Optimal Social Insurance Design: UI Benefit Levels

1 Overview

- optimal insurance design, application: UI benefit level
- optimal benefit level trades off distortion in behavior (moral hazard cost) with consumption smoothing benefit
- moral hazard:
  - search less intensively for new job
  - more likely to loose job
  - choose too risky jobs
  - too picky about taking new job (but maybe better match)
- moral hazard cost is captured by elasticity of unemployment duration w.r.t. benefit rate
- consumption smoothing: reduced drop in consumption upon unemployment as a function of UI benefit level
- value of consumption smoothing depends on risk aversion
- all these reduced form parameters can be estimated to derive optimal benefit level

2 Bailey-Chetty Model

2.1 Setup

- risk-averse worker with consumption utility $u(c)$
- initially employed, exogenous probability of becoming unemployed $p$
- once unemployed, regain employment with probability $q$ at utility cost $h(q)$
\[ c = \begin{cases} 
    w - \tau & \text{if employed (probability } 1 - p) \\
    b & \text{if unemployed (probability } p(1 - q)) \\
    w & \text{if reemployed (probability } pq) 
\end{cases} \]

- \( \tau \) is tax to finance benefit \( b \)
- \( w \) is exogenous
- \( \tau \) and \( b \) must satisfy the government budget constraint
  \[ p(1 - q)b = \tau(1 - p) \]  \hspace{1cm} (1)

2.2 First Best

- suppose we could pick \( b, \tau \) and also dictate the worker’s search effort \( q \)
- solve
  \[ \max_{b, \tau, q} \{(1 - p)u(w - \tau) + p[qu(w) + (1 - q)u(b) - h(q)]\} \]
  s.t. (1)
- FOCs
  \[ u'(w - \tau) = u'(b), \]
  which implies \( b = w - \tau \) and thus full insurance

2.3 Second Best

- now suppose the government can set the policy parameters \( b, \tau \), but it cannot dictate \( q \)
- instead, given \( b, \tau \), the worker chooses \( q \) to maximize expected utility
- worker’s problem:
  \[ V(b, \tau) \equiv \max_q \{(1 - p)u(w - \tau) + p[qu(w) + (1 - q)u(b) - h(q)]\} \]
- solution \( q^*(\tau, b) \), which solves the FOC
  \[ h'(q) = u(w) - u(b), \]
so that, in fact, $q^*(b)$ only

- substitute the worker’s best response $q^*(b)$ into the government budget constraint (1) to get

$$\tau^*_B(b) = \frac{p}{1 - p} (1 - q^*(b)) b$$

(2)

- $\tau^*_B(b)$ is the tax rate that is required to finance a benefit level $b$, taking into account the behavioral effect of $b$ on $q$

- second best problem:

$$\max_b V(b, \tau^*_B(b)),$$

(3)

which yields the second best optimal benefit level

- how to compute it?

- calibration:
  - choose some $u(c)$ using reasonable degree of risk aversion
  - choose $p$ from data based on current policy
  - choose $h(q)$ so that resulting $q^*(b)$ matches data, given $u(c)$
  - note: need to observe pairs $(q^*, b)$ for that
  - with these inputs, can compute optimal $b, \tau$

- alternative: try to derive optimality condition directly in terms of observable quantities

- FOC of (3):

$$V_b(b, \tau^*_B(b)) + V_{\tau}(b, \tau^*_B(b)) \tau^*_B(b) = 0$$

(4)

- by the envelope theorem,

$$V_b(b, \tau) = (1-q)pu'(b)$$

$$V_{\tau}(b, \tau) = -(1-p)u'(w - \tau)$$

(5)

and from the budget constraint (1),

$$\tau^*_B(b) = \frac{p}{1 - p} \left[ 1 - q^*(b) - q^*(b)b \right]$$

(6)
substituting (5) and (6) in rearranged FOC (4) yields

\[
\tau_{\beta}'(b) = -\frac{V_b}{V_{\tau}}
\]

\[\Leftrightarrow 1 - q^{*}(b) - q'^{*}(b)b = (1 - q^{*}(b)) \frac{u'(b)}{u'(w - \tau)} \]

\[\Leftrightarrow u'(w - \tau) \left[ 1 - \frac{q'^{*}(b)b}{1 - q^{*}(b)} \right] = u'(b) \quad (7)
\]

RHS: marginal benefit of raising consumption when unemployed (increase \(b\))

LHS: marginal cost of raising the tax in the employed state to cover increase in \(b\)

- direct utility cost is \(u'(w - \tau)\)
- additional indirect cost (through budget constraint) from behavioral response of \(q^*\) in square brackets

we can rearrange (7) once again to get

\[
\frac{u'(b) - u'(w - \tau)}{u'(w - \tau)} = -\frac{q'^{*}(b)b}{1 - q^{*}(b)} \equiv \epsilon_{1-q,b}(b) \quad (8)
\]

Bailey formula:

- LHS: consumption smoothing benefit of social insurance
- RHS: moral hazard cost as measured by the elasticity of unemployment probability \(1 - q\) w.r.t. the benefit level \(b\)

based on (8), can do “reduced form” evaluation of benefit:

- specify \(u(c)\) (effectively risk aversion)
- get \(\epsilon_{1-q,b}\) from data
- get consumption levels of employed and unemployed, \(c_e, c_u\) from data, where (here) \(c_e = w - \tau\) and \(c_u = b\)
  - in fact, not very different from calibration above

special case: CRRA utility with

\[u(c) = \frac{c^{1-\gamma}}{1-\gamma} \Rightarrow u'(c) = c^{-\gamma}\]
Bailey formula (8) becomes

\[
\left( \frac{c_u}{c_e} \right)^{-\gamma} = 1 + \varepsilon_{1-q,b}(b),
\]

so we only need \(c_u/c_e\) (percent consumption drop upon unemployment) to measure the consumption smoothing benefit

more generally, can do Taylor approximation

\[
u'(c_u) - u'(c_e) \approx u''(c_e)(c_u - c_e),
\]

assuming that \(u''\) (prudence) and thus the precautionary savings effects are small (see Chetty, 2006, for a third order approximation taking into account precautionary effects, suggests that this matters quantitatively)

then the LHS of the Bailey formula (8) becomes

\[
\frac{u'(c_u) - u'(c_e)}{u'(c_e)} \approx \frac{u''(c_e)}{u'(c_e)} c_e - c_u = \gamma(c_e) \frac{c_e - c_u}{c_e},
\]

which yields the approximate Bailey formula

\[
\frac{\Delta c}{c_e} \gamma(c_e) = \varepsilon_{1-q,b}(b)
\]

LHS further decomposes the consumption smoothing benefit in two components:

- relative drop in consumption upon unemployment
- degree of (relative) risk aversion at \(c_e\)

another useful way of rewriting the formula is in terms of the replacement rate \(r \equiv c_u/c_e:\)

\[
1 \frac{1}{1 - r} = \frac{\gamma}{\varepsilon'}
\]

i.e. the optimal replacement rate is increasing in risk aversion and decreasing in the duration elasticity

we will consider how to estimate these inputs into the formula later, but first consider extensions of the model and thus of the formula (10)
2.4 Extensions – Chetty (JPubE 2006)

- key message: optimality formula (10) is robust to dropping various simplifying assumptions in the Bailey model

- consider e.g. the following dynamic model: time is $t \in [0, 1]$, no discounting, zero interest

- no savings, hand-to-mouth workers

- $t = 0$: unemployed with probability $p$, employed with probability $1 - p$

- then those workers who became unemployed choose the time $t = 1 - q$ at which they become reemployed, at utility cost $h(q)$

- preferences become

$$\int_{0}^{1-q} u(b) dt + \int_{1-q}^{1} u(w) dt - h(q) = (1 - q)u(b) + qu(w) - h(q)$$

upon unemployment

- same with government budget constraint $p \int_{0}^{1-q} b dt = (1 - p)\tau$

- same model as previously, $1 - q$ is now just duration of unemployment

- same Bailey formula, $\epsilon_{1-q,b}$ is now just duration elasticity

- let’s consider yet another dynamic variant of the model

- suppose workers can perfectly save and borrow at rate $r = 0$

- since there is no duration risk (workers can control unemployment duration perfectly), after getting unemployed, they equalize their consumption across unemployment and reemployment at

$$c = qw + (1 - q)b$$

- ex ante expected utility becomes

$$V(b, \tau) \equiv \max_{q} (1 - p)u(w - \tau) + p [u(qw + (1 - q)b) - h(q)]$$

and solution $q^*(b)$
• this is now different than before

• but: FOC for optimal benefit will be the same as before since we get the same envelope conditions, with now

\[ V_\tau(b, \tau) = -(1-p)u'(w - \tau) \]
\[ V_b(b, \tau) = p(1-q)u'(c_u) \]

• we get the Bailey formula as before with now

\[ c_u = qw + (1-q)b \]

rather than \( c_u = b \)

• thus Bailey formula is robust to introducing saving

• "sufficient statistics" for welfare (Chetty JPubE 2006, AER 2009)

• Chetty (2006) explores various further extensions, e.g.

  – liquidity constraints just increase consumption drop therefore increase consumption smoothing benefit

  – utility from leisure reduces elasticity of unemployment w.r.t. benefits

  – many periods/unemployment spells

  – heterogeneity

• in the latter cases, get weighted averages of consumption drop/risk aversion/elasticity in Bailey formula

2.5 Problems

• disadvantage compared to “structural” approach (calibration): only tells us in what direction to move with \( b \), not informative about finding global optimum

• if estimate \( \Delta c, \varepsilon \) from data, away from the optimum, Bailey formula (10) will hold as an inequality for given \( \gamma \). This tells us about the direction of a welfare improving reform locally. But \( \Delta c, \varepsilon \) are endogenous to policy, so as we move away from the policy underlying the data, these “sufficient statistics” would change. We cannot treat them as fixed.
• does not help in computing total welfare gains
• moreover, even if could estimate $\Delta c$ and $\epsilon$ precisely, results depend crucially on risk aversion $\gamma$, which yields wide bands for optimal benefits
• aggregate uncertainty: parameter inputs vary over the business cycle
• weighted time/state averages of elasticities/marginal utilities in general formula may be hard to estimate/construct
• so far: no unemployment duration risk, once unemployed, agent can perfectly determine duration $1 - q$
• could just pay out fixed amount on the day when become unemployed, rather than constant flow of benefits (severance pay)
• would get rid of moral hazard (in terms of reemployment), could achieve first best!
• ignored key practical issues:
  – duration risk
  – moral hazard from becoming unemployed
• only looked at benefit level so far, not optimal duration of benefits (important policy instrument)
• will consider in more detail when consider dynamics of UI more explicitly, i.e. optimal time path of benefits over duration

2.6 UI over the Business Cycle with Frictional Labor Markets
• how should UI benefits be adjusted over the business cycle, e.g. in recessions?
• during a recession, it’s much harder to find a job, so how does that affect the duration elasticity $\epsilon$?
• Landais, Michaillat, Saez (2011)
• key: endogenous (and cyclical) labor market tightness
• general equilibrium effects and search externality: get additional (corrective) terms in the formula, since want to discourage search in frictional labor market
• workers don’t take into account that by searching harder and finding a job, they make it harder for the other unemployed to find a job (extreme case: rat race with fixed number of jobs out there)

• need to distinguish two elasticities:
  – micro-elasticity $\varepsilon^m$: the percentage increase in unemployment $1 - q$ when $b$ increases by 1%, taking into account the jobseekers’ reduction in search effort but keeping labor market tightness fixed. Can be estimated by measuring the reduction in the job-finding probability of an individual unemployed worker whose unemployment benefits are increased, keeping the benefits of all other workers constant
  – macro-elasticity $\varepsilon^M$: percentage increase in unemployment when $b$ increases by 1%, assuming that both search effort and tightness adjust. Can be estimated by measuring the increase in aggregate unemployment following a general increase in unemployment benefits.

• modified Bailey-Chetty formula:

$$\frac{1}{1 - r} = \frac{\gamma}{\varepsilon^M} + \left(\frac{\varepsilon^m}{\varepsilon^M} - 1\right) C,$$

where $C > 0$ is a term that includes the elasticity of the disutility of effort $h(q)$, $\gamma$, etc.

• collapses back to (11) if $\varepsilon^m = \varepsilon^M$

• otherwise, two departures:
  – what matters in the first term is the macro-elasticity $\varepsilon^M$, not the micro-elasticity $\varepsilon^m$, which is usually used for calibration (see below). Only $\varepsilon^M$ captures the budgetary cost of UI from higher (aggregate) $b$.
  – a second term, involving $\varepsilon^m / \varepsilon^M$ appears on the RHS. It’s the Pigouvian correction term taking into account the effect of a change in $b$ on labor market tightness.

• corrective term is positive if $\varepsilon^m > \varepsilon^M$. Show that this is true in a standard model of frictional labor markets: decrease in $b$ leads to more individual search effort and therefore a lower (individual) probability of unemployment, as captured by $\varepsilon^m$. When more unemployed find jobs, however, this decreases the job-finding rate
of the remaining unemployed (labor market tightness falls). As a consequence, the equilibrium reduction in unemployment following an aggregate increase in search effort ($\epsilon^M$) is smaller than the decrease in the individual probability of unemployment following an increase in individual search efforts ($\epsilon^m$).

- in particular, corrective term justifies UI even if there is no motive for insurance ($\gamma = 0$). Want to discourage search by providing UI since search imposes negative externality on other job-seekers when labor market tightness is endogenous.
- argue that $\epsilon^m$ is roughly acyclical, but $\epsilon^M$ is lower in recessions than in booms, which would justify more generous UI replacement rates in recessions

3 Implementing the Bailey-Chetty Formula

key inputs:

- elasticity of unemployment duration w.r.t. benefit rate
- consumption smoothing effect w.r.t. benefit rate

3.1 UI Benefits and Unemployment Durations

- estimates suggest elasticity $\sim .4$ to $.8$ (see Krueger and Meyer, 2002, NBER WP 9014, for a useful review)

- much larger than standard estimates of male labor supply elasticities
  - intensive versus extensive margin
  - short run versus long run
  - heterogeneity (typical unemployed may be just different than typical employed)
  - UI literature uses natural experiments, cleaner source of variation

- UI affects many dimensions of labor supply
  - probability of job loss
  - spousal labor supply
  - entry into labor market (entitlement effect)
  - choice of occupation/sector (unemployment risk)
how to estimate elasticity?

- regress
  \[ D_i = \beta b_i + \varepsilon_i \]
  where \( D_i \) is the unemployment duration of individual \( i \) and \( b_i \) is the benefit rate

- problem: \( b_i \) is a function of the earnings history of individual \( i \), which is likely to also affect \( D_i \)

- regress
  \[ D_i = \beta_1 b_i + \beta_2 P_i + \varepsilon_i \]
  where \( P_i \) captures individual \( i \)'s earnings history

- \( b_i \) is a nonlinear function of past earnings. This would assume that any nonlinear function of earnings affects unemployment durations only through benefits.

- approaches to deal with endogeneity of UI benefits:
  - regression discontinuity (e.g. when UI benefits depend on age cutoff, Bender/Schmieder/von Wachter, 2011)
  - “natural experiments:” legal variation across states and across time in benefit rates (see handout)
    * across US states
      \[ D_{is} = \beta_1 b_s + \beta_2 P_{is} + X_{is} \beta_3 + \varepsilon_{is} \]
    * change in benefit rate within a state
      \[ D_{it} = \beta_1 b_t + \beta_2 P_{it} + X_{it} \beta_3 + \varepsilon_{it} \]
    where \( b_s \) or \( b_t \) is e.g. the maximum (or average) benefit in state \( s \) at time \( t \)
    * change in benefits within a state that affects some groups, but not others. Unaffected group can serve as control, difference in difference estimate (see handout).
    * change in benefits in some states but not in others. States that change at different times serve as controls for each other.

- many studies using such cross-state variation since UI is determined at the state level
• preferable regression: do not just use e.g. maximum benefit $b_s$ in regression as above (get attenuation bias), but in fact do 2SLS using e.g. maximum benefit in state $s$ as instrument for the individual benefit $b_{is}$

• potential problem: UI parameters in a state may be endogenous to labor market characteristics in that state

• even better: construct “simulated instrument” (Gruber, AER 1997):
  - take entire national sample of unemployed
  - run it through the UI system of each state $s$ at time $t$ and compute average replacement rate
  - gives exogenous average replacement rate for that state at that time (since we held the sample fixed)
  - effectively uses all the information about the program (e.g. not just the average benefit)

• how to interpret (uncompensated) elasticity: liquidity versus moral hazard effects (Chetty, JPE 2008)
  - liquidity effect: higher UI benefit may increase durations because the unemployed don’t have to rush back to work as they are unable to smooth their consumption by saving/borrowing
  - not distortionary, since this effect would also emerge from a lump-sum severance payment
  - moral hazard effect: UI raises duration as it subsidizes leisure (a substitution effect)
  - uncompensated elasticity includes both effects, but moral hazard cost is captured only by the second effect
  - shows that roughly 60% of the effect of UI benefits on durations is due to the liquidity effect (comparing unemployed with high and low amounts of assets at the beginning of the spell separately, and the effect of lump-sum severance payments)
  - suggests higher optimal UI benefits, or alternative policy provisions to make sure the unemployed have access to liquidity, such as loans, UI savings accounts (Feldstein, AER 2005)
• macro- versus micro-elasticity: most of the evidence is about $\varepsilon^m$, but recent effort to disentangle the two.

3.2 UI Benefits and Consumption Smoothing

• Gruber (AER 1997):
  \[ \Delta \log c_{is} = \beta_1 b_{is} + X_{is} \beta_2 + \varepsilon_{is} \]
  where $b_{is}$ gets instrumented as discussed above
• we expect $\Delta \log c < 0$, so $\beta_1 > 0$ indicates a consumption smoothing effect of UI
• measure of $c$: food expenditure
• problems:
  – consumption versus expenditure (Aguiar and Hurst, JPE 2005)
  – how to interpret consumption smoothing benefit? Even if we found no effect (because people smooth their consumption even in the absence of UI through other channels), UI may be a more efficient means of achieving consumption smoothing. Examples:
    * spousal labor supply (Cullen and Gruber, JLE 2000)
    * precautionary savings (Engen and Gruber, JME 2001)
    * reduce investment in children’s education (with bad long run implications). Bad form of consumption smoothing that gets mitigated by UI. Just looking at consumption smoothing effect may understate benefits of UI system.

3.3 Putting It Together

• Gruber (1997) puts estimates of consumption smoothing effects and elasticity together to compute optimal benefit level using Bailey formula, for different levels of risk aversion

• results: optimal replacement rate

\[
\begin{array}{c|cc}
\gamma & \varepsilon = .9 & \varepsilon = .6 \\
1 & 0 & 0 \\
2 & 2.3\% & 29.4\% \\
3 & 29.4\% & 47.5\% \\
4 & 43.0\% & 56.6\%
\end{array}
\]
• surprisingly low, compared with actual average replacement rate around (50%). For this to be optimal, would need high risk aversion or low moral hazard cost

• but Chetty (JPE 2008): in view of liquidity/moral hazard decomposition, there is a positive welfare gain from increasing UI benefits even starting from the current replacement rate around 50%