Labor Supply
Mostly drawn from Keane and Rogerson (2011 NBER)

Cullen Roberts

Chicago Economics

February 25, 2013
Outline

1. Introduction
   - Main Points
   - Setting - related literature

2. Old Labor Supply Literature
   - Structural Model and Identification
   - Results
   - Criticism

3. New Structural Labor Supply Literature
   - Human Capital
   - Borrowing Constraints
   - Optimization Frictions - Raj Chetty
   - Extensive Margin

4. Non-structural labor supply and related work
   - Taxable Income Elasticity
   - Marginal Tax Rates?
What is an elasticity? (pt 1)

\[ \ln h = \alpha + \beta \ln w + \varepsilon \]

Running this regression, \( \beta \) would be an elasticity. Facile. (TIE)
- Labor supply is affected by: timing of the tax, how the tax money, tax progressivity, etc... (Shimer Fall 2011)
- Without a model, regressions estimates are of limited utility.

Structural Approach: **Structural Parameters** are those invariant over some counterfactuals, such as different policy environments
- With the model, they can be used for counterfactual analysis
- Finding models and their parameters is how one *understands* the economy.
- Structure also warns of endogeneity and suggests instruments
What is an elasticity? (pt 2)

- In some models, a parameter or set of parameters will map to the a specific elasticity.
  - But changing the model slightly will change how those parameters map onto the elasticity.
  - Be careful about calling the parameter the elasticity

- **Labor supply response** can vary over the life cycle (models with human capital).
Who Cares?

Labor supply matters because of

- Deadweight loss from taxation
- Changes in income distribution due to taxation
- Business Cycles - unemployment, hours worked, output
  - Calibrating macro models more generally (warning!)
- Relevant for macro theories: balanced growth requires income effect offsets substitution effect.
- Human capital accumulation (related to DWL and changes in income)
1980’s Foundational Labor Supply Literature

Three big papers:
- Browning, Deaton, and Irish (1985)
- Altonji (1986)

Except for superficial details, do not differ in model estimated but in identification strategy.

“Prime-age” males – only intensive margin. (verify age range)
- Not considering female labor supply – Heckman and MaCurdy (1982) – where extensive margin and selection crucial
Benchmark Model (1)

Agent maximizes

$$\sum_{a=0}^{T} \beta^a \left[ \frac{1}{1-\frac{1}{\eta}} c_a^{1-\frac{1}{\eta}} - \frac{\alpha}{1+\frac{1}{\gamma}} h_a^{1+\frac{1}{\gamma}} \right]$$

Subject to a lifetime budget constraint:

$$\lambda \left[ T + \sum_{a=0}^{T} (1-\tau) \beta^a e_a h_a w - \sum_{a=0}^{T} \beta^a c_a \right]$$

Assuming perfect capital markets and a steady state macroeconomy where wages are determined by the MPL of aggregate production function plus an efficiency units multiplier $e_a$ (so a person’s wage may change over their life-cycle). $\beta = \frac{1}{1+r}$ tacitly assumed.
Notation

\( a \)  age  
\( T \)  date of death or government transfer (bad notation)  
\( \tau \)  tax rate  
\( \beta \)  discount rate  
\( e_a \)  efficiency units - scales wage rate  
\( c_a \)  consumption  
\( h_a \)  labor  
\( w \)  aggregate wage in steady state  
\( \lambda \)  MU of lifetime income  
\( \eta \)  parameter for consumption: \( c_a^{1 - \frac{1}{\eta}} \)  
\( \gamma \)  parameter on labor: \( h_a^{1 + \frac{1}{\gamma}} \). "The elasticity"  
\( \alpha \)  parameter premultiplying labor in utility function.
Benchmark Model (2)

FOC’s:

\[ [c_a] : c_a^{\frac{-1}{\eta}} = \lambda \]

\[ [h_a] : \alpha h_a^{\frac{1}{\gamma}} = \lambda (1 - \tau) e_a w \]

Take log’s of the second FOC:

**structural model:** \( \ln h_a = \gamma \ln e_a + b \)

where

\[ b \equiv \gamma \left[ \ln \lambda + \ln (1 - \tau) + \ln (w) - \ln \alpha \right] \]

Can first difference to remove \( \lambda \). Then \( \Delta \ln h_a = \gamma \Delta \ln we_a + \Delta b \).
Question: normalizing $h_a$

- We have not given $h_a$ any units.
  - We could normalize this to whatever we want to - in fact we must choose the units - hrs/yr or hrs/wk.
  - Are these normalizations innocuous?

- Consider case where $\gamma \to 0$ holding everything else constant (frisch elasticity small).
  - Is it problematic this implies $h_a$ must equal one?
Answer: normalizing $h_a$

- $\alpha$ can adjust.
- Let $\tilde{h}_a = \kappa h_a$ where $\kappa$ is some renormalization. Then $\tilde{\alpha} = \left(\frac{1}{\kappa}\right)^{\frac{1}{\gamma}}$.
- That is, renormalizing $h$ will change our estimate of $\alpha$ (which shows up in the constant of the undifference logged equation) but the analysis will otherwise be unaffected.
Benchmark Model (3): Elasticities

- Can solves the model for fundamental elasticities.
- Consider three counterfactuals:
  1. wages change exogenously and in one period only;
  2. wages change exogenously in all periods;
  3. taxes are changed in all periods and the proceeds returned lump-sum.
- For the first, the income effect will be negligible.
- For (2) & (3) we need to know how the lagrange multiplier is affected.
Benchmark Model (4): $\lambda$

- Combine FOC’s:

$$h_a = \left(\frac{e_a}{e_0}\right)^\gamma h_0$$

- PDV of income equals PDV of consumption.

- Assume perfect credit markets and $\beta = \frac{1}{1+r}$. Then consumption is constant over the lifecycle.

  - Therefore, let $c_a \equiv \bar{c}wh_0$. If tax revenue were “discarded”, then $c_a = \bar{c}wh_0(1 - \tau)$
Benchmark Model (5): $\lambda$

- Consider FOC on consumption and perform algebraic manipulations, where $D = 0$ for compensated.

\[
\ln \lambda = -\frac{1}{\eta} \left[ \ln \bar{c} + \ln w + \ln h_0 + D \ln (1 - \tau) \right] \\
= -\frac{1}{\eta} \left[ \ln \bar{c} + \ln w + \gamma \ln \left( \frac{e_0}{e_a} \right) + \ln h_a \right]
\]

\[
\ln h_a = \\
\gamma \ln e_a + \gamma \left[ -\frac{1}{\eta} \left[ \ln \bar{c} + \ln w + \gamma \ln \left( \frac{e_0}{e_a} \right) + \ln h_a + D \ln (1 - \tau) \right] \right] \\
+ \ln (1 - \tau) + \ln (w) - \ln \alpha
\]
Benchmark Model (6): \( \lambda \)

\[
\ln h_a = \frac{\gamma}{\eta + \gamma} [(\eta - 1) \ln w - \eta \ln \alpha - \ln \bar{c} - \gamma \ln e_0] \\
+ \left[ \frac{\gamma \eta}{\eta + \gamma} - D \frac{\gamma}{\eta + \gamma} \right] \ln (1 - \tau) + \gamma \ln e_a
\]

- Can solve for policy counterfactuals using this model
Benchmark Model (7): 3 counterfactuals

1. \( \frac{d \ln h_a}{d \ln e_a} = \gamma + \gamma \frac{d \ln \lambda}{d \ln e_a} \) let \( \frac{d \ln \lambda}{d \ln e_a} \approx 0 \). So \( \gamma \) is the Frisch elasticity.

2. \( \frac{d \ln h_a}{d \ln (1-\tau)} \bigg|_{D=0} = \frac{\gamma \eta}{\eta + \gamma} \). This elasticity puts you on the original budget constraint; the compensated or “hicks elasticity”.

3. \( \frac{d \ln h_a}{d \ln (1-\tau)} \bigg|_{D=1} = \frac{\gamma (\eta - 1)}{\eta + \gamma} \)
The three big elasticities

1. $\varepsilon_F$ Frisch - instantaneous change in tax rate. No income effects because it is over such a short time frame it won’t effect permenant income. All substitution effect.

2. $\varepsilon_C$ Compensated - tax revenue is refunded lump-sum. Some people call this the Hicksian, but we are compensated to be on the original budget set, not the original indifferance curve. If we were on the original indifferance curve, then there would be no income effects – this is the same thing as the Frisch in the additively separable benchmark model! Rather, income effects are intermediate.

3. $\varepsilon_M$ Marshallian - tax revenue is ’thrown away’ or spent on a good which enters into the agent’s utility function in an additively separable way.
The three big elasticities - aside

For two good model,

\[ \eta_x = \frac{p_x \frac{\lambda}{s_x} - U_{xy} \eta_y \frac{s_y}{s_x} \frac{1}{p_y}}{U_{xx}} \]

- With additive separability in x and y, \( \lambda \) held constant same thing as no income effect
- Thus, with an additively separable utility function, the TRUE hicksian elasticity is the frisch elasticity.
Relative Size of elasticities

In the benchmark model...

\[ |\varepsilon_F| > |\varepsilon_C| > |\varepsilon_M| \]

Because

\[ \gamma > \frac{\gamma \eta}{\eta + \gamma} > \frac{\gamma (\eta - 1)}{\eta + \gamma} \]

Assuming \( \eta \in (1, \infty) \), \( \gamma \in (0, \infty) \).
Relative Size of elasticities - Intuition

With increase in taxes,

- Substitution effect makes people work less.
- Income effect makes people work more.
- For log utility of consumption ($\eta = 1$) they perfectly offset for Marshallin elasticity.
- Income effect is smaller for compensated elasticity and smallest for Frisch elasticity.
- All elasticities are the same only when there is no income effect on consumption
  - quasilinear utility with consumption numeraire, $\eta \to \infty$
Question: DWL of taxation?

- Posit a Marshallian elasticity of zero, representative agent, a flat tax.
  - What is the DWL of taxation.
  - How does this depend on whether there is a lump sump transfer?
Answer: DWL of taxation

- Trick question
- If government expenditures enter into utility in additively separable way, then no distortion.
  - (labor supply perfectly inelastic).
- Marshallian elasticity is irrelevant for lump-sum transfers.
  - Income effect is in no way identified by mere fact that income effect offsets substitution effect.
- Government projects that do not have income effect are less distortionary and hence less expensive!
Optimal Tax

Figure: Keane (2010), drawing from Brewer, Saez, and Shepard (2008)

Table 2: Revenue Maximizing Flat Tax Rates given different Labor Supply Elasticities

<table>
<thead>
<tr>
<th>Elasticity ((e))</th>
<th>Optimal Tax Rate ((\tau))</th>
</tr>
</thead>
<tbody>
<tr>
<td>( g=0 )</td>
<td>( g=0.5 )</td>
</tr>
<tr>
<td>2.0</td>
<td>33%</td>
</tr>
<tr>
<td>1.0</td>
<td>50%</td>
</tr>
<tr>
<td>0.67</td>
<td>60%</td>
</tr>
<tr>
<td>0.5</td>
<td>67%</td>
</tr>
<tr>
<td>0.3</td>
<td>77%</td>
</tr>
<tr>
<td>0.2</td>
<td>83%</td>
</tr>
<tr>
<td>0.1</td>
<td>91%</td>
</tr>
<tr>
<td>0.0</td>
<td>100%</td>
</tr>
</tbody>
</table>

- Gov't objective: taxes + \( g \) times after tax wages: \( \tau w + g (1 - \tau) w \).
  - Government redistributes income lump sum.
  - Government puts weight \( g \) on taxpayer's after tax income.
  - \( g = 0 \) means government wants to maximize transfers.

- Brewer, Saez, and Shepard (2008): \( \tau^* = \frac{1 - g}{1 - g + e} \), \( e \) is hicksian
\[ \Delta \ln h_{it} = -\alpha + \gamma \Delta \ln w_{it} (1 - \tau_t) - \frac{1}{\eta} \ln \beta (1 + r_t) + \gamma \Delta X_{it} + \gamma \xi_{it} + \gamma \Delta \varepsilon_{it} \]

- \( \tau_{it} \) tax rate may vary over time and persons
- \( X_{it} \) controls for exogenous shifts in tastes for work
- \( \xi_{it} \) surprise part of the change in \( \lambda \)
- \( \varepsilon_{it} \) taste shocks

**Unexpected changes in income may affect permanent income:**
\[ \Delta \ln w_{it} (1 - \tau_t) \]

- need instrument. Want \( z \perp \xi, \varepsilon \), measurement error but otherwise \( z \perp \Delta \ln w_{it} \).
- Wages changes don’t affect \( \lambda \) if they are expected.
MaCurdy: Empirical Model

- MaCurdy uses education and work experience as instruments for wage changes
- Plausibly orthogonal to taste shocks?
- Plausibly orthogonal to measurement error?
- Concerns outside the model (specification error)? This is the rest of the presentation
The papers differ in their identification strategies, chiefly dealing with $\lambda$

- MaCurdy differences out $\lambda$ and uses instruments that affect expected wage growth
- We are focusing on MaCurdy

- Altonji (1986) uses the levels equation but proxies for $\lambda$:

$$\lambda_{t-1} = E_{t-1} \left[ \frac{\lambda_t}{R_{t-1,t}} \right]$$

so

$$\frac{\lambda_t}{R_{t-1,t}} = \lambda_{t-1} + e_{\lambda_t}$$

- Thus past labor supply proxies for the current MU of income (related to Heckman and MaCurdy (1980) treating $\lambda_t$ as a fixed effect)

$$\ln \lambda_{a-1} = \ln \alpha + \frac{1}{\gamma} \ln h_{a-1} - \ln [(1 - \tau_{a-1}) e_{a-1}]$$

- However, Altonji uses consumption as his proxy (PSID)

- Browning et al. use the British Family Expenditure Surveys from 1970-77 (not panel!)

- He differences the equation and uses consumption data
### Table 7: Summary of Elasticity Estimates for Males

<table>
<thead>
<tr>
<th>Authors of Study</th>
<th>Year</th>
<th>Marshall</th>
<th>Hicks</th>
<th>Frisch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static Models</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kosters</td>
<td>1969</td>
<td>-0.09</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>Achenfelter-Heckman</td>
<td>1973</td>
<td>-0.16</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>Boskin</td>
<td>1973</td>
<td>-0.07</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>Hall</td>
<td>1973</td>
<td>n/a</td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td>8 British studies*</td>
<td>1976-83</td>
<td>-0.16</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>8 NIT studies*</td>
<td>1977-84</td>
<td>0.05</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>Burtless-Hauserman</td>
<td>1978</td>
<td>0.00</td>
<td>0.00-13</td>
<td></td>
</tr>
<tr>
<td>Wales-Woodland</td>
<td>1979</td>
<td>0.14</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td>Hausermann</td>
<td>1981</td>
<td>0.00</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>Blomquist</td>
<td>1983</td>
<td>0.08</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>Blomquist-Hansson-Busewitz</td>
<td>1990</td>
<td>0.12</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>MacCurdy-Green-Paarsch</td>
<td>1990</td>
<td>0.00</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>Triest</td>
<td>1990</td>
<td>0.05</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>Van Soest-Woitrz-Kapteyn</td>
<td>1990</td>
<td>0.19</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td>Ecklof-Sackien</td>
<td>2000</td>
<td>0.05</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>Dynamic Models</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MacCurdy</td>
<td>1981</td>
<td>0.08b</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>MacCurdy</td>
<td>1983</td>
<td>0.70</td>
<td>1.22</td>
<td>6.25</td>
</tr>
<tr>
<td>Browning-Deaton-Irtich</td>
<td>1985</td>
<td></td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>Bundell-Walker</td>
<td>1986</td>
<td>-0.07</td>
<td>0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>Altongi</td>
<td>1986</td>
<td>-0.24</td>
<td>0.11</td>
<td>0.17</td>
</tr>
<tr>
<td>Altongi</td>
<td>1986</td>
<td></td>
<td>0.31</td>
<td></td>
</tr>
<tr>
<td>Bover</td>
<td>1989</td>
<td>0.00</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>Altug-Miller</td>
<td>1990</td>
<td></td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>Angrist</td>
<td>1991</td>
<td></td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>Zilnik-Kniezner</td>
<td>1999</td>
<td>0.12</td>
<td>0.13</td>
<td>0.16</td>
</tr>
<tr>
<td>Pistaferri</td>
<td>2003</td>
<td>0.51b</td>
<td></td>
<td>0.70</td>
</tr>
<tr>
<td>Imui-Keane</td>
<td>2004</td>
<td>0.20c</td>
<td>0.66a</td>
<td>3.0-2.75f</td>
</tr>
<tr>
<td>Zilnik-Kniezner</td>
<td>2005</td>
<td>-0.47</td>
<td>0.33</td>
<td>0.54</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>0.04</td>
<td>0.30</td>
<td>0.83</td>
</tr>
</tbody>
</table>

* = Average of the studies surveyed by Pencavel (1986)

b = Effect of surprise permanent wage increase

c = Using MaCurdy Method #1
d = Using first difference hours equation
e = Approximation of responses to transitory wage increase based on model sir
f = Age range
The Frisch elasticity is small in most of the papers that have used this approach but not all! Keane calls this “the uncanny gap”.

The other elasticities must be smaller (based on this model!)

Based on this literature, there has been a consensus that the labor supply response to taxes is small, at least on intensive margin. (pencavel, 1986)

Keep in mind this is focusing on prime-age men.
Criticism

1. instruments violate exclusion restriction
   1. (model should account for human capital)
   2. Altonji (1986) was aware of this – I believe it wasn’t addressed because of tractability.

2. instruments are somewhat weak
   1. standard errors on parameter estimates typically do not exclude large elasticities.

3. Frisch elasticity is only partially identified with small wage shocks if there are adjustment costs

4. Consumption is steeply upward sloping for young workers, which model does not allow
   1. suggests precautionary savings or borrowing constraints
   2. Consider $\gamma \ln \beta (1 + r_t)$ in MaCurdy’s model.
Compensation included Value of Added Human Capital

- Related to Mincer Model, Ben-Porath, Generalized Ben-Porath, etc.
- On the job training creates systematic measurement error.
  - Observed wage: \( \frac{\text{income}}{\text{work} + \text{training}} \)
  - True wage: \( \frac{\text{income}}{\text{work}} \)
- Learning by doing ties human capital accumulation to labor decisions.
  - The observed wage ignores the value of human capital being accumulated.
  - Taxes on labor are taxes on investment – potentially large distortions
- Both models imply very different responses to tax changes.
Fig. 13.—Married males, some or all college (13-16 years), not self-employed, nonfarm, not enrolled in school (1966 1/1,000 census tape).
Table 3B in Heckman (1976) JPE

<table>
<thead>
<tr>
<th>Schooling Class</th>
<th>Dollar Investment</th>
</tr>
</thead>
<tbody>
<tr>
<td>12 years of school</td>
<td>976</td>
</tr>
<tr>
<td>13–16 years of school</td>
<td>2,068</td>
</tr>
<tr>
<td>16 years of school</td>
<td>2,492</td>
</tr>
</tbody>
</table>

Keane & Rogerson: young workers’ productivity exceeded wage by 54 percent due to time spent investing in skills

I couldn’t find this number in Heckman’s paper, but it isn’t too far off from 1+ 2,068 / 3000 based on the figure and table

Bottom line: using experience as an instrument ignores why wages are rising

For many workers, human capital accumulation is a major component of compensation during phases of life cycle
Keane and Imai (2004) - Summary

- Add human capital accumulation via LBD to something akin to benchmark model
- Marginal disutility of labor = wage + present value of change in human capital stock
  - Effects of temporary tax changes are smaller
  - Effects of permanent tax changes are large (taxing capital now)
- Parameters no longer analytically map to labor supply elasticities
- Estimate via simulated maximum likelihood on NLSY79 white males (w/ asset data)
  - non-linear bellman equations (need utility function to estimate $\gamma$)
  - measurement error, productivity and MU(c) shocks
Keane and Imai (2004) - 1.a: heuristic

- **Heuristic**: LBD captured by

  \[ w_{t+1} = \left( 1 + \kappa \sum_{j=1}^{t-1} h_{t-j} \right) w_1 \]

- Each unit of labor increases subsequent marginal product by a constant \( \kappa \).
  - Non-depreciating, linear.
- Add this to the benchmark model. Otherwise, things stay the same.
Keane and Imai (2004) - 1.b: heuristic

Now FOC for labor is different:

\[ [h_t] : \alpha h_t^{\frac{1}{\gamma}} = \lambda (1 - \tau_t) w_t + \lambda \mathbb{E}_t \left[ \sum_{j=t+1}^{A} \frac{1}{(1 + r)^{t-a}} \kappa (1 - \tau_t) w_1 h_t \right] \]

Combine with consumption FOC to yield MRS:

\[ \alpha h_t^{\frac{1}{\gamma}} c_t^{\frac{1}{\eta}} = (1 - \tau_t) w_t + \mathbb{E}_t \left[ \sum_{j=t+1}^{A} \frac{\kappa (1 - \tau_t) w_1 h_t}{(1 + r)^{t-a}} \right] \]
Keane and Imai (2004) - 2.a: model

Actual functional forms are different. General model

preferences:  \[ E_t \left[ \sum_{\tau=t}^{T} \beta^\tau [u(c_{\tau}, \tau) - v(h_{\tau}, \varepsilon_{2,\tau})] \right] \]

inter-temp budg. constraint:  \[ A_{t+1} = (1 + r) A_t + W_{t,s} h_t - C_t \]

wage rate:  \[ W_{t,s} = R_s K_t \]

human capital dynamics:  \[ K_{t+1} = g(h_t, K_t, t) \varepsilon_{1,t+1} \]

value function:  \[ V_{t,s}(A_t, K_t, \varepsilon_{2,t}) = \max_{c_t, h_t} \{ u(c_t, t) - v(h_t, \varepsilon_{2,t}) + \beta E_t[V_{t+1,s+t}(1 + r) A_t + R_s K_t h_t - C_t, g(h_t, K - t, t) \varepsilon_{1,t+1}, \varepsilon_{2,t+1}] \} \]

- MU of consumption time varying. \( \varepsilon_{2,\tau} \) is shock to disutility of labor. \( \varepsilon_{1,t+1} \) is shock to human capital accumulation (incorrectly termed wage shock in Keane & Rogerson).
Question

- MU of consumption is time-varying.
  - What fact makes this necessary?
  - What are alternatives for modelling this?
Consumption profile is upside-down U-shaped
Credit constraints and precautionary savings both can account for this.
Power function utility specialization:

\[ u(c_\tau, \tau) = A(\tau) \frac{C_{t}^{a_1}}{a_1} \]

\[ v(h_\tau, \varepsilon_{2\tau}) = \varepsilon_{2\tau} b \frac{h_{\tau}^{a_2}}{a_2} \]

Then

\[ \gamma \equiv \text{i.e.s.} \equiv \frac{1}{a_2 - 1} \]
many empirical articles analyzing intertemporal labor supply behavior, such as Shaw (1989) or Hotz et al (1988), use a translog function of consumption and leisure as the utility function. Although this approach has the advantage of being locally flexible, none of the parameters can be straightforwardly interpreted as describing the intertemporal elasticity of substitution in labor supply. Hence, from their estimation results, it is difficult to draw any conclusions about how much people substitute labor intertemporally, unless one simulates their estimated models.

The same could be said of Imai and Keane!

You need to simulate their model to understand the lifetime labor supply response to a tax change.
Keane and Imai (2004) - 2.c: model

Specialize law of motion and human capital production

\[ g(K, h, t) = k_0 + \delta K + G(K, h, t) \]

\[ G(K, h, t) = A_0 (1 + A_1(t - 19))(B_1 + K) [(h + d_1)^\alpha - B_2(h + d_1)] \]

NLSY79
Human capital production function captures stylized facts:

1. Wages are increasing in previous period’s hours worked
2. Slope increases current wages – complementarity in LBD \((B_1 + K)\)
3. Lower bound in human capital production at \(h = 0\). \(d_1\)
4. Very large hours have a negative or no effect. \(-B_2(h + d_1)\)

Seems to be consistent with LBD.
Keane & Imai (2004) - 2.f: model

- Assume lognormality in random variables, like error

\[ \ln(\varepsilon_i) \sim \mathcal{N}\left(-\frac{1}{2} \sigma_i^2, \sigma_i\right), \quad i = 1, 2 \]

- Terminal value function.
  - Pick one that is concave, increasing, allows negativity in assets (!), and is less sensitive with \( \phi \).

\[ V_{T+1}(A_{T+1}) = \begin{cases} 
3\ln(A_{T+1} + \phi) - 1 - 3\ln(\phi) & A_{T+1} > 0 \\
\left(\frac{A_{T+1} - \phi}{\phi}\right)^3 & A_{T+1} \leq 0
\end{cases} \]

- NLSY had data only to age 39, so likelihood of \( \phi \) very flat. Set to 10,000.
Parameters interacted with (exogenous) education:

1. preference parameters $b$, $C_0$, $C_1$, $C_2$
2. human capital production function parameters $K_0$, $\delta$, $A_0$, $A$, and $\alpha$

“Measurement error”

1. IID
2. Allowed in true wage $W_{t,s} = R_s K_t$, labor supply, and assets
3. Measurement error only partly identified – could be transitory shocks.
4. Called preference and productivity shocks if persistent

Simulated maximum likelihood (our new friend)
# Keane & Imai (2004) - 3.a: Results

<table>
<thead>
<tr>
<th>Utility Function Parameters</th>
<th>Disutility of labor</th>
<th>Consumption utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_2$</td>
<td>Disutility of labor curvature</td>
<td>$0.2617 (5.728 \times 10^{-4})$</td>
</tr>
<tr>
<td>$b$</td>
<td>$b_{nh}$ Non-high school</td>
<td>$0.01700 (5.968 \times 10^{-5})$</td>
</tr>
<tr>
<td>$b_h$</td>
<td>High school graduate</td>
<td>$0.5859 (0.01080)$</td>
</tr>
<tr>
<td>$b_c$</td>
<td>Some college</td>
<td>$0.5241 (0.003821)$</td>
</tr>
<tr>
<td>$b_{cg}$</td>
<td>College graduate</td>
<td>$0.5175 (0.01022)$</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>Std. error of disutility shock</td>
<td>$0.5460 (0.01967)$</td>
</tr>
</tbody>
</table>

### Table 4

**Estimation Results**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_2$</td>
<td>Disutility of labor curvature</td>
<td>1.2618 (8.504 \times 10^{-4})</td>
</tr>
<tr>
<td>$b_{nh}$</td>
<td>Non-high school</td>
<td>$1.831 \times 10^{-5} (1.891 \times 10^{-7})$</td>
</tr>
<tr>
<td>$b_h$</td>
<td>High school graduate</td>
<td>$1.651 \times 10^{-5} (5.801 \times 10^{-8})$</td>
</tr>
<tr>
<td>$b_c$</td>
<td>Some college</td>
<td>$1.623 \times 10^{-5} (1.041 \times 10^{-7})$</td>
</tr>
<tr>
<td>$b_{cg}$</td>
<td>College graduate</td>
<td>$1.745 \times 10^{-5} (2.041 \times 10^{-7})$</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>Std. error of disutility shock</td>
<td>0.01156 (6.748 \times 10^{-4})</td>
</tr>
</tbody>
</table>
**Keane & Imai (2004) - 3.b: Results**

<table>
<thead>
<tr>
<th>Production Function Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta$</td>
<td></td>
</tr>
<tr>
<td>$\delta_n$</td>
<td>Non-high school</td>
</tr>
<tr>
<td>$\delta_h$</td>
<td>High school graduate</td>
</tr>
<tr>
<td>$\delta_c$</td>
<td>Some college</td>
</tr>
<tr>
<td>$\delta_{cg}$</td>
<td>College graduate</td>
</tr>
<tr>
<td>$k_0$</td>
<td>Non-high school</td>
</tr>
<tr>
<td>$k_{0h}$</td>
<td>High school graduate</td>
</tr>
<tr>
<td>$k_{0c}$</td>
<td>Some college</td>
</tr>
<tr>
<td>$k_{0cg}$</td>
<td>College graduate</td>
</tr>
<tr>
<td>$A_0$</td>
<td>Non-high school</td>
</tr>
<tr>
<td>$A_{0h}$</td>
<td>High school graduate</td>
</tr>
<tr>
<td>$A_{0c}$</td>
<td>Some college</td>
</tr>
<tr>
<td>$A_{0cg}$</td>
<td>College graduate</td>
</tr>
<tr>
<td>$A_1$</td>
<td>Non-high school</td>
</tr>
<tr>
<td>$A_{1h}$</td>
<td>High school graduate</td>
</tr>
<tr>
<td>$A_{1c}$</td>
<td>Some college</td>
</tr>
<tr>
<td>$A_{1cg}$</td>
<td>College graduate</td>
</tr>
<tr>
<td>$\alpha$</td>
<td></td>
</tr>
<tr>
<td>$\alpha_n$</td>
<td>Non-high school</td>
</tr>
<tr>
<td>$\alpha_h$</td>
<td>High school</td>
</tr>
<tr>
<td>$\alpha_n$</td>
<td>Some college</td>
</tr>
<tr>
<td>$\alpha_{cg}$</td>
<td>College graduate</td>
</tr>
<tr>
<td>$B_2$</td>
<td>$- B_2(h + d_1)$</td>
</tr>
<tr>
<td>$B_1$</td>
<td>Additive constant in capital term $B_1 + K$</td>
</tr>
<tr>
<td>$\sigma_0$</td>
<td>Std. error of wage shock</td>
</tr>
<tr>
<td>$d_1$</td>
<td>Additive constant in hours term $h + d_1$</td>
</tr>
</tbody>
</table>

(continued)
### Keane & Imai (2004) - 3.c: Results

<table>
<thead>
<tr>
<th>Mean Initial Assets</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{A}$ Mean initial assets when the starting age is 20</td>
<td>3250.8 (458.6)</td>
</tr>
<tr>
<td>$\bar{A}$ Mean initial assets when the starting age is after 20</td>
<td>7190.4 (631.1)</td>
</tr>
<tr>
<td>$V_{\bar{A}}$ Std. error, initial assets</td>
<td>2218.7 (241.3)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Measurement Error Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_{\xi_0}$ Initial period wage$^b$</td>
<td>0.4909 (0.003626)</td>
</tr>
<tr>
<td>$\sigma_{\xi_1}$ Wage$^c$</td>
<td>0.4643 (0.001333)</td>
</tr>
<tr>
<td>$\sigma_{\xi_2}$ Hours$^d$</td>
<td>590.7 (2.156)</td>
</tr>
<tr>
<td>$\sigma_{\xi_{31}}$ Asset$^e$</td>
<td>2623.5 (178.5)</td>
</tr>
<tr>
<td>$\sigma_{\xi_{32}}$ Asset</td>
<td>948.8 (11.98)</td>
</tr>
</tbody>
</table>

**NOTES:** Standard errors are in parentheses.

1. $g(K, h, t) = A_0 (1 + A_1 (t - 19)) (B_1 + K) [h + d_1(t - 19) - B_2 (h + d_1)] + \delta K + k_0$.
2. $K_0^{D} = K_0 \xi_0$, $\ln(\xi_0) \sim N(0, \sigma_{\xi_0})$.
3. $K_0^{D} = K_0 h_0 \xi_0^{D}$, $\ln(\xi_{0,0}) \sim N(0, \sigma_{\xi_{0,0}})$.
4. $K_0^{D} = K_0 h_0 \xi_0^{D}$, $\ln(\xi_{0,0}) \sim N(0, \sigma_{\xi_{0,0}})$.
5. $h_0^{D} = h_0 + \xi_{2,0}$, $\xi_{2,0} \sim N(0, \sigma_{\xi_{2,0}})$.
6. $A_0^{D} = A_0 + \xi_{3,0}$, $\xi_{3,0} \sim N(0, \sigma_{\xi_{3,0}})$, $\sigma_{\xi_{3,0}} = \sigma_{\xi_{3,0}} + \sigma_{\xi_{3,2}}(t - 19)$.
## Table 2

Effects of an Unexpected 5% Permanent Tax Increase in Imai-Keane (all changes in %)

<table>
<thead>
<tr>
<th>Age</th>
<th>Age 25</th>
<th>Age 30</th>
<th>Age 35</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hours</td>
<td>Wage</td>
<td>Assets</td>
</tr>
<tr>
<td>25</td>
<td>-2.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>-2.9</td>
<td>-0.4</td>
<td>+19.8</td>
</tr>
<tr>
<td>35</td>
<td>-3.2</td>
<td>-0.7</td>
<td>+26.3</td>
</tr>
<tr>
<td>40</td>
<td>-3.8</td>
<td>-1.0</td>
<td>+14.5</td>
</tr>
<tr>
<td>45</td>
<td>-5.1</td>
<td>-1.3</td>
<td>+6.9</td>
</tr>
<tr>
<td>50</td>
<td>-7.9</td>
<td>-2.0</td>
<td>+2.6</td>
</tr>
<tr>
<td>55</td>
<td>-13.3</td>
<td>-3.6</td>
<td>-0.4</td>
</tr>
<tr>
<td>60</td>
<td>-19.3</td>
<td>-7.5</td>
<td>-3.0</td>
</tr>
<tr>
<td>65</td>
<td>-29.2</td>
<td>-11.6</td>
<td>-3.8</td>
</tr>
</tbody>
</table>
Keane & Imai (2004) - 3.e Results

1. $\gamma = 3.82$

2. Simulated hours response to a temporary wage increase of 2
   ranged from 0.6% for young individuals to 4% for those near
   retirement
   - human capital accumulation changes incentives to work now
   - human capital accumulation magnifies income effect of temporary
     tax change

3. Can macro-economists 'plug-in' this parameter estimate? I don’t
   think so...
Bias in the benchmark approach

1. Simulate data using the estimated model
2. Estimate the benchmark model as MaCurdy 1981 or Altonji 1986
   - They use quadratic and twice lagged wages as instruments (not education?!?)
3. Assess the bias
### Table 6

**ML, OLS, IV Results**

<table>
<thead>
<tr>
<th>Age</th>
<th>Method</th>
<th>$b_2$</th>
<th>$a_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>20–36</td>
<td>ML</td>
<td>3.82 (0.0124)</td>
<td>1.26 (0.000850)</td>
</tr>
<tr>
<td></td>
<td>Simulated data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20–64</td>
<td>OLS</td>
<td>-0.444 (0.00248)</td>
<td>-1.25 (0.0126)</td>
</tr>
<tr>
<td>20–64</td>
<td>3SLS</td>
<td>0.971 (0.0941)</td>
<td>2.03 (0.0999)</td>
</tr>
<tr>
<td>20–56</td>
<td>OLS</td>
<td>-0.301 (0.00247)</td>
<td>-2.33 (0.0273)</td>
</tr>
<tr>
<td>20–56</td>
<td>3SLS</td>
<td>1.13 (0.259)</td>
<td>1.89 (0.204)</td>
</tr>
<tr>
<td>20–46</td>
<td>OLS</td>
<td>-0.270 (0.00280)</td>
<td>-2.70 (0.0382)</td>
</tr>
<tr>
<td>20–46</td>
<td>3SLS</td>
<td>0.537 (0.217)</td>
<td>2.86 (0.753)</td>
</tr>
<tr>
<td>20–36</td>
<td>OLS</td>
<td>-0.293 (0.00366)</td>
<td>-2.41 (0.0426)</td>
</tr>
<tr>
<td>20–36</td>
<td>3SLS</td>
<td>0.325 (0.256)</td>
<td>4.08 (2.42)</td>
</tr>
<tr>
<td>NLSY79 data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20–36</td>
<td>OLS</td>
<td>-0.231 (0.00659)</td>
<td>-3.33 (0.124)</td>
</tr>
<tr>
<td>20–36</td>
<td>3SLS</td>
<td>0.260 (0.0769)</td>
<td>4.85 (1.14)</td>
</tr>
</tbody>
</table>

*Note: Delta method is used to calculate the standard errors for $b_2$ for ML and $a_2$ for OLS and 3SLS results. Std. errors are in parentheses. Instruments: const, experience (which is age-19), experience squared, twice lagged wage.*
### Table 7

OLS, IV results with cleaned data

<table>
<thead>
<tr>
<th>Age</th>
<th>Method</th>
<th>$b_2$</th>
<th>$a_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>20–36</td>
<td>ML</td>
<td>3.82 (0.0124)</td>
<td>1.26 (0.000850)</td>
</tr>
<tr>
<td></td>
<td>Simulated data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20–64</td>
<td>OLS</td>
<td>-0.168 (0.0299)</td>
<td>-4.94 (1.056)</td>
</tr>
<tr>
<td></td>
<td>3SLS</td>
<td>1.21 (0.196)</td>
<td>1.83 (0.134)</td>
</tr>
<tr>
<td>20–56</td>
<td>OLS</td>
<td>-0.171 (0.0895)</td>
<td>-4.85 (3.067)</td>
</tr>
<tr>
<td></td>
<td>3SLS</td>
<td>1.68 (0.569)</td>
<td>1.60 (0.202)</td>
</tr>
<tr>
<td>20–46</td>
<td>OLS</td>
<td>-0.171 (0.00590)</td>
<td>-4.87 (0.203)</td>
</tr>
<tr>
<td></td>
<td>3SLS</td>
<td>0.655 (0.521)</td>
<td>2.53 (1.22)</td>
</tr>
<tr>
<td>20–36</td>
<td>OLS</td>
<td>-0.177 (0.00775)</td>
<td>-4.65 (0.248)</td>
</tr>
<tr>
<td></td>
<td>3SLS</td>
<td>0.0750 (0.576)</td>
<td>14.3 (102)</td>
</tr>
<tr>
<td>NLSY79 data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20–36</td>
<td>OLS</td>
<td>-0.232 (0.00607)</td>
<td>-3.31 (0.113)</td>
</tr>
<tr>
<td></td>
<td>3SLS</td>
<td>0.142 (0.0731)</td>
<td>8.03 (3.61)</td>
</tr>
</tbody>
</table>

Note: Delta method is used to calculate the standard errors for $b_2$ for ML and $a_2$ for OLS and 3SLS results. Std. errors are in parentheses. Instruments: const, experience (which is age-19), experience squared, twice hourly wage.
### Table 8
ML, OLS, IV Results of Altonji Estimation

<table>
<thead>
<tr>
<th>Age</th>
<th>Method</th>
<th>$b_2$</th>
<th>$a_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>20–36</td>
<td>ML</td>
<td>3.82 (0.0124)</td>
<td>1.26 (0.000850)</td>
</tr>
<tr>
<td>Simulated data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20–64</td>
<td>OLS</td>
<td>−0.206 (0.00459)</td>
<td>−3.85 (0.108)</td>
</tr>
<tr>
<td></td>
<td>3SLS</td>
<td>2.81 (0.0804)</td>
<td>1.36 (0.0102)</td>
</tr>
<tr>
<td>20–56</td>
<td>OLS</td>
<td>−0.272 (0.00274)</td>
<td>−2.68 (0.0371)</td>
</tr>
<tr>
<td></td>
<td>3SLS</td>
<td>−0.218 (0.0306)</td>
<td>−3.59 (0.646)</td>
</tr>
<tr>
<td>20–46</td>
<td>OLS</td>
<td>−0.246 (0.00299)</td>
<td>−3.07 (0.0494)</td>
</tr>
<tr>
<td></td>
<td>3SLS</td>
<td>0.322 (0.0467)</td>
<td>4.11 (0.451)</td>
</tr>
<tr>
<td>20–36</td>
<td>OLS</td>
<td>−0.281 (0.00380)</td>
<td>−2.56 (0.0481)</td>
</tr>
<tr>
<td></td>
<td>3SLS</td>
<td>0.476 (0.182)</td>
<td>3.10 (0.803)</td>
</tr>
</tbody>
</table>

**Note:** Delta method is used to calculate the standard errors for $b_2$ for ML and $a_2$ for OLS and 3SLS results. Std. errors are in parentheses. **Instruments:** const, experience (which is age-19), experience squared, experience cubed.
Keane & Imai (2004) - 4.e: Comparison

- $\gamma$ estimated with MaCurdy method considerably lower than true $\gamma$
- $\gamma$ estimated with MaCurdy method on simulated data is considerably larger than $\gamma$ estimated by MaCurdy on NLSY79 data.
Domeij and Floden (2006) - cautionary reminder

- credit constraints, precautionary savings, complementarity of consumption and leisure, and time-varying tastes all produce similar life-cycle dynamics.

- For this reason they are often difficult to distinguish.
Credit constraints may bias estimates of $\gamma$ downwards.

- Borrowing crucial for smoothing consumption after negative wage shocks.
- Household may save, faces non-negativity constraint on assets.
- $\hat{\gamma}$ biased downwards.

- Faced with temporary negative wage shock and perfect capital markets, agent would reduce hours worked borrow against future earning.
- However, if one cannot borrow and does not have a stock of wealth one can smooth consumption only by working more.
- That is, we would be estimating the marshallian elasticity for that period, which could even be negative.
- If credit-constrained low-wealth agents are prevalent in the data, they will bias downward estimates of the aggregate.
Domeij and Floden (2006) - 1.b: Heuristic

Consider the MaCurdy empirical approach:

\[ \Delta \ln h_{it} = -\alpha + \gamma \Delta \ln w_{it} (1 - \tau_t) - \frac{1}{\eta} \ln \beta (1 + r_t) + \gamma \Delta X_{it} + \gamma \xi_{it} + \gamma \Delta \varepsilon_{it} \]

- MaCurdy assumes \( \beta (1 + r_t) = 1 \), so \( -\frac{1}{\eta} \ln \beta (1 + r_t) \) drops out.
- Instead, let \( r_t \) be endogenous.
  - Let \( r_t \) increase enough to prevent borrowing.
  - Same thing as a credit constraint.
- Then \( r_t \) will be negatively correlated with \( \Delta \ln w_{it} \)
  - With steeply increasing wages, you may want to borrow to smooth consumption.
- Using instruments for expected wage growth would bias \( \hat{\gamma} \) downward.
Domeij and Floden (2006) - 1.c: Do Credit Constraints Exist?

- Deaton (1991) and Rull (1997) show that most Americans have low asset levels.
- Can Americans borrow? As we saw this is a difficult problem.
- Credit constraints should be more severe for consumption than they are for education.
  - Consumption debts can be dissolved in bankruptcy, making them riskier for lenders.
  - Government subsidizes education loans.
Domeij and Floden (2006) - 2.a: Model

- Utility same as in benchmark model
- Non-negativity constraint in assets:
  \[ A_{it+1} = (1 + r)[A_{it} + w_{it} h_{it} - c_{it}], \quad A_{it} \geq 0 \]
- Stochastic process for wages
  \[ \ln w_{it} = \psi_t + z_{it}, \quad z_{it} = \rho z_{it-1} + \varepsilon_{it} \]
  - Specifying a wage process is not necessary for MaCurdy (1981) estimation, but becomes necessary after departing from benchmark assumptions.
Question: Why Wage Process Now Needed?
Answer: Why Wage Process Now Needed?

- Marginal utility of income depends on medium run asset level
- One may have assets to last one year but not several if smoothing consumption in response to a negative wage shock
- Agent needs to know how persistent wage shocks are so he ascertain these medium-term events.
Domeij and Floden (2006) - 2.c: Model

- A higher expected wage increase the next period will increase the value of borrowing this period, which will increase the hours worked this period.
- This is an omitted variable.
- This omitted variable will be postively correlated with the wage change.
- Thus, the coefficient on the wage change will be the sum of the frisch elasticity minus this (positive) omitted variable –
  - frisch elasticity is biased downwards when using conventional expected wage change instrumental variables.
Calibrate \( \{ \rho, \sigma_{\epsilon}, \sigma_{\psi} \} = \{0.90, 0.21, 0.34\} \) Floden and Linde (2001)

\{ \eta, \gamma \} = \{ \frac{2}{3}, 0.50 \}


\( \hat{\gamma} = 0.23 \) in full sample

\( \hat{\gamma} = 0.44 \) in subsample with positive assets.

\( \hat{\gamma} = 0.50 \) in subsample with assets above sample mean

\( \hat{\gamma} = -0.09 \) in subsample below 0.1 of the sample mean
Domeij and Floden (2006) - 4.a: Estimation

- PSID, male household heads
- Using MaCurdy estimation strategy on full sample (1277 obs.): \( \hat{\gamma} = 0.42 \) (outlier?)
- Using MaCurdy estimation on subsample with liquid wealth \( \geq 1 \) mo. income: \( \hat{\gamma} = 1.28 \) (std. error 1.15)
Wages follow exogenous stochastic process
Workers know typical wage path, but are wary of shocks
Precautionary motive incentivizes young workers to save, and not borrow against future income
Similarly, precautionary motive incentivizes young workers to work more while young than otherwise would.
Superimposing lifecycle wage path and lifecycle labor path will mislead about $\gamma$. 
Question: Precautionary Savings

Review of first-year material: What generates precautionary savings? Will our benchmark model generate it?
Answer: Precautionary Savings

- Convexity of *marginal* utility generates precautionary savings
  \[ U'(c_1) = \mathbb{E} \left[ \beta U' ((1 + r)(A - c_1 + \varepsilon)) \right] \]

- Exponent less than one implies convex marginal utility for power functions
  - Our benchmark model assumes this.
Utility function differs from benchmark model:

\[ U(c_{it}, l_{it}) = \frac{(c_{it}^{\theta} l_{it}^{1-\theta})^{1-\frac{1}{\eta}}}{1 - \frac{1}{\eta}} \]

- CRS production in consumption and leisure of commodity.
- Curvature of utility in commodity is decreasing in \( \eta \).
- More curvature reduces substitution of commodity across periods.
  - With low \( \eta \), consumption increases in periods with high labor to smooth commodity.
  - Low \( \eta \) also increases precautionary savings (marginal utility more convex in composite).
Calibration:
- Attanasio and Weber (1995) for utility function parameters
  - $\eta = \frac{1}{2.2}$ and $\theta = 0.4$
- use Meghir and Pistaferri (2004) wage process estimates
  - standard deviation of permanent shocks 0.18
  - transitory shocks 0.17.
Low (2005) - Summary 4

- Low didn’t use this model to estimate any labor supply parameters.
- However, precautionary savings causes the hours profile to be much flatter than it otherwise would be.
- This is the source of bias in MaCurdy type education/experience instrumental regressions.
Raj Chetty (2012) - 1a: Summary

- Agents reoptimize iff gains exceed fixed adjustment cost $\delta$
  - Adjustment cost could be technological, psychic, or “behavioral” (missperception and optimization error)
- Adjustment benefits are second and higher order components of linear approximation to utility function
  - Essentially, this is an underidentified version of the benchmark model.
    - Given estimate of $\gamma$ using the benchmark model, there exists a correspondence from $\delta$ to structural $\tilde{\gamma}$
      \[
      \tilde{\gamma} = g(\delta, \gamma)
      \]
  - Empirical estimate of $\gamma$ could be small whether true frisch elasticity $\tilde{\gamma}$ is large or small.
    - Like Manski Bounds – pessimistic about ability to identify something, so use extreme cases to bound estimate.
- Adjustment costs affect observed elasticity in short run. Cross-country comparisons may be less affected (!).
Raj Chetty (2012) - 2a: Model

- Suppose people are at their optimum before a tax change.
  - After the tax change, agents will adjust iff gains exceed optimization friction
  - Frictions could be thought of as:
    - Psychic costs,
    - Inattentiveness (tax changes not salient),
    - Unmodelled costs of searching for a new job or travelling for a new job.

- Utility function is same as benchmark model, but consumption is specialized as numeraire
Raj Chetty (2012) - ...: Model

- Utility evaluated at the optimum is:

  \[ U(h^* \mid \tau_t) = \frac{1}{1 + \gamma} \left[ \frac{1}{\alpha} \right]^\gamma [(1 - \tau_t) w]^{1+\gamma} \]

- Changes in utility due to after tax change
  \[ \Delta (1 - \tau) = (1 - \tau) - (1 - \tilde{\tau}) \] can be decomposed into that do to
  - change in tax holding hours fixed at old optimum
  - change in hours holding tax fixed at new optimum:

  \[ U\left(\tilde{h}^* \mid \tilde{\tau}\right) - U(h^* \mid \tau) = [U(h^* \mid \tilde{\tau}) - U(h^* \mid \tau)] + \left[U\left(\tilde{h}^* \mid \tilde{\tau}\right) - U(h^* \mid \tilde{\tau})\right] \]

- For small tax changes the second term is zero (envelope theorem). Second term is second-order.
- The derivative \( \frac{dU}{d(1 - \tau)} = wh^* \) capture first term
Raj Chetty (2012) - ...: Model

- Second term is captured by the second and higher order terms of taylor approximation. That is,

\[
U \left( \tilde{h}^* \mid \tilde{\tau} \right) - U (h^* \mid \tilde{\tau}) = \frac{1}{2} \gamma \frac{wh_t^*}{(1 - \tau)} (\tau - \tilde{\tau})^2 + o(\tau - \tilde{\tau})
\]

- Approximate utility gain due to adjusting hours as

\[
\left[ U \left( \tilde{h}^* \mid \tilde{\tau} \right) - U (h^* \mid \tilde{\tau}) \right] = \frac{1}{2} \gamma \frac{wh_t^*}{(1 - \tau)} (\tau - \tilde{\tau})^2
\]

- Utility loss from failing to adjust consumption is linear in \(\gamma\).
  - Exe. if the tax falls from 33% to 30%, and \(\gamma = 1\), then the utility drops by approximately 0.1 percent of consumption.
  - If the adjustment cost \(\delta\) exceeds this, then there is no adjustment.
Raj Chetty (2012) - …: Model

The agent will not adjust if

\[- \frac{1}{2} U''(h^*) (\tilde{h}^* - h^*)^2 \approx \left[ U(\tilde{h}^* | \tilde{\tau}) - U(h^* | \tilde{\tau}) \right] < \delta w h_t^* (1 - \tau_t)\]

Algebraic manipulations:

\[ U''(h^*) = -\frac{a}{\gamma} x^{\frac{1}{\gamma} - 1} \]

\[ h^* x^{\frac{1}{\gamma} - 1} (\tilde{h}^* - h^*)^2 < 2 \frac{\gamma}{a} \delta w h_t^* (1 - \tau_t) \]

\[ a h^* x^{\frac{1}{\epsilon}} = w (1 - \tau) \]

\[ h^* x^{\frac{1}{\gamma} - 1} (\tilde{h}^* - h^*)^2 < 2 \gamma \delta h_t^{* \frac{1}{\gamma} - \frac{1}{\gamma}} \]

\[ \frac{\tilde{h}^* - h^*}{h^*} < [2 \gamma \delta]^{\frac{1}{2}} \]
Dilemma

- $\gamma$ small: optimal hours will not adjust much even if $\delta = 0$.
- $\gamma$ large: adjustment cost increasing in $\gamma$ – still may not adjust.
Assume $\delta = 1\%$

“...even though MaCurdy (1981) estimates an intensive margin elasticity of only 0.15, his estimate is consistent with a structural elasticity as large as $\gamma = 5.63$. The reason is that MaCurdy’s estimates are identified from changes in wage rates of approximately 10%, which is not big enough to overcome small frictions.” - Chetty 2012
Raj Chetty (2012) - ...: Proposition

- Does this model reconcile micro with macro?
- “Intuitively, macroeconomic comparisons are more likely to overcome frictions because they analyze steady state behavior and because they induce coordinated change in work patterns.” -chetty (2012)
- I am still digesting this.
**Raj Chetty (2012) - ...: Application**

**Figure:**

<table>
<thead>
<tr>
<th>Study</th>
<th>Identification</th>
<th>$\beta$</th>
<th>SE($\beta$)</th>
<th>$t$</th>
<th>$A_{Bay}^{1→2}$</th>
<th>$\sigma_2$</th>
<th>$\sigma_1$</th>
<th>$\sigma_0$</th>
<th>55% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Ho (1981)</td>
<td>Life-cycle wage variation, 1947-1976</td>
<td>0.15</td>
<td>0.39</td>
<td>0.03</td>
<td>0.80</td>
<td>0.04</td>
<td>1.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Ho and Keyes (1985)</td>
<td>U.S. ETTC experience, 1984-1996, men</td>
<td>0.20</td>
<td>0.07</td>
<td>0.07</td>
<td>0.00</td>
<td>15.39</td>
<td>0.00</td>
<td>15.39</td>
<td></td>
</tr>
<tr>
<td>3. Ho and Keyes (1985)</td>
<td>U.S. ETTC experience, 1984-1996, women</td>
<td>0.09</td>
<td>0.07</td>
<td>0.07</td>
<td>0.00</td>
<td>15.37</td>
<td>0.00</td>
<td>15.37</td>
<td></td>
</tr>
<tr>
<td>4. Ho and Keyes (1985)</td>
<td>U.K. tax reform, 1978-1992</td>
<td>0.14</td>
<td>0.09</td>
<td>0.23</td>
<td>0.01</td>
<td>1.76</td>
<td>0.00</td>
<td>2.04</td>
<td></td>
</tr>
<tr>
<td>5. Ho and Keyes (1985)</td>
<td>Life-cycle wage, tax variation 1978-1987</td>
<td>0.15</td>
<td>0.07</td>
<td>0.39</td>
<td>0.03</td>
<td>0.80</td>
<td>0.04</td>
<td>0.96</td>
<td></td>
</tr>
<tr>
<td>A. Hour Elastics</td>
<td>Mean observed elasticity</td>
<td>0.15</td>
<td>0.05</td>
<td>0.45</td>
<td>0.15</td>
<td>0.92</td>
<td>0.10</td>
<td>1.04</td>
<td></td>
</tr>
<tr>
<td>6. Ho and Keyes (1985)</td>
<td>Ireland 1987 noon tax year</td>
<td>0.20</td>
<td>0.07</td>
<td>0.20</td>
<td>0.00</td>
<td>15.39</td>
<td>0.00</td>
<td>15.39</td>
<td></td>
</tr>
<tr>
<td>7. Ho and Keyes (1985)</td>
<td>U.S. tax reform, 1980-1994, men</td>
<td>0.14</td>
<td>0.07</td>
<td>0.04</td>
<td>0.00</td>
<td>3.51</td>
<td>0.00</td>
<td>3.64</td>
<td></td>
</tr>
<tr>
<td>8. Ho and Keyes (1985)</td>
<td>U.S. tax reform, 1980-1994, women</td>
<td>0.14</td>
<td>0.07</td>
<td>0.04</td>
<td>0.00</td>
<td>3.51</td>
<td>0.00</td>
<td>3.64</td>
<td></td>
</tr>
<tr>
<td>9. Ho and Keyes (1985)</td>
<td>Chicago housing vacancy turnover</td>
<td>0.14</td>
<td>0.03</td>
<td>0.36</td>
<td>0.02</td>
<td>0.84</td>
<td>0.01</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>10. Ho and Keyes (1985)</td>
<td>Chicago housing vacancy turnover</td>
<td>0.14</td>
<td>0.03</td>
<td>0.36</td>
<td>0.02</td>
<td>0.84</td>
<td>0.01</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>11. Ho and Keyes (1985)</td>
<td>Chicago housing vacancy turnover</td>
<td>0.14</td>
<td>0.03</td>
<td>0.36</td>
<td>0.02</td>
<td>0.84</td>
<td>0.01</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>12. Ho and Keyes (1985)</td>
<td>Chicago housing vacancy turnover</td>
<td>0.14</td>
<td>0.03</td>
<td>0.36</td>
<td>0.02</td>
<td>0.84</td>
<td>0.01</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>13. Ho and Keyes (1985)</td>
<td>Chicago housing vacancy turnover</td>
<td>0.14</td>
<td>0.03</td>
<td>0.36</td>
<td>0.02</td>
<td>0.84</td>
<td>0.01</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>14. Ho and Keyes (1985)</td>
<td>Chicago housing vacancy turnover</td>
<td>0.14</td>
<td>0.03</td>
<td>0.36</td>
<td>0.02</td>
<td>0.84</td>
<td>0.01</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>15. Ho and Keyes (1985)</td>
<td>Chicago housing vacancy turnover</td>
<td>0.14</td>
<td>0.03</td>
<td>0.36</td>
<td>0.02</td>
<td>0.84</td>
<td>0.01</td>
<td>0.95</td>
<td></td>
</tr>
</tbody>
</table>

*Note: Table I shows a subset of the data from the study.*
Raj Chetty (2012) - Application

Figure:

<table>
<thead>
<tr>
<th>Study</th>
<th>Identification</th>
<th>b</th>
<th>s.e.(b)</th>
<th>Marg(1 - γ)</th>
<th>a1</th>
<th>a2</th>
<th>a3</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>C. Top Income Elasticity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. Anker and Carroll (1995)</td>
<td>U.S. Tax Reform Act of 1986</td>
<td>0.57</td>
<td>0.12</td>
<td>0.37</td>
<td>1.53</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. Chetty (1999)</td>
<td>U.S. Tax Reform Act of 1986</td>
<td>1.00</td>
<td>0.15</td>
<td>0.37</td>
<td>2.14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19. Saez (2000)</td>
<td>U.S. tax reform 1986-2000</td>
<td>0.50</td>
<td>0.18</td>
<td>0.30</td>
<td>1.77</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B. Merit/Cross Section

<table>
<thead>
<tr>
<th>Study</th>
<th>Identification</th>
<th>b</th>
<th>s.e.(b)</th>
<th>Marg(1 - γ)</th>
<th>a1</th>
<th>a2</th>
<th>a3</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>21. Petroni (2004)</td>
<td>Cross-country tax variation, 1970-1996</td>
<td>0.46</td>
<td>0.09</td>
<td>0.42</td>
<td>0.18</td>
<td>1.20</td>
<td>0.10</td>
<td>1.41</td>
</tr>
<tr>
<td>22. Davis and Hoxby (2005)</td>
<td>Cross-country tax variation, 1970-1996</td>
<td>0.30</td>
<td>0.08</td>
<td>0.58</td>
<td>0.07</td>
<td>0.57</td>
<td>0.01</td>
<td>0.76</td>
</tr>
<tr>
<td>23. Blau and Kahn (2007)</td>
<td>U.S. wage variation, 1983-2000</td>
<td>0.31</td>
<td>0.04</td>
<td>1.00</td>
<td>0.19</td>
<td>0.53</td>
<td>0.18</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Unified bounds using panels (A) and (B)     0.28 | 0.64 | 0.23 | 0.61
Minimum bounds (μ⁺⁻)                        0.30
Unified bounds using all pairs               0.47 | 0.51 | 0.23 | 0.59
Minimum bounds (μ⁺⁻)                        0.50

*Bounds on structural elasticities are shown using data from existing studies. Column 3 shows the point estimate of the observed elasticity, column 4 shows the standard error, column 5 shows the size of the test of marginal tax wage, and column 6 shows identification. Columns 6 and 7 show the lower and upper bounds on the structural elasticity, calculated using Proposition 5. Column 8 gives a 95% confidence interval for a consistent estimator due to Abadie and Imbens (2008). See Appendix B for sources and details of the underlying calculations in columns 3-5.
Assume $\gamma = 0.5$ (frisch elasticity of 0.5 if no frictions)
Raj Chetty (2012) - Criticisms

- MaCurdy uses expected wage changes in his estimation
  - Presumably agents would consider this when planning work
    - Planning would mitigate adjustment costs
  - Expected changes are by their very nature salient
- Downward bias in $\hat{\gamma}$ hinges upon agents being at optimum before after-tax wage change.
- Chetty’s model, though plausible is underidentified
  - partial identification is the essence of his correspondence $\tilde{\gamma} = g(\delta; \gamma)$
  - Only if readers believe in non-trivial “$\delta$-class” adjustment costs should reader believe conventional estimates of $\gamma$ are (potentially) biased.
  - Furthermore, technological adjustment costs, psychic costs, and salience cannot be distinguished
  - Chetty spins his paper as behavioral – probably because it’s cool right now; this is pure hand-waving
Extensive Margins


- Thus far we have focused on prime-age males.
- We have ignored women, young workers, and old workers because of selection problems and the extensive margin (Heckman and MaCurdy, 1982).
- The main point of the early work on the extensive margin:
  - With indivisible labor, decision to exit labor force is unrelated to $\gamma$.
  - However, estimating $\gamma$ from a representative agent model on aggregate data will capture this extensive margin.
  - In the extreme, an individual’s frisch could be zero while the representative agent’s frisch is infinite.
  - This could help to explain why macro estimates of $\gamma$ tend to be much larger than micro estimates.

This margin is crucial to discerning the laffer curve, however I am not giving it any more attention.
Feldstein (1995)

- Elasticity of **taxable income** – what is reported on tax forms
  - May be sufficient for DWL
- Treasury department panel of > 4,000
- Estimates “elasticity” between 1 and 2 using law change and dif-in-dif
  - Sample restricted to those paying higher income taxes
- Non-structural, leaves many questions unanswered:
  - Tax avoidance, such as shifting income across time or utilizing charities?
  - Tax evasion?
  - Changes in labor supply or occupational choice?
  - How does counterfactual analysis depend on government transfers? Progressivity?
- Although this approach allows for no definite conclusions, it may suggest the benchmark estimates underestimate the true labor supply response.
Do People Respond to marginal tax rates?

- Do people actually respond to marginal tax rates?
- This was given attention in the 1980’s
  - If people’s preferences are distributed continuously, it would seem there should be mass points of wage income at tax kink points
  - Evidence on this is mixed
- More attention has been paid to this of late
  - Behavioral is cool right now
  - If people respond to average tax rates rather than marginal tax rates, then progressive taxes are no more distortionary than flat taxes
Hausman (1985) discussed whether there are mass points at tax-change induced kinks in budget constraints.
Saez (2010)

- Find evidence of bunching around first kink point of EITC
  - Chiefly among self-employed
  - May largely be due to “tax evasion”
- Find evidence of bunching around first income tax bracket where tax liability starts
- No further evidence of bunching.
- How much power do his tests have? What is alternative model?
Saez (2010)

Figure:

Panel A. Incidence curves and bunching

- Individual $L$ incidence curves
- Individual $H$ incidence curves

Additional notes on the graph:

- Individual $L$ decreases $z^*$ before and after the tax.
- Individual $H$ decreases $z^* + \Delta z^*$ before and after the tax.
- $\Delta z^*/z^* = \theta \Delta y / (1 - \theta)$ with $\theta$ compensated elasticity.
Saez (2010)

Figure:

**Figure 1. Bunching Theory**
Saez (2010)

Figure:
Figure: Saez (2010)
Saez (2010)

Figure 6: Taxable Income Density, 1960-1969: Bunching around First Kink
Saez (2010)

- Taxe rates politicians use are endogenous to population’s distribution of wages
  - Simply cross-sectional histograms are dangerous!
- In principle one could estimate taxable income elasticities from the bunch points
- Saez does so; I do not report them.
Electricity natural experiment

- In a Californian city, two utilities charge complicated marginal electricity rates
- These vary over time, and there is a curvaceous border between two utilities’ regions
- RDD
- Data fits both model with average taxes in FOC better than marginal taxes in FOC
  - However, cannot reject marginal taxes in FOC

Problems: electricity bills for these utilities VERY complicated. electricity bills a small share of income.

- It probably isn’t worth it to put in the effort to understand the electricity bills.
- However, it probably IS worth it to put in effort to understand tax rates
References I

Joseph Altonji. 
Intertemporal substitution in labor supply: Evidence from micro data. 

Martin Browning, Angus Deaton, and Margaret Irish. 
A profitable approach to labor supply and commodity demand over the life cycle. 
*Econometrica, 1985.*

Feldstein. 
Effect of marginal tax rates on taxable income: a panel study on the 1986 tax reform act. 
*Journal of Political Economy, 1995.*
References II

Jonathan Gruber and Emanuel Saez.
Elasticity of taxable income: evidence and implications.
*Journal of Public Economics, 2002.*

James Heckman.
A life-cycle model of earnings, learning, and consumption.

James Heckman and T. MaCurdy.
A life cycle model of female labor supply.

Koichiro Ito.
Do consumers respond to marginal or average price? evidence from nonlinear electricity pricing.
*University of California, Berkely, Job Market Paper, 2010.*
References III

Keane and Rogerson.
Reconciling micro and macro labor supply elasticities: A structural perspective.
_NBER Working Paper, 2011._

Michael Keane.
Labor supply and taxes: A survey.
_NBER, 2010._

Hamish Low.
Self insurance in a life-cycle model of labour supply and savings.
_Review of Economic Dynamics, 2005._

Thomas MaCurdy.
An empirical model of labor supply in a life-cycle setting.
References IV

Pencavel.
Labor supply of men: a survey.

*handbook of labor economics, 1986.*