

Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods

Raj Chetty
UC-Berkeley and NBER

September 2008

MOTIVATION

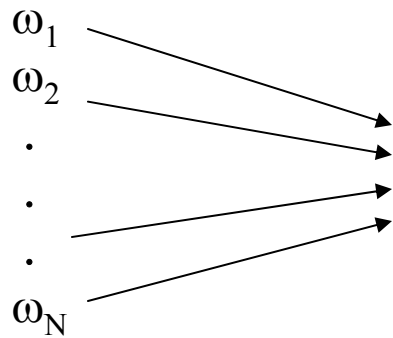
- Two competing paradigms for policy evaluation and welfare analysis: “structural” vs. “reduced-form”
- Structural approach generally involves two steps: estimate primitives of a model and then simulate effects of policies on welfare
 - Critique: difficult to identify full primitive structure without implausibly strong assumptions.
- Reduced-form: estimate statistical relationships using transparent, exogenous sources of variation for identification (“treatment effects”)
 - Critique: Estimates not useful for welfare analysis because they are not deep parameters; endogenous to policy regime (Lucas 1976, Heckman and Vytlačil 2005)

SUFFICIENT STATISTICS

- Past decade of work in public economics provides a strategy that bridges the gap between the two methods
 - Idea: Instead of primitives, identify “sufficient statistics” for welfare analysis that can be estimated using reduced-form methods
 - Any set of primitives (ω) consistent with sufficient statistics (β) generates the same value of welfare gain (dW/dt)

THE SUFFICIENT STATISTIC APPROACH

**Structural
Primitives**



ω =prefs.,
constraints

ω not uniquely
identified

**Sufficient
Statistics**



$$\beta = f(\omega, t)$$
$$y = \beta_1 X_1 + \beta_2 X_2 + \varepsilon$$

β identified using
program evaluation

**Welfare
Change**

$$\frac{dW}{dt}(t)$$

dW/dt used for
policy analysis

Intellectual History

- Idea that it is adequate to estimate “sufficient statistics” to answer some questions dates to Marschak (1954) and Koopmans (1954)
 - But applied to a wide range of policy questions only in past decade
 - 1950-70s – simple structural models fit to macro and micro data
 - 1980s: concerns about identification of non-linear structural models with heterogeneity (e.g. Ashenfelter 1978, LaLonde 1985)
 - Reduced-form, quasi-experimental methods (e.g. Angrist 1990, Card and Krueger 1995; Imbens and Wooldridge 2008)
 - 1990s: Large body of “program evaluation” estimates developed
- Most recent literature integrates program evaluation estimates with structural models to make statements about welfare

OBJECTIVES OF THIS TALK

1. Codify main steps and concepts of sufficient statistic approach
2. Review a set of applications, showing how several independent papers are variants on this theme
3. Discuss benefits and costs of this strategy vs. structural methods

Advantages: identify fewer parameters, weaker modelling assumptions

Disadvantages: only local welfare analysis, “black box” (no evaluation of model)

→ Sufficient statistic methods provide a useful complement to (rather than a substitute for) structural methods in future work

OUTLINE

1. Harberger (1964) and Extensions
2. General Framework
3. Application 1: Taxation
4. Application 2: Social Insurance
5. Application 3: Behavioral Models
6. Summary of Advantages/Disadvantages

HARBERGER (1964)

- Precursor to modern sufficient statistic literature: Harberger's partial-equilibrium analysis of deadweight cost of taxation
- Simple model that is useful to build intuition about more sophisticated applications discussed later
- Objective: calculate excess burden (EB) of tax. How much extra revenue could be raised by lump sum taxation, keeping utility constant?
- To simplify exposition, ignore income effects (quasilinear utility)

Setup

- Individual endowed with Z units of good y (numeraire)
- Normalize price of y to 1
- Firms convert numeraire y into J other consumption goods (x_1, \dots, x_J)
- Producing x_j units of good j requires $c_j(x_j)$ units of numeraire
 - Let $c(x) = \sum c_j(x_j)$ denote total cost of producing vector x
 - Production perfectly competitive
- Let $p = (p_1, \dots, p_J)$ denote prices of produced goods
- Government levies a unit tax t on good 1
 - Goal: measure efficiency cost of this tax (social surplus from transactions that do not occur because of tax)

- Consumers take prices as given and solve

$$\begin{aligned} & \max_{x,y} u(x_1, \dots, x_J) + y \\ & \text{s.t. } px + tx_1 + y = Z \end{aligned}$$

- Representative firm takes p as given and solve

$$\max_x px - c(x)$$

- Two problems define demand supply fns. Equilibrium:

$$x^D(p) = x^S(p)$$

- Social welfare: sum of utility, profits, and tax revenue

$$\begin{aligned} W(t) &= \left\{ \max_x u(x) + Z - tx_1 - px \right\} + \left\{ \max_x px - c(x) \right\} + tx_1 \\ &= \left\{ \max_x u(x) + Z - tx_1 - c(x) \right\} + tx_1 \end{aligned}$$

Calculation of Excess Burden

- Structural method: Estimate J good supply + demand system and recover $u(x)$ and $c(x)$
 - Ex: use Stone-Geary or AIDS and CES production functions
 - Or non parametric methods to recover preferences and technology as in Hausman (1981) and Hausman and Newey (1994)
- Econometric challenge: simultaneity.
 - Need $2J$ instruments to identify supply and demand in J markets.

Harberger Approach

- Private sector choices are made to maximize term in curly brackets (private surplus) in social welfare function

$$W(t) = \{ \max_x u(x) + Z - tx_1 - c(x) \} + tx_1$$

- Envelope conditions for (x_1, \dots, x_J) yield simple formula:

$$\frac{dW}{dt} = -x_1 + x_1 + t \frac{dx_1}{dt} = t \frac{dx_1}{dt}$$

- Tax induces changes in x and p , but these responses do not have a first-order effect on private surplus b/c of optimization
 - Loss in surplus determined purely by difference between WTP and cost of good x_1 (triangle between demand and supply)
- dx_1/dt is a “sufficient statistic” for calculating dW/dt .
- Do not need to identify primitives, simplifying identification.

Heterogeneity

- Benefit of sufficient statistic approach is particularly evident in a model that permits heterogeneity across individuals
- N agents with wealth Z_i and utility functions

$$u^i(x^i) + y$$

- Social welfare:

$$W(t) = \left\{ \sum_{i=1}^N \max_{x^i} [u^i(x^i) + Z^i - tx_1^i] - c(x) \right\} + t \sum_{i=1}^N x_1^i$$

- Structural method requires estimation of demand systems for all agents
- Sufficient statistic formula is *unchanged* – still need only slope of aggregate demand dx_1/dt

$$\frac{dW}{dt} = - \sum_{i=1}^N x_1^i + \sum_{i=1}^N x_1^i + t \frac{d \sum_{i=1}^N x_1^i}{dt} = t \frac{dx_1}{dt}$$

Discrete Choice

- Now suppose individuals can choose only one of the J products
 - E.g. car models, modes of transportation, or neighborhoods
- Each product j characterized by a vector of K observable attributes

$$x_j = (x_{1j}, \dots, x_{Kj})$$

and an unobservable attribute ζ_j

- Agent i 's utility from choice j is

$$u_{ij} = v_{ij} + \varepsilon_{ij}$$

with $v_{ij} = Z^i - p_j + \zeta_j + \phi^i(x_j)$

- Let P_{ij} denote probability i chooses product j , P_j total expected demand for product j , and $c_j(P_j)$ cost of production

- Assume ε_{ij} has a type 1 extreme value distribution (mixed logit)
- Then probability individual i chooses product j is

$$P_{ij} = \frac{\exp(v_{ij})}{\sum_j \exp(v_{ij})}$$

and consumer i 's expected surplus is

$$S_i(p_1, \dots, p_J) = E \max(u_{i1}, \dots, u_{iJ}) = \log(\sum_j \exp v_{ij})$$

- Aggregating over consumers and including producer profits gives

$$W = \sum_i \log(\sum_j \exp(v_{ij})) + pP - c(P)$$

- Structural approach to policy analysis: identify ϕ_i and $c(P)$ using methods e.g. in Berry (1994) or BLP (1995)
- Sufficient statistic: two examples
 1. Tax on good 1. Then easy to establish that

$$\frac{dW}{dt} = t \frac{dP_1}{dt}$$

2. Tax on all products in the market.

$$\frac{dW}{d\tau} = \tau \sum_j P_j \frac{dP_j}{d\tau} = \tau \frac{dE_P}{d\tau}$$

where E_p = total expenditure on products in the market

- Do not need to estimate substitution patterns within market
- Microeconomic demands not smooth but expected welfare is \rightarrow use similar envelope conditions

GENERAL FRAMEWORK

- Modern sufficient statistic approach builds on Harberger's idea
- First present a general framework that nests papers in this literature
 - Explains why identification of a few sufficient statistics is adequate to answer many questions
 - Provides a “recipe” for deriving such formulas in future work
- Abstractly, many government (price intervention) policies amount to levying a tax t and paying a transfer $T(t)$
 - Redistributive taxation: transfer to another agent
 - Social insurance: transfer between states
 - Excess burden: transfer used to finance lump sum grant
- Develop a rubric to calculate dW/dt using sufficient statistics

Step 1: Specification of General Structure of Model

- Government levies tax on x_1 and pays transfer $T(t)$ in units of x_N
- Utility: $U(x_1, \dots, x_N)$
- Constraints: $G_1(x, t, T), \dots, G_M(x, t, T)$
- Private sector takes t and T as given and solves

$$\max U(x_1, \dots, x_J)$$

$$\text{s.t. } G_1(x, t, T) = 0, \dots, G_M(x, t, T) = 0$$

- Note that this nests case of competitive production, with $U(x) = u(x) - c(x)$ because decentralized eq. maxes total surplus

Step 2: Multiplier Representation for dW/dt

- Social welfare:

$$W(t) = \max_{x_1, \dots, x_J} U(x_1, \dots, x_J) + \sum_{m=1}^M \lambda_m G_m(x, t, T)$$

- Using envelope thm., welfare gain from increasing t is:

$$\frac{dW}{dt} = \sum_{m=1}^M \lambda_m \left\{ \frac{\partial G_m}{\partial T} \frac{dT}{dt} + \frac{\partial G_m}{\partial t} \right\}$$

- Key unknowns: multipliers λ_m
- Other parameters known mechanically from constraints

Step 3: Map Multipliers to Marginal Utilities

- Multipliers can typically be recovered by exploiting first-order conditions
- Private sector optimization requires

$$u'(x_i) = -\sum_{m=1}^M \lambda_m \frac{\partial G_m}{\partial x_i}$$

- Under a mild technical condition on structure of constraints, we obtain

$$\frac{dW}{dt} = k_T \frac{dT}{dt} u'(x_N) - k_t u'(x_1)$$

where k_t , k_T are known functions of equilibrium quantities

- Problem reduced to recovery of only **two marginal utilities** rather than full structure (U, G)

Step 4: Recover Marginal Utilities from Observed Choices

- Final step in deriving formula is to back out the two marginal utilities from empirically observable choices
 - No canned procedure; applications illustrate various methods
 - Usual trick: marginal utilities appear in first-order-conditions for choices → back out from high-level elasticities of behavior
 - Can generally express formulas in terms of parameters that can be estimated using program-evaluation methods
- Harberger example: $u'(x_N) = 1$ given quasilinearity. To recover $u'(x_1)$, use first order condition

$$u'(x_1) = p_1 + t$$

Step 5: Empirical Implementation

$$\frac{dW}{dt}(t) = f\left(t, \frac{dx_1}{dt}, \frac{dx_1}{dZ}, \frac{dx_2}{dt}, \frac{dx_2}{dZ}, \dots\right)$$

- Problem in empirical implementation: program evaluation studies estimate $\Delta x_1/\Delta t$ not dx_1/dt where $\Delta t = t_1 - t_0$
- This can give information about change in welfare from t_0 to t_1
- Three options:
 1. Bound $W(t_1) - W(t_0)$
 2. Take a linear approximation to demands to calculate $W(t_1) - W(t_0)$
 3. Estimate $x(t, Z)$ non-parametrically if data and variation permit
 - Analogous to estimating full distribution of MTEs (Heckman and Vytlacil 2005)

Step 6: Structural Evaluation

- Find a vector of ω that matches sufficient statistics, assess plausibility
- Run three types of simulations:
 1. Compare exact simulated welfare gain with that implied by sufficient statistic formula given approximations made in step 5
 2. Simulate how sufficient statistics vary with t ; if highly non-linear, take into account in empirics
 3. Use structural model to guide out-of-sample extrapolations and solve for globally optimal policy.
- This step is often not implemented, but is critical to make reliable out-of-sample predictions using sufficient statistic methods
 - Problems with sufficient statistic approach evident when structural primitives are not assessed.

APPLICATION 1: TAXATION

Feldstein, Martin. "The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act," *Journal of Political Economy*, 1995.

Diamond, Peter. "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates," *American Economic Review*, 1998.

Saez, Emmanuel. "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 2001.

Other Sufficient Statistic References:

Piketty (Revue Francaise 1997)

Gruber and Saez (JPubE 2002)

Goulder and Williams (JPE 2003)

Chetty (AEJ-EP forthcoming)

Feldstein (1995, 1999)

- Following Harberger, large literature in labor estimated effect of taxes on hours worked to assess efficiency costs of taxation
- Feldstein observed that labor supply involves multiple dimensions, not just choice of hours: training, effort, occupation
- Taxes also induce inefficient avoidance/evasion behavior
- Structural approach: account for each of the potential responses to taxation separately and then aggregate
- Feldstein's solution: elasticity of taxable income with respect to taxes is a sufficient statistic for calculating deadweight loss

Setup

- Government levies linear tax t on reported taxable income
- Agent makes N labor supply choices: l_1, \dots, l_N
- Each choice l_i has disutility $\psi_i(l_i)$ and wage w_i
- Agents can shelter e of income from taxation by paying cost $g(e)$
- Taxable Income (TI) is

$$TI = \sum_{i=1}^N w_i l_i - e$$

- Consumption is given by taxed income plus untaxed income

$$x_N = (1 - t)TI + e$$

- Social welfare:

$$W(t) = \{(1 - t)TI + e - g(e) - \sum_{i=1}^N \psi_i(l_i)\} + t \cdot TI$$

- Totally differentiating $W(t)$ gives

$$\frac{dW}{dt} = \frac{dTI}{dt} + \frac{de}{dt} (1 - g'(e)) - \sum_{i=1}^N \psi'_i(l_i) \frac{dl_i}{dt}$$

- Use first order conditions to measure marginal utilities:

$$g'(e) = t$$

$$\psi'_i(x_i) = (1 - t)w_i$$

- Substituting into dW/dt yields Feldstein's formula:

$$\frac{dW}{dt} = t \frac{dTI}{dt}$$

- Intuition: marginal social cost of reducing earnings through each margin is equated at optimum \rightarrow irrelevant what causes change in TI.

- Simplicity of identification in Feldstein's formula has led to a large literature estimating elasticity of taxable income
- Problem: since primitives are not estimated, assumptions never tested
- Chetty (2008) questions validity of assumption that $g'(e) = t$
 - Costs of many avoidance/evasion behaviors are transfers to other agents in the economy, not real resource costs
 - In a model that permits such transfer costs, Feldstein's formula is invalid because of externality associated with sheltering
 - Instead, EB depends on weighted average of taxable income and total earned income elasticities
- Practical importance: even though reported taxable income is highly sensitive to tax rates for rich, efficiency cost may not be large!
- A structural approach would not have run into this problem because $g(e)$ would have been identified.

Saez (2001)

- Saez characterizes optimal tax rates in Mirrlees' (1971) model using high-level sufficient statistics
 - Multiple policy instruments, continuum of heterogeneous agents
 - Levy tax $T(z)$ at income level $z \rightarrow$ net-of-tax income: $z - T(z)$
- Mirrlees characterized optimal tax rates in terms of primitives that entered complex first-order-conditions
 - Offers little intuition about key determinants of $T(z)$
 - Simulations reach variable conclusions depending on primitives

Mirrlees Model

- Individuals choose labor supply to maximize

$$u(c, l) = c - \psi(l)$$
$$\text{s.t. } c = wl - T(wl)$$

- Government chooses tax schedule $T(z)$ to maximize social welfare

$$W(T(z)) = \int_0^{\infty} \tilde{G}(u(c(w, T), wl(w, T))) dF(w)$$

subject to resource and IC constraints

$$G_1(c, z, T) = \int_0^{\infty} z(w, T) dF(w) - \int_0^{\infty} c(w, T) dF(w) - E = 0$$

$$G_2(c, z, T) = (1 - T'(z))w - \psi'(l(w)) = 0$$

- Diamond and Saez obtain following formula for optimal tax $T(z)$:

$$\frac{T(z)}{1-T(z)} = \frac{1}{\varepsilon(z)zh(z)} \int_z^{\infty} (1 - g(z'))h(z')dz'$$

- Elasticities $\varepsilon(z)$, density $h(z)$, and marginal utility $g(z)$ at each point of income distribution together determine optimal tax rate
 - Marginal social welfare weights taken as exogenous to system
- **Not** an explicit formula for optimal tax; can only be used as a test.
- To compute optimal tax system, Saez calibrates structural primitives to match three parameters that enter formula and then simulates $T(z)$
 - Optimal income tax schedule is inverse-U-shaped, with a large lump sum grant and marginal rates ranging from 50-80%.
- Illustrates power of combining sufficient statistic and structural approaches to do more than marginal welfare analysis

APPLICATION 2: SOCIAL INSURANCE

Gruber, Jonathan. “The Consumption Smoothing Benefits of Unemployment Insurance,” *American Economic Review*, 1997.

Chetty, Raj. “Moral Hazard vs. Liquidity and Optimal Unemployment Insurance,” *Journal of Political Economy*, 2008.

Shimer, Robert and Ivan Werning. “Reservation Wages and Unemployment Insurance,” *Quarterly Journal of Economics*, 2007.

Other Sufficient Statistic References:

Chetty (JPubE 2006)

Einav, Finkelstein, and Cullen (2008)

Finkelstein, Luttmer, and Notowidigdo (2008)

Chetty and Saez (2008)

Static Model of Social Insurance (Baily 1978)

- Two states: high and low (unemployed, sick ,etc.).
- Income in high state: $A + w_h$; in low state: $A + w_l$
- Consumption in high state: c_h ; in low state: c_l
- Agent can control probability of high state via effort e at cost $\psi(e)$
 - Reflects search effort, investment in health, etc.
 - Choose units so that probability of high state is $p(e)=e$.
- Imperfect private insurance: individuals can transfer $\$z$ from high state to low state via informal risksharing at cost $q(b_p)$
 - $\$1$ increase in $c_l \rightarrow \$(1-e)/e+q(b_p)$ reduction in c_h
- Social insurance: government pays a benefit b in low state financed by a tax $t(b)=b \cdot (1-e)/e$

- Social welfare:

$$W(b) = eu(A + w_h - \frac{1-e}{e}b_p - q(b_p) - t(b)) + (1-e)u(A + w_l + b_p + b) - \psi(e)$$

- Marginal welfare gain has marginal-utility representation:

$$\frac{dW}{db} = (1-e)\left\{u'(c_l) - \left(1 + \frac{\varepsilon_{1-e,b}}{e}\right)u'(c_h)\right\}$$

- To convert to money-metric, compare welfare gain of increasing insurance program and wage bill in high state:

$$M_W(b) = \frac{\frac{dW}{db}(b)/(1-e)}{\frac{dW}{dw_h}(b)/e} = \frac{u'(c_l) - u'(c_h)}{u'(c_h)} - \frac{\varepsilon_{1-e,b}}{e}$$

Sufficient Statistics

Recent literature on social insurance makes two contributions:

1. Shows that formula holds in a general class of **dynamic** models: arbitrary choices and constraints, stochastic wages, heterogeneity

- Distills analysis to two parameters; structural models often forced to assume no borrowing or private insurance

2. Recovers marginal utility gap from choice data:

$$\frac{u'(c_l) - u'(c_h)}{u'(c_h)} = f(\text{observables})$$

- Gruber (1997): consumption
- Shimer and Werning (2007): reservation wages
- Chetty (2008): liquidity and substitution effects in effort

Gruber (1997)

- Quadratic approximation to utility function yields

$$\frac{u'(c_l) - u'(c_h)}{u'(c_h)} = \gamma \frac{\Delta c}{c_h}(b)$$

- Gruber estimates a linear consumption function:

$$\frac{\Delta c}{c_h}(b) = \alpha + \beta b$$

- Plugging back into Baily's formula yields

$$\frac{dW}{db} = (\alpha + \beta b)\gamma - \frac{\varepsilon_{1-e,b}}{e}$$

- Gruber estimates $\alpha = 0.24$, $\beta = -0.28$ using consumption data from PSID and state-level changes in UI benefits in the U.S.
- With $\gamma = 1$, obtains $dW/db < 0$; extrapolating to lower benefit levels, concludes that optimal benefit near 0.
- Limitation: value of γ is highly debated; may vary with context.

Chetty (2008)

- Uses comparative statics of effort choice (e) to back out marginal utils.
- First order condition for effort:

$$\psi'(e) = u(c_h) - u(c_l)$$

- Effects of cash grant (e.g. severance pay) and higher benefit level:

$$\partial e / \partial A = \{u'(c_h) - u'(c_l)\} / \psi''(e) \leq 0$$

$$\partial e / \partial b = -u'(c_l) / \psi''(e)$$

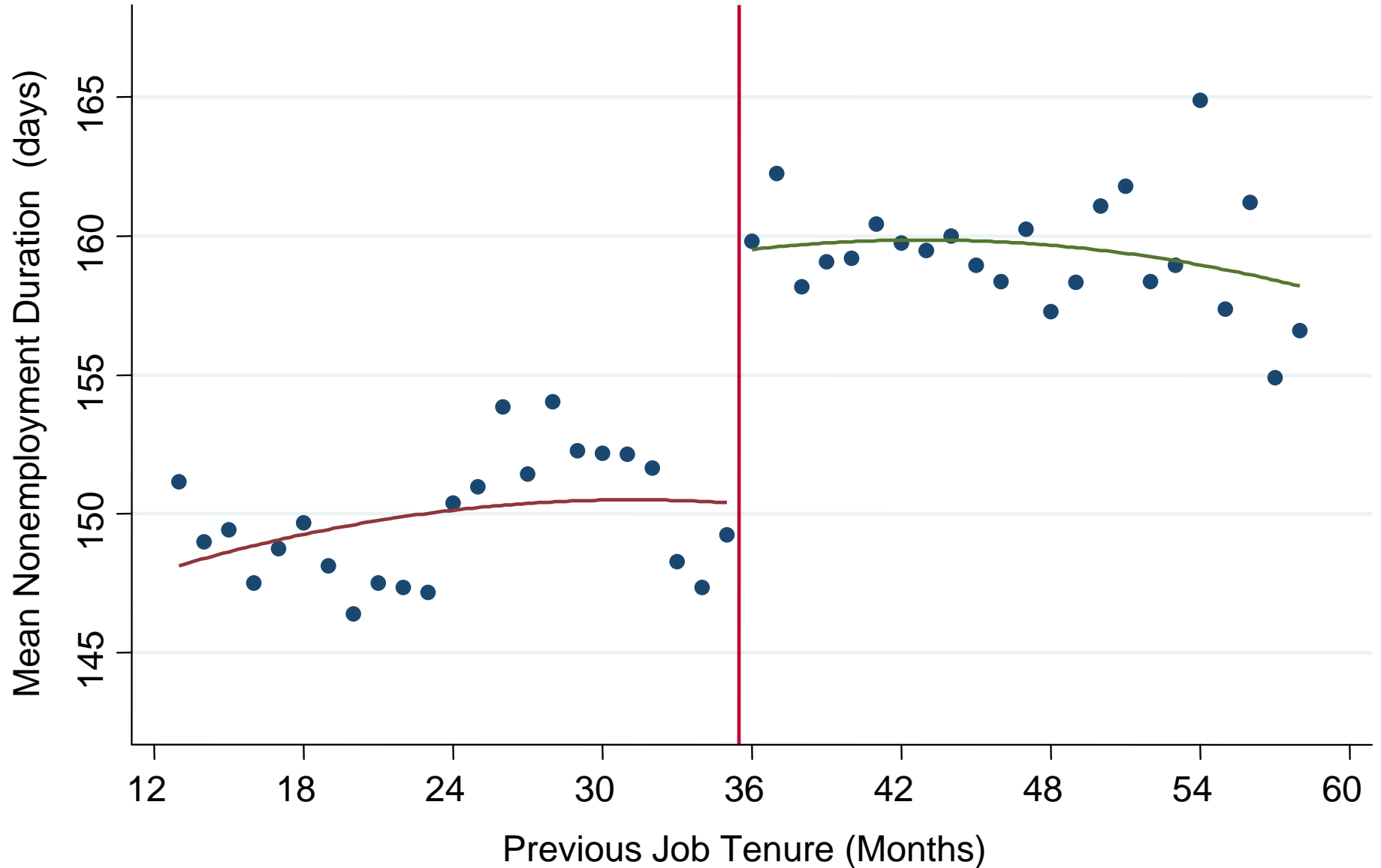
- It follows that

$$\frac{u'(c_l) - u'(c_h)}{u'(c_h)} = \frac{-\partial e / \partial A}{\partial e / \partial A - \partial e / \partial b}$$
$$\Rightarrow M_w(b) = \frac{-\partial e / \partial A}{\partial e / \partial A - \partial e / \partial b} - \frac{\varepsilon_{1-e,b}}{e}$$

- Liquidity effect (de/dA) measures completeness of private insurance; moral hazard effect (de/dw_h) measures efficiency cost of insurance.

Card, Chetty and Weber (QJE 2007)

Effect of Severance Pay on Nonemployment Durations in Austria

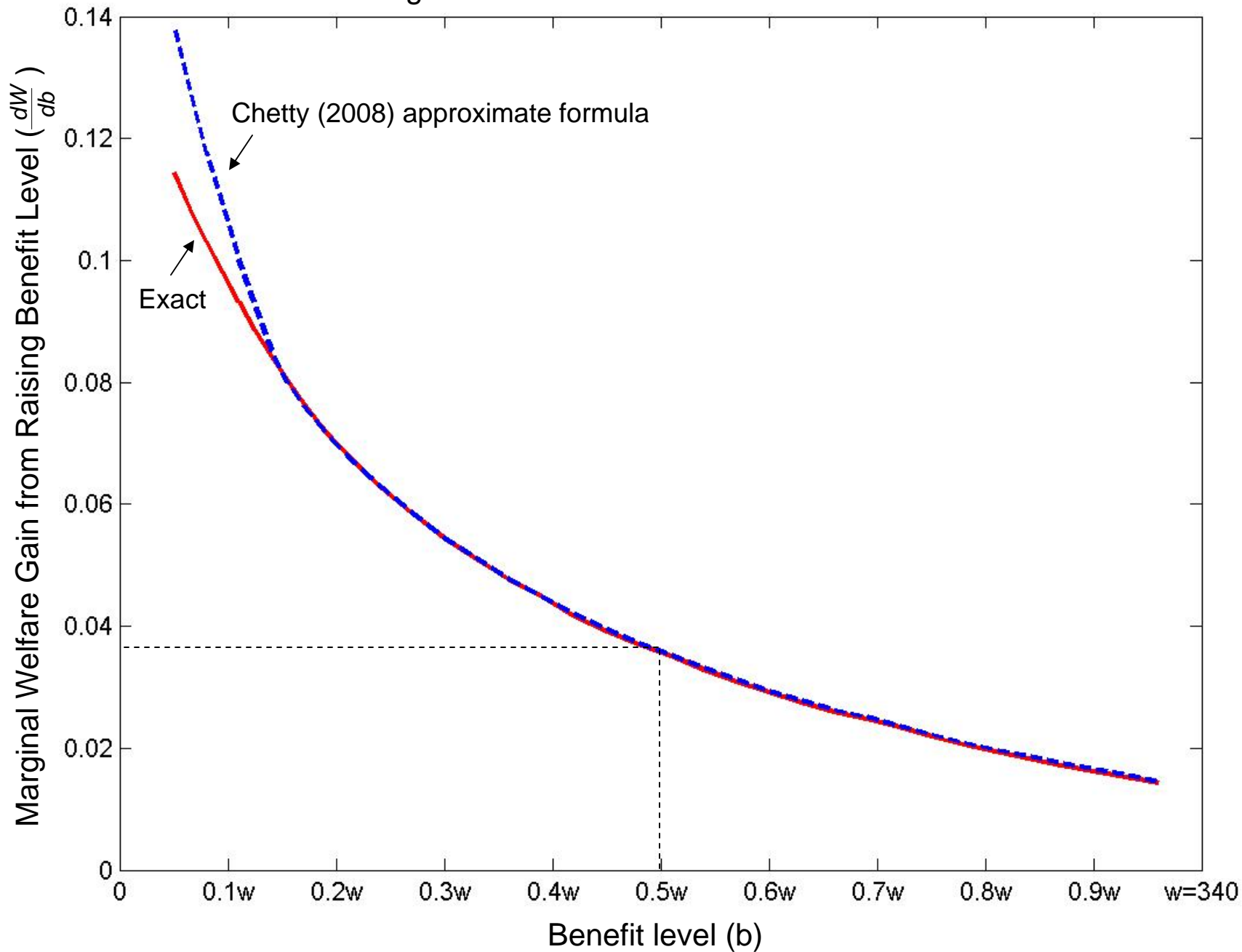


Calibration of Chetty (2008) formula

- Chetty estimates de/dA_0 and de/db using quasi-experimental variation in UI laws and severance payments.
 - Plugging estimates into formula for dW/db , Chetty calculates dW/db
 - Welfare gain from raising weekly benefit level by 10% from current level in U.S. (50% wage replacement) is \$5.9 bil = 0.05% of GDP
 - Uses structural model calibrated to match sufficient statistics to assess policy implications
 - dW/db falls rapidly with b , suggesting we are near optimum
 - Also conducts simulations of other policies, e.g. provision of liquidity through loans.
- Combine structural and sufficient stat. approaches to extent beyond marginal policy analysis in a credible manner.

Figure 1

Marginal Welfare Gain vs. Initial Benefit Level



APPLICATION 3: BEHAVIORAL MODELS

Chetty, Raj, Adam Looney, and Kory Kroft. “Salience and Taxation: Theory and Evidence.” *American Economic Review*, forthcoming.

Bernheim, Douglas and Antonio Rangel. “Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics,” *Quarterly Journal of Economics*, forthcoming.



Orig. Tag

Exp. Tag

Figure 2a

Per Capita Beer Consumption and State Beer Excise Taxes

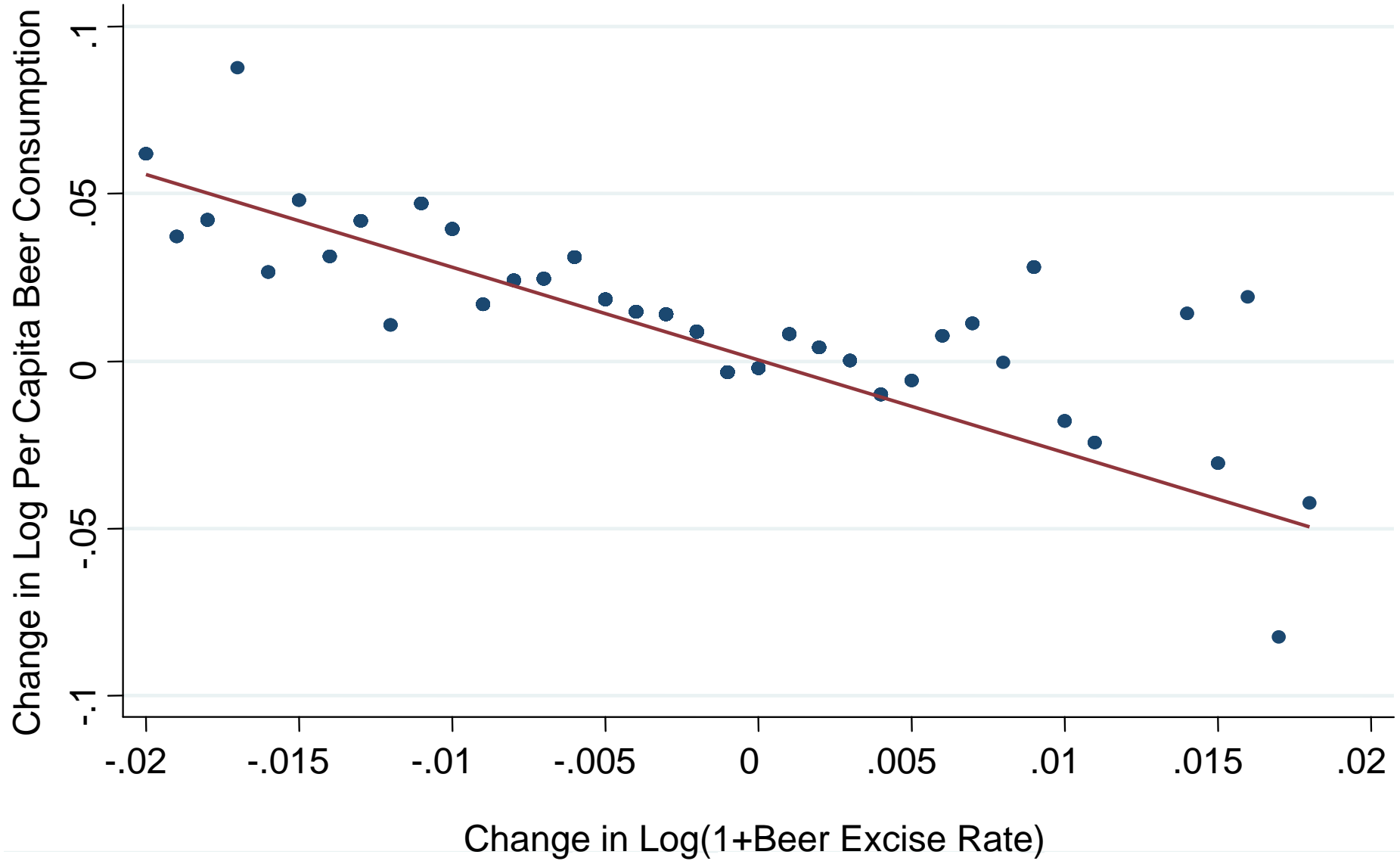
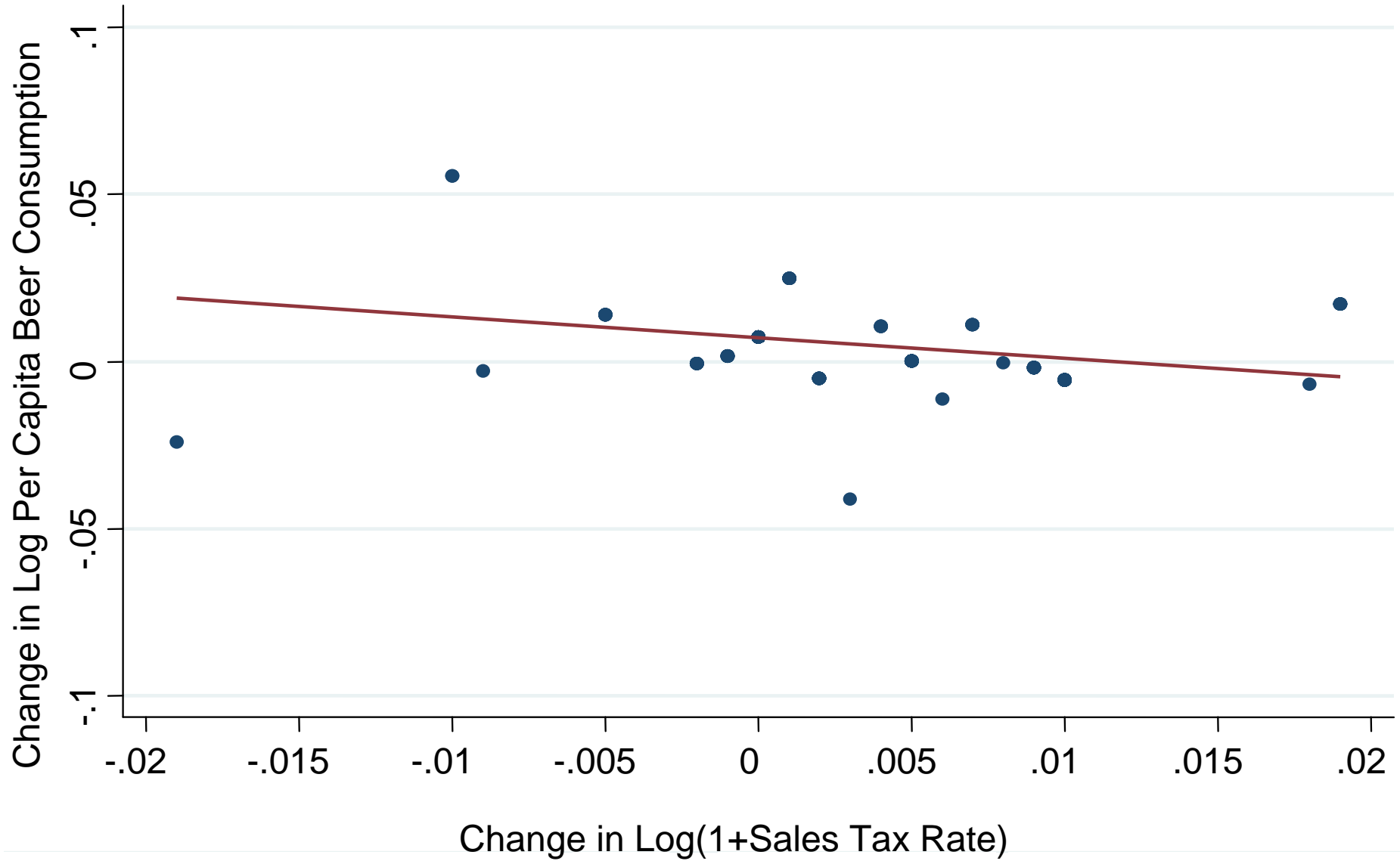


Figure 2b

Per Capita Beer Consumption and State Sales Taxes



Chetty, Looney, Kroft (2008): Welfare Analysis in Behavioral Models

- Existing results on optimal tax/transfer policy are based on models inconsistent with preceding evidence.
 - Need an alternative method of analyzing welfare consequences (incidence, efficiency costs) in view of evidence to make progress.
- Objective: Develop formulas for incidence and efficiency costs of taxes that allow for salience effects
- Many potential positive models for salience effects (cognitive costs, heuristics, psychological factors); difficult to distinguish
 - Therefore develop a method of welfare analysis that does not rely on a specific positive model of optimization errors

Setup

- Two goods, x_1 and x_2 ; normalize price of x_2 to 1
- Good x_2 untaxed. Government levies a tax t on x_1 ; tax not included in the posted price (not salient).
- Representative consumer has quasilinear utility:

$$U(x_1) = u(x_1) + Z - (p + t)x_1$$

- Key deviation from standard neoclassical model: do **not** assume that x_1 is chosen to maximize $U(x_1)$
- Instead, take demand $x_1(p, t)$ as an empirically estimated object, permitting $dx_1/dp \neq dx_1/dt$
- Place no structure on demand functions except for feasibility:

$$(p + t)x_1(p, t) + x_2(p, t, Z) = Z$$

Calculation of Excess Burden

- Social welfare function to calculate excess burden:

$$W(p, t) = \{u(x_1) + Z - (p + t)x_1\} + T(t)$$

- No envelope condition for $x_1 \rightarrow$ need to deviate from standard rubric
- Totally differentiate $W(t)$ to obtain

$$\frac{dW}{dt} = [u'(x_1) - p] \frac{dx_1}{dt}$$

- Challenge: identifying $u'(x_1)$ when agents do not optimize perfectly
 - In neoclassical model, know that $u'(x_1) = p+t$ from f.o.c.
 - One strategy: specify structural model of how x_1 deviates from optimal choice, then back out $u'(x_1)$
 - Alternative: make a general assumption to narrow class of models and identify sufficient stats.

Preference Recovery Assumption

A1 When tax inclusive prices are fully salient, the agent chooses the same allocation as a fully optimizing agent:

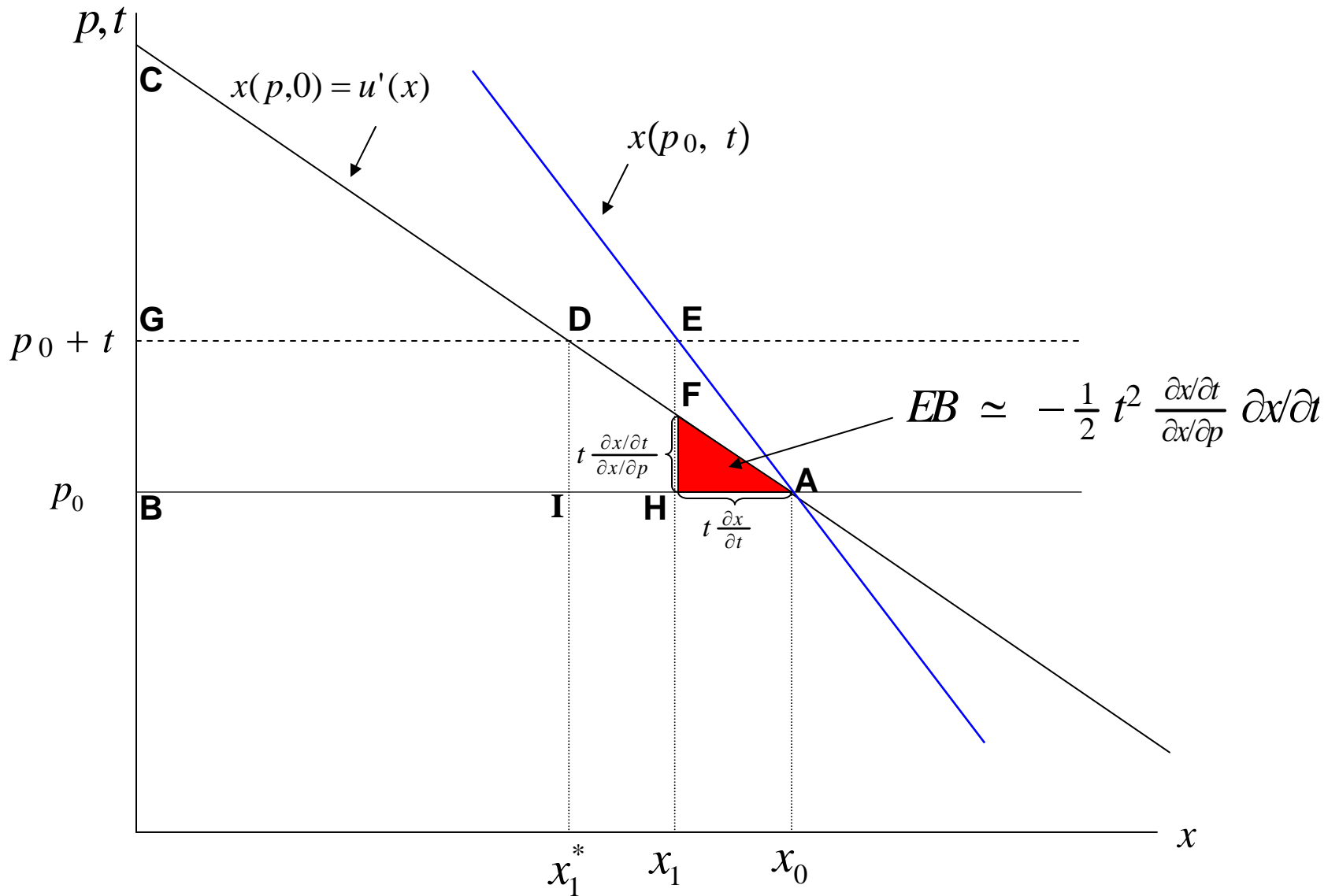
$$x_1(p, 0) = x_1^*(p, 0) = \arg \max u(x_1(p, 0)) + Z - px(p, 0)$$

→ Two steps in efficiency calculation:

1. Use **price-demand** $x(p, 0)$ to recover utility as in standard model
 2. Use **tax-demand** $x(p, t)$ to calculate $W(t)$ and EB
- Easy to illustrate graphically in case of quasilinear utility

Figure 4

Excess Burden with Quasilinear Utility and Fixed Producer Prices



Formula for EB with Optimization Errors

- When utility is quasilinear, excess burden of a small tax t is

$$EB \simeq -\frac{1}{2} t^2 \theta \frac{\partial x}{\partial t} \quad \text{where} \quad \theta = \frac{\partial x}{\partial t} / \frac{\partial x}{\partial p}$$

- Simple modification of Harberger formula: price (or wage) and tax elasticities are together sufficient statistics
- Similar simple modification of standard formula for tax incidence
- Formula permits arbitrary optimization errors w.r.t. taxes, but requires optimization w.r.t. prices

SUFFICIENT STATISTIC VS. STRUCTURAL APPROACHES

Advantages:

1. Simplifies identification: permits focus on estimating dx_1/dt using transparent, design-based methods (e.g. experiments)
 - Can therefore be implemented with fewer assumptions than structural method (e.g. arbitrary heterogeneity)
2. Can be applied when positive model unclear

Disadvantages:

1. Can only be used for local welfare analysis around observed policies unless paired with structural model
2. “Black box”: welfare analysis never “theory free.”
 - Primitives not identified → cannot determine if assumptions consistent with data (Feldstein 1995, Gruber 1997)

Combining Structural and Sufficient Statistic Methods

- Sufficient statistic formulas should be used in combination with structural methods
 - Evaluate structural model by testing whether its prediction for marginal welfare gains match sufficient statistic prediction
 - Use structural model for overidentification tests of validity of general model used to derive suff stat. formula
 - Calibrate structural model to match key moments for welfare
 - Make out-of-sample predictions (e.g. optimal policy) guided by structural model
- Can pick a point on interior of continuum between program evaluation and fully structural work.

POTENTIAL APPLICATIONS

- **[Labor]** Training programs, minimum wage
 - Lee and Saez (2008) – optimal minimum wage is a function of employment elasticity w.r.t. minimum wage
- **[Macro]** Intertemporal behavior, growth models
 - Aguiar and Hurst (JPE 2005) vs Scholz et al. (JPE 2006): identify key moments for calibrations.
- **[IO]** Analysis of competition policy, regulation
 - Challenge: allowing for strategic interactions and non-marginal changes