OPTIMAL TRANSFER PROGRAMS: TARGETING, TAGGING, AND ORDEALS

Henrik Jacobsen Kleven

London School of Economics

Lecture Notes for MSc Public Economics (Ec426): Lent term 2009

AGENDA

1. The different types of social programs.

2. The moral hazard cost of low-income support.

3. Reducing moral hazard through better targeting:

   (a) Tagging [Akerlof, AER 1978]

   (b) In-kind benefits, goods subsidies [Nichols-Zeckhauser, AER 1982]

   (c) Ordeals [Nichols-Zeckhauser, AER 1982; Besley-Coate, AER 1992]

4. Lessons from U.S. welfare reform efforts.

TYPES OF SOCIAL PROGRAMS: WHO HAVE ACCESS?

Universal programs: equal access to benefits for all citizens.

Means-tested programs: access is restricted by income and assets.

Categorical programs: access restricted by personal characteristics (kids, marital status, immigration status, disability, old age).

Many programs are both means-tested and categorical, e.g. programs targeted to low-income single mothers (e.g., TANF in the US).

TYPES OF SOCIAL PROGRAMS: WHAT KIND OF BENEFIT?

Cash programs: provide cash benefits (e.g., TANF in the US).

In-kind programs: provide goods such as health care (e.g., Medicaid in the US), food (e.g., Food Stamps in the US) and housing.

The principle of consumer sovereignty seems to suggest that cash is always better than goods.

But this argument may be wrong. To see why, we have to consider the moral hazard cost—or efficiency cost—of social programs.

A MEANS-TESTED CASH PROGRAM

A typical means-tested cash program provides benefits

$$B = \begin{cases} G - t \cdot wh & wh \leq G/t \\ 0 & wh > G/t \end{cases}$$

where G is an income guarantee, t is a phase-out rate, wh is earnings.

If $t = 1$, the program brings everyone below the G-level (e.g. the poverty line) up to G, but provides no support above G.

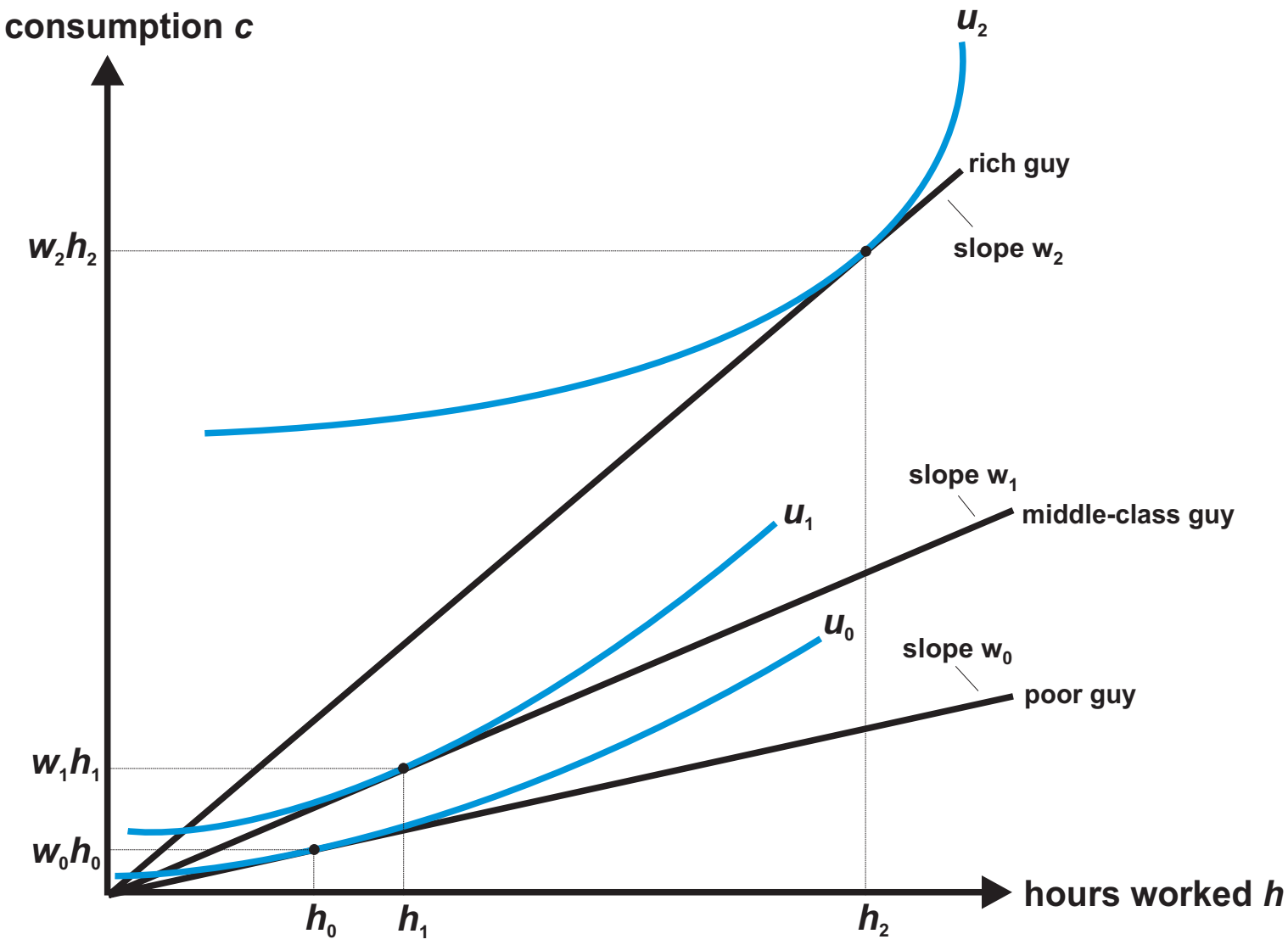
In the phase-out range $wh \leq G/t$, the individual budget constraint is

$$c \leq wh + B \quad \Leftrightarrow \quad c \leq (1 - t)\, wh + G$$

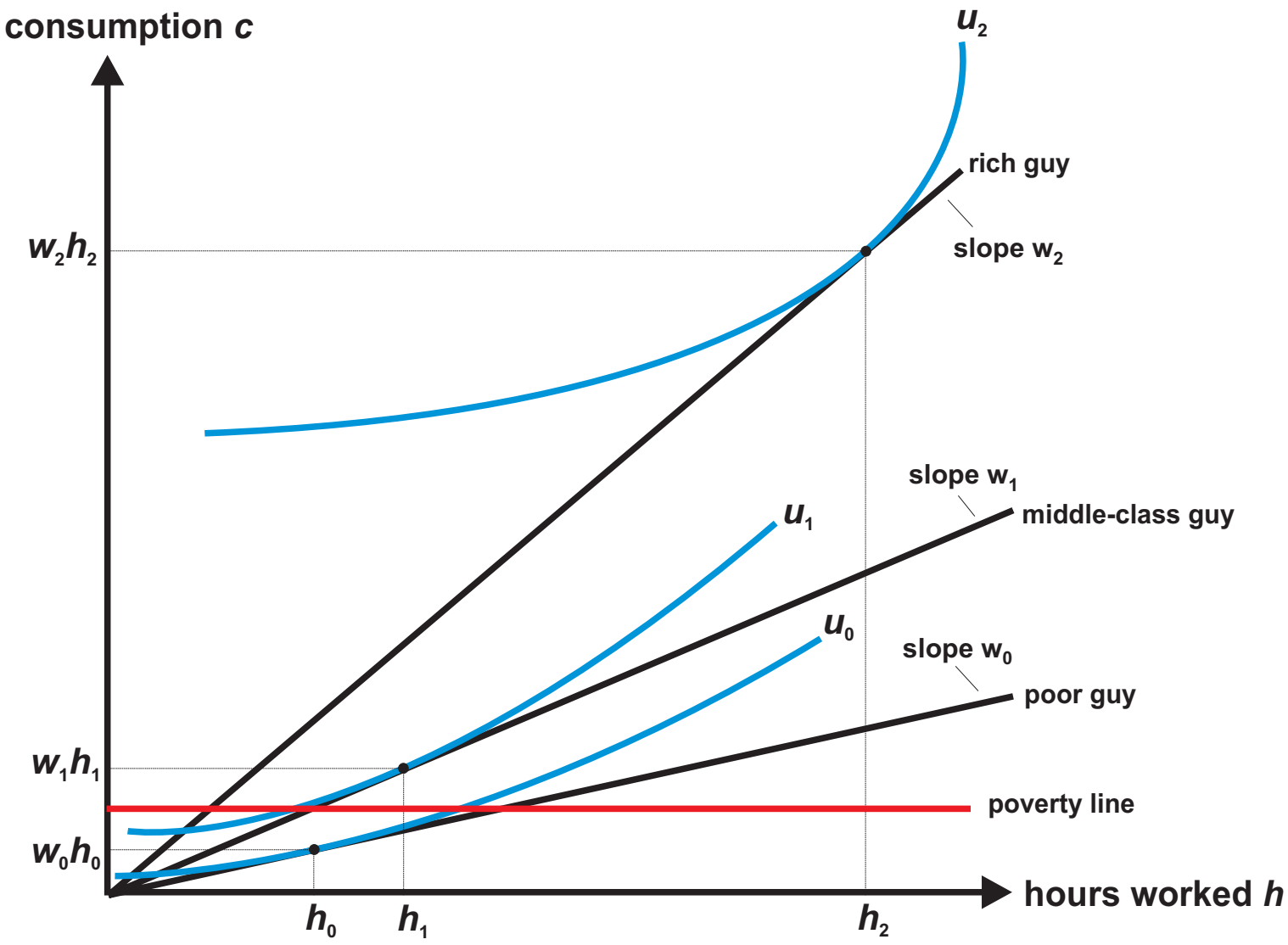
The benefit phase-out works like a tax. The benefit scheme gives an incentive to reduce labor supply so as to qualify for a higher benefit.

LABOR SUPPLY RESPONSES TO A MEANS-TESTED WELFARE PROGRAM

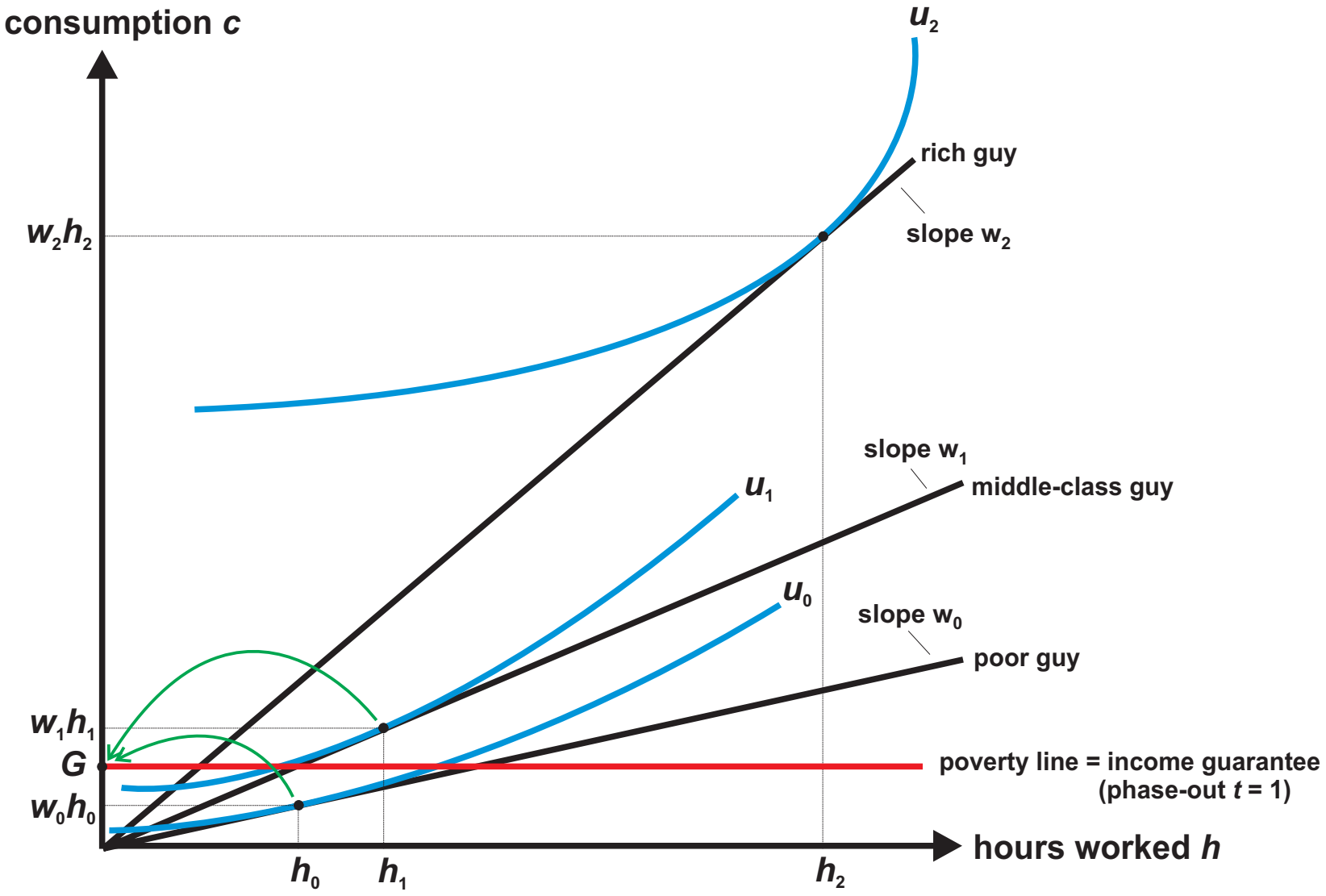
LABOR SUPPLY RESPONSES TO A MEANS-TESTED WELFARE PROGRAM



LABOR SUPPLY RESPONSES TO A MEANS-TESTED WELFARE PROGRAM



LABOR SUPPLY RESPONSES TO A MEANS-TESTED WELFARE PROGRAM



OKUN'S LEAKY BUCKET AND MORAL HAZARD

The income redistribution process is like a leaky bucket: we are carrying money from the rich to the poor, but some of the money leaks out.

Two types of **moral hazard** create the leak:

1. The not-so-poor masquerade as poor to qualify for higher benefits
   $\rightarrow$ make social programs more expensive.

2. The rich masquerade as not-so-rich to avoid some of the income tax
   $\rightarrow$ reduce revenue for social programs

These revenue leaks imply that a redistributive tax-transfer scheme imposing a 1$ cost on the rich will be able to give less than $1 to the poor. The difference is the moral hazard cost.

THE INFORMATIONAL CONSTRAINT

A **first-best** benefit scheme would be based on **earnings capacity / ability** (w) which is immutable for the individual.

But w is known only to the individual—the gov't observes earnings wh (**asymmetric information**).

Because earnings is a choice variable, an earnings-based welfare program induces high-ability individuals to reduce earnings and masquerade as low-ability individuals. Hence, the program is only **second-best**.

How can we better target truly low-ability individuals given the informational constraint?

INCREASING TARGET EFFICIENCY

1. Move from earnings-based to categorical programs (**tagging**). [Akerlof, AER 1978].

2. Move from cash to **in-kind programs** and **goods subsidies**. [Nichols & Zeckhauser, AER 1982].

3. Use **ordeal mechanisms** (pure deadweight costs on recipients). [Nichols & Zeckhauser, AER 1982; Besley & Coate, AER 1992].

TAGGING

If we can identify individual characteristics that are

1. **observable** to the government
2. **negatively correlated** with earnings capacity
3. **immutable** for the individual (unresponsive to incentives)

then targeting benefits to such characteristics is optimal.

Criteria 1 makes this form of targeting feasible, criteria 2 ensures that it redistributes from high- to low-ability, and criteria 3 ensures that there is no moral hazard associated with this redistribution.

AN AKERLOF-TYPE MODEL

A population consists of individuals with identical preferences and heterogeneous abilities (= wage rates) w.

A share γ of the population—group A—are (visibly) disabled and have $w = 0$.

The remaining share—group B—are not (visibly) disabled. This group is distributed on $(\underline{w}, \bar{w})$ according to density $f(w)$.

The gov't cannot observe the ws directly, but it can observe type-A disability.

AN AKERLOF-TYPE MODEL

The gov't operates a Negative Income Tax (NIT) system giving a lump sum transfer G financed by a constant marginal tax rate t. The transfer can potentially be targeted according to disability, i.e. $G = \{G_A, G_B\}$.

We assume a quasi-linear utility function

$$u = (1 - t)\, wh + G - d\,(h)\,,$$

where h is hours worked, and disutility of working $d\,(h)$ is an increasing, convex function ($d' > 0$, $d'' > 0$) normalized so that $d\,(0) = 0$.

INDIVIDUAL OPTIMIZATION

Optimal labor supply is characterized by

$$d'(h) = (1-t)\,w \quad \text{if} \;\; d'(0) \leq (1-t)\,w; \quad \text{otherwise } h = 0.$$

Group A have $w = 0$ and hence choose the corner solution $h_A = 0$ giving indirect utility $v_A = G_A$.

Group B choose $h_B = h_B((1-t)\,w) \geq 0$ giving indirect utility $v_B = (1-t)\,wh_B + G_B - d(h_B)$.

GOVERNMENT OPTIMIZATION

A system with $G_A = G_B$ is universal, whereas a system with $G_A \neq G_B$ is categorical (tagging).

The gov't chooses G_A, G_B and t to maximize social welfare

$$W = \gamma \Psi (v_A) + (1 - \gamma) \int_{\underline{w}}^{\bar{w}} \Psi (v_B) f (w) \, dw$$

subject to a gov't budget constraint

$$(1 - \gamma) \int_{\underline{w}}^{\bar{w}} twh_B f (w) \, dw - \gamma G_A - (1 - \gamma) G_B = 0.$$

The Lagrange multiplier associated with the gov't budget is λ.

OPTIMAL TRANSFER POLICY

The FOC for G_A is given by

$$\gamma\Psi'(G_A) - \gamma\lambda = 0.$$

The FOC for G_B is given by

$$(1-\gamma)\int_{\underline{w}}^{\bar{w}} \Psi'(v_B)\, f(w)\, dw - (1-\gamma)\,\lambda = 0.$$

Hence, the optimal policy satisfies

$$\Psi'(G_A) = \int_{\underline{w}}^{\bar{w}} \Psi'(v_B)\, f(w)\, dw.$$

THE OPTIMALITY OF TAGGING

We have $v_B > G_B$ for those with $h_B > 0$ (revealed preference). Since $\Psi(.)$ is strictly concave, we then have $\Psi'(v_B) < \Psi'(G_B)$. (For those with $h_B = 0$, we have $v_B = G_B$ and $\Psi'(v_B) = \Psi'(G_B)$). Then,

$$\Psi'(G_A) = \int_{\underline{w}}^{\bar{w}} \Psi'(v_B) f(w) dw < \Psi'(G_B).$$

This inequality can only be satisfied by having $G_A > G_B$.

Universal benefits ($G_A = G_B$) are dominated by categorical benefits to the disabled ($G_A > G_B$).

THE OPTIMAL MARGINAL TAX RATE

The FOC for t can be written as

$$\frac{t}{1-t} = -\frac{1}{\lambda} \cdot \frac{cov\left(\Psi'\left(v_B\right), wh\right)}{\varepsilon \cdot E\left[wh|B\right]} > 0,$$

where $E\left[wh|B\right]$ denotes average earnings in group B and ε is the labor supply elasticity.

The covariance $cov\left(\Psi'\left(v_B\right), wh\right)$ is always negative, since wh is increasing in w, v_B is increasing in w, and $\Psi\left(.\right)$ is concave.

The numerator reflects preferences for equality. The denominator reflects the concern for efficiency.

TAGGING IN PRACTICE

We have a theoretical case for targeting benefits according to individuals characteristics that are (1) observable, (2) negatively correlated with earnings capacity, and (3) immutable.

But are there characteristics satisfying all three criteria in practice?

Potential candidates: (a) blindness and disability, (b) single motherhood.

Single motherhood is widely used as a tagging device in many countries. It satisfies 1 and 2, but does it satisfy 3? It has been accused by conservatives of destroying the stable family.

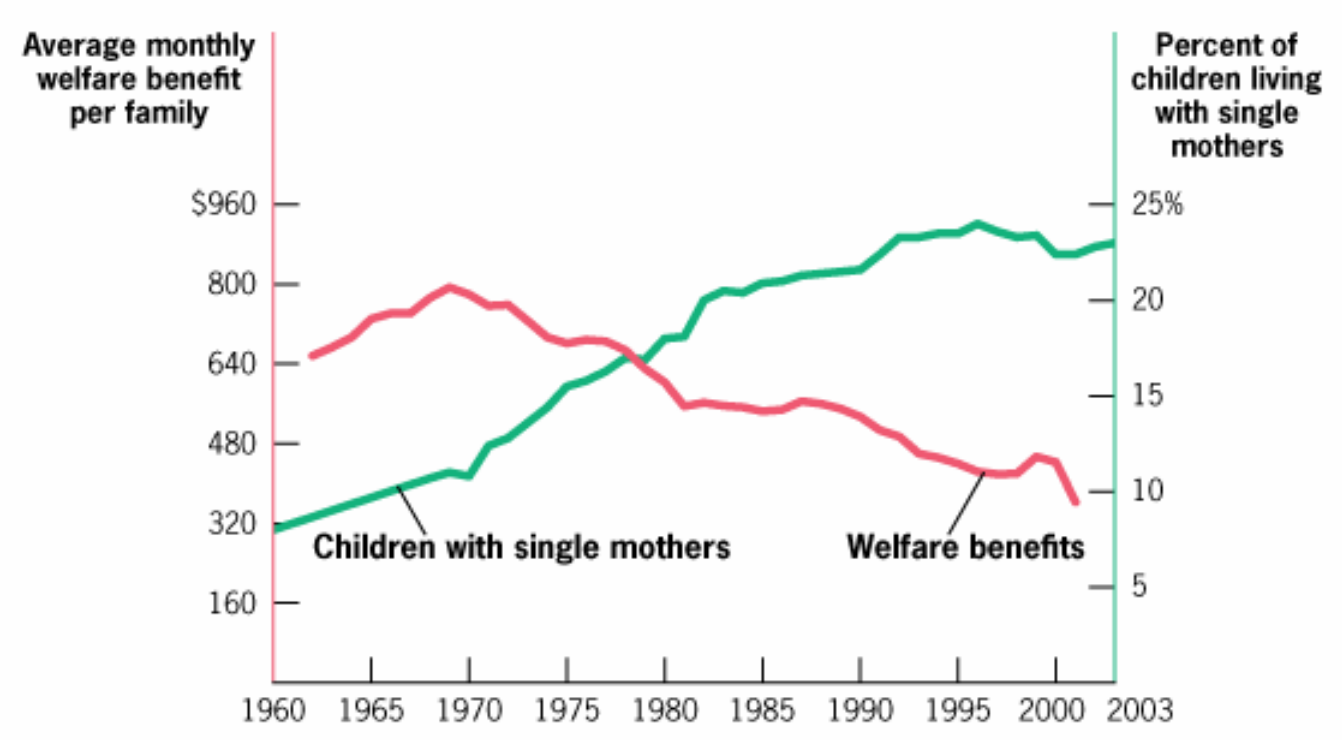
IS SINGLE MOTHERHOOD CAUSED BY WELFARE INCENTIVES?

Evidence suggests that the effect is either very small or non-existent:

1. U.S. **time series evidence** shows that single motherhood has been increasing since the 70s, whereas welfare benefits have been declining. But it is difficult to draw causal interpretations from time series.

2. U.S. **state/time variation** in welfare benefits and single motherhood. Single motherhood does not grow (significantly) more in states that are raising benefits relative to others (Moffitt, JEL 1992; Blank, JEL 2002).

WELFARE BENEFITS AND SINGLE MOTHERHOOD OVER TIME IN THE UNITED STATES



Source: Gruber (2007)

IS DISABILITY A GOOD TAGGING DEVICE?

Classification errors in social insurance programs:

TYPE I ERROR	a truly eligible individual applies for benefits but is rejected
TYPE II ERROR	a truly ineligible individual applies for benefits and is accepted (moral hazard)

Benítez-Silva, Buchinsky, Rust (NBER 2004): for the U.S. DI program, the type II error rate is 22% and the type I error rate is 62%.

These numbers suggest that neither the observability nor the unchangeability criteria are completely satisfied for DI.

NICHOLS & ZECKHAUSER (1982): PUTTING RESTRICTIONS ON RECIPIENTS

The idea of tagging is to use observables to identify the least able. An alternative approach is to set up a system inducing **self-revelation**.

Nichols and Zeckhauser (1982) analyze ways of inducing self-revelation through restrictions on recipients.

Restrictions can be placed on (a) the choice of income, (b) the choice of consumption, or (c) the allocation of time.

The general theme is that such restrictions can improve target efficiency if they are more costly to imposters than to intended recipients.

A SIMPLE MODEL

Two individuals: high-ability w_H and low-ability w_L. One common utility function $u(c_i, h_i)$ for $i = L, H$. Consumption c_i equals earned income $w_i h_i \equiv y_i$ plus a net transfer.

The gov't has a concave social welfare function and wishes to redistribute from H to L, but it cannot observe abilities.

The tax-transfer schedule has to be based on income and takes the following form: if income is less than $\bar{y}$, receive transfer T. If income is greater than $\bar{y}$, pay tax T.

INCENTIVE COMPATIBILITY

If the scheme is too generous, H masquerades as L by choosing income $\bar{y}$. In this case, there would be no redistribution, only inefficient choices.

The optimal tax-transfer scheme must satisfy incentive compatibility

$$u\left(y_H^* - T, \frac{y_H^*}{w_H}\right) \geq u\left(\bar{y} + T, \frac{\bar{y}}{w_H}\right) \tag{IC}$$

where y_H^* is the optimal choice of H conditional on not masquerading.

The optimal tax system redistributes as much as possible, while ensuring that (IC) is satisfied.

OPTIMAL TRANSFER ELIGIBILITY

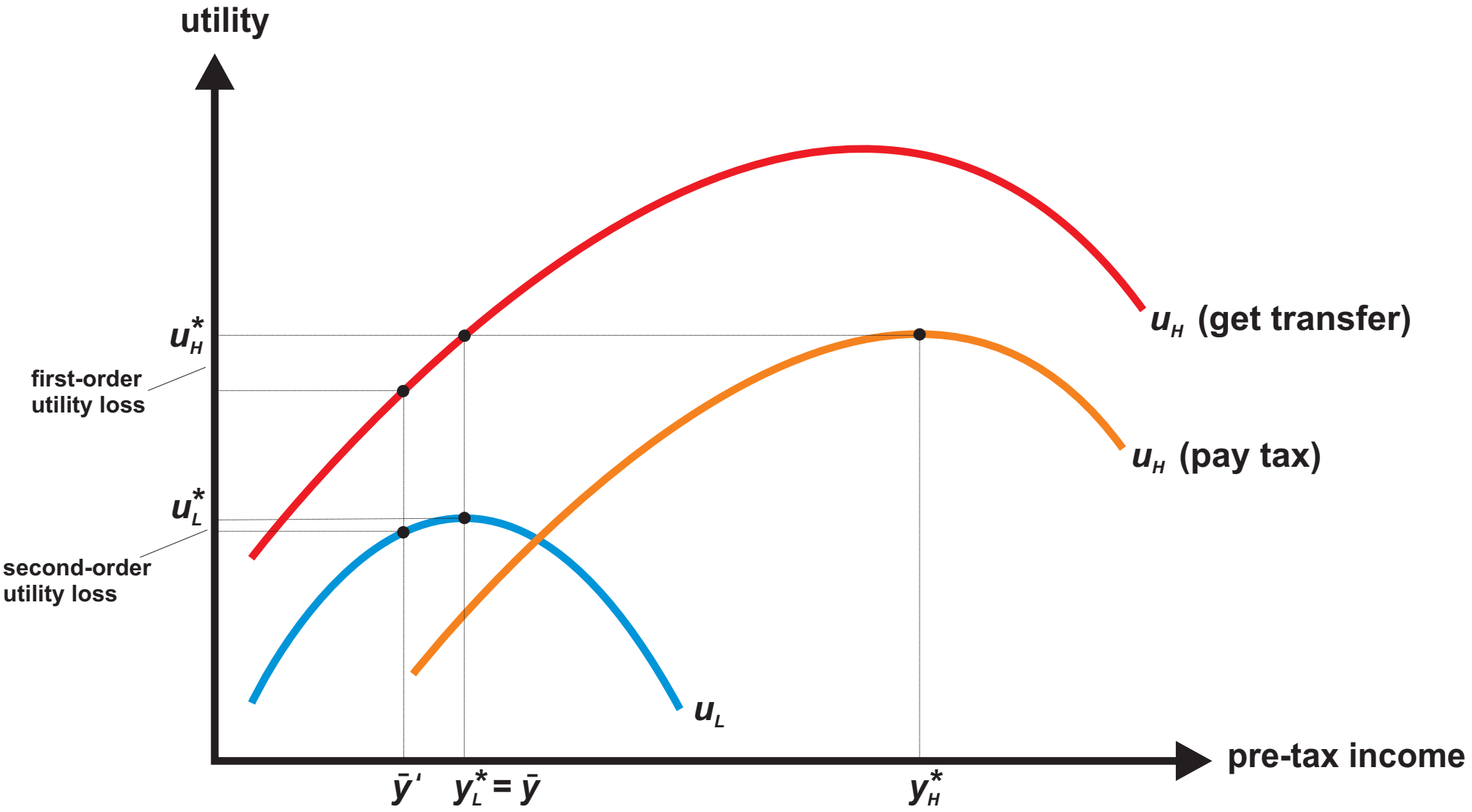
A natural solution is $\bar{y} = y_L^*$. This reduces $\bar{y}$ as far below y_H^* as possible (to prevent moral hazard) without hurting L.

But this is not optimal because, from (IC), if we reduce $\bar{y}$ further, we relax the incentive compatibility constraint and can increase T.

As we move $\bar{y}$ slightly below y_L^*, the cost to L is second-order but the cost to H (if he masquerades) is first-order. Room to increase T (a first-order gain for L) without H masquerading.

This income restriction satisfies some productive efficiency to increase target efficiency.

THE OPTIMALITY OF RESTRICTING RECIPIENTS' INCOME



IN-KIND BENEFITS AND GOODS SUBSIDIES

Can target efficiency be improved further by restricting other choices than income? What about restricting consumption choices by linking redistribution to specific goods?

There is a strong a priori argument against doing this: consumer sovereignty. The utility gain from receiving a bundle of goods is never higher than the cash value of those goods.

But, although this argument applies to each recipient *individually*, it may not apply for a program *as a whole*.

EXTENDING THE MODEL

Allow for two consumption goods, x and z, the prices of which are p_x and p_z.

The utility function is $u(x_i, z_i, h_i, w_i)$ for $i = L, H$. We allow for ability to impact utility directly.

The tax-transfer schedule imposes a tax T above income $\bar{y}$ and gives a transfer T at or below $\bar{y}$.

This system has been optimized as prescribed above, i.e. $\bar{y} < y_L^*$ and H *marginally* prefers not to masquerade.

DEMAND IF MASQUERADING

If H masquerades, his budget is $p_x x_H + p_z z_H = \bar{y} + T$ and his utility is $u\left(x_H, z_H, \frac{\bar{y}}{w_H}, w_H\right)$. Utility is maximized under the budget with respect to x_H and z_H.

Optimal demand for x is then $x_H^* = x_H\left(p_x, p_z, \bar{y} + T, \frac{\bar{y}}{w_H}, w_H\right)$.

For the low-ability individual, we have $x_L^* = x_L\left(p_x, p_z, \bar{y} + T, \frac{\bar{y}}{w_L}, w_L\right)$.

Demand for x (and z) may depend on ability either through the argument $\frac{\bar{y}}{w_i}$ (substitutability with leisure) or through the argument w_i (direct preference-dependence on ability).

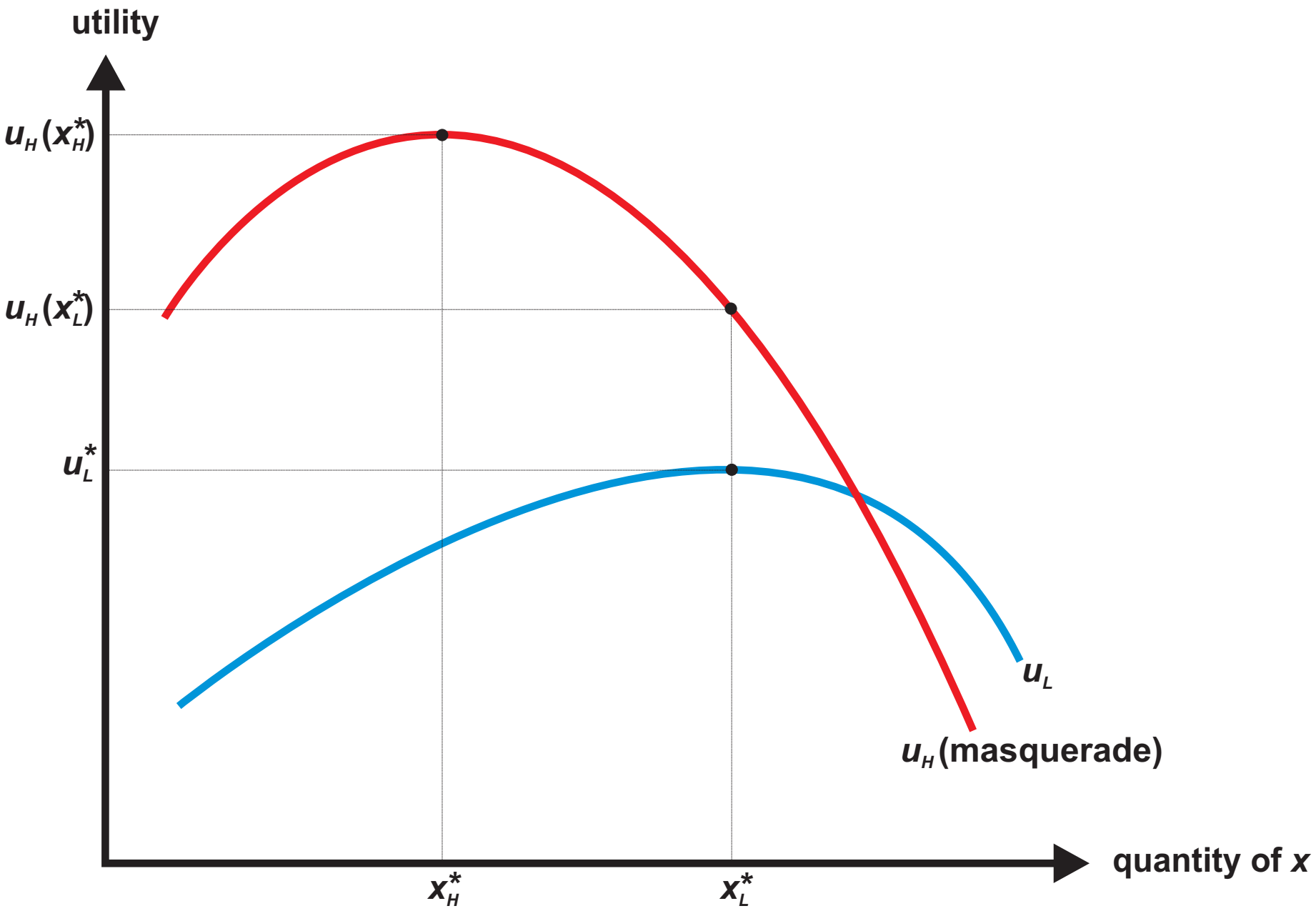
OPTIMALITY OF IN-KIND TRANSFER

If $x_L^* > x_H^*$, so that demand varies with ability *conditional on income*, then x is an **indicator good**.

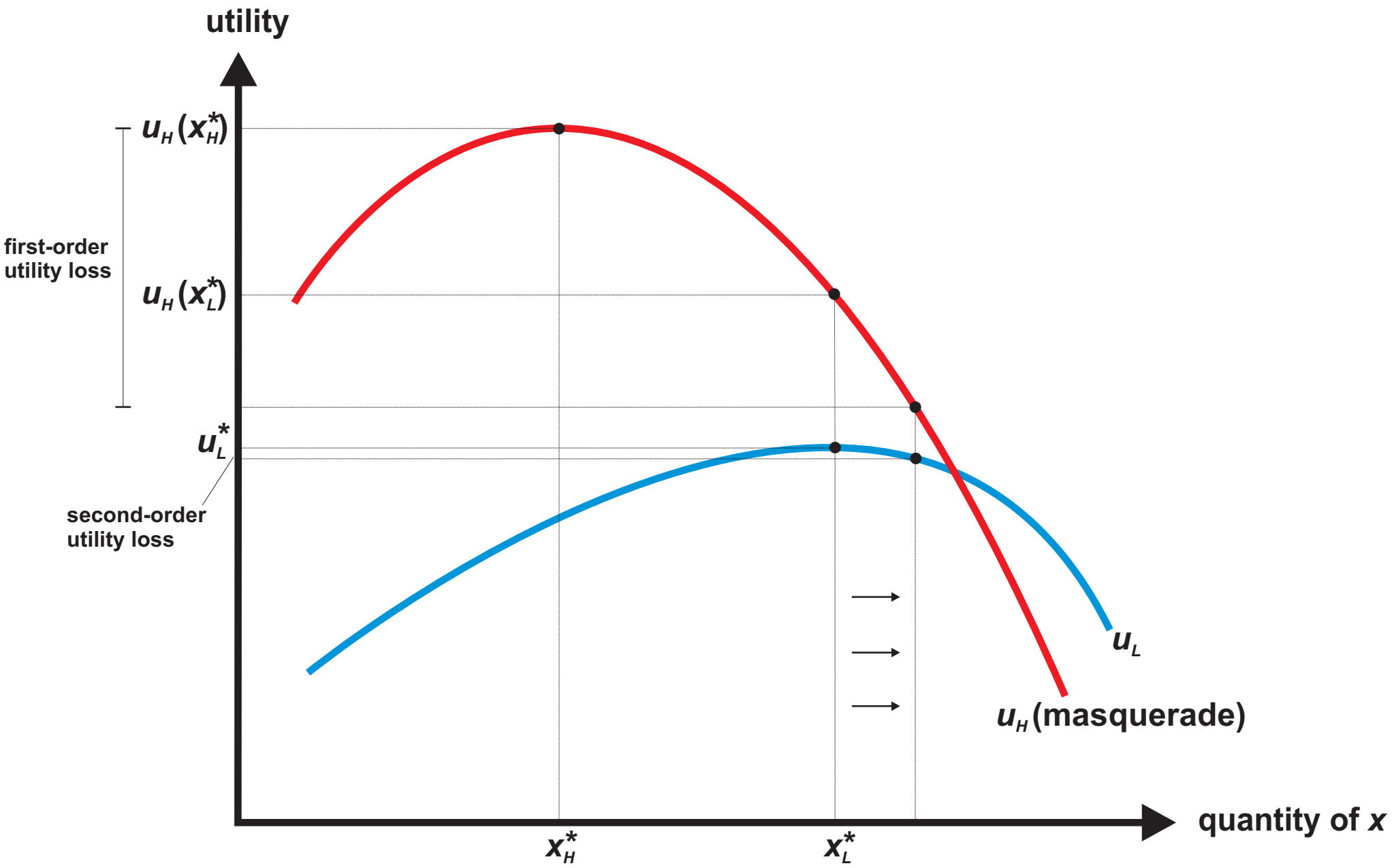
If such a good exists, we can improve target efficiency by linking redistribution to this good. The argument is a first-order vs second-order argument as before.

If we convert part of the cash transfer into a transfer of x at a level slightly above x_L^*, we hurt L a little bit but we hurt H more (if he masquerades). Room to increase T (a first-order gain for L) without H masquerading.

THE OPTIMALITY OF IN-KIND TRANSFERS



THE OPTIMALITY OF IN-KIND TRANSFERS



WHEN CAN IN-KIND BENEFITS AND SUBSIDIES IMPROVE TARGET EFFICIENCY?

1. If differences in demand across rich and poor are due only to income effects (luxuries versus necessities), in-kind transfers cannot improve target efficiency.

2. If, *conditional on income*, demand varies with ability, in-kind transfers can improve target efficiency.

Potential candidates: housing projects, medical care.

ORDEALS

An ordeal is a pure deadweight cost on recipients.

Examples of ordeals are:

1. Work and training requirements (workfare). Used in U.S. TANF.
2. Tedious and complex administrative procedures. Common in many social programs (Currie, 2004).
3. Demeaning qualification tests.
4. In-hospital requirements in health insurance. Used in U.S. Medicaid.

ORDEALS, CONT'D

Ordeals can serve a screening function if (i) the utility gain from transfers is lower for the non-deserving, and/or if (ii) the utility cost of the ordeal is higher for the non-deserving.

Two types of recipients in a cash welfare program: low-ability individuals and high-ability but "lazy" individuals.

Introduce a time loss in the program, e.g. work requirements or waiting lines. The time loss will be more costly to the high-ability/lazy individuals.

THE COST OF ORDEALS: INCOMPLETE TAKE UP

A problem with many social programs, especially in the U.S., is that not all eligibles take up benefits. Take-up rates can be far lower than 100% in some programs.

Three possible explanations: (*i*) welfare stigma, (*ii*) imperfect information, (*iii*) transaction costs associated with take up.

The third explanation is related to the presence of ordeals. Currie (2004) suggests that transaction costs due to ordeals may be an empirically important factor for incomplete take up.

U.S. WELFARE REFORM

The Personal Responsibility and Work Opportunity Reconciliation Act of 1996 (PRWORA), or the welfare reform law:

1. Cash welfare was changed from matching grants to lump-sum **block grants** from the federal gov't to state gov'ts.

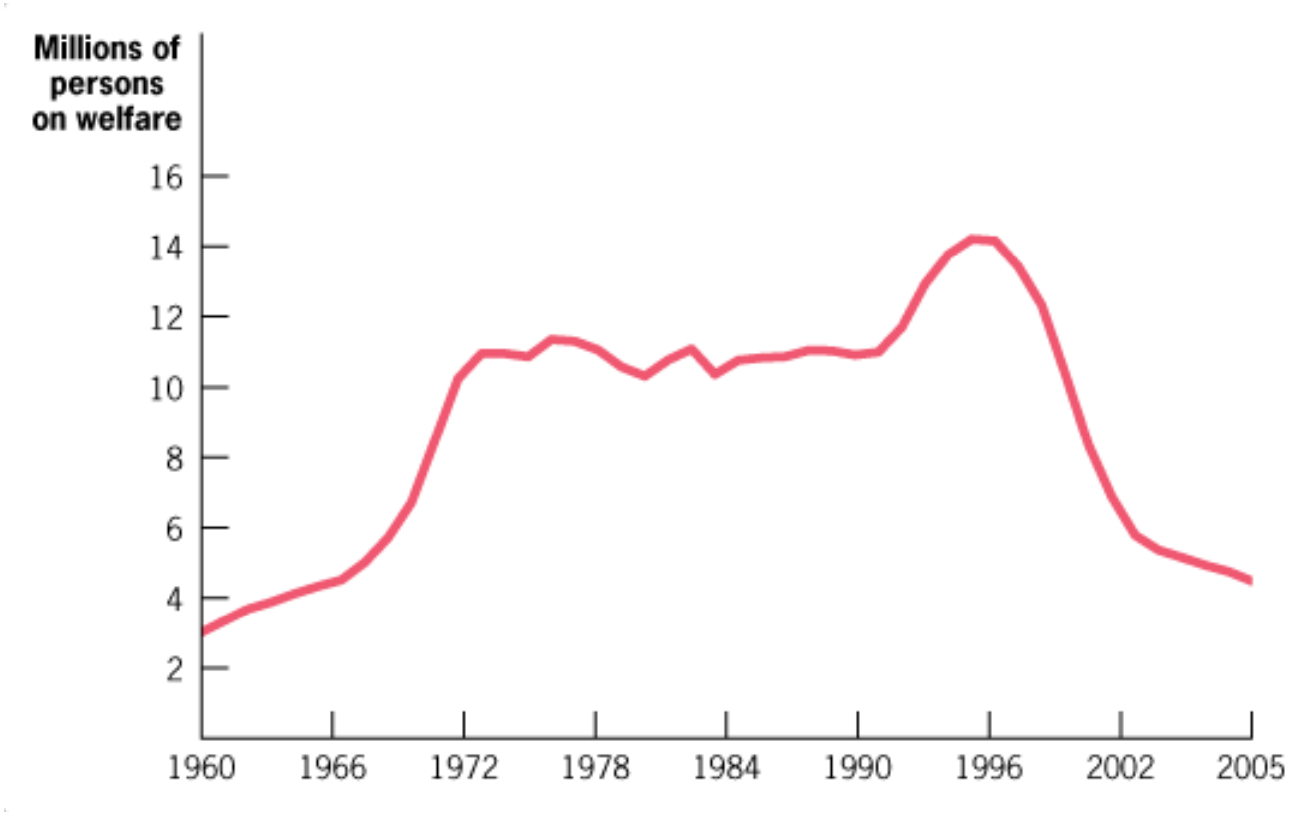
2. States were given **greater discretion** in designing programs.

3. **Work requirements** were imposed on welfare recipients.

4. **Time limits** were imposed on welfare recipients.

5. New efforts to **limit unwed motherhood** were introduced.

WELFARE CASELOADS OVER TIME IN THE U.S.



Source: Gruber (2007)

THE IMPACT OF WELFARE REFORM: TIME SERIES

Over the 1990s, caseloads fell, labor force participation soared, and the poverty rate declined. But time series evidence is not compelling:

1. The caseload fall began in 1994 (and the participation increase even earlier), i.e. *before* the welfare reform took effect.

2. The 1990s saw one of the largest economic booms in the U.S. history.

3. The welfare reform was passed shortly after two tax reform acts (1990 and 1993) which also aimed at bringing low-income single mothers into work through expansions of the EITC.

It is very difficult to separate the effects from welfare reform, tax reform, and the business cycle.

QUASI-EXPERIMENTAL ESTIMATIONS

Use the reform as a natural experiment by comparing a treatment group T affected by the reform to a control group C not (or less) affected, but experiencing similar macro economic shocks. Two possibilities:

1. Compare single mothers (T) to either married mothers (C_1) or single women without children (C_2).

2. Use the fact that some states (but not others) were allowed to experiment with certain aspects of the reform before the national law was in place.

These approaches suggest that the welfare reform did have positive effects, but was far from the whole story (see Blank, JEL 2002).