$1000 increase in private income has the same effect as a \$1000 increase in Fed block grant: both shift the budget in the same way and lead to the same outcome

Example: \$1000 private income increase leads to \$800 more in private consumption and \$200 more in local taxes and public spending. \$1000 extra fed grant leads to \$200 extra in public good spending and \$800 cut in local taxes and hence \$800 extra in private consumption

Similarly, with multiple public goods (e.g., schools and police), an extra \$1000 Fed grant for school has the same effect on schools and police than a \$1000 Fed grant for police

Money is fungible: only total resources matter for the allocation across private good and public goods at the local level

THE FLYPAPER EFFECT

Hines and Thaler JEP'95 found that the crowd-out of state spending by federal spending is low and often close to zero

Economist Arthur Okun described this as the flypaper effect because “the money sticks where it lands” instead of replacing state spending

But evidence is based on correlation [not necessarily causation as states that get grants maybe the ones that like spending the most]

Recent studies show that there is a flypaper effect in the short-run but that there is substantial crowd-out from block grants in the long-run

REDISTRIBUTION IN ACTION: SCHOOL FINANCE EQUALIZATION

School finance equalization: Laws that mandate redistribution of funds across communities in a state to ensure more equal financing of schools.

Without school finance equalization, huge disparity in property tax base and hence school funding (per pupil) across areas (example from Bay Area: Lafayette is very wealthy, Richmond is poor)

Many states (including California) impose equalization: pool local taxes at state level and redistribute them across districts

Equalization often imposed by courts without thinking carefully about economic consequences

REDISTRIBUTION IN ACTION: SCHOOL FINANCE EQUALIZATION

Tax price: For school equalization schemes, the amount of revenue a local district would have to raise in order to gain \$1 more of spending. Examples:

- 1) With no equalization, the tax price is \$1 (local govt keeps all its revenue)
- 2) With perfect equalization, the tax price is infinite (raising local revenue has zero impact on local spending)
- 3) Suppose state takes 50% of local revenue and redistributes it equally across cities: the tax price is \$2 (city needs to raise \$2 in local revenue to be able to keep \$1 extra)

CALIFORNIA SCHOOL EQUALIZATION

In 1960s-1970s, California used to have one of the best public school systems in the nation, now it has one of the worst

California used to have no school finance equalization and hence big disparities across areas

1976: Serrano vs. Priest case: Supreme court ruled that disparities above a threshold were unconstitutional

⇒ Wealthy districts forced to give all their tax revenue above the threshold to the common pool to fund poor districts

⇒ local government has no incentive to raise taxes ⇒ taxes and school funding fall in rich districts

⇒ Property taxes no longer able to fund schools adequately

10.3

APPLICATION: School Finance Equalization and Property Tax Limitations in California

If residents perceived that property taxes were “too high” in California, why did they wait until 1978 to lower them?

- Proposition 13 actually a response to school finance equalization in California.
- Taxes no longer financed local school spending; just taxes, rather than prices. Tax price became infinite.
- Voters were happy to limit property taxes once those taxes no longer brought them any benefit.

CONCLUSION

Higher levels of government are able to implement redistribution across lower levels of government either directly with means-tested programs for individuals (such as Food stamps) or through grants to local governments (such as Medicaid)

In the same way that 100% tax on individuals is counterproductive, 100% confiscation of local taxes (as in the most drastic school equalization schemes) is counterproductive (leads local governments to abandon their taxes)

Higher level government should have primary responsibility for redistribution

REFERENCES

Jonathan Gruber, Public Finance and Public Policy, Fourth Edition, 2012
Worth Publishers, Chapter 10

Hines, James R., and Richard H. Thaler. "Anomalies: The flypaper effect."
The Journal of Economic Perspectives 9.4 (1995): 217-226.(web)

Hoxby, Caroline M. "Does Competition among Public Schools Benefit Students and Taxpayers?." The American Economic Review 90.5 (2000): 1209-1238.(web)

Rothstein, Jesse. "Does Competition Among Public Schools Benefit Students and Taxpayers? Comment." The American Economic Review 97.5 (2007): 2026-2037.(web)

Tiebout, Charles M. "A pure theory of local expenditures." The Journal of Political Economy 64.5 (1956): 416-424.(web)

Undergraduate Public Economics

Danny Yagan
UC Berkeley

Lecture 15
Education

Education

Education is one of the largest public goods provided by government

Approximately 5.5% of GDP or 1/6 of government expenditure in the US

About 80% of spending done at the state and local level

Focus of an extensive body of research in the rapidly expanding field of economics of education

Why Should the Government Be Involved in Education?

Ex-ante not obvious because education does not look like a public good

1) Returns to education are largely private

2) Education is (at least partially) excludable and rival

⇒ we should expect students to invest roughly the optimal amount in their own education

Why Should the Government Be Involved in Education?

Four motives for government intervention:

- 1) Externalities (productivity spillovers, crime, citizenship)
- 2) Family failures: Divergence between parent and child preferences (some parents may not take good care of their children)
- 3) Borrowing constraints (poor but talented students may not be able to borrow against future earnings to get an education)
- 4) Individual failures: young people might not do what is in their long-run interest due to self-control problems or lack of information

1) Externalities of education on crime and voting

$$Crime_i = \alpha + \beta Educ_i + \varepsilon_i$$

Observational regression comparing the educated vs. not-educated likely biased because propensity to crime ε_i is negatively correlated with $Educ_i$.

Lochner and Moretti (2004) use as instrument changes in state compulsory attendance laws: State T increases compulsory attendance from 9 to 10 years at time t , State C does not.

Can look at effect on education (“first-stage”) and then look at effect on crime (“second-stage”) using Difference-in-difference

They show that an extra year of schooling reduces incarceration rates significantly

0.1 pct point decline for white males relative to a mean of 1%

0.3 pct point decline for black males relative to mean of 3%

⇒ Gap in schooling between whites and blacks accounts for more than 1/4 difference in crime rates

Social return to education exceeds private return by 25% based purely on reduction in crime

Moretti, Mulligan, Oreopoulos (2003) find positive effects of education on likelihood of voting using same strategy

2) Divergences between parent and child preferences

Hard to find direct evidence

Duflo (2003) shows evidence that grandmothers spend more than grandfathers on female grandchildren

Duflo (2003) uses pension reform in 1992 in South-Africa giving all Blacks (65+) a minimum pension when household income is low (before, only whites could get the pension under Apartheid)

Duflo (2003) finds that pension availability improves the weight for height Z-score of female grandchildren (nutrition improves) but only when a grandmother gets the pension (and not when a grandfather does)

⇒ Parents preferences matter for kids outcomes

TABLE 3. Effect of the Old-Age Pension Program on Weight for Height: OLS and 2SLS Regressions

Variable	OLS						2SLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>Girls</i>							
Eligible household	0.14 (0.12)	0.35* (0.17)	0.34* (0.17)				
Woman eligible ^a				0.24* (0.12)	0.61* (0.19)	0.61* (0.19)	1.19* (0.41)
Man eligible ^b				-0.011 (0.22)	0.11 (0.28)	0.056 (0.19)	-0.097 (0.74)
Observations	1574	1574	1533	1574	1574	1533	1533
<i>Boys</i>							
Eligible household	0.0012 (0.13)	0.022 (0.22)	0.030 (0.24)				
Woman eligible ^a				0.066 (0.14)	0.28 (0.28)	0.31 (0.28)	0.58 (0.53)
Man eligible ^b				-0.059 (0.22)	-0.25 (0.34)	-0.25 (0.35)	-0.69 (0.91)
Observations	1670	1670	1627	1670	1670	1627	1627
<i>Control variables</i>							
Presence of older members ^c	No	Yes	Yes	No	Yes	Yes	Yes
Family background variables ^d	No	No	Yes	No	No	Yes	Yes
Child age dummy variables ^e	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*Significant at the 5 percent level.

Note: The instruments in column 7 are woman eligible and man eligible (the first stage is in table A-1). Standard errors (robust to correlation of residuals within households and heteroscedasticity) are in parentheses.

^aIn column 7 this variable is replaced by a dummy for whether a woman receives the pension.

^bIn column 7 this variable is replaced by a dummy for whether a man receives the pension.

^cPresence of a woman over age 50, a man over age 50, a woman over age 56, a man over age 56, and a man over age 61.

^dFather's age and education; mother's age and education; rural or metropolitan residence (urban is the omitted category); size of household; and number of members ages 0-5, 6-14, 15-24, and 25-49.

^eDummy variables for whether the child was born in 1991, 1990, or 1989.

Source: Author's calculations.

Source: Duflo (2003)

3) Borrowing Constraints: effects of loans

If there are no borrowing constraints (and individuals are rational), current resources should not matter for educational decisions: invest in education only if PDV benefits $>$ costs

Empirical evidence shows that availability of loans do matter suggesting that borrowing constraints are an issue

Solis (2013) studies the effects of guaranteed loans on college attendance in Chile

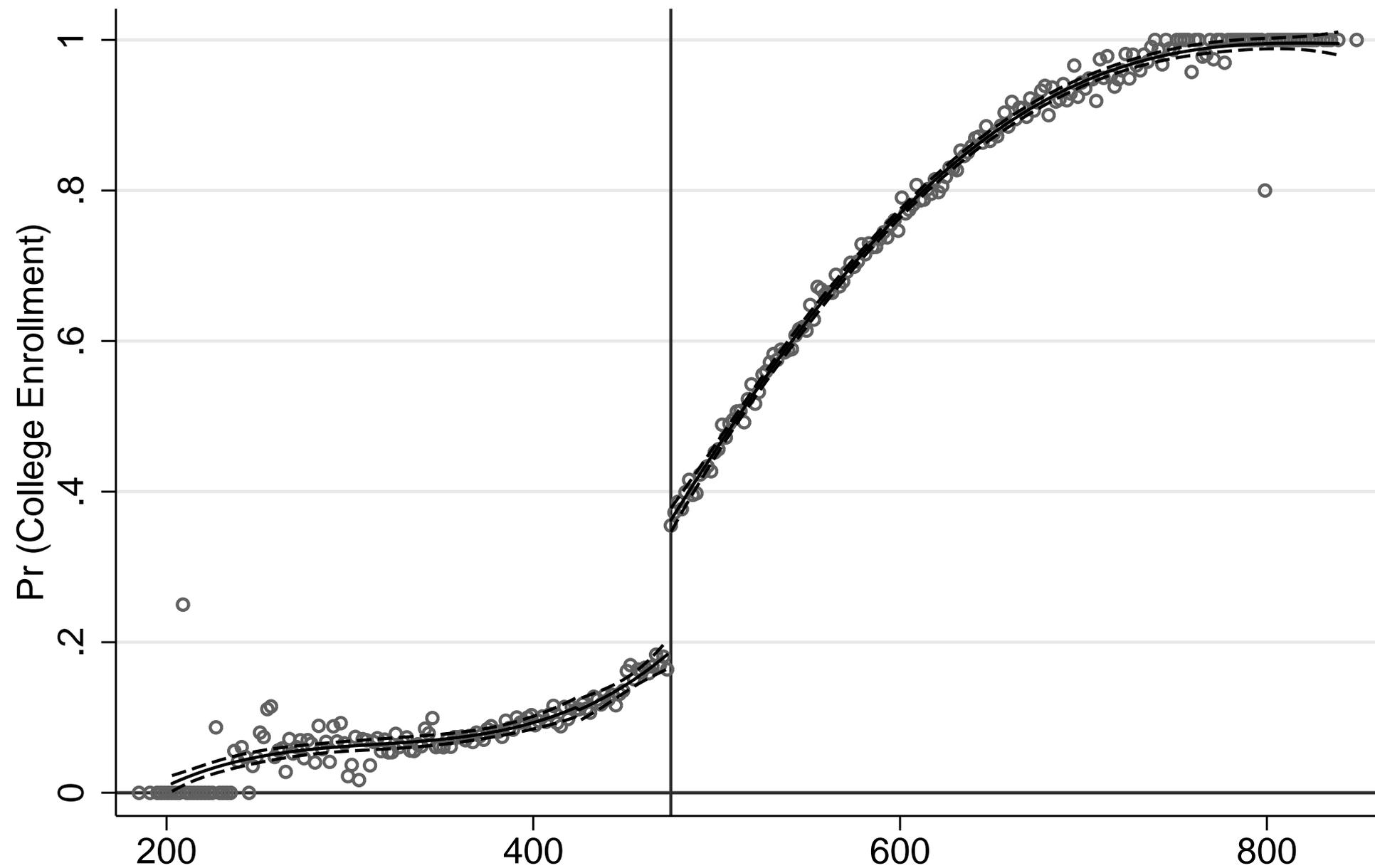
Guaranteed loan is available if test score of student (equivalent of SAT for Chile) is above threshold equal to 475.

Regression discontinuity design: does discontinuity in loan availability translate into discontinuity in college attendance?

YES

⇒ Very compelling evidence that loan availability matters

College Enrollment; All Years; bw=2



PSU score
Source: Solis (2013)

4) Behavioral motives (individual failures): high-school

Rational education decision should be based on comparing returns to education (higher wage later in life) vs. costs of education (tuition and time) \Rightarrow Requires that young individuals know the return to education

Jensen (2010) shows that simply presenting information about rates of return to education changes behavior

1) He uses survey data for eighth-grade boys in the Dominican Republic

Finds that the perceived returns to secondary school are extremely low, despite high measured returns

2) Then carries out randomized field experiment: Students at randomly selected schools given information on the higher measured returns completed on average 0.20 more years of school over the next four years than those who were not.

EFFECT OF PROVIDING INFORMATION ABOUT RETURNS TO COLLEGE IN DOMINICAN REPUBLIC

	<i>Δ Implied Return (Self)</i>		<i>Returned Next Year</i>			<i>Completed Secondary</i>		<i>Years of Schooling</i>	
	(1)	(2)	(3)	(4)	(5)	(7)	(8)	(9)	(10)
Treatment	366 (29)	366 (29)	.039 (.025)	.041 (.023)		.020 (.024)	.023 (.020)	.18 (.098)	.20 (.083)
Log (income per capita)		30.0 (48)		.075 (.042)			.21 (.044)		.75 (.16)
School Performance		1.1 (13)		.011 (.010)			.019 (.008)		.085 (.035)
Father's education		-26 (33)		.082 (.029)			.061 (.029)		.28 (.12)
Interviewed									.014 (.027)

Source: Jensen 2010

4) Behavioral motives (individual failures): university

Hoxby-Avery '12 shows that high-achieving US students (top 2% of SAT scores) from disadvantaged backgrounds apply to weaker colleges [than other high-achieving US students]

Even though top schools offer generous financial aid to talented students from disadvantaged backgrounds

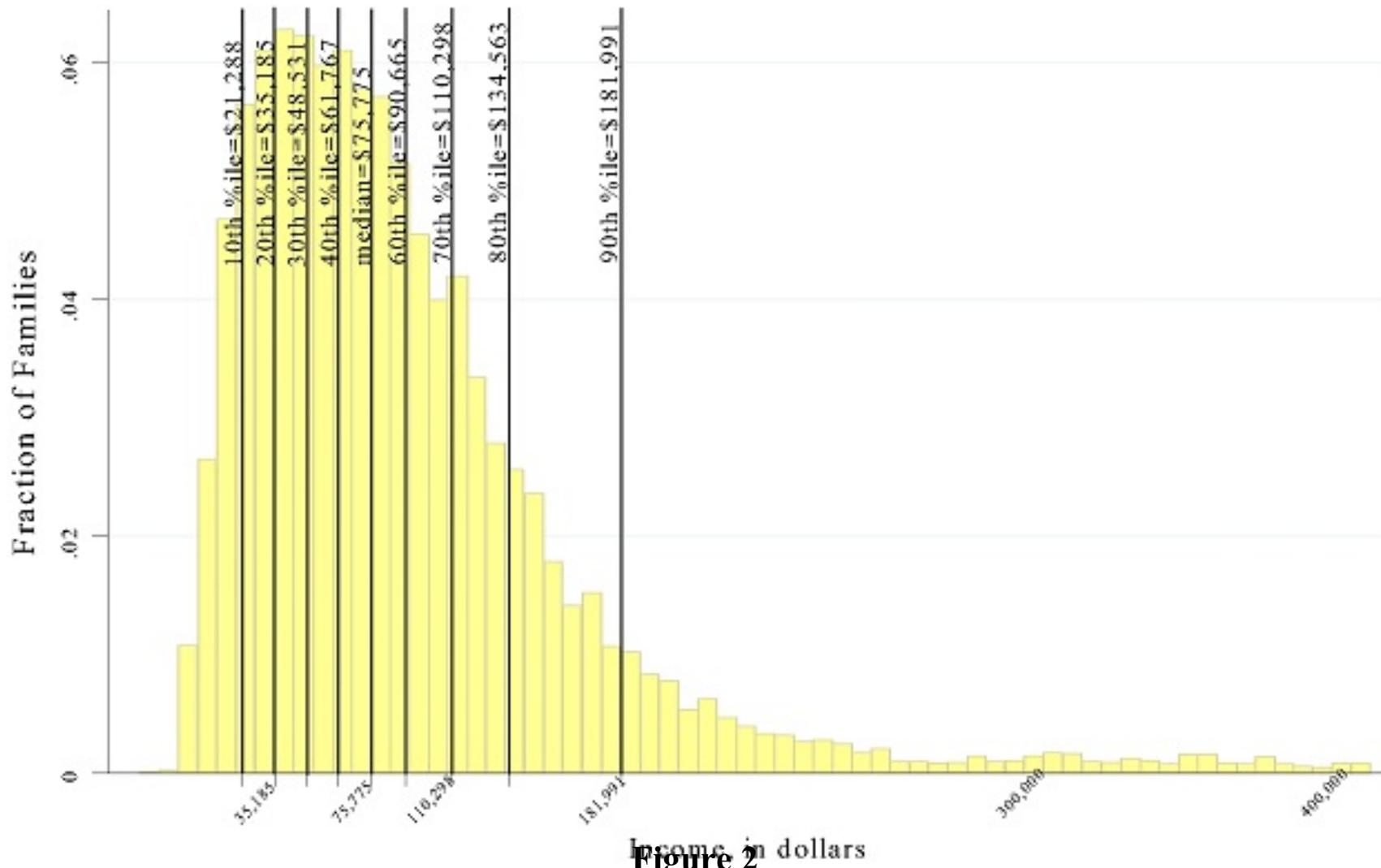
Mechanism: poor talented kids in “nowhere” schools do not get good advice from family/local counselors ⇒ End up at local college (often paying more than they would at top college)

⇒ Informational failure prevents poor but talented kids to exploit their potential

Hoxby-Turner '13 does randomized experiment providing personalized mailing info to talented students (relevant suggested applications, net-cost calculator) ⇒ Significant effect on number and quality of applications

Figure 1

Distribution of Family Income Among Families with a Child in the 12th Grade, 2008



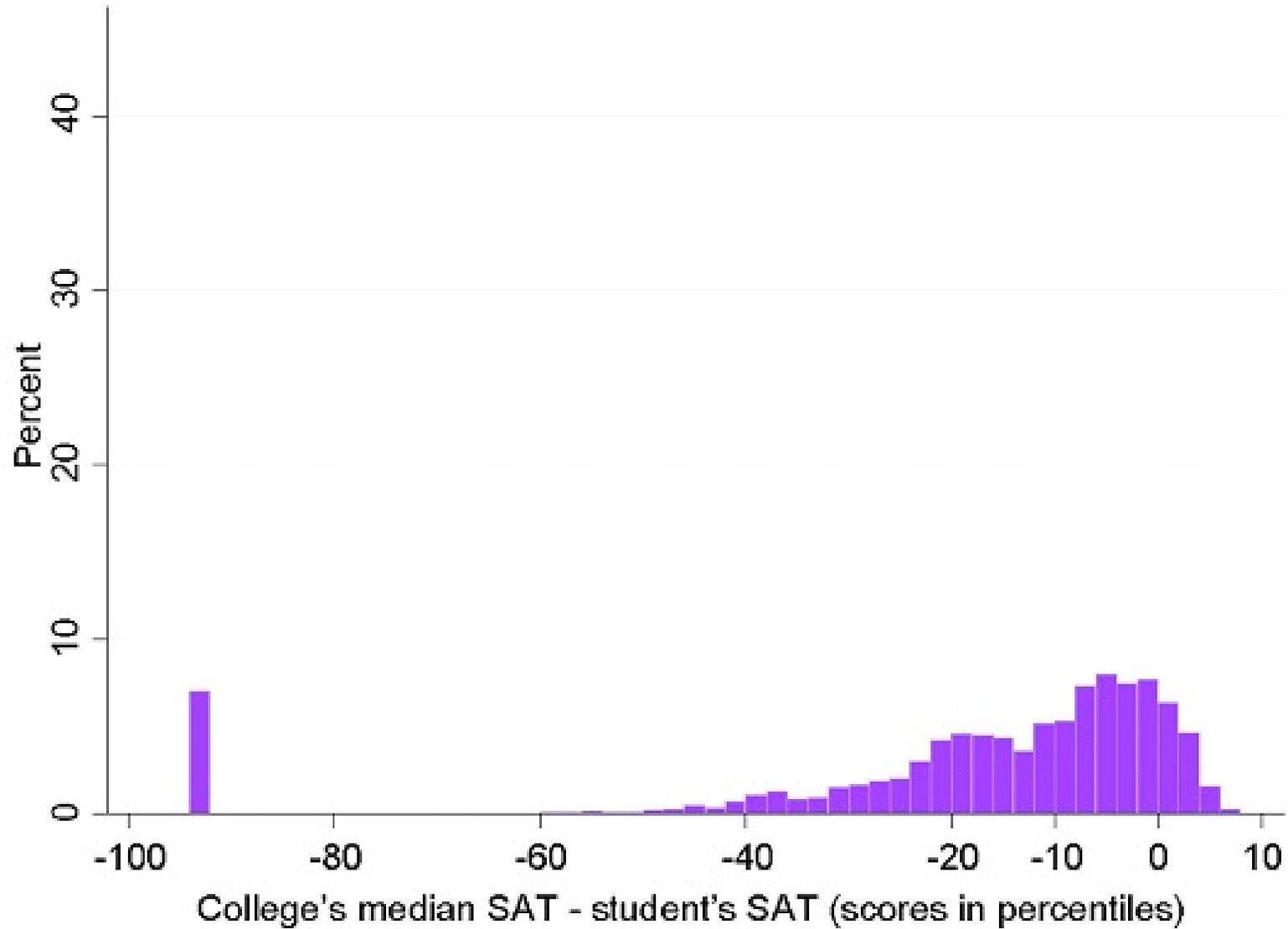
Source: Hoxby, C. M., & Avery, C. 2012

Table 1
College Costs and Resources by Selectivity

Selectivity (Barron's)	Out-of-Pocket Cost for a Student at the 20 th Percentile of Family Income (includes room and board)	Comprehensive Cost (includes room and board)	Instructional Expenditure per Student
most competitive	6,754	45,540	27,001
highly competitive plus	13,755	38,603	13,732
highly competitive	17,437	35,811	12,163
very competitive plus	15,977	31,591	9,605
very competitive	23,813	29,173	8,300
competitive plus	23,552	27,436	6,970
competitive	19,400	24,166	6,542
less competitive	26,335	21,262	5,359
some or no selection, 4- year	18,981	16,638	5,119
private 2-year	14,852	17,822	6,796
public 2-year	7,573	10,543	4,991
for-profit 2-year	18,486	21,456	3,257

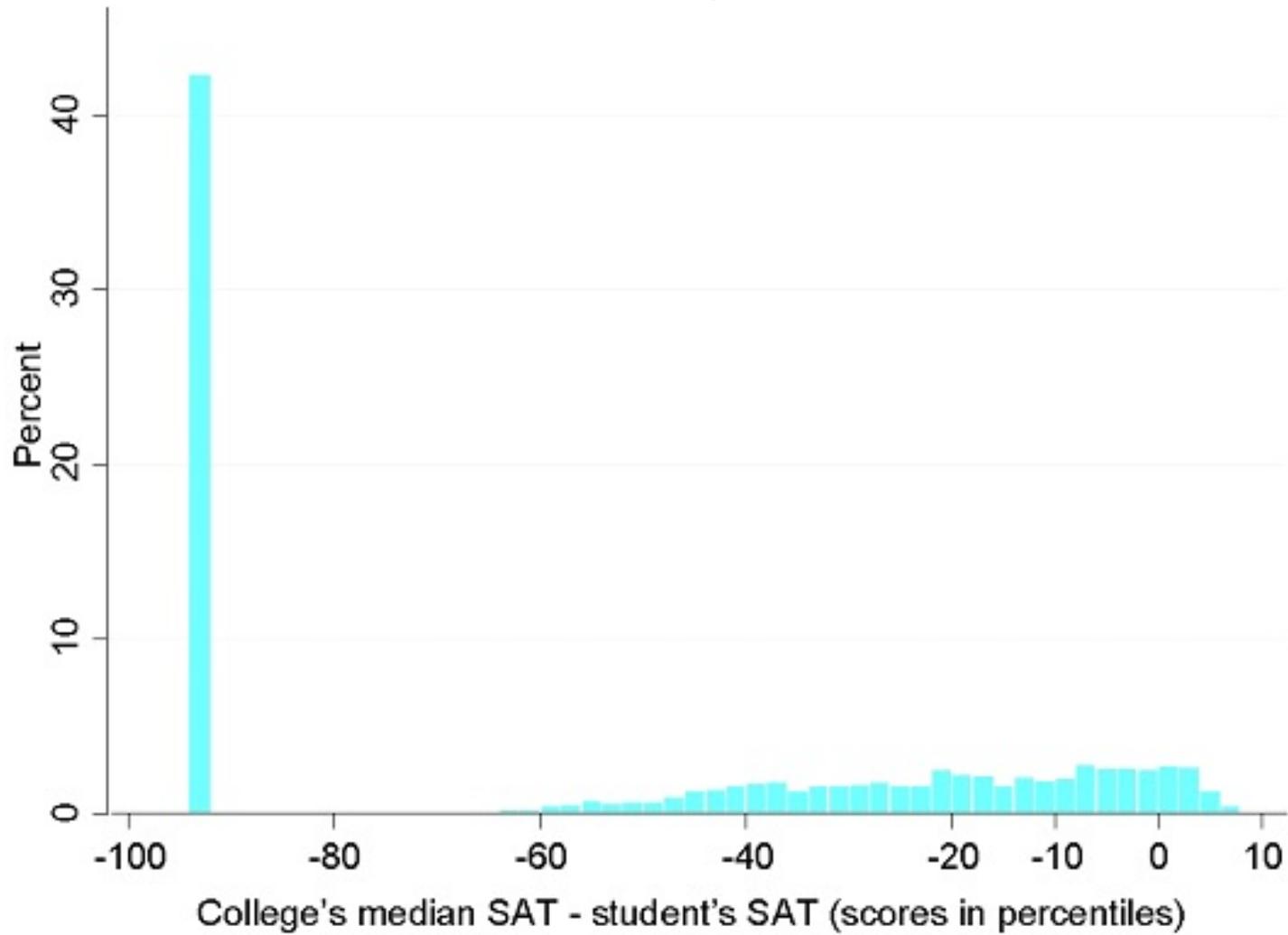
Notes: The sources are colleges' net cost calculators for the out-of-pocket cost column and IPEDS for the remaining columns. The net cost data were gathered for the 2009-10 school year by the authors, for the institutions at the very competitive and more selective levels. For the institutions of lower selectivity, net cost estimates are based on the institution's published net cost calculator for the year closest to 2009-10--never later than 2011-12. Net costs are then reduced to approximate 2009-10 levels using the institution's own room and board and tuition net of aid numbers from IPEDS, for the relevant years.

Figure 8
High Income Students' Portfolios of College Applications
(1 student = weight of 1)



Source: Hoxby, C. M., & Avery, C. 2012

Figure 10
Low Income Students' Portfolios of College Applications
(1 student = weight of 1)



Source: Hoxby, C. M., & Avery, C. 2012

Two Approaches to Improving Education

A) Individual-based interventions

- 1) Provide vouchers for K-12 schools
- 2) Provide subsidies and loans to individuals for college costs

B) Improving the production process

- 1) Charter schools
- 2) Direct improvements in education production function, e.g. teacher quality and personnel policies
- 3) Letting for-profit schools compete with existing public/non-profit schools

PUBLIC SCHOOLS AND VOUCHERS

K-12 education is provided for free in **public schools** in the US (funded by taxes)

If parents send child to private school, they have to pay private school tuition and do not get refunded for their taxes paid toward public school

⇒ Strong incentives to use public schools

Educational vouchers: A fixed amount of money given by the government to families with school-age children, who can spend it at any type of school, public or private.

A voucher effectively refunds parents for taxes paid if they do not use public schools ⇒ Puts public and private schools in competition

RATIONALES FOR VOUCHERS

(1) Consumer Sovereignty: Vouchers allow families to more closely match their educational choices with their tastes.

(2) Competition: Vouchers allow the education market to benefit from the competitive pressures that make private markets function efficiently.

PROBLEMS WITH EDUCATIONAL VOUCHERS

- 1) Vouchers Will Lead to Segregation:** Vouchers have the potential to reintroduce segregation along many dimensions, such as race, income, or child ability.
- 2) Vouchers Benefit kids from richer background:** The government would pay a portion of the private school costs that students and their families are currently paying themselves
- 3) The Education Market May not Be Competitive:** A large fraction of parents do not actively search the best possible school for their kids

Estimating the Effects of Voucher Programs

Rouse (1998) studied the effect of the Milwaukee voucher program on the achievement of students who used their vouchers to finance a move to private schools

1) She noted that one cannot directly compare students who do and do not use vouchers, since they may differ along many dimensions \Rightarrow This selective use of vouchers would bias any comparison between the groups.

2) Oversubscribed schools had to select randomly from all applicants, using a lottery (generates a quasi-experiment) \Rightarrow Comparing lottery winners to losers, finds slight improvement in math scores (no difference in reading)

In the United States, about 10% of students are enrolled in private schools, a proportion that doubles or triples in the low-income developing world

Angrist et al. 2002 shows that lottery voucher program had strong positive effects on education in Colombia

External validity issue: voucher lottery strategy estimates effects of vouchers on families motivated to use them (entered the lottery). Unmotivated parents might not be affected by vouchers.

Charter Schools

Some school districts have not offered vouchers for private schools but have instead allowed students to choose freely among public schools.

Charter schools: Schools financed with public funds that are not usually under the direct supervision of local school boards or subject to all state regulations for schools. Have more flexibility to recruit teachers / adjust hours / curriculum

Estimating the Effects of Charter Schools

Oversubscribed charter schools also use a lottery to assign admissions

Generates randomized experiment allowing to estimate the causal effect of charter schools by comparing lottery winners and lottery losers

Angrist, Pathak, Walters AEJ'13 carry out a comprehensive analysis of charter schools effects in Massachusetts

Find that urban charter schools boost achievement well beyond that of urban public school students, while non-urban charters reduce achievement from a higher baseline

⇒ Charter schools can have a positive or negative impact depending on what they do

Most effective approach to education: focus on instruction time, pupil comportment, selective teacher hiring, and focus on traditional math and reading skills.

School Accountability

Making schools accountable for student performance can provide incentives for schools to increase the quality of the education they offer.

Accountability programs can have two unintended effects:

- 1) they can lead schools and teachers to “teach to the test.”
- 2) schools can manipulate the pool of test takers and the conditions under which they take tests to maximize success.

No Child Left Behind (Key Bush administration program in education):

Evidence that it had small positive effects on test-scores but this could be due primarily to “teach to the test” effects

Unlikely to be re-authorized in its old form

MEASURING THE RETURNS TO EDUCATION

Returns to education: The benefits that accrue to society when students get more schooling or when they get schooling from a higher-quality environment.

Effects of education levels on productivity

There is a large literature that shows that more education leads to higher wages in the labor market:

$$Earnings_i = \alpha + \beta Education_i + \varepsilon_i$$

There is substantial controversy, however, over the implications of this correlation ($\beta > 0$).

Effects of Education on Earnings

1) Education as Human Capital Accumulation

human capital: A person's stock of skills, which may be increased by education

In that scenario, education raises earnings because it improves productivity

2) Education as a Screening Device

screening: A model that suggests that education provides only a means of separating high-ability from low-ability individuals and does not actually improve skills.

In that scenario, education raises individual earnings but it does not improve productivity (rat-race)

MEASURING THE RETURNS TO EDUCATION

Policy Implications

Under the human capital model, government would want to support education or at least provide loans to individuals so that they can get more education and raise their productivity.

Under the screening model, however, the government would *not* want to support more education for any given individual.

Differentiating the theories

Most of the returns to education reflect accumulation of human capital rather than screening

Clark-Martorell '14 show that barely getting a high-school degree in Texas has no visible impact on later earnings

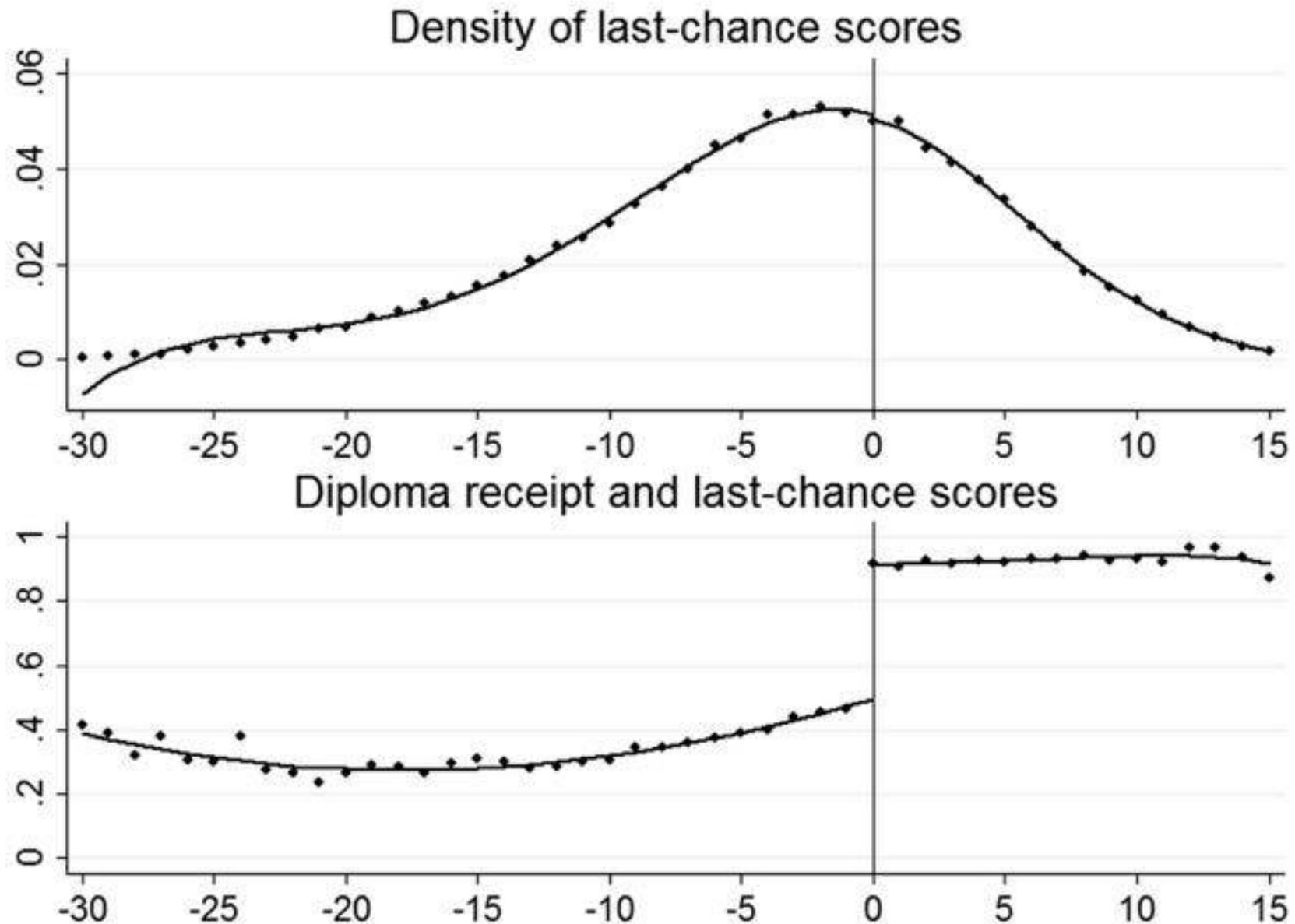


FIG. 1.—Last-chance exam scores and diploma receipt. The graphs are based on the last-chance sample. See table 1 and the text. Dots are test score cell means. The scores on the x -axis are the minimum of the section scores (recentered to be zero at the passing cutoff) that are taken in the last-chance exam. Lines are fourth-order polynomials fitted separately on either side of the passing threshold.

Source: Clark and Martorell JPE'14

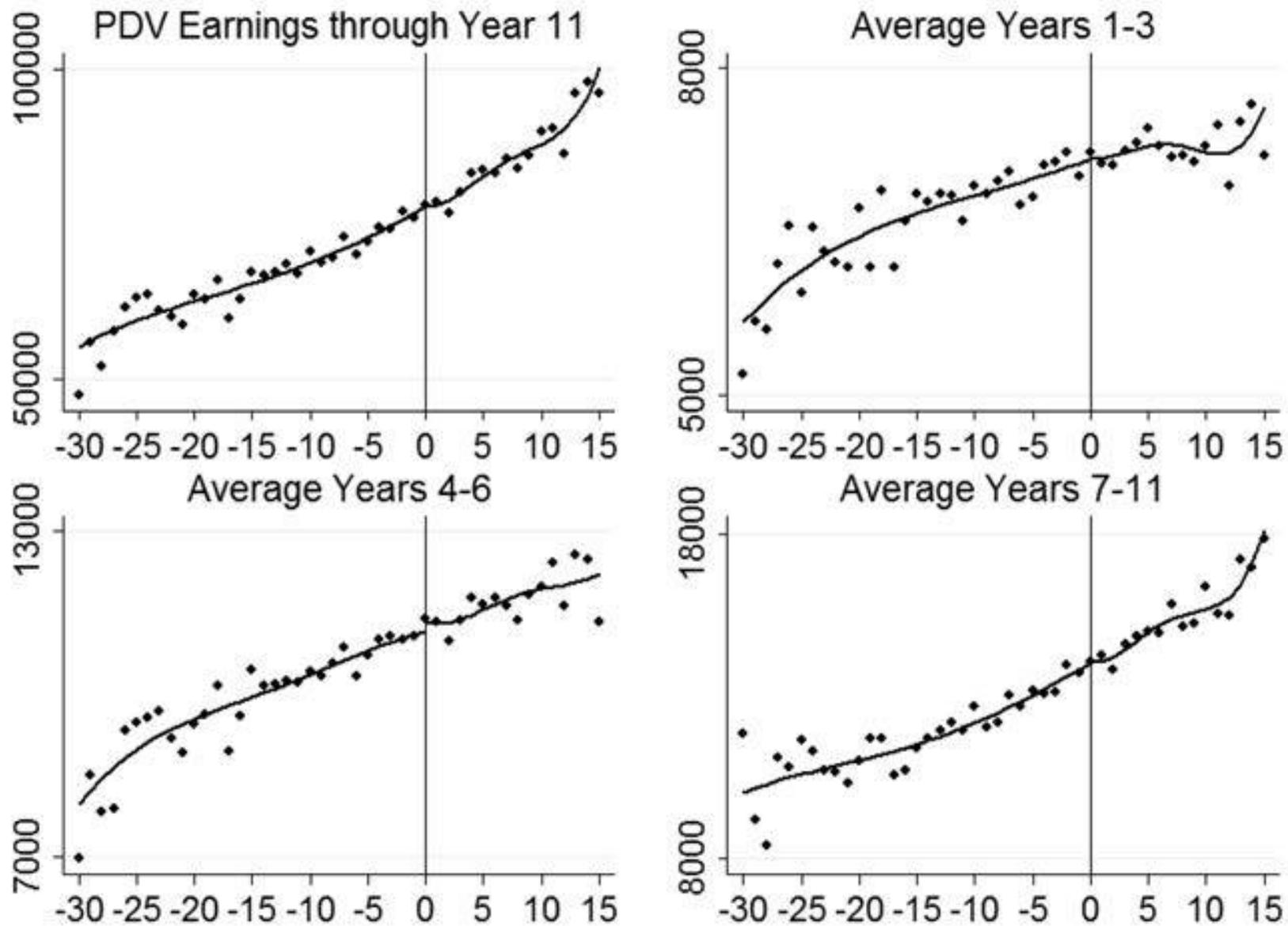


FIG. 2.—Earnings by last-chance exam scores. The graphs are based on the last-chance samples. See table 1 and the text. Dots are test score cell means. The scores on the x -axis are the minimum of the section scores (recentered to be zero at the passing cutoff) that are taken in the last-chance exam. Lines are fourth-order polynomials fitted separately on either side of the passing threshold.

Source: Clark and Martorell JPE'14

Evidence on the Returns to Education and Screening

Basic observational approach:

$$Earnings_i = \alpha + \beta Education_i + \varepsilon_i$$

Amounts to comparing the earnings of people with different education.

Issue: ability to earn ε_i might be correlated with education choices

Two methods try to control for this bias in estimating the true human capital effects of education

- 1) Control for underlying ability by adding variables (e.g. SAT score) in the regression so that any remaining effect of education represents true productivity effects (omitted variable bias remains a concern)
- 2) Find exogenous variation in education (e.g., policy change induces more education for some group but not for another group)

Although all of these approaches have some limitations, the result of the analysis is surprisingly consistent: each additional year of education raises wages by 7-10%

THE IMPACT OF SCHOOL QUALITY

A number of approaches have been taken to estimate the impact of school quality on student test scores.

Two approaches have been used to address this issue: experimental data, and quasi-experimental using policy changes

Findings suggest that the outcomes of efforts to improve school quality can be very dependent on the approach taken to improvements

Estimating the Effects of Class Size and Teacher Quality

The state of Tennessee implemented Project STAR in 1985, randomly assigning 11,000 students (grades K–3) to small classes (13–17 students) or regular classes (22–25 students)

Krueger and Whitmore 2001 shows positive effects of small class size on test scores

Chetty-Friedman-Hilger-Saez-Schanzenbach-Yagan 2011 linked students to college enrollment and adult earnings data: finds positive effect on college but (noisy) zero effect on adult earnings

Note: kids and teachers also randomly assigned across classes: strong class effects are visible (due to teachers or peers) and they have long-term effects on college and earnings

Figure 2a: Effect of Class Size on College Attendance by Year

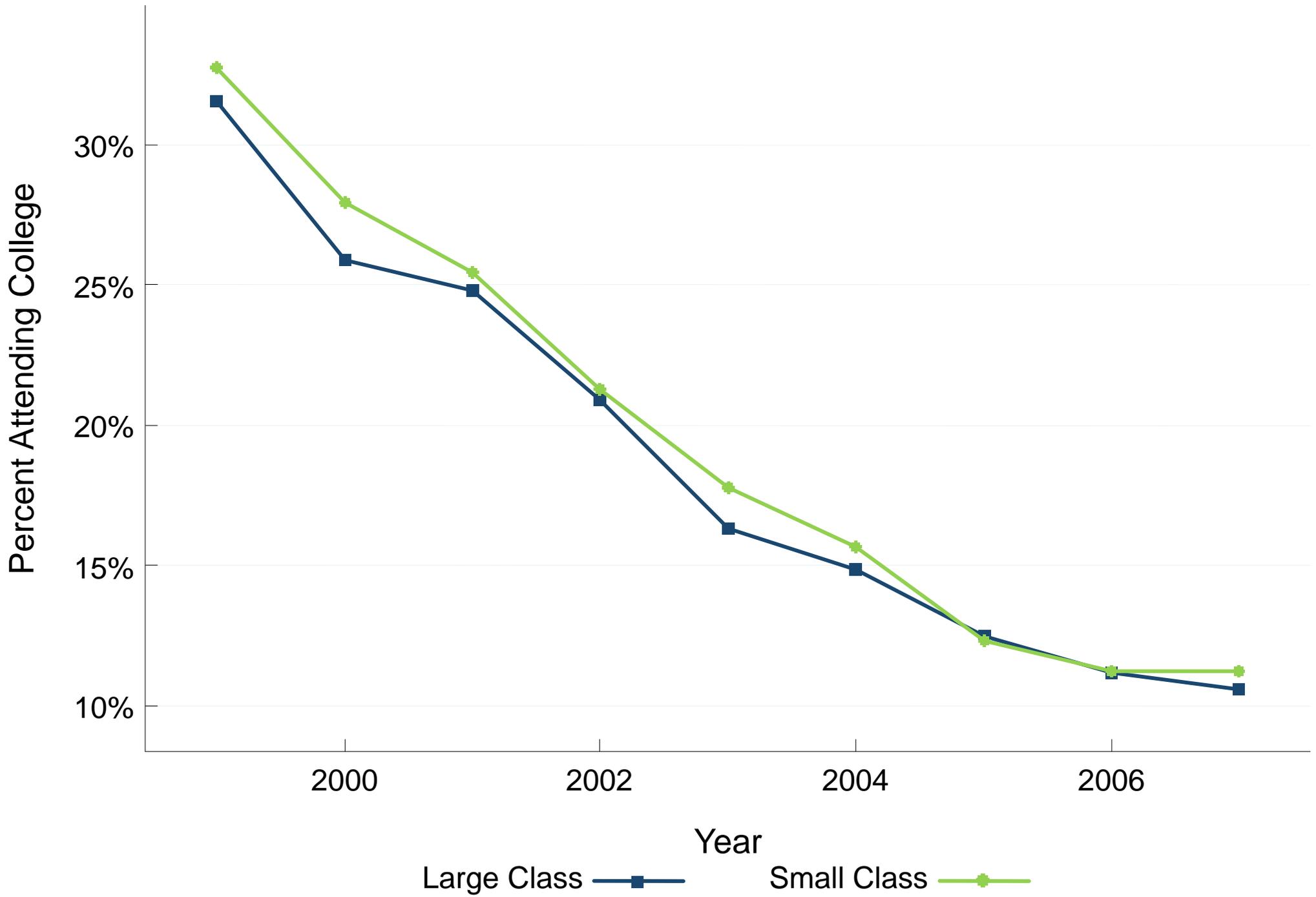


Figure 2c: Effect of Class Size on Wage Earnings by Year

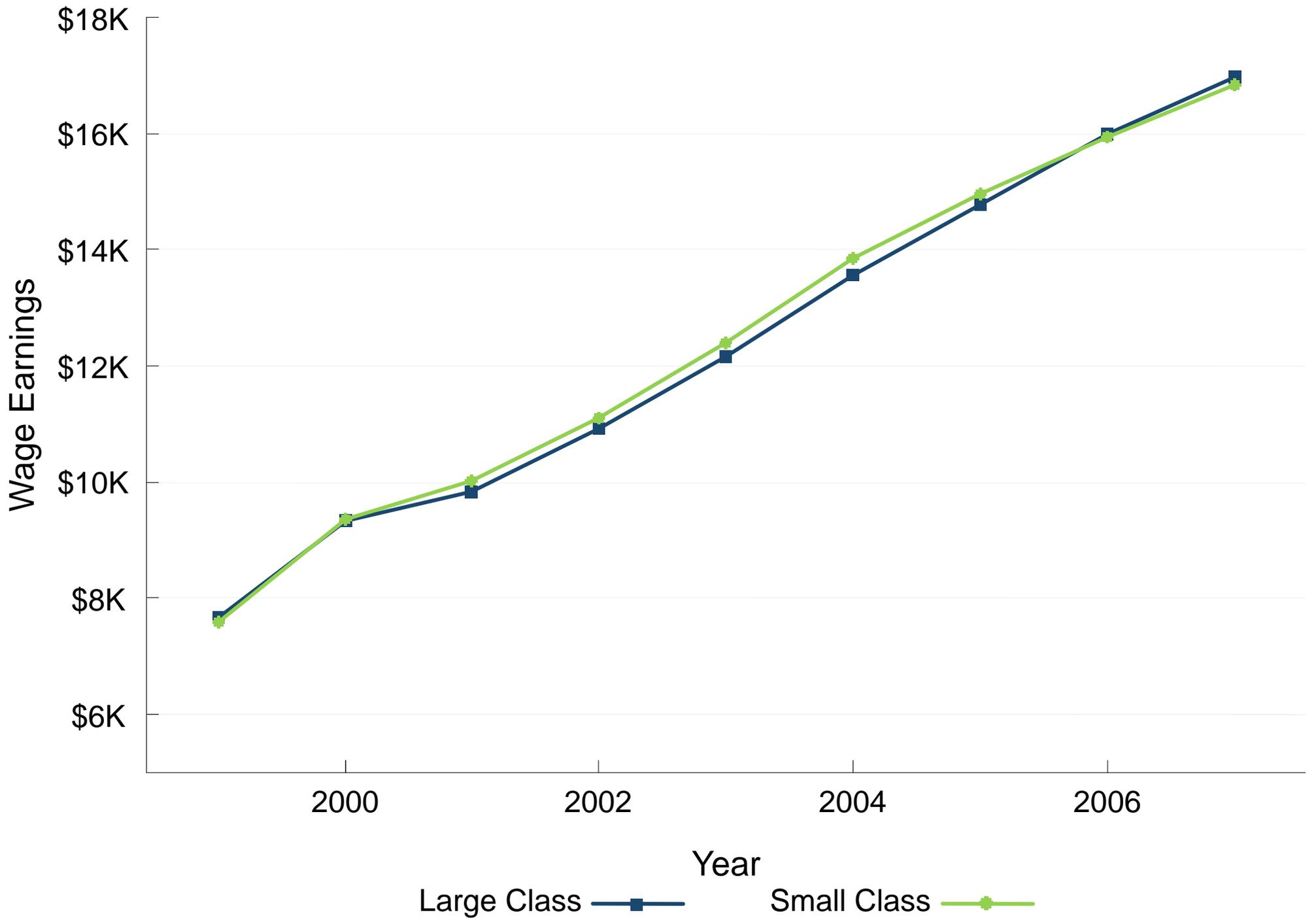
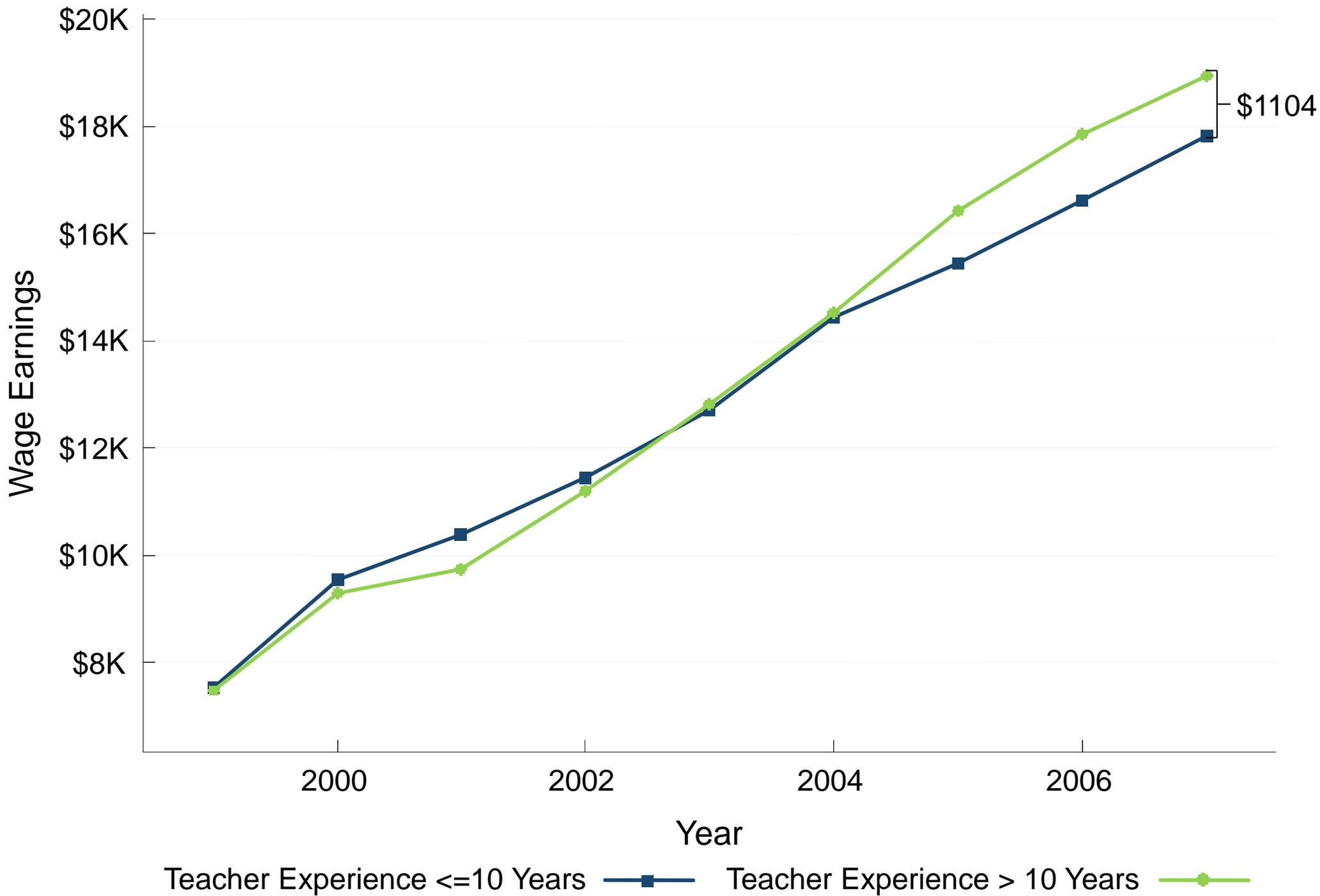


Figure 3c: Effect of Teacher Experience on Earnings by Year



Current Government Role in Higher Education

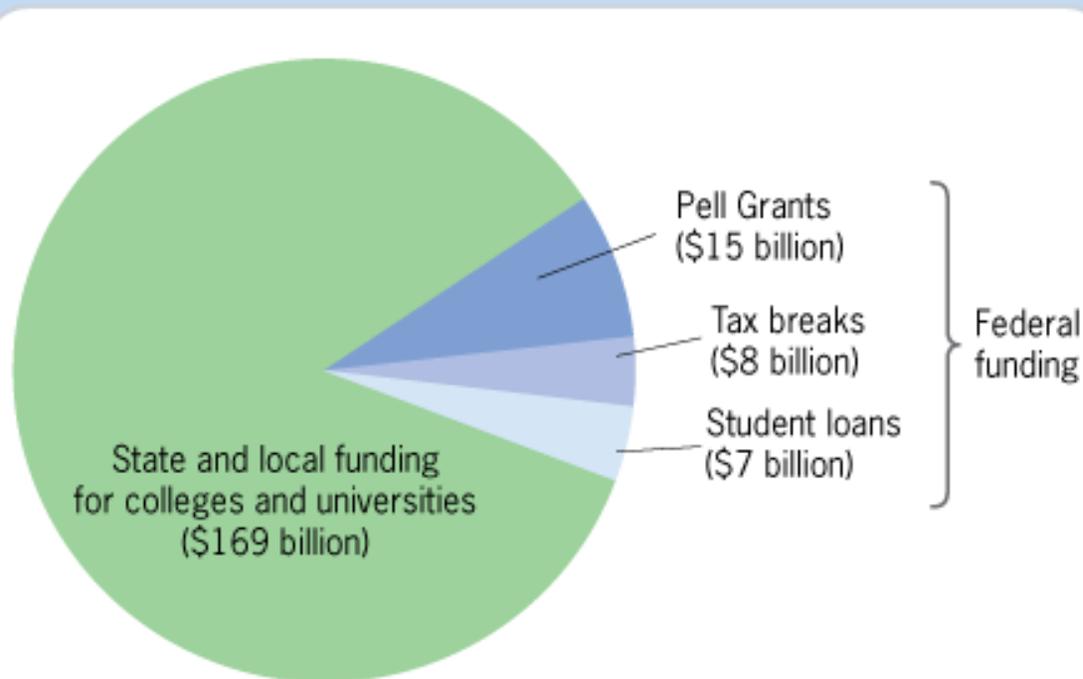
- 1. State Provision:** The primary form of government financing of higher education is direct provision of higher education through locally and state-supported colleges and universities.
- 2. Pell Grants:** Subsidy to higher education administered by the federal government that provides grants to low-income families to pay for their educational expenditures.
- 3. Loans:** **(a) direct student loans:** Loans taken directly from the Department of Education. **(b) guaranteed student loans:** Loans taken from private banks for which the banks are guaranteed repayment by the govt.
- 4. Tax Relief:** Tax credits for higher education tuition costs

11.5

The Role of the Government in Higher Education

Current Government Role

■ FIGURE 11-4



Government Spending on Higher Education • Eighty-five percent of the roughly \$199 billion the government spends annually on higher education is in the form of state and local funding for colleges and universities. The remainder is split among Pell Grants, tax breaks, and student loans.

What Is the Market Failure in Higher Education?

If individuals are rational, the borrowing constraint market failure can be addressed solely with government supported loans

However, if individuals are not rational (self-control problems, myopia, lack of information), even government supported loans might not be enough to motivate individuals to acquire higher education

⇒ Direct tuition subsidies might be more effective

⇒ Direct help with applications

Effects of tuition aid on attending college

Dynarski (2003) studies elimination of Social Security Administration (SSA) program to provide tuition aid to students with parents deceased or disabled SSA beneficiaries in 1982

DD analysis: Compare college attendance of kids with deceased father (Treatment group) to kids with father alive (control group), before 1982 vs. after 1982

Finds very large 20 percentage point impact of program on college attendance \Rightarrow Small changes in tuition costs have dramatic impact suggesting that borrowing constraints matter

EFFECT OF SSA COLLEGE AID ON PROBABILITY OF ATTENDING COLLEGE

TABLE 2—OLS, EFFECT OF ELIGIBILITY FOR
STUDENT BENEFITS ON PROBABILITY
OF ATTENDING COLLEGE BY AGE 23

	(1) Difference- in-differences	(2) Add covariates
Deceased father × before	0.182 (0.096)	0.219 (0.102)
Deceased father	−0.123 (0.083)	Y
Before	0.026 (0.021)	Y

Effects of cash allowance on attending college in France

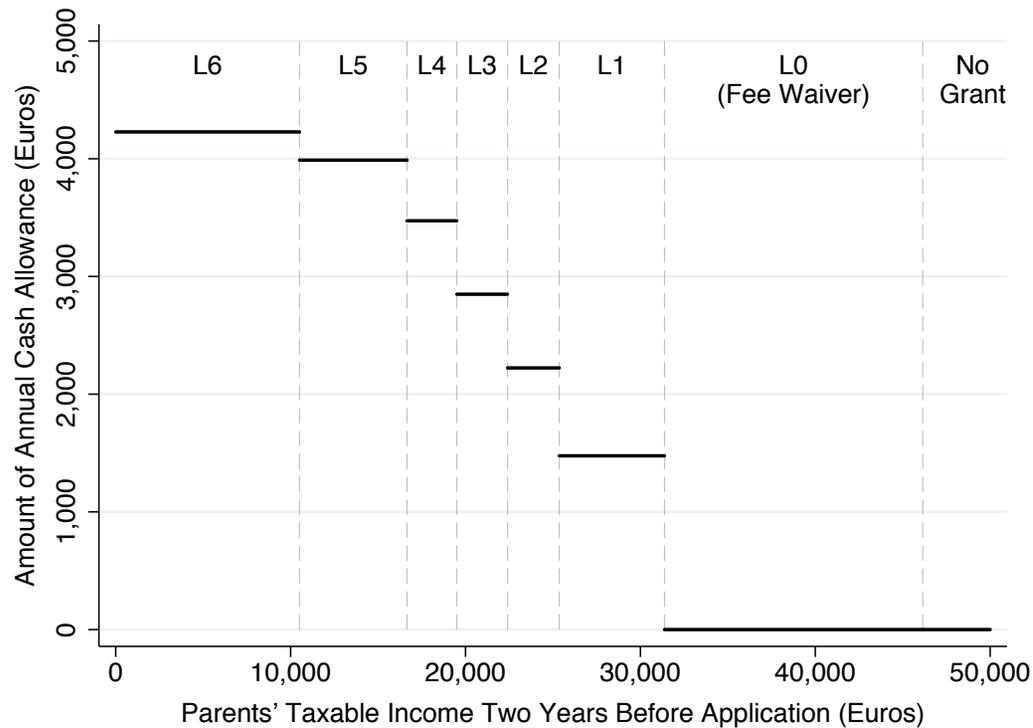
Fack and Grenet (2014) study the effects of aid to students based on parental income in France

Level of aid is a discontinuous function of parental income

Regression discontinuity design: does the discontinuity in aid translate into a discontinuity in college attendance? YES

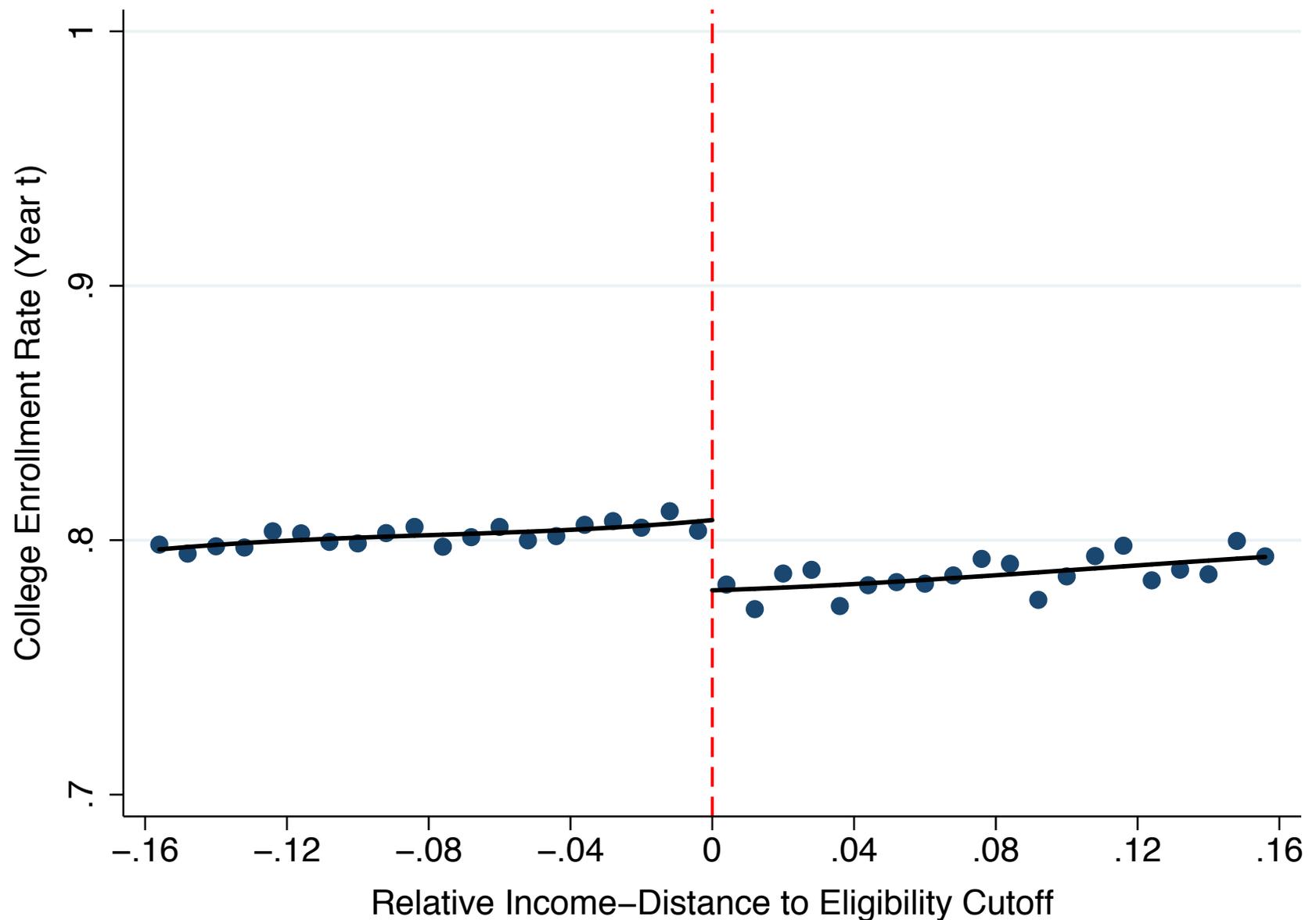
⇒ Very compelling evidence that financial aid for higher education matters

Figure 2: Amount of Annual Cash Allowance Awarded to Applicants with an FNA Score of 3 Points, as Function of their Parents' Taxable Income



Notes: The figure shows the amount of annual cash allowance awarded in 2009 to BCS grant applicants with a family needs assessment (FNA) score of 3 points (median value), as a function of their parents' taxable income two years before the application. Applicants eligible for a level 0 grant qualify for fee waivers only. Applicants eligible for higher levels of grant qualify for fee waivers and an annual cash allowance, the amount of which varies with the level of grant: 1,476 euros (level 1), 2,223 euros (level 2), 2,849 euros (level 3), 3,473 euros (level 4), 3,988 euros (level 5) and 4,228 euros (level 6). Income thresholds and allowance amounts are expressed in 2011 euros.

(b) €1,500 Allowance (L1/L0 Cutoffs)



Source: Fack and Grenet (2014)

Effects of college application tutoring

Carrell-Sacerdote NBER'13 carry out a field experiment in New Hampshire high-schools

College students from Dartmouth help senior high-schoolers to apply to college (weekly meetings in Winter semester)

Randomization within high schools: select only 50% of seniors

Find large positive impact on women (+15 points likelihood of enrolling in college) but no effects on men

Also find a cash bonus for applying to colleges without tutorial does not have any impact

Role of Government in supply of Higher Education

Private non-profit universities have inelastic supply (e.g., fixed student bodies at top schools such as Harvard)

Historically, expansion of supply was carried out by public institutions (state universities and community colleges): Example: 1960 Master plan for California with 3-tier system (Community, State, UC)

Government push also central to increase attendance: GI Bill after WWII/Korea War increased college attendance by 15-20 points for men born 1921-1933 (Stanley QJE'03)

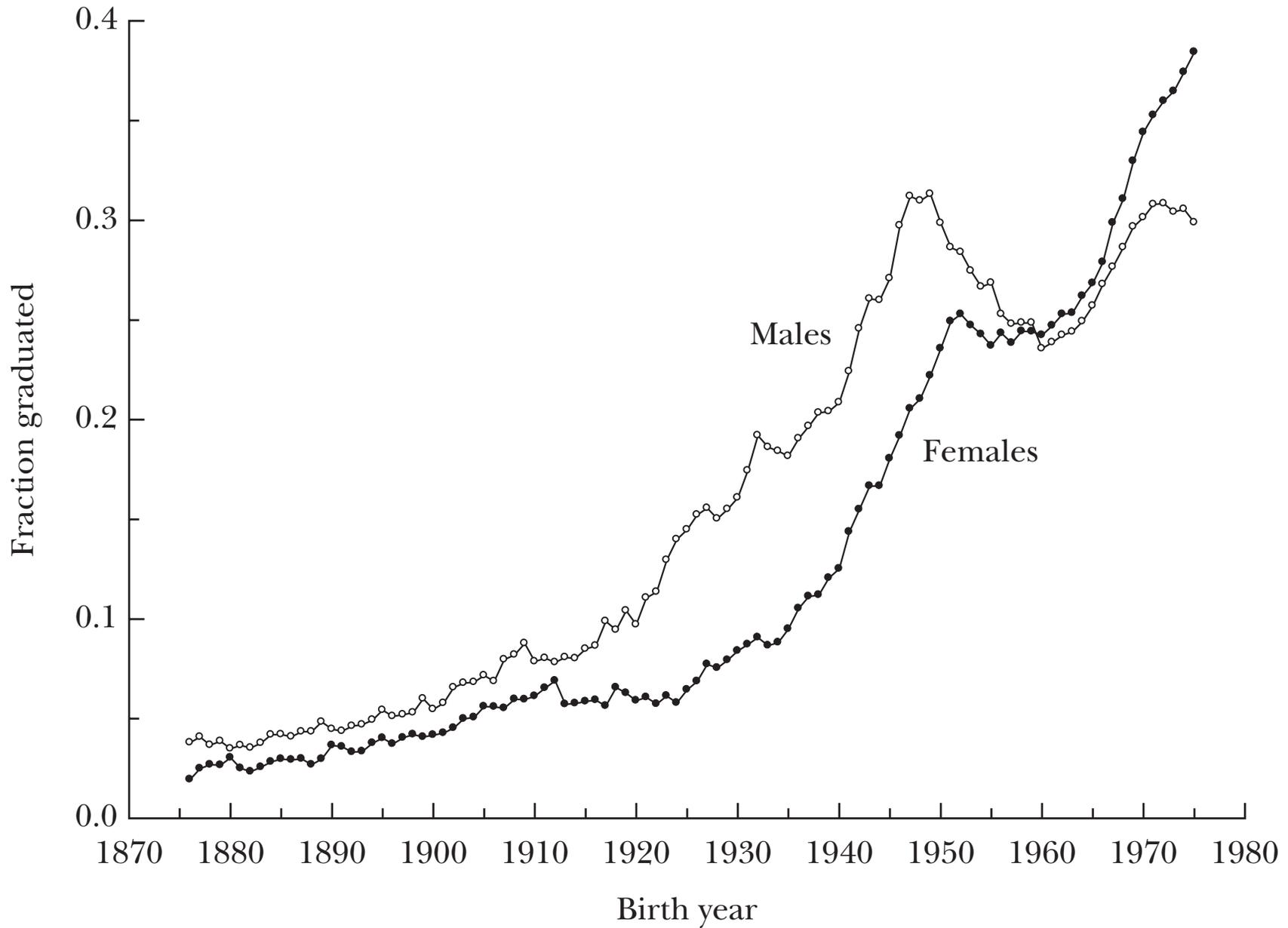
Recently, states have retreated and supply has been provided by for-profit schools (get about 10% of total enrollment today)

Deming-Goldin-Katz '12 show that for-profit schools provide little benefits, charge a lot, and are savvy at exploiting Fed Pell Grants and saddle students with debt

⇒ Symptom of market failure due to individual failures/lack of information

Figure 1

College Graduation Rates (by 35 years) for Men and Women: Cohorts Born from 1876 to 1975



Sources: 1940 to 2000 Census of Population Integrated Public Use Micro-data Samples (IPUMS).

REFERENCES

Jonathan Gruber, *Public Finance and Public Policy*, Fourth Edition, 2012
Worth Publishers, Chapter 11

Angrist, Joshua, et al. "Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment." *The American Economic Review* (2002).(web)

Angrist, Joshua, Parag A. Pathak, and Christopher R. Walters. "Explaining Charter School Effectiveness." *American Economic Journal: Applied Economics*, 5.4 (2013): 1-27.(web)

Carrell, Scott E., and Bruce Sacerdote. "Late Interventions Matter Too: The Case of College Coaching New Hampshire." No. w19031. National Bureau of Economic Research, 2013.(web)

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. "How does your kindergarten classroom affect your earnings? Evidence from Project STAR." *The Quarterly Journal of Economics* 126, no. 4 (2011): 1593-1660.(web)

Clark, Damon and Paco Martorell "The Signaling Value of a High School Diploma" *Journal of Political Economy* 122(2), 282-318. (web)

Deming, David, Claudia Goldin, and Lawrence F. Katz. "The For-Profit Postsecondary School Sector: Nimble Critters or Agile Predators?" *Journal of Economic Perspectives* (2012), vol. 26(1), pages 139-64. (web)

Duflo, Esther, "Grandmothers and Granddaughters: Old?Age Pensions and Intrahousehold Allocation in South Africa" *World Bank Economic Review*, 17(1) (2003): 1–25.(web)

Dynarski, Susan M. "Does Aid Matter? Measuring the Effect of Student Aid on College Attendance and Completion." *The American Economic Review* 93.1 (2003): 279-288.(web)

Fack, Gabrielle, and Julien Grenet, "Improving College Access and Success for Low-Income Students: Evidence from a Large French Need-based Grant Program," *American Economic Journal: Applied Economics*, 2014. (web)

Goldin, Claudia, Lawrence Katz, Ilyana Kuziemko "The Homecoming of American College Women: The Reversal of the College Gender Gap", *Journal of Economic Perspectives*, 2006. (web)

Hoxby, Caroline, and Christopher Avery. "The Missing"One-Offs": The Hidden Supply of High-Achieving, Low Income Students." No. w18586. National Bureau of Economic Research, 2012.(web)

Hoxby, Caroline, and Sarah Turner. "Expanding College Opportunities for High-Achieving, Low Income Students." SIEPR Discussion Paper No. 12-014 (2013).(web)

Jensen, Robert. "The (perceived) returns to education and the demand for schooling." *The Quarterly Journal of Economics* 125.2 (2010): 515-548.(web)

Krueger, Alan B., and Diane M. Whitmore. "The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from Project STAR." *The Economic Journal* 111.468 (2001): 1-28.web

Lochner, Lance, and Enrico Moretti. "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports." *American Economic Review* 94.1 (2004): 155-189.(web)

Milligan, Kevin, Enrico Moretti, and Philip Oreopoulos. "Does education improve citizenship? Evidence from the United States and the United Kingdom." *Journal of Public Economics* 88.9 (2004): 1667-1695.(web)

Rouse, Cecilia Elena. "Private school vouchers and student achievement: An evaluation of the Milwaukee parental choice program." *The Quarterly Journal of Economics* 113.2 (1998): 553-602.(web)

Solis, Alex "Credit Access and College Enrollment", Working Paper, Uppsala University, 2013 (web)

Stanley, Marcus. "College education and the midcentury GI Bills." *The Quarterly Journal of Economics* 118.2 (2003): 671-708.(web)

Undergraduate Public Economics

Danny Yagan
UC Berkeley

Lecture 16

Social Insurance: The New Function of
Government

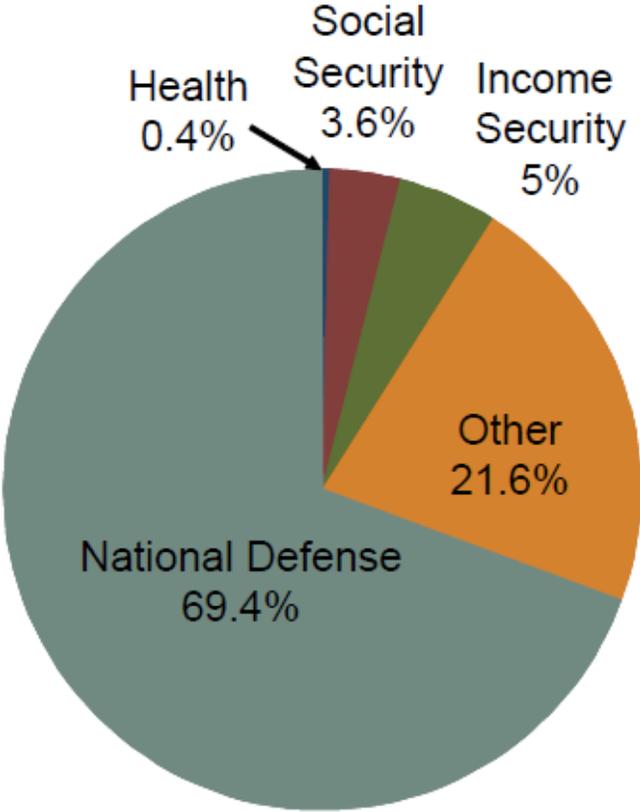
DEFINITION

Social insurance programs: Government interventions in the provision of insurance against adverse events:

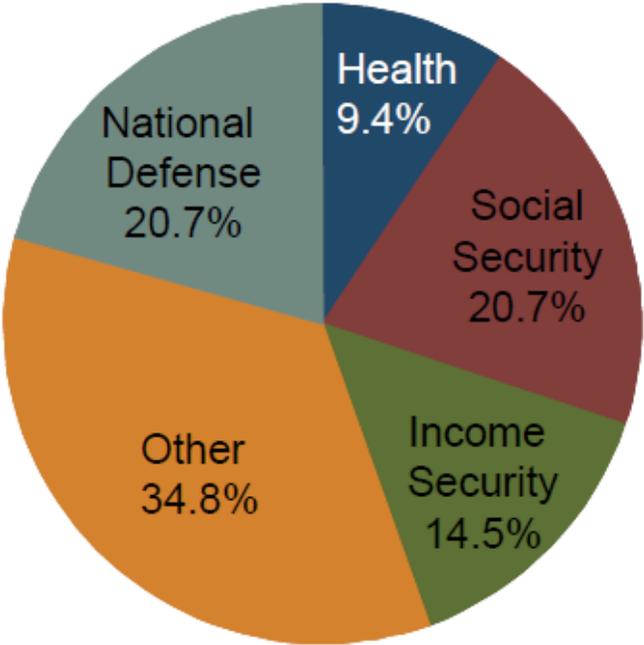
Examples: (a) health insurance (Medicaid, Medicare), (b) retirement and disability insurance (Social Security), (c) unemployment insurance

Growth in government over the 20th century is mostly due to the growth of social insurance (health and retirement benefits)

Motivation: Modern Governments do Social Insurance



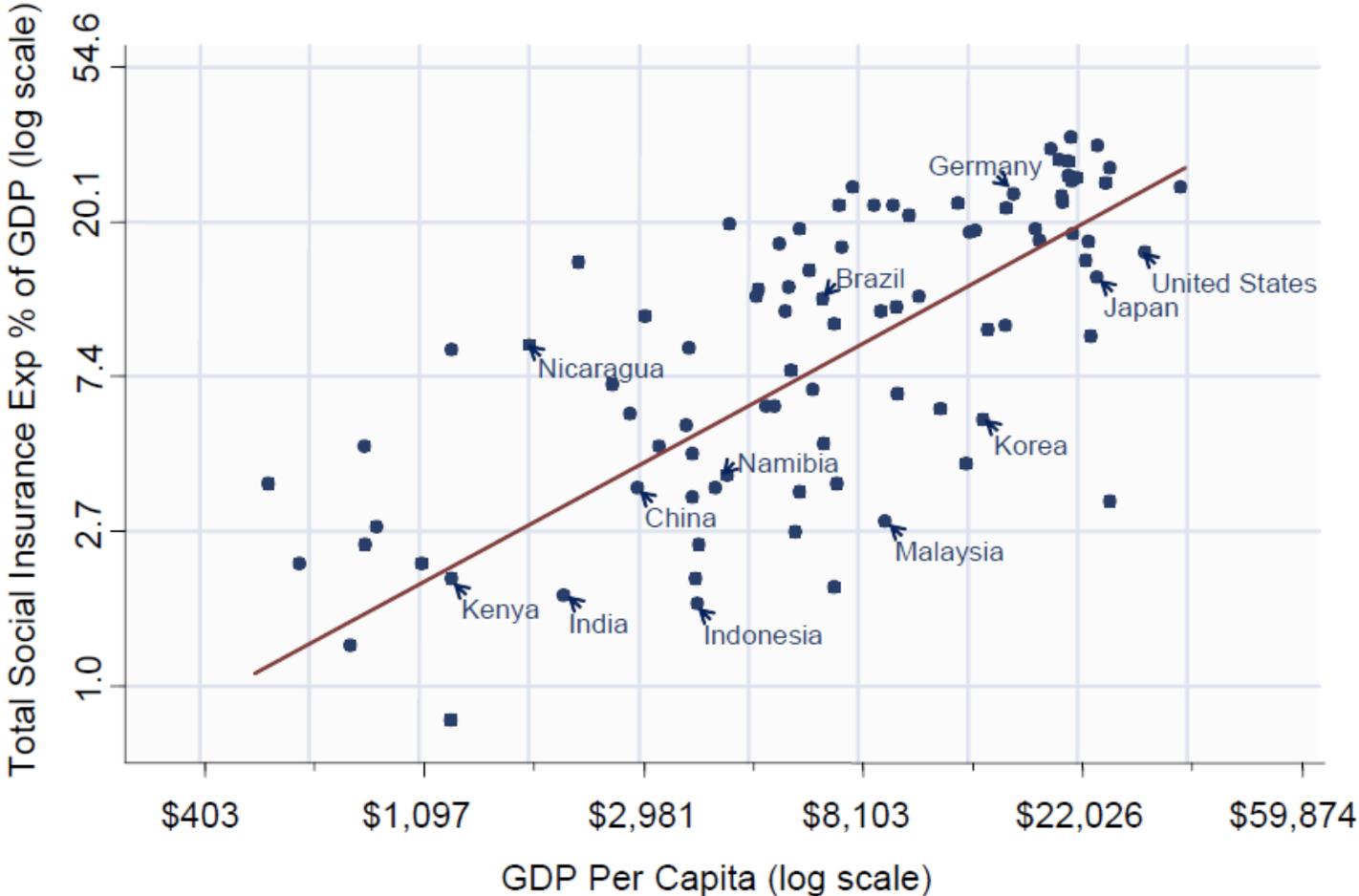
1953



2008

Source: Chetty; Office of Management and Budget, historical tables

Motivation: Modern Governments do Social Insurance



Source: Chetty-Finkelstein (2013) (1996 data)

What Is Insurance?

Insurance premiums: Money that is paid to an insurer so that an individual will be insured against adverse events.

A sampling of private insurance products that exist in the United States includes:

- Health insurance
- Auto insurance
- Life insurance
- Casualty and property insurance

EXPECTED UTILITY MODEL

Utility function $U(c)$ increasing in consumption c and concave in consumption c : $U'(c) > 0$ and $U''(c) < 0$

Expected utility model: Individuals want to maximize expected utility defined as the weighted sum of utilities across states of the world, where the weights are the probabilities of each state occurring.

If q is probability of adverse event, expected utility is written as:

$$EU = (1-q) * U(\text{consumption with no adverse event}) + q * U(\text{consumption with adverse event})$$

Actuarially fair premium: Insurance premium that is set equal to the insurer's expected payout.

EXPECTED UTILITY MODEL

Person has income W (regardless of health)

Person is sick with probability q

If sick, person incurs medical cost d to get better

Insurance contract: pay premium p always, and receive payout b only if sick

Expected utility:

$$EU = (1 - q)U(W - p) + qU(W - p - d + b)$$

Expected profits of insurers: $EP = p - qb$

Competition among insurers $EP = 0 \Rightarrow b = p/q$

This is called **actuarially fair** insurance

EXPECTED UTILITY MODEL

Individual chooses the level of premiums p to maximize:

$$EU = (1 - q)U(W - p) + qU(W - d - p + p/q)$$

First order condition:

$$0 = dEU/dp = -(1 - q)U'(W - p) + q[-1 + 1/q]U'(W - d - p + p/q)$$

$$\Rightarrow U'(W - p) = U'(W - d - p + p/q)$$

$\Rightarrow W - p = W - d - p + p/q$ (because U is concave and hence U' is strictly decreasing and hence invertible)

$$\Rightarrow 0 = -d + p/q \Rightarrow p = d \cdot q$$

This implies that the person is perfectly insured: consumption is the same in both states and equal to $W - d \cdot q$

Intuition: with concave utility, marginal utility decreases and it is always desirable to reduce consumption in high income states to increase consumption in low income states

Introducing heterogeneity in risk across individuals

Suppose now that there are two types of individuals: sickly and healthy
Sickly have $q = q_S$ and Healthy have $q = q_H$ with $q_S > q_H$

First scenario: Symmetric Information: Insurance companies and individuals can observe q_H vs. q_S types (for example, could be age status)

Then insurance companies will charge 2 policies, each actuarially fair:

$p_S, b_S = p_S/q_S$ for the sickly

$p_H, b_H = p_H/q_H$ for the healthy

Each type will still choose to buy perfect insurance $b_S = b_H = d$ and
 $p_S = q_S d, p_H = q_H d$

Sickly always consume $W - q_S d$

Healthy always consume $W - q_H d$

Private insurance does not equalize incomes across types only within types

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

What if $W - q_S d < 0$? Sickly person cannot afford insurance and dies (or starves) if sick

Introducing heterogeneity in risk across individuals

Second scenario: Asymmetric Information: Insurance companies cannot observe q_H vs. q_S types but individuals do (for example, diet or family health history)

If insurance companies charge the same two policies as before

$p_S = q_S d, b_S = d$ for the sickly

$p_H = q_H d, b_H = d$ for the healthy

Then everybody wants to buy the healthy insurance which is cheaper \Rightarrow Insurance company will make losses \Rightarrow cannot be an equilibrium [this is called **Adverse Selection**]

Two equilibrium possibilities:

1) Pooling equilibrium: Insurance companies offer a contract based on average risk [good deal for sickly, mediocre deal for healthy but better than no insurance]

2) Separating equilibrium: Insurance companies offer two contracts: one expensive contract with full insurance for the sickly, one cheap contract with partial insurance for the healthy: each type self-select into its contract \Rightarrow Outcome not efficient as healthy as under-insured

Adverse Selection

Adverse selection is when individuals know more about their risk level than the insurer and hence individuals with higher risk are more likely to purchase insurance.

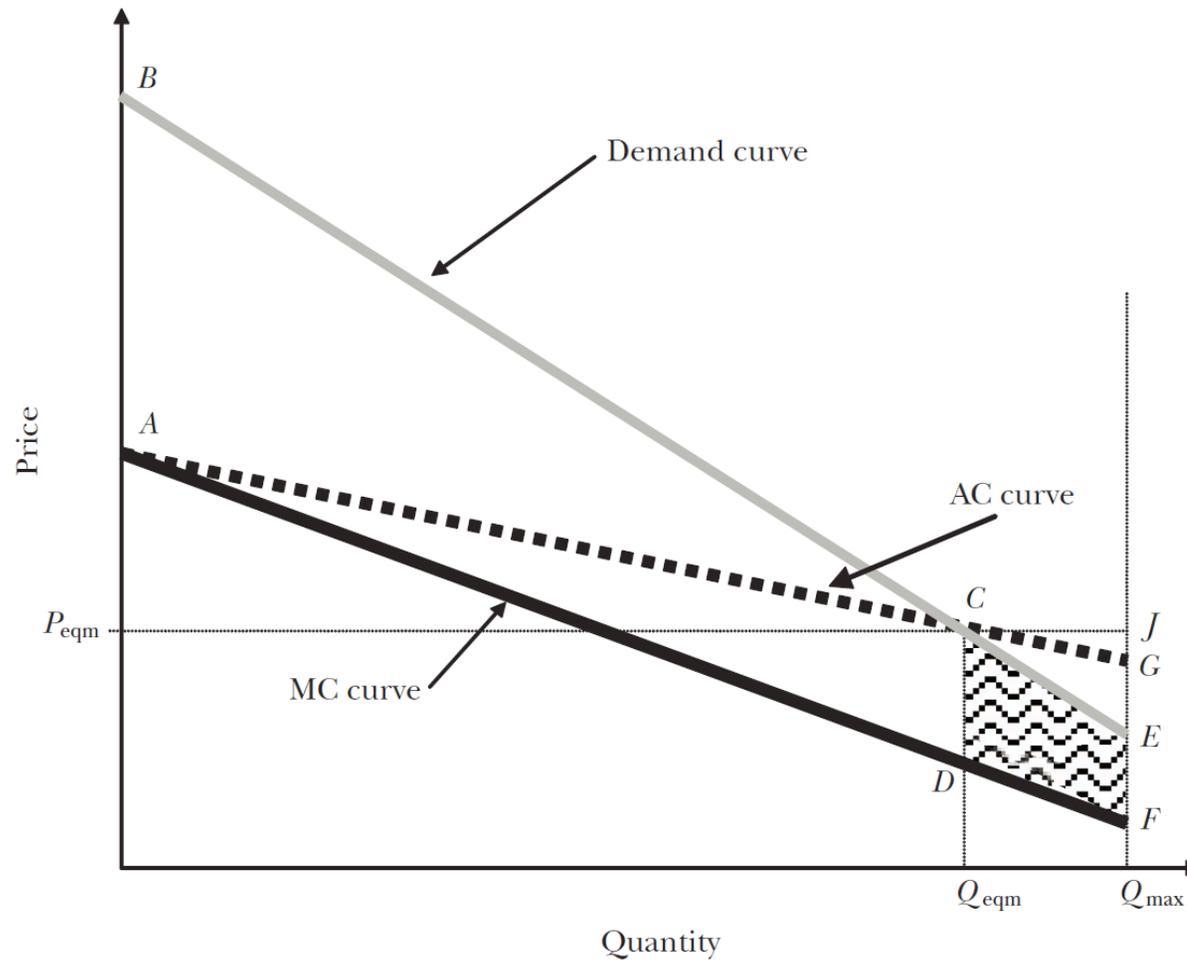
Example: people with high risk of getting sick more likely to buy health insurance than people with low risk of getting sick (if insurers cannot discriminate)

With adverse selection, market of health insurance can unravel (“death spiral”):

Health insurance is offered at average fair price, bad deal for low risk people and hence only high risk people buy it \Rightarrow insurers make losses \Rightarrow insurers raise the price further \Rightarrow only very high risk people buy it \Rightarrow insurers make losses again \Rightarrow no insurance contract is offered at all even though everybody wants full actuarially fair insurance

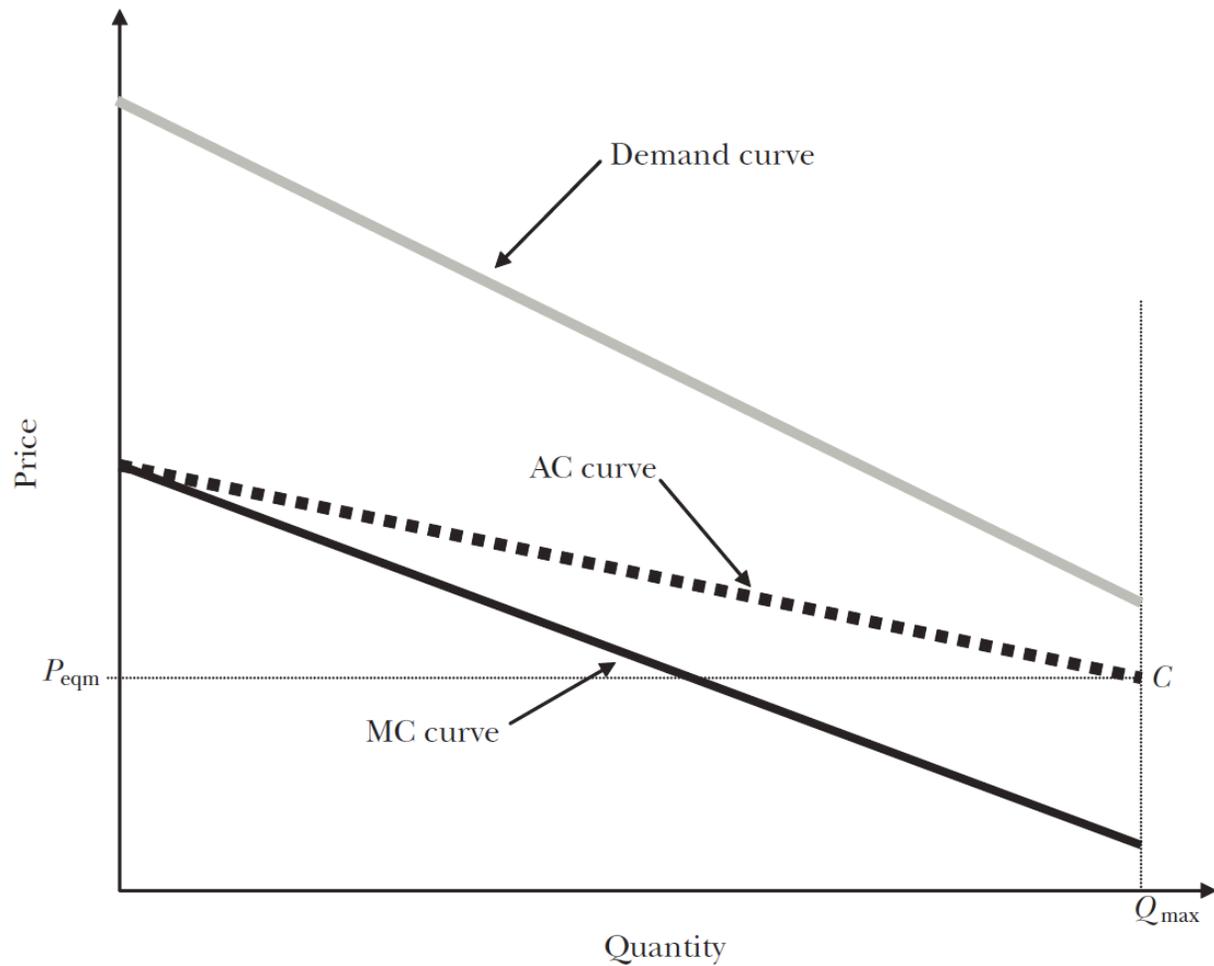
This inefficiency (market failure) arises because of **asymmetric information**

Akerlof (1970): Some insured, some not (efficiency loss)



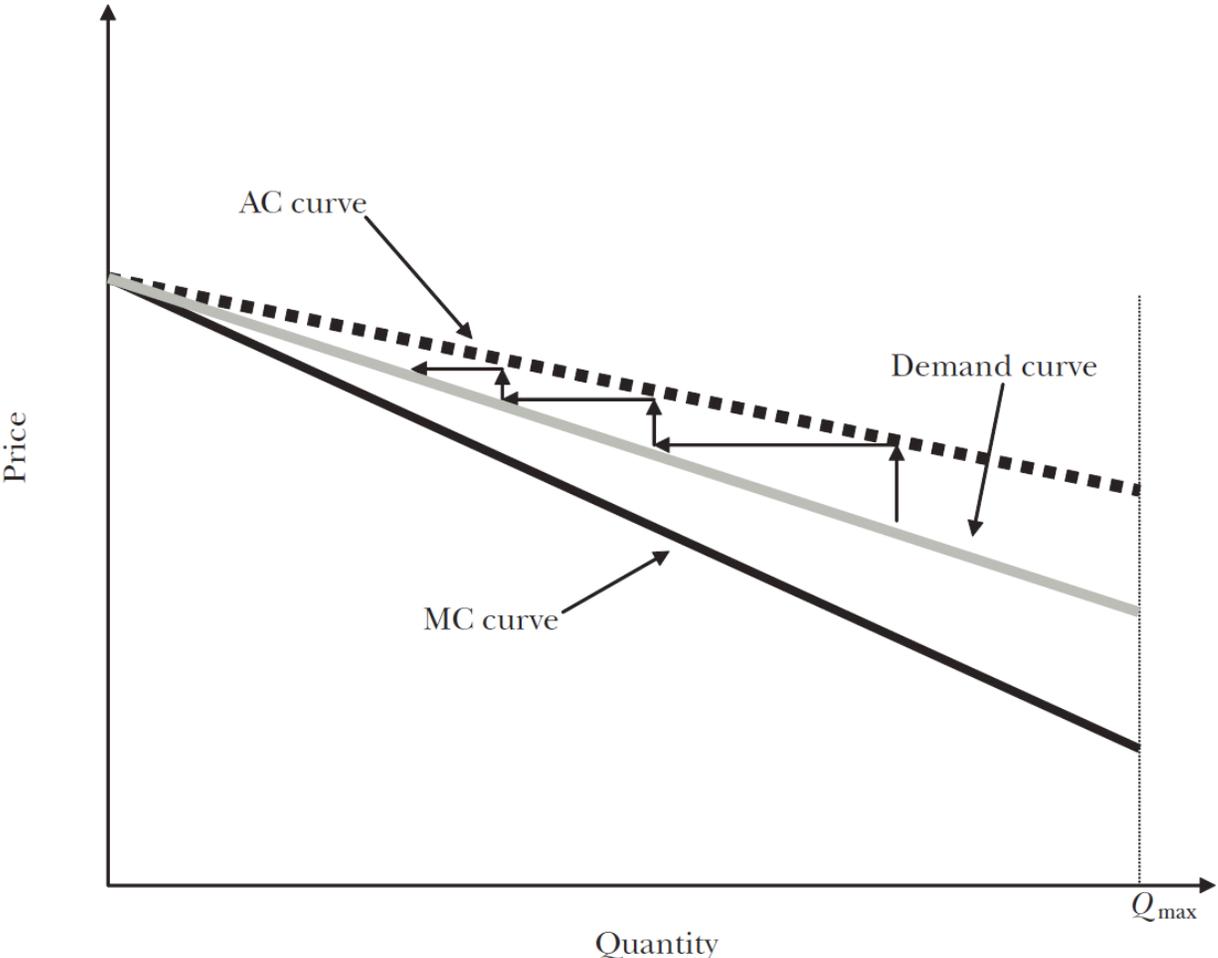
Source: Einav-Finkelstein (2011)

Akerlof (1970): Everyone insured (no efficiency loss)



Source: Einav-Finkelstein (2011)

Akerlof (1970): No one insured (large efficiency loss)



Source: Einav-Finkelstein (2011)

How Does the Government Address Adverse Selection?

The government can address adverse selection and improve market efficiency but this involves redistribution

Natural solution is to impose a **mandate**: everybody is required to purchase insurance \Rightarrow If price is the same for everybody, the low risk end up subsidizing the high risks

From a social perspective, being high risk (e.g. having a sickly constitution) is rarely consequence of individual choices \Rightarrow Society might want to compensate individuals for this

\Rightarrow Explains why all OECD countries (except US until Obamacare) have adopted universal health insurance

Obamacare (a) forbids insurers from charging based on pre-existing conditions, (b) mandates that everybody needs to get insurance, (c) subsidizes health insurance for low income families

WHY SOCIAL INSURANCE: OTHER REASONS

Redistribution: Private insurers cannot provide insurance against pre-existing conditions so those with high risk have to pay more: society may want to compensate high risk people (as being high risk is often not the fault of the person)

⇒ Universal health insurance funded by taxation effectively redistributes from low-risk people to high-risk people

Externalities

Your lack of insurance can be a cause of illness for me, thereby exerting a negative physical externality.

Example: flu shots protect the individual who gets it from the flu but indirectly protects others (as the flu is very contagious)

WHY SOCIAL INSURANCE: OTHER REASONS

Individual Failures

Individuals may not appropriately insure themselves against risks if the government does not force them to do so (myopia, lack of information, self-control problems)

If individuals understand their own failures, they will support social insurance (e.g., Medicare Health Insurance for elderly is very popular)

If individuals really want to be myopic, they will oppose govt social insurance (paternalism)

Administrative Costs

The administrative costs for Medicare are less than 2% of claims paid. Administrative costs for private insurance average about 12% of claims paid.

High administrative costs arise because private insurers try to screen away sickly customers and steal healthy customers from competitors. Individuals may also not understand well products and hence be sensitive to flashy advertisements.

MORAL HAZARD

Moral hazard: Adverse actions taken by insured individuals in response to insurance against adverse outcomes.

Example: If you receive unemployment benefits replacing lost wages, you may not search as much for a new job \Rightarrow Insurance reduces incentives to remedy adverse events

Moral Hazard exists with both private and social insurance as long as insurer cannot perfectly monitor the person insured \Rightarrow Insurers do not offer perfect insurance

The existence of moral hazard problems creates the *central trade-off of social insurance*: insurance is desirable for consumption smoothing but insurance can create moral hazard

[similar to the problem of optimal income taxation equity-efficiency trade-off]

MORAL HAZARD

What Determines Moral Hazard?

- How hard it is to observe whether the adverse event has happened
- How easy it is to change behavior in get into or stay in the adverse event

Moral Hazard Is Multidimensional: In examining the effects of insurance, three types of moral hazard play a particularly important role:

- 1) Reduced precaution against entering the adverse state (example: auto insurance)
- 2) Increased odds of staying in the adverse state (example: unemployment insurance)
- 3) Increased expenditures when in the adverse state (example: health insurance)

⇒ Moral hazard increases the cost of providing insurance

PUTTING IT ALL TOGETHER: OPTIMAL SOCIAL INSURANCE

Optimal social insurance trades-off two considerations:

1) The benefit of social insurance is the amount of consumption smoothing provided by social insurance programs

2) The cost of social insurance is the moral hazard caused by insuring against adverse events

⇒ Optimal social insurance systems should partially, but not completely, insure individuals against adverse events.

CONCLUSION

Asymmetric information in insurance markets has two important implications:

1) It can cause adverse selection in private insurance provision hence the need for **social** insurance

2) It can cause moral hazard (as insurer cannot perfectly monitor behavior), hence the need to limit generosity of insurance

The ironic feature of asymmetric information is, therefore, that it simultaneously motivates and undercuts the rationale for government intervention through social insurance.

REFERENCES

Jonathan Gruber, Public Finance and Public Policy, Fourth Edition, 2012
Worth Publishers, Chapter 12

Finkelstein, Amy and Liran Einav. "Selection in Insurance Markets: Theory and Empirics in Pictures", Journal of Economic Perspectives, 2011 (web)

Undergraduate Public Economics

Danny Yagan
UC Berkeley

Lecture 17 Social Security

RETIREMENT PROBLEM

Life-Cycle: Individuals ability to work declines with aging but individuals continue to live after they are unwilling/unable to work

Standard Life-Cycle Model Prediction: Absent any government program, rational individual would save while working to consume savings while retired

Optimal saving problem is extremely complex: uncertainty in returns to saving, in life-span, in future ability/opportunities to work, in future tastes/health

In practice: When govt was small before 20th century \Rightarrow Many people worked till unable to (often till close to death) and then were taken care of by family members [US elderly poverty rate very high before Social Security]

GOVT INTERVENTION IN RETIREMENT POLICY

Actual Retirement Programs: All OECD countries implement substantial government funded retirement programs (substantial share of GDP around 6-8%, US smaller around 4%), started in first part of 20th century and have been growing.

Common structure:

Individuals pay social security contributions (payroll taxes) while working and receive retirement benefits when they stop working till the end of their life (annuity)

Extension of the earlier family model: it's no longer your own working kids who take care of you in old age but all workers in the country

In the United States, the public retirement program is called **Social Security**

SOCIAL SECURITY: PROGRAM DETAILS

How Is Social Security Financed?

Almost all workers in the United States pay the Federal Insurance Contributions Act (FICA) tax on their earnings.

Tax is 12.4% of earnings (6.2% paid by employer, 6.2% paid by employees) up to a cap of \$118,500 (in 2015)

Who Is Eligible to Receive Social Security?

A person must have worked and paid this payroll tax for 40 quarters (10 years) over their lifetime, and must be of age 62 or older.

SOCIAL SECURITY: PROGRAM DETAILS

How Are Social Security Benefits Calculated?

Annuity payment: A payment that lasts until the recipient's death.

The amount of this annuity payment is a function of the recipient's average (taxable) earnings over the person's 35 highest earning years where each month's earnings are expressed in today's dollars (AIME = average indexed monthly earnings)

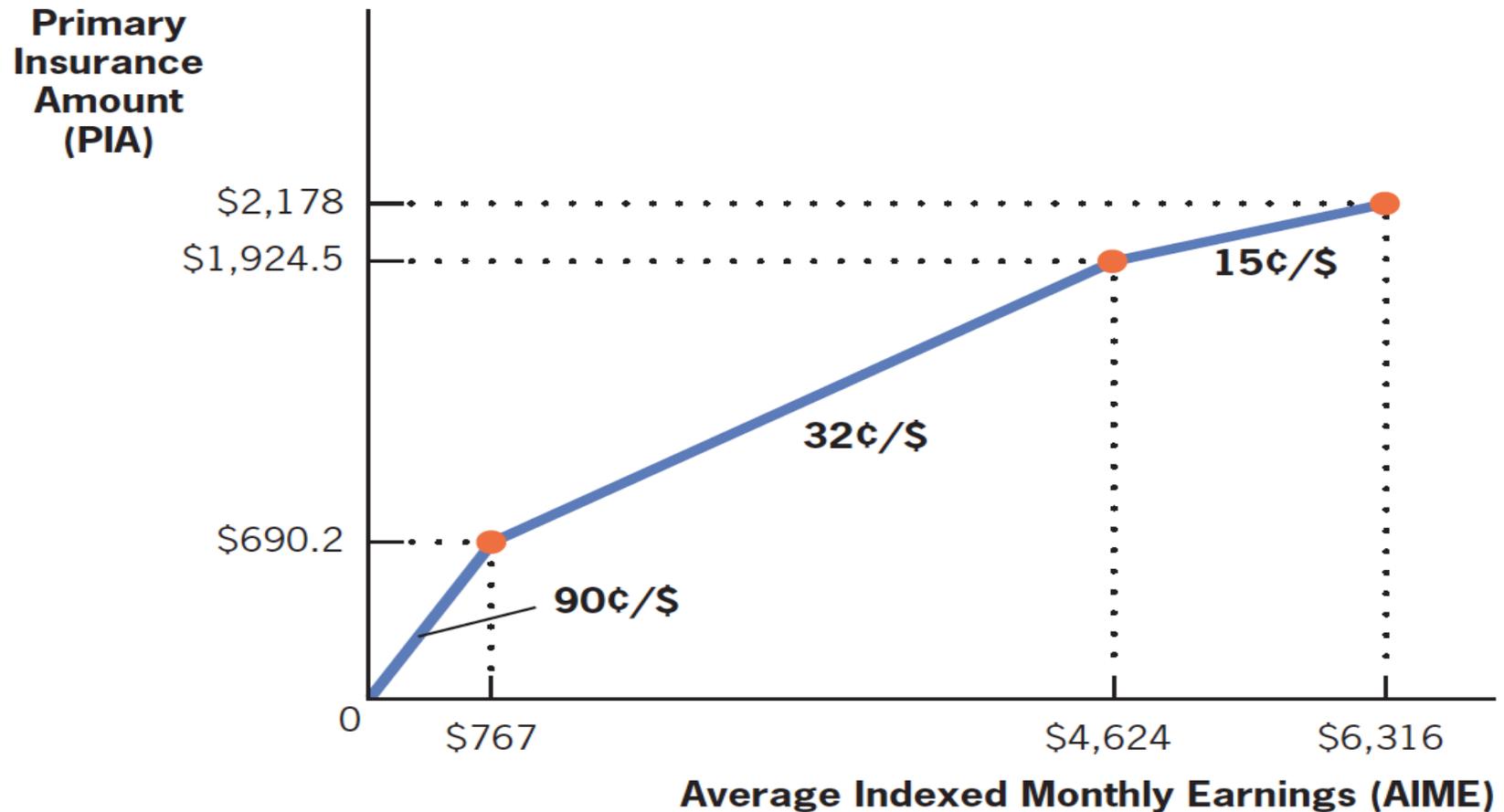
Once benefits start for a given person, they are indexed to price inflation once every year ("real" annuity)

APPLICATION: Why Choose 35 Years?

- Using the 35 highest years reflects multiple concerns.
 - No penalty for low-earning years early in career.
 - Not too large a benefit for high earning years late in career.
- Too short a window leads to abuse:
 - Bus driver working 25-hour shifts to maximize pension payment.
 - Brazilian public employees receiving promotions right before retirement.

13.1

Social Security Benefits as a Function of Earnings



How Are Social Security Benefits Paid Out?

Full Benefits Age (FBA): The age at which a Social Security recipient receives full retirement benefits (Primary Insurance Amount): currently 66 (used to be 65 and is increasing to 67)

Early Entitlement Age (EEA): The earliest age at which a Social Security recipient can receive reduced benefits: currently 62

If you claim benefits 1 year before FBA, you get 8% less in annual benefits (permanently), if you claim 2 years before FBA, you get 16% less in annual benefits, etc.

You get 8% more in benefits if you claim 1 year after FBA. Benefits automatically paid at 70.

SOCIAL SECURITY: PROGRAM DETAILS

Can You Work and Receive Social Security?

The *earnings test* reduces the benefits of 62 to 65-year olds by \$0.50 for each dollar of earnings they have above about \$15K

Not really a tax because later benefits are increased (as if you had retired later) but most people don't understand the system

Are There Benefits for Family Members?

- Spouses of claimants (get own benefits or 50% of primary earner benefits, whichever is biggest)
- Children of deceased workers.
- Spouses who survive a Social Security recipient

SOURCES OF RETIREMENT INCOME

- 1) Govt provided retirement benefits (US Social Security): For 2/3 of US retirees, SS is more than 50% of income. 1/3 of elderly households depend almost entirely (90%+) on SS.
- 2) Home ownership: 75% of US elderly are homeowners
- 3) Employer pensions (tax favored): about 50% of elderly US households have employer pension benefits. Two types:
 - a) Traditional: defined benefit and mandatory: **employer** carries full risk [in sharp decline, many in default]
 - b) New: defined contribution and elective: 401(k)s, **employee** carries full risk
- 4) Extra savings through non-tax favored instruments: significant only for wealthy minority [=10% of retirees]

FUNDED VS. UNFUNDED PROGRAMS

Two forms of retirement programs:

1) Unfunded (pay-as-you-go): benefits of current retirees are paid out of contributions from current workers [generational link]

current benefits = current contributions

2) Funded: workers contributions are invested in financial assets and will pay for benefits when they retire [no generational link]

current benefits = past contributions + market returns on past contributions

Social security (as most public retirement systems) is unfunded

Most private pension plans (such as 401(k)s) are funded

FUNDED VS UNFUNDED SYSTEMS

1) **Funded system:** each generation gets a market return r on contributions: $\text{benefits} = \text{tax you paid} \cdot (1 + r)$

2) **Unfunded system:** 1st generation of retirees gets free benefits when the system starts

For later generations: pay tax (for older generation) and you get benefits from younger generation

Implicit return on taxes is the sum of population growth n and real wage growth (per worker) g

$\text{benefits} = \text{tax paid} \cdot (1 + n)(1 + g) \simeq \text{tax paid} \cdot (1 + n + g)$

What Is Social Security and How Does It Work?

How Does Social Security Work Over Time?

How Social Security Redistributes Income

■ TABLE 13-1

Social Security in a Two-Period World

Period	Number of Young Workers	Earnings Per Young Worker	Taxes Paid Per Young Worker	Total Taxes Paid	Number of Old Retirees	Benefits to Old Retirees	Taxes Paid by Old Retirees	Rate of Return
1	100	\$20,000	0	0	0	0	—	—
2	105	\$21,000	\$2,100	\$220,500	100	\$2,205	0	Infinite
3	110	\$22,050	\$2,205	\$242,550	105	\$2,310	\$2,100	10%
4	115	\$23,153	\$2,315	\$266,225	110	\$2,420	\$2,205	10%
5	121	\$24,310	0	0	115	0	\$2,315	Negatively infinite

FUNDED VS UNFUNDED SYSTEMS

Unfunded system is always desirable when $n + g > r$ (Diamond 1965): an economy with $n + g > r$ is called **dynamically inefficient** and introducing an unfunded system makes a Pareto improvement

US economy: Annual $n = 1\%$ and $g = 1\%$ [$n + g$ was higher in 1940-1970]. $r \simeq 5\%$. In general $r > n + g$ in practice.

Note that r is much more risky than $n + g$: risk adjusted market rate of return should be lower than average market rate r but still higher than $n + g$

Funded system delivers higher returns because it does not deliver a free lunch to 1st generation

Choice between funded vs. unfunded system is an **inter-generational redistribution trade-off**

How Does Social Security Redistribute in Practice?

Social Security Wealth (SSW): The expected present discounted value of a person's future Social Security payments minus the expected present discounted value of a person's payroll tax payments.

SSW is computed as follows:

- Calculate the entire future stream of benefits that a person expects to receive before he or she dies.
- Use a discount rate to calculate the present discounted value (PDV) of that stream of benefits.
- Calculate the entire future stream of social security taxes that a person expects to pay before he or she dies.
- Compute the PDV of that stream of taxes.
- Take the difference between these two to get the SSW.

13.1

How Does Social Security Redistribute in Practice? SSW for a Single Male

Earnings	Turns 65 in 1960	Turns 65 in 1995	Turns 65 in 2030
Low earner	\$26,100	\$12,500	-\$4,100
Average earner	35,500	-5,100	-56,200
High earner	35,800	-41,100	-248,500

- Redistribution from younger to older cohorts due to:
 - First cohort didn't pay in until 1937.
 - Payroll tax has increased over time.

13.1

How Does Social Security Redistribute in Practice?

Some examples of how SSW varies within groups that are the same ages include the following:

- Females have more SSW than males because they live longer.
- Married couples have more SSW than single people.
- Single-earner couples have more SSW than two-earner couples.
- The gains to the poor relative to the rich from Social Security are overstated because the length of life rises with income.

RATIONALES FOR SOCIAL SECURITY

A. Individual Failure

Without a public program, people won't save enough for their own retirement because of myopia, self-control problems, information (how much to save, how to invest savings)

Popularity of Social Security suggests that people understand their own failures and the need for government intervention

B. Adverse selection in the annuities market

The longer a person lives, the less money the insurer makes from an annuity contract

⇒ People with short life expectancy less likely to buy annuities

This could lead to such a high price for annuities that most potential buyers would not want to buy them

MODEL: RATIONAL VS. MYOPIC SAVERS

1) Rational individuals: [draw graph]

$$\max u(c_1) + \delta u(c_2) \text{ st } c_1 + s = w \text{ and } c_2 = s \cdot (1 + r)$$

$$\Rightarrow c_1 + c_2/(1 + r) = w$$

FOC: $u'(c_2)/u'(c_1) = 1/[(1 + r)\delta]$, let s^* be optimal saving

Example: If $\delta = 1$ and $r = 0$ then $s^* = w/2$ and $c_1 = c_2 = w/2$

2) Myopic individuals:

$$\max u(c_1) \text{ st } c_1 + s = w \text{ and } c_2 = s \cdot (1 + r) \Rightarrow c_1 = w \text{ and } s = c_2 = 0$$

MODEL: RATIONAL VS. MYOPIC SAVERS

Social welfare is always $u(c_1) + \delta u(c_2)$

Govt imposes forced saving tax τ such that $\tau = s^*$ and benefits $b = \tau \cdot (1 + r)$. Cannot borrow against b [as in current Social Security]

1) Rational individual unaffected: adjusts s one-to-one so that outcome unchanged [rational unaffected as long as $\tau \leq s^*$]: 100% crowding out of private savings by forced savings

$c_1 = w - (s^* + s')$ and $c_2 = (s^* + s') \cdot (1 + r) \Rightarrow$ choosing s' is equivalent to choosing $s = s^* + s'$, rational person chooses $s' = 0$

2) Myopic individual affected (0% crowding out): new outcome maximizes Social Welfare

Forced savings is a good solution: does not affect those responsible, affects the myopic individuals in the socially desired way

MODEL: COMMENTS

Universal forced savings vs. means-tested program

An alternative to forced savings is to just have a means-tested program for poor elderly (who did not save) and financed by tax on everybody

2 drawbacks of means-tested program:

- a) Transfer from responsible individuals to myopic individuals
- b) Incentives to under-save to get means-tested pension

⇒ Forced savings program does not penalize responsible people and is likely to be more stable politically

Crowd-Out Effect of Social Security on Savings

The effect of Social Security on private savings has been the subject of a large number of studies over the past 30 years

To measure the impact of Social Security on savings, there must be a way to compare people with different levels of Social Security benefits who are otherwise identical

In the United States, Social Security is a national program that applies to almost all workers; very similar people usually have very similar benefits. Recent studies have provided evidence on the impact of Social Security-like programs on private savings in Italy.

Italian Reforms in 1992 substantially reduced the benefits, and thus future SSW, for younger workers in the public sector, while reducing much less the benefits of older workers and those in the private sector.

Studies estimate that about 30–40% of the reduction in SSW was offset by higher private savings.

Evidence for Myopia and adequate savings

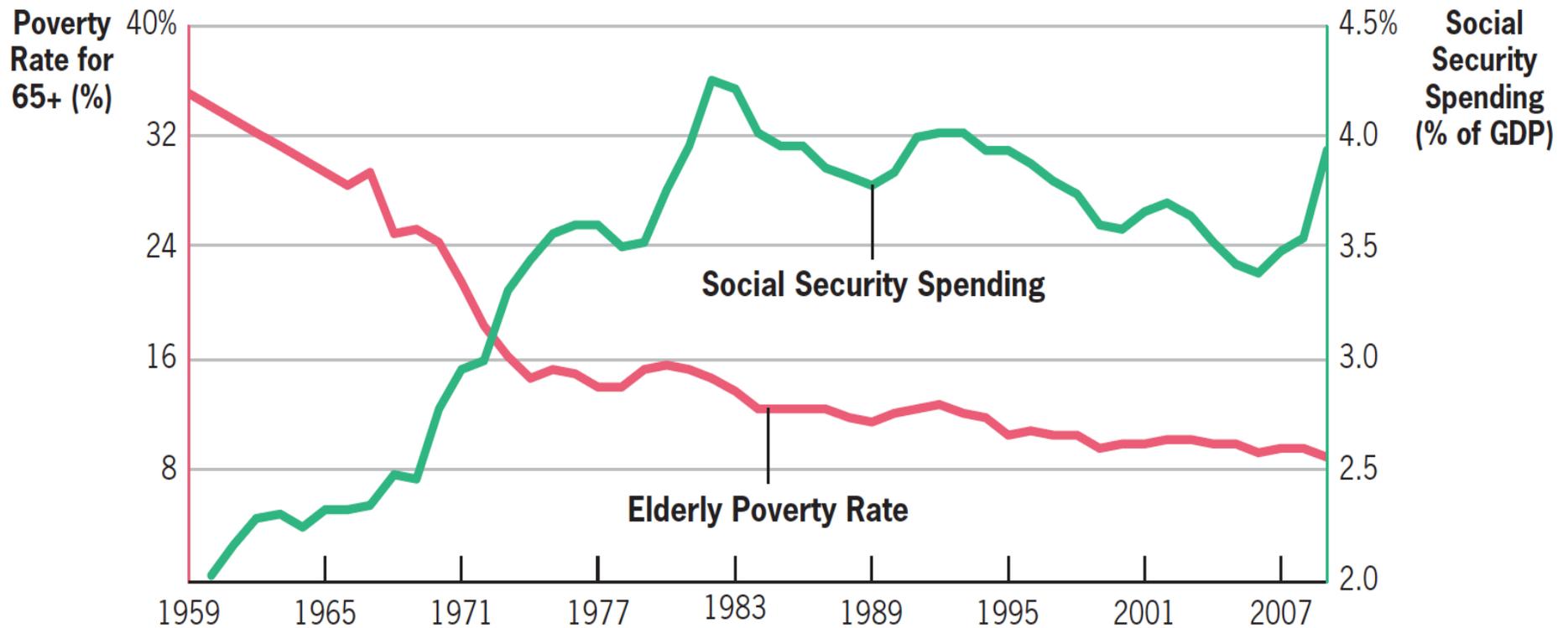
1) Diamond JpubE 1977: old age poverty has fallen as Social Security expanded. Poverty for other groups has not fallen nearly as much

2) Fall in consumption **during** retirement: Hamermesh (1984) shows that consumption falls by 5% per year for the elderly [consumption is not smooth but not necessarily suboptimal]

3) Fall in consumption **at** retirement: Bernheim, Skinner, Weinberg (2001) show that drop in consumption is significant for all groups except the wealthiest [consistent with myopia]

13.2

Living Standards of the Elderly, 1959–2009



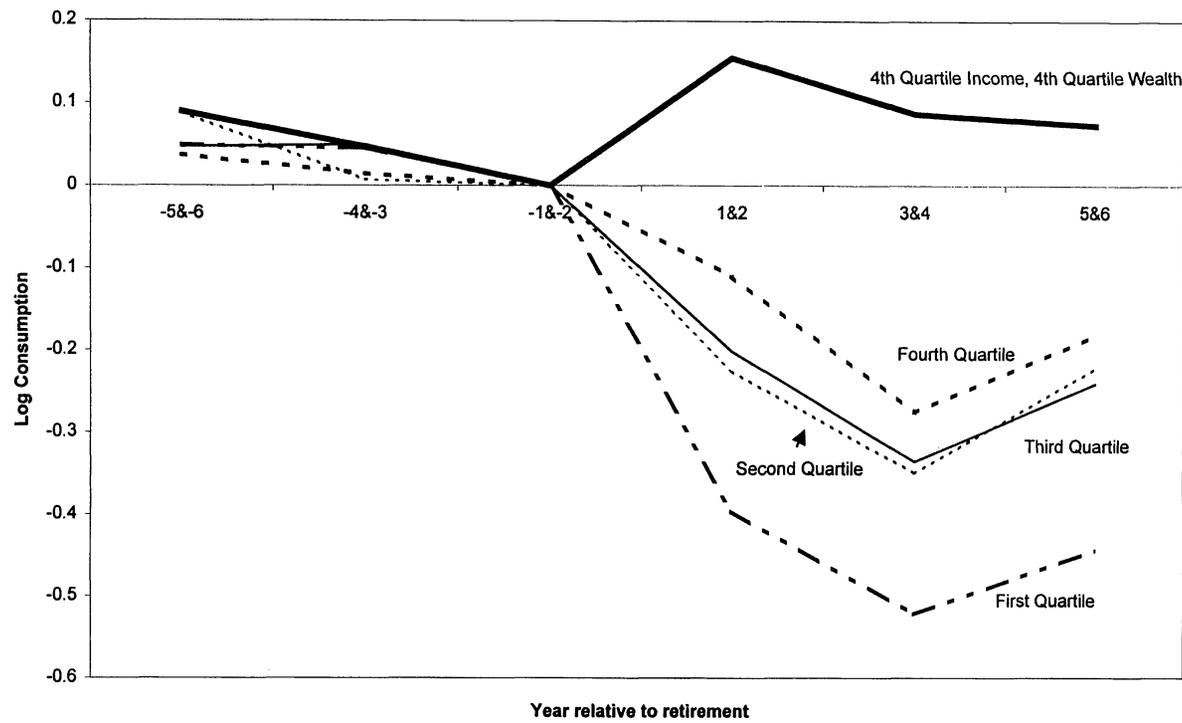


FIGURE 4. CHANGE IN CONSUMPTION AT RETIREMENT, BY WEALTH QUARTILE

Consumption drop at retirement: Aguiar-Hurst JPE05

Starting point: Empirically, consumption falls with retirement...but studies use expenditures as measure of consumption

Aguiar-Hurst JPE05 show that it is important to differentiate between consumption and expenditures. Further, the paper provides new information on the complementarity of consumption and leisure after retirement.

- 1) Confirm that expenditure on food falls by 17% at retirement but
- 2) time spent on home production rises by 60%
- 3) all measures of caloric intake, vitamin intake, meat quality, etc. do not drop at retirement (find that caloric intake falls when getting unemployed, surprising)

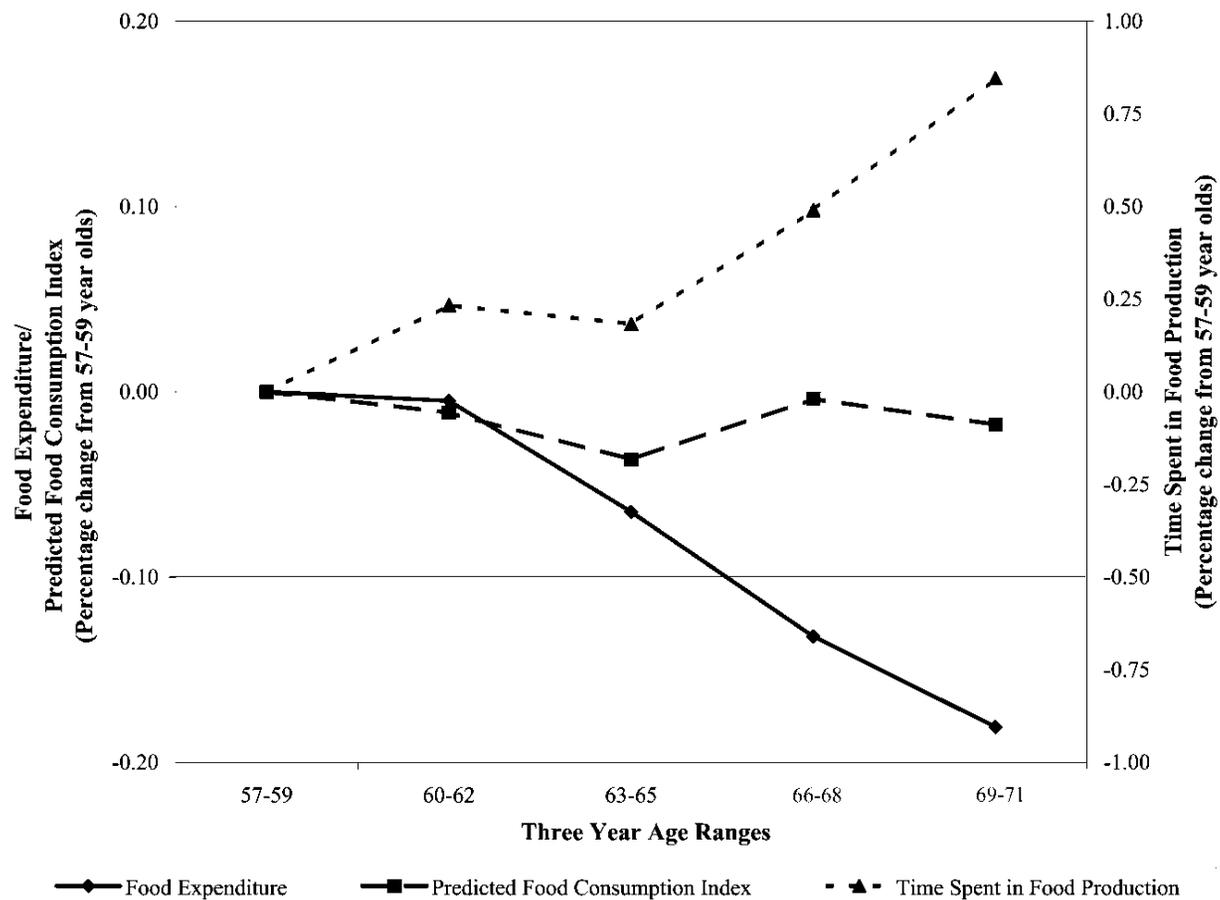


FIG 1.—Percentage change in food expenditure, predicted food consumption index, and time spent on food production for male household heads by three-year age ranges. Data are taken from the pooled 1989–91 and 1994–96 cross sections of the CSFII, excluding the oversample of low-income households. The sample is restricted to male household heads (1,510 households). All series were normalized by the average levels for household heads aged 57–59. All subsequent years are the percentage deviations from the age 57–59 levels. See Sec. IV for details of data and derivation of food consumption index

SOCIAL SECURITY AND RETIREMENT: THEORY

If a 62-year-old worker works until 63, instead of retiring at 62 and claiming her Social Security benefits, three things happen through the Social Security system:

- 1) She pays an extra year of payroll taxes on her earnings.
- 2) She receives one year less of Social Security benefits.
- 3) She gets a higher Social Security benefit level through the actuarial adjustment (8% extra permanently per year of delay)

Adjustment is called **actuarially fair** if those 3 effects cancel out in PDV (US system has been reformed to be close to fair on average)

SOCIAL SECURITY AND RETIREMENT: THEORY

Two key elements of a social security system may affect retirement behavior:

1) Availability of benefits at **Early Retirement (Entitlement) Age (EEA)**: (62 in US)

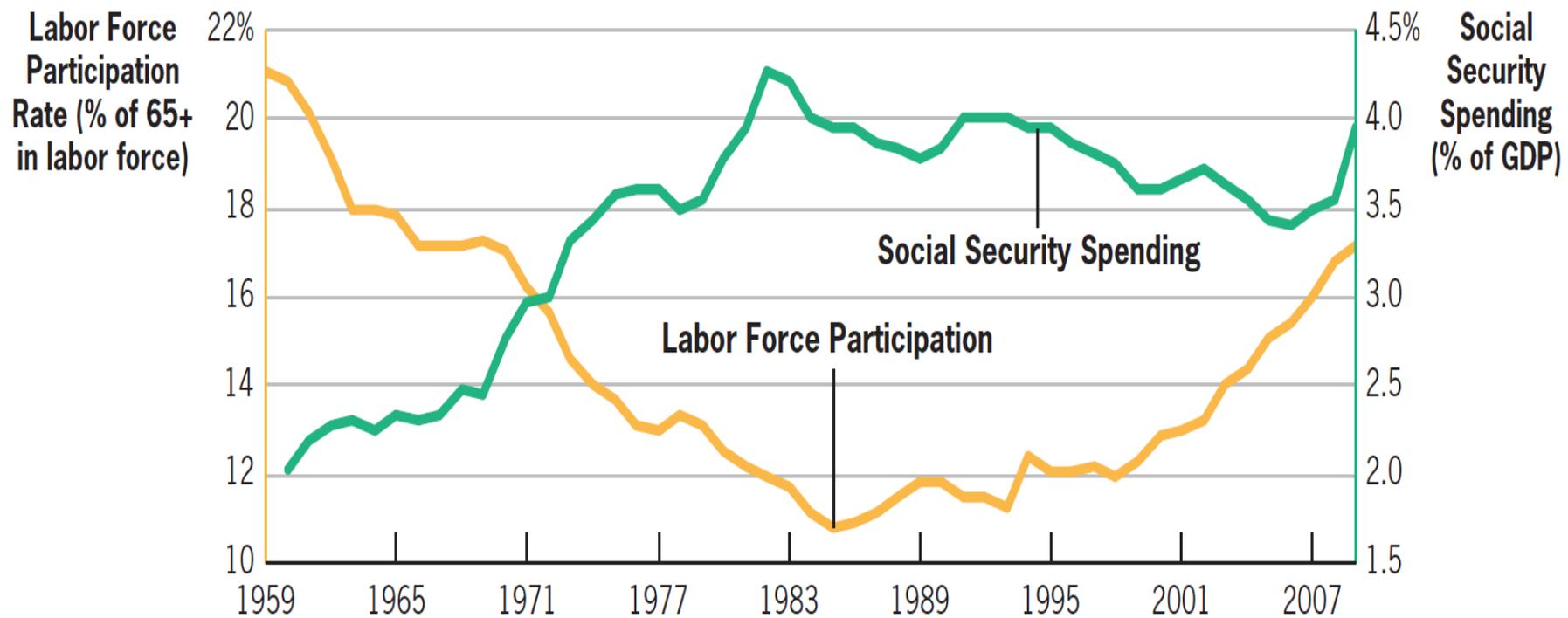
Those effects arise because of myopia or lack of information [a rational individual is not affected by EEA because he/she can use own savings while retired till he/she reaches age 62]

2) Non-actuarially fair adjustments of benefits for those retiring after the EEA:

If benefits are not adjusted in a fair way, they can create a huge implicit tax on work (US used to have very little adjustment, naturally not based on mortality rate)

13.3

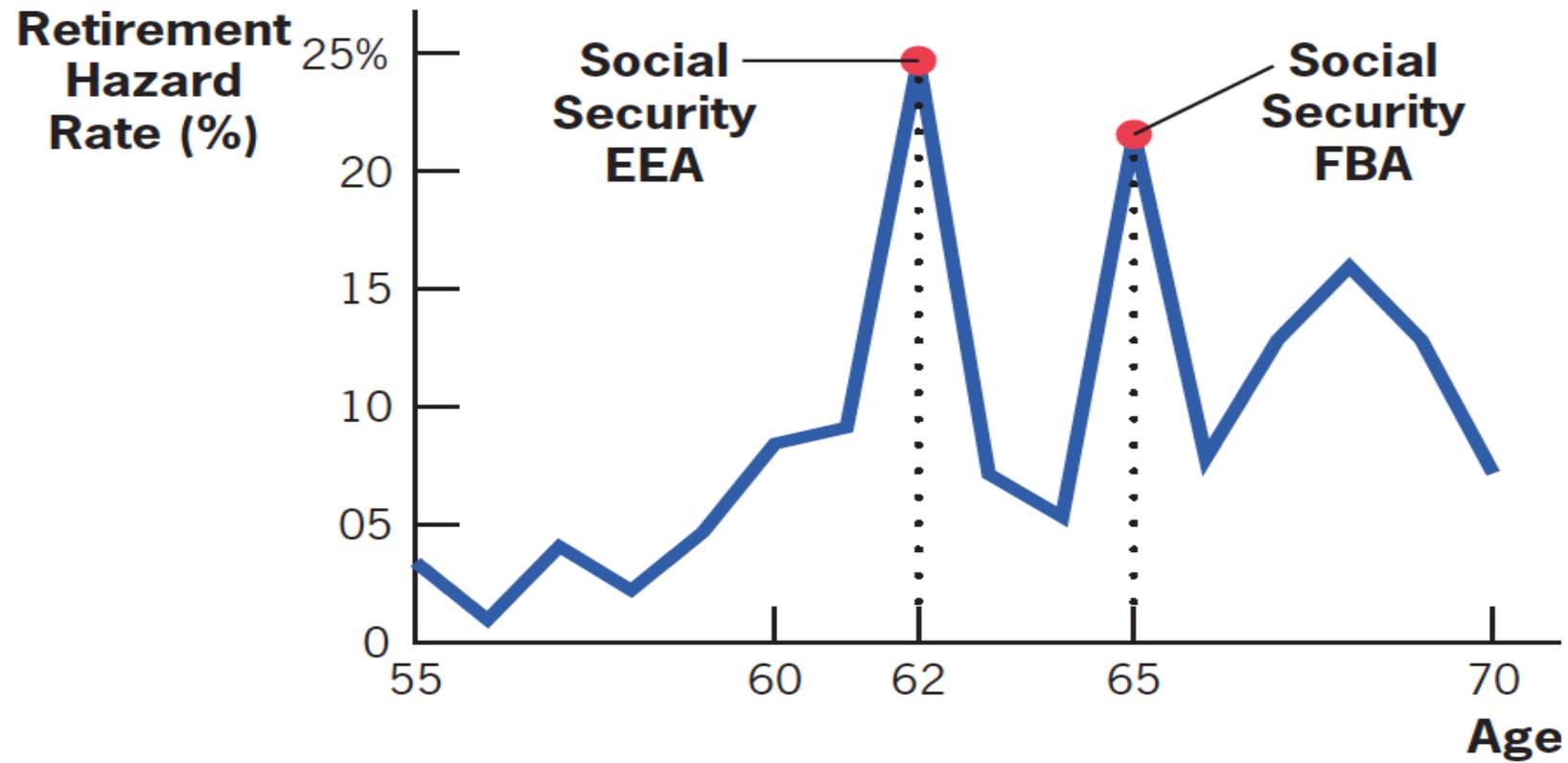
Elderly Work and Social Security, 1959–2009



13.3

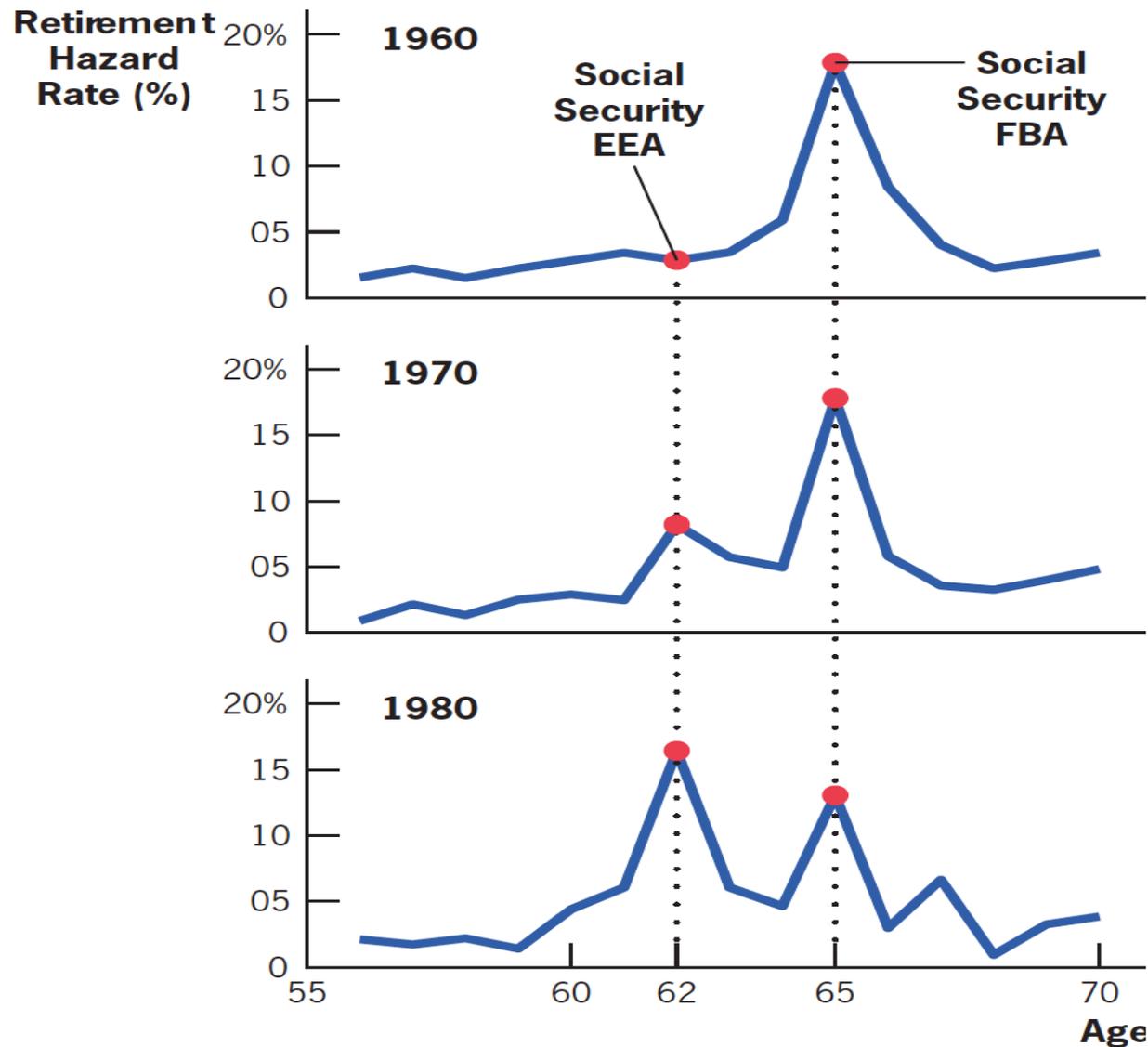
Spike in Retirement Hazard at EEA

- **Retirement hazard rate:** The percentage of workers retiring at a certain age.



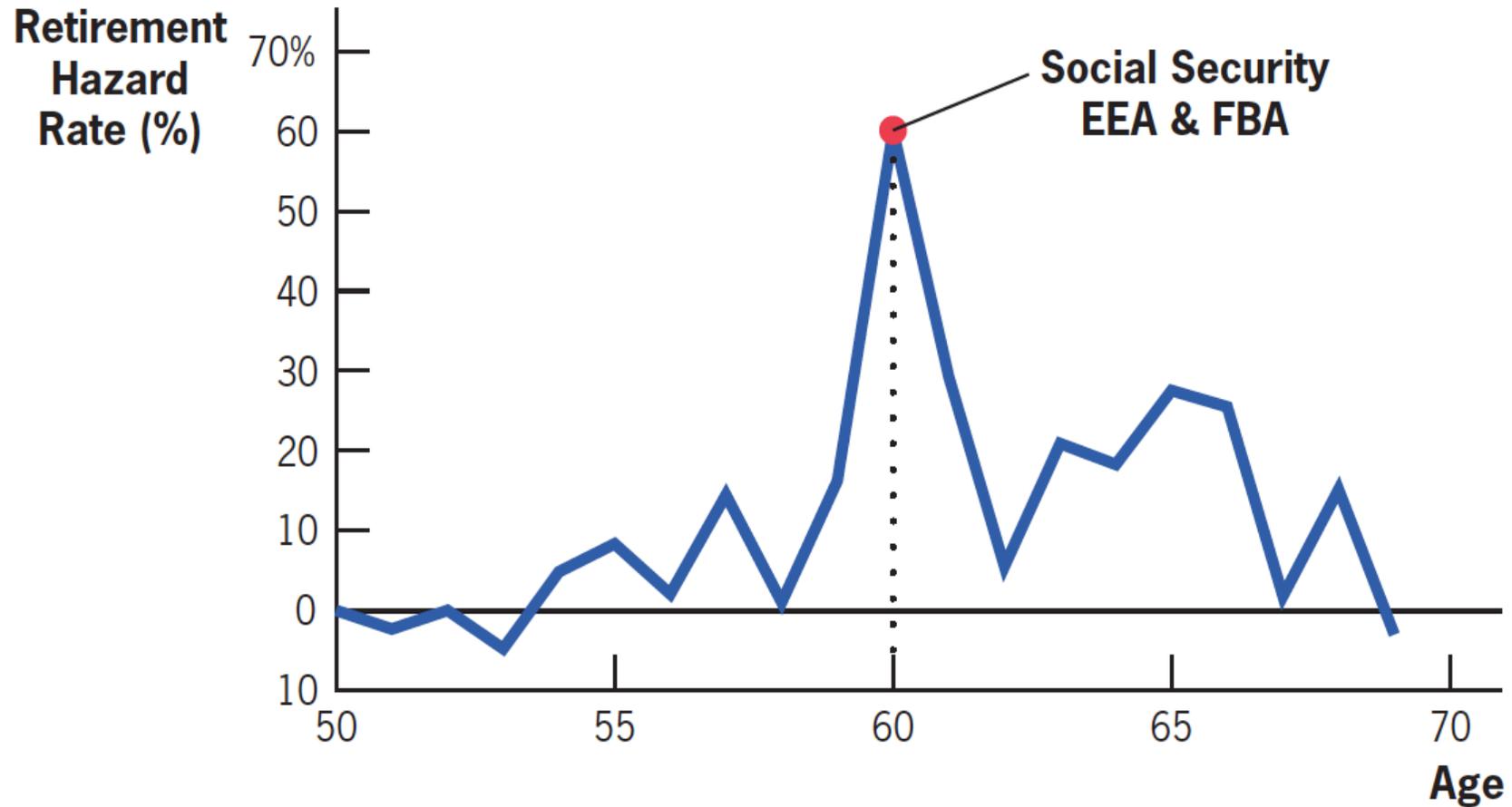
13.3

Spike in Retirement Hazard at EEA



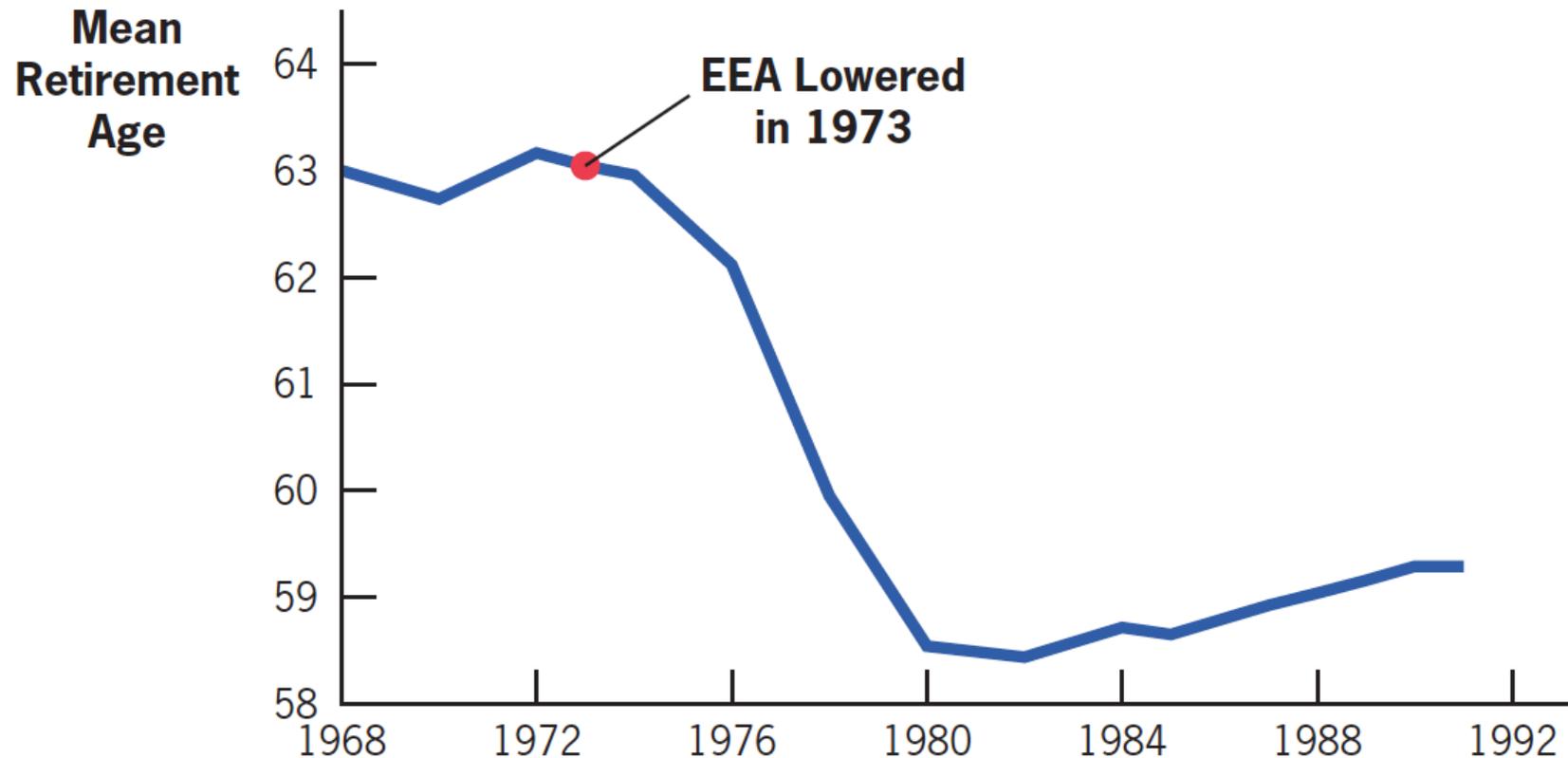
13.3

Retirement Hazard Rate in France



13.3

Evidence: Retirement Age in Germany, 1968–1992



- Retirement age lowered from 65 to 60 in 1973.

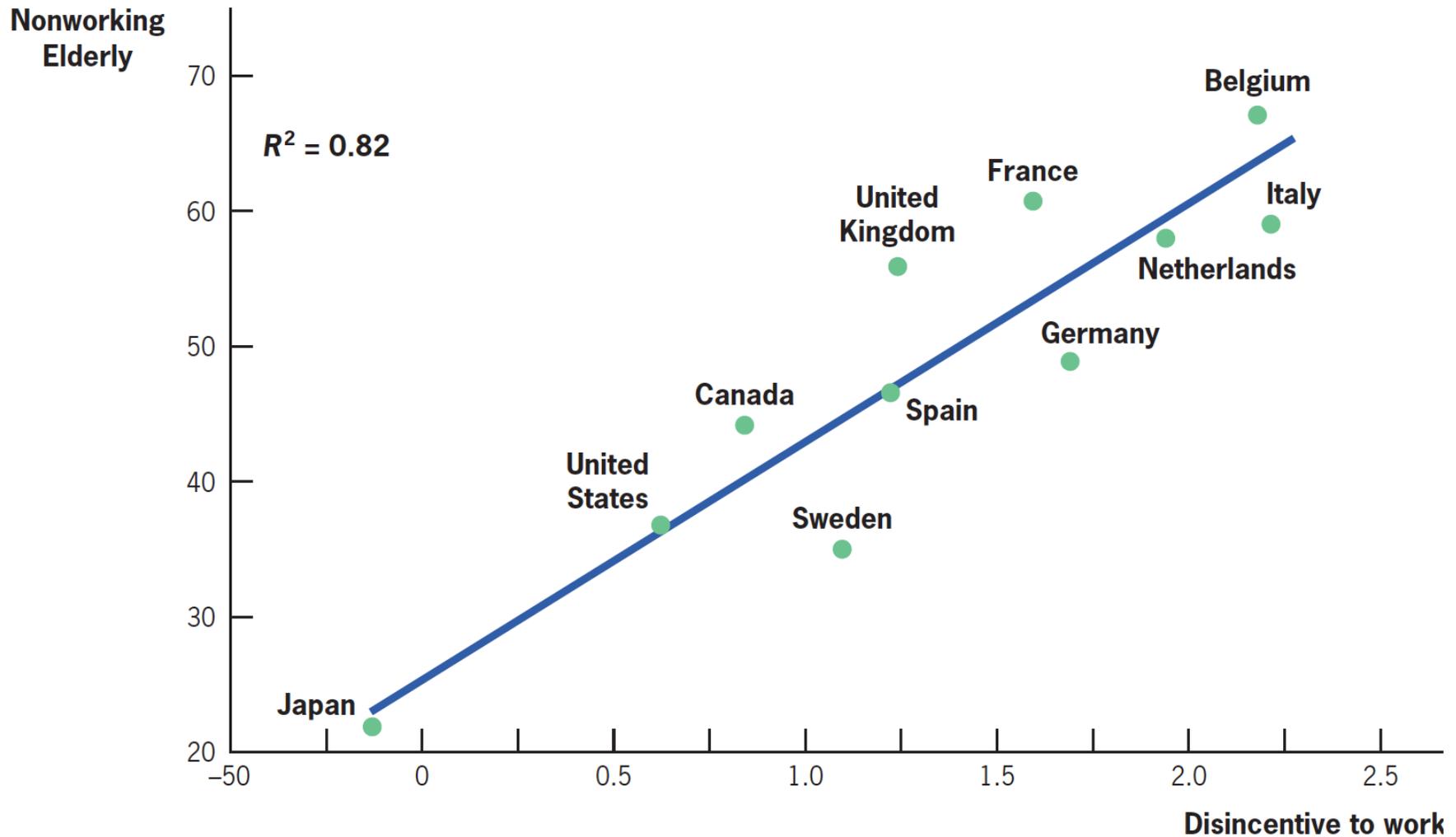
13.3

APPLICATION: Implicit Social Security Taxes and Retirement Behavior

- Gruber and Wise (1999) calculated the implicit tax from Social Security for a series of countries.
- Across countries, there is a great deal of variation in the implicit tax rate.
 - Implicit tax close to zero for 62-year-olds in the United States.
 - 91% in the Netherlands.
- And countries with higher taxes have less elderly labor force participation.

13.3

Implicit Social Security Taxes and Retirement Behavior



Social Security and Retirement: Implications

Evidence suggests that it is potentially very costly to design Social Security systems that allow very early retirement and/or penalize additional work beyond the retirement age.

Adjusting systems to more fairly reward work at old ages can mitigate much of the efficiency cost of Social Security

It seems better to have an early retirement age that is not too low and provide disability benefits to those who truly cannot work and haven't yet reached the early retirement age

Social Security Reform: Problems with Current System

Rate of return $n + g$ has declined from over 3% to about 2% due to:

1) n : Retirement of baby boom large cohorts born 1945-1965:

2) Increase in life expectancy at retirement age

Note: top half of individuals (in terms of lifetime earnings) has seen large life expectancy gains while bottom half life expectancy has stagnated in recent decades

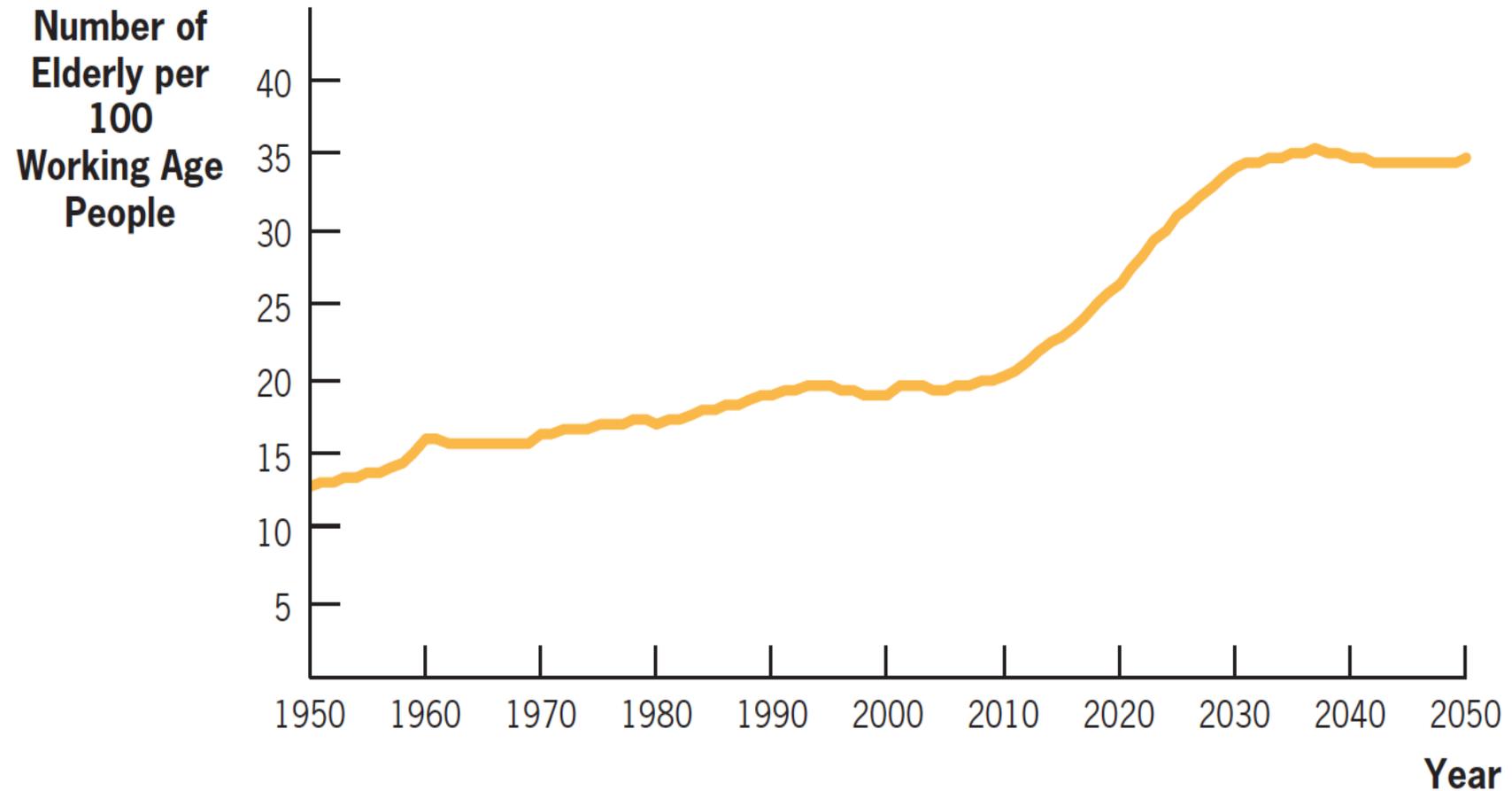
1)+2) imply number of elderly per working age person increases from .15 in 1960 to .35 in 2030

3) g : Slower productivity growth since 1975 (from 2% to 1%)

System requires adjusting taxes or benefits to remain in balance

13.4

Social Security Reform



1983 GREENSPAN COMMISSION

Demographic changes are predictable, so 1st reform was implemented in 1983 (designed to solve budget problems over next 75 years)

- 1) Increased payroll taxes to build a trust-fund
- 2) Increased retirement age in the future (from age 65 to 67)

Trust fund invested in Treasury Bills (Fed gov debt):

$$TF_{t+1} = TF_t \cdot (1 + i) + SSTax_t - SSBen_t$$

Trust fund is now peaking around (\$2.5 Tr), will be exhausted by 2040, taxes will then cover about 75% of promised benefits

Requires new adjustment: increase payroll tax rate now by 1.7 percentage points or wait till 2040 and then increase tax by 3.5 pp (from 12.4% to 15.9%, relatively small)

13.4

APPLICATION: The Social Security Trust Fund and National Savings

- In theory, one benefit of the partial funding of Social Security through the build-up of the trust fund is an increase in national savings.
- The trust fund is “off budget,” not supposed to be part of budget discussion.
- But typically the government reports the deficit/surplus from the “unified budget,” which incorporates off-budget categories.
- Makes it easy to treat trust fund as an asset, avoid fixing the deficit.

SOCIAL SECURITY REFORM OPTIONS

- 1) Increase contributions: increase tax rate or earnings cap
- 2) Reduce benefits: straight cut not politically feasible: a) Index retirement age to life expectancy, b) Index benefits to chained-CPI instead of CPI after retirement, c) Make benefits fully taxable for income tax
- 3) Means-tested benefits: bad for savings incentives and could make program politically unstable [a program for the poor is a poor program]. Could explain unusual conservative support.
- 4) Invest Trust Fund in higher yield assets (such as stock-market, as proposed by Clinton in 1990s). Advantages: higher return on average and government can be a long-term investor.
- 5) Major reform: privatization

SOCIAL SECURITY PRIVATIZATION

Two components:

1) Funding the system

2) Create individual accounts (like private employer 401k pensions)

current benefits = past contributions + market return

Controversial academic and policy debate, a number of countries have privatized their social security systems (Chile, Mexico, UK)

Main proponent: Feldstein (Harvard), Main critic: Diamond (MIT)

Pro argument: get higher return on contributions $r > n + g$, increase capital stock and future wages.

SOCIAL SECURITY PRIVATIZATION ACCOUNTING

Exactly the reverse of pay-as-you-go calculations:

1) First generation loses as they need to fund current retirees and own contributions. All future generations gain [generational redistribution]

2) If govt increases debt to pay for current retirees: future generations get higher return on contributions but need to re-pay higher govt debt \Rightarrow Complete wash for all generations

\Rightarrow Only way funding generates real changes is by hurting some transitional generations which have to double pay

ADDITIONAL PRIVATIZATION ISSUES

- 1) Risk: individuals bear investment risk (stock market fluctuates too much relative to economy) and cannot count on defined level of benefits [Privatization needs to include minimum pension provision]
- 2) Annuitization: hard to impose in privatized system because of political constraints [hard to force sick person to annuitize her wealth] ⇒ Some people will exhaust benefits before death and be poor in very old age [looming problem with 401(k)s]
- 3) Lack of financial literacy: Individuals do not know how to invest. Complicated choice, govt can do it for people more efficiently
- 4) Administrative costs: privatized systems (Chile, UK) admin costs very high (1% of assets) due to wasteful advertisement by mutual funds bc of lack of financial literacy

Evidence on Lack of Financial Literacy

401(k) private pensions in the US offer strong evidence of lack of financial literacy

1) Default effects: opt-in vs. opt-out have enormous effects on 401(k) enrollment [Madrian and Shea QJE'01]

2) $1/N$ investment choices of 401(k) contributions: many people invest contributions by dividing them equally into investment options (regardless of the options)

3) People often invest 401(k) in company stock which is extremely risky (Enron). Strong evidence of default effects in investment choices as well

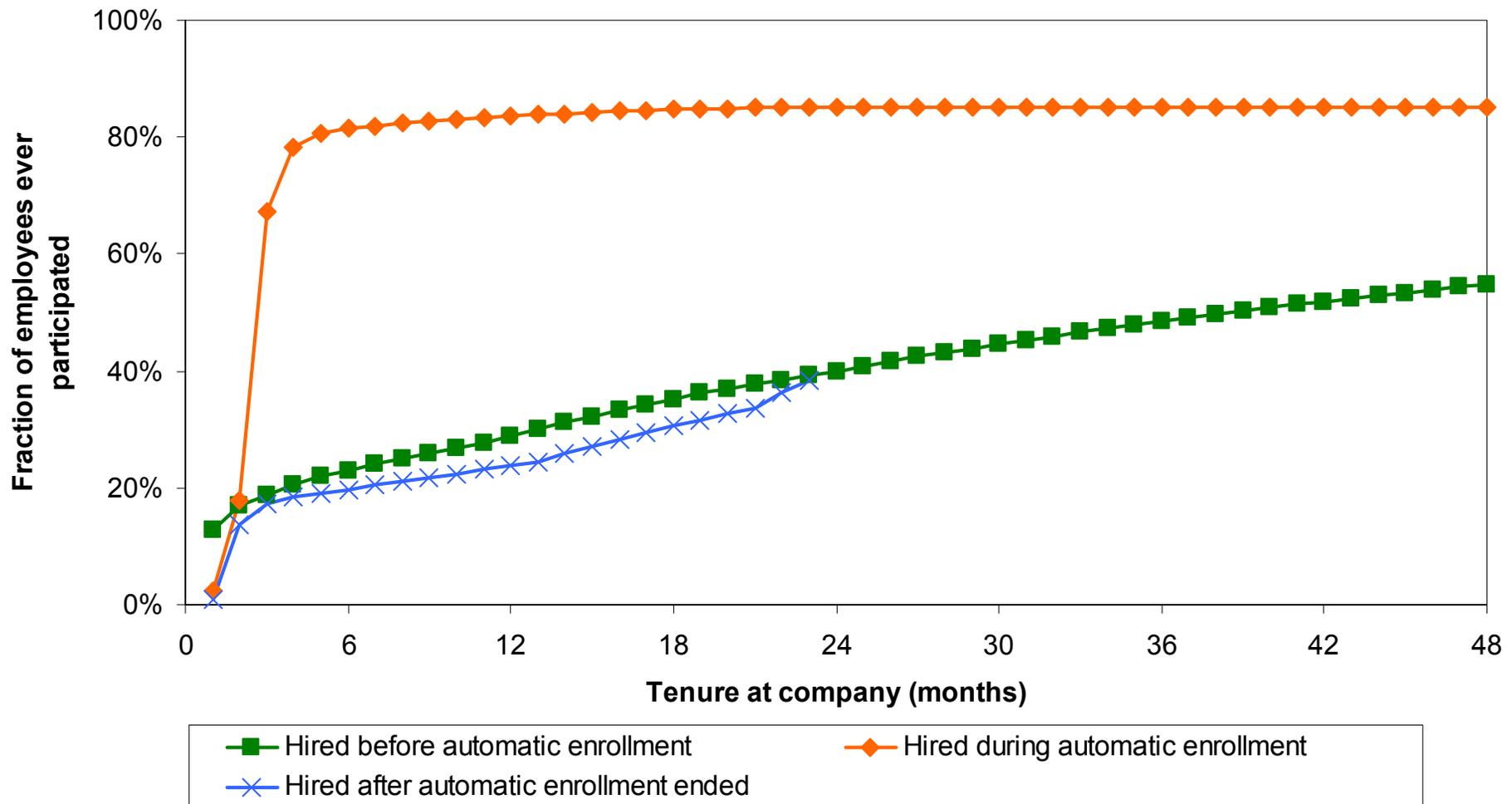
4) Evidence that financial education and advice has impacts on savings decisions (Thaler and Benartzi JPE '04: Saving More Tomorrow experiment).

⇒ May be much better to force people to save via mandatory social security system than rely on individual rationality

Automatic enrollment effect

Automatic enrollment dramatically increases participation.

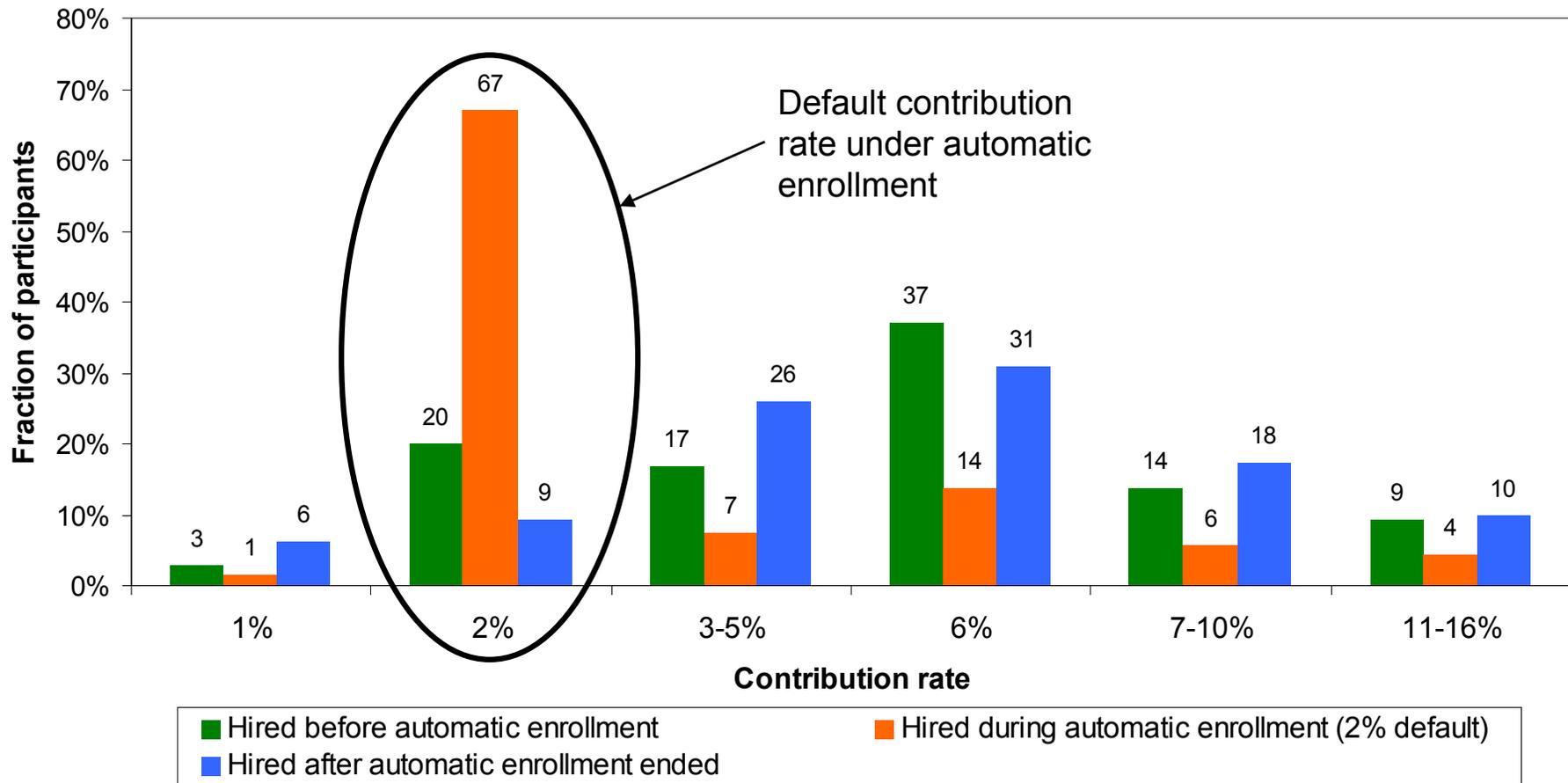
401(k) participation by tenure at firm: Company B



Automatic enrollment effect

Employees enrolled under automatic enrollment cluster at the default contribution rate.

Distribution of contribution rates: Company B



The Flypaper Effect in Individual Investor Asset Allocation (Choi, Laibson, Madrian 2007)

Studied a firm that used several different match systems in their 401(k) plan.

I'll discuss two of those regimes today:

Match allocated to employer stock and workers can reallocate

- Call this “default” case (default is employer stock)

Match allocated to an asset actively chosen by workers; workers *required* to make an active designation.

- Call this “no default” case (workers must choose)

Economically, these two systems are identical.

They both allow workers to do whatever the worker wants.

Consequences of the two regimes

	<u>Balances in employer stock</u>	
	Default ES	No Default
Own Balance in Employer Stock	24%	20%
Matching Balance in Employer Stock	94%	27%
Total Balance in Employer Stock	56%	22%

CONCLUSION

Social Security is the largest social insurance program in the United States, and the largest single expenditure item of the federal government

Key reasons for existence of social security programs: likely irrational undersaving by some individuals for retirement, and adverse selection in the annuities market

Social Security faces a long-run financing problem requiring to increase taxes or cut benefits in the long-run

The question of how to resolve this problem will be one of the most contentious sources of political debate for at least the first part of the twenty-first century

REFERENCES

Jonathan Gruber, Public Finance and Public Policy, Fourth Edition, 2012
Worth Publishers, Chapter 13

Aguiar, M. and E. Hurst “Consumption vs Expenditure.”, Journal of Political Economy, Vol. 113, 2005, 919-948.(web)

Bernheim, B. Douglas, Jonathan Skinner, and Steven Weinberg. “What accounts for the variation in retirement wealth among US households?.” American Economic Review (2001): 832-857.(web)

Choi, James, David Laibson, and Brigitte Madrian. “The flypaper effect in individual investor asset allocation.” (2007).(web)

Diamond, P. “National Debt in a Neoclassical Growth Model”, American Economic Review, Vol. 55, 1965, 1126-1150. (web)

Diamond, Peter A. “A framework for social security analysis.” Journal of Public Economics 8.3 (1977): 275-298.(web)

Gruber, Jonathan, and David A. Wise, eds. “Social security and retirement around the world.”[Book] University of Chicago Press, 2008.(web)

Hamermesh, Daniel S. “Consumption during retirement: the missing link in the life cycle.” The Review of Economics and Statistics 66.1 (1984): 1-7.(web)

Madrian, Brigitte C., and Dennis F. Shea. "The power of suggestion: Inertia in 401 (k) participation and savings behavior." *The Quarterly Journal of Economics* 116.4 (2001): 1149-1187.(web)

Thaler, Richard H., and Shlomo Benartzi. "Save more tomorrow: Using behavioral economics to increase employee saving." *Journal of Political Economy* 112.S1 (2004): S164-S187.(web)

Undergraduate Public Economics

Danny Yagan
UC Berkeley

Lecture 18

Unemployment Insurance, Disability
Insurance, and Workers' Compensation

INSTITUTIONAL FEATURES

Unemployment insurance, workers' compensation, and disability insurance are three social insurance programs in the United States, and they share many common features.

Unemployment insurance (UI): A federally mandated, state-run program in which payroll taxes are used to pay benefits to unemployed workers laid off by employers.

Disability insurance (DI): A federal program in which a portion of the Social Security payroll tax is used to pay benefits to workers who have suffered a medical impairment that leaves them permanently unable to work.

Workers' compensation (WC): State-mandated insurance, which firms generally buy from private insurers, that pays for medical costs and lost wages associated with an on-the-job injury.

Comparison of the Features of UI, DI, and WC

Characteristic	UI	DI	WC
Qualifying Event	Job loss, job search	Disability	On-the-job injury
Duration	26-65 weeks	Indefinite	Indefinite (if verified)
Difficulty of verification	Job loss: easy Search: impossible	Somewhat difficult	Very difficult
Average after tax replacement rate	47%	60%	89%
Variation across states	Benefits and other rules	Only disability determination	Benefits and other rules

Unemployment Insurance

Unemployment insurance is a major social insurance program in the U.S.

Substantial size: \$50 bn/year in normal times (\$150bn/year during Great Recession)

Macroeconomic importance in stabilization/stimulus

Like other social programs, triggered by an event

In this case, involuntary job loss

Controversial debate about unemployment benefits

Benefit: helps people in a time of need

Cost: reduces incentive to search for work while unemployed

What is the optimal design of UI system given this tradeoff?

Institutional Features of Unemployment Insurance

UI is a federally mandated, state-run program

Although UI is federally-mandated, each state sets its own parameters on the program.

This creates a great deal of variation across states

Useful as a “laboratory” for empirical work

⇒ UI is a heavily studied program

Financing of UI Benefits

1) UI is financed through a payroll tax on employers:

⇒ an employee will not see a deduction for UI on his or her paycheck.

This payroll tax averages 1-2% of earnings

2) UI is partially experience-rated on firms

⇒ the tax that finances the UI program rises as firms have more layoffs, but not on a one-for-one basis

Eligibility Requirements and Benefits

- 1) Individuals must have earned a minimum amount over the previous year.
- 2) Unemployment spell must be a result of a layoff, rather than from quitting or getting fired for cause (easy to check)
- 3) Individual must be actively seeking work and willing to accept a job comparable to the one lost (hard to check)

These eligibility requirements mean that not all of the unemployed actually collect benefits.

Even among eligible, 1/2 do not take up the UI benefit (Lack of information about eligibility, stigma from collecting a government handout, or transaction costs)

UI Benefits

UI benefits are a function of previous earnings

These benefits vary by state.

The replacement rate is the amount of previous earnings that is replaced by the UI system.

$$R = B/W$$

Replacement rates vary from 35% to 55% of earnings, and UI is treated as taxable income.

14.1

Unemployment Benefit Schedule for Michigan



UI Benefits Duration

In general, one can collect UI for 6 months.

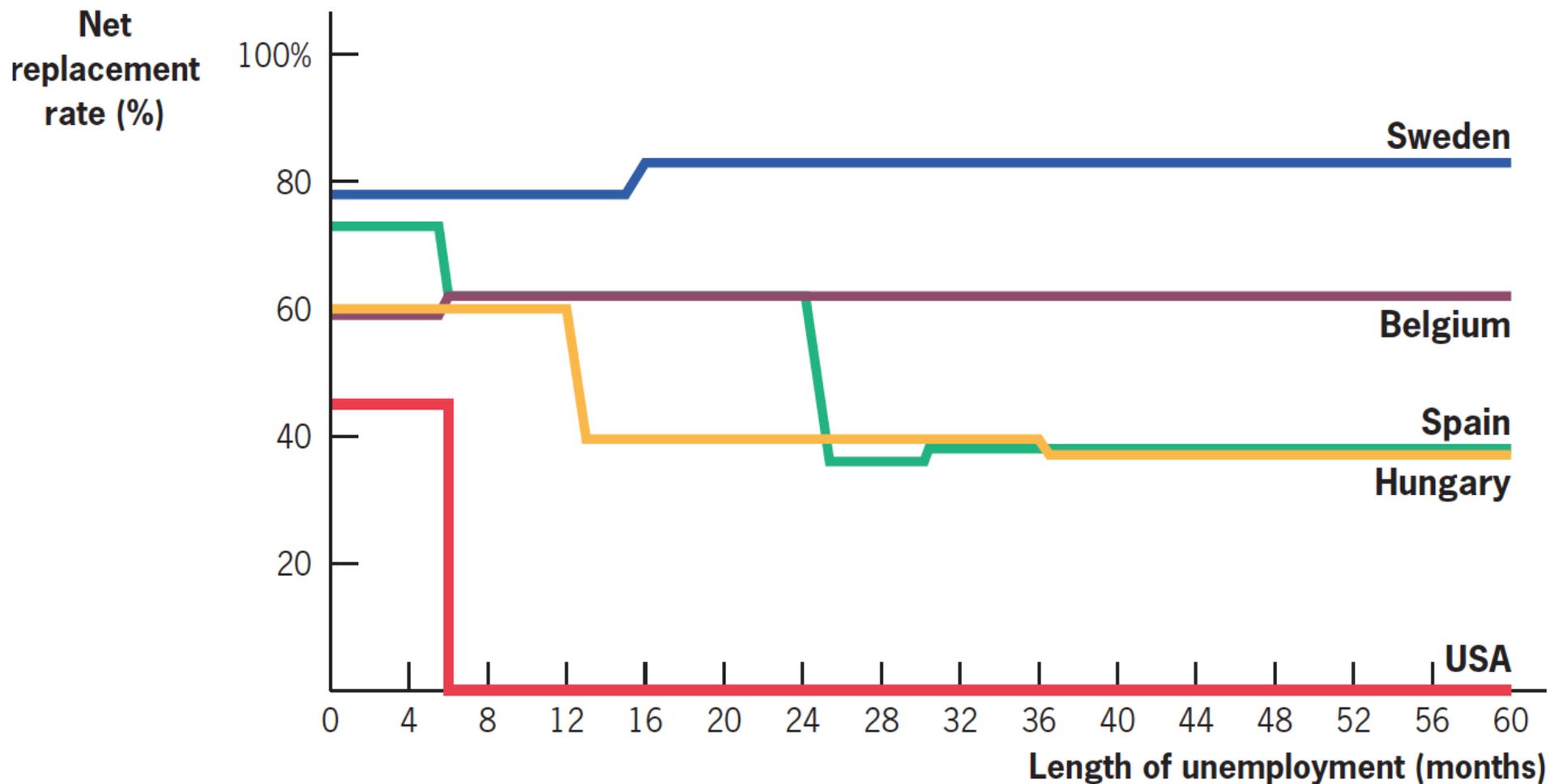
In recessions, benefits are automatically extended to 9 months or 12 months

In deep recessions, benefits can be further extended (23 months in 2008-13)

Duration of UI benefits typically much higher in European countries

14.1

APPLICATION: The Duration of Social Insurance Benefits around the World



Analysis of Optimal Unemployment Insurance

Which system is the best?

First need to define what we mean by “best”—what is the objective function?

Typical objective considered by economists: maximize agent’s welfare

In this case, because there is uncertainty, welfare is given by expected utility

Use a formal mathematical model to tackle the problem and get a number for the optimal benefit

Expected Utility Model

Individual's expected utility:

$$EU = (1 - p)u(c_e) + pu(c_u) = (1 - p)u(w - t) + pu(b)$$

p : probability of being unemployed

c_e = consumption when employed,

c_u = consumption when unemployed

w = wage when working

t = tax used to finance program,

b = UI benefit

Government needs to balance budget (taxes fund benefits):

$$t = (p/(1 - p)) \cdot b$$

Optimal UI with no moral hazard

No moral hazard means that p is not affected by UI

Plugging in govt. budget constraint, rewrite individual's expected utility as:

$$EU = (1 - p)u(w - (p/(1 - p))b) + pu(b)$$

Government's problem: find b that maximizes EU .

Optimal benefit b^* will be b such that: $c_u = c_e$

This is **full insurance** (as we saw earlier in class)

Optimal UI with moral hazard

With moral hazard, p increases with b as more generous benefits deter job search and hence increase unemployment

Government now chooses b to maximize EU but taking into account that p is a function of b in the budget constraint

$$EU = (1 - p)u(w - [p(b)/(1 - p(b))]b) + pu(b)$$

Get new formula:

$$\frac{u'(c_u) - u'(c_e)}{u'(c_e)} = \frac{1}{1 - p} \varepsilon_{p,b} \text{ with } \varepsilon_{p,b} = \frac{b}{p} \cdot \frac{dp}{db}$$

$\varepsilon_{p,b} > 0$ is the elasticity of unemployment rate with respect to benefits (captures size of moral hazard effects)

Now $0 < c_u < c_e < w$: partial insurance is optimum. Optimum level increases with curvature of $u(\cdot)$ but decreases with elasticity $\varepsilon_{p,b}$.

Empirical Estimation of Effects of UI

Moral hazard in UI is thought to manifest itself in the duration of the unemployment spell

Economists ask whether the unemployed find jobs more slowly when benefits are higher

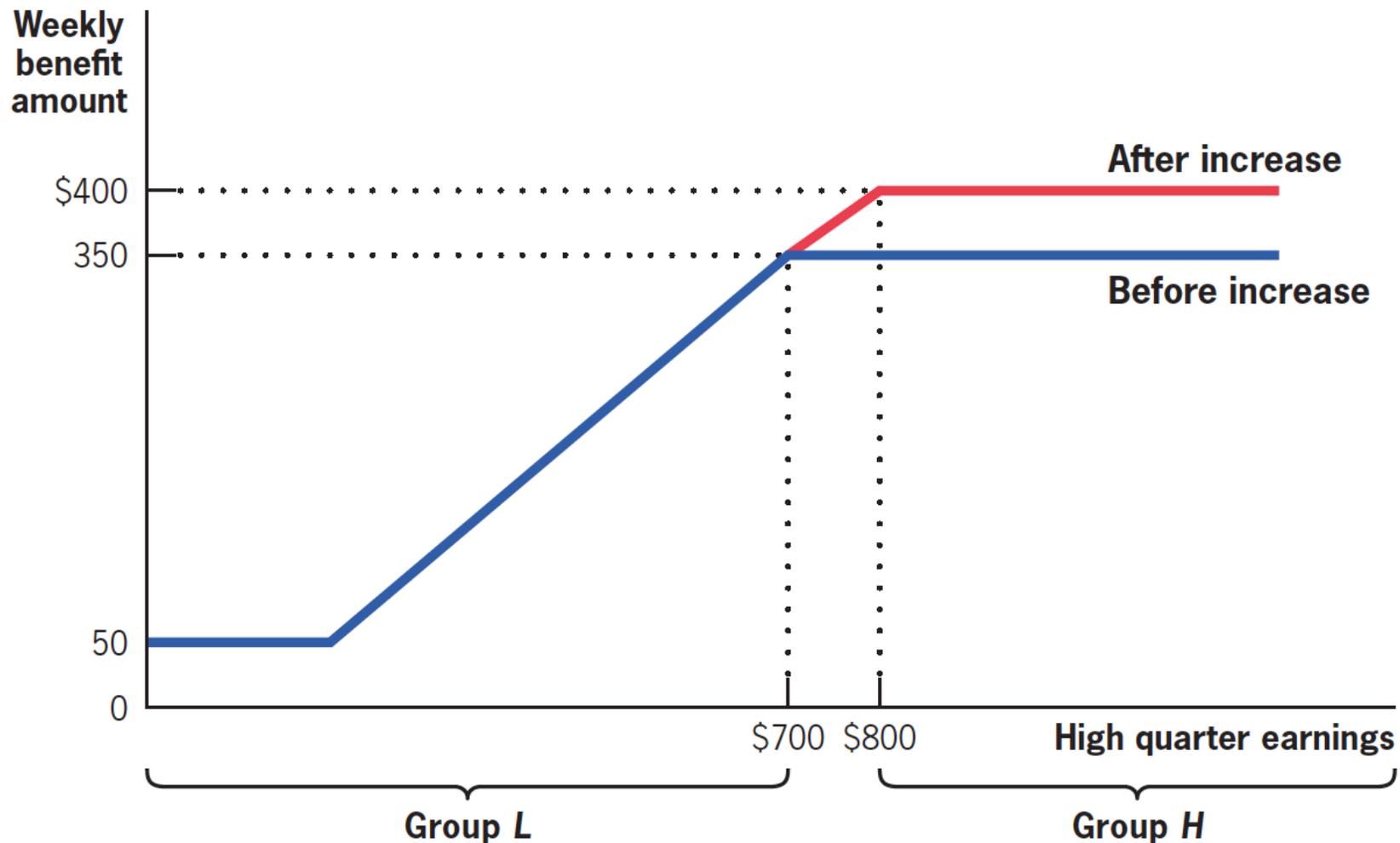
Key challenge: need to use quasi-experiments to identify these effects

One common empirical approach (Meyer 1990): difference-in-difference

Exploit changes in UI laws that affect a “treatment” group and compare to a “control” group

14.3

EVIDENCE: Moral Hazard Effects of Unemployment Insurance



Empirical Estimation of Effects of UI: Evidence

Meyer (1990) and many others implement this method using data on unemployment durations in the U.S. and state-level reforms

General finding: benefit elasticity of 0.4-0.6

10% rise in unemployment benefits leads to about a 4-6% increase in unemployment durations.

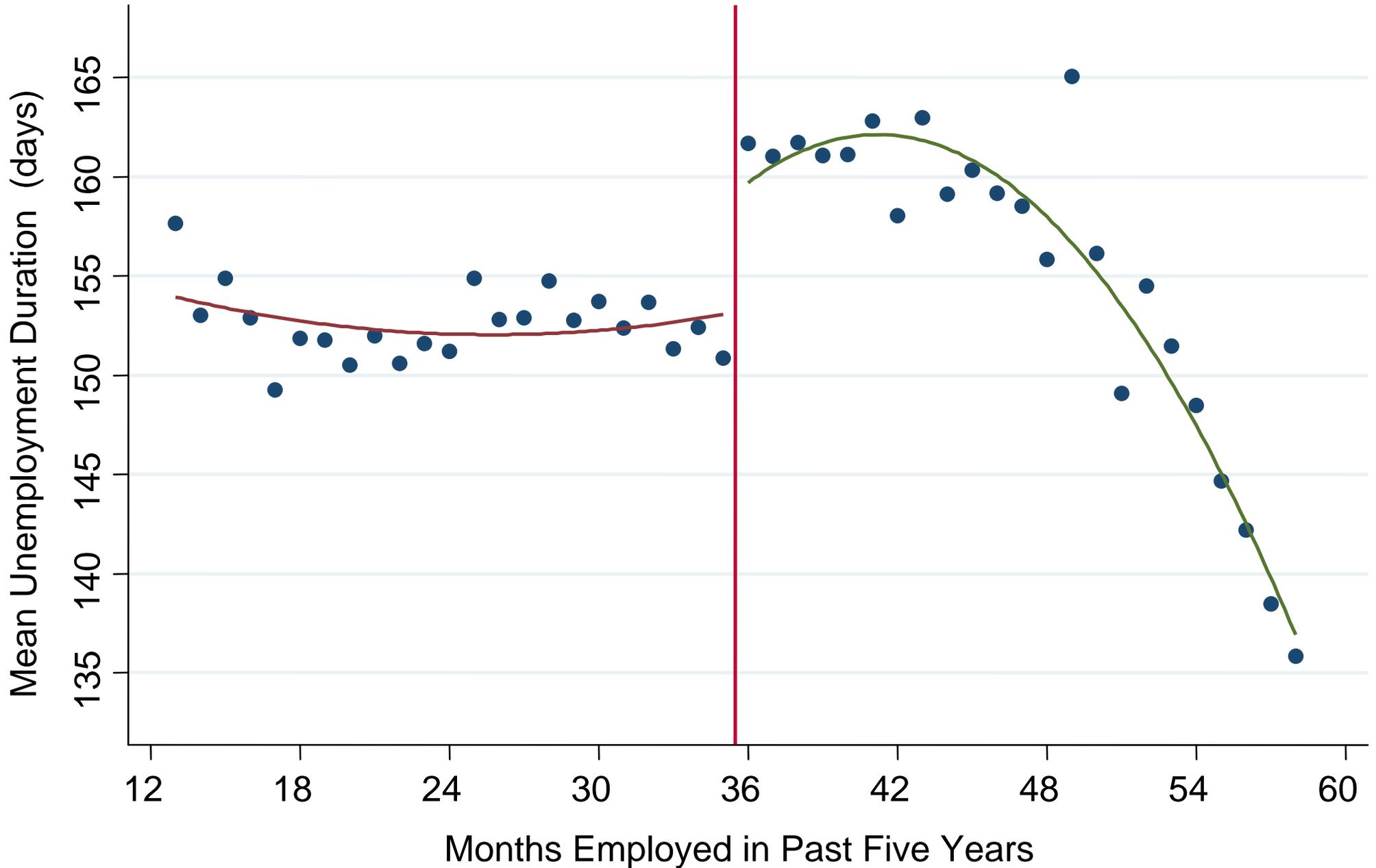
More recent empirical approach: **regression discontinuity**

Card-Chetty-Weber (2007) use the fact that in Austria, you get up to 30 weeks of benefits when you have been employed for 36+ months in last 5 years (instead of up to 20 weeks)

Can look at duration of unemployment based on how long you have worked in last 5 years \Rightarrow Finds somewhat smaller elasticity around 0.3

Card, Chetty, Weber (2007)

Effect of Benefit Extension on Unemployment Durations



Evidence on Consumption-Smoothing

Difference-in-difference strategy has been used to examine how UI benefits affects consumption

Gruber (1997) finds that consumption falls on average when people lose their job by about 10-15%

\$1 increase in UI benefits increases consumption by 30 cents

Much less than 1-1 because savings behavior changes, spousal labor supply, borrowing from friends, etc. (this is called self-insurance)

Does UI have Long-Term Benefits?

Another potential benefit of UI, neglected in simple model above: improvements in **match quality**

Are people forced to take worse jobs because they have to rush back to work to put food on the table?

E.g. engineer starts working at McDonalds.

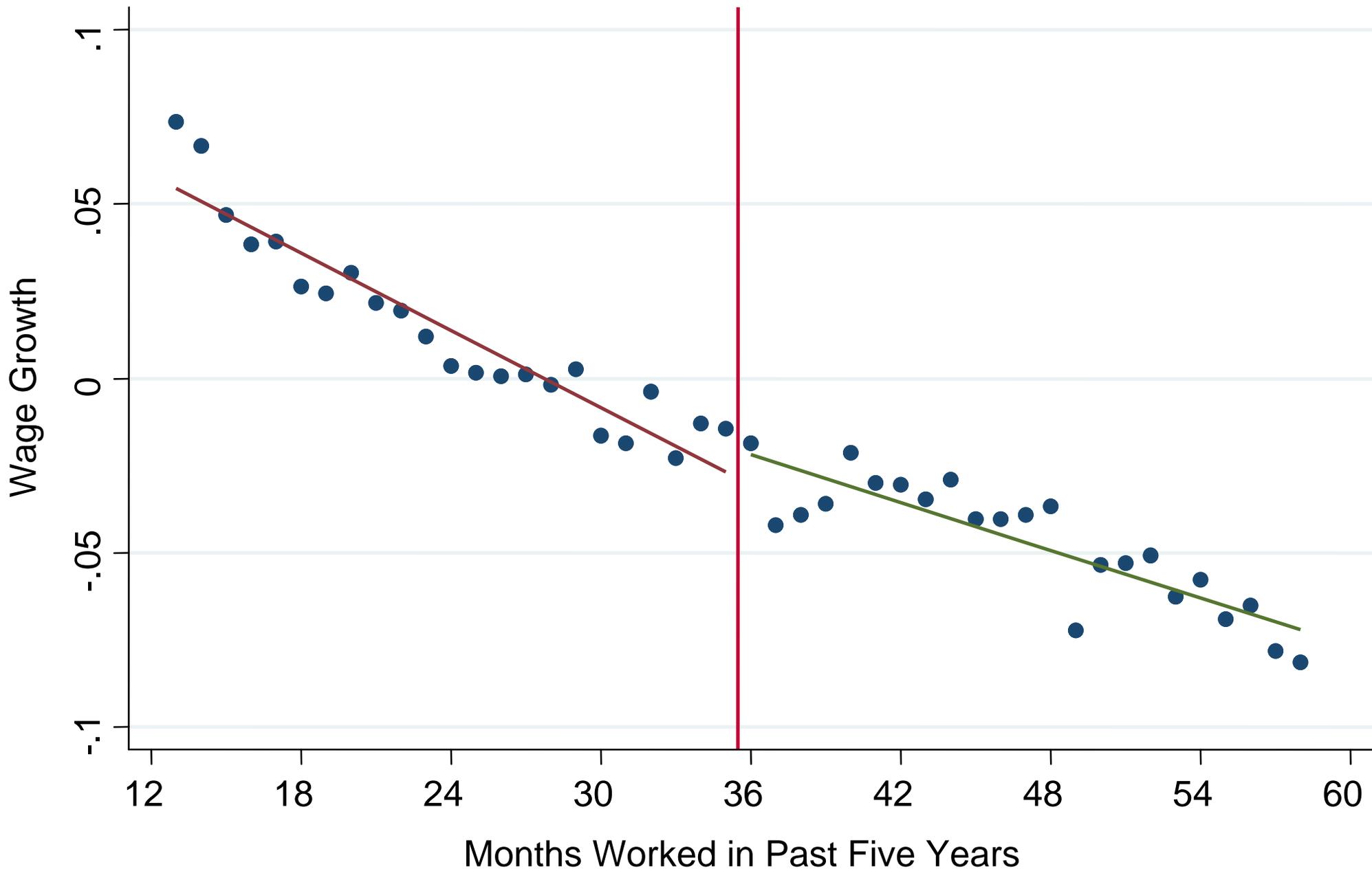
Can examine this using similar data

Look at whether people who got higher benefits and took longer to find a job are better off years later

Card-Chetty-Weber (2007) exploit again the **regression discontinuity** and find no long-term match benefit on subsequent wage or subsequent job duration

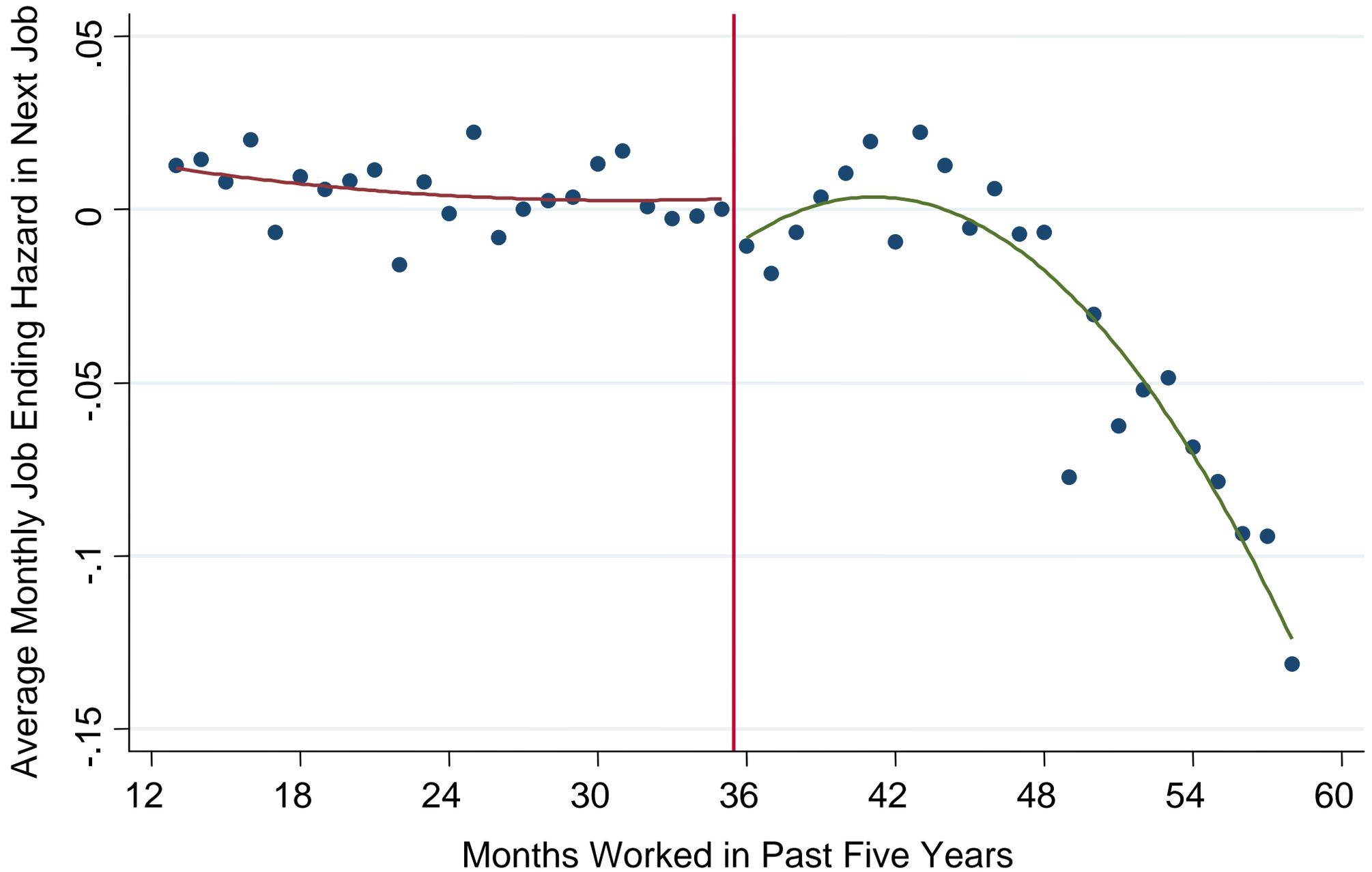
Card, Chetty, Weber (2007)

Effect of Extended Benefits on Subsequent Wages



Card, Chetty, Weber (2007)

Effect of Extended Benefits on Subsequent Job Duration



Summary of Empirical Findings on UI

1. Higher benefit level \Rightarrow longer unemployment durations
(moral hazard cost)

2. Higher benefit level \Rightarrow more consumption while unemployed
(consumption smoothing benefit)

3. UI benefits have no beneficial effects on long-term job outcomes

\Rightarrow Model implies that providing some UI is desirable but UI replacement rate should be only around 50% based on those empirical findings

UI and Firm Behavior: Experience Rating

Effect of UI on firms comes from experience rating

Perfect experience rating for firms: if firm A lays a person off, firm A pays taxes to make up for the costs firm A imposes on the UI system

Employers with a lot of layoffs get taxed more

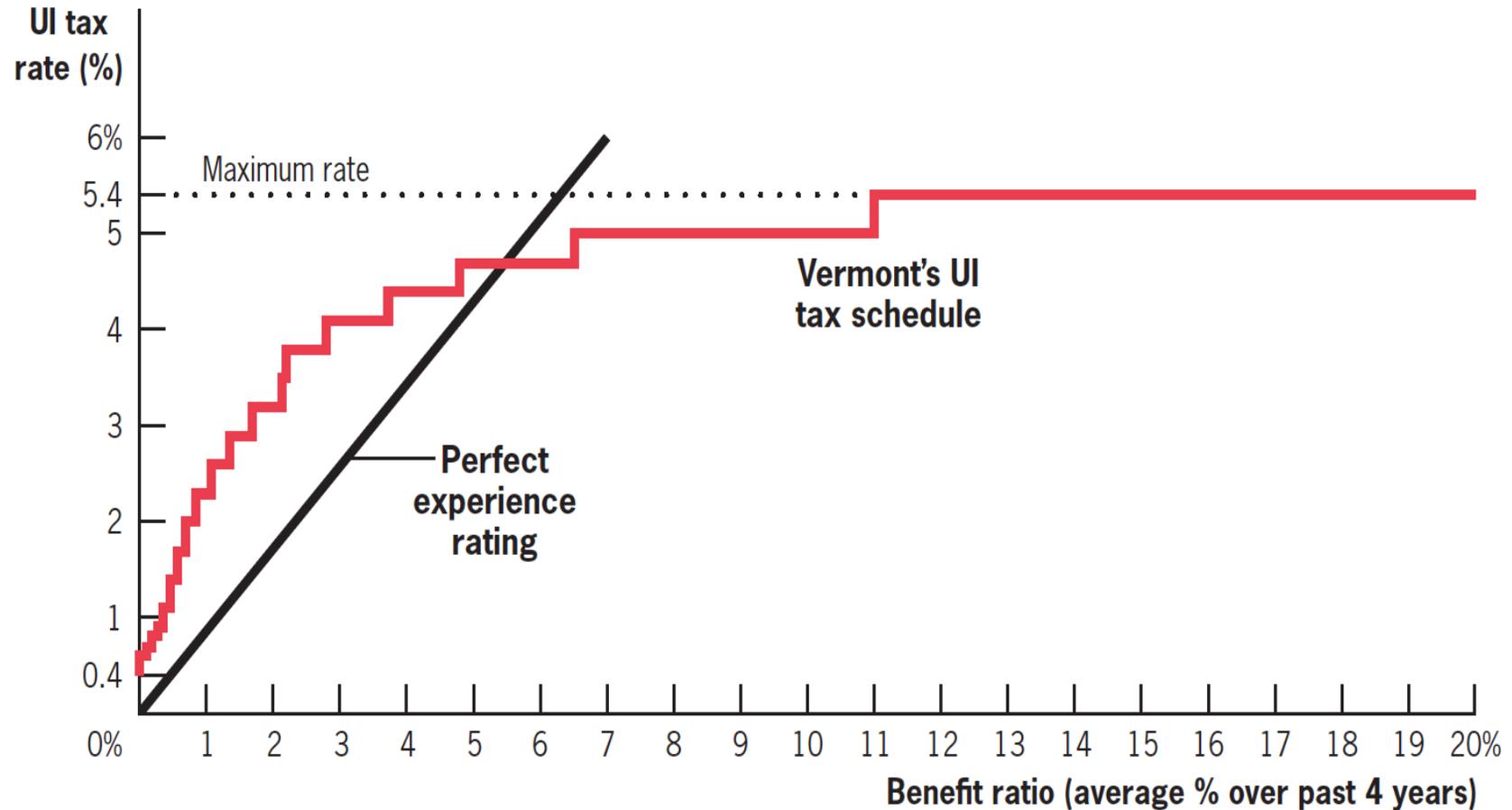
In practice, we have partial experience rating

Firm does not fully pay an additional tax each time it lays off a worker.

Payroll taxes rise less than one-for-one with layoffs because of cap on UI tax.

14.4

Partial Experience Rating in Vermont



Partial Experience Rating and Temporary Layoffs

Partial experience rating subsidizes firms with high layoff rates.

Firms and workers may make a joint decision whether to place the worker on temporary layoff, with a promise of being hired back later.

UI system makes this a partially paid vacation.

With partial experience rating, government ends up sharing in the cost of the vacation.

Partial Experience Rating and Layoffs: Evidence

Empirical studies have examined state systems with different degrees of experience rating using difference-in-difference methods

They find that partial experience rating increases the rate of temporary layoffs.

Partial experience rating alone can account for one- third of all temporary layoffs in the U.S.

The Benefits of Partial Experience Rating

What is the benefit of partial experience rating?

Fully experience rated UI would “hit firms while they are down.”

Similar to “consumption smoothing” for workers

Conceptually, tradeoff is similar to that we discussed for individual workers

But for firms, smoothing benefits are weaker

Firms have collateral - should be easier to get a loan

Most economists agree that UI should be fully experience rated

Should UI Benefits be Extended during Recessions?

US extends UI benefits during recessions. Extensions ended in 2014 (controversial policy debate)

1) Social Justice: Harder to find jobs in recessions \Rightarrow being unemployed is less of a choice \Rightarrow Extending benefits is desirable

2) Efficiency: In recessions, the job market is too slack [too hard to find jobs, too easy for firms to find workers].

a) If longer UI benefits decrease slack in labor market then longer UI benefits desirable [this is the case if UI benefits stimulate aggregate demand or if job seekers compete for a fixed number of jobs in recession, this is the left-wing view]

b) If longer UI benefits increase slack in labor market then shorter UI benefits desirable [this is the case if longer UI benefits increase the bargaining power of workers and hence increase wages further reducing labor demand, this is the right-wing view]

Economists try to tell apart a) from b) using empirical evidence

DISABILITY INSURANCE

Disability is conceptually close to retirement: some people become unable to work before old age (due to accidents, medical conditions, etc.)

All advanced countries offer public disability insurance almost always linked to the public retirement system

Disability insurance allows people to get retirement benefits before the “Early Retirement Age” if they are unable to work due to disability

US DISABILITY INSURANCE

- 1) Federal program funded by OASDI payroll tax, pays SS benefits to disabled workers under retirement age.
- 2) Program started in 1956 and became more generous over-time (age 50+ condition removed, definition of disability liberalized, replacement rate has grown)
- 3) Eligibility: Medical proof of being unable to work for at least a year, Need some prior work experience, 5 months waiting period with no earnings required (screening device)
- 4) Social security examiners rule on applications. Appeal possible for rejected applicants. Imperfect process with big type I and II errors (Parsons AER'91) ⇒ Scope for Moral Hazard
- 5) DI tends to be an absorbing state (very few go back to work)

US DISABILITY INSURANCE

- 1) In 2010, about 8m DI beneficiaries (not counting widows+children), about 5% of working age (20-64) population
- 2) Very rapid growth: In 1960, less than 1% of working age pop was on DI
- 3) Growth particularly strong during recessions: early 90s, late 00s

Key question: Are DI beneficiaries unable to work? or are DI beneficiaries not working because of DI.

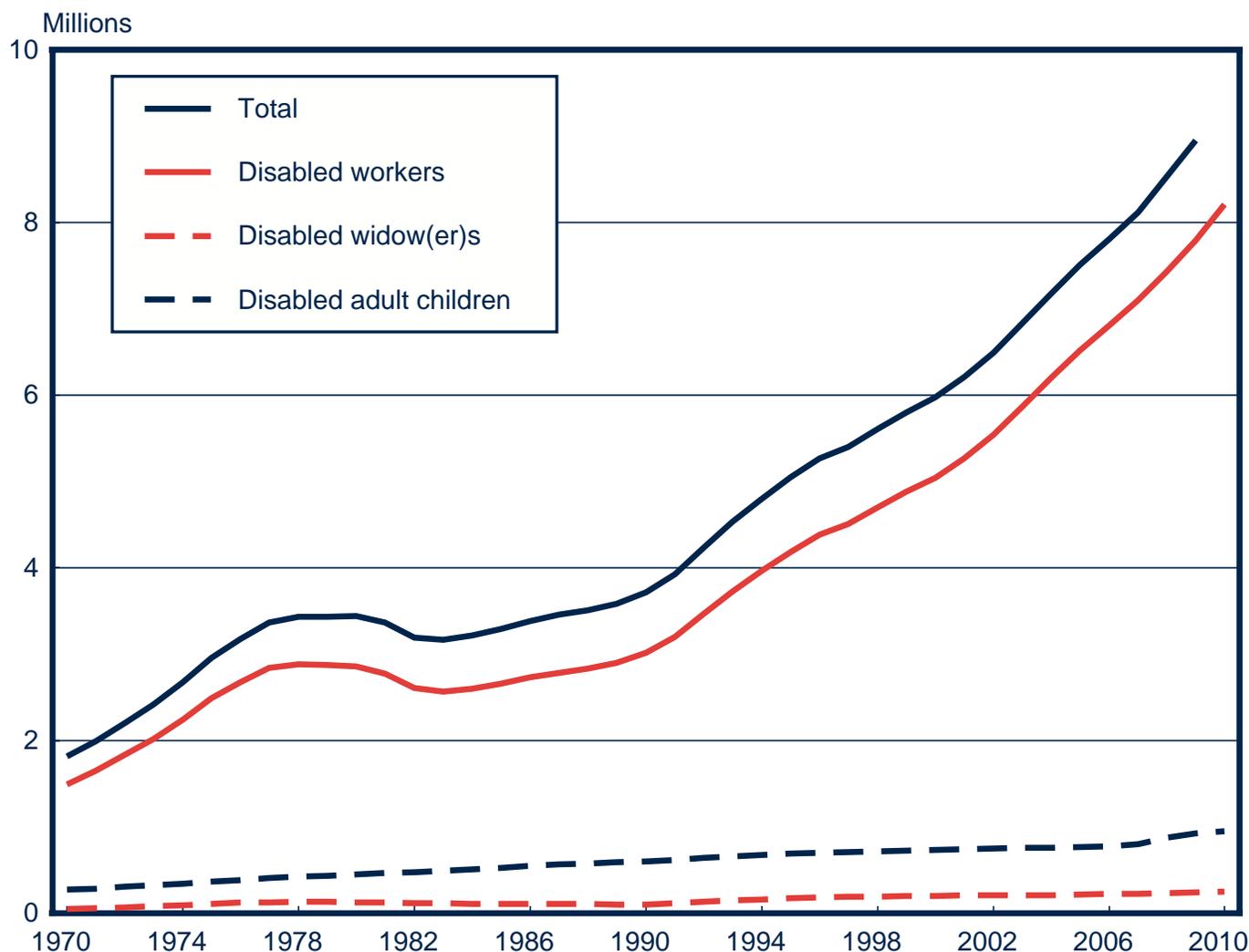
Beneficiaries in Current-Payment Status

Chart 2.

Source: SSA DI annual report

All Social Security disabled beneficiaries in current-payment status, December 1970–2010

The number of disabled workers grew steadily until 1978, declined slightly until 1983, started to increase again in 1984, and began to increase more rapidly beginning in 1990. The growth in the 1980s and 1990s was the result of demographic changes, a recession, and legislative changes. The number of disabled adult children has grown slightly, and the number of disabled widow(er)s has remained fairly level. In December 2010, slightly over 8.2 million disabled workers, over 949,000 disabled adult children, and just under 245,000 disabled widow(er)s received disability benefits.

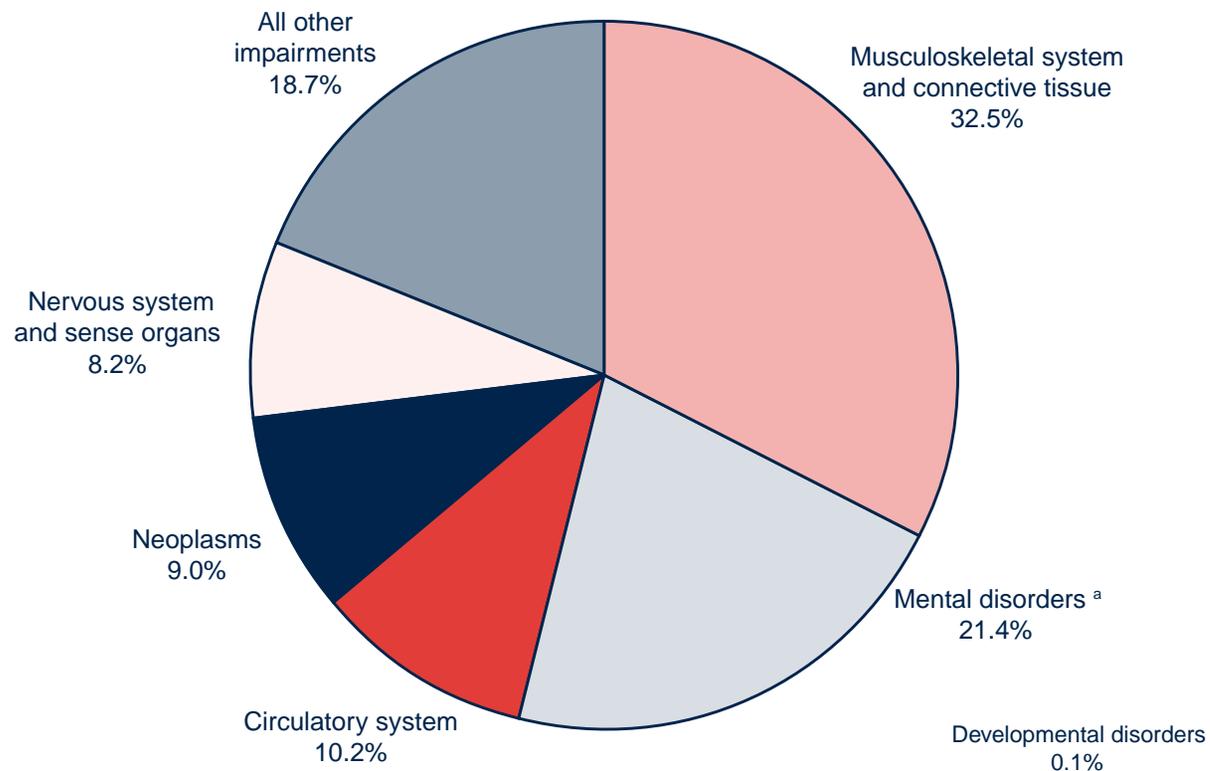


SOURCE: Table 3.

Chart 10.
Disabled-worker awards, by selected diagnostic group, 2010

Source: SSA DI annual report

In 2010, 1,026,988 disabled workers were awarded benefits. Among those awardees, the most common impairment was diseases of the musculoskeletal system and connective tissue (32.5 percent), followed by mental disorders (21.4 percent), circulatory problems (10.2 percent), neoplasms (9.0 percent), and diseases of the nervous system and sense organs (8.2 percent). The remaining 18.7 percent of awardees had other impairments.



US DISABILITY INSURANCE

Detecting disability is challenging, particularly for back injuries and mental health conditions

One way to quantify difficulty in assessment: audit study

Take a set of disability claims that was initially reviewed by a state panel

One year later, resubmit them to the panel as anonymous new claims.

Compare decisions on the **same** cases

⇒ Substantial evidence of Type I errors (incorrect rejection of a disabled person) and Type II errors (letting a non-disabled person on the program)

TABLE 1—REASSESSMENTS OF INITIAL
SOCIAL SECURITY DETERMINATIONS

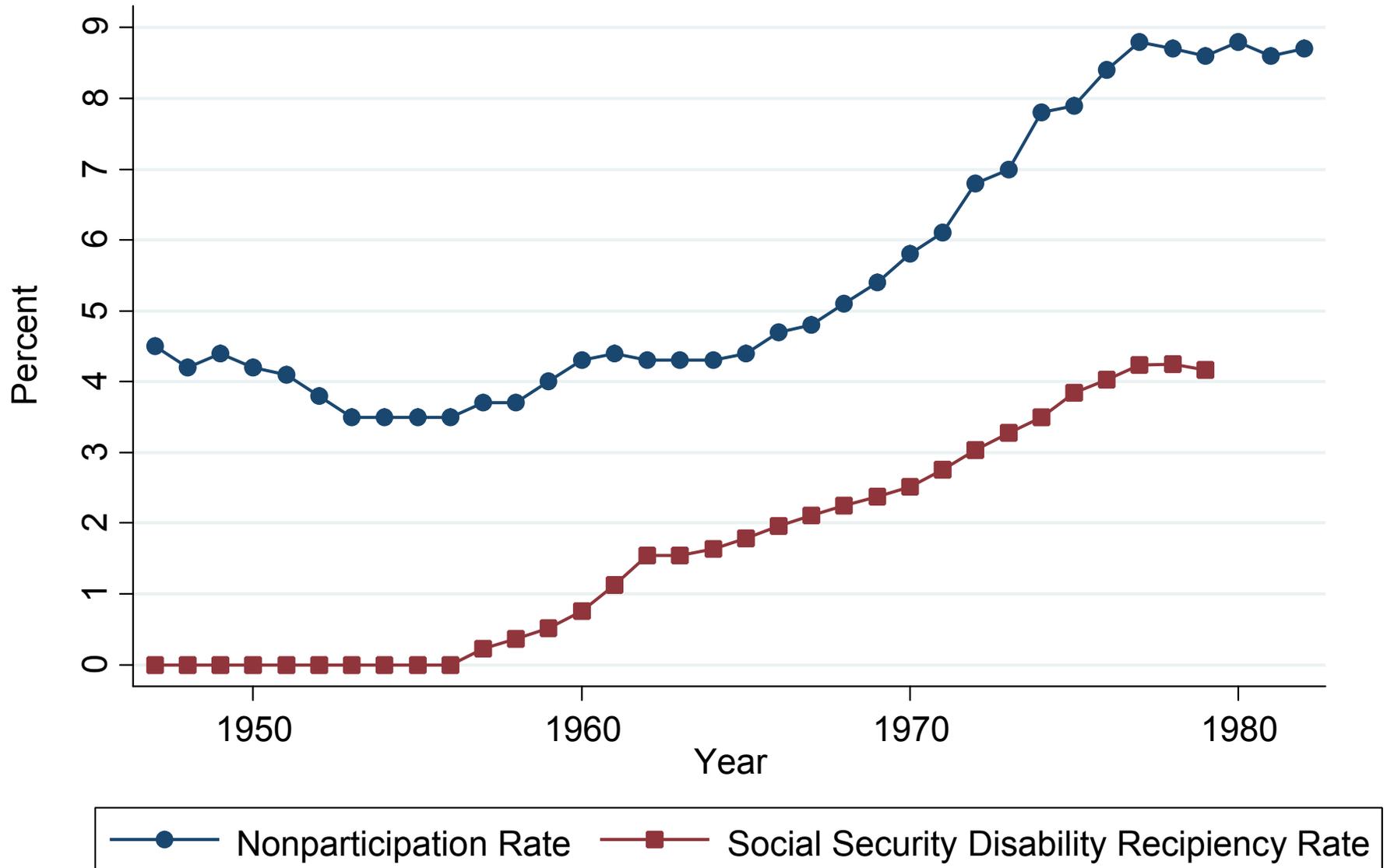
A. *Bureau of Disability Insurance Review One Year
After Initial Determination (Percentages):*

BDI assessment	Initial determination	
	Allowance	Denial
Allowance	78.8	21.1
Denial	22.5	77.5

Note: The sample sizes are 250 initial allowances and 248 initial denials.

Source: Smith and Lilienfeld (1971 p. 195).

Nonparticipation and Reciprocity Rates, Men 45-54 Years Old



Source: Parsons 1984 Table A1

DI Empirical Effects: Observational Studies

Parallel growth of DI recipients and non-participation rates among men aged 45-54 but causality link not clear

Cross-Sectional Evidence (Parsons '80): Does potential DI replacement rate have an impact on labor force participation (LFP) decision?

Uses cross-sectional variation in potential replacement rates

Survey data on men aged 45-59 from 1966-69

OLS regression

$$LFP_i = \alpha + \beta DIrepreate_i + \varepsilon_i$$

Large effect that can fully explain decline in LFP among men 45+

DI EMPIRICAL EFFECTS: OBSERVATIONAL STUDIES

Issues with Cross-Sectional Evidence:

- 1) $DIreprate_i$ depends on wages (higher for low wage earners) and likely to be correlated with ε_i (likelihood to become truly disabled)
- 2) Impossible to control fully for wages in regression because all variation in $DIreprate_i$ is due to wages
- 3) Bound AER'89 replicates Parson's regression on sample that never applied to DI and obtains similar effects implying that the OLS correlation not driven by UI

DI EMPIRICAL EFFECTS: REJECTED APPLICANTS

Bound AER'89 proposes a technique to bound effect of DI on LFP rate

Uses data on LFP on (small sample of) rejected applicants as a counterfactual

Idea: If rejected applicants do not work, then surely DI recipients would not have worked \Rightarrow Rejected applicants' LFP rate is an upper bound for LFP rate of DI recipients absent DI

Results: Only 30% of rejected applicants return to work and they earn less than half of the mean non-DI wage

\Rightarrow at most 1/3 of the trend in male LFP decline can be explained by shift to DI

Von Waechter-Manchester-Song AER'11 replicate Bound using full pop SSA admin data and confirm his results

TABLE 2—EMPLOYMENT, EARNINGS, AND OTHER CHARACTERISTICS OF REJECTED DISABILITY INSURANCE APPLICANTS

	1972			1978		
	Population	Rejected Applicants	Beneficiaries	Population	Rejected Applicants	Beneficiaries
Labor Supply						
Percent Employed	77.7	32.6	3.2	69.3	28.7	2.3
Percent Worked 71/77	91.9	45.0	7.5	86.7	40.4	5.5
Percent Full Year (≥ 50 Weeks) ^a	76.8	47.4	31.4	83.5	41.2	22.2
Percent Full Time (≥ 35 Hours) ^a	95.4	75.9	25.0	92.4	79.6	38.3
Earnings Among Positive Earners						
Median Annual Earnings, 71/77 ^b	\$9000	\$4000	\$700	\$14000	\$5300	\$1000

DI EMPIRICAL EFFECTS: REJECTED APPLICANTS

Maestas-Mullen-Strand AER'13 obtain causal effect of DI on LFP using natural variation in DI examiners' stringency and large SSA admin data linking DI applicants and examiners

Idea: (a) Random assignment of DI applicants to examiners and (b) examiners vary in the fraction of cases they reject \Rightarrow Valid instrument of DI receipt

Result 1: DI benefits reduce LFP of applicants by 28 points \Rightarrow DI has moderate impact (consistent with Bound AER'89)

Result 2: DI has heterogeneous impact: small effect on those severely impaired but big effect on less severely impaired

Tough judges marginal cases unlikely to work without DI, lenient judges marginal cases somewhat likely to work without DI

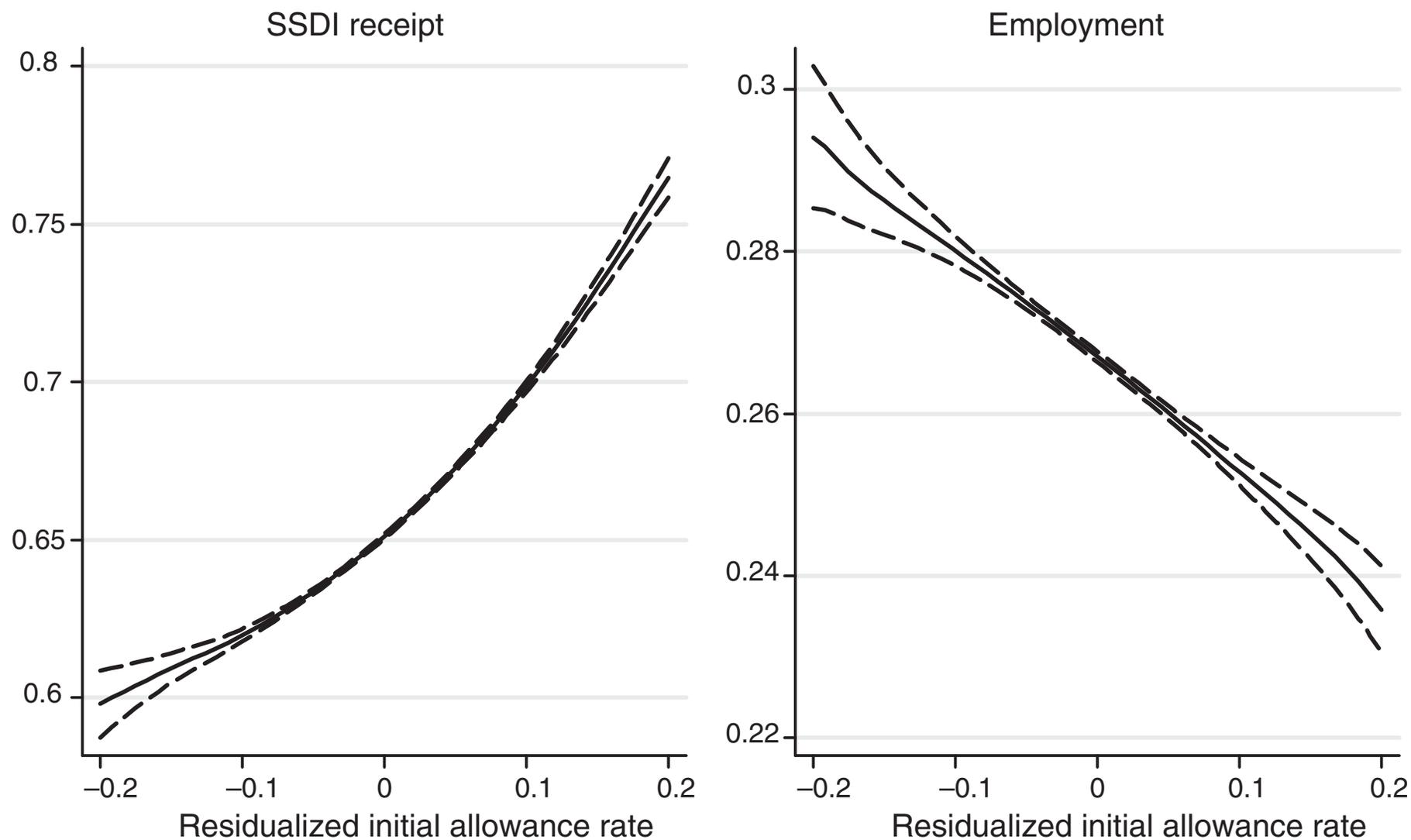


FIGURE 4. SSDI RECEIPT AND LABOR SUPPLY BY INITIAL ALLOWANCE RATE

Notes: Ninety-five percent confidence intervals shown with dashed lines. Employment measured in the second year after the initial decision. Bandwidth is 0.116 for DI and 0.130 for labor force participation.

Source: DIODS data for 2005 and 2006

Workers Compensation: Institutional Features

Workers compensation is insurance for injuries on the job, mainly temporary injuries that prevent work (short-term)

Workers Compensation is a state-level program

Two components: medical and indemnity

Indemnity payment replaces roughly two-thirds of lost wages.

Unlike UI, WC payments are untaxed, leading to a higher replacement that is near 90% on average.

Substantial variation across states in benefit levels

Workers Compensation (WC): Institutional Features

1) Workers comp is a mandated benefit; no explicit tax but firms required by law to provide this benefit to workers

Most firms choose to buy coverage from private insurers

Premiums are more tightly experience rated than UI because they are determined by private sector

Insurance companies charge high-risk firms more.

2) Important feature of WC: no-fault insurance.

When there is a qualifying injury, WC benefits paid regardless of whether the injury was the worker's or the firm's fault.

Idea: reduce inefficiency of tort system (legal costs) by having fixed rules and not worrying about liability

Moral Hazard in Workers? Compensation

Moral hazard in WC can manifest itself in reported injuries, injury durations, and types of injuries reported.

E.g. easier to report back pain—very hard to verify

Huge issue in CA—companies pay very high workers comp rates

Governor Schwarzenegger reform in 2004 cut benefits sharply, claiming to reduce injuries and “open CA for business”

Is it true that there is substantial moral hazard?

Again, consider several pieces of evidence

Strategy 1: Timing of injuries. “Monday effect” (faking weekend injuries into work injuries)

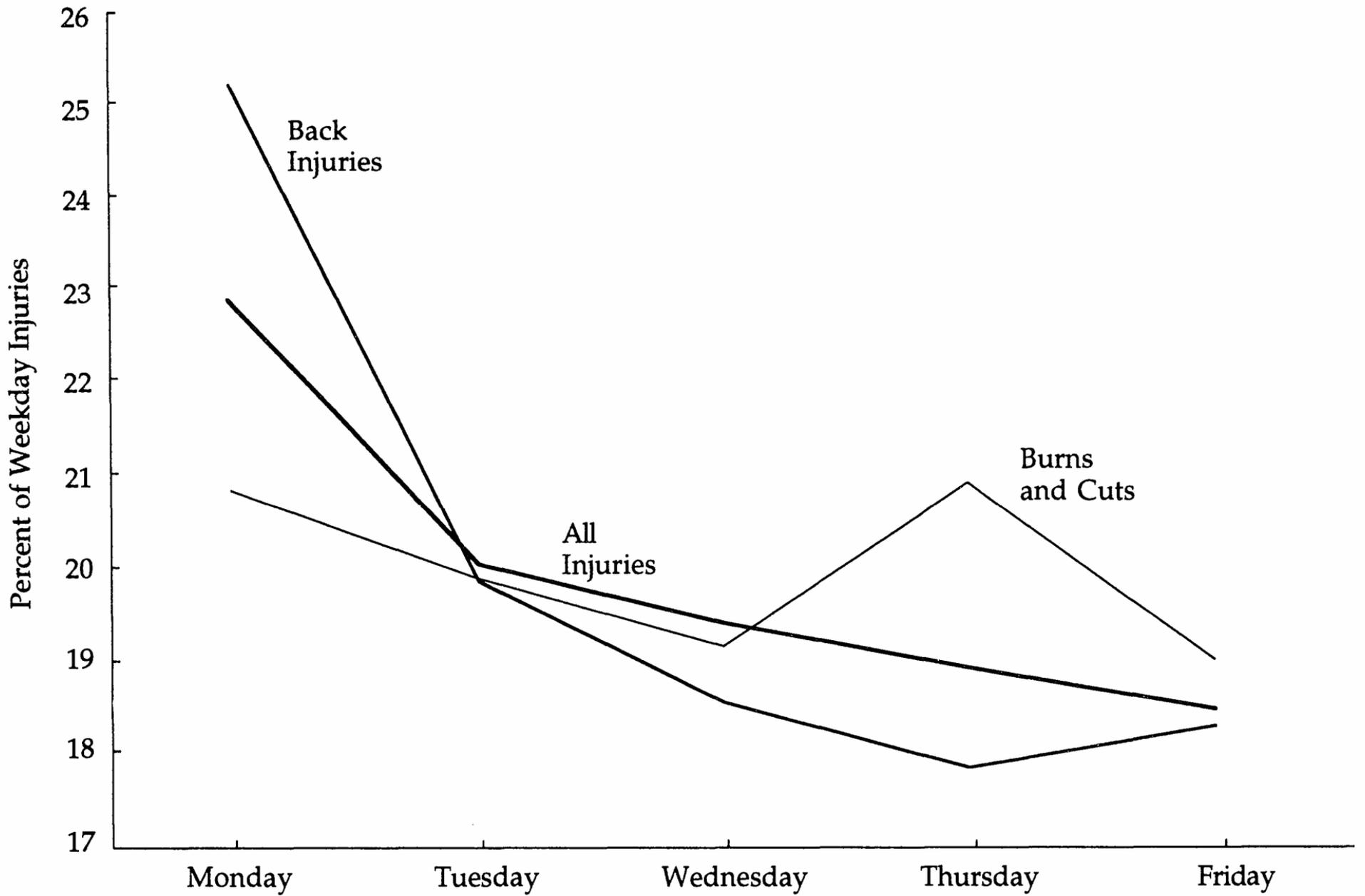


Figure 1. Distribution of Weekday Injuries.

Source: Card and McCall 1996

Moral Hazard in Workers? Compensation

Strategy 2: examine effect of workers comp benefit levels on durations using a diff-in-diff strategy (Meyer, Viscusi, Durbin 1995)

Reforms in Kentucky and Michigan that increased benefits for high-earning workers (but not low-earning workers) in late 1980s

Compare changes in injury durations and medical costs for high-earners vs. low earners in those states before and after reform

Variable	Kentucky			Michigan		
	Before increase (1)	After increase (2)	Percentage change (3)	Before increase (4)	After increase (5)	Percentage change (6)
Maximum benefit (\$)	131.00	217.00	65.65	181.00	307.00	69.61
Replacement rate, high earnings (percent)	32.70 (0.25)	51.02 (0.37)	56.02 (1.65)	30.01 (0.35)	44.15 (0.48)	47.14 (2.33)
Replacement rate, low earnings (percent)	66.42 (0.20)	66.66 (0.22)	0.36 (0.44)	66.64 (0.24)	66.35 (0.30)	-0.45 (0.58)

Source: Meyer, Viscusi, Durbin 1995

TABLE 4—KENTUCKY AND MICHIGAN: DURATION AND MEDICAL COSTS OF TEMPORARY TOTAL DISABILITIES DURING THE YEARS BEFORE AND AFTER BENEFIT INCREASES

Variable	High earnings		Low earnings		Differences		Difference in differences
	Before increase (1)	After increase (2)	Before increase (3)	After increase (4)	[(2)–(1)] (5)	[(4)–(3)] (6)	[(5)–(6)] (7)
Mean duration (weeks)							
Kentucky	11.16 (0.83)	12.89 (0.83)	6.25 (0.30)	7.01 (0.41)	1.72 (1.17)	0.76 (0.51)	0.96 (1.28)
Michigan	14.76 (2.25)	19.42 (2.67)	10.94 (1.09)	13.64 (1.56)	4.66 (3.49)	2.70 (1.90)	1.96 (3.97)
Median duration (weeks)							
Kentucky	4.00 (0.14)	5.00 (0.20)	3.00 (0.11)	3.00 (0.12)	1.00 (0.25)	0.00 (0.16)	1.00 (0.29)
Michigan	5.00 (0.45)	7.00 (0.67)	4.00 (0.22)	4.00 (0.28)	2.00 (0.81)	0.00 (0.35)	2.00 (0.89)
Median medical cost (dollars)							
Kentucky	393.51 (19.29)	411.49 (22.72)	238.96 (8.48)	254.40 (9.11)	17.98 (29.80)	15.44 (12.44)	2.55 (32.30)
Michigan	689.73 (77.30)	765.00 (134.53)	390.63 (32.80)	435.00 (33.09)	75.27 (155.16)	44.38 (46.59)	30.89 (162.00)

Source: Meyer, Viscusi, Durbin 1995

Moral Hazard in Workers' Compensation

Result: 10% increase in WC benefit raises out-of-work duration due to injury by 4%

Again, need to weigh this against benefits to reach policy conclusions

Give people more time to heal after injury without rushing them back to work

Higher consumption while out of work

No evidence yet on these issues

CONCLUSION

Individuals clearly value the consumption smoothing provided by social insurance programs

In each case there are moral hazard costs associated with the provision of the insurance

Empirical analyses of all three programs can be used to inform policy makers' decisions as program reforms move forward

REFERENCES

Jonathan Gruber, Public Finance and Public Policy, Fourth Edition, 2012
Worth Publishers, Chapter 14

Bound, John. "The Health and Earnings of Rejected Disability Insurance Applicants." *American Economic Review* 79.3 (1989): 482-503.(web)

Card, David, and Brian P. McCall. "Is Workers' Compensation Covering Uninsured Medical Costs? Evidence from the "Monday Effect"." *Industrial and Labor Relations Review* (1996): 690-706.:(web)

Card, David, Raj Chetty, and Andrea Weber. "The Spike at Benefit Exhaustion: Leaving the Unemployment System or Starting a New Job?." *The American Economic Review* 97.2 (2007): 113-118.(web)

Gruber, Jonathan. "The Consumption Smoothing Benefits of Unemployment Insurance." *The American Economic Review* (1997).(web)

Gruber, Jonathan. "Disability Insurance Benefits and Labor Supply." *The Journal of Political Economy* 108.6 (2000): 1162-1183.(web)

Gruber, Jonathan. "Covering the Uninsured in the United States." *Journal of Economic Literature*, 43.3 (2008): 571-606.(web)

Maestas, Nicole, Kathleen Mullen and Alexander Strand "Does Disability Insurance Receipt Discourage Work? Using Examiner Assignment to

Estimate Causal Effects of SSDI Receipt”, American Economic Review, 103(5), 2013, 1797-1829. (web)

Meyer, Bruce D. “Unemployment Insurance and Unemployment Spells.” Econometrica 58.4 (1990): 757-782.(web)

Meyer, Bruce D., W. Kip Viscusi, and David L. Durbin. “Workers’ compensation and injury duration: evidence from a natural experiment.” The American Economic Review (1995): 322-340.(web)

Parsons, Donald O. “The decline in male labor force participation.” The Journal of Political Economy (1980): 117-134.(web)

Parsons, Donald O. “The health and earnings of rejected disability insurance applicants: comment.” The American Economic Review 81.5 (1991): 1419-1426.(web)

Smith, Richard Thomas, and Abraham M. Lilienfeld. “The social security disability program: an evaluation study.”[Book] No. 39. US Social Security Administration, Office of Research and Statistics, 1971.

Von Wachter, Till, Jae Song, and Joyce Manchester. “Trends in employment and earnings of allowed and rejected applicants to the social security disability insurance program.” American Economic Review 101.7 (2011): 3308.(web)

Undergraduate Public Economics

Danny Yagan
UC Berkeley

Lecture 19 Health Insurance

MOTIVATION

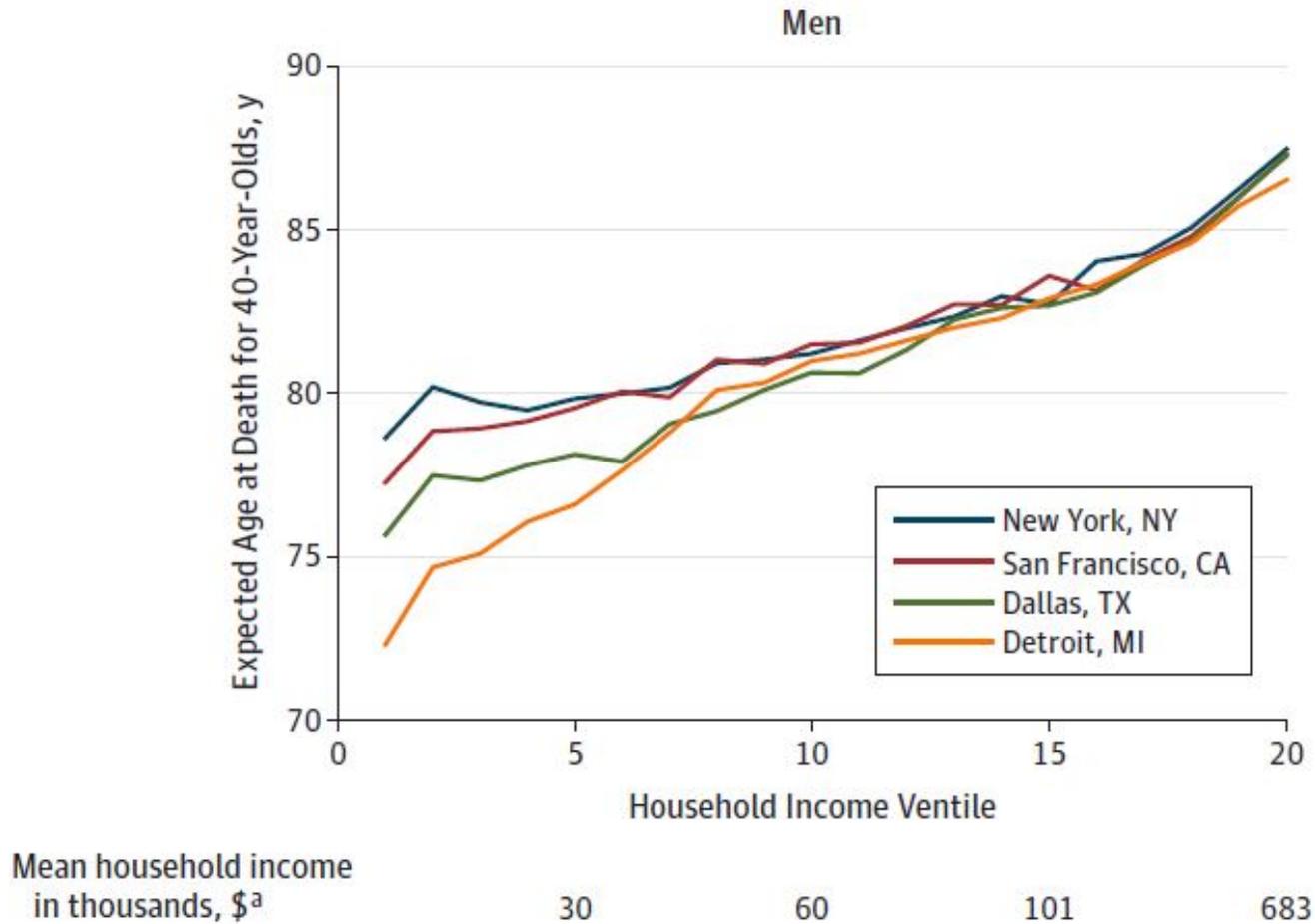
Huge benefits from the U.S. health care system, but big issues: (a) US health care is very expensive (18% of GDP relative to 10% in other OECD countries), (b) costs growing quickly, (c) significant fraction of population is uninsured

There are enormous disparities in medical outcomes across demographic groups in the US

U.S. is only major industrialized nation that does not provide universal access to health care for its citizens

Recent ObamaCare law (Affordable Care Act) has drastically reduced the number of uninsured from 50m (in 2013) to 25m (in 2017+). May be repealed.

LIFE EXPECTANCY BY INCOME, GEOGRAPHY



Source: Chetty-Stepner-Abraham-Lin-Scuderi-Turner-Bergeron-Cutler 2016.
40-year-olds, race-adjusted.

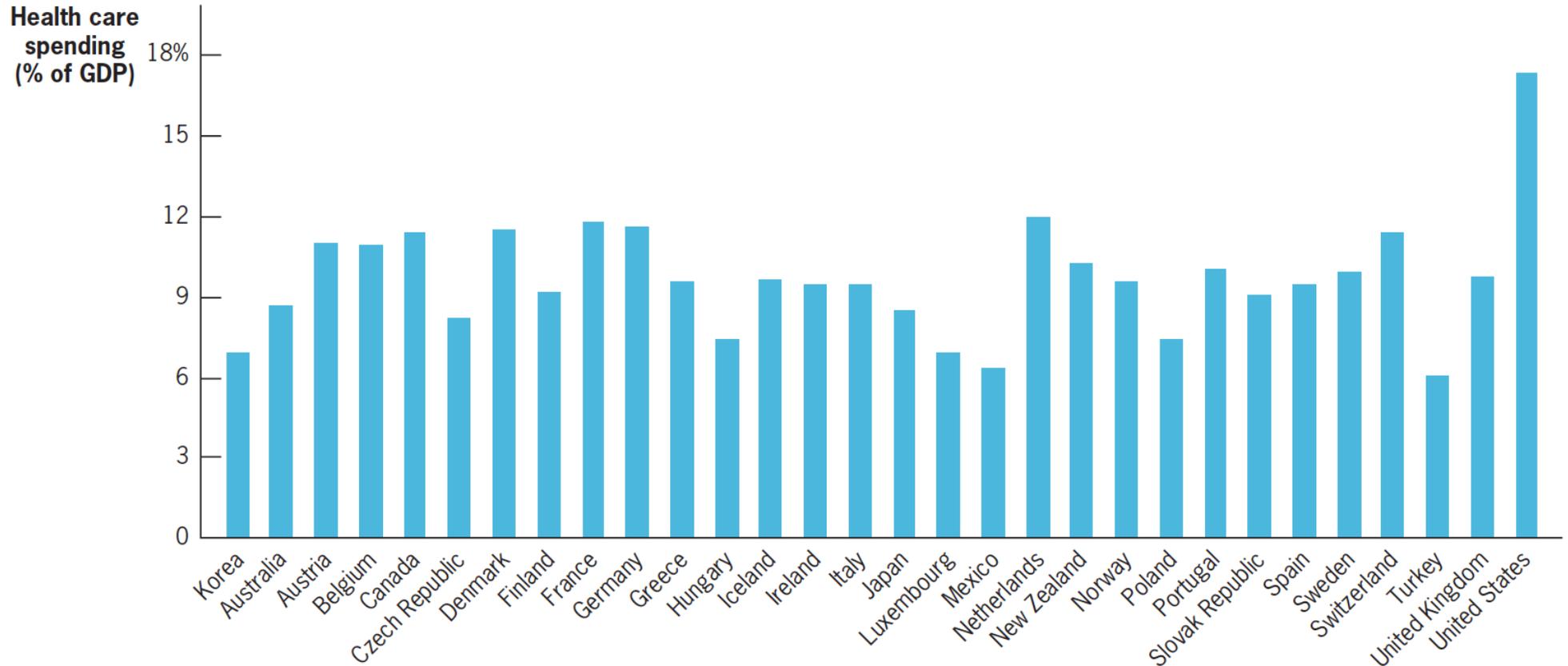
LIFE EXPECTANCY BY INCOME, GEOGRAPHY



Source: Chetty-Stepner-Abraham-Lin-Scuderi-Turner-Bergeron-Cutler 2016.
40-year-olds, race-adjusted.

15.1

Healthcare Spending in the OECD Nations



UNIVERSAL HEALTH INSURANCE

All OECD countries (except the US) provide universal health care insurance funded by progressive income taxation:

Individuals who get sick can have health care paid for by the government

Government either directly controls doctors/hospitals (like National Health Service in the UK) or government reimburses private health care providers (“single-payer” like in France)

Government controls costs and limits health-care over-consumption through:

- 1) Regulation (govt picks allowed treatments based on cost effectiveness, bargains for prices, rations care)
- 2) Patient co-payments (patients share part of the cost)

US HEALTH INSURANCE

US has a mix of public and private insurance:

1) Government provided insurance

(a) Medicare for the elderly (65+), (b) Medicaid for the poor, (c) Veterans benefits (TRICARE/CHAMPVA)

2) Privately provided insurance:

(a) Employer provided health insurance (large), (b) Direct private purchase (small), (c) ObamaCare exchanges

3) Uninsured: (1/6) of population before ObamaCare

In the US, health insurance typically restricts treatments on effectiveness (not cost effectiveness) \Rightarrow Huge incentives for health providers to supply new expensive treatments

15.1

Americans' Source of Health Insurance Coverage, 2010

	People (millions)	Population %
<i>Private</i>	201.0	64.0
Employment-based	176.3	55.3
Direct purchase	26.8	9.8%
<i>Public</i>	87.4	31.0
Medicare	43	14.5
Medicaid	42.6	15.9
TRICARE/CHAMPVA	11.6	4.2
<i>Uninsured</i>	46.2	16.3

WHY EMPLOYERS PROVIDE PRIVATE INSURANCE

1) risk pooling: The goal of all insurers is to create *large insurance pools with a predictable distribution of medical risk.*

2) tax incentive: employer provided health insurance is a non-taxable form of compensation for employees (not subject to payroll taxes or individual income tax)

⇒ Fiscally advantageous to get insurance through employer (non-taxable) than to purchase it directly as an individual (with after-tax income)

15.1

Illustrating the Tax Subsidy

	Jim	Peter
Wage	30	30
Employer health insurance spending	0	5
Pre-tax wage	30	25
After-tax wage	20	16.67
Personal health spending	4	0
After-tax, after-health spending income	16	16.67

NONGROUP INSURANCE

Nongroup direct insurance market: The market through which individuals or families buy insurance directly rather than through a group, such as the workplace.

Before ObamaCare in the nongroup market: (A) those in the worst health (publicly known pre-existing conditions) were charged enormous prices, (B) those in good health also faced high prices due to **adverse selection** [you privately know you're low-risk → you don't buy]

ObamaCare: Addressed A by forbidding pricing/ discrimination based on preexisting conditions and addressed B by mandating health insurance

MEDICARE

Started in 1965 as a universal health insurance system for the elderly and nonelderly on disability insurance.

Federal program that provides health insurance to all people over age 65 and disabled

Every citizen who has worked for 10 years (or their spouse) is eligible.

Financed with an uncapped payroll tax totaling 2.9%

Physician reimbursement fairly generous (but not as high as private insurance)

MEDICAID

Provides health care for the poor (means-tested benefit)

Financed from general revenues

Targets welfare recipients, low income kids and elderly (for non-Medicare costs such as long-term care)

70% of recipients are mothers/kids but 66% of expenditure goes to long-term care for elderly.

Doctor reimbursement low \Rightarrow some docs refuse Medicaid

Big variation across states in Medicaid generosity (costs are shared between state/feds)

Program eligibility criteria have been expanded over time (higher incomes allowed, including ObamaCare in most states)

16.3

The Medicare Program

The largest public health insurance program in the United States is Medicare.

■ TABLE 16-2

Medicaid and Medicare

	Medicaid	Medicare
Eligibles	Families on welfare Low-income children, pregnant women Low-income elderly, disabled	Retirees and spouses 65 and older Certain disabled individuals under 65 People with kidney failure (requiring dialysis or transplant)
Premiums	None	Hospital coverage: none Physician coverage: \$66.60 per month Prescription drug coverage: Variable
Deductibles/copayments	None (or very small)	Hospital coverage: \$1,068 deductible for first 60 days Physician coverage: \$135 deductible, 20% coinsurance Prescription drug coverage: Variable
Services excluded	None (or very minor)	Prescription drugs (until 2006) Routine checkups, dental care, nursing home care, eyeglasses, hearing aids, immunization shots
Provider reimbursement	Very low	Moderate (but falling)

Medicaid provides health insurance for low-income individuals, covering a wide range of health services at little cost to those individuals. Medicare provides health insurance for those age 65 and over, covering many, though not all, health services at some cost to those individuals.

Is Universal Health Care Desirable?

People face different health risks (pre-existing conditions) ⇒
Those facing high health risks face very high insurance costs
in private market

Should the government insure people for health risks? Re-
distribution view: Yes. Adverse selection view: Yes. Moral
hazard / fairness view: No if health risks due to choices (diet,
exercise) instead of genetics or accidents. Leviathan view: No
("price of freedom").

Virtually all OECD countries answer yes and provide universal
health care.

In all cases (private and public), health insurance needs to deal
with moral hazard (over-provision, over-consumption)

Optimal Health Insurance: Consumer Side

As with other insurance, optimal generosity determined by the trade off between consumption-smoothing benefit and moral hazard cost.

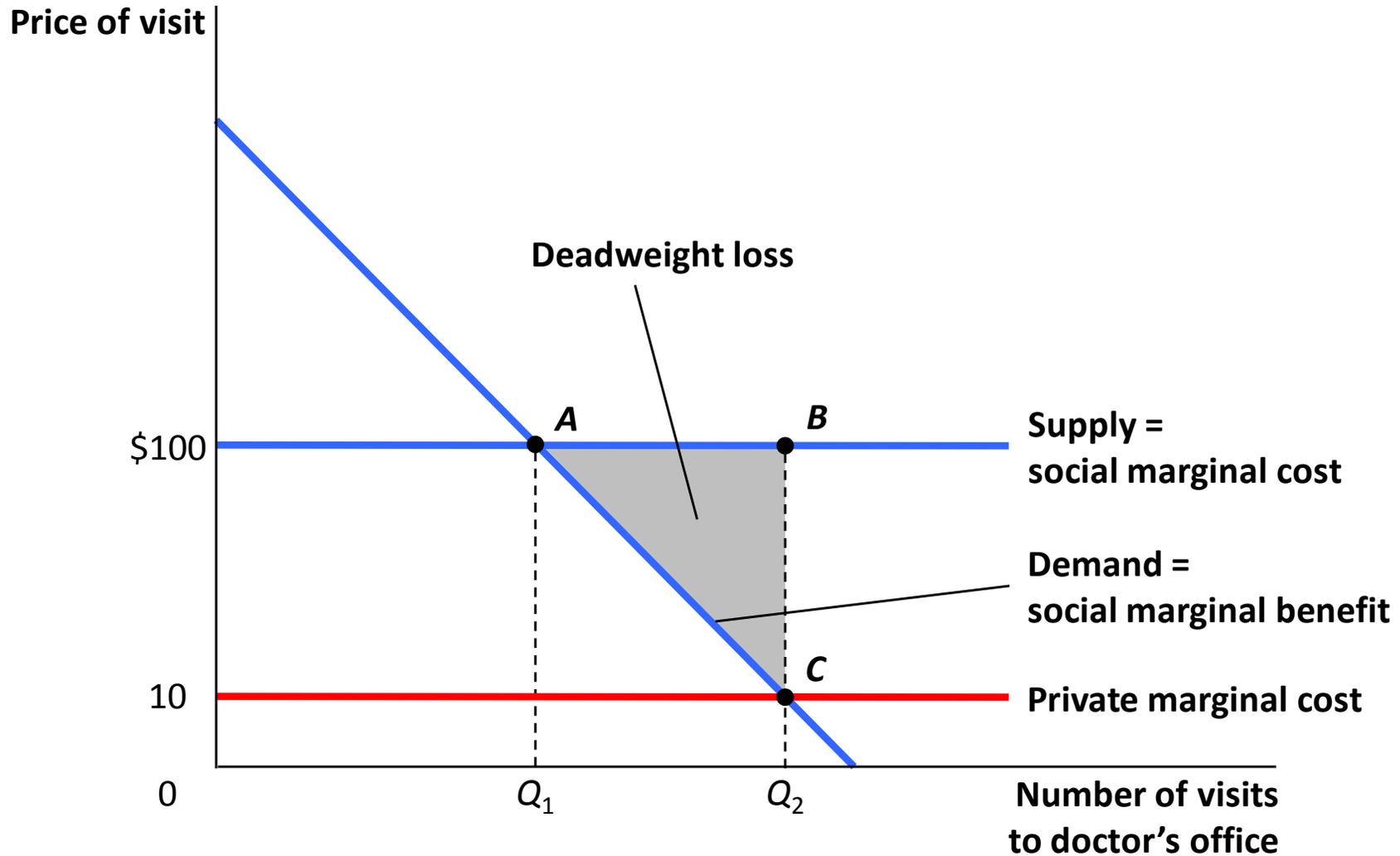
Consumption when sick = $c_s < c_h$ = consumption when healthy

Insurance raises c_s and lowers $c_h \Rightarrow$ higher expected utility if risk averse.

Moral hazard: overconsumption of healthcare because insured individual pays only a fraction of health care costs when he/she is sick. Fraction paid by individual is called the **co-payment**

15.2

Moral Hazard Costs of Health Insurance for Patients



How Elastic Is the Demand for Medical Care? The RAND Health Insurance Experiment

The best evidence on the elasticity of demand for medical care comes from one of the most ambitious social experiments in U.S. history: the RAND Health Insurance Experiment (HIE) in late 1970s

\$150m expenditure involving 6000 people tracked over 3 years

Random assignment of health plans with different co-payment parameters: Copayment rates from 0% to 95%.

All families given \$1000 to participate, so no one was made worse off from the experiment.

The RAND Health Insurance Experiment: Results

Medical care demand is somewhat price sensitive: individuals who were in the free care plan used 46% more care than those paying 95% of their medical costs.

Overall, 10% rise in the price of medical care to individuals \Rightarrow use 2% less care (elasticity = .2). Medical utilization not very sensitive to price but distortion still large due to very low co-payment rates in most insurance programs

Those who used more health care due to the lower price did not, on average, see a significant improvement in their health.

For those who were chronically ill and don't have sufficient income to easily cover co-payments, there was some deterioration in health.

Oregon Medicaid Health Insurance Experiment

- In 2008, Oregon had a limited Medicaid budget \Rightarrow used lottery to select individuals on waitlist to be given a chance to apply for Medicaid insurance coverage
- 30,000 “lottery winners” (treatment group) out of 90,000 participants (lottery losers are control group)

Not all winners received coverage. Some non-winners later received insurance on their own.

But it is still the case that winning the lottery increases probability of having health insurance by 29 percentage points

- Finkelstein et al. (2012) use lottery as instrument to estimate causal effect of insurance coverage itself

Two way to report the results:

ITT (intention to treat): just compare winners and losers

LATE (local average treatment effect): Inflate estimates by $1/[\text{difference in fraction insured between winners and losers}] = 1/.29 = 3.5$

Oregon Medicaid Health Insurance Experiment

- Data sources: admin data from hospitals, credit reporting data, and survey responses regarding utilization, health, and financial outcomes
- Key results: winning the Medicaid lottery leads to:
 - 1) higher health care utilization (including primary and preventive care as well as hospitalizations)
 - 2) lower out-of-pocket medical expenditures and medical debt (including fewer bills sent to collection agencies for unpaid debt)
 - 3) better self-reported physical and mental health

Table V: Health Care Utilization (Survey Data)

	Extensive Margin (Any)				Total Utilization (Number)			
	Control Mean	ITT	LATE	p-values	Control Mean	ITT	LATE	p-values
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Prescription drugs currently	0.637 (0.481)	0.025 (0.0083)	0.088 (0.029)	[0.002] {0.005}	2.318 (2.878)	0.100 (0.051)	0.347 (0.176)	[0.049] {0.137}
Outpatient visits last six months	0.574 (0.494)	0.062 (0.0074)	0.212 (0.025)	<0.0001 {<0.0001}	1.914 (3.087)	0.314 (0.054)	1.083 (0.182)	<0.0001 {<0.0001}
ER visits last six months	0.261 (0.439)	0.0065 (0.0067)	0.022 (0.023)	[0.335] {0.547}	0.47 (1.037)	0.0074 (0.016)	0.026 (0.056)	[0.645] {0.643}
Inpatient Hospital admissions last six months	0.072 (0.259)	0.0022 (0.0040)	0.0077 (0.014)	[0.572] {0.570}	0.097 (0.4)	0.0062 (0.0062)	0.021 (0.021)	[0.311] {0.510}
<i>Standardized treatment effect</i>		0.050 (0.011)	0.173 (0.036)	<0.0001		0.040 (0.011)	0.137 (0.038)	[0.0003]
<i>Annual spending^a</i>					3,156	226 (108)	778 (371)	[0.037]

Table VIII: Financial Strain (Survey Data)

	Control Mean	ITT	LATE	p-values
	(1)	(2)	(3)	(4)
Any out of pocket medical expenses, last six months	0.555 (0.497)	-0.058 (0.0077)	-0.200 (0.026)	[<0.0001] {<0.0001}
Owe money for medical expenses currently	0.597 (0.491)	-0.052 (0.0076)	-0.180 (0.026)	[<0.0001] {<0.0001}
Borrowed money or skipped other bills to pay medical bills, last six	0.364 (0.481)	-0.045 (0.0073)	-0.154 (0.025)	[<0.0001] {<0.0001}
Refused treatment bc of medical debt, last six months	0.081 (0.273)	-0.011 (0.0041)	-0.036 (0.014)	[0.01] {0.01}
<i>Standardized treatment effect</i>		-0.089 (0.010)	-0.305 (0.035)	[<0.0001]

Table IX: Health

	Control Mean (1)	ITT (2)	LATE (3)	p-values (4)
Panel A: Administrative data				
Alive	0.992 (0.092)	0.00032 (0.00068)	0.0013 (0.0027)	[0.638]
Panel B: Survey Data				
Self reported health good / very good / excellent (not fair or poor)	0.548 (0.498)	0.039 (0.0076)	0.133 (0.026)	[<0.0001] {<0.0001}
Self reported health not poor (fair, good, very good, or excellent)	0.86 (0.347)	0.029 (0.0051)	0.099 (0.018)	[<0.0001] {<0.0001}
Health about the same or gotten better over last six months	0.714 (0.452)	0.033 (0.0067)	0.113 (0.023)	[<0.0001] {<0.0001}
# of days physical health good, past 30 days*	21.862 (10.384)	0.381 (0.162)	1.317 (0.563)	[0.019] {0.018}
# days poor physical or mental health did not impair usual activity, past 30 days*	20.329 (10.939)	0.459 (0.175)	1.585 (0.606)	[0.009] {0.015}
# of days mental health good, past 30 days*	18.738 (11.445)	0.603 (0.184)	2.082 (0.64)	[0.001] {0.003}
Did not screen positive for depression, last two weeks	0.671 (0.470)	0.023 (0.0071)	0.078 (0.025)	[0.001] {0.003}
<i>Standardized treatment effect</i>		0.059 (0.011)	0.203 (0.039)	[<0.0001]

Source: Finkelstein et al. 2012

Table X: Potential Mechanisms for Improved Health (Survey Data)

	Control Mean (1)	ITT (2)	LATE (3)	p-values (4)
Panel A: Access to care				
Have usual place of clinic-based care	0.499 (0.500)	0.099 (0.0080)	0.339 (0.027)	[<0.0001] {<0.0001}
Have personal doctor	0.490 (0.500)	0.081 (0.0077)	0.280 (0.026)	[<0.0001] {<0.0001}
Got all needed medical care, last six months	0.684 (0.465)	0.069 (0.0063)	0.239 (0.022)	[<0.0001] {<0.0001}
Got all needed drugs, last six months	0.765 (0.424)	0.056 (0.0055)	0.195 (0.019)	[<0.0001] {<0.0001}
Didn't use ER for non-emergency, last six months	0.916 (0.278)	-0.0011 (0.0043)	-0.0037 (0.015)	[0.804] {0.804}
<i>Standardized treatment effect</i>		0.128 (0.0084)	0.440 (0.029)	[<0.0001]

Source: Finkelstein et al. 2012

Consumption-Smoothing Benefits

Consumption-smoothing benefits bigger for large shocks

Some events, like a check-up, are minor and predictable

Others, like a heart attack, are expensive and unpredictable.

Insurance is much more valuable for expensive, unpredictable events

Small shocks lead to small fluctuations in marginal utility

Also less moral hazard for large, unpredictable shocks

⇒ Optimal policy: large deductibles and very generous coverage for “catastrophes”

Application: Medicare Prescription Drug Benefit

Starting in 2006, Medicare “Part D” covers drug expenses.

In return for a monthly premium, this program pays for

0% of the drug costs up to \$250

75% of the costs for the next \$2,250

0% of the costs for the next \$3,600 (“donut hole”)

95% of the costs above \$5,100

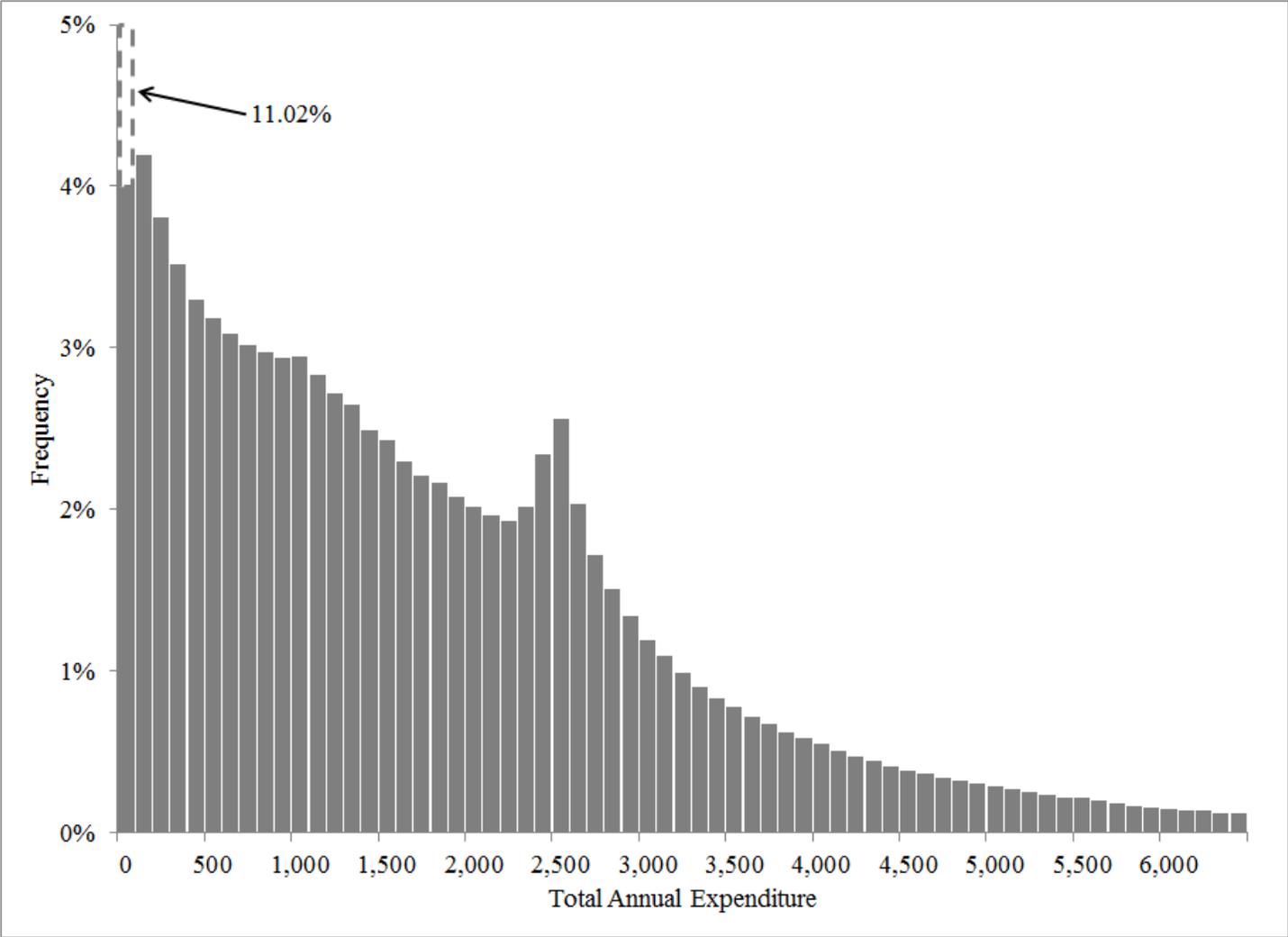
Middle bracket with 75% refund: exactly opposite of optimal design!

Rationale: political. Help the most people in this way (but do not maximize expected welfare).

ObamaCare eliminates the “donut hole”

Einav, Finkelstein, Schrimpf (2013) show that individuals bunch at kink where 75% subsidy stops \Rightarrow Moral hazard response

Figure 2: Annual spending distribution (in 2008)



The figure displays the distribution of total annual prescription drug spending in 2008 for our baseline sample. Each bar represents the set of people that spent up to \$100 above the value that is on the x-axis, so that the first bar represents individuals who spent less than \$100 during the year, the second bar represents \$100-200 spending, and so on. For visual clarity, we omit from the graph the 3% of the sample whose spending exceeds \$6,500. The kink location (in 2008) is at \$2,510. N =1,251,969.

Source: Einav, Finkelstein, Schrimpf (2013)

Estimating Health Benefits

Another approach of evaluating benefits of a health insurance program: look directly at health outcomes instead of consumption-smoothing benefit

How to implement this?

Simply comparing those enrolled in Medicaid to those not enrolled will suffer from bias.

Factors such as income and health status will bias the results.

Series of studies by Currie and Gruber: use Medicaid expansions and diff-in-diff strategy to evaluate value of programs

16.2

EVIDENCE: Using State Medicaid Expansions to Estimate Program Effects

Eligibility for all Children, by State

Year	Missouri Eligibility	Michigan Eligibility
1982	12%	20%
2000	76%	34%

Eligibility for Children by age in Washington, D.C.

Year	Age 13	Age 0
1982	18%	48%
2000	59%	56%

Effect of Medicaid Expansions on Health

Currie and Gruber find that these reductions in the number of uninsured had positive effects on health outcomes in pregnancies

1) Utilization of health services increased: Early prenatal care visits rose by more than 50%

2) Health care outcomes improved: Infant mortality declined by 8.5% due to the expansions in Medicaid for pregnant women.

⇒ Highly cost-effective policy.

Costs Per Life Saved of Various Regulations

Regulation concerning ...	Year	Agency	Cost per life saved (\$ millions)
Childproof lighters	1993	CPSC	\$0.1
Food labeling	1993	FDA	0.4
Reflective devices for heavy trucks	1999	NHTSA	0.9
Medicaid pregnancy expansions	1996	Currie & Gruber	1.0
Children's sleepware flammability	1973	CPSC	2.2
Rear/up/should seatbelts in cars	1989	NHTSA	4.4
Asbestos	1972	OSHA	5.5
<i>Value of statistical life</i>			7.0
Benezene	1987	OSHA	22
Asbestos ban	1989	EPA	78
Cattle feed	1979	FDA	170
Solid waste disposal facilities	1991	EPA	100,000

Effect of Medicare on Health

Medicare becomes available when you turn 65 \Rightarrow Can do a **regression discontinuity design** to see what happens when you cross age 65 threshold. Two recent papers use this strategy:

1) Card-Dobkin-Maestas “The Impact of Nearly Universal Insurance Coverage on Health Care Utilization and Health: Evidence from Medicare” AER 2008

Examines impacts across groups; with an interest in evaluating impacts on inequality in utilization

2) Card-Dobkin-Maestas “Does Medicare Save Lives?” QJE 2009

Examines impacts on outcomes (mortality following hospital admission)

Basic idea is to draw graphs of outcomes based on age for various groups

The discontinuity at 65 captures **short-term** changes in health care utilization and mortality from shift from < 65 to > 65

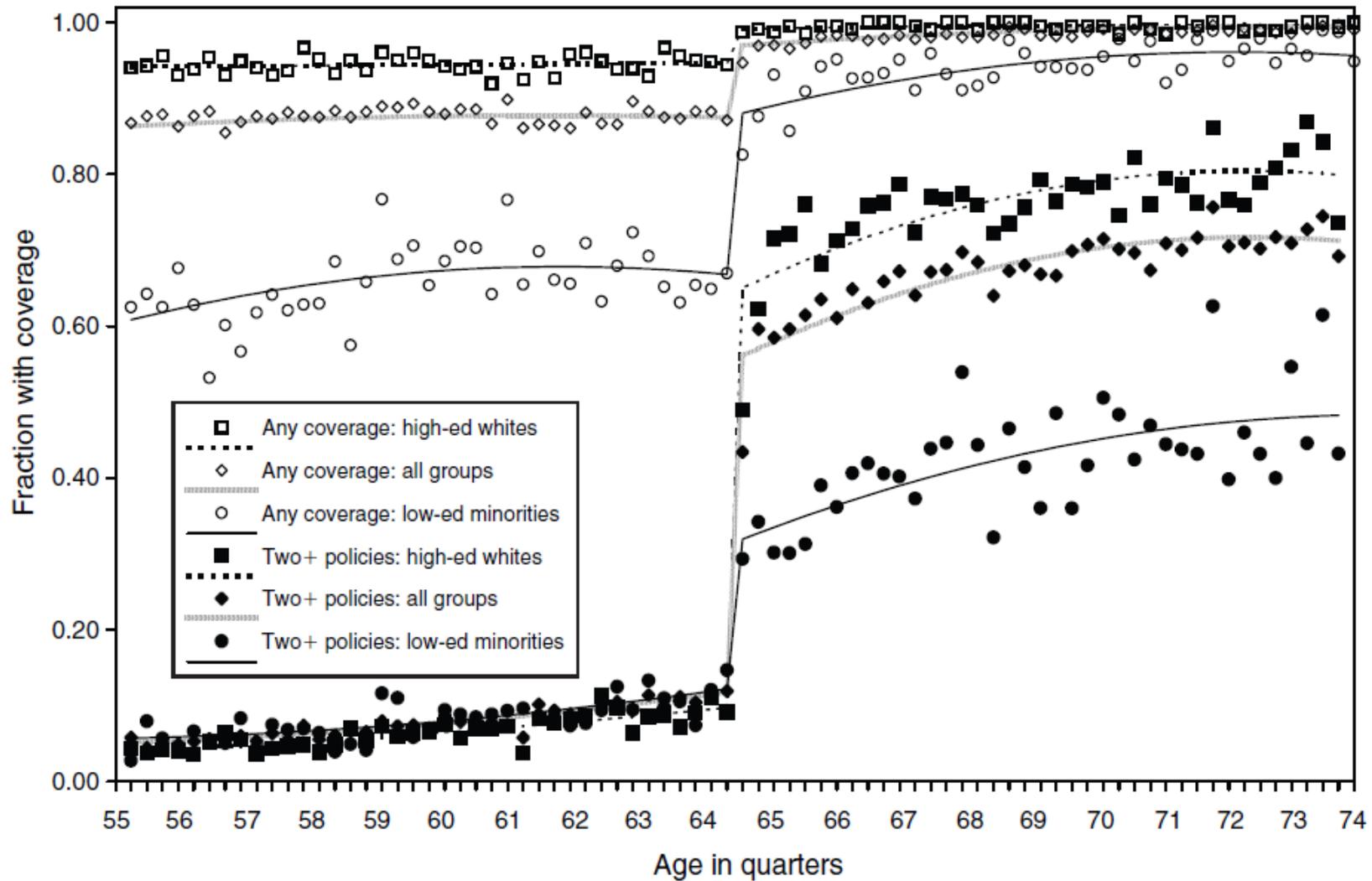


FIGURE 1. COVERAGE BY ANY INSURANCE AND BY TWO OR MORE POLICIES, BY AGE AND DEMOGRAPHIC GROUP

First stage: sharp increase in coverage; more for disadvantaged
 (From NHIS; age measured in quarters) FIGURE 1

Hospital discharge data (CA, FL, NY 1992-2002), ages 60-70

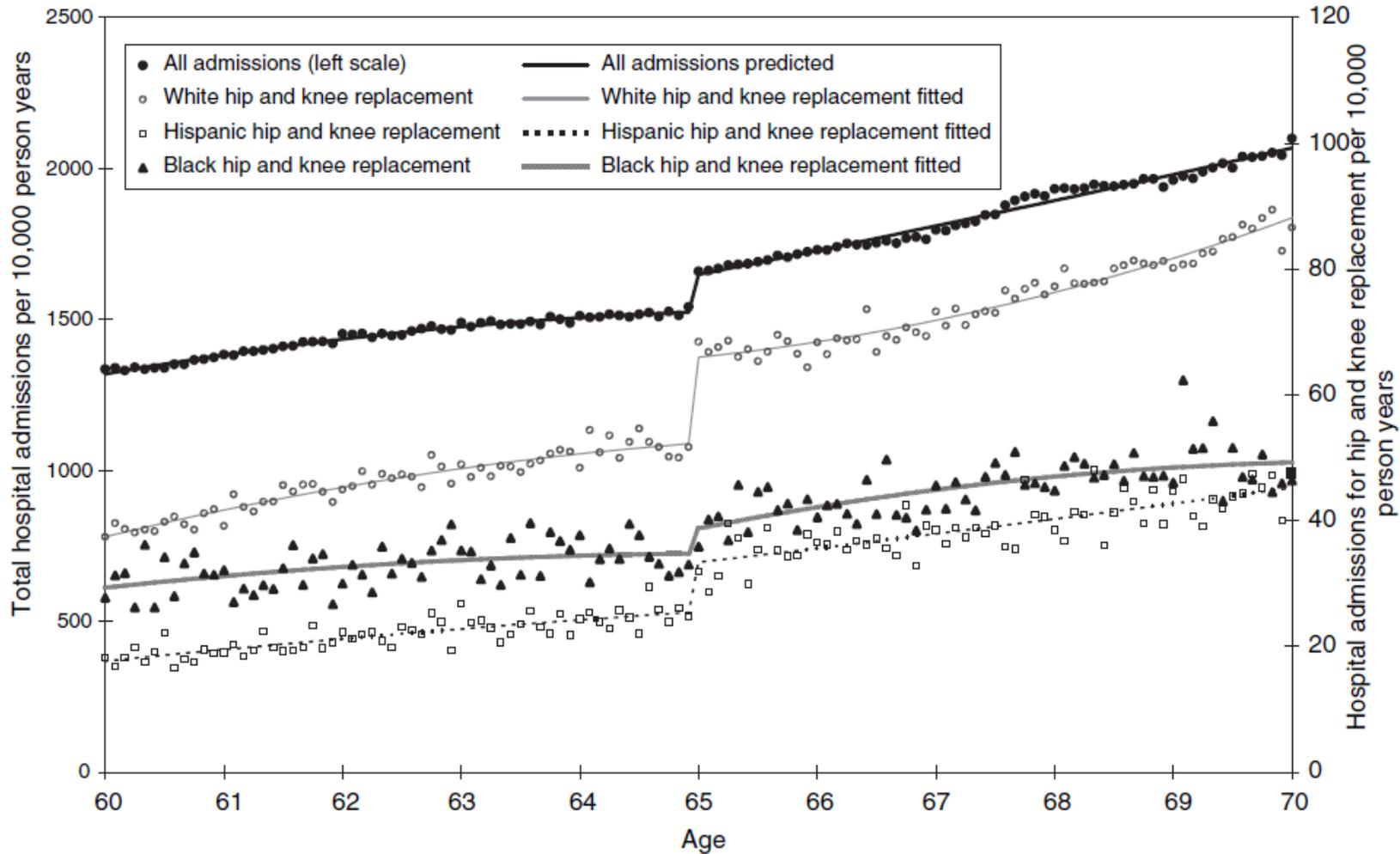


FIGURE 3. HOSPITAL ADMISSION RATES BY RACE/ETHNICITY

Increase is driven by discretionary medical care, diagnostic heart treatments.

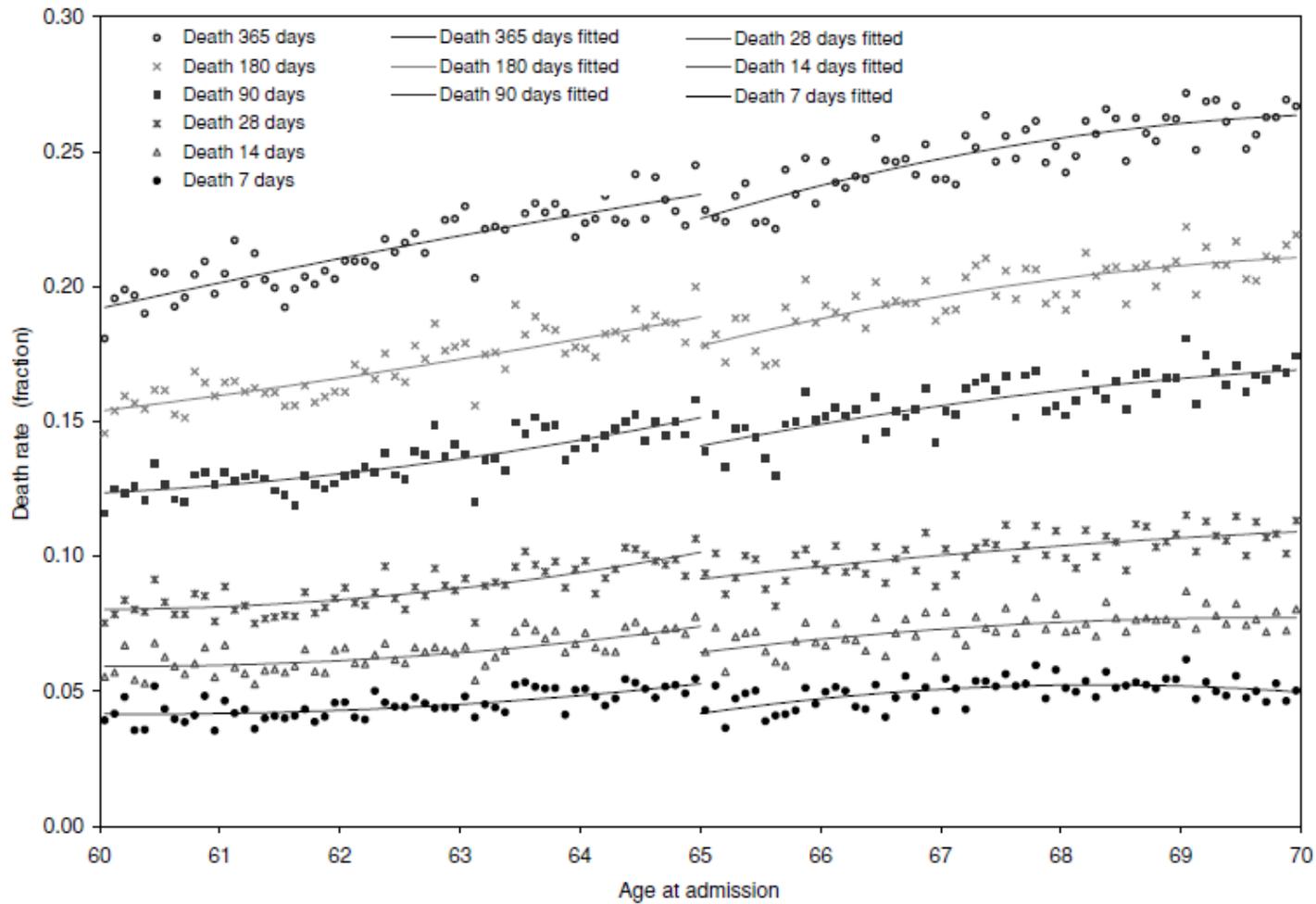


FIGURE VI
Patient Mortality Rates over Different Follow-Up Intervals

Nontrivial decrease in mortality.

Effects of Medicare on Health

1) Big increase in health insurance coverage, especially for disadvantaged groups

2) Big increase in health care utilization

3) Visible decrease in mortality after admission for conditions requiring Emergency Room immediate hospitalization (i.e. conditions for which likelihood of going to hospital is same before 65 and after 65)

⇒ Medicare health insurance saves lives, relative to private alternative

Effects of Insurance on Health Outcomes

How to reconcile two results: Medicaid and Medicare spending improves health but RAND experiment spending did not?

The studies examine different parts of the “medical effectiveness curve.”

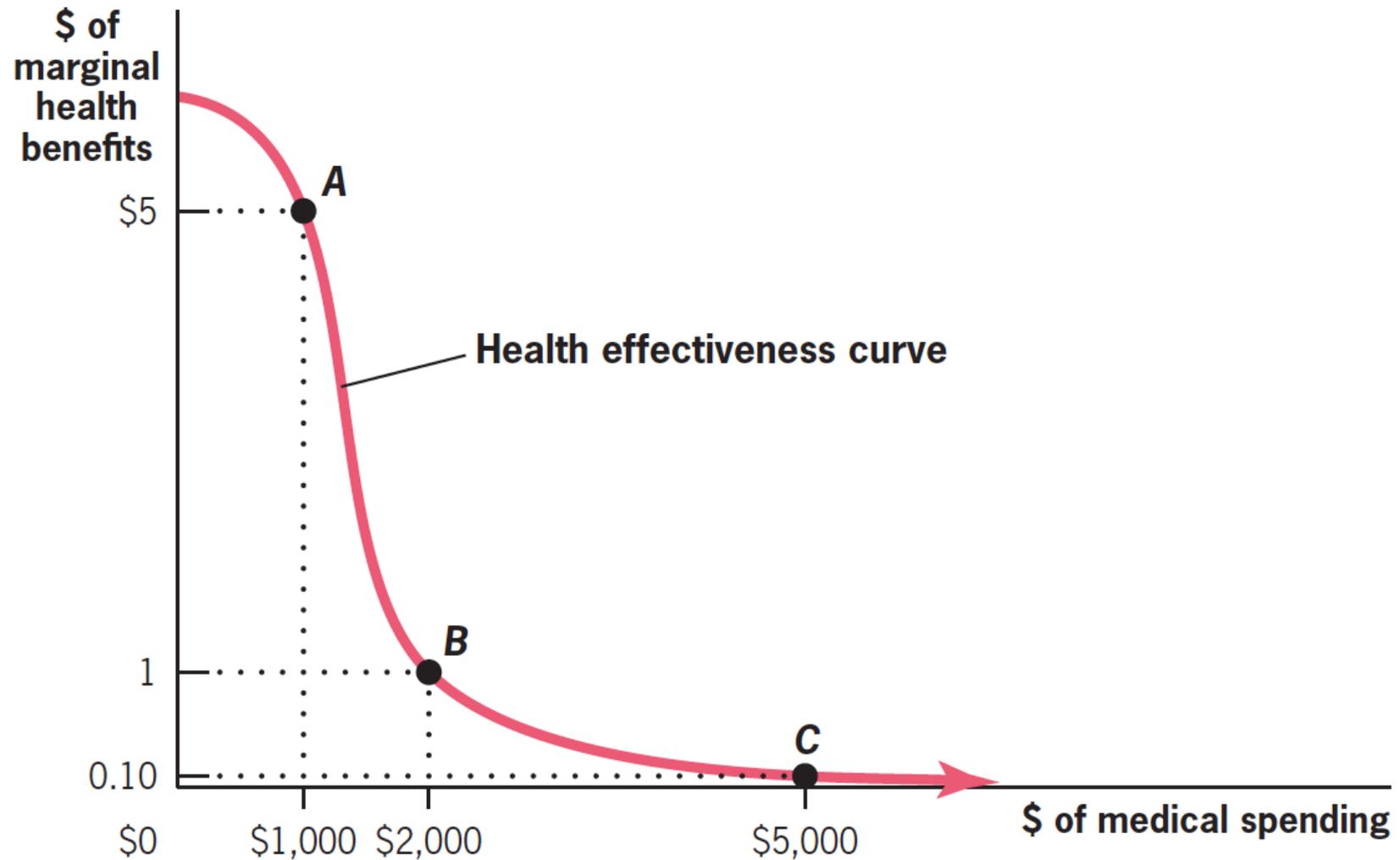
Moving individuals from uninsured to having some insurance has an important positive effect on health

Adding to the generosity of current insurance does not seem to cause significant changes on health

⇒ US health insurance system leaves many uninsured but provides overly generous care to the insured

15.2

The “Flat of the Curve”



Optimal Health Insurance: Provider Side

Preceding analysis of optimal insurance assumes patient makes entire healthcare decision

This assumed a passive doctor, in the sense that doctor provides whatever treatment patient requested

Clearly reality is closer to the opposite!

Incorporating supply side issues is critical in understanding health insurance

Question: choice of payment schemes for physician

Retrospective (fee-for-service) vs. prospective (diagnosis based fixed payments)

Optimal Health Insurance: Provider Side

Intuition: if patient doesn't choose level of care, healthcare may be inefficiently high

If physician is compensated for all costs \Rightarrow it is in his interest to do lots of procedures (e.g. too many C-section births)

Optimal Health Insurance: Provider Side Model

Payment for physician services is $P = \alpha + \beta \cdot c$

α =fixed cost payment for a given diagnosis

β =payment for proportional costs c (tests, nurses)

Various methods of payment (α, β) :

1. Fee-for-service ($\alpha = 0, \beta > 1$): No fixed payment for practice, but insurance company pays full cost of all visits to doctor + a surcharge.
2. Diagnosis based payment ($\alpha > 0, \beta = 0$): varying by type and # of patients but not services rendered

Optimal Health Insurance: Provider Side

General trend has been toward higher α , lower β

Private market has shifted from FFS to HMO (Health Maintenance Organizations) capitation schemes [where insurer pays a fixed amount per patient no matter what the health costs are]

Medicare/Medicaid shifted in 1980s to a prospective payment scheme.

Tradeoff: lower β provides incentives for doctors to provide less services. But they may provide too little!

⇒ Lower costs, but complaints of lower quality of care. Deep trade-off.

Evidence: Payment Schemes and Physician Behavior

1) In 1983, Medicare moved from retrospective reimbursement to prospective reimbursement.

2) **Prospective payment system (PPS)** is Medicare's system for reimbursing hospitals based on nationally standardized payments for specific diagnoses.

All diagnoses for hospital admissions were grouped into Diagnosis Related Groups (DRGs).

Government reimbursed a fixed amount per DRG. More severe DRGs received higher reimbursement.

Evidence: Payment Schemes and Physician Behavior

Cutler (1993) finds that PPS led to:

1. A reduction in treatment intensity. For example, the average length of hospital stay for elderly patients fell by 1.3 days.
2. No adverse impact on patient outcomes despite the reduction in treatment intensity.

Evidence that doctors put some weight on profits

Suggests they are practicing “flat of the curve” medicine: too much treatment before.

3. Cost growth slowed dramatically in the five years after PPS but then accelerated again.

Evidence: Payment Schemes and Physician Behavior

Why did costs accelerate? PPS not a perfect capitation scheme:

- 1) DRG creep: although the price per diagnosis was fixed, hospitals reacted by changing the DRG categorization (“up-coding”)
- 2) The design of the DRGs used actual treatments (e.g., a person with heart trouble might be assigned the DRG “pace-maker implantation” or “coronary bypass”).
- 3) This effectively creates a retrospective reimbursement system

Technology Growth and Health Care Cost Growth

- 1) Health care technology contributes to rising survival rates
- 2) Many new technologies have modest health benefits and are very costly and yet are adopted because Medicare/Private insurance accept any health effective treatment
⇒ fuels the development of new technologies (e.g. scans) which lead to growing costs and over-treatment, even if not worth social cost
- 3) Countries with lowest costs adopt only socially cost effective treatments (reduces costs while typically having small effects on health outcomes)
- 4) US health care system spends a lot on the insured (where marginal value of care is small) and spends little on the uninsured (where marginal value of care is high)

OBAMACARE (AFFORDABLE CARE ACT)

System begun in 2014 (follows the Romney Care model of Massachusetts):

- 1) Free insurance for low-income families [Medicaid expansion in most states, up to 138% of FPL (federal poverty line), 90% for by federal government]
- 2) Subsidized insurance for middle-income families [subsidized health insurance in ObamaCare exchanges for families 100%-400% of FPL]
- 3) ObamaCare exchanges in which insurance companies cannot condition premiums on pre-existing conditions or current health

OBAMACARE (AFFORDABLE CARE ACT)

4) Firm mandate: employers with 50+ employees must offer insurance to 30+ hour employees [else they pay a tax]

5) Individual mandate: individuals must have insurance [else they pay a tax]

6) Surtax on high incomes, tax on expensive “Cadillac” employer-based plans

⇒ Avoids universal health insurance (single-payer) but expands each of three major types pre-existing insurance: employer-provided, Medicaid, and non-group (exchanges), funded by progressive income taxation

LEGAL CHALLENGES TO OBAMACARE

1) Is the individual mandate constitutional? [July 2012]

Ruling: yes, but Feds cannot force States to expand Medicaid
⇒ Many states (including TX, FL) decided to opt-out of the Medicaid expansion [even though Fed was paying 90%]

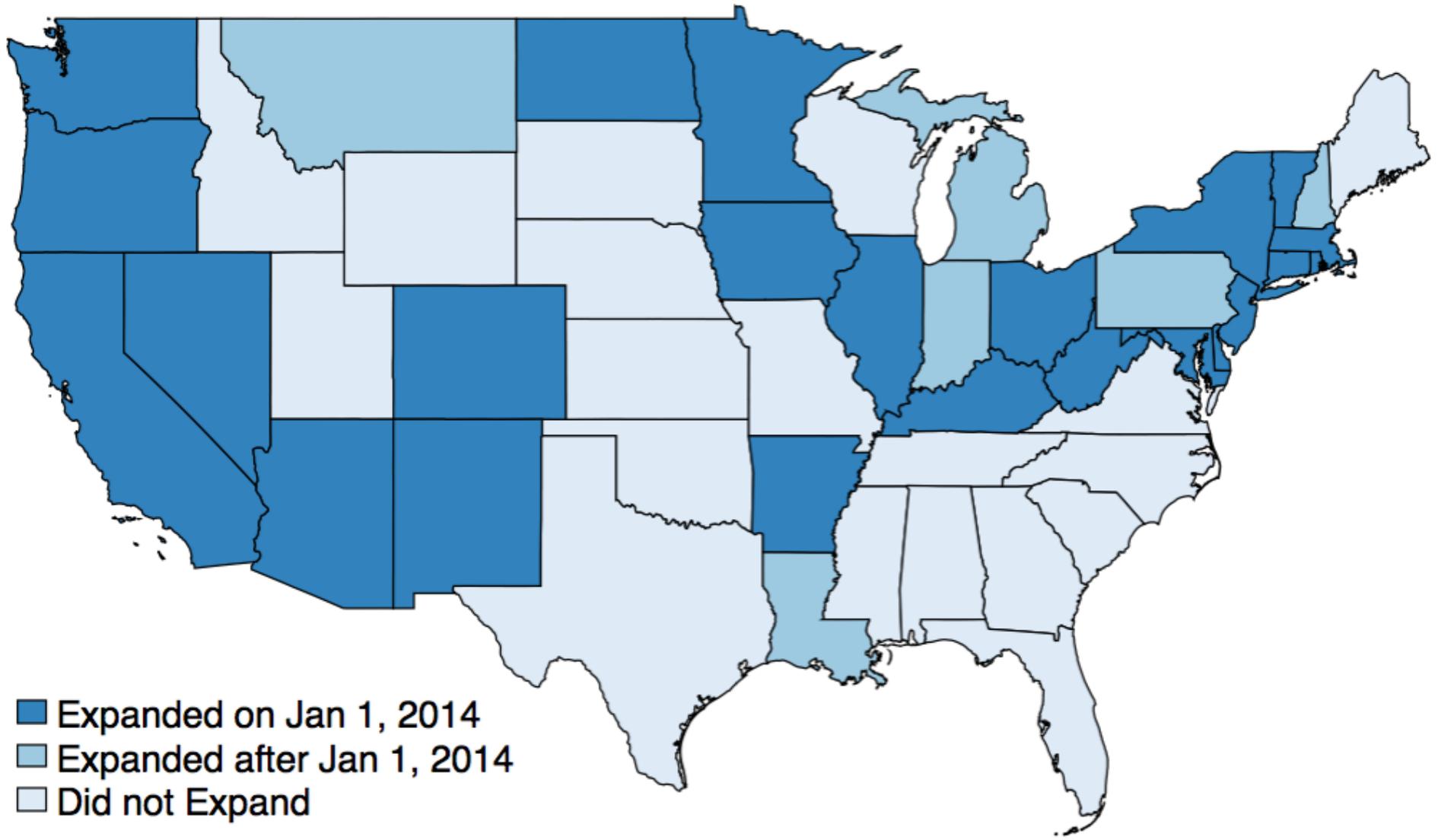
Consequence: There is a coverage gap because people below poverty line cannot access subsidized exchanges [States seem to be moving to accept Medicaid expansion]

2) Can the Feds set up exchanges if states don't do it themselves? [July 2015]

Ruling: yes, so subsidized exchanges did not disappear in many states ⇒ People were not priced out and there was no adverse selection death spiral in which only the high-risk buy

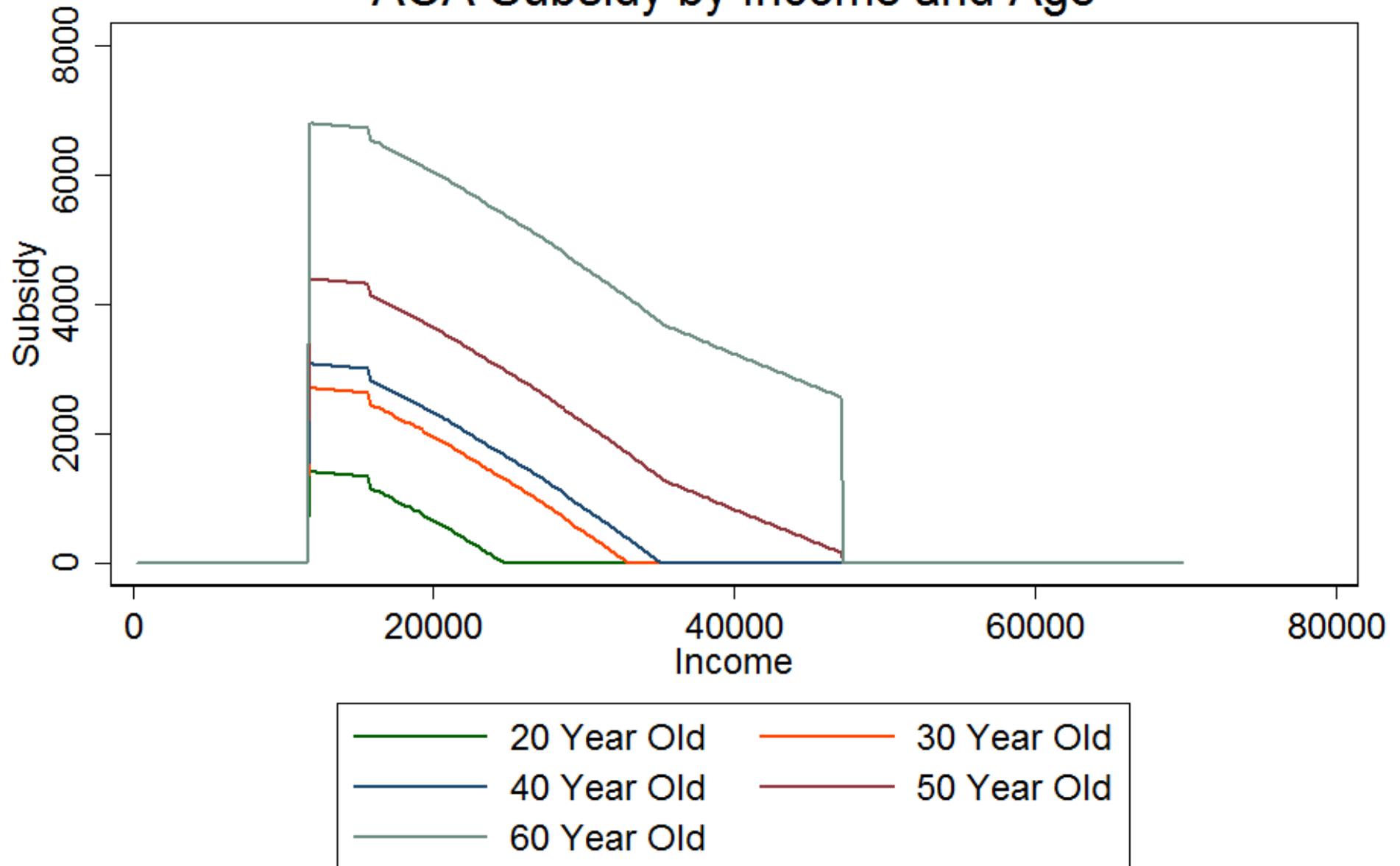
Graphs from Duggan-Goda-Jackson (2017)

State Medicaid Expansion Decisions



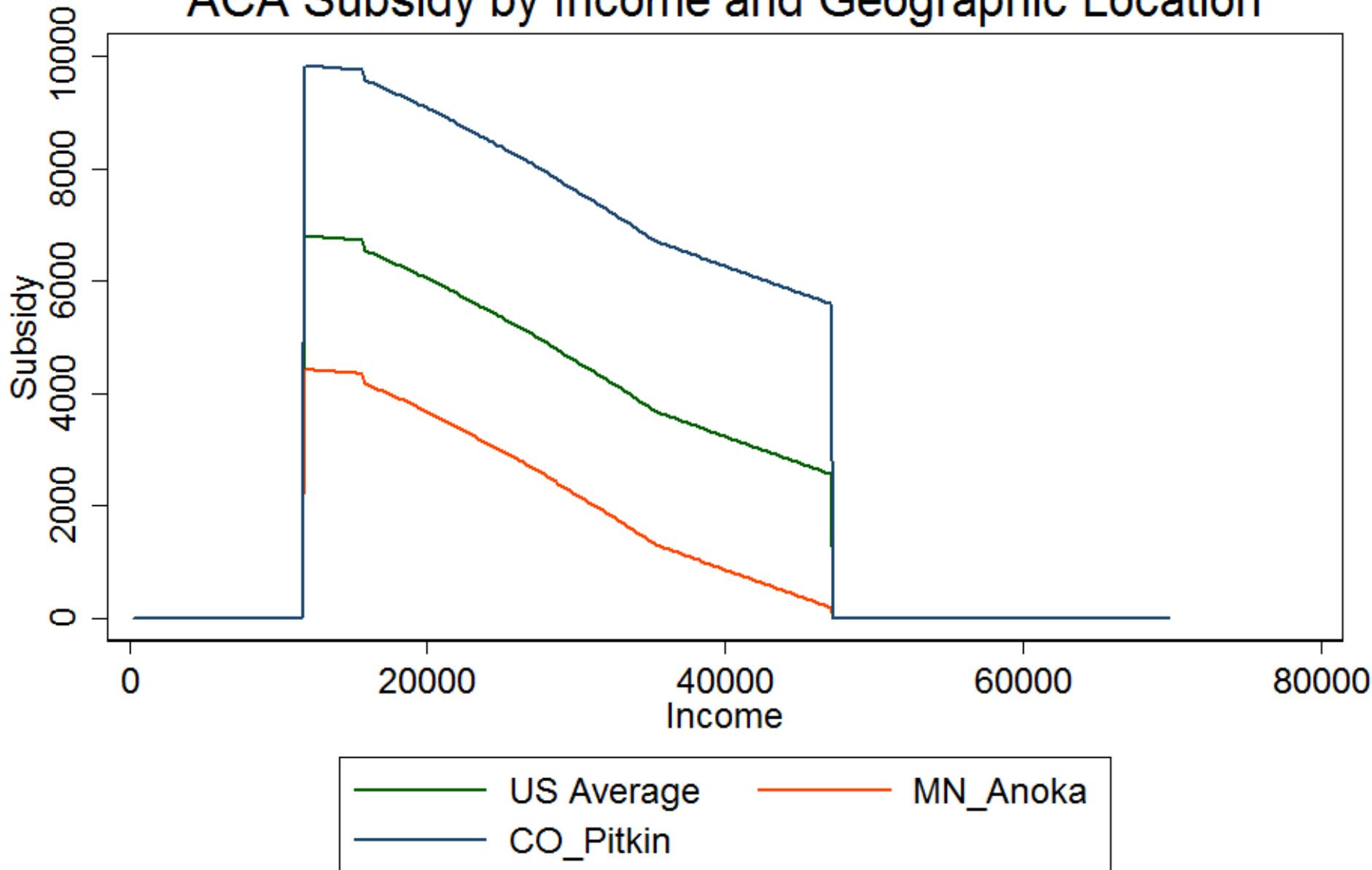
Source: Duggan-Goda-Jackson (2017)

ACA Subsidy by Income and Age



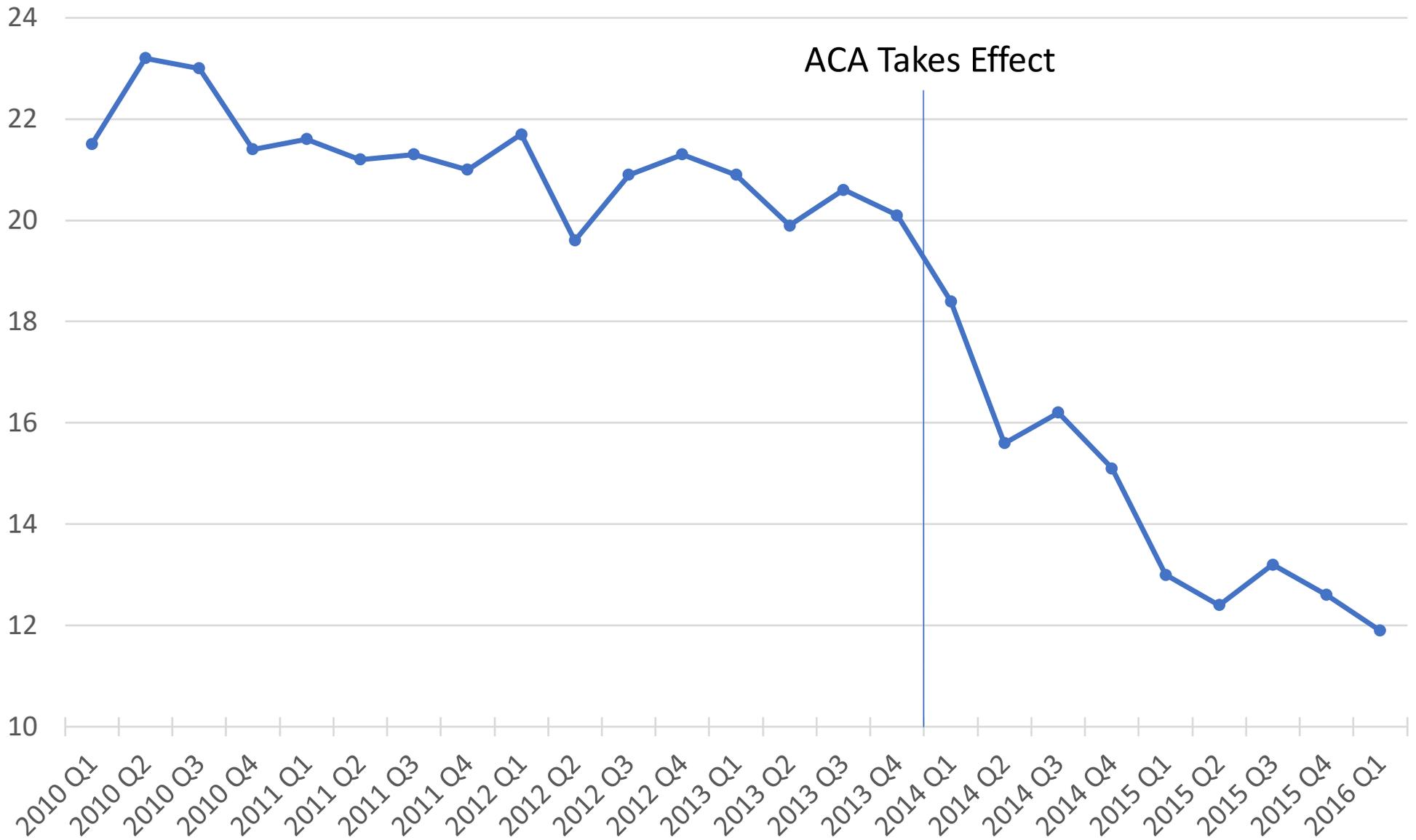
Note: Subsidy calculated for single person in 2015 for all scenarios.

ACA Subsidy by Income and Geographic Location



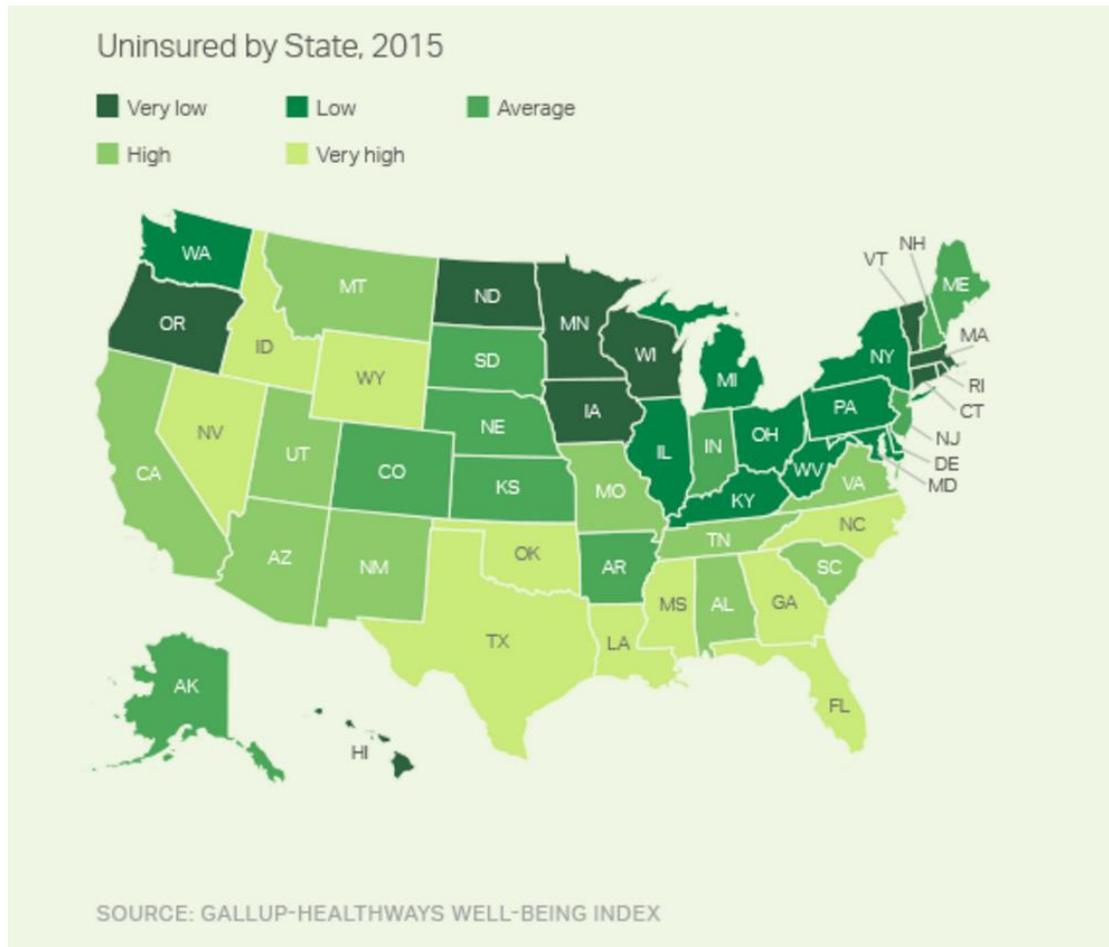
Note: Subsidy calculated for single 60-year-old in 2015 for all scenarios.

Percentage of Persons 18-64 Uninsured, January 2010 - March 2016



Source: NCHS. National Health Interview Survey Early Release of Quarterly Estimates, retrieved September 12, 2016.

Coverage Gains Vary by State



State	% Uninsured		Expanded Medicaid
	2013	2015	
California	21.6	11.8	Yes
Colorado	17.0	10.3	Yes
Florida	22.1	15.7	No
Illinois	15.5	8.7	Yes
Kentucky	20.4	7.5	Yes
Massachusetts	4.9	3.5	Yes
New York	12.6	8.6	Yes
Oregon	19.4	7.3	Yes
Texas	27.0	22.3	No
Virginia	13.3	12.6	No

16.6

The Massachusetts Experiment with Incremental Universalism

- Striking results:
 - MA uninsurance rate 3%, compared to 18% nationally.
 - Half of the increase in coverage from Medicaid or government subsidized plans.
 - Premiums in the non-group market have fallen by half relative national trends.
 - Costs of the reform roughly consistent with projections.

THE CONTINUED UNINSURED

Fraction of individuals uninsured should fall by 50% with ObamaCare [from 1/6 of population in 2013 down to 1/12 eventually]. Remaining uninsured:

- 1) Undocumented immigrants (no access to Medicaid, ObamaCare) \simeq 10m
- 2) Low-income people for whom insurance is still not worth it relative to emergency care alternative [especially poor in states that didn't expand Medicaid]
- 3) Low-risk people for whom insurance is still not worth it relative to emergency care alternative with tax

IS OBAMACARE KILLING JOBS?

Less need to work if you get health insurance for free

Earlier evidence: 2005 Tennessee Medicaid disenrollment increased employment [Garthwaite Gross Notowidigdo 2014]

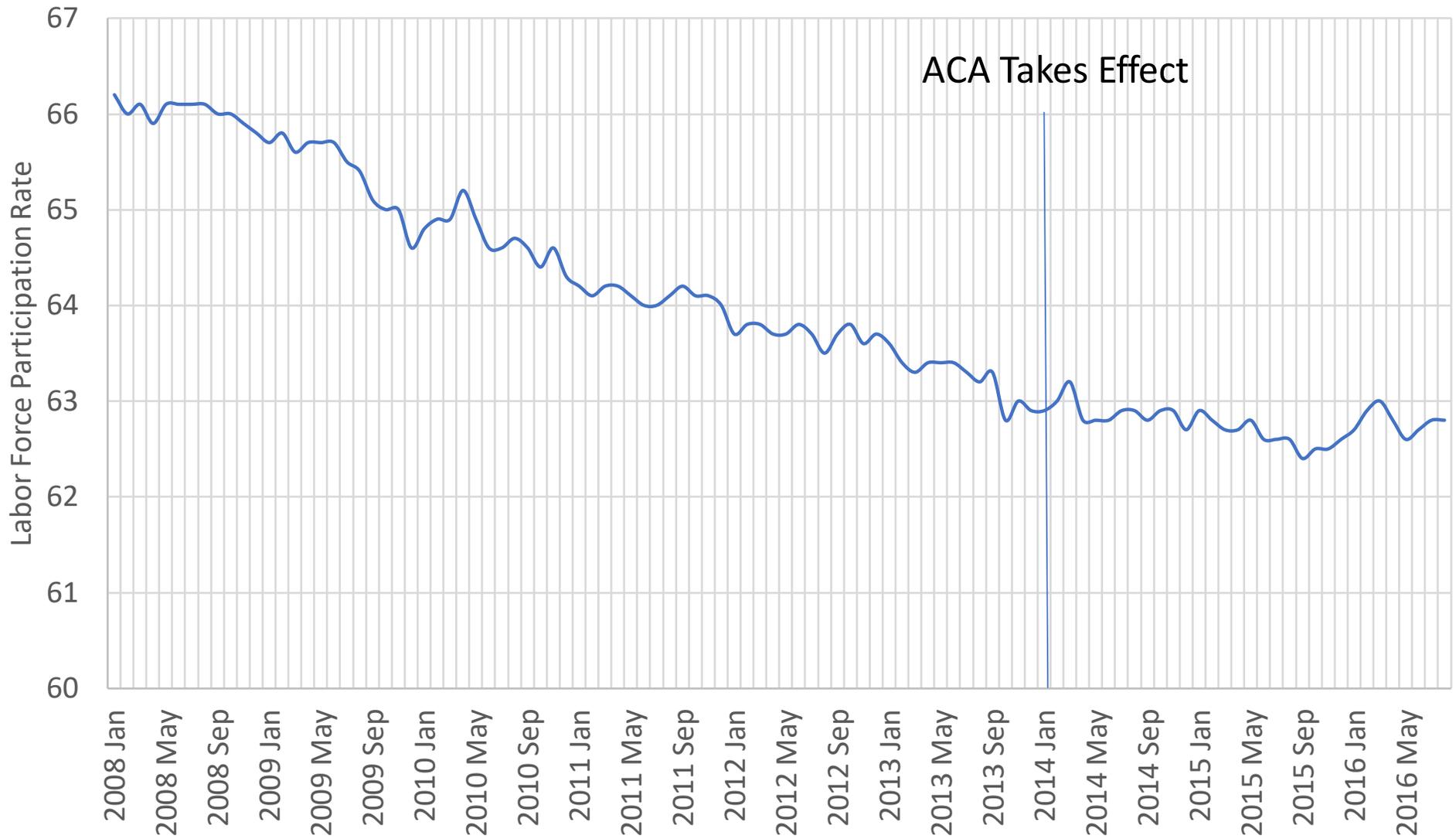
Huge public debate on whether ObamaCare is “job killer” by:

a) Reducing value to work

b) Forcing employers to offer insurance that workers don't value \$1-for-\$1 [effective tax, especially for min wage workers]

c) Firms can offer less insurance if they downsize or hire part-time

Labor Force Participation Rate Seasonally Adjusted Ages 16 and over



ARE PREMIUMS SKYROCKETING?

2016 Pres. campaign issue: Rising ObamaCare premiums

Employer-provided insurance (most people): Standard increase

Exchanges: Big increase from 2016, but 2016 was surprisingly cheap. Now at prices CBO had originally projected.

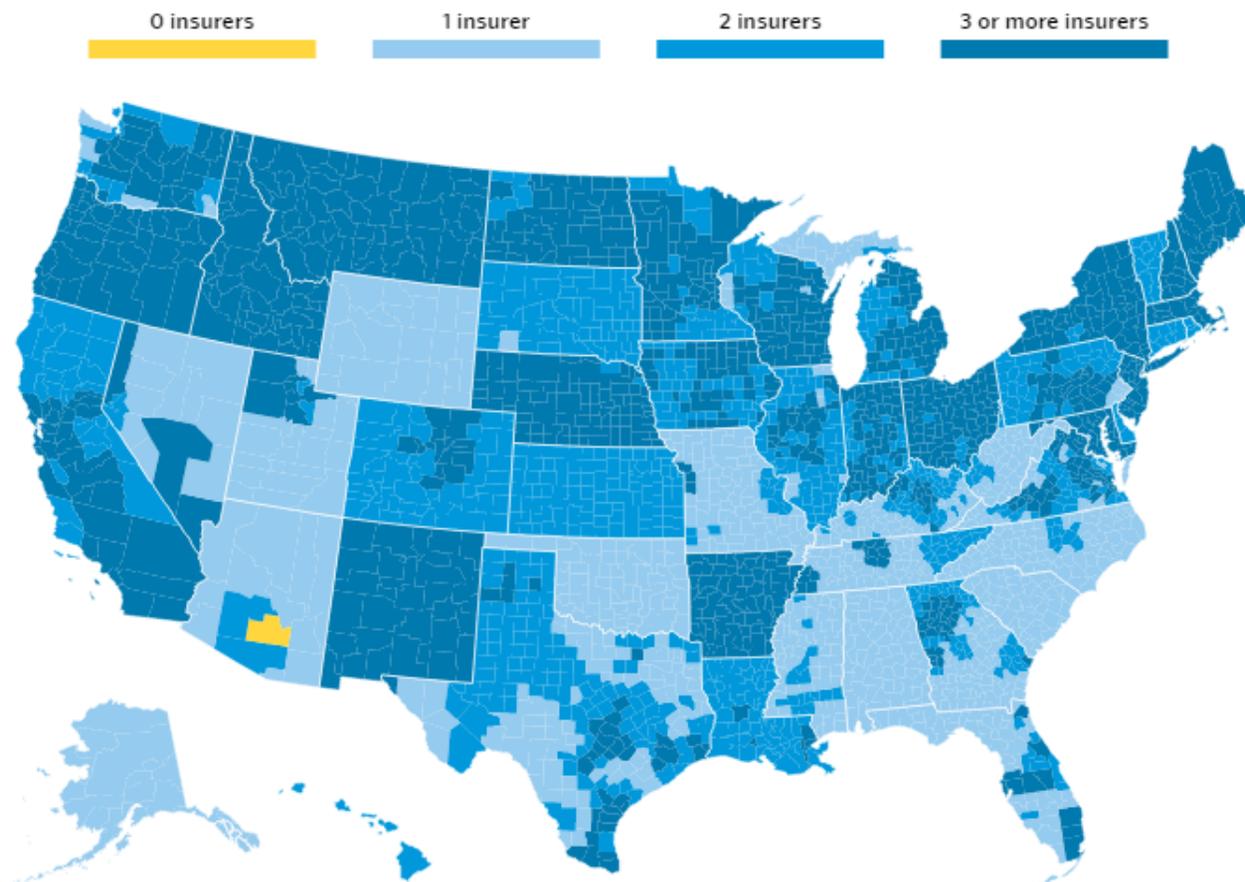
Possible explanation: Insurers deliberately under-priced premiums [lost money] in first year in order to gain market share [hoping to keep customers forever], then raised prices

Future: Probably stable. But possible death spiral if healthy start dropping out.

Either way: Need to avoid monopoly pricing in rural areas [Cabral-Geruso-Mahoney 2015]

Possible Trouble on the Horizon for the Exchanges in 2017

Estimated number of insurers participating in Affordable Care Act exchanges by county, 2017



Source: Duggan-Goda-Jackson (2017)

THE FUTURE (1/2)

Republicans promise to “repeal and replace” ObamaCare. But ObamaCare is structured the way it is for a reason:

Suppose you want to ban insurance companies from using pre-existing conditions in insurance pricing

→ death spiral, unless you have an individual mandate

→ millions can't afford insurance, unless you have subsidies

That's ObamaCare! (no use of pre-existing conditions + mandate + subsidies)

THE FUTURE (2/2)

So really the options are ObamaCare, a return to the pre-ObamaCare world, or single-payer. Questions to guide the choice:

- 1) Preferences: Do we want more or less redistribution from rich to low- and middle-class? Do we believe in redistribution from healthy/low-risk to sick/high-risk?
- 2) Centralization: Patchwork of policies with private insurers [ObamaCare] vs. Universal coverage [Medicare, rest of OECD]
- 3) Cost control under universal coverage: Private cost-benefit [Hippocratic Oath] vs. Social cost-benefit, public management innovations [Government Performance Lab]

REFERENCES

Jonathan Gruber, Public Finance and Public Policy, Fourth Edition, 2012
Worth Publishers, Chapter 15 and Chapter 16

Card, David, Carlos Dobkin, and Nicole Maestas. "The impact of nearly universal insurance coverage on health care utilization and health: evidence from Medicare." *American Economic Review* 98.5 (2008): 2242-2258.(web)

Cabral, Marika, Michael Geruso, Neale Mahoney. "Does Privatized Health Insurance Benefit Patients or Producers? Evidence from Medicare Advantage." 2015.

Card, David, Carlos Dobkin, and Nicole Maestas. "Does Medicare save lives?." *The Quarterly Journal of Economics* 124.2 (2009): 597-636.(web)

Currie, Janet, and Jonathan Gruber. "The technology of birth: Health insurance, medical interventions, and infant health." No. w5985. National Bureau of Economic Research, 1997.(web)

Cutler, David M. "The incidence of adverse medical outcomes under prospective payments." No. w4300. National Bureau of Economic Research, 1993.(web)

Duggan, Mark, Gopi Shah Goda, Emilie Jackson. "The Effect of the Affordable Care Act on Coverage and Labor Market Outcomes." Stanford working paper.

Einav, Liran, Amy Finkelstein, Paul Schrimpf "The Response of Drug Expenditures to non-linear Contract Design: Evidence from Medicare Part D," NBER Working Paper 19393, 2013. (web)

Finkelstein, Amy, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Heidi Allen, and Katherine Baicker. "The Oregon Health Insurance Experiment: Evidence from the First Year." *The Quarterly Journal of Economics* 127, no. 3 (2012): 1057-1106.(web)

Garthwaite, Craig, Tal Gross, and Matthew Notowidigdo. "Public Health Insurance, Labor Supply, and Employment Lock." *Quarterly Journal of Economics*, 2014.