Comments are welcome at poterba@mit.edu.
Lecture 1: Tax Analysis and Incidence

Review of partial equilibrium tax analysis and discussion of ways to use this basic building block to connect to other developments in applied economics.

Assume: Market initially in equilibrium at price $p$. Government introduces a per-unit tax at rate $\tau$ per unit of the good purchased. Tax is rebated in lump-sum form ("differential incidence"). How does this affect producers and consumers?

Market Equilibrium:

\[ D(p + \tau) = S(p). \]  (1)

$p = \text{producer price}$ of the taxed commodity
$\tau = \text{"specific tax" of amount } \tau$
$q = \text{the consumer price} \ (q = p + \tau)$.

"Legal Incidence is Irrelevant." It does not matter whether the tax is collected from producers or consumers. $D(q) = S(q - \tau)$ vs. $D(p + \tau) = S(p)$. 
A tax at rate $\tau$ per unit of the taxed good is an example of a specific tax. The tax amount is independent of the producer price of the good. Examples:

* federal (18.4 ¢) and state gasoline excise taxes
  (average 18.2¢ per gallon nationally, 23.5¢ in MA)
* federal tax on cigarettes ($1.01 per pack)
* federal tax on distilled spirits ($13.50 per gallon).

In contrast, ad valorem sales taxes are levied as a percentage of the value of each transaction. Examples:

* Massachusetts sales tax which is levied at a rate of 6.25% on purchases other than food, prescription drugs, fuel, electricity, and clothing costing less than $175.00.

To analyze a specific tax: find $\frac{dp}{d\tau}$ from (1):
\[
\frac{dp}{d\tau} = \frac{D'}{S' - D'}.
\]  

(2)

Rewrite in terms of elasticities (in principle we need compensated elasticities, but Marshallian are used in many applications, invoking small income effects). Multiply numerator and denominator by \(q/D = (p+\tau)/D\) which yields

\[
\frac{dp}{d\tau} = \frac{\eta_D}{p + \tau} \frac{\eta_S - \eta_D}{p}.
\]  

(3)

where \(\eta_D = \frac{D'(p+\tau) \cdot (p+\tau)}{D(p+\tau)}\) and \(\eta_S = \frac{S'(p) \cdot p}{S(p)}\). When there are no prior taxes and the specific tax is small, we can evaluate (3) at the point \(\tau = 0\) so

\[
\frac{dp}{d\tau} = \frac{\eta_D}{\eta_S - \eta_D}.
\]  

(4)
Special Cases:

i) \( \eta_S = 0 \): Inelastic supply. \( \frac{dp}{d\tau} = -1 \). The consumer price is unaffected by the tax. “Producers bear the tax.”

ii) \( \eta_D = 0 \): Inelastic demand. \( \frac{dp}{d\tau} = 0 \). The producer price is fixed, consumer prices rise by the full amount of the tax. “Consumers bear the tax.”

iii) \( \eta_S = \infty \): Infinitely elastic supply. \( \frac{dp}{d\tau} = 0 \). The consumer price adjusts by the full amount of the tax and again “consumers bear the tax.” This may describe the situation in a "small open economy" that levies a tax on a single good.
Efficiency Analysis

The diagram depicting the deadweight loss (DWL) of a tax helps motivate the corresponding algebra. The area of the DWL triangle is \((1/2)\tau(Q_0 - Q_1)\). The area is positive, but by convention we set DWL < 0.
We can write the (negative of) the area of the triangle as:

\[ \text{DWL} = \frac{1}{2} \tau dQ = \frac{1}{2} \frac{dQ}{d\tau} \cdot \tau^2. \]

This is sometimes called the “Harberger Triangle.” Note that \(dQ\) denotes the change in the quantity of the good transacted as a result of the specific tax \(\tau\).

Continuing the analysis of a new and small tax, we can set \(d\tau = \tau\) in (4) so that

\[ dQ = S'(p) \cdot dp = \frac{Q}{p} \cdot \tau \cdot \frac{\eta_S \eta_D}{\eta_S - \eta_D}. \quad (5) \]

DWL is therefore

\[ \text{DWL} = \frac{\eta_S \eta_D}{\eta_S - \eta_D} \cdot \frac{Q}{p} \cdot \tau^2. \quad (6) \]
A more revealing expression normalizes the deadweight burden by the revenue raised, \( R = \tau Q \):

\[
\frac{DWL}{R} = \frac{1}{2} \cdot \frac{\eta_S \eta_D}{\eta_S - \eta_D} \cdot \left( \frac{\tau}{p} \right). \tag{7}
\]

For policy analysis it is often more helpful to consider the marginal DWL per unit of revenue raised, \( \partial DWL/\partial R = (\partial DWL/\partial \tau)/(\partial R/\partial \tau) \). We can find this ratio by differentiating (6) to obtain \( \partial DWL/\partial \tau \), and by differentiating

\[
R = \tau^*Q \tag{8}
\]

to find the denominator \( (\partial R/\partial \tau) \). Since that \( Q = D(p+ \tau) \), we find \( \partial R/\partial \tau = Q + \tau^*Q'(p+ \tau) \) which at \( \tau = 0 \) is just \( Q \). Dividing the derivative of (6) by \( Q \) yields:

\[
\frac{\partial DWL}{\partial R} = \frac{\eta_S \eta_D}{\eta_S - \eta_D} \cdot \left( \frac{\tau}{p} \right) = 2 \cdot (DWL/R). \tag{9}
\]
Several Observations About DWL:

(i) DWL is increasing in both $\eta_s$ and $\eta_d$. If either elasticity is zero, there is no deadweight burden. This is the basis for claims that an efficient tax system places high tax rates on inelasticity supplied and demanded goods.

(ii) DWL as a share of revenue raised is increasing in the ad valorem tax rate $\tau/p$. This is the basis for claims that roughly equal tax rates across goods are desirable and that high tax rates lead to high deadweight losses relative to revenue yield.

(iii) If $\eta_s = \infty$ for each of N goods, and there are no cross-price demand effects for these goods, then if the tax rate is equal for all taxed goods, the sum of the deadweight losses in all markets will be a sales-weighted arithmetic mean of the demand elasticities for the taxed goods. Raising the tax rate on goods with lower demand elasticities, and lowering the tax rate on goods with higher elasticities, while keeping total revenue constant, will reduce the sum of deadweight burdens. This is a simple form of an "optimal tax" problem.
Tax Incidence in the One-Market Setting

Incidence analysis - the study of how the burden of a tax is divided between different market participants - corresponds, in the case of an excise tax in a single market, to the division between producers and consumers. In the absence of pre-existing taxes, we can approximate the producers’ burden as the number of units sold times the change in the producer price (Qdp), and the consumers’ burden as the number of units purchased times the change in the consumer price (Qdq). (By convention to make both values negative, since dp < 0 and dq > 0, we set consumers burden equal to -Qdq.)

Producers’ burden is

\[ B^P = S(p) \frac{dp}{d\tau} \tau = \left( \frac{\eta_D}{\eta_S - \eta_D} \right) \cdot Q\tau. \]  

(10)

Note when \( \eta_D = 0 \) the producer bears none of the tax. In the same notation, the consumers’ burden is
$$B^C = -D(q) \frac{dq}{d\tau} \tau = -D(q) \left[ \frac{dp}{d\tau} + 1 \right] \cdot \tau = -\left( \frac{\eta_s}{\eta_s - \eta_D} \right) Q \tau \quad (11)$$

When $\eta_s = 0$ the consumer bears no burden.

**Illustrative Application: Cigarette Taxation**

Partial equilibrium tax analysis can be used to compute the efficiency costs and the distributional effects of the cigarette excise tax. In 2012, average retail price of a pack of cigarettes was about $6.00. U.S. consumers purchased approximately 15 billion packs of cigarettes. In 2000, annual consumption was 24 billion packs; rising tax burdens on tobacco are one of the factors that are widely cited as a factor in this decline.

Federal excise tax: $1.01/pack
Average state cigarette excise tax: $1.49
Average state sales tax: $.20
Average producer price of cigarettes: $3.30 (includes about $0.60 paid to distributors/retailers and $2.70 paid to manufacturer; treat $3.30 as producer price). Note that the total excise tax on cigarettes is about $2.70/pack.

In some states the taxes are much higher; in Massachusetts, for example, the total tax is $3.52. In New York City, the total state and local tax burden, including a $1.50/pack city tax, is $5.85/pack. In Pennsylvania the state and local tax burden is $1.60 -- where would you shop?

National average: $\tau/p \approx 0.82 \ (2.70/3.30)$. Tobacco Tax Revenue ($\tau Q$) about $40$ billion/year.
Elasticity Assumptions to Implement Partial Equilibrium Analysis:

**Demand:** Very horizon-sensitive. At "intermediate" (five year?) horizon, estimates suggest $\eta_D \approx -0.50$. Illustrative empirical work: Hu, Ren, Keeler, and Bartlett (1995, *Health Economics*) estimate an overall elasticity of -0.46, which is the combined effect of an elasticity of -0.22 for consumption by those who smoke and -0.33 for the decision of whether or not to smoke at all.

**Supply:** If the U.S. were a small open economy, one might assume a supply elasticity of infinity, since the country might be a price-taker in the world market. In practice the supply curve of cigarettes to the U.S. economy is likely to be upward sloping. Limited empirical evidence on supply elasticity, so evaluate DWL with different values of $\eta_S$. 
Efficiency and Distributional Calculations: Cigarette Excise Tax
Parameter Assumptions
\[ \eta_D = -0.50, \eta_S = \infty \quad \eta_D = -0.50, \eta_S = 1 \]

<table>
<thead>
<tr>
<th></th>
<th>DWL/Revenue</th>
<th>Marginal DWL</th>
<th>Producers’ Burden</th>
<th>Consumers' Burden</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.210</td>
<td>0.420</td>
<td>0</td>
<td>$13.3B</td>
</tr>
<tr>
<td></td>
<td>0.137</td>
<td>0.274</td>
<td>$13.3B</td>
<td>$26.7B</td>
</tr>
</tbody>
</table>

Recall that \( \tau Q = \) revenue is about $40B. While these calculations are illustrative, even in the simple case of the market for cigarettes they omit both conceptual and empirical richness. We’ll explore them in turn.
Building from the Intuitive DWL Analysis: Next Conceptual Steps

* Compensated Elasticities. Deadweight burdens depend on compensated not uncompensated elasticities. For a good that is a small share of income this may not be an important distinction but for larger budget shares it matters. Public finance economists recognize that a tax can be imposed, quantity can remain the same, but there can be a large deadweight loss.

* Multi-Market /General Equilibrium Effects. Goulder and Williams (2003 JPE) point out that a change in an excise tax in one market can matter through its effect in other markets (i.e. by changing demands). They base their analysis on a familiar expression from Harberger’s original work:

\[
DWL = .5 t^2 \sum_j \frac{dX_j}{dt_k} + \sum_{j \neq k} t_j \frac{dX_j}{dt_k}
\]  

(12)

Most analyses ignore the second term, the effects in other markets, often because of lack of information on \( \frac{dX_j}{dt_k} \). But these effects can be important,
particularly when the excise tax has a non-trivial effect on the real product wage. Consumer price index \( q^* \) is a budget-share \((s_j)\) weighted average of consumer prices:

\[
q^* = \sum_j s_j^* p_j (1 + t_j)
\]

(13)

The real product wage is \( w(1-t_y)/q^* \), so raising an excise tax on a good with a non-trivial budget share can distort the labor market by reducing the real product wage. Goulder and Williams suggest that the under-estimate for range of elasticities we have used could be on the order of 50%. Their general point is that when raising a tax in one market has important effects in other markets, we need to consider them.

* Pre-existing Market Distortions. Examples: nominal price rigidities, unionized workers who raise production costs and create factor market distortions, uninformed consumers, imperfect competition. Small taxes in perfectly competitive markets yield second-order efficiency costs. The same taxes in markets with other distortions can have first-order effects – triangles become trapezoids.
An example is the role of imperfect competition in tobacco markets. Barnett, Keeler, and Hu (JPubE 1995) show that the retail price increase associated with a federal excise tax is larger than the reaction to a state or local tax, suggesting price coordination. This means changes in tax rates could have effects on degree of market power or facilitate collusion - with additional welfare effects.

*Externalities*. When the consumption of a good has effects that are not priced and therefore not considered in purchase decisions, the partial equilibrium efficiency analysis will not capture all of the efficiency consequences of an excise tax. It is even possible that if the external costs of a good's consumption are large enough, raising the tax may lead to an efficiency gain. A key public finance insight on externality-correction is that the optimal tax may not equal the sum of the marginal damages created by consumption of a good. Atkinson and Stern (REStudies, 1974) pointed out that when taxing the externality-creating good also has effects on revenue -- for example, taxing gasoline raises commuting costs and lowers labor supply, which reduces labor income taxes -- we need to consider the "fiscal externality" and the shadow cost of raising the foregone revenue associated with this behavioral response.
Estimates of the external cost of smoking have been made using various strategies, and they vary substantially. A brief taxonomy of potential external costs associated with cigarette consumption:

* Costs that smokers impose on other participants in group health and life insurance programs, and in social insurance programs: $0.33 per pack 1995 (estimate from Congressional Research Service)

* Secondhand smoke external effects: controversial, upper estimate $0.70/pack.

* Maternal smoking increases likelihood of low-birthweight baby: $0.40-$0.70 per pack

* Reductions in workplace productivity of non-smokers, cleaning charges, fires... somewhat difficult to determine external components
Empirical Challenges in the Analysis of Cigarette Taxation:

1. How to Choose the Price Elasticity of Demand: What Horizon?

- Need to distinguish short run vs. long run (effect of changes in quantity for current smokers vs. changes in number of smokers)
- Example: Chris Carpenter & Phil Cook, NBER WP 2007, study National Youth Risk Behavior Survey (YRBS) 1991-2005 using state tax rate changes. 100k high school kids. Price elasticity of smoking participation is -0.56 for high school kids.
- Price elasticity appears to be largest at lower income levels, so burden of higher taxes may fall more on higher income smokers: Income Quartile: 1 (-1.09 elasticity), 2 (-.70), 3 (-.53), 4(-.39).
- Recognize very short run “hoarding” effects (California 50 cent per pack increase example, increase in residential fires)
2. Smuggling.

If higher tax rates increase smuggling then effect on revenue will be estimated incorrectly (this matters for optimal tax calculation). Also we’ll mis-estimate the demand effects and the welfare effects. (Gruber, Sen, Stabile, JHE, 2003): huge interstate differences in tax rates, estimate of 6% of total U.S. cigarette consumption not paying taxes (smuggled), costs states about $1 billion/year in foregone revenue (role of Indian reservations); note in Canada when taxes rose sharply, "legal exports" rose from 1.5% of sales to 50% of sales. (Cigarettes were being shipped to US and illegally re-imported).

How elastic is smuggling, what are revenue effects, how does smuggling affect estimates of price elasticity?

* Interesting work by Fisman and Wei, NBER WP 2003, comparing composition of goods crossing the China-Hong Kong border and the effects of differential tariffs.

* Related example (Marion and Muhlegger, JPE 2008) of diesel fuel and #2 home heating oil. Dying one product to promote compliance increased sales of diesel fuel 26%! 
3. Distribution of Burden from Tobacco Taxes

Consumption of cigarettes accounts for a larger budget share at low income levels. Poterba AER 1989:

<table>
<thead>
<tr>
<th>Quantile</th>
<th>Income</th>
<th>Consumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowest</td>
<td>4.6%</td>
<td>2.2%</td>
</tr>
<tr>
<td>Second</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>Third</td>
<td>1.3</td>
<td>1.7</td>
</tr>
<tr>
<td>Fourth</td>
<td>0.9</td>
<td>1.3</td>
</tr>
<tr>
<td>Fifth (top)</td>
<td>0.5</td>
<td>0.7</td>
</tr>
</tbody>
</table>

Lifetime versus annual measures of distributional burdens may give different results.
Behavioral Economics and Tax Policy:

General Insights and Illustration Using the Cigarette Excise Tax Example

1. Bounded Self-Control. Addiction might be one consequence that matters for analysis of cigarette taxation. Procrastination, failing to save for retirement are other related behaviors that attract interest from tax policy.

2. Imperfect Optimization. Might be due to limited attention, which generates "salience" effects, might be limited computational capacity, which leads to "schmeduling" and "mental accounts".

3. "Nonstandard preferences" such as endowment effects. Consequences for tax design and analysis are not yet clear.

What aspects of tax analysis are affected by behavioral economics considerations?

1. Modeling behavioral responses to taxes

2. Analyzing the welfare effects of tax policy -- if the quantity of the taxed good doesn't change when taxes rise, what is the welfare cost?

3. Optimal tax theory -- when defaults matter, when individuals can become addicted, can tax policy play a larger role in shifting behavior?

General problem: Neoclassical public finance relies on revealed preference arguments to support welfare analysis. Behavioral economics often questions the interpretation of demand choices - hence a challenge for standard public economics analysis.
Example One: Addiction.

Two potential "models" for this potential self-control problem:

* Becker-Murphy (JPE 1988) "rational addiction" model. Each consumer has a stock of “addiction capital” S in each period; c is the consumption of the addictive good. Utility each period:

\[ U_t = u(c_t, S_t, x_t) \]  \hspace{1cm} (14)

where \( x \) = composite for all other goods, \( Uc > 0, Us < 0 \), and

\[ S_t = (1-d)*S_{t-1} + c_t \]  \hspace{1cm} (15)

The consumer maximizes

\[ V_t = \Sigma_{j=0,T} (1+\beta)^j u(c_{t+j}, S_{t+j}, x_{t+j}). \]  \hspace{1cm} (16)

Consumers are forward looking and recognize that if they smoke at \( t \), it will be more difficult not to smoke at \( t+j \). But some consumers still decide to smoke. This is not a decision that is made for irrational reasons, but rather it reflects a
fully rational calculation of the lifetime costs and benefits. A key empirical prediction of this model is that expected future increases in cigarette prices (or taxes) should reduce current smoking. In such rational addiction models the case for government intervention is eliminated.

*Irrational model with time-inconsistent preferences (examples include Gruber-Koszegi (QJE 2001, JPubE 2004)). Time-inconsistent consumers do not recognize future health costs. Example of potential preference structure: hyperbolic discounting. Per-period utility can still be given by (14) but now the consumer maximizes

\[
V_t = u(c_t, S_t, x_t) + \lambda \sum_{j=1,T} (1+\beta)^{-j} u(c_{t+j}, S_{t+j}, x_{t+j}).
\]

(17)

\(\lambda \neq 1\) induces time-inconsistent plans: consumers fail to fully recognize the future effects of smoking. Tax policy could play a role in reducing current demand for cigarettes and thereby lead to welfare improvement. What evidence would support this model? i) laboratory evidence for time inconsistency; ii) evidence that many smokers wish they never started, or would like to stop; iii) many smokers spend money on commitment devices to stop; iv) Gruber-Mullainathan study, "likely smokers" are happier in high tobacco tax states.
Example Two: Tax Salience.

Another recent strand of behavioral research, emphasizing limited decision-making capacity, partly motivated by interesting empirical findings suggesting that consumers respond differently to taxes depending on context.

Chetty, Looney, Kroft (2009 AER) study difference between excise tax (e, included in good's posted price) and sales tax (t, charged at register). Standard neoclassical focuses on $q = p(1 + t)(1+e)$. Derivatives of demand w.r.t. $p$, $(1+t)$ and $(1+e)$ should be identical. Are they?

\[
\Delta \ln (\text{Beers Per Capita}) = -0.91*\Delta \ln (1+e) - 0.01*\Delta \ln(1+t)
\]

\[
(0.17) \quad (0.30)
\]

Prima facie rejection of standard model? Some ingenuity is required to explain this pattern in the standard neoclassical model.

Second source of evidence: randomized controlled trial (RCT) by putting labels in grocery store showing tax-inclusive prices for some goods; result is decline in
demand, implying labeling alone matters. Suggests potential importance of costs of processing information and general lack of attention.

Third, related, empirical finding: Finkelstein (2009 QJE) compares rate of change of bridge & tunnel tolls before and after adoption of "EZ Pass". Fares increase afterward, consistent with less consumer or voter visibility once charged automatically.

These empirical findings suggest that demand may not be $x(p + t, y)$ but rather $x(p, t, y)$. One simple parametric way to analyze this possibility is to write the demand curve as $x(p + \theta t, y)$. In this case

$$\frac{dp}{d\tau} = \theta \frac{\eta_D}{\eta_S - \eta_D}.$$  (19)

where $\eta_D$ denotes $x_p^*(p+t)/ x$, the "usual" demand elasticity, in this case defined as the elasticity with respect to the producer price. In the empirical work described above, $\theta \approx 0.35$ in posted prices example, and $\theta \approx 0$ for the beer excise tax.
Implications of Imperfect Salience ($\theta < 1$):

* Incidence: If a tax has low salience ($\theta \approx 0$) then consumers bear most of the burden (low salience is like low elasticity)

* Legal Incidence may matter after all - which side of the market the tax is collected on can become consequential

* Deadweight burden analysis is more complex. The budget constraint still must be satisfied. If the taxed good doesn't respond to the tax, what does? There must be adjustments elsewhere, and those are the source of the distortions. Empirical analysis of efficiency requires evidence of what demands are affected by the not-salient tax.
A Different Direction: Beyond Approximations: Exact Welfare Analysis

The consumers' burden measures above are an approximation to a utility-based measure of the costs of a tax-induced price change. They suffer from all the limitations of Marshallian consumer surplus measures. A more theoretically grounded measure such as the **compensating variation** can overcome these shortcomings:

\[
CV(q_1, q_0, u_0) = e(q_1, u_0) - e(q_0, u_0).
\]

In this expression \(e(q_1, u_0)\) denotes the expenditure function evaluated at utility level \(u_0\) and prices \(q_1\). The CV measures the cost to a consumer of a shift in prices from \(q_0\) to \(q_1\). There are several related measures that can be used to evaluate consumer welfare, as this diagram shows:

Diamond and McFadden (1974 *Journal of Public Economics*) adapt the compensating variation to measure the deadweight burden of a tax-induced price change from \(q_0\) to \(q_1 = q_0 + \tau\). They define the DWL of a tax vector \(\tau\) as
\[ CV(q_0 + \tau, q_0, u_0) - R(q_0 + \tau, q_0, u_0) \quad (21) \]

where \( R(q_0 + \tau, q_0, u_0) \) denotes the compensated revenue associated with a vector of specific taxes \( \tau \). \( R(q_0 + \tau, q_0, u_0) \) is defined using Hicksian demands:

\[ R(q_0 + \tau, q_0, u_0) = \sum_j \tau_j h_j(q_0 + \tau, u_0). \quad (22) \]

To implement the Diamond-McFadden approach, one needs estimates of the expenditure function. Most empirical work delivers estimates of Marshallian demand curves but not the underlying utility functions that are needed to construct \( e(q_1, u_0) \) and \( h(q_1, u_0) \). Hausman (AER 1981) demonstrates that in many cases it is possible to recover preferences from commonly-used demand specifications, and he then argues that there is no excuse for using “triangle approximations” in place of exact expressions for welfare loss. Consider as an example a linear demand curve for good \( x_1 \) (note that in this demand curve \( q_1 \) denotes the consumer price of good 1, not a vector of post-tax prices):

\[ x_1 = \gamma + \alpha q_1 + \delta y. \quad (23) \]
where $y$ is household income. What preferences would generate this demand function? We need to solve the partial differential equation that follows from Roy’s Identity

$$
\frac{\partial v(q, y)}{\partial q_1} - \frac{\partial v(q, y)}{\partial y} = x_1(q, y) = \gamma + \alpha q_1 + \delta y
$$

(24)

to find $v(q, y)$, the indirect utility function. In this case,

$$
v(q, y) = e^{-\alpha_1} \left[ y + \frac{\alpha}{\delta} q_1 + \frac{\alpha}{\delta^2} + \frac{\gamma}{\delta} \right].
$$

(25)

The analogous expenditure function is

$$
e(q, u) = e^{\alpha_1} u - \frac{\alpha}{\delta} q_1 - \frac{\alpha}{\delta^2} - \frac{\gamma}{\delta}.
$$

(26)
It is straightforward to check that (26) implies (23).

Another common demand specification is the double logarithmic model:

$$\log x_1 = \gamma + \alpha \log q_1 + \delta \log y.$$  \hspace{1cm} (27)

In this case, the indirect utility function is

$$v(q, y) = -e^\gamma \cdot \frac{q_i^{1+\alpha}}{1+\alpha} + \frac{y^{1-\delta}}{1-\delta}.$$  \hspace{1cm} (28)

The corresponding expenditure function is

$$e(q, u) = \left[ (1-\delta) \left( u + e^\gamma \frac{q_i^{1+\alpha}}{1+\alpha} \right) \right]^{\frac{1}{1-\delta}}.$$  \hspace{1cm} (29)

From estimates of linear or double-log demand curves, it is straightforward to find the associated preferences and then to compute welfare measures such as CV and the Diamond-McFadden measure of efficiency loss.
Lecture 2: The Elasticity of Taxable Income: Why it Matters and How We Estimate It

Why Focus on "Elasticity of Taxable Income" (ETI)?

i) Captures not just hours of work but other changes on margins of adjustment such as effort, structure of pay, occupation and career choice, non-labor income, tax evasion and avoidance.

ii) This elasticity may determine the deadweight burden of the tax system provided there are no "fiscal externalities" in which another tax base depends in part on the size of this tax base, such as a link between the corporate and the personal income tax bases. Chetty (AEJ Policy 2009) also shows that if individuals mis-estimate the cost of tax evasion, or if some "costs" of avoidance are really transfers, the deadweight loss analysis will also fail.

iii) ETI links easily to calculation of "Revenue Maximizing Tax Rate" $\tau^*$. 
Simple Analytics of ETI:

Let \( z \) = taxable income; in the simplest case

\[
    z = w*l
\]  \hspace{1cm} (1)

but more generally \( z \) would subtract deductions and could include non-labor income. The key parameter of interest will be the elasticity of \( z \) with respect to the "keep ratio" \((1-\tau)\):

\[
    e = \frac{dz}{d(1-\tau)} \cdot \frac{(1-\tau)/z}{(1-\tau)/z}
\]  \hspace{1cm} (2)

If we write the revenue yield of the tax system with taxable income as a function of the keep ratio, then

\[
    R = \tau z (1-\tau).
\]  \hspace{1cm} (3)

The ETI is often discussed in the context of the top tax bracket, and the revenue collected from taxpayers in this bracket. Assume that there are \( N_{\text{top}} \) taxpayers in the top bracket, which begins at taxable income \( z_{\text{top}} \), and that \( z_m \) is the mean
income of these taxpayers. If we raise the top marginal tax rate, which applies to income above $z_{\text{top}}$, the revenue yield absent any behavioral response would be

$$dR_{\text{top, no response}} = N_{\text{top}}(z_m - z_{\text{top}})d\tau.$$  \hfill (4)

The revenue change (loss) as a result of behavioral response to this tax rate increase is

$$dR_{\text{top, due to response}} = N_{\text{top}}\tau z_m = - N_{\text{top}}\tau e z_m d\tau/(1 - \tau).$$  \hfill (5)

The total revenue effect is the sum of (4) and (5):

$$dR_{\text{top}} = N_{\text{top}}(z_m - z_{\text{top}})\{1 - e[z_m/(z_m - z_{\text{top}})]\tau/(1 - \tau)\}d\tau.$$  \hfill (6)

It is convenient to define $a = z_m/(z_m - z_{\text{top}})$. This ratio, which depends on the shape of the upper tail of the income distribution, is constant for the Pareto distribution. (See Diamond and Saez, 2011 JEP paper on "The Case for a Progressive Tax," for a useful discussion.)
With this simplification, if we ask what tax rate will maximize revenue from the taxpayers in the top tax bracket (leaving aside changes in taxable income that might lead to shifts in $N_{top}$), we set (6) equal to zero which implies

$$1 - e^a [\tau/(1 - \tau)] = 0$$

or

$$\tau^* = 1/(1+e^a).$$

In the special case of a single-bracket tax system, $z_{top} = 0$ and $a = 1$, so $\tau^*$ is just $1/(1+e)$. This simple relationship contributes to the interest in this parameter; more generally, it is a key determinant of the revenue yield of tax increases. Notice that since $a \geq 1$, the revenue maximizing flat rate is always higher than the revenue-maximizing top marginal rate. This is because raising the top rate raises money only from taxpayers in the top bracket, but it distorts their behavior by the same amount as a flat tax change of the same magnitude would.
The ETI parameter can also be linked to the marginal deadweight burden per dollar of additional revenue. Assuming a small change in the top marginal rate, the envelope theorem allows us to write the loss in welfare from this tax increase for each individual as \(- (z_m - z_{top}) \times d\tau\). But the change in revenue collected, \(dR_{top}\), is the sum of \((dR_{top, no\ response})\) and \(dR_{top, due\ to\ response}\). This means that the additional burden per dollar of revenue (with a negative sign to make the result positive) is \(-dR_{top, due\ to\ response}/dR\). Using (5) and (6) we can write this as:

\[
\text{Excess Burden} = - \frac{\{\tau e^{*}z_m/(1- \tau)\}/\{(z_m - z_{top})\} \{1 - e^{*}a^{*}/(1- \tau)\}}{z_m - z_{top}} = - \frac{\{\tau e^{*}a\} \{1 - e^{*}a^{*}/(1- \tau)\}}{1 - \tau - e^{*}a^{*}\tau}.
\]

\(9\)

Note that this is the result that the earlier discussion of "fiscal externalities" and other considerations may call into question.
ETI Literature: Survey by Saez, Slemrod, Giertz (2012) is very helpful

1. Time Series Evidence: Long time series evidence from Saez and Piketty:

Note sharp increase in top 0.1% share in 1988, first year of low marginal tax rates after 1986 tax reform (Feenberg-Poterba).
2. Cross-Sectional and Panel Data Evidence:

a. Lindsey (1987) "Income Projection Method" uses tax return information on the composition of income by income class in 1979, and the change in aggregate income from various sources in subsequent years, to "forecast" the distribution of income by source in 1982, 1983, and 1984. Let k denote an income range, and j denote an income source (wages, for example). Then

\[ P_{j,k,1984} = \left( \frac{Z_{j,1984}}{Z_{j,1979}} \right)^* \left( \frac{Z_{j,k,1979}}{Z_{j,1979}} \right) \]  

(10)

Lindsey compared actual and predicted income taxes, and found more taxes than expected at high income levels - just where marginal tax rates fell most.

<table>
<thead>
<tr>
<th>Year</th>
<th>&gt; $200K</th>
<th>$50-200K</th>
<th>$30-50K</th>
<th>&lt; $30&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980</td>
<td>+0.2%</td>
<td>+1.9%</td>
<td>+2.8%</td>
<td>-0.2%</td>
</tr>
<tr>
<td>1982</td>
<td>+18.6</td>
<td>+3.7</td>
<td>+3.6</td>
<td>+0.4</td>
</tr>
<tr>
<td>1984</td>
<td>+55.7</td>
<td>+7.6</td>
<td>+1.4</td>
<td>-5.6</td>
</tr>
</tbody>
</table>

Estimate of e for highest income group was greater than 2.
b. Regression analysis allowing for individual effects

Feldstein (1995 JPE) uses panel data on individual tax returns (example of the key role of administrative tax return data in empirical public finance) to estimate

\[ \ln \left( \frac{Z_{i,2}}{Z_{i,1}} \right) = \alpha + \beta \ln \left( \frac{1-\tau_{i,2}}{1-\tau_{i,1}} \right) + X_i \gamma + \varepsilon_i \]  

(11)

where variation in marginal tax rates comes from TRA86.

<table>
<thead>
<tr>
<th>1985 ( \tau_i )</th>
<th>Sample Size</th>
<th>( \Delta \ln(1-\tau_i) )</th>
<th>( \Delta \ln Z_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>22</td>
<td>800</td>
<td>9.0%</td>
<td>9.4%</td>
</tr>
<tr>
<td>28</td>
<td>713</td>
<td>16.3</td>
<td>3.9</td>
</tr>
<tr>
<td>45</td>
<td>45</td>
<td>30.9</td>
<td>12.4</td>
</tr>
<tr>
<td>49</td>
<td>35</td>
<td>41.2</td>
<td>27.1</td>
</tr>
<tr>
<td>50</td>
<td>22</td>
<td>44.0</td>
<td>18.4</td>
</tr>
</tbody>
</table>

Implied estimate of \( \varepsilon \):

"high vs. medium" marginal tax rate comparison: 1.04
"highest vs. high" marginal tax rate comparison: 3.05
"highest vs. medium" marginal tax rate comparison: 2.14
Problem with this approach: mean reversion or underlying trends
  • if mean reversion, then folks who are rich will get poorer – if tax rates on
    rich are falling, understates tax response
  • if underlying trends towards more inequality, and tax rates on rich are
    falling, then overstate tax response

c. Repeated Cross Sections of Tax Returns

Gruber and Saez (2003) estimate a **compensated** elasticity of taxable income by
controlling for income. Their basic specification, where i denotes an individual
tax return in year 1 or year 2 and is:

\[
\ln \left( \frac{Z_{i,2}}{Z_{i,1}} \right) = \alpha + e \ln \left( \frac{1-\tau_{i,2}}{1-\tau_{i,1}} \right) + \gamma \ln \left( \frac{Z_{i,2} - T_{i,2}}{Z_{i,1} - T_{i,1}} \right) + X_i \gamma + \epsilon_i
\]  

They estimate this equation using repeated cross sections of U.S. income tax

**Broadest Income Concept:**  \( e = 0.30 \) \( (0.12) \)
**Taxable Income:**  \( e = 0.47 \) \( (0.19) \)
Note here include both uncompensated tax effect and income effect – income effect can in principle be identified separately as tax reforms can change individual incomes and MTRs differently.

Problem is endogeneity of tax rate – so instrument with: \[ \frac{(1-\tau^*_i,2)/(1-\tau_{i,1})}{(1-\tau_{i,2})/(1-\tau_{i,1})} \], where the former uses period 1 incomes in period 2 tax code - that is, instead of holding bracket constant (as in Feldstein), hold the actual income constant – just change in tax rate for fixed income (note this IV fails if tax code didn’t change). Controls include non-linear function of base period income – so that control for both mean reversion and omitted trends correlated with base period income.

Conclusions:
• income effects negligible
• overall ETI of 0.4 – more modest than Feldstein
• driven totally by highest income tax payers – 0.6 for top, 0 for bottom
• broad income elasticity much lower – 0.12 overall, 0.17 for highest income earners
Extension: Kopczuk (2005) – if no income is deductible, elasticity is 0.12 – but that as more income is deductible, elasticity rises one for one with share deductible.

Giertz (2007) uses richer data than Gruber-Saez to replicate their analysis over a longer period: 1979-2001. He estimates ETI of 0.3 – lower in 1990s – partly because fewer folk were itemizing. Gross income elasticity is same throughout time period at about 0.15.

Chetty et al. (2009) suggest that adjustment costs are key and that longer run effects may be larger (e.g. since bunching is much larger around larger kinks). Costs of adjustment is an emerging important topic in behavioral response to taxation.
Key Questions Regarding Interpretation of ETI:

1. Short-Run Behavioral Responses vs. Long-Run Shifts
   - example of 1986 Tax Reform Act of 1986 capital gains rate increase
   - 1992 acceleration of bonuses on Wall Street
   - SR effects can be large, LR much harder to evaluate
   - best LR evidence points to ETI around 0.3 or 0.4 -- implies revenue-maximizing top income tax rate of around 60%

2. "Fiscal Externalities": when is income shifted from one part of fiscal system to another?
   - example of Subchapter C / Subchapter S corporations. When $\tau_{\text{personal}} > \tau_{\text{corporate}}$ individuals choose to locate income within corporations, avoid direct earnings but when sign flips, income migrates INTO the personal income tax base (no underlying change in behavior)
   - must consider all types of income: standard example is turning labor income into various forms of capital income

3. Evasion: embed tax paying decisions in optimizing model of tax evasion
4. Heterogeneity in Responses:
   - Most taxable income effects seem to flow from deduction behavior
   - High income households seem more responsive than lower income counterparts (different lessons for top bracket and EITC?)
   - History-dependence: cutting rates after many years of high rates may have different effects than cutting further after low rates
   - Structure of entire tax code can matter - tax avoidance opportunities
Illustrating the Use of ETI: The Revenue Effects of the “Fiscal Cliff Fix”

For a taxpayer with taxable income of $500K in 2012, tax rates were 35% federal, (say) 6% state deductible at 35% so \( .35 + .06 \times (1-.35) = 0.389 \).

In 2013: 39.6% federal income + 0.9 federal Medicare + .03*(39.6) phase-out of itemized deductions + 6*(1-.396) state = 0.453

“Keep Ratio”: 2012 = 1-.389 = 0.611, 2013 = 1-.453 = 0.547

Ratio is .895, \( \ln (.547/.611) = -0.11 \).

If ETI = 0.25, that implies a 2.78% decline in taxable income

DATA (from 2007): look at married joint filers

46.1 M returns, $4.17T in taxable income (compare 110.5M returns, 6.1T taxable income total)
Returns with more than $500K taxable income: 871K returns (top 1.8%), 1.33T taxable income. This means 0.897T taxable income above $500K.

“Naïve” revenue increase: 6.7% increase in MTR (4.6+.9+1.2) above $500K so $60B increase

ETI effect: reduce $1.33T taxable income by 2.78% or $37B. This is taxed at 41.7% federal so .417*37B = $15.4B foregone revenue.

About one quarter of tax increase is lost to behavioral response.
"Bunching Estimators" for ETI and Other Elasticities (Saez, AEJ: Policy 2010):
Empirical Question: How much "excess mass" do we observe at the kink point? Subject to assumptions on shape of distribution of taxpayers without the kink point, the size of the kink can be related to the elasticity of taxable income at this point. Point Estimate: \( e \approx 0.2 \) at "first kink," smaller (near zero) elsewhere.
Emerging evidence on difference between wage earners (with W-2 reporting of income) and self-employed individuals:

Figure 4. Earnings Density and the EITC: Wage Earners versus Self-Employed
Lecture 3: Tax Expenditures

Overview: In practice tax systems rarely tax the entire “base” that they are designed for. Example: “C Efficiency” and Value Added Taxes. “C Efficiency” denotes fraction of NIPA consumption that is covered by the VAT base.

<table>
<thead>
<tr>
<th>Country</th>
<th>C-Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canada</td>
<td>74%</td>
</tr>
<tr>
<td>France</td>
<td>49</td>
</tr>
<tr>
<td>Germany</td>
<td>55</td>
</tr>
<tr>
<td>Italy</td>
<td>41</td>
</tr>
<tr>
<td>New Zealand</td>
<td>98</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>41</td>
</tr>
</tbody>
</table>

Source: R. de Mooij and M. Keen, NBER WP 17913, March 2012.
Compare two potential tax bases: Consumption vs. Income. The former excludes income from capital. The latter, “Haig Simons income tax base” includes all accretions to wealth: labor income, capital income such as dividends and interest payments, and unrealized capital gains. (Choice of tax base and tax reform will be discussed in detail in a later lecture.)

The current U.S. income tax is a “hybrid” income tax. The base is not as broad as a Haig-Simons income tax – it excludes some items (imputed rent on homes, employer provided health insurance), allows a range of deductions (charity, state and local tax payments, medical expenses above a threshold), and offers credits for some other activities (Child Tax Credit for families with children, Earned Income Tax Credit for low income families with earnings, credits for investment in certain energy-efficient technologies). It also taxes many types of capital income (income in pension accounts and capital gains for example) at rates below the tax rate on wage income.

Current fiscal environment has generated great interest in raising revenue by broadening the tax base and limiting deductions. Precisely how they are restricted has important consequences for both efficiency and redistribution.
“Tax Expenditure” is a phrase coined by Stanley Surrey of Harvard Law School and US Treasury to reflect expenditure that takes place through the tax system.

Base-Broadening Strategies:
- Eliminate existing deductions, exclusions, and credits (DECs)
  - Eliminate exclusion for employer provided health insurance (EHI)
  - Eliminate exclusion for interest on state & local government bonds
  - Eliminate mortgage interest deduction
- Retain but limit existing DECs
  - Cap on total deductions and/or exclusions
  - Income-related clawbacks
  - Cap on total tax reduction from DECs (Feldstein/MacGuineas Plan)
  - Limit rate for deductions to less than taxpayer’s marginal rate

Analyzing Alternative Strategies:
- How would each reform affect incentives, including those for labor supply and other times of income generation?
- How responsive are taxpayer choices to variation in the after-tax price of various activities (health insurance, housing, charity)?
- What are the efficiency costs of tax deductions and exclusions?
## Tax Expenditures with Largest Revenue Cost, FY2013 ($Billion)

<table>
<thead>
<tr>
<th>Description</th>
<th>Amount ($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employer-Provided Health Insurance</td>
<td>147.8</td>
</tr>
<tr>
<td>Pension Contributions &amp; Earnings</td>
<td>147.0</td>
</tr>
<tr>
<td>Reduced Tax Rates on Capital Gains &amp; Dividends</td>
<td>110.4</td>
</tr>
<tr>
<td>Mortgage Interest Deduction</td>
<td>89.6</td>
</tr>
<tr>
<td>Earned Income Credit</td>
<td>58.1</td>
</tr>
<tr>
<td>Charitable Giving</td>
<td>46.9</td>
</tr>
<tr>
<td>Carry-over Basis on Capital Gains at Death</td>
<td>43.9</td>
</tr>
<tr>
<td>State/Local Income Taxes</td>
<td>27.8</td>
</tr>
<tr>
<td>Capital Gains Exclusion for Principal Residences</td>
<td>26.1</td>
</tr>
<tr>
<td>Child Tax Credit</td>
<td>25.7</td>
</tr>
<tr>
<td>State/Local Property Taxes</td>
<td>22.8</td>
</tr>
</tbody>
</table>


Note: Some tax expenditures involve changes in the timing of tax payments: pension contributions, carry-over basis for capital gains.
Reforms that affect the “after-tax price” of various activities currently associated with deductions, exclusions, and credits. In addition, some tax expenditure reforms could affect incentives for earning income. Consider:

- How does reform affect $\frac{d(Tax\ Liability)}{d(Earnings)}$?
- How does reform affect marginal price of goods purchased with after-tax earnings?

Reforms that do not raise the marginal tax rate on earnings:

- Dollar cap on total value of deductions, or deductions and exclusions. This changes the after-tax price of the affected goods – those that were deductible or were excluded – but changes the tax rate that applies to various sources of income only to the extent that the deduction cap, by raising the amount of taxable income that corresponds to any amount of pre-tax earnings, leads to a higher marginal tax rate on earnings.
- Disallowing specific deductions or exclusions: Same as for dollar cap.
Reform that does raise the marginal tax rate on earnings: Phase-out of eligibility for deductions.

Let $D =$ total value of deductions for a taxpayer, and “traditional” tax liability:

$$T = T(Y - D)$$  \hspace{1cm} (1)

without income-related phase out. Now tax liability with phase-out floor is

$$T = T(Y - \min[D, D - 0.03*\max\{Y-F, 0\}])$$ \hspace{1cm} (2)

$F$ is the income floor at which deductions begin to phase out.

“Traditional” case: marginal after-tax wage rate is

$$\frac{d[w*h - T(w*h - D)]}{dh} = w*(1-T'(Y-D)).$$ \hspace{1cm} (3)
With “Phase Out Above Floor”:

\[
d[w*h - T(w*h - \min[D, D - 0.03*\max\{w*h-F, 0\}])]/dh \\
= w*(1-T'(w*h-D)) \text{ if } w*h < F \\
= w*\{1-T'(w*h-D+.03*[w*h-F])*(1.03)\} \text{ if } w*h \geq F. \quad (4)
\]

In the phase-out case the effective marginal tax rate rises by a factor of 1.03 relative to the no-phase out case.

There are even reforms that could lower the marginal tax rate on earnings and other sources of income. One example is the Feldstein-MacGuineas (FM) plan, which limits tax savings as a percentage of taxable income. This means that as a taxpayer's earnings or other income rises, the taxpayer is able to claim a larger amount of deductions and exclusions -- hence a reduction in the effective MTR.
To illustrate this point, imagine an FM-style plan in which each taxpayer is limited to a reduction of 5 percent of tax liability from deductions. (The 5 percent parameter is just illustrative.) In this case:

\[ T = \max \{ T(w*h - D), T(w*h) - (0.05/T'(w*h))*w*h \} \quad (2') \]

Now under the assumption that small changes in hours of work do not change the taxpayer’s marginal tax rate (\( T''(w*h) = 0 \)) the after-tax wage is:

\[
\frac{d[w*h - \max \{ T(w*h - D), T(w*h) - (0.05/T'(w*h))*w*h \}]}{dh} \\
= w*(1-T’(w*h - D)) \text{ if } T(w*h) - T(w*h - D) < (0.05*w*h)/T’(w*h) \\
= w*(1-T’(w*h)+(0.05/T’(w*h))) \text{ otherwise} \quad (3')
\]
Even tax expenditure reforms that don’t affect the marginal tax burden on earnings may affect the after-tax real product wage. As in the discussion of general equilibrium incidence effects assume that there are J consumption goods with producer prices $p_j$. Some goods are taxed at rates $t_j$; purchases of others qualify for a tax deduction at rate $T’$. The price index for the consumer is:

$$ q^* = \sum j s_j \ast p_j (1 + t_j - T’j \ast I_j \in D) $$

where $I_{j \in D}$ is an indicator variable that equals 1 when purchases of good $j$ qualify for a tax deduction. Since the real product wage in this case is $w(1-T’)/q^*$, when the tax base is broadened by eliminating all tax deductions, the price index rises and the real after tax product wage falls (ceteris paribus) by:

$$ \sum j s_j \ast p_j (1 + t_j) / \sum j s_j \ast p_j (1 + t_j - T’j \ast I_j \in D) - 1. $$

As in the Goulder-Williams analysis of excise taxes above, the resulting change in the real product wage can affect labor supply. This underscores the importance of computing how a change in tax expenditures may affect the prices of various commodities and thereby cause efficiency losses.
Tax Exclusion for Employer-Provided Health Insurance (EHI):

Largest tax expenditure, both income and consumption tax bases would include it. Rationale: encourages provision of group insurance in the workplace, thereby correcting a market failure due to asymmetric information. Open question: Would EHI still dominate even without the tax subsidy? There is limited evidence on the extensive margin, whether firms would change their offer of EHI if the price changed. Gruber and Lettau (2004) find virtually no elasticity of insurance offering with respect to the tax price for large and medium-sized firms, substantial (price elasticity = -0.69) sensitivity for small firms.

What about the intensive margin? How does the “generosity” of insurance offered vary with the tax treatment? Gruber and Lettau (2004) estimate the price elasticity of demand for insurance on the intensive margin is about -0.7.

How large is the tax subsidy? Consider the choice between one dollar of cash wages and one dollar of employer-provided health insurance. Firm cost of one dollar of cash wage payment:

\[ 1 + \tau_{ss} + \tau_{mc} \]  

(7)
where $\tau_{ss}$ denotes the employer share of the payroll tax for Social Security (currently 6.2% on wages of less than $110,100), and $\tau_{mc}$ denotes the employer share of the Medicare payroll tax (currently 1.45% of wages).

After-tax income to individual when firm pays $1 of gross wages:

$$1 - \tau_{fed} - \tau_{state} - \tau_{ss} - \tau_{mc}$$

(8)

where $\tau_{fed}$ denotes the taxpayer’s marginal federal income tax rate, $\tau_{state}$ the corresponding state income tax rate, reduced if the taxpayer itemizes for federal income tax purposes by the marginal federal rate at which deductions can be claimed, $\tau_{ss}$ is the employee’s payroll tax rate for Social Security (usually 6.2%, but 4.2% in 2011 and 2012), and $\tau_{mc}$ denotes the employee’s Medicare payroll tax (1.45% in most cases).

If the employer devotes one dollar to the provision of EHI, the employee receives an insurance policy that costs one dollar. If the employer devotes the same dollar to providing wage income, the employee receives:

After-tax price of EHI: 

$$\frac{(1 - \tau_{fed} - \tau_{state} - \tau_{ss} - \tau_{mc})}{(1 + \tau_{ss} + \tau_{mc})}.$$  

(9)
For plausible values of the income tax parameters, even for a low-bracket taxpayer, facing $\tau_{\text{fed}} = 0.15$ and $\tau_{\text{state}} = 0.04$, the after-tax price of wage income is 0.681 and the tax “wedge” is 0.319. For a high-bracket individual, someone for whom $\tau_{\text{fed}} = .35$, the Social Security payroll tax is not likely to apply, but the increase in the federal marginal tax rate is enough to generate a larger tax wedge: 0.413.

Note an important assumption here is how wages would respond to changes in net insurance costs for firms (evidence suggests complete pass through, although perhaps not for each individual: there may be grouping in pass through, i.e. young workers, older workers). Jon Gruber studied the pass-through of costs for a new state-imposed mandate for maternity coverage in California; workers in age groups for whom the mandate was likely to increase employer insurance costs experienced lower wages as a result of the mandate.
Gruber (2011) estimates the distribution of tax benefits associated with the income and payroll tax exclusions for EHI. These benefits are highly skewed toward higher incomes.

<table>
<thead>
<tr>
<th>Income Decile</th>
<th>Tax Saving Amount ($Billion 2009)</th>
<th>Percentage of Tax Saving</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (lowest)</td>
<td>$2.0B</td>
<td>1%</td>
</tr>
<tr>
<td>2</td>
<td>2.0B</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>$6.0B</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>$13.0B</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>$20.0B</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>$26.0B</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>$34.0B</td>
<td>13</td>
</tr>
<tr>
<td>8</td>
<td>$42.0B</td>
<td>16</td>
</tr>
<tr>
<td>9</td>
<td>$55.0B</td>
<td>21</td>
</tr>
<tr>
<td>10 (highest)</td>
<td>$62.0</td>
<td>24</td>
</tr>
</tbody>
</table>

Source: Gruber (2011 National Tax Journal).

Why this pattern? Greater coverage, higher MTR at higher incomes.
What about reform options such as capping the amount of employer provided insurance that can be excluded?

Gruber (2011) uses a micro-simulation model to evaluate reforms to current tax rules that are less extreme than complete inclusion of EHI in the federal individual income tax base. Complete income tax repeal in his model yields $157B. By comparison, capping the exclusion at the median employer plus employee contribution has a more modest revenue effect: $29B. Capping the employer contribution alone is estimated to raise $22B. Gruber estimates that repealing the exclusion for EHI would reduce the number of individuals covered by employer-provided health insurance by 15 million (relative to 152 million who are covered under the status quo) and reduce employer spending by $172B relative to a baseline of $563B. Capping the exclusion would reduce the number covered by 1 million, and reduce employer spending on EHI by $17B.
Tax Rate Endogeneity and the Estimation of Behavioral Responses

A general problem of estimating the distortions created by tax expenditures is developing credible estimates of the key elasticities of supply and demand. One of the central challenges is the endogeneity of “tax prices” in standard demand models. Another problem is separating temporary and permanent effects of tax rate changes, either when the tax code changes or when individuals experience time-varying tax rates. The endogeneity problem can be illustrated for the case in which the tax code subsidies the purchase of a particular good, for example charitable giving. Assume the underlying demand model is log-linear:

\[
\ln G_i = \alpha + \beta \ln Y_i + \gamma \ln (1-\tau_i) + \delta X_i + \varepsilon_i \tag{10}
\]

The key parameter is the elasticity of demand, \( \gamma \). We would typically expect \( \gamma < 0 \). The marginal tax rate \( \tau_i = T_i'(Y_i - D_i - G_i) \) determines the price of giving; the tax function depends on gross income minus other deductions minus the deduction for charitable gifts. The problem is that \( \varepsilon_i \) is correlated with \( G_i \), which in the presence of a nonlinear tax schedule may be correlated with \( \tau_i \). Larger values of error term translate into larger deductions, hence (if tax schedule is progressive) lower marginal tax rate, hence larger value of \( (1-\tau_i) \). This induces a
spurious positive correlation between $G_i$ and $(1-\tau_i)$ leading to an upward bias in the estimates of $\gamma$. Since we expect $\gamma < 0$, this bias will tend to yield under-estimates of the price elasticity of demand.

How do we solve this? (i) One strategy is to use variation other than $G_i$ to generate variation in $\tau_i$. This is done by constructing the "first dollar marginal tax rate" $\tau_i^* = T_i'(Y_i - D_i)$ for all taxpayers. Note $\tau_i^*$ is correlated with $\tau_i$ but it is NOT affected by the spurious correlation channel noted above. Some studies estimate reduced form regressions replacing $\tau_i$ with $\tau_i^*$ in the regression equation; a better strategy is to use $(1-\tau_i^*)$ as an instrumental variable for $(1-\tau_i)$ uses IV. (ii) The alternative strategy is to find a change in tax structure that can be combined with panel data on taxpayers to estimate the effect of marginal tax rates on charitable giving. This strategy in effect uses the change in marginal tax rates due to the reform as an instrumental variable for the change in the taxpayer’s observed marginal tax rate.
One example of a study that uses the second approach is Randolph (1995 JPE), “Dynamic Income, Progressive Taxes, and the Timing of Charitable Contributions." This study estimates an Almost Ideal Demand System (AIDS) with current and future income, current and future tax price variables to explain charitable giving; dependent variable is share of income devoted to charitable gifts.

Sample of 12000 taxpayers, six years of panel data (1979, 80, 83, 84, 85, 88). Spans significant change in marginal tax rates (TRA86) so there is "transitory" tax rate variation. Few demographic variables on tax returns (married, number of exemptions, age (sometimes age > 65 dummy variable). Estimation sample: 51,146 returns.

<table>
<thead>
<tr>
<th>Elasticity Measure</th>
<th>Income</th>
<th>Tax Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;Current&quot; (no transitory/perm distinction)</td>
<td>0.82 (0.01)</td>
<td>-1.21 (0.07)</td>
</tr>
<tr>
<td>Transitory</td>
<td>0.58 (0.01)</td>
<td>-1.55 (0.06)</td>
</tr>
<tr>
<td>Permanent</td>
<td>1.14 (0.01)</td>
<td>-0.51 (0.06)</td>
</tr>
</tbody>
</table>
Lecture 4: Tax Treatment of Retirement Saving

Reasons for Studying Taxation and Saving:

1. Intertemporal choices are an important potential margin of distortion and several high-profile efficiency analyses suggest capital income should not be taxed: Atkinson-Stiglitz, Chamley-Judd. More recent work (Diamond-Banks) calls this into question. This relates to a long-standing debate on consumption vs. income taxation.

2. "Distribution tables" depend on tax treatment of capital income because it is highly skewed.

3. Potential growth linkage. High saving countries tend to be high growth countries. Taxes may affect saving rate and ultimately the steady-state level of capital. In steady state,

\[ sf(k) = (n + \delta)k \]  

where \( s = \) saving rate. This means steady state consumption is
\[ c = (1-s)f(k) = f(k) - (n + \delta)k. \quad (2) \]

This is maximized when \( f'(k) = (n + \delta) \). Taxing saving can affect \( s \), can move \( k \) toward or away from this "golden rule". Remember two caveats:

* To change the capital stock, the government has a powerful instrument - deficit policy - that does not operate primarily through creating incentives for private behavior.

* Saving = investment in closed economy, but this equivalence breaks down in the open economy.

4. Aging baby boomers have drawn interest to question of "retirement saving adequacy" and hence a focus on personal saving (not just national saving).
Role of Empirical Work in Studying Taxation and Household Saving

* Describe Stylized Facts about Household Saving Behavior

* Help Determine Which of Three Models (Lifecycle, Dynastic Altruism, Precautionary) is “Right”. (Are these distinct models? LCH can be augmented with precautionary demand for wealth or with bequest motives.)

* Calibrate Specific Models for Studying Behavioral Responses to Tax-Induced Changes in Rates of Return or Other Aspects of Saving Environment. Examples: Intertemporal Elasticity of Substitution (IES) Determines Distortion in Consumption Profile When Rate of Return Changes; Shape of “Marginal Utility of Bequest” Function Determines Response to an Estate Tax.
Saving Decisions Take Place in a Complex Institutional Environment

* In the “standard textbook model,” investors can earn rate of return \( r \), borrow and lend at the same rate.

* In reality, investors face different borrowing and lending rates, both before and after tax; the tax treatment of income from saving and investing differs depending on the particular asset the individual is holding; saving can take place in a “tax deferred account” (like IRA) or in a traditional taxable setting
Balance Sheet for the U.S. Household Sector, 2012:Q2

<table>
<thead>
<tr>
<th><strong>Assets</strong></th>
<th>$76.1 Trillion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real Estate</td>
<td>19.1</td>
</tr>
<tr>
<td>Other Tangible Assets</td>
<td>5.1</td>
</tr>
<tr>
<td>Financial Assets</td>
<td>51.9</td>
</tr>
<tr>
<td>Deposits</td>
<td>8.7</td>
</tr>
<tr>
<td>Taxable Bonds</td>
<td>3.0</td>
</tr>
<tr>
<td>Tax-Exempt Bonds</td>
<td>1.8</td>
</tr>
<tr>
<td>Corporate Stock &amp; Mutual Fund Shares</td>
<td>14.3</td>
</tr>
<tr>
<td>Noncorporate Business Equity</td>
<td>7.7</td>
</tr>
<tr>
<td>Pension Fund Reserves (incl. 401(k) &amp; IRA)</td>
<td>13.7</td>
</tr>
<tr>
<td>Other</td>
<td>2.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Liabilities</strong></th>
<th>13.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortgages</td>
<td>9.6</td>
</tr>
<tr>
<td>Other</td>
<td>3.9</td>
</tr>
</tbody>
</table>

| **Net Worth**                     | 62.7           |

Source: Federal Reserve Board, Flow of Funds Accounts of the United States Second Quarter 2012, Table B.100.
Margins of Distortion in Household Saving Behavior

* How much to save

* Asset allocation: Which assets to hold, how much to allocate to each (relates to puzzle of limited stock ownership; stocks vs. bonds, tax-exempt bonds vs. taxable bonds)

* Asset location: Which assets to hold in taxable accounts, which to hold in tax-deferred accounts

* Asset sale and purchase decisions: Trading decision is affected by capital gains taxation

* Leverage: How much to borrow, and in what form
Stylized Facts About Household Wealth Holdings (2007 SCF)

* Incomplete Portfolios: 91% of families have transaction accounts, 21% hold stocks outside retirement accounts, 50% hold retirement accounts, 94% have some financial assets, 86% own at least one car; 70% own a house

* Very Skewed: Top 1% hold about 50% of financial wealth, 40% of net worth including tangible assets; Top 10% about 80% of financial, 70% of total. Median family net worth $120,300, mean $556,300.

* Many Households Have No Wealth: 30% negative net financial wealth, 20% negative net worth

* Limited Liquid Wealth: Hall (2012): 58% of earnings to 74% of households with less than two months of earnings in liquid form

* Wealthy have proportionately less owner-occupied housing, more equity and "alternative asset classes." Business equity is > 1/3 of wealth for top 1%.

* Inheritance accounts for substantial share of household wealth (half?)
The Complexity of Capital Income Taxation

* Saving Accounts, CDs, Treasury bonds: Interest income, taxed at ordinary tax rate

* Stocks: Dividends (taxed in recent years at 15%, 20% in 2013 +) and capital gains (taxed at realization at 15% if long term (> 1 year) in 2012, 20% starting in 2013)

* Tax-Exempt Bonds: Untaxed but face an "implicit tax" through lower return in market

* Equity Mutual Funds: Dividends taxed like stocks; capital gains taxed as realized by the fund (not the investor)

* Tax-Deferred Accounts (non-Roth IRAs, 401(k)s): Pre-tax contributions, returns untaxed until funds are withdrawn from the account

* Tax-Deferred Accounts (Roth): After-tax contributions, no further tax on return or withdrawals
Taxes and Saving: Standard Theory Offers Ambiguous Prediction

– Substitution Effect Makes Future Consumption More Expensive So Increases Current C

– Income Effect (two period model with endowment given in first period) Household is Poorer (rise in price of future C) so Current C Should Decline

– “Human Wealth Effect” – PDV of Future Labor Income or Other Receipts Rises Which can Increase Current Consumption
Are Households Target Savers? Role for Precautionary Saving

* Substantial Risk of Late-Life Expenses (CBO projections for 65-year-olds in 2010: 45% will use a nursing home, 25% for at least a year. Average nursing home costs: $187/day for semi-private room, $209/day for private room.)

* Hubbard/Skinner/Zeldes (JPE 1994) model precautionary savings demand and transfer programs. Key insight: wealth-tested transfer programs can provide a strong disincentive for low-income households to save. They offer: i) a potential explanation for low levels of saving observed for many households; ii) an investigation of how social insurance programs affect saving; iii) explicit modeling of uncertainty that may affect households
Fig. 2.—Optimal consumption with a welfare "floor"

Households maximize:

\[ V = \sum D_t (1+\delta)^{-t} C_t^{\gamma}/\gamma \]  

(3)

\( D_t \) = probability of survival to year \( t \).

\( A_t \) = assets in period \( t \)

\[ A_t = A_{t-1}(1+r) + E_t + T(E_t, M_t, A_{t-1}(1+r)) - M_t - C_t \]  

(4)

\[ T = \max \{0, C_{\text{floor}} + M_t - A_{t-1}(1+r) - E_t\} \]  

(5)

\( C_{\text{floor}} \) is a consumption floor set by government transfer programs (Medicaid, food stamps, public housing)

Two stochastic shocks: \( M_t \) and \( E_t \). Key question: How persistent are the shocks.

Findings: saving nothing can be optimal for some low-income households with high likelihood of relying on transfer programs.
Economics of Tax-Deferred Accounts

1. How do these accounts work? How important are they? Do they transform the income tax into a consumption tax for many households? (“hybrid tax”)

2. Does the availability of these accounts raise personal saving? Does it raise national saving? (Substitution is the key question)

3. How does the structure of these accounts affect saving decisions?
Basic Structure of Taxable vs. Tax-Deferred Accounts

Three relevant tax rates: $\tau_0$ while earn income that might be saved, $\tau_1$ while earning investment returns, $\tau_2$ when withdrawing assets to finance consumption. Nomenclature: can use “T” for taxable, “E” for exempt in each period:

<table>
<thead>
<tr>
<th>Type of Account</th>
<th>When Wages Earned</th>
<th>When Balance Grows</th>
<th>When Account Balance Drawn Out</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional Savings</td>
<td>T</td>
<td>T</td>
<td>E</td>
</tr>
<tr>
<td>Traditional IRA</td>
<td>E</td>
<td>E</td>
<td>T</td>
</tr>
<tr>
<td>Roth IRA</td>
<td>T</td>
<td>E</td>
<td>E</td>
</tr>
</tbody>
</table>
Four Types of Accounts:

Traditional taxable account:

\[ V_{\text{taxable},T} = (1-\tau_0) e^{(1-\tau_1) r_T} \]  \hspace{1cm} (6)

“Traditional” Individual Retirement Account (IRA) or 401(k) Plan

\[ V_{\text{IRA},T} = (1-\tau_2) e^{r_T} \]  \hspace{1cm} (7)

“Roth” IRA or 401(k)

\[ V_{\text{Roth IRA},T} = (1-\tau_0) e^{r_T} \]  \hspace{1cm} (8)

“Nondeductible” IRA (when income > threshold for traditional deductible IRA)

\[ V_{\text{Non-deductible IRA},T} = (1-\tau_0) e^{r_T} - \tau_2 [(1-\tau_0) e^{r_T} - (1-\tau_0)] \]  \hspace{1cm} (9)
Budget Set with Restricted Contribution to IRA:

slope = 1+r(1-τ)

slope = 1+r
Accumulation in Traditional & Tax-Deferred Saving Accounts

Case 1: \( r=0.06, t = 0.33 \)

<table>
<thead>
<tr>
<th>Account Type</th>
<th>10 Years</th>
<th>30 Years</th>
<th>50 Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taxable</td>
<td>1.00</td>
<td>2.24</td>
<td>5.00</td>
</tr>
<tr>
<td>Deductible or Roth IRA</td>
<td>1.22</td>
<td>4.05</td>
<td>13.46</td>
</tr>
<tr>
<td>Non-Deductible IRA</td>
<td>1.04</td>
<td>2.94</td>
<td>9.24</td>
</tr>
</tbody>
</table>

Case 2: \( r=0.02, t = 0.33 \)

<table>
<thead>
<tr>
<th>Account Type</th>
<th>10 Years</th>
<th>30 Years</th>
<th>50 Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taxable</td>
<td>0.77</td>
<td>1.00</td>
<td>1.31</td>
</tr>
<tr>
<td>Deductible or Roth IRA</td>
<td>0.82</td>
<td>1.22</td>
<td>1.82</td>
</tr>
<tr>
<td>Non-Deductible IRA</td>
<td>0.77</td>
<td>1.04</td>
<td>1.44</td>
</tr>
</tbody>
</table>
The Institutional Details (background reading!)

- **Traditional “Deductible” IRA**
  - Fully deductible contributions for incomes below $58K (single), $92K (married joint) in 2012
  - Partial Phase-out of deductibility (58-68K, 92-112K)
  - No tax on income accruing within IRA account
  - Fully taxable as ordinary income when withdrawn
  - “Penalty Tax” of 10% if withdrawn before 59 ½
  - Contribution limit: $5500 plus $1000 if over 50 (“catch up”)
  - Required Minimum Distributions (RMDs) if owner is over 70 ½
  - Balance from a pension account can be “rolled over” to an IRA when retire or leave employment

- **Roth IRA**
  - No deduction for contributions, no taxation of withdrawals
  - $5500 (+1000) contribution limit but in after tax dollars (so like contribution $5500/(1-\tau_0) dollars to a regular IRA)
  - No restrictions on withdrawals while contributor is alive; RMDs apply after death
• 401(k) plans
  - Employer Sponsored Plans – key difference from IRAs
  - Tax-deductible contributions (although there are now Roth-401(k) variants at some firms)
  - Plans often include employer match so value at withdrawal is $V_{401(k),T} = (1+m)\cdot(1-\tau_T)\cdot e^{rT}$ where $m =$ employer match rate
  - Withdrawal rules similar to IRAs; RMDs after age 70 ½
  - Contribution limits much higher than IRAs: $17,000$ in 2012 plus $5000$ catch-up if over 50
  - No phase-outs with income
  - “Hardship withdrawals” if need assets while still working; also loan provisions

Roll-over rules: If income is < $100K owner can convert Regular IRA to Roth IRA (note special 2010-12 provision: no limit on income for conversion)
Operation of 401(k)s and IRAs in U.S.

* 1980 to 2011: From 3/4 of pension contributions to defined benefit (DB) plans to 3/4 to defined contribution (DC) plans like 401(k)s

* DC Plan and IRA assets in 2011: $9.4T ($17.9T total retirement assets)

* Some future retirees will have lifetime exposure to 401(k)s

* Accumulation potential: a married couple contributing 8% of salary for 30 Years, with 50/50 allocation and historical (pre-2008) equity returns, median Balance at 65 for median earner is $468,000; 25th percentile: $289,000; 75th percentile $706,000

*Actual Balances: SIPP Data

<table>
<thead>
<tr>
<th>Year</th>
<th>Mean</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999</td>
<td>$66,660</td>
<td>$24,844</td>
</tr>
<tr>
<td>2003</td>
<td>$80,592</td>
<td>$43,127</td>
</tr>
<tr>
<td>2007</td>
<td>$137,430</td>
<td>$76,946</td>
</tr>
</tbody>
</table>
Retirement Accounts and U.S. Households, 2011:

* 31%: neither employer-sponsored retirement plan nor IRA
* 8%: IRA, no employer-sponsored plan
* 30%: only employer-sponsored plan, no IRA
* 31%: employer sponsored plan and IRA

Effect of IRA & 401(k) Eligibility on Wealth Accumulation

Earliest Studies of IRAs

- Discovered that Many Households Had Very Little Financial Wealth So Little Opportunity for Substitution
- 1986 SIPP Data (Venti & Wise): Contributors with IRA Assets of $7000 (median) have Non-IRA assets of $13,500; Non-Contributors Medial Non-IRA Financial Assets of $1000.
- Conflicting Evidence on IRAs (but little cross-sectional variation in eligibility for IRAs)
- Emphasize Difference Between Limit Contributors and Those Contributing Less than the Limit Amount

Margins for 401(k) or IRA Accumulation to Crowd Out Other Wealth:

- Non-IRA, Non-401(k) Financial Assets
- Other Pension Assets (Defined Benefit Plans)
- Housing Equity (Borrowing Against Home to Fund 401(k) Plan)
The 401(k) “Eligibility Experiment”

- Since firms choose whether to offer 401(k) plans, eligibility varies across households
- What explains decision to offer 401(k)?
  - Historical firm provision of profit-sharing plan
  - Median voter outcome reflecting preferences of workers at the firm
  - Do firms with 401(k)s reduce availability of other benefits?
  - Firm age, composition of workforce - younger firms, more 401(k)s
  - Worker screening device: does desire to work at a 401(k)-employer signal “low discount rate” worker?
- Exogeneity of 401(k) eligibility: not a randomized trial, but not like universal-eligibility IRAs
- Participation conditional on eligibility: 36% in bottom decile, 65% at median, 85% at top decile

Basic specification (Poterba/Venti/Wise *JPubE* 1995 and subsequent studies) on repeated cross-sections with varying 401(k) eligibility

\[ A_{a,i} = \alpha_a + X_i \beta_a + E_i \gamma_a + u_{a,i} \] (10)
Allow for $\beta_a$ and $\gamma_a$ (already asset-specific) to vary by income of the household head. Thus the “eligibility effects” associated with $E_i$ are different for high and low incomes.

### Estimated Eligibility Effects ($\gamma_a$): SIPP, Total Family Financial Assets

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 10</td>
<td>1581</td>
<td>2061</td>
<td>1378</td>
<td>2033</td>
</tr>
<tr>
<td>10-20</td>
<td>1902</td>
<td>2404</td>
<td>1997</td>
<td>4045</td>
</tr>
<tr>
<td>20-30</td>
<td>2624</td>
<td>4206</td>
<td>2558</td>
<td>5499</td>
</tr>
<tr>
<td>30-40</td>
<td>4605</td>
<td>9062</td>
<td>3256</td>
<td>8683</td>
</tr>
<tr>
<td>40-50</td>
<td>6726</td>
<td>12588</td>
<td>6206</td>
<td>14470</td>
</tr>
<tr>
<td>50-75</td>
<td>14108</td>
<td>24384</td>
<td>10080</td>
<td>26093</td>
</tr>
<tr>
<td>&gt; 75</td>
<td>30971</td>
<td>57348</td>
<td>29842</td>
<td>51080</td>
</tr>
</tbody>
</table>

Subsequent Research Focuses on Addressing Potential Endogeneity and Heterogeneity of Households:

1. Propensity Score Methods (Dan Benjamin (JPubEc 2003)
   - SIPP 1990, 25-64, not self-employed
   - Group “like households” based on "propensity scores" (estimated probability of treatment given covariates) rather than observed attributes
   - Eligibles are more likely to own homes, have a DB plan, be married, have two earners, have higher income

Estimated Effect of 401(k) Eligibility on Household Wealth

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Households</td>
<td>3434</td>
<td>2738</td>
</tr>
<tr>
<td></td>
<td>(2444)</td>
<td>(576)</td>
</tr>
<tr>
<td>Households w/Wealth &lt;100K</td>
<td>1795</td>
<td>1818</td>
</tr>
<tr>
<td></td>
<td>(640)</td>
<td>(424)</td>
</tr>
</tbody>
</table>
   • Instrument for 401(k) participation using 401(k) eligibility
   • Allow Flexible Effects at Different Income Levels
   • Instruments for 401(k) Participation Using Eligibility
   • Cannot Reject Zero Effect of 401(k) Participation at Lowest Income Levels, But Positive Effect of Participation at Higher Income Levels
   • Evidence of Heterogeneity within Most Income Groups (but not highest)
   • Larger Impact of 401(k) contributions on total financial assets for low- and moderate-income households

3. "Active vs. Passive Savers": Chetty, Friedman, Leth-Petersen, Nielsen, Olsen
   NBER WP 18565 2012
   • Use Danish tax records data that include wealth data so can track flow contributions to retirement plans and wealth accumulation
   • Study employer changes that shift amount being contributed to pension plan: very little evidence of household saving response
   • Also examine changes in tax incentives for contributing to particular accounts: only about 15% of taxpayers seem to respond to tax subsidies
What Explains 401(k) Contributions?

Variables with Some Predictive Power
- Financial Sophistication of Participant (Education as Proxy)
- Employer Match
- “Behavioral” Factors

- Firm that shifted from “opt-in” to “opt-out” structure of 401(k) plan. No changes to budget constraint facing employees.
- Participation Rate in 401(k) Before Opt-Out Plan: 57% at start of employment, 64% after 3-5 years, 83% for 20+ year employees
- Participation Rate After Opt-Out: 86% for new employees same tenure mix
- Why is this finding so important: Saving is a first-order decision for households (compare “book of the month club”) and it appears to be sensitive to framing and other considerations
Changing Evolution of Default Policies
- Initially Money Market Funds (no risk for employer – can’t lose money)
- Now “Target Date Funds” that focus on automatic age-related shifts in equity exposure
- Key Role of “Safe Harbor” Provisions in Allowing Employers to Offer Default Allocations that Involve Some Risk

Other Applications of Behavioral Economics to Studying Saving Choices:
- Failure of households to take advantage of match rates even when they can withdraw immediately
- Small number of rebalancing transactions for most households: Samuelson/Zeckhauser: Median lifetime rebalancing transactions at TIAA-CREF is zero
- Important social learning effects (Duflo/Saez QJE: study librarians and their decisions with regard to 401(k) plan – if existing workers in “social group” contribute more, new workers do, too)
Behavioral Economics in Action: Designing Default Programs

- Benartzi - Thaler Save More Tomorrow (SMART) plan
- Generic default plan: How should the default contribution rate and asset allocation be set? Do participants assume that the default allocation has been set because it is "optimal"?
- Should the default include some equity exposure? Too safe a default will fail to build wealth, too risky exposes employers to potential liability and may discourage contributors
- Welfare analysis of "X% of Salary" default rule:
  * Some who would not contribute at all now contribute X%
  * Some who would have contributed less than (more than) X% now contribute X%
  * Some who would have made different asset allocations now choose the default allocation
  * Welfare depends on elasticities of participation, contribution level with respect to default, and on distribution of individuals pre-default across different contribution levels
Lecture 5: Asset Markets and Incidence

Many taxes are levied on durable assets (houses, physical capital such as buildings and equipment, patents, natural resource stocks). These are traded assets with prices - changes in current and future taxes can affect them.

Asset value changes are a key component of incidence, contributing to gains and losses for existing asset holders. Our goal is to compute the price change by analyzing the present discounted value of future tax payments and the future rental income stream associated with that asset.

Best illustrations are in markets with long-lived assets: corporate capital, land, housing are examples. Focus on owner-occupied homes following Poterba QJE 1984.

Equilibrium condition in housing market:

Value of asset services per unit of housing: $R(H)$

$R'(H) < 0$ reflects diminishing value of housing capital
Period-to-period equilibrium requires that $R(H)$ equal the investor cost of holding the asset for one period. This cost includes the depreciation of the asset (exponential at rate $\delta$ per period) and the opportunity cost of funds. We denote this as $(1-\tau)r$: the after-tax interest rate. There might also be a risk premium (we’ll set this to zero).

Investor cost also includes any capital gain or loss to holding the asset. A capital gain reduces the investor cost of holding the asset; a capital loss increases it.

$q_{H,t} = \text{asset price at the start of period } t$

$q_{H,t+1} - q_{H,t} = \text{capital gain or loss during period } t$

Equilibrium Condition:

$$R(H_t) = q_{H,t} (r(1-\tau)+\delta) - (q_{H,t+1} - q_{H,t}).$$

(1)
Consider a tax on each house that takes the form of a required payment $T_t$ in period $t$. Because houses are long-lived assets, even taxes that will not be levied until future years can depress prices today. In each period when the tax is levied, the equilibrium condition becomes

$$R(H_t) - T_t = q_{H,t} \left(1 + r(1 - t) + \delta\right) - q_{H,t+1}. \quad (2)$$

In general, $q_{H,t+1}$ is unknown at time period $t$. Assume perfect foresight; one could also embed this analysis in a stochastic model of price determination.

To find the $q_H$ from (2), solve forward by rewriting (2) as:

$$q_{H,t} = \frac{R(H_t) - T_t + q_{H,t+1}}{1 + r(1 - t) + \delta}. \quad (3)$$

Solving recursively by substituting for $q_{H,t+1}$ yields, after multiple substitutions,
\[ q_{H,t} = \sum_{i=0}^{S} \frac{R(H_{t+i}) - T_{t+i}}{(1 + r(1 - \tau) + \delta)^{i+1}} + \frac{q_{H,t+S}}{(1 + r(1 - \tau) + \delta)^{S+1}}. \]  

(4)

We impose a transversality condition to rule out an “exploding” asset price:

\[ \lim_{S \to \infty} \frac{q_{H,t + S}}{(1 + r(1 - \tau) + \delta)^{S+1}} = 0. \]  

(5)

With this condition we can see how a stream of tax liabilities \( \{T_t\} \) will affect the price of houses:

\[ q_{H,t} = \sum_{i=0}^{\infty} \frac{R(H_{t+i})}{(1 + r(1 - \tau) + \delta)^{i+1}} - \sum_{i=0}^{\infty} \frac{T_{t+i}}{(1 + r(1 - \tau) + \delta)^{i+1}}. \]  

(6)

The second term is the present discounted value of current and future tax payments.

If the stock of housing is fixed, so \( H_{t+i} = H_t \) for all \( i \), then from (6) we can determine \( dq_{H,t}/dT_{t+i} \). When the housing stock is endogenous, however, changes
in future tax policies will also affect current and future investment, hence \{H_t\}. These changes will generally offset the effect of taxes on house prices. When taxes rise, thereby depressing prices, housing construction will decline. That will reduce the future stock of housing capital, thereby raising the rental value of a unit of housing services, thereby helping to raise prices.

To formalize this we need to model the supply function for new construction:

\[
I_t = \psi(q_{H,t})
\]  \hspace{1cm} (7)

Where \(I_t\) denotes gross construction of new housing. The net change in the housing stock is given by

\[
H_{t+1} - H_t = \psi(q_{H,t}) - \delta H_t.
\]  \hspace{1cm} (8)

The corresponding equation for the evolution of house prices from (2) is

\[
q_{H,t+1} - q_{H,t} = (r(1 - \tau) + \delta)q_{H,t} - R(H_t) + T_t.
\]  \hspace{1cm} (9)
Equations (8) and (9) define a two-equation system of difference equations in two variables: \((q_H, H)\). To analyze how the value of \(q_{H,t}\) responds to a shock to \(\{T_t\}\), we analyze the stability properties of this system of difference equations using phase diagram methods.

The figure below shows the loci on which \(q_H\) and \(H\) are respectively constant. The steady state is defined by:

\[
\psi(q_H) = \delta H \quad \text{and} \quad (r(1-\tau)+\delta)q_H = R(H) - T. \tag{10}
\]

When the system is out of equilibrium, for example when a tax shock moves the \((dq_H/dt)\) curve, there is a unique stable path that leads to the equilibrium point; conditional on a value of \(H\), there is only one value of \(q_H\) that will result in the system evolving back to the equilibrium.
One can link this diagram to the basic partial equilibrium incidence diagram we have used before. Note that even when the $dH/dt = 0$ locus is horizontal, so there is a fixed long-run supply price of housing, a leftward shift in the $dq_H/dt = 0_{\text{initial}}$ locus induced for example by a new tax on housing would create a transitory decline in house prices that would burden existing house owners.
To illustrate how this framework can be used to evaluate changes to housing tax policy, consider the user cost of owner-occupied housing in the current tax law:

\[
c = (1 - \tau_y) \left( \lambda r_M + (1 - \lambda) r_{Alt} + \beta \right) + m + (1 - \tau_y) \tau_{prop} - \pi_e
\]  

(11)

where \( \tau_y \) is the marginal income tax rate for mortgage interest and property tax deductions as well as investment income, \( r_M \) is the mortgage interest rate, \( \lambda \) is the home's loan-to-value ratio, \( r_{Alt} \) is the return on the alternative assets that the household might invest in if not using equity for a house, \( \beta \) denotes the pre-tax housing risk premium, \( m \) is the combined cost of depreciation and maintenance, \( \tau_{prop} \) is the property tax rate, and \( \pi_e \) is the expected rate of nominal house price appreciation. Plausible parameter values for 2012 might be: \( \tau_y = 0.25, r_M = r_{Alt} = 0.04, \lambda = 0.75, \beta = m = 0.02, \tau_{prop} = 0.015, \pi_e = 0.02 \). This implies \( c = 0.05625 \).

What if we eliminate the federal income tax deduction for the property tax? Now \( c' = 0.06 \). This is an increase of 6.7 percent; with fixed \( H \) and a price elasticity of demand of -1, house prices would drop by 6.7 percent. A calibrated rational expectations model suggests -3.7 percent, indicating the importance of future changes in asset supply and associated rent changes.
Role of Empirical Work in Studying Asset Market Incidence:

(i) document importance of asset price adjustments;
(ii) estimate elasticity of demand for the durable asset (this is calibrating R(H));
(iii) estimate supply curve $\psi(q_H)$

Proposition 13 in California: Property Tax Reform (K. Rosen, JPE 1982)

<table>
<thead>
<tr>
<th>Community</th>
<th>Property Tax Rate FY1978</th>
<th>Property Tax Rate FY1979</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berkeley</td>
<td>13.75</td>
<td>5.04</td>
</tr>
<tr>
<td>Pleasanton</td>
<td>12.83</td>
<td>5.60</td>
</tr>
<tr>
<td>Novato</td>
<td>10.62</td>
<td>4.81</td>
</tr>
<tr>
<td>Menlo Park</td>
<td>8.64</td>
<td>4.09</td>
</tr>
<tr>
<td>Atherton</td>
<td>8.81</td>
<td>4.14</td>
</tr>
</tbody>
</table>

\[ \Delta \text{House Price} = 7.28 \times \text{Tax Saving} + \text{Home Attributes} \times \beta \]  \hspace{1cm} (12)
Related study: David Cutler (AER 1988) on impact of TRA86 on corporations (variation came from types of assets firms used and changes in tax treatment, level of debt - since tax rate was reduced, and related factors). Modest evidence that firms whose tax burdens would rise saw declines in share prices.

General Approach: "Event Study Methodology"

Return on Security $i$ at time $t$:

**CAPM:**

$$R_{it} = \alpha_i + \beta_i * R_{M,t} + \varepsilon_{it} \quad (13)$$

**Fama-French Factors:**

$$R_{it} = \alpha_i + \sum \beta_{ij} * R_{j,t} + \varepsilon_{it} \quad (14)$$

**ADD:** Variable capturing news of policy change. Let $I_t$ denote indicator variable for days on which the probability of tax reform rises.

Key appeal of asset markets: well-measured variables that may capture some information on impact of taxation.

Key challenge: measuring change in expected policy outcome.
Some Useful Additional Resources for Tax Economists:

Urban-Brookings Tax Policy Center
(http://www.taxpolicycenter.org/)

Congressional Budget Office
(www.cbo.gov)

Emmanuel Saez lecture notes at Berkeley
(http://elsa.berkeley.edu/~saez/course/course.html)

Raj Chetty’s on-line lecture series at Harvard
(http://obs.rc.fas.harvard.edu/chetty/public_lecs.html)

Handbook of Public Economics: Volumes 1-4 in print, volume 5 presentations at Burch Center website, UC-Berkeley:
(http://elsa.berkeley.edu/~burch/confhand.htm)