Graduate Public Economics

Optimal Labor Income Taxes/Transfers

Emmanuel Saez

TAXATION AND REDISTRIBUTION

Key question: Do/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

Modern governments raise large fraction of GDP in taxes (30-50%) and spend significant fraction of GDP on transfers

FACTS ON US TAXES AND TRANSFERS

References: Comprehensive description in Gruber undergrad textbook (taxes/transfers) and Slemrod-Bakija (taxes)

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, interest from state+local bonds, unrealized capital gains

FEDERAL US INCOME TAX

Taxable income = AGI - personal exemptions - deduction

personal exemption = $ 3650 * # family members (in 2010)

deduction is max of standard deduction or itemized deductions

Standard deduction is a fixed amount depending on family structure ($11.4K for couple, $5.7K for single in 2010)

Itemized deductions: mortgage interest payments, charitable giving, state and local income taxes paid, medical expenses (above 7.5% of income)

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

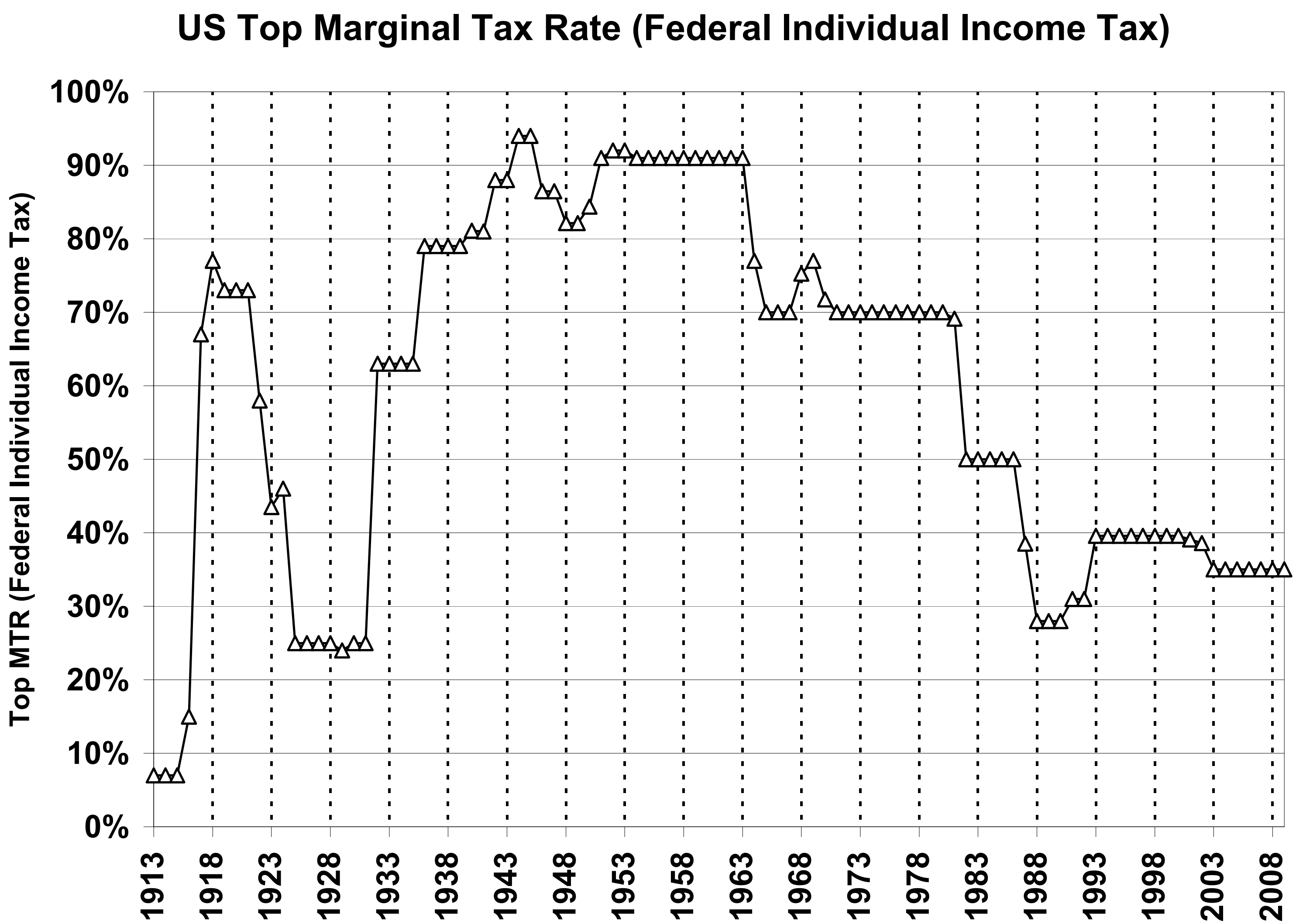
FEDERAL US INCOME TAX: TAX BRACKETS

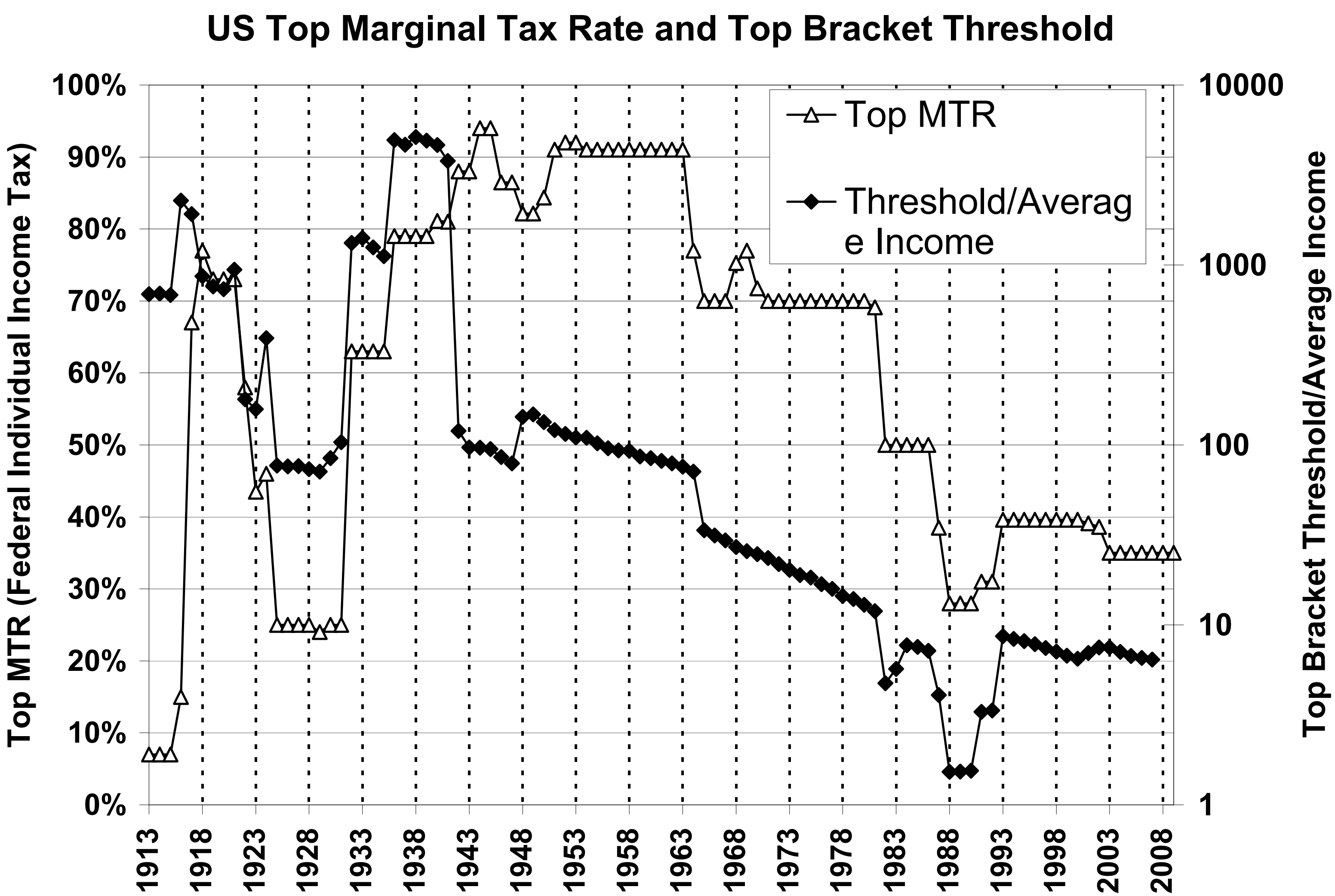
Tax $T(z)$ is piecewise linear and continuous function of taxable income z with constant marginal tax rates (MTR) $T'(z)$ by brackets

In 2009, 6 brackets with MTR 10%,15%,25%,28%,33%,35% (top bracket for z above $373K), indexed on price inflation

Lower preferential rates (up to a max of 15%) 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)
Top MTR (Federal Individual Income Tax)
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
1913
1918
1923
1928
1933
1938
1943
1948
1953
1958
1963
1968
1973
1978
1983
1988
1993
1998
2003
2008


US Top Marginal Tax Rate and Top Bracket Threshold
Top MTR (Federal Individual Income Tax)
Top Bracket Threshold/Average Income
Top MTR
Threshold/Average Income
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
10000
1000
100
10
1
1913
1918
1923
1928
1933
1938
1943
1948
1953
1958
1963
1968
1973
1978
1983
1988
1993
1998
2003
2008

FEDERAL US INCOME TAX: AMT AND CREDITS

Alternative minimum tax (AMT) is a parallel tax system (quasi flat tax at 28%) with fewer deductions: actual tax =max($T(z), AMT$) (hits 2-3% of tax filers in upper middle class)

Tax credits: Additional reduction in taxes

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

(2) **Refundable** (can reduce taxes below zero, i.e., be net transfers): EITC (earned income tax credit, up to $5000, working families with kids), Child Tax Credit ($1000 per kid, partly refundable)

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 tax filers get a tax refund as withholdings $>$ net taxes owed

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

Information + 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

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

BOTTOM LINE ON ACTUAL TAXES/TRANSFERS

1) Based on current income, family situation, and disability (retirement) status $\Rightarrow$ Strong link with **current ability to pay**

2) Some allowances made to reward / encourage certain behaviors: charitable giving, home ownership, savings, energy conservation, and more recently work (refundable tax credits such as EITC)

3) Provisions pile up overtime making tax/transfer system more and more complex until significant simplifying reform happens (such as US Tax Reform Act of 1986)

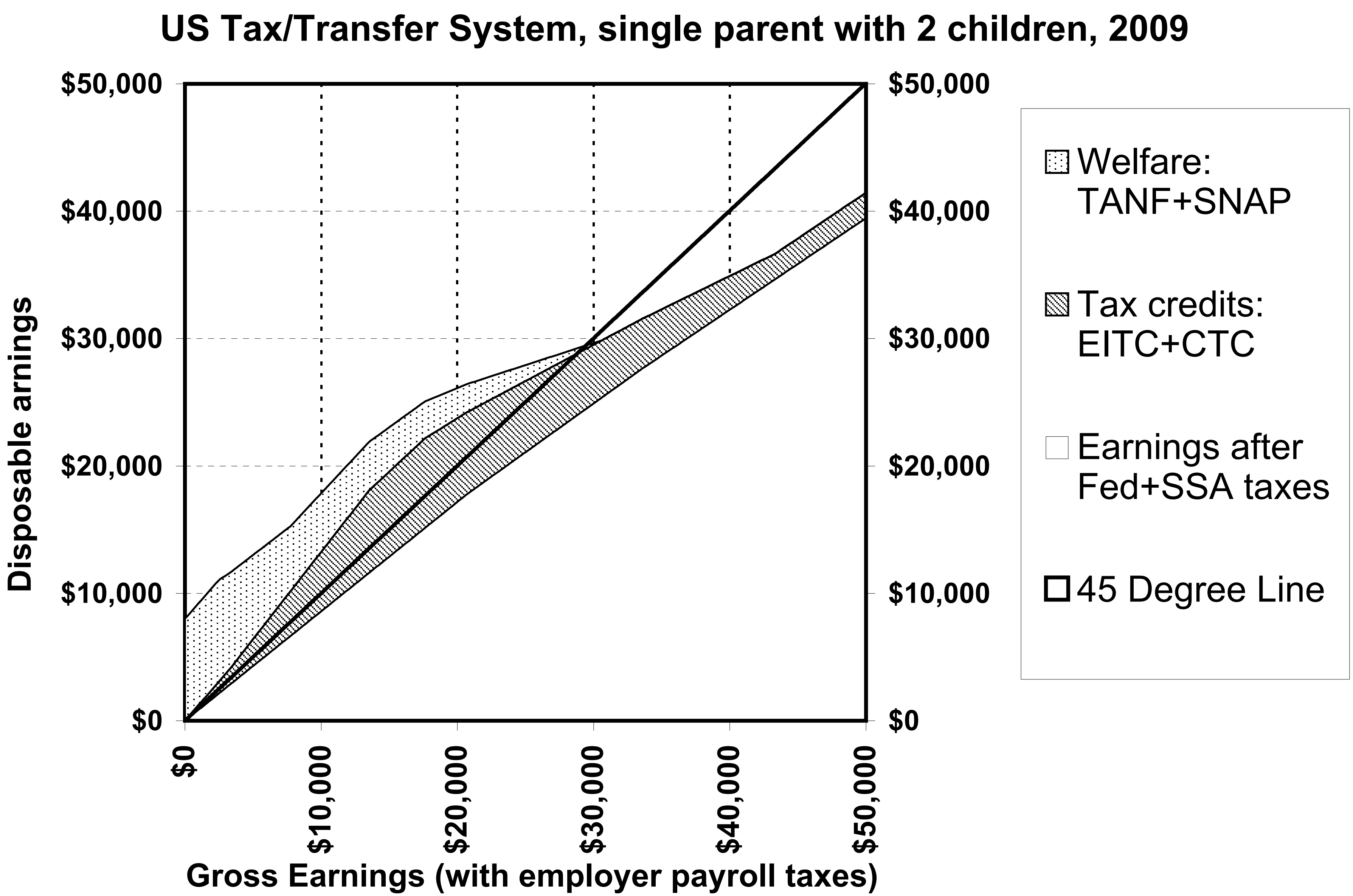
KEY CONCEPTS FOR TAXES/TRANSFERS

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: $z-T(z) = -T(0)+z-[T(z)-T(0)] = -T(0)+z\cdot(1-\tau_p)$ (extensive labor supply response)

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



US Tax/Transfer System, single parent with 2 children, 2009
Disposable arnings
$50,000
$40,000
$30,000
$20,000
$10,000
$0
$0
$10,000
$20,000
$30,000
$40,000
$50,000
Gross Earnings (with employer payroll taxes)
Welfare: TANF+SNAP
Tax credits: EITC+CTC
Earnings after Fed+SSA taxes
45 Degree Line

OPTIMAL TAXATION: SIMPLE MODEL WITH NO BEHAVIORAL RESPONSES

Utility $u(c)$ strictly increasing and concave

Same for everybody where c is after tax income.

Income is z and is fixed for each individual, $c = z - T(z)$ where $T(z)$ is tax on z. z has density distribution $h(z)$

Government maximizes **Utilitarian** objective: $\int_0^{\infty} u(z - T(z))h(z)dz$

subject to **budget constraint** $\int T(z)h(z)dz \geq E$ (multiplier λ)

SIMPLE MODEL WITH NO BEHAVIORAL RESPONSES

Form lagrangian: $L = [u(z - T(z)) + \lambda T(z)]h(z)$

FOC $T(z)$: $0 = \partial L/\partial T(z) = [-u'(z - T(z)) + \lambda]h(z) \Rightarrow u'(z - T(z)) = \lambda \Rightarrow z - T(z) =$ constant for all z.

$\Rightarrow c = \bar{z} - E$ where $\bar{z} = \int zh(z)dz$ average income.

100% marginal tax rate. Perfect equalization of after-tax income.

Utilitarianism with decreasing marginal utility leads to perfect egalitarianism.

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

$\Rightarrow$ Optimal income tax theory incorporates behavioral responses (Mirrlees REStud '71)

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

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

2ND WELFARE THEOREM FALLACY

Suppose individuals differ in their ability to earn

2nd Welfare Theorem: Any Pareto Efficient outcome can be reached by (1) Suitable redistribution of initial endowments [individualized **lump-sum** taxes based on ability and not behavior], (2) Then letting markets work freely

$\Rightarrow$ No conflict between efficiency and equity

In reality, redistribution of initial endowments is not feasible (information pb) and govt needs to use **distortionary** taxes and transfers based on income and consumption to redistribute

$\Rightarrow$ Real conflict between efficiency and equity

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

MIRRLEES OPTIMAL INCOME TAX MODEL

1) Standard labor supply model: Individual maximizes $u(c,l)$ subject to $c = wl - T(wl)$ where c consumption, l labor supply, w wage rate, $T(.)$ nonlinear income tax $\Rightarrow$ taxes affect labor supply

2) Individuals differ in ability w: w distributed with density $f(w)$.

3) Govt social welfare maximization: Govt maximizes $SWF = \int G(u(c,l))f(w)dw$ ($G(.)\uparrow$ concave) subject to

(a) budget constraint $\int T(wl)f(w)dw \geq E$ (multiplier λ)

(b) individuals' FOC $w(1-T')u_c + u_l = 0$

MIRRLEES MODEL RESULTS

Optimal income tax trades-off redistribution and efficiency (as tax based on w only not feasible) $\Rightarrow T(.) < 0$ at bottom (transfer) and $T(.) > 0$ further up (tax) [full integration of taxes/transfers]

Mirrlees formulas complex, only a couple fairly general results:

1) $0 \leq T'(.) \leq 1$, $T'(.) \geq 0$ is non-trivial (rules out EITC) [Seade '76]

2) Marginal tax rate $T'(.)$ should be zero at the top (if skill distribution bounded) [Sadka-Seade]

3) If everybody works and lowest $wl > 0$, $T'(.) = 0$ at bottom

BEYOND MIRRLEES

Mirrlees '71 has had a profound impact on information economics: models with asymmetric information in contract theory

Discrete 2-type version of Mirrlees model developed by Stiglitz JpubE '82 with individual FOC replaced by Incentive Compatibility constraint [high type should not mimick low type]

Till late 1990s, Mirrlees results not closely connected to empirical tax studies and little impact on tax policy recommendations

Since late 1990s, Diamond AER'98, Piketty '97, Saez ReStud '01 have connected Mirrlees model to practical tax policy / empirical tax studies

INTENSIVE LABOR SUPPLY ELASTICITY CONCEPTS

$\max u(c, z)$ st $c = z(1-\tau) + R$, $u \uparrow c$ consumption), $u \downarrow z$ earnings (labor effort). R is virtual income and τ marginal tax rate.

FOC $(1-\tau)u_c + u_z = 0 \Rightarrow$ Marshallian labor supply $z = z(1-\tau, R)$

Uncompensated elasticity: $\varepsilon^u = [(1-\tau)/z]\partial z/\partial(1-\tau)$.

Income effects: $\eta = (1-\tau)\partial z/\partial R \leq 0$.

Substitution effects: Hicksian labor supply: $z^c(1-\tau, u)$, defines a compensated elasticity $\varepsilon^c > 0$ (subst. effects).

Slutsky equation: $\partial z^c/\partial(1-\tau) = \partial z/\partial(1-\tau) - z\partial z/\partial R \Rightarrow \varepsilon^c = \varepsilon^u - \eta$

LAFFER CURVE

With a constant tax rate τ, total reported income Z depends on $1-\tau$ (net-of-tax rate)

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

Top of the Laffer Curve corresponds to tax rate τ^* maximizing tax revenue: inefficient to have $\tau > \tau^*$

$0 = R'(\tau^*) = Z - \tau^* dZ/d(1-\tau) \Rightarrow$

$\tau^* = 1/(1+e)$ where $e = [(1-\tau)/Z]dZ/d(1-\tau)$ is the elasticity of reported income with respect to the net-of-tax rate

OPTIMAL TOP INCOME TAX RATE (SAEZ '01)

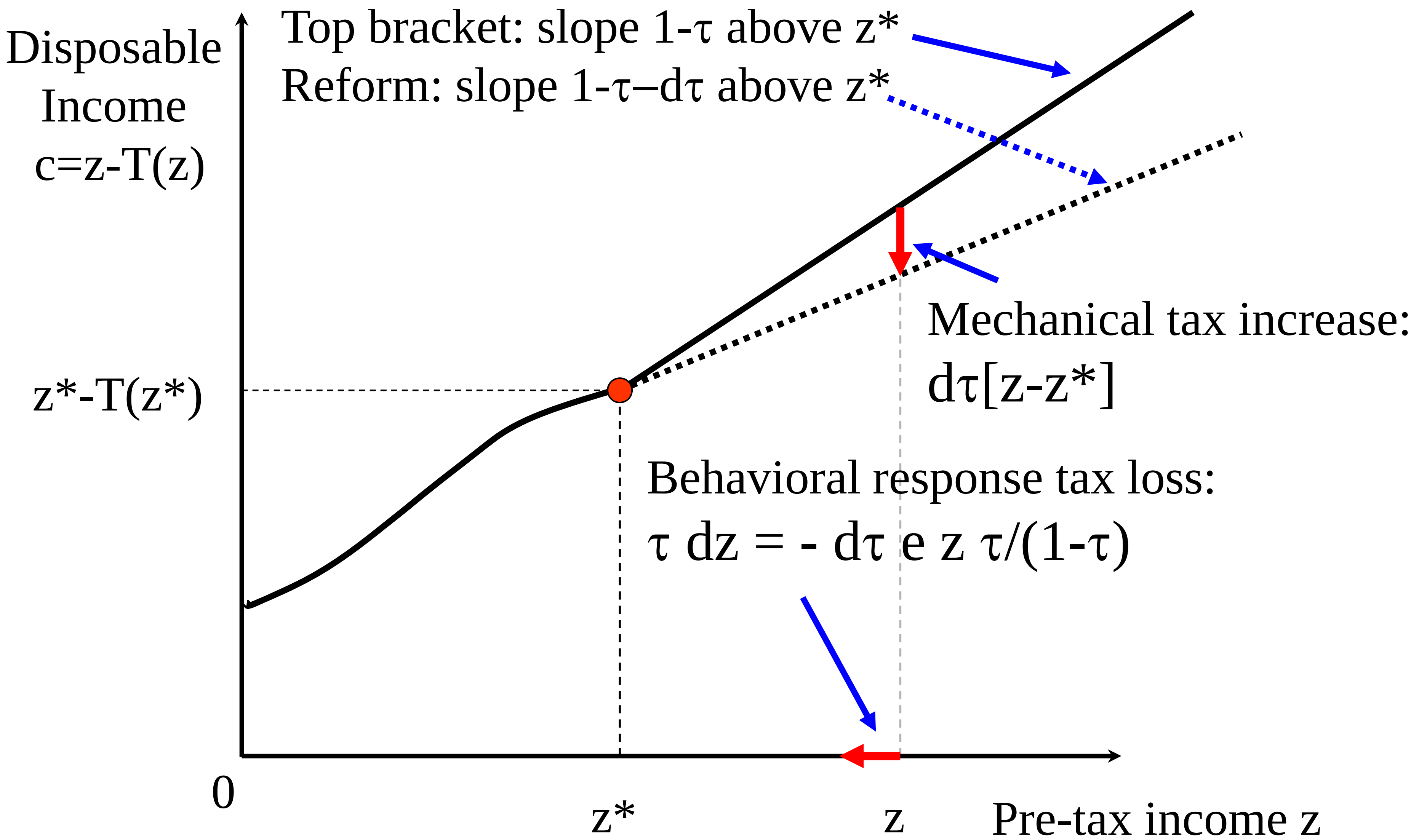
Consider constant MTR τ above fixed z^*. Goal is to derive optimal τ

Elasticity of taxable income literature (Saez, Slemrod, Giertz JEL '09) estimates ε

Assume that N individuals above z^*. Denote by $z^m(1-\tau)$ their average income [depends on net-of-tax rate $1-\tau$], with elasticity $e = [(1-\tau)/z_m] \cdot dz_m/d(1-\tau)$

Note that e is a mix of income and substitution effects, Saez '01 shows that $e = \varepsilon^c + \eta/a$ so that $\varepsilon^u < e < \varepsilon^c$.

[a is $z^m/(z^m - z^*) > 1$ as we will see]



Disposable Income
c=z-T(z)
Top bracket: slope 1-τ above z*
Reform: slope 1-τ–dτ above z*
z*-T(z*)
Mechanical tax increase:
dτ[z-z*]
Behavioral response tax loss:
τ dz = - dτ e z τ/(1-τ)
0
z*
z
Pre-tax income z

OPTIMAL TOP INCOME TAX RATE

Consider small $d\tau > 0$ reform above z^*.

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

$$dM = N \cdot [z^m - z^*]d\tau$$

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

$$dB = N\tau dz^m = -N\tau \frac{dz^m}{d(1-\tau)} d\tau = -N\frac{\tau}{1-\tau} \cdot \frac{1-\tau}{z^m} \frac{dz^m}{d(1-\tau)} \cdot z^m d\tau$$

$$\Rightarrow dB = -N\frac{\tau}{1-\tau} \cdot e \cdot z^m d\tau$$

3) **Welfare effect:**

Money-metric utility loss is dM by envelope theorem: govt values marginal consumption of rich at $0 \leq \bar{g} < 1$: $dW = -\bar{g}dM$ [formally $\bar{g} = \int_{z^*}^{\infty} G'(u) \cdot u_c \cdot h(z)dz/((1-H(z))\lambda)$]

NOTE ON WELFARE EFFECT OF TAX REFORM

Indirect utility: $V(1-\tau, R) = \max_z u(z(1-\tau) + R, z)$ where R is virtual income intercept

Reform: $d\tau$ and $dR = z^* d\tau$:

$dV = u_c \cdot [-z d\tau + dR] = -u_c \cdot [z - z^*] d\tau$

$[z - z^*] d\tau$ is the mechanical increase in taxes

Envelope theorem: no effect of dz because z is already chosen to maximize utility

OPTIMAL TOP INCOME TAX RATE

$$dM + dW + dB = Nd\tau \left\{ (1 - \bar{g})[z^m - z^*] - e\frac{\tau}{1-\tau} z^m \right\}$$

Optimal τ such that $dM + dW + dB = 0 \Rightarrow$

$$\frac{\tau}{1-\tau} = \frac{(1-\bar{g})(z_m/z^* - 1)}{e \cdot z_m/z^*}$$

Optimal $\tau \downarrow \bar{g}$ [redistributive tastes]

Optimal $\tau \downarrow$ with e [efficiency]

Optimal $\tau \uparrow z_m/z^*$ [thickness of top tail]

ZERO TOP RATE RESULT

Suppose top earner earns z^T, and second top earner earns z^S, then $z^m = z^T$ when $z^* > z^S \Rightarrow z^m/z^* \to 1$ when $z^* \to z^T \Rightarrow$

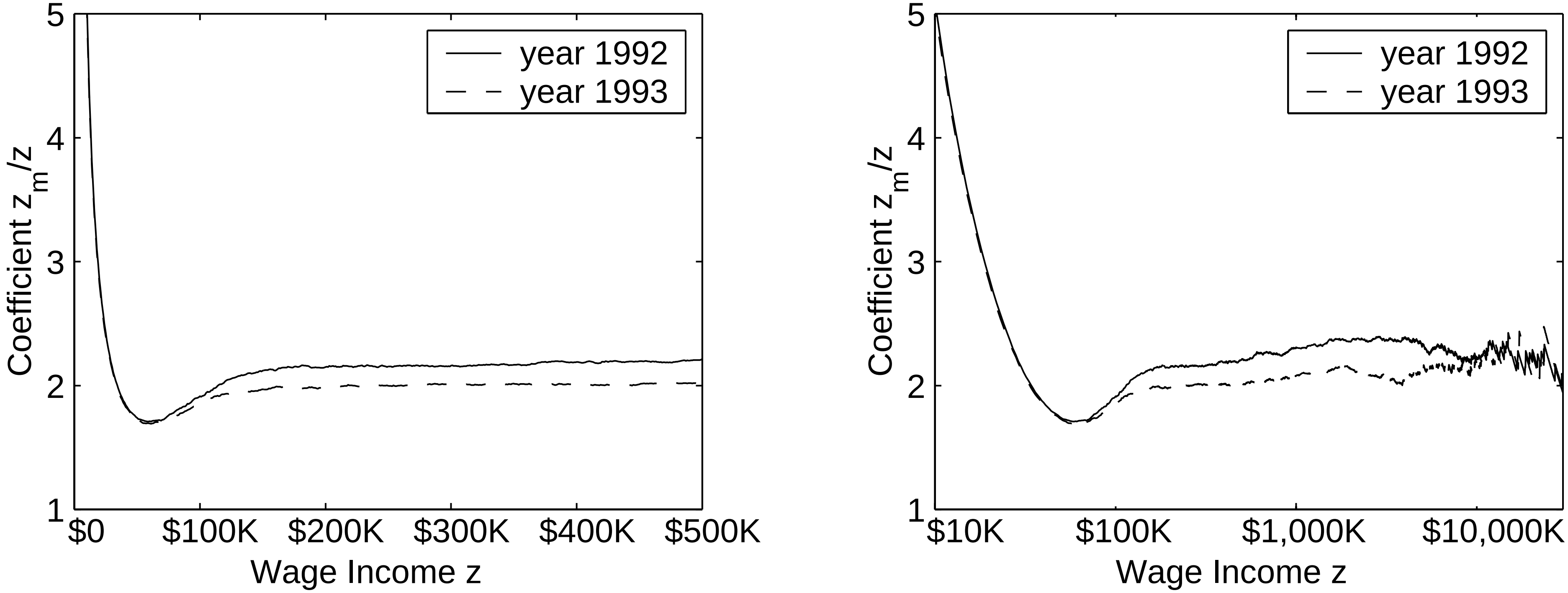
$dM = Nd\tau[z^m - z^*] << dB = Nd\tau e\frac{\tau}{1-\tau}z^m$ when $z^* \to z^T$

Intuition: extra tax applies only to earnings above z^* but behavioral response applies to full $z^m \Rightarrow$

Optimal τ should be zero when z^* close to z^T (Sadka-Seade zero top rate result)

Result applies only to top earner: if $z^T = 2 \cdot z^S$ then $z^m/z^* = 2$ when $z^* = z^S$

FIGURE 2 − Ratio mean income above z divided by z, z_m/z, years 1992 and 1993



OPTIMAL TOP INCOME TAX RATE

Empirically: z^m/z^* very stable above $z^* = \$200K$

Pareto distribution $1 - F(z) = (k/z)^a$, $f(z) = a \cdot k^a/z^{1+a}$, with a Pareto parameter

$$z^m(z^*) = \frac{\int_{z^*}^{\infty} z f(z) dz}{\int_{z^*}^{\infty} f(z) dz} = \frac{\int_{z^*}^{\infty} z^{-a} dz}{\int_{z^*}^{\infty} z^{-a-1} dz} = \frac{a}{a-1} \cdot z^*$$

a measures *thinness* of top tail of the distribution [log-normal has $a = \infty$ but empirically $a \in (1.5, 2.5)$]

$$\tau = \frac{1 - \bar{g}}{1 - \bar{g} + a \cdot e}$$

TAX REVENUE MAXIMIZING TAX RATE

Utilitarian criterion with $u_c \to 0$ when $c \to \infty \Rightarrow \bar{g} \to 0$ when $z^* \to \infty$

Rawlsian criterion $\Rightarrow \bar{g} = 0$ for any $z^* > \min(z)$

In the end, $\bar{g}$ reflects the value that society puts on marginal consumption of the rich

$\bar{g} = 0 \Rightarrow$ Tax Revenue Max Rate $\tau = 1/(1 + a \cdot e)$ (upper bound on top tax rate)

Example: $a = 2$ and $e = 0.5 \Rightarrow \tau = 50\%$

Laffer linear rate is a special case with $z^* = 0$, $z^m/z^* = \infty = a/(a-1)$ and hence $a = 1$, $\tau = 1/(1+e)$

EXTENSIONS AND LIMITATIONS

1) Model includes only intensive earnings response. Extensive earnings responses [entrepreneurship decisions, migration decisions] $\Rightarrow$ Formulas can be modified

2) Model does not include **fiscal externalities**: part of the response to $d\tau$ comes from **income shifting** which affects other taxes $\Rightarrow$ Formulas can be modified

3) Model does not include **classical externalities**: (a) charitable contributions, (b) positive spillovers (trickle down) [top earners underpaid], (c) negative spillovers [top earners overpaid]

Classical general equilibrium effects on prices are NOT externalities and do not affect formulas [Diamond-Mirrlees AER '71, Saez JpubE '04]

MIGRATION EFFECTS

Migration issues are particularly important at the top end (brain drain). Some theory papers (Mirrlees '82). No great empirical work (on individual side).

Migration depends on average tax rate. Define $P(z-T(z)|z)$ fraction of z earners in the country: Elasticity

$$\eta^m = \frac{z-T(z)}{P}\frac{\partial P}{\partial(z-T(z))}$$

Tax revenue maximizing formula becomes:

$$\tau = \frac{1}{1+a\cdot e+\bar{\eta}^m}$$

Note: $\bar{\eta}^m$ depends on size of jurisdiction: large for cities, zero worldwide $\Rightarrow$ (1) Redistribution easier in large jurisdictions, (2) Tax coordination across countries $\uparrow$ ability to redistribute (big issue currently in EU)

FISCAL EXTERNALITY EFFECTS

Behavioral response to income tax comes not only from reduced labor supply but also shifts to other forms of income or activities

Critical distinction is whether shift is toward **untaxed activities** (leisure, untaxed fringe benefits, perks) vs. **taxed activities** (deferred compensation, shift to corporate income tax base)

Shifts to untaxed activities do not change analysis because individuals optimize (Feldstein REStat '99)

Shifts to taxed activities create a **fiscal externality** which affects analysis (Saez-Slemrod-Giertz JEL' 09)

FISCAL EXTERNALITY EFFECTS

Go back to small reform $d\tau > 0$ above z^*: Reduction in individual tax base $dz^m < 0$. Assume fraction s of this reduction comes from shift to taxed activities (with average tax rate t).

Fiscal externality $dE = -t \cdot s \cdot N dz^m > 0$. Optimum $dM + dW + dB + dE = 0 \Rightarrow$

Tax Revenue Maximizing Rate:

$$\tau = \frac{1 + s \cdot t \cdot a \cdot e}{1 + a \cdot e}$$

Income shifting tax loopholes inefficient: closing loopholes can reduce the taxable income elasticity and increase redistributive power of govt

CLASSIC EXTERNALITIES

1) Classic externalities require additional Pigouvian correction on top of the regular optimal income tax (Sandmo ’75, Cremer-Gahvari-Ladoux JpubE ’98). Best to target directly externality if possible

3a) If top pay = marginal productivity, then no externalities, standard theory.

3b) If top pay $<$ marginal productivity (e.g., unions divert surplus from top to bottom workers or firm insurance)$\Rightarrow$ labor supply of top earners has positive externality and optimal tax rate should be lower

3c) If top pay $>$ marginal productivity (e.g., executives skim their companies)$\Rightarrow$ skimming is a negative externality for shareholders, tax on top pay may mitigate the externality

GENERAL NON-LINEAR INCOME TAX $T(z)$

(1) Lumpsum grant given to everybody equal to $-T(0)$

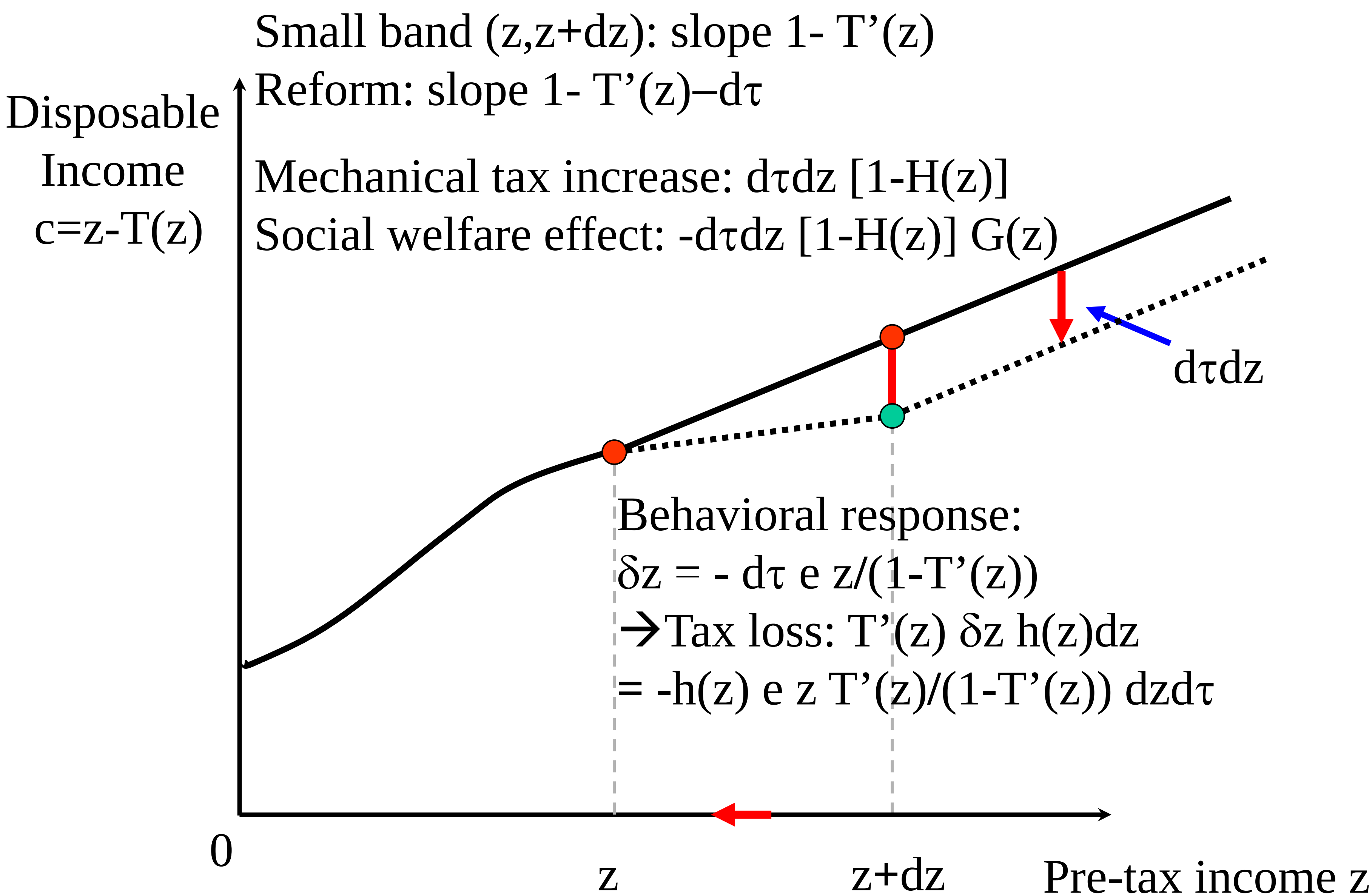
(2) Marginal tax rate schedule $T'(z)$ describing how (a) lumpsum grant is taxed away, (b) how tax liability increases with income

Let $H(z)$ be the income CDF [population normalized to 1] and $h(z)$ its density [endogenous to $T(.)$]

Let $g(z)$ be the social marginal value of consumption for taxpayers with income z in terms of public funds [formally $g(z) = G'(u) \cdot u_c/\lambda$]: no income effects $\Rightarrow \int g(z)h(z)dz = 1$

Redistribution valued $\Rightarrow g(z) \downarrow$ with z

Let $G(z)$ the *average* social marginal value of c for taxpayers with income above z [$G(z) = \int_z^\infty g(s)h(s)ds/(1 - H(z))$]


Small band (z,z+dz): slope 1- T'(z)
Reform: slope 1- T'(z)–dτ
Disposable
Income
c=z-T(z)
Mechanical tax increase: dτdz [1-H(z)]
Social welfare effect: -dτdz [1-H(z)] G(z)
dτdz
Behavioral response:
δz = - dτ e z/(1-T'(z))
→Tax loss: T'(z) δz h(z)dz
= -h(z) e z T'(z)/(1-T'(z)) dzdτ
0
z
z+dz
Pre-tax income z

GENERAL NON-LINEAR INCOME TAX

Assume away income effects $\varepsilon^c = \varepsilon^u = e$ [Diamond AER'98 shows this is the key theoretical simplification]

Consider small reform: increase T' by $d\tau$ in small band z and $z + dz$

Mechanical effect $dM = dz d\tau (1 - H(z))$

Welfare effect $dW = -dz d\tau (1 - H(z))G(z)$

Behavioral effect: substitution effect δz inside small band $[z, z + dz]$: $dB = h(z)dz \cdot T' \cdot \delta z = -h(z)dz \cdot T' \cdot d\tau \cdot z \cdot e_{(z)}/(1 - T')$

Optimum $dM + dW + dB = 0$

GENERAL NON-LINEAR INCOME TAX

$$\frac{T'(z)}{1-T'(z)} = \frac{1}{e_{(z)}} \left(\frac{1-H(z)}{zh(z)} \right) [1-G(z)]$$

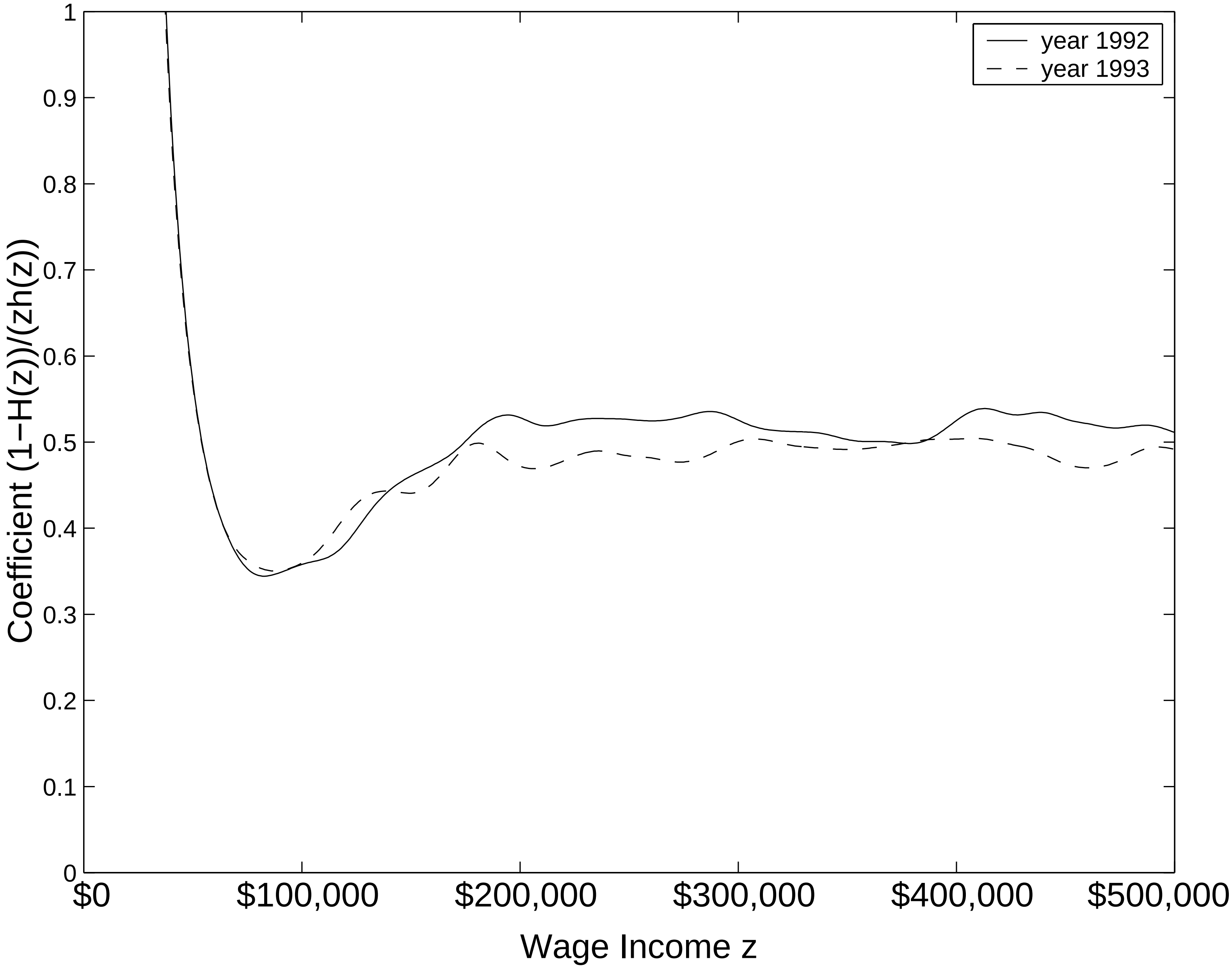
1) $T'(z) \downarrow e_{(z)}$ (elasticity efficiency effects)

2) $T'(z) \uparrow (1-H(z))/(zh(z))$ (shape of distribution efficiency effects)

3) $T'(z) \downarrow G(z)$ (redistributive tastes)

Asymptotics: $G(z) \to \bar{g}$, $(1-H(z))/(zh(z)) \to 1/a$, $e_{(z)} \to e \Rightarrow$ Recover top rate formula $\tau = (1-\bar{g})/(1-\bar{g}+a \cdot e)$

FIGURE 4 − Hazard Ratio (1−H(z))/(zh(z)), years 1992 and 1993



NEGATIVE MARGINAL TAX RATES NEVER OPTIMAL

Suppose $T' < 0$ in band $[z, z + dz]$

Increase T' by $d\tau > 0$ in band $[z, z + dz]$: $dM + dW > 0$ and $dB > 0$ because $T'(z) < 0$

$\Rightarrow$ Desirable reform

$\Rightarrow$ $T'(z) < 0$ cannot be optimal

NUMERICAL SIMULATIONS

$H(z)$ [and also $G(z)$] endogenous to $T(.)$. Calibration method (Saez Restud '01):

Specify utility function (e.g. constant elasticity):

$$u(c, z) = c - \frac{1}{1 + \frac{1}{e}} \cdot \left(\frac{z}{n}\right)^{1+\frac{1}{e}}$$

Individual FOC $\Rightarrow z = n^{1+e}(1 - T')^e$

Calibrate the exogenous skill distribution $F(n)$ so that, using **actual** $T'(.)$, you recover **empirical** $H(z)$

Use Mirrlees '71 tax formula (expressed in terms of $F(n)$) to obtain the optimal tax rate schedule T'.

NUMERICAL SIMULATIONS

$$\frac{T'(z(n))}{1-T'(z(n))} = \left(1+\frac{1}{e}\right)\left(\frac{1}{nf(n)}\right)\int_n^\infty \left[1-\frac{G'(u(m))}{\lambda}\right] f(m)dm,$$

Iterative Fixed Point method: start with T'_0, compute $z^0(n)$ using individual FOC, get $T^0(0)$ using govt budget, compute $u^0(n)$, get λ using $\lambda = \int G'(u)f$, use formula to estimate T'_1, iterate till convergence

Fast and effective method (Brewer-Saez-Shepard '09)

NUMERICAL SIMULATION RESULTS

$$\frac{T'(z)}{1-T'(z)} = \frac{1}{e_{(z)}} \left(\frac{1-H(z)}{zh(z)} \right) [1-G(z)]$$

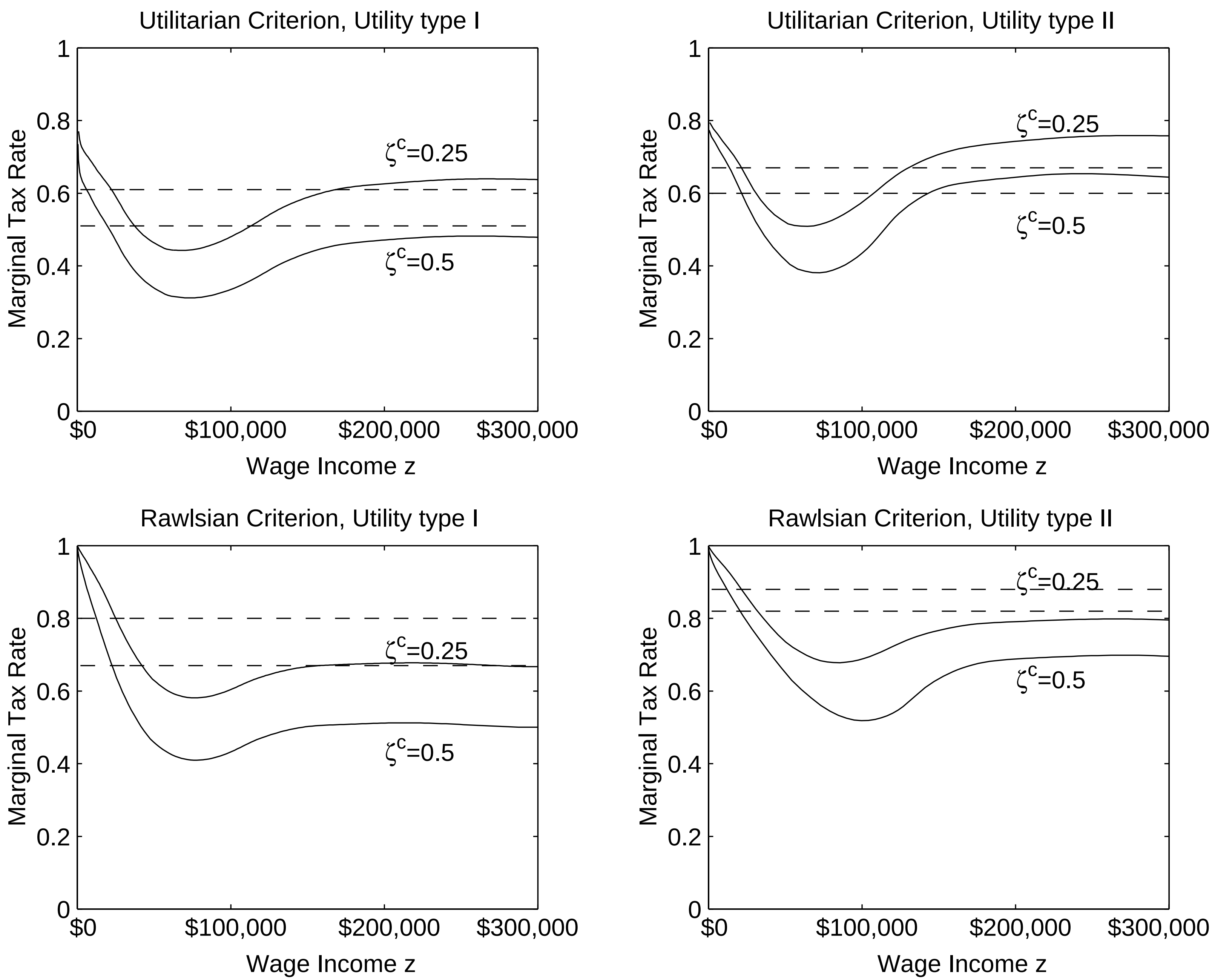
1) Take utility function with e constant

2) $(1-H(z))/(zh(z))$ is U-shaped empirically

3) $1-G(z) \uparrow$ with z from 0 to 1 ($\bar{g}=0$)

$\Rightarrow$ Numerical optimal $T'(z)$ is U-shaped with z: reverse of the general results $T'=0$ at top and bottom [Diamond AER'98 gives theoretical conditions to get U-shape]

FIGURE 5 − Optimal Tax Simulations



OPTIMAL TRANSFERS: MIRRLEES MODEL

Mirrlees model predicts that optimal transfer at bottom takes the form of a "Negative Income Tax":

1) Lumpsum grant $-T(0)$ 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 so intensive response does not generate large output losses

EXTENSIONS

1) Income effects can be introduced (Saez Restud '01). Keeping $\varepsilon_c(z)$ and $g(z)$ constant: Higher income effects $\Rightarrow$ Higher $T'(z)$ for high incomes

2) Inverted problem: use current $T(z)$ and $H(z)$ to back out welfare weights $g(z)$ [very sensitive to assumptions on $e_{(z)}$]

3) Pareto Efficient taxation (Werning '07): any tax schedule such that $g(z) \geq 0$ for all z is Pareto Efficient (and conversely)

If $g(z) < 0$ in some range, can design a tax reform that keeps utilities constant and raises tax revenue [tax system is locally on the wrong side of the Laffer curve]

COMMODITY VS. INCOME TAXATION

Suppose we have K consumption goods $c = (c_1, .., c_K)$ with pre-tax price $p = (p_1, .., p_K)$. Individual h has utility $u^h(c_1, .., c_K, z)$

Key question: Can government increase SWF using differentiated commodity taxation $t = (t_1, .., t_K)$ (after tax price $q = p + t$) in addition to nonlinear Mirrlees income tax on earnings z?

In practice, govt (a) exempts some goods (food, education, health) from sales tax or value-added-tax, (b) imposes additional excise taxes on some goods (cars, gasoline, luxury goods)

$\max_{t,T(.)} SWF \geq \max_{t=0,T(.)} SWF$ because more instruments cannot hurt

ATKINSON-STIGLITZ THEOREM

Famous Atkinson-Stiglitz JpubE' 76 shows that

$$\max_{t,T(.)} SWF = \max_{t=0,T(.)} SWF$$

(i.e, commodity taxes not useful) under two assumptions on utility functions $u^h(c_1, .., c_K, z)$

1) Weak separability between $(c_1, .., c_K)$ and z in utility

2) Homogeneity across individuals in the sub-utility of consumption $v(c_1, .., c_K)$ [does not vary with h]

$$u^h(c_1, .., c_K, z) = U^h(v(c_1, .., c_K), z)$$

Original proof was based on optimum conditions, new straightforward proof by Laroque EL '05, and Kaplow JpubE '06.

ATKINSON-STIGLITZ THEOREM PROOF

Let $V(y, p+t) = \max_c v(c_1, .., c_K)$ st $(p+t)\cdot c \leq y$ be the indirect utility of consumption c [common to all individuals]

Start with $(T(.), t)$. Let $c(t)$ be consumer choice.

Replace $(T(.), t)$ with $(\bar{T}(.), t = 0)$ where $\bar{T}(z)$ such that $V(z - T(z), p+t) = V(z - \bar{T}(z), p) \Rightarrow$ Utility $U^h(V, z)$ and labor supply choices z unchanged for all individuals.

Attaining $V(z - \bar{T}(z), p)$ at price p costs at least $z - \bar{T}(z)$

Consumer also attains $V(z - \bar{T}(z), p) = V(z - T(z), p + t)$ when choosing $c(t) \Rightarrow z - \bar{T}(z) \leq p \cdot c(t) = z - T(z) - t \cdot c(t)$

$\Rightarrow \bar{T}(z) \geq T(z) + t \cdot c(t)$: the government collects more taxes with $(\bar{T}(.), t = 0)$

ATKINSON-STIGLITZ INTUITION

With separability and homogeneity, conditional on earnings z, consumption choices $c = (c_1, .., c_K)$ do not provide any information on ability

$\Rightarrow$ Differentiated commodity taxes $t_1, .., t_K$ create a tax distortion with no benefit $\Rightarrow$ Better to do all the redistribution with the individual income tax

Note: With weaker linear income taxation tool (Diamond-Mirrlees AER '71, Diamond JpubE '75), need stronger assumptions on preferences (linear Engel curves, Deaton EL'81) to obtain no commodity tax result

Unless Engel curves are linear, commodity taxation can be useful to "non-linearize" the tax system

WHEN A-S ASSUMPTIONS FAIL

Thought experiment: force high ability people to work less and earn only as much as low ability people: if higher ability consume more of good k than lower ability people, then taxing good k is desirable. Happens when:

1) High ability people have a relatively higher taste for good k (independently of income) [indirect tagging]

2) Good k is positively related to leisure (consumption of k increases when leisure increases keeping after-tax income constant) [tax on holiday trips, subsidy on computers and work related expenses]

In general Atkison-Stiglitz assumption is a good starting place for most goods $\Rightarrow$ Zero-rating on some goods under VAT for redistribution is inefficient and administratively burdensome [Mirrlees review]

ATKINSON-STIGLITZ AND TAX ON SAVINGS

Standard two period model (w=wage rate in period 1, retired in period 2)

$$u^h(c_1, c_2, z) = u(c_1) + \frac{u(c_2)}{1+\delta} - b(z/w)$$

δ is the discount rate, $b(.)$ is the disutility of effort, budget $c_1 + c_2/(1 + r(1 - t_K)) \leq z - T(z)$

Aktinson-Stiglitz implies that savings taxation t_K (equivalent to tax on c_2) is useless in the presence of an optimal income tax if δ is the same for everybody

If low ability people have higher δ [empirically plausible] then savings tax $t_K > 0$ is desirable (Saez JpubE '02)

Diamond-Spinnewijn '09 consider nonlinear savings tax

ATKINSON-STIGLITZ AND TAX ON SAVINGS II

Conjecture to verify:

Suppose now that labor supply decision is about retirement age [length of work life vs. retirement life]

Savings are used for retirement consumption

$\Rightarrow$ Retirement consumption is positively related to leisure [high skill person retiring earlier and earning life-time like a low skilled person needs to save more to finance smooth consumption profile]

$\Rightarrow$ Retirement savings should be taxed

OPTIMAL TRANSFERS: MIRRLEES MODEL

Mirrlees model predicts that optimal transfer at bottom takes the form of a "Negative Income Tax":

1) Lumpsum grant $-T(0)$ 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 so intensive response does not generate large output losses

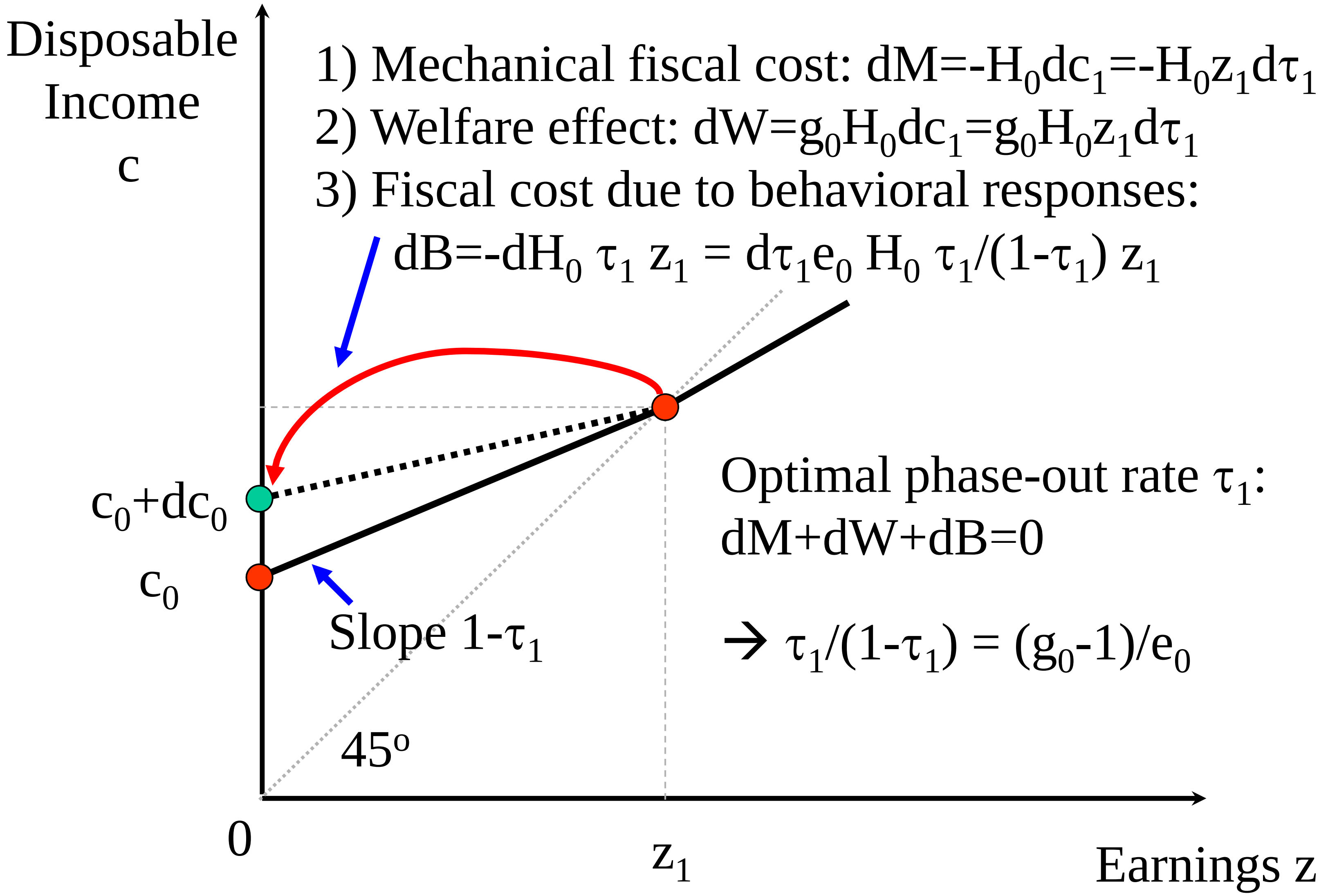
Reform: Increase τ_1 by $d\tau_1$ and c_0 by $dc_0=z_1 d\tau_1$

1) Mechanical fiscal cost: $dM=-H_0 dc_1=-H_0 z_1 d\tau_1$

2) Welfare effect: $dW=g_0 H_0 dc_1=g_0 H_0 z_1 d\tau_1$

3) Fiscal cost due to behavioral responses:

$dB=-dH_0\ \tau_1\ z_1 = d\tau_1 e_0\ H_0\ \tau_1/(1-\tau_1)\ z_1$



Optimal phase-out rate τ_1:

$dM+dW+dB=0$

$\rightarrow \tau_1/(1-\tau_1) = (g_0-1)/e_0$

OPTIMAL TRANSFERS: PARTICIPATION RESPONSES

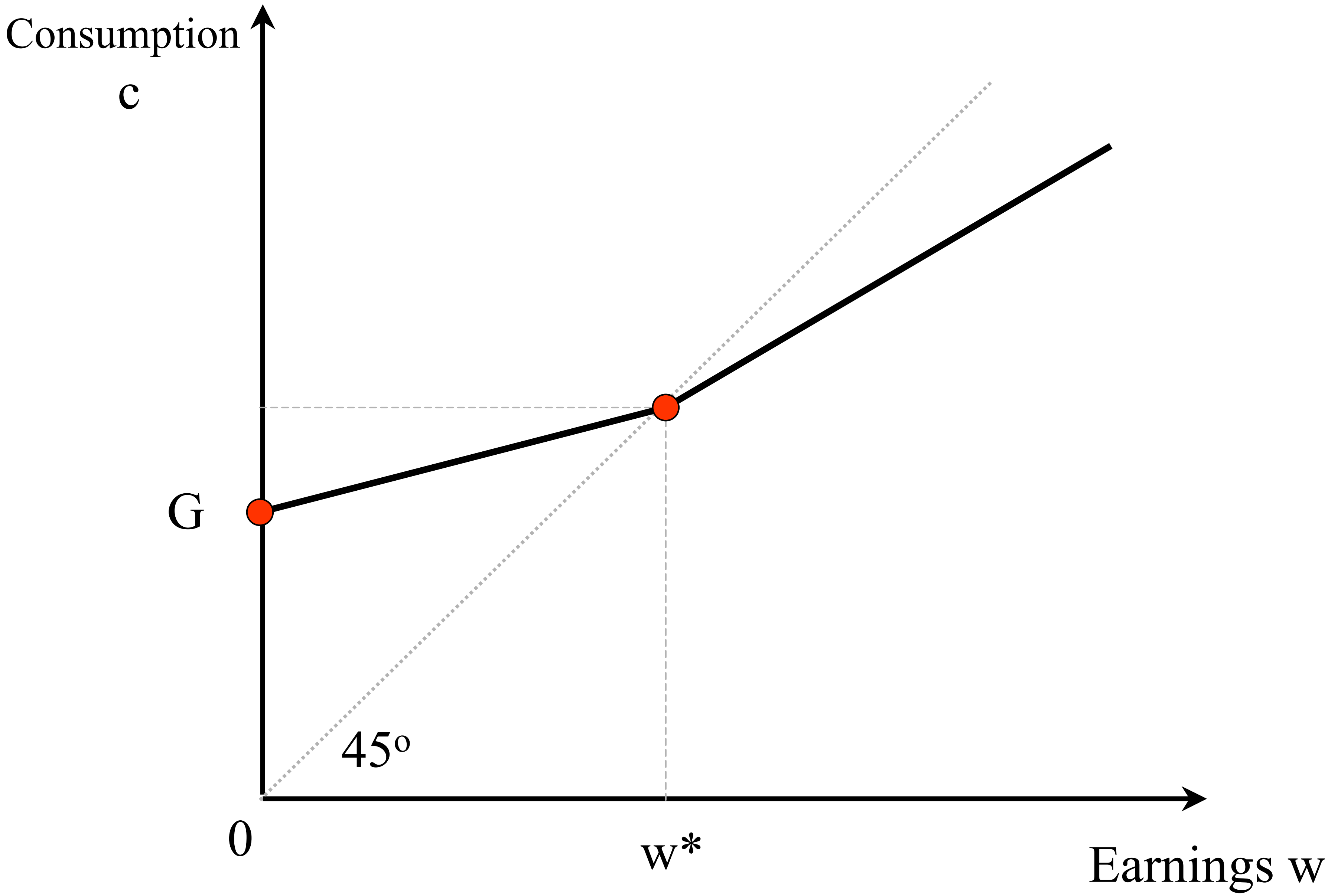
Empirical literature shows that participation labor supply responses [due to fixed costs of working] are large at the bottom [much larger and clearer than intensive responses]

Diamond JpubE'80, Saez QJE'02, Laroque EMA'05 incorporate such extensive labor supply responses in the optimal income tax model

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

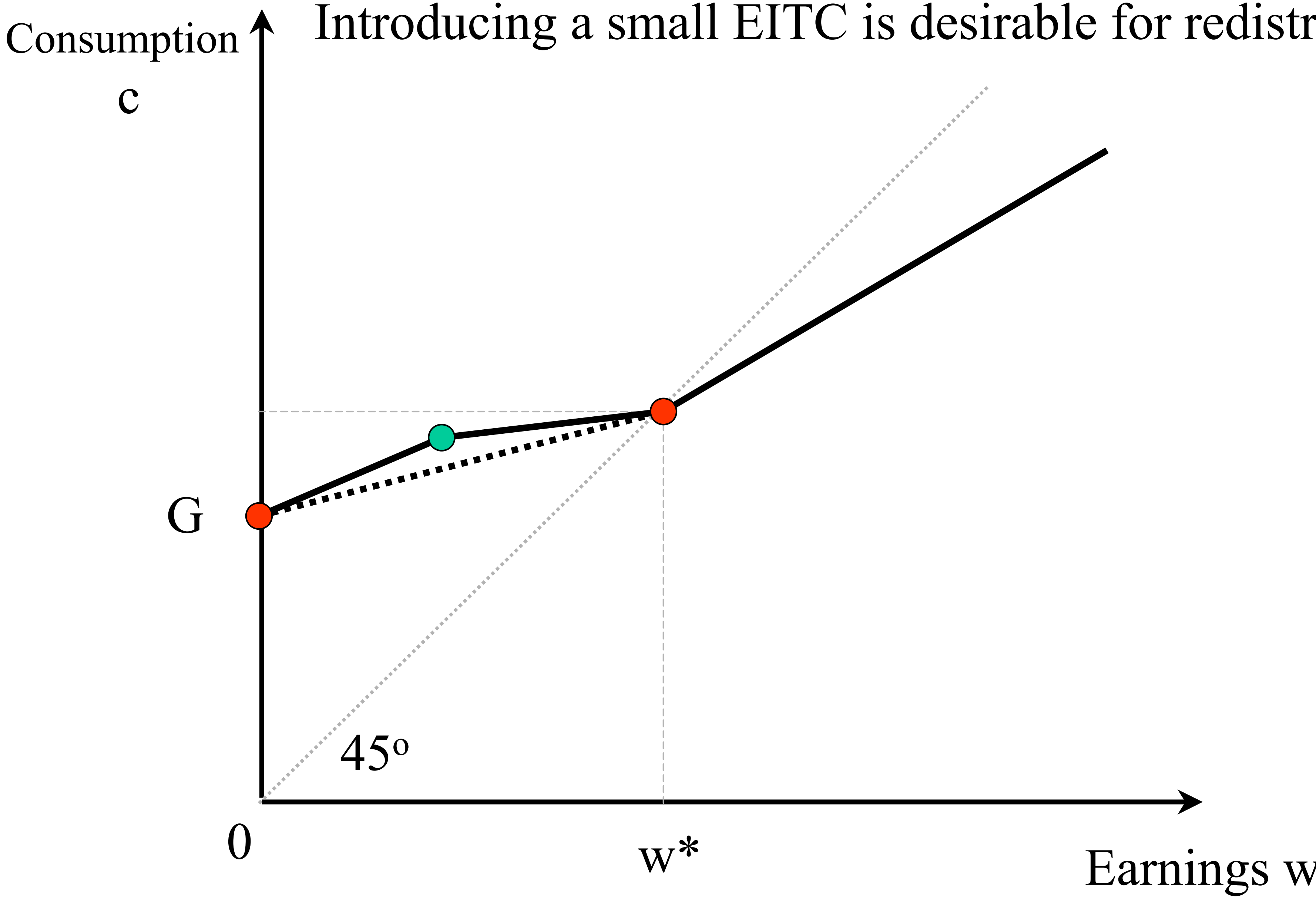
Key result: in-work subsidies with $T'(z) < 0$ (such as EITC) become optimal when labor supply responses are concentrated along extensive margin and social marginal welfare weight on low skilled workers > 1.

Starting from a Means-Tested Program



Starting from a Means-Tested Program

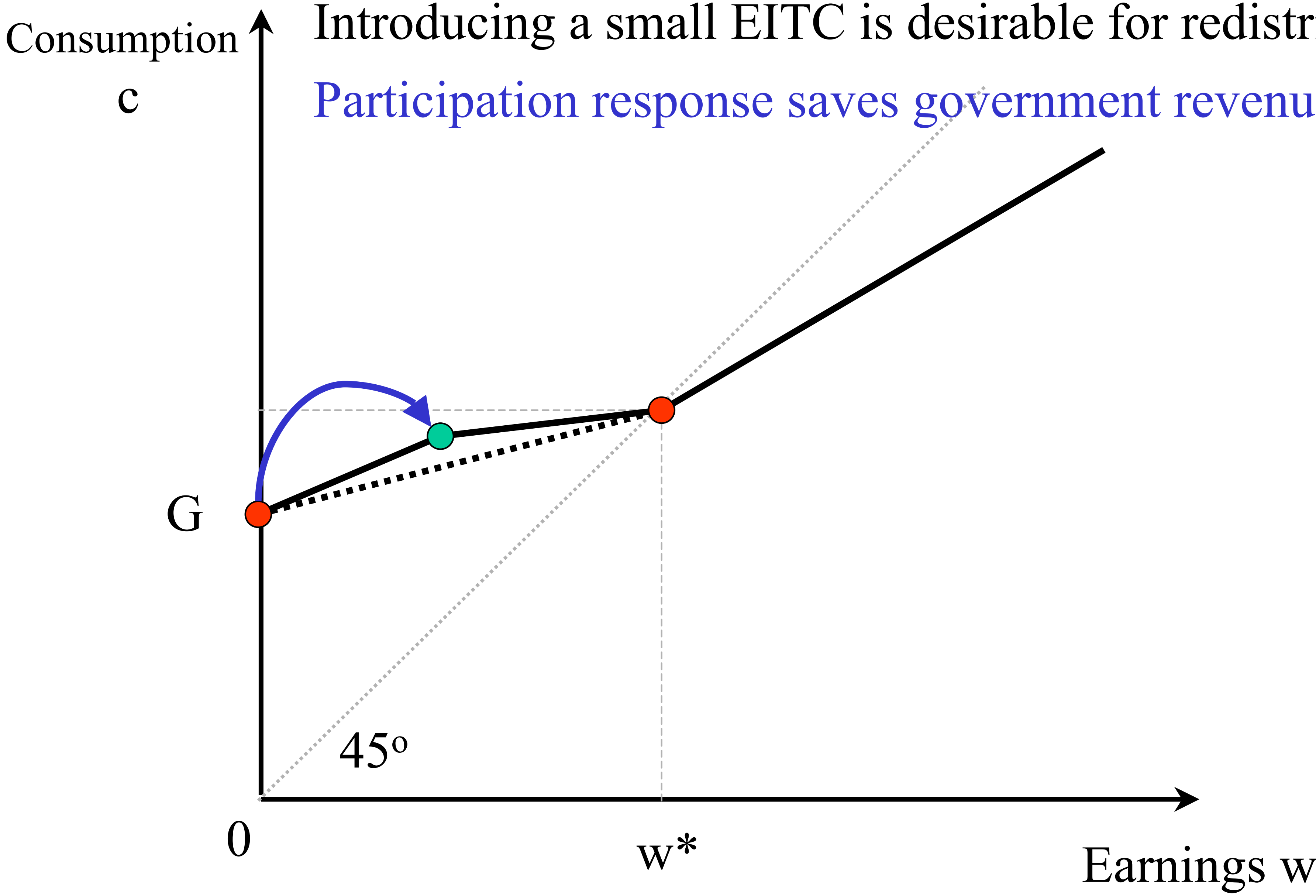
Introducing a small EITC is desirable for redistribution



Starting from a Means-Tested Program

Introducing a small EITC is desirable for redistribution

Participation response saves government revenue



SAEZ QJE'02 PARTICIPATION MODEL

Model with discrete earnings outcomes: $w_0 = 0 < w_1 < ... < w_I$

Tax/transfer T_i when earning w_i, $c_i = w_i - T_i$

Participation labor supply: Skill i individual compares c_i and c_0 when deciding to work $\Rightarrow$ Participation tax rate τ_i such that $c_i - c_0 = w_i \cdot (1 - \tau_i)$

$\Rightarrow$ In aggregate, fraction $h_i(c_i - c_0)$ of population earns w_i

Participation elasticity $e_i = (c_i - c_0)/h_i \cdot \partial h_i/\partial(c_i - c_0)$

Social Welfare function is summarized by social marginal welfare weights at each earnings level $g_i \downarrow i$, and average to one $\sum_i g_i h_i = 1$ (if no income effects)

Figure 3a: Optimal Tax/Transfer Derivation

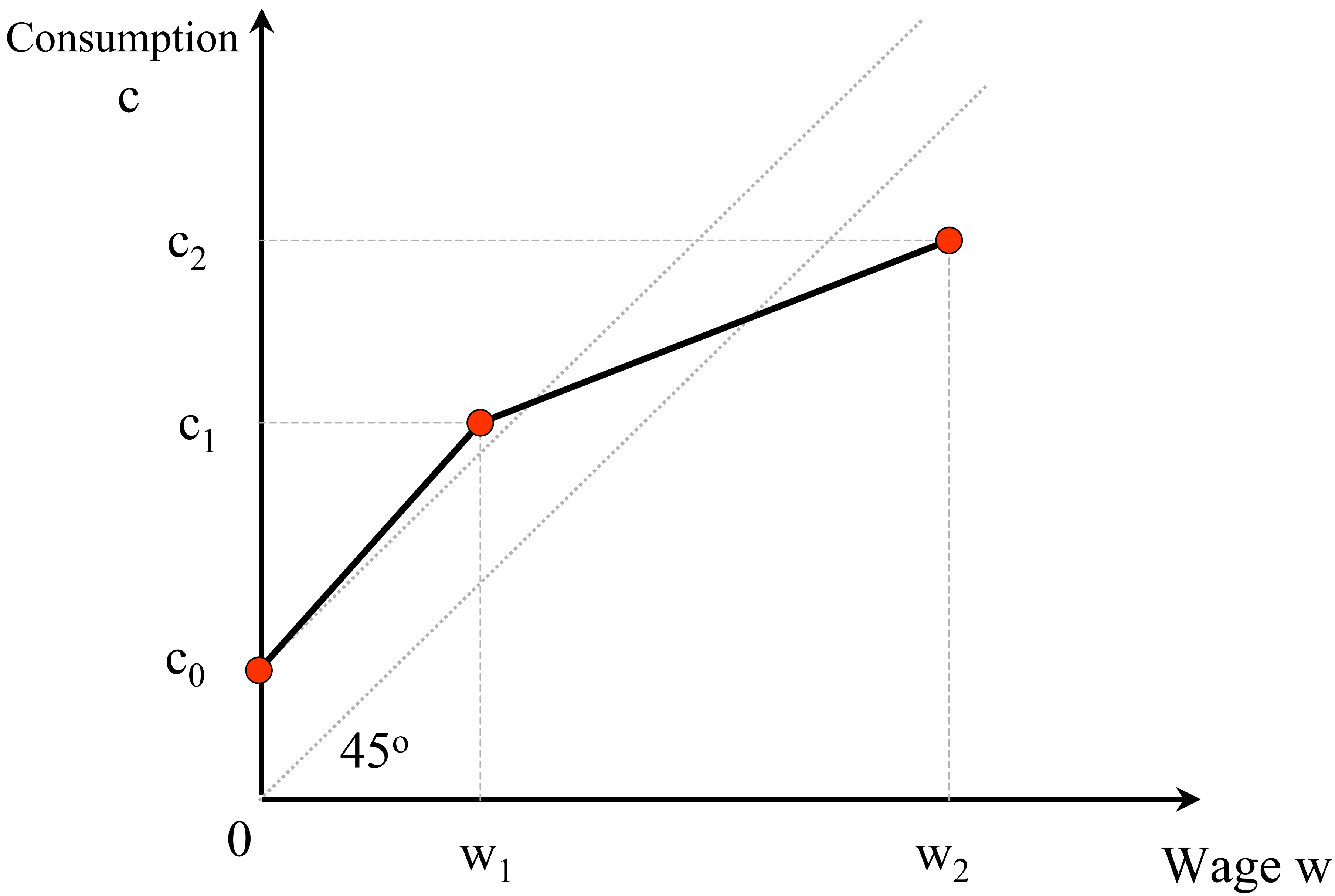


Figure 3a: Optimal Tax/Transfer Derivation (assuming $g_1>1$)

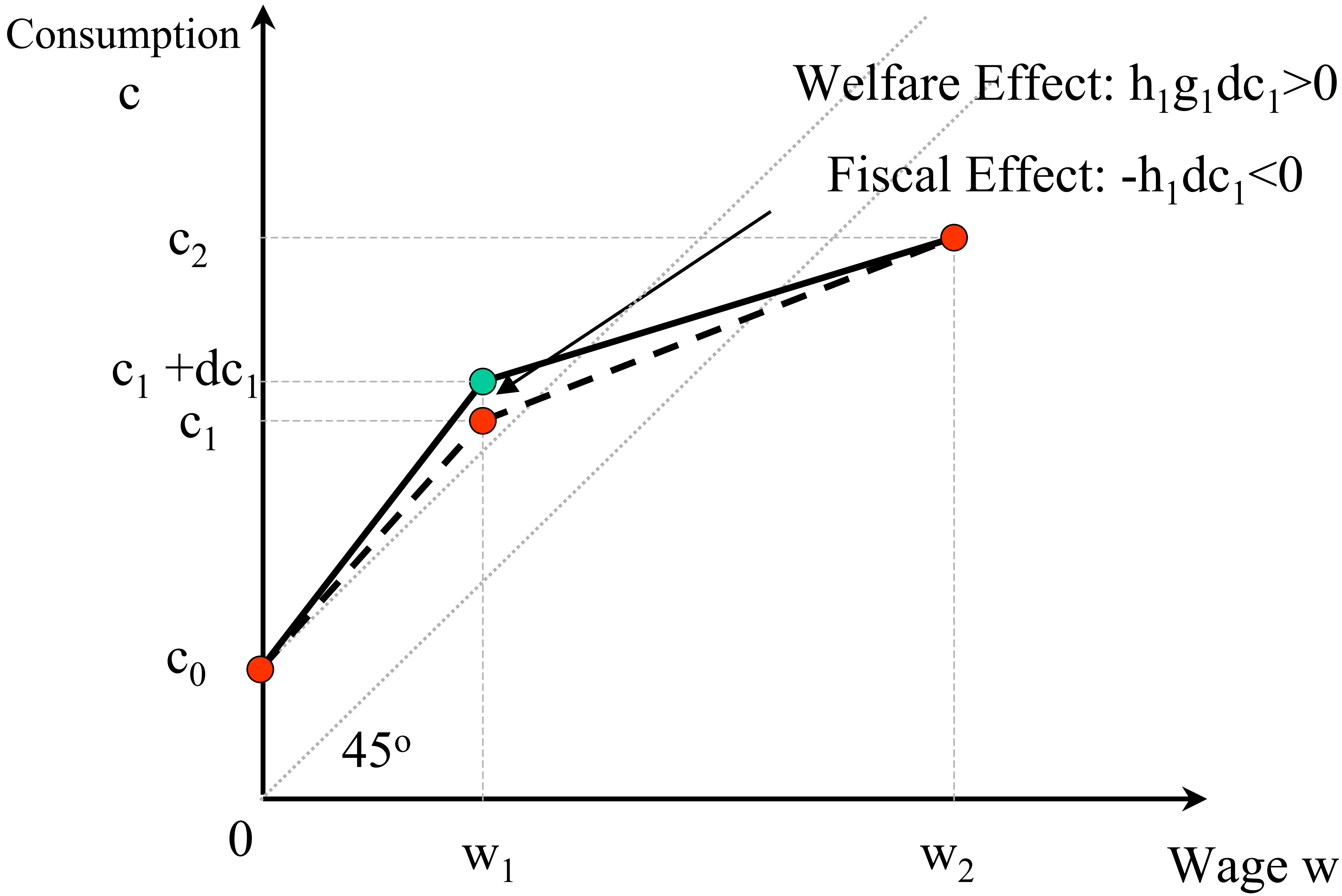


Figure 3a: Optimal Tax/Transfer Derivation (assuming $g_1>1$)

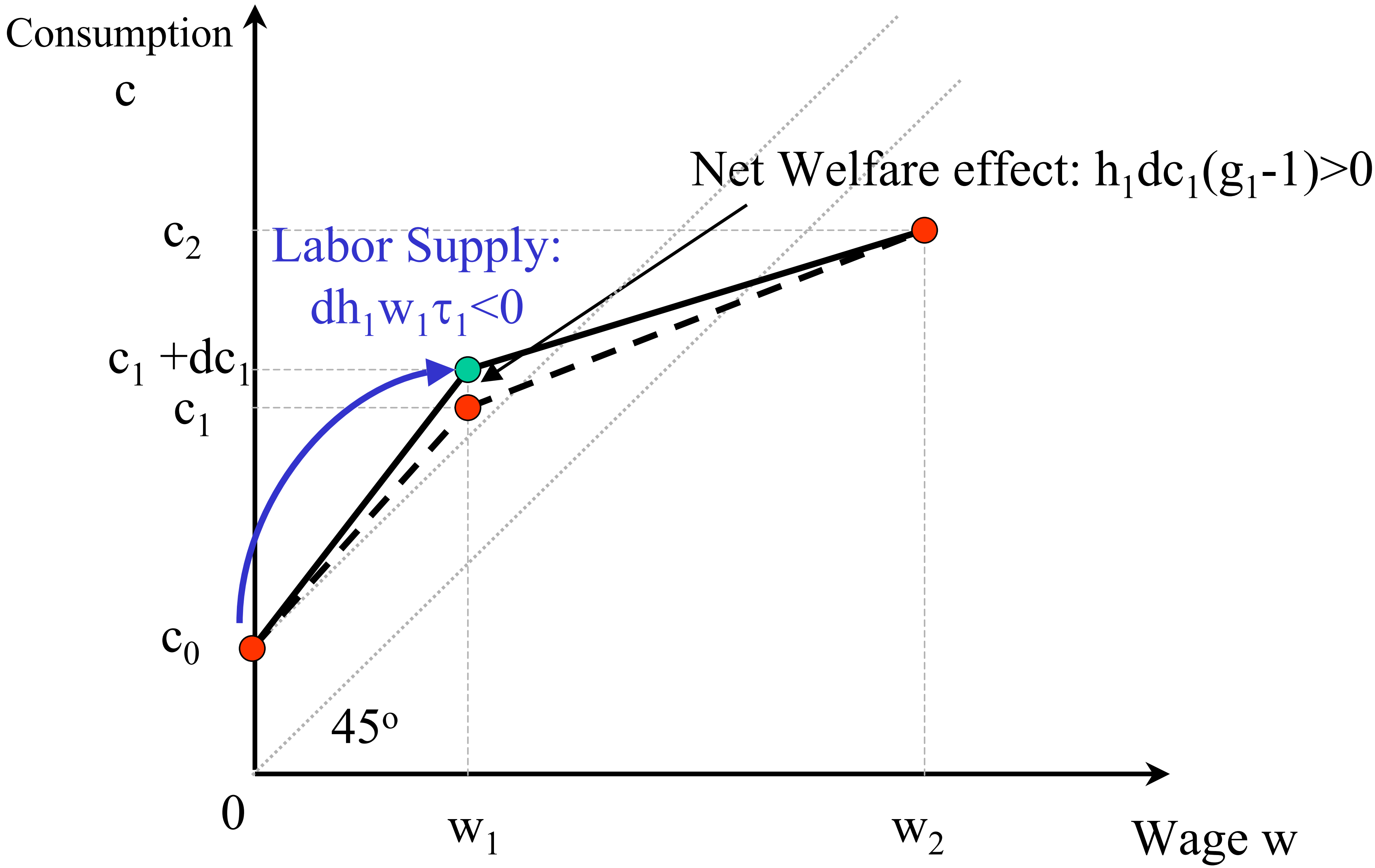
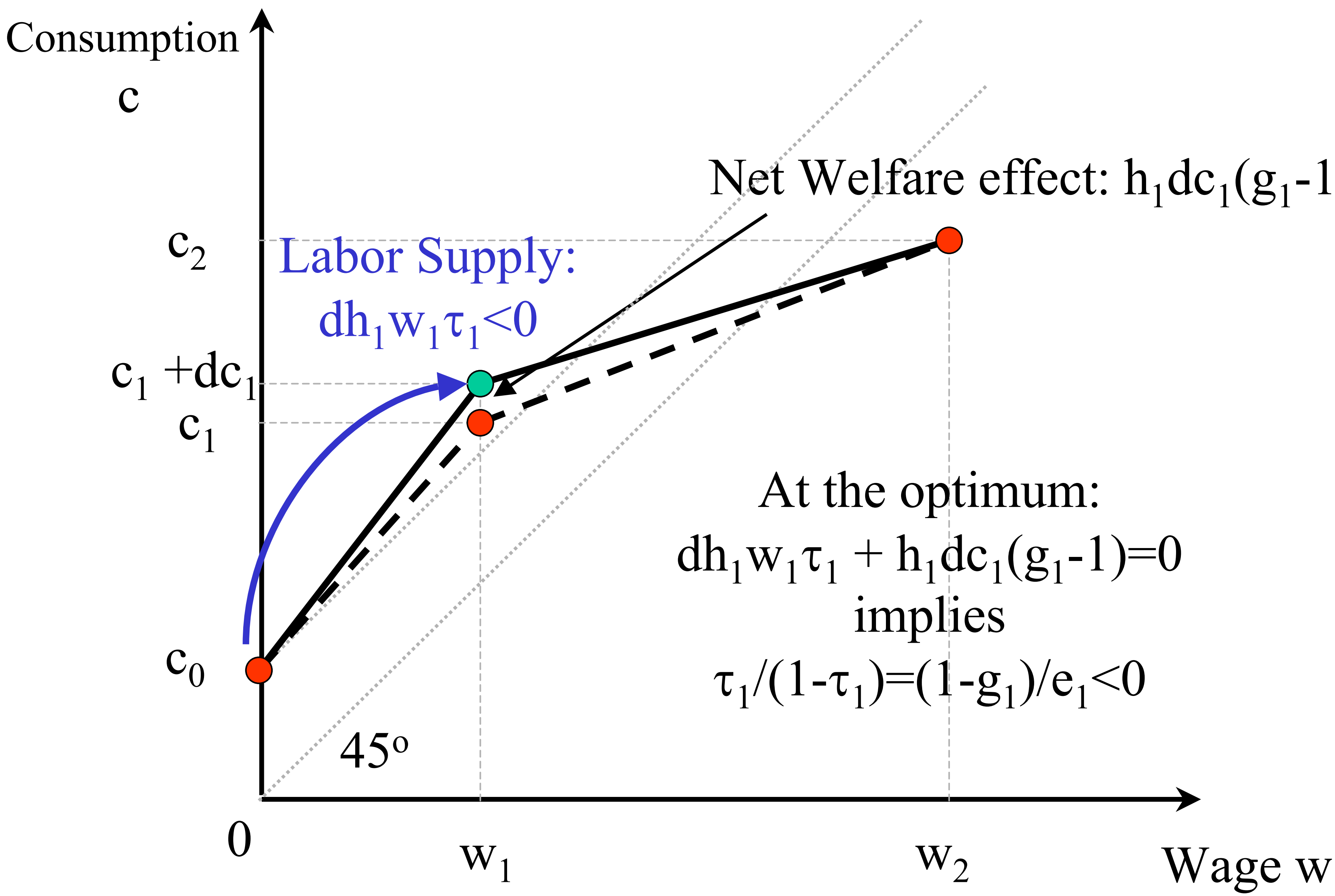


Figure 3a: Optimal Tax/Transfer Derivation (assuming $g_1 > 1$)



SAEZ QJE'02: OPTIMAL TAX DERIVATION

Small reform $dc_i = -dT_i > 0$. Three effects:

1) Mechanical Change in tax revenue $dM = h_i dT_i$

2) Behavioral Effect: $dh_i = -e_i h_i dT_i/(c_i - c_0) \Rightarrow$ Tax loss: $dB = -(T_i - T_0)dh_i = -e_i h_i dT_i(T_i - T_0)/(c_i - c_0)$

3) Welfare Effect: each worker in job i looses dT_i so welfare loss $dW = -g_i h_i dT_i$ [No first order welfare loss for switchers]

FOC: $dM + dB + dW = 0 \Rightarrow$

$$\frac{\tau_i}{1-\tau_i} = \frac{T_i - T_0}{c_i - c_0} = \frac{1}{e_i}(1 - g_i)$$

$g_1 > 1 \Rightarrow T_1 - T_0 < 0 \Rightarrow$ in-work subsidy

ACTUAL TAX/TRANSFER SYSTEMS

1) Transfer programs used to be of the traditional form with high phasing-out rates (sometimes above 100%) $\Rightarrow$ No incentives to work (even with modest elasticities)

2) In-work benefits have been introduced and expanded in OECD countries since 1980s (US EITC, UK Family Credit, etc.) and have been politically successful $\Rightarrow$ (a) Redistribute to low income workers, (b) improve incentives to work

OPTIMAL TRANSFERS IN RECESSIONS (GUESS)

1) The models we have covered consider only voluntary unemployment [people compare costs of work vs. benefits of work and can find a job if they want to]. Reasonable approximation during good times with low involuntary unemployment

2) During recessions (such as US in 2008-2009), many unemployed would like to work but cannot find a job

$\Rightarrow$ Labor supply participation responses shut down during recession [unemployed cannot find jobs, workers do not want to abandon jobs]

$\Rightarrow$ Redistribution becomes close to lumpsum [no efficiency costs while labor supply is frozen]

$\Rightarrow$ Redistributing more to non-workers during recessions is efficient [justification for extending unemployment benefits during recessions]

TAGGING

We have assumed that $T(z)$ depends only on earnings z.

In reality, govt can observe many other characteristics X also correlated with ability [gender, race, age, disability, family structure, height,...] and set $T(z, X)$. Two theory results:

1) If characteristic X is **immutable** then redistribution across the X groups will be complete [until average social marginal welfare weights are equated across X groups]

2) If characteristic X can be manipulated [behavioral response or cheating] but X correlated with ability then taxes will still depend on both X and z.

References: Akerlof AER'78 (welfare), Nichols-Zeckhauser AER'82 (welfare), Weinzierl '08 (age), Mankiw-Weinzierl '09 (height), Kaplow '07 (chap 7)

TAGGING WITH IMMUTABLE CHARACTERISTICS

Consider X binary immutable (**T**alls vs. **S**horts)

With $T(z)$ independent of X, Talls have higher ability on average $\Rightarrow$ Average social marginal welfare weights $\bar{g}^T < \bar{g}^S \Rightarrow$ Transfer from Talls to Shorts is desirable (surtax on Talls which finances an allowance on Shorts)

Optimal height transfers should be up to the point where $\bar{g}^T = \bar{g}^S$

Mankiw-Weinzierl '09 compute the optimal $T^{Tall}(z)$ and $T^{Short}(z)$ based on calibrated mode: optimal transfer $T^{Tall}(z) - T^{Short}(z)$ not trivial ($\simeq$ 10% of income)

Importantly: They show that you can get a (very modest) **Pareto** improvement using taxes on height and income instead of only income

PROBLEM WITH TAGGING

In practice public would oppose height based redistribution because height does not cause high earnings $\Rightarrow$

1) **Horizontal Equity** concerns [people with same "ability-to-pay" should pay the same tax] impose constraints on feasible policies [not captured by utilitarian framework]

2) Constrained optimization analysis [$T(z)$ instead of $T(z, X)$] remains valid even with heterogeneity in preferences

3) In practice $T(z, X)$ depends on X only when X is **directly** related to welfare [family structure, # kids, medical expenses] or ability to earn [disability status] ("ability-to-pay" intuition)

IN-KIND REDISTRIBUTION

Significant fraction of actual transfers are in-kind and often rationed (health care, education, public housing, nutrition subsidies) [care not cash San Francisco reform]

1) **Rational Individual perspective:**

(a) In-kind transfer is **tradeable** at market price $\Rightarrow$ in-kind equivalent to cash

(b) In-kind transfer **non-tradeable** $\Rightarrow$ in-kind inferior to cash.

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) Under standard welfarist objective: Efficiency considerations in a 2nd best context

EFFICIENCY OF IN-KIND REDISTRIBUTION

Depends on what income tax tools are available:

1) No income tax: Income z not observable (devo countries) $\Rightarrow$ In-kind provision or subsidies for necessities desirable

2) Linear tax model (Ramsey): Guesnerie-Roberts EMA'84 $\Rightarrow$ rationing goods encouraged by the tax system is desirable [and forcing consumption of goods discouraged by tax]

3) Nonlinear income tax: Under Atkinson-Stiglitz assumption [weak-separability and homogeneity $U^h(v(c_1,..,c_K),z)$] $\Rightarrow$ Any distortion (quota, rationing, subsidy) involving c choices not desirable provided $T(z)$ optimal

If good c_k related to leisure/ability [soup kitchen with queuing requirement] then A-S fails and in-kind redistribution possibly desirable even with optimal $T(z)$

IMPOSING ORDEALS ON TRANSFER RECIPIENTS

Many actual transfer programs impose requirements on beneficiaries (complex application, job search, training, or work requirements) and hence have low take-up (often $< 50\%$)

1) If social objective is welfarist and income z observable: ordeals unlikely to be desirable:

Compare ordeal to benefit cut: (a) only benefit cut saves money mechanically, (b) both reduce welfare of recipients, (c) both reduce take-up [good fiscally]

Need implausible sorting effects for ordeal to be desirable [e.g., ordeal does not hurt much deserving beneficiaries and discourages undeserving take-up, conditional on z]

2) Non-welfarist objective [such as poverty alleviation] or income z not observable: then ordeal can be desirable [Besley-Coate AER'92]

WORK RESTRICTIONS AND MINIMUM WAGE

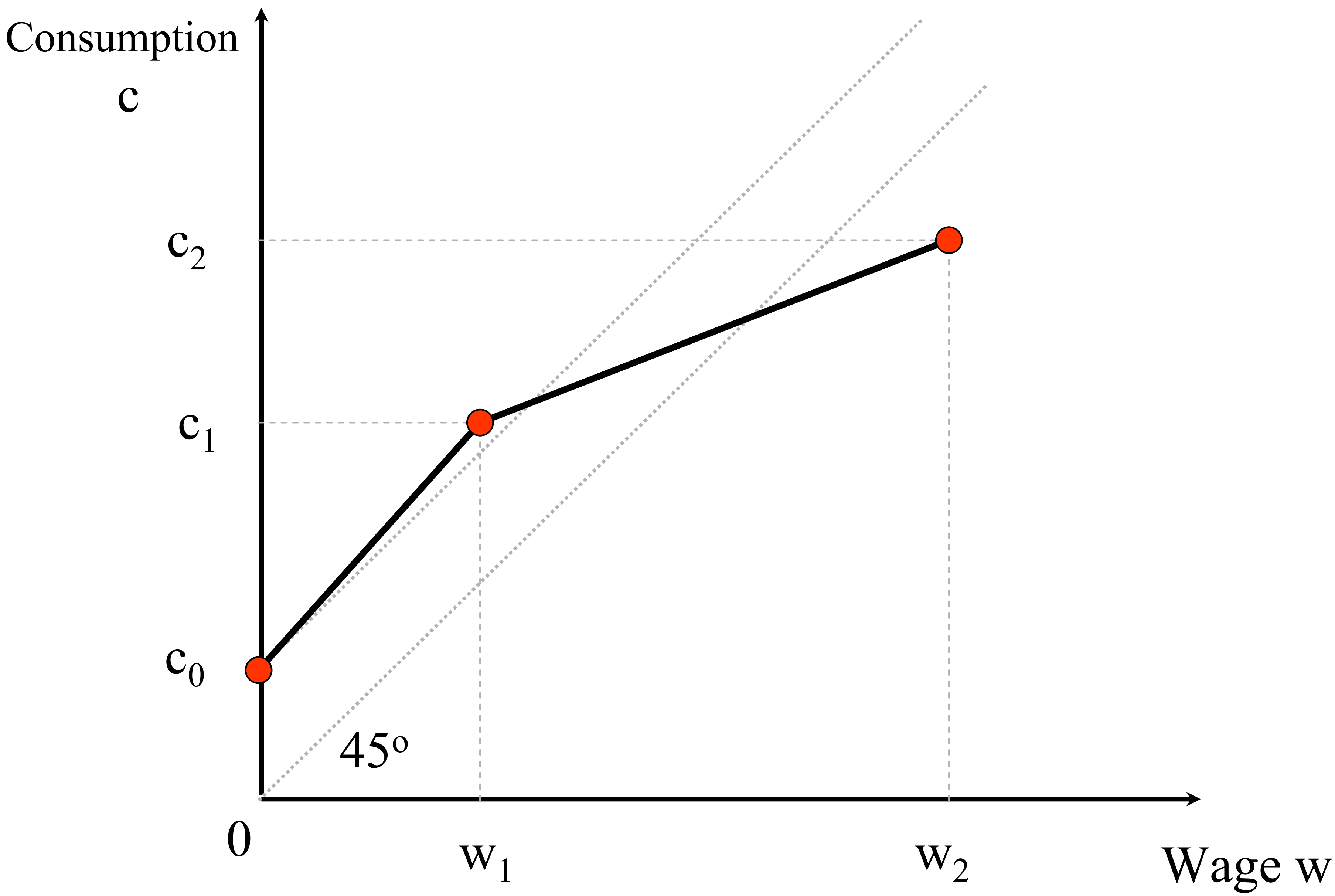
Minimum wage creates rationing of low skilled work. Could minimum wage be desirable on top of nonlinear tax/transfer?

Lee and Saez '08 use a job choice model [Saez QJE '02 with endogenous wages]. Two results:

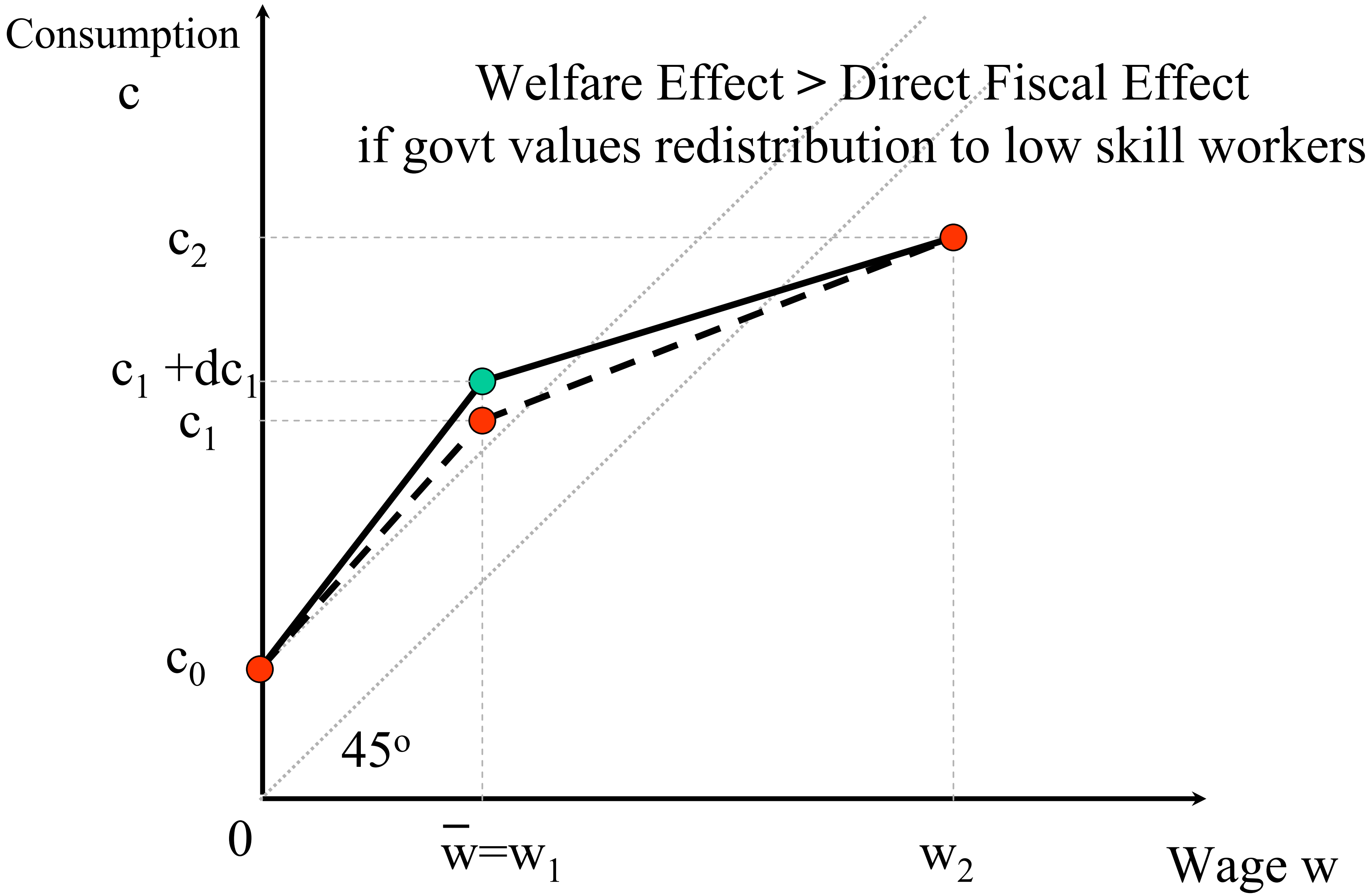
1) Minimum wage desirable if (a) govt wants to redistribute to low skilled workers ($g_1 > 1$) and (b) rationing created by min wage is **efficient**

2) If labor supply responses along extensive margin only then minimum wage with positive tax rate on low skilled work $\tau_1 > 0$ is 2nd best Pareto inefficient [delivers strong policy reform prescription]

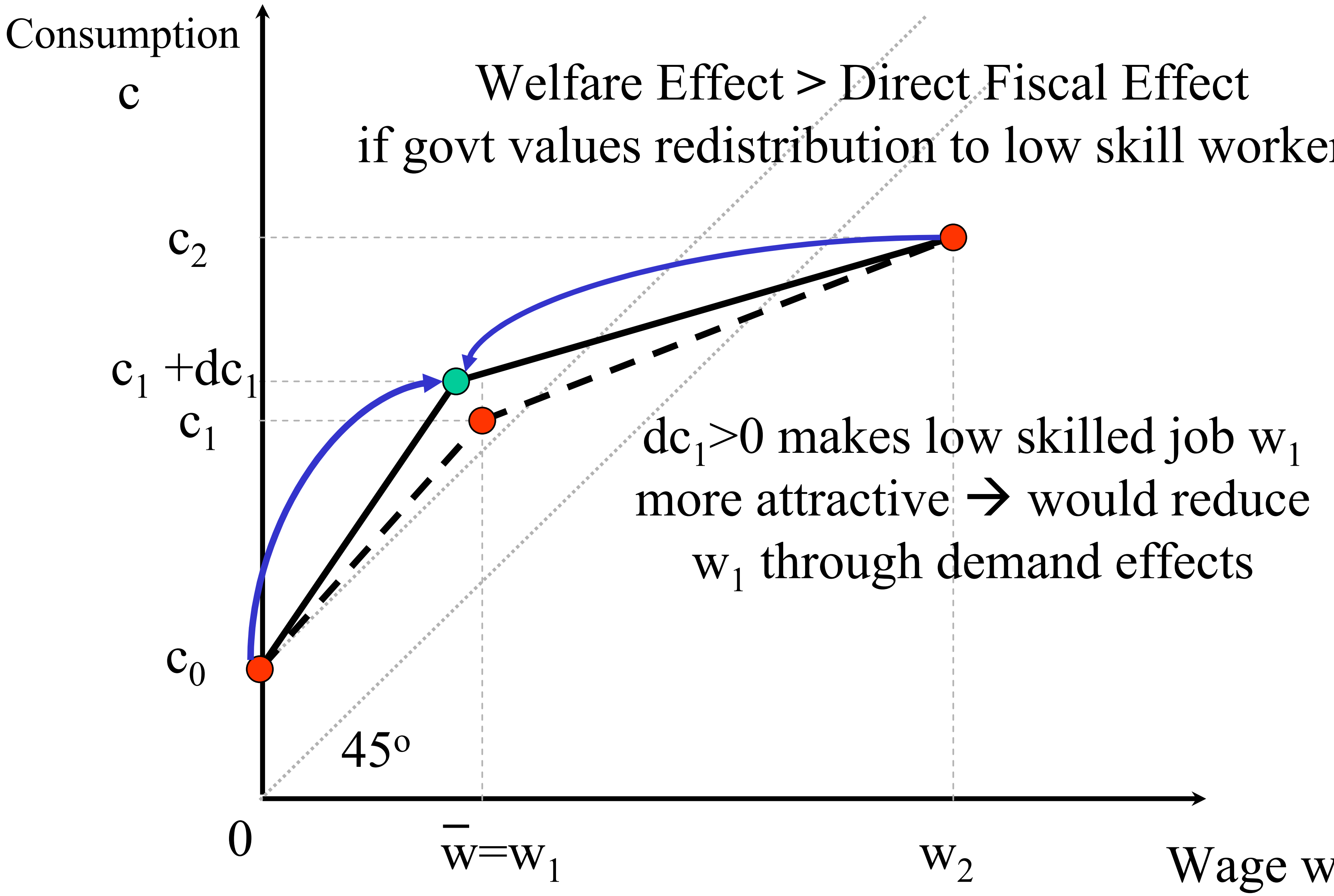
2. Optimal Tax/Transfer System (no min wage)



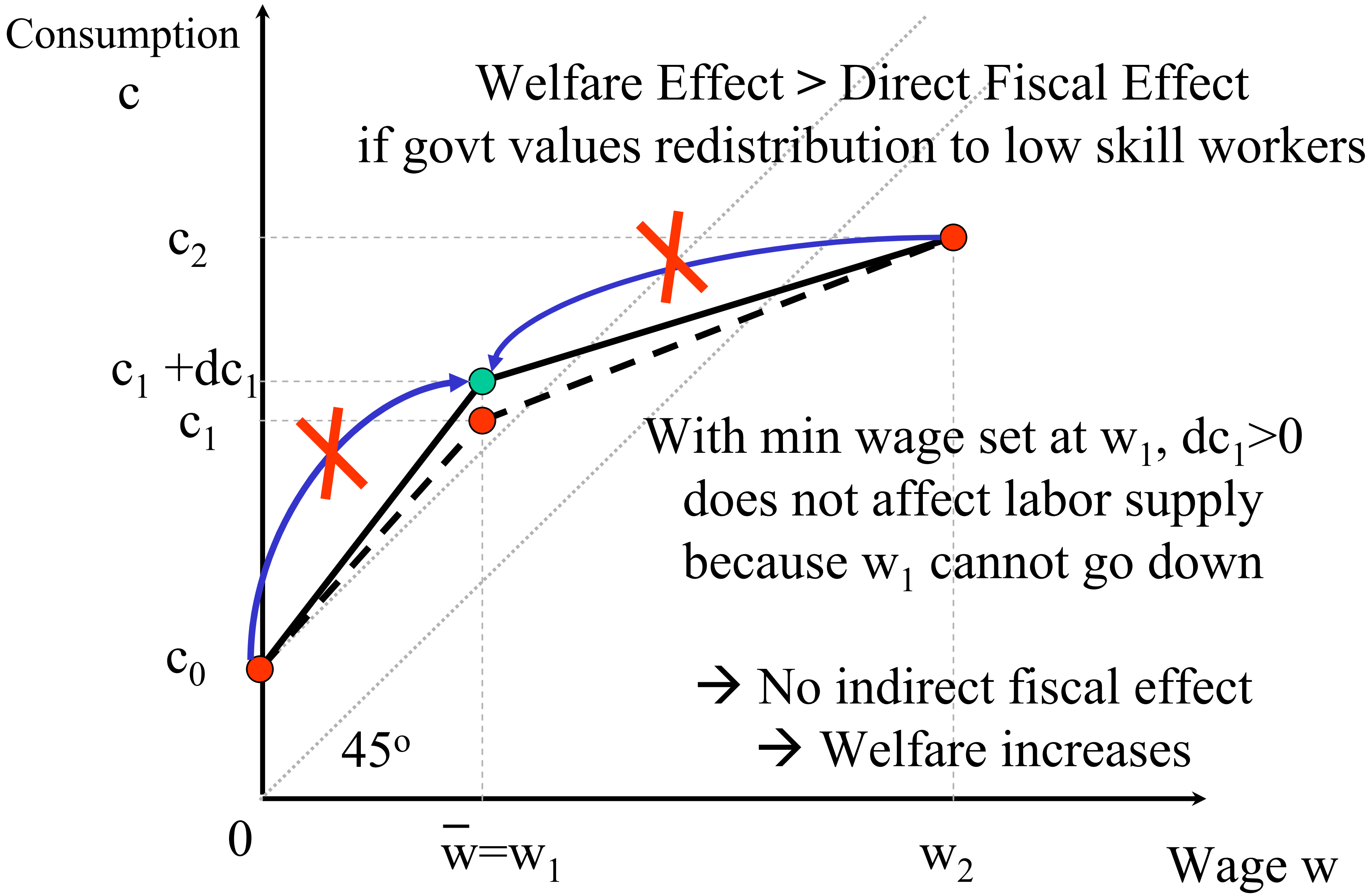
2. Set Min wage $\bar{w}=w_1$ and increase c_1 by dc_1



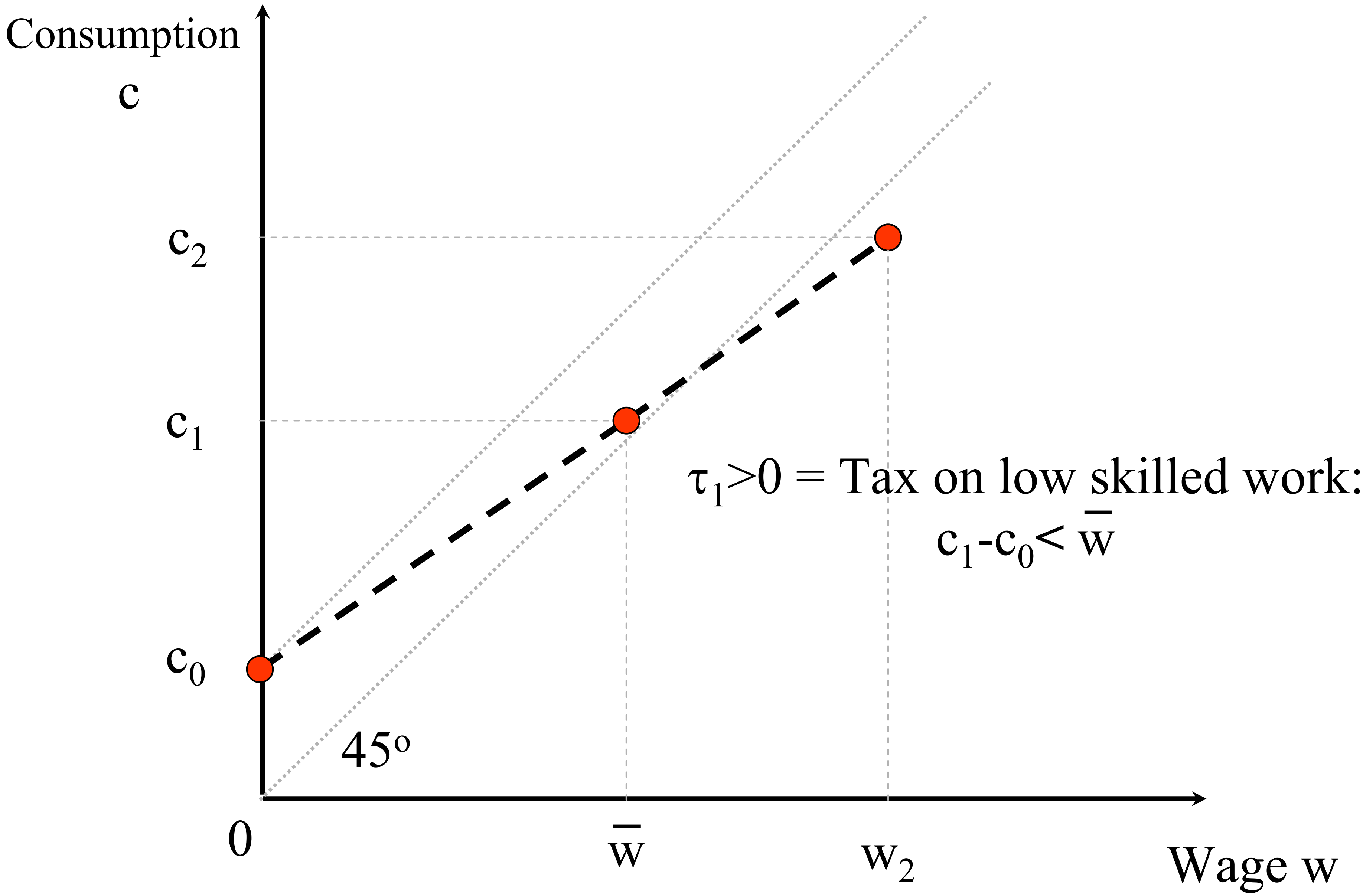
2. Desirability of Min Wage with Optimal Taxes



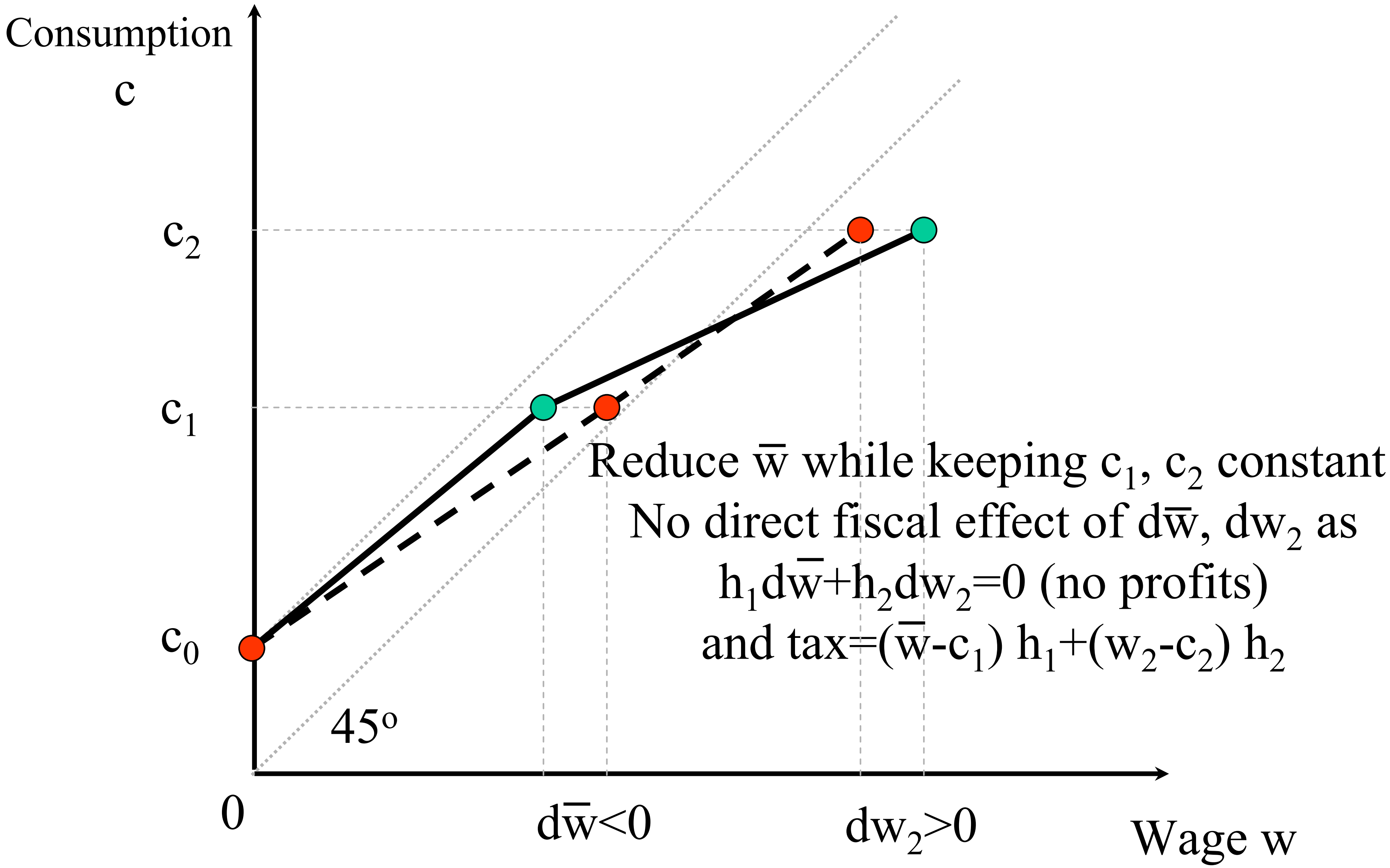
2. Desirability of Min Wage with Optimal Taxes



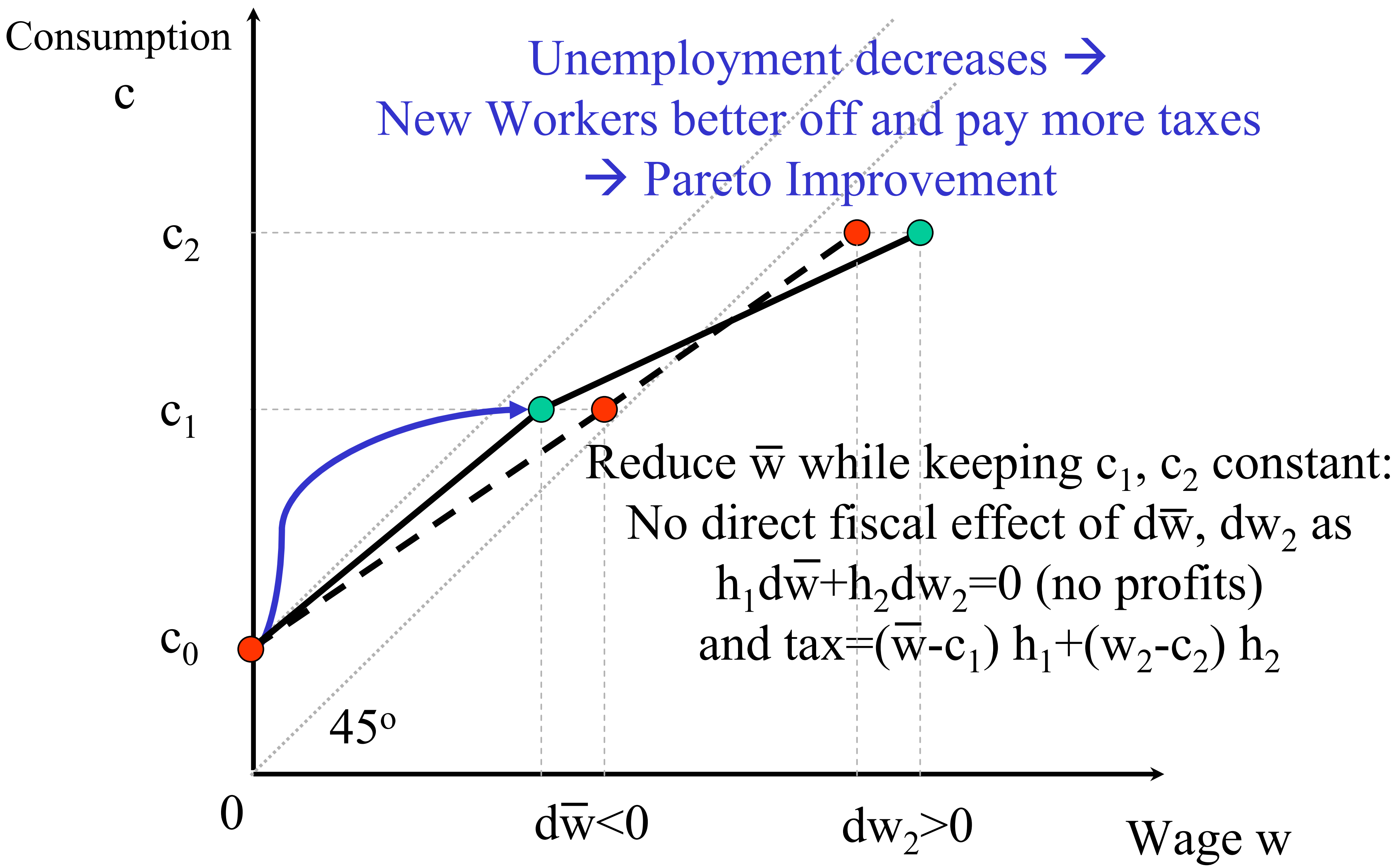
3. Pareto Improving Policy when $\tau_1 > 0$ and min wage binds



3.Pareto Improving Policy when τ_1>0 and min wage binds



3. Pareto Improving Policy when $\tau_1 > 0$ and min wage binds



FAMILY TAXATION: MARRIAGE AND CHILDREN

Two important issues in policy debate:

1) Marriage: What is the optimal taxation of couples vs. singles? Should secondary earnings be treated differently?

2) Children: What should be the net transfer (transfer or tax reduction) for family with children (as a function of family income and structure)?

Theoretical literature is not great in part because utilitarian framework is not fully satisfactory

TAXATION OF COUPLES

1) Economies of scale and sharing in consumption within families $\Rightarrow$ Welfare best measured by family income relative to size [$\equiv$ **normalized income**]

$\Rightarrow$ Taxes/Transfers should be based on family income which can create a marriage penalty / subsidy

Note: Impossible to have a tax/transfer system that (1) is family income based, (2) has marriage neutrality, (3) is progressive (i.e., not strictly linear)

2) If marriage responds to tax/transfer differential $\Rightarrow$ better to reduce marriage penalty, i.e., move toward individualized system

Particularly important when hard to observe cohabitation can substitute for marriage (Scandinavian countries)

TAXATION OF COUPLES

3) Labor supply of secondary earners more elastic than labor supply of primary earner $\Rightarrow$ Secondary earnings should be taxed less (standard Ramsey intuition, Boskin-Sheshinski JpubE'83)

Labor supply elasticity differential is $\downarrow$ as earnings gender gap $\downarrow$

4) Welfare effect of spousal earnings $\downarrow$ primary earnings $\Rightarrow$ Transfer for having a non-working spouse [=tax on secondary earnings] $\downarrow$ primary earnings [Kleven-Kreiner-Saez EMA'09]

In OECD countries: income tax systems have become **individual based** but means tested transfers have remained **family based**

TRANSFERS OR TAX CREDITS FOR CHILDREN

1) Children reduce **normalized 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 $T_{kid}(z) \uparrow$ with income z?

Pro: they reduce normalized income most for upper earners [e.g., France computes taxes as $N \cdot T(z/N)$ where N is # family members, kids count as .5 $\Rightarrow T_{kid}(z) \uparrow z$].

Cons: lower earners need child transfers most [most OECD countries have means-tested transfers conditional on number of kids $\Rightarrow T_{kid}(z) \downarrow z$, US has $T_{kid}(z)$ inverted U-shape due to EITC and Child Tax Credit]

TRANSFERS OR TAX CREDITS FOR CHILDREN

3) Family does not make decisions as a single unit (Chiappori): transfers to mothers has bigger effects on children's consumption than transfers to fathers [Lundberg JHR, Duflo WBER '99]

4) Children create externalities [positive: retirement programs, negative: global warming]. If fertility responds to transfers, case for subsidizing/taxing children [Europe vs. China]

5) Child care cash costs are positively related to work $\Rightarrow$ Such costs should be subsidized by Atkinson-Stiglitz [often they are in practice]

CHILDREN AND LIMITS OF UTILITARIAN MODEL

If fertility decisions unrelated to children tax/transfers $\Rightarrow$ Social marginal utility should be equated across families with 0 children, families with 1 child, etc.

If ability uncorrelated with children $\Rightarrow$ Families with kids will get fully compensating transfers

If ability positively correlated with children $\Rightarrow$ Families with kids might be taxed more heavily [as in the height tax case]

Seems an absurd model to think about transfers for children $\Rightarrow$ Need to come up with more realistic alternative