Optimal Taxation with Behavioral Agents

Emmanuel Farhi and Xavier Gabaix

Harvard

January 23, 2020
Our Paper

- Behavioral version of three pillars of optimal taxation theory:
  - Ramsey (linear taxation to raise revenues and redistribute)
  - Pigou (linear taxation to correct for externalities)
  - Mirrlees (nonlinear taxation to raise revenues and redistribute)

- Unified treatment of behavioral biases with sufficient statistics:
  - misperceptions of taxes
  - “internalities”
  - mental accounts, etc...
Outline

- Behavioral price theory
- Behavioral optimal tax formulas (Ramsey, Pigou, Mirrlees)
- Concrete lessons by specializing model
- Additional results (Diamond-Mirrlees, Atkinson-Stiglitz...)
Example: Decision vs. Experienced Utility

- Decision utility $u^s$ and experience utility $u$

- Agent behavior

$$c(q, w) = \arg\max_c u^s(c) \text{ s.t. } q \cdot c \leq w$$

- Ex. internalities from temptation, hyperbolic discounting...
Example: Misperception

- True prices $\mathbf{q}$ and perceived prices $\mathbf{q}^s(\mathbf{q}, w)$
- Agent behavior (Gabaix 2014)

$$
c(\mathbf{q}, w) = \arg \max_{\mathbf{c} \in \mathbb{R}^n} |\mathbf{q}^s(\mathbf{q}, w) u(\mathbf{c}) \text{ s.t. } \mathbf{q} \cdot \mathbf{c} = w
$$

i.e.

$$
u'(c(\mathbf{q}, w)) = \lambda \mathbf{q}^s(\mathbf{q}, w) \text{ with } \lambda \text{ such that } \mathbf{q} \cdot c(\mathbf{q}, w) = w
$$

- Implications:
  - “trade-off” according to perceived relative prices $\frac{u'_{c_1}}{u'_{c_2}} = \frac{q_1^s}{q_2^s}$
  - budget constraint satisfied $\mathbf{q} \cdot \mathbf{c} = w$
General Model: Behavioral Price Theory

- Two primitives:
  - Marshallian demand function $c(q, w)$ with $q \cdot c(q, w) = w$
  - "experienced" utility function $u(c)$

- Indirect utility function $v(q, w) = u(c(q, w))$

- Misoptimization wedge $\tau^b = q - \frac{u_c(c(q, w))}{v_{w}(q, w)}$

- Slutsky matrix $S^C_j(q, w) = c_{qj}(q, w) + c_{w}(q, w)c_j(q, w)$

- Behavioral Roy identity $\frac{v_{qj}(q, w)}{v_{w}(q, w)} = -c_j - \tau^b \cdot S^C_j$
Mapping to the General Model: Concrete Examples

- **Decision vs. experienced utility model:**
  - misoptimization wedge $\tau^b = \frac{u^c}{v^s_w} - \frac{u_c}{v_w}$
  - $\tau_i^b > 0$ for “tempting” goods
  - Slutsky $S_{ij} = S^s_{ij}$

- **Misperception model:**
  - misoptimization wedge $\tau^b = q - q^s$
  - $\tau_i^b > 0$ for goods with non-salient taxes
  - Slutsky $S^H_{ij} = \sum_k S_{ik} \frac{\partial q^s_k(q, w)}{\partial q_j}$
Many-Person Ramsey (Diamond 1975)

▶ Social objective function

\[ L(\tau) = W(\nu^h(p + \tau, w)) + \lambda \sum_h [\tau \cdot c^h(p + \tau, w) - w] \]

▶ Optimal tax formula

\[ 0 = \frac{\partial L(\tau)}{\partial \tau_i} = \sum_h [(\lambda - \gamma^h)c_i^h + \lambda (\tau - \tilde{\tau}^{b,h}) \cdot S_i^{C,h}] \]

▶ Sufficient statistics:

▶ social marginal welfare weight \( \beta^h = W_{v^h} \nu^h_w \)

▶ social marginal utility of income \( \gamma^h = W_{v^h} \nu^h_w + \lambda \tau \cdot c^h_w \)

▶ substitution elasticities \( S_i^{C,h} \)

▶ weighted misoptimization wedge \( \tilde{\tau}^{b,h} = \frac{\beta^h}{\lambda} \tau^{b,h} \)
Many-Person Ramsey (Diamond 1975)

- Optimal tax formula

\[ 0 = \frac{\partial L(\tau)}{\partial \tau_i} = \sum_h \left[ (\lambda - \gamma^h)c_i^h + \lambda (\tau - \tilde{\tau}^{b,h}) \cdot S_i^{C,h} \right] \]

- Three terms:
  - mechanical \( (\lambda - \gamma^h)c_i^h \)
  - substitution \( \lambda \tau \cdot S_i^{C,h} \)
  - misoptimization \( -\lambda \tilde{\tau}^{b,h} \cdot S_i^{C,h} \)

- Additional condition if lump sum taxes \( \sum_h (\lambda - \gamma_h) = 0 \)
Assume symmetric Slutsky matrices $S_{ij}^{C,h} = S_{ji}^{C,h}$.

Then tax formula expressible in “discouragement” form:

$$\frac{-\sum_{h,j} \tau_j S_{ij}^{C,h}}{c_i} = 1 - \frac{\bar{\gamma}}{\lambda} - \text{cov} \left( \frac{\gamma^h}{\lambda}, \frac{Hc_i^h}{c_i} \right) - \frac{\sum_{h,j} \tilde{\tau}_j^b S_{ij}^{C,h}}{c_i}$$
Pigou (Sandmo 1975)

- Externality $\xi = \xi((c^h))_{h=1...H}$, indirect utility $v^h(q,w,\xi)$

- Optimal tax formula

$$0 = \frac{\partial L(\tau)}{\partial \tau_i} = \sum_{h}[(\lambda - \gamma^{\xi,h})c^h_i + \lambda(\tau - \tau^{\xi,h} - \tilde{\tau}^{b,h}) \cdot S^{c,h}_i]$$

where $\tau^{\xi,h}$ traditional externality wedge

- General model NOT subsumed by traditional theory of externalities
Nudges

- Nudge $\chi$: influences demand $c(q, w, \chi)$, possibly utility $u(c, \chi)$, but not budget $q \cdot c = w$

- Ex. decision utility $u^s(c)$, perceived price $q^s,*(q, w)$, nudgeability $\eta \geq 0$

- Agent behavior

\[ c(q, w, \chi) = \arg \max_{c} u^s(c) \text{ s.t. } q \cdot c \leq w \]

i.e.

\[ u'^s(c) = \Lambda B^s_c(q^s, c, \chi) \text{ with } \Lambda \text{ such that } q \cdot c(q, w, \chi) = w \]

- Nudge as a tax $B^s(q, c, \chi) = q^s,*(q, w) \cdot c + \chi \eta c_i$

- Nudge as an anchor $B^s(q, c, \chi) = q^s,*(q, w) \cdot c + \eta |c_i - \chi|$
Optimal Nudges

- Optimal nudge formula

\[ 0 = \frac{\partial L}{\partial \chi} = \sum_h [\lambda (\tau - \tau^{\xi,h} - \tilde{\tau}^{b,h}) \cdot c^h_\chi + \beta^h u^h_{\chi v^h_w}] \]

- Integrates nudges in canonical optimal taxation framework
Taking Stock

- So far:
  - general taxation motive
  - general behavioral biases
  - generalize canonical optimal tax formulas
  - sufficient statistics approach

- Now:
  - specialize model: behavioral bias, taxation motive
  - concrete lessons for taxes
Ramsey: Inverse Elasticity Rule

- Representative agent with quasilinear utility

\[ u(c) = c_0 + \sum_{i>0} u^i(c_i) \]

- Misperception of taxes \( \tau_i^s = m_i \tau_i \) (salience)

- Social objective, limit of small taxes (\( \Lambda = \lambda - 1 \) small)

\[ L(\tau) = -\sum_i \frac{1}{2}(\tau_i^s)^2 \psi_i y_i + \Lambda \sum_i \frac{\tau_i}{p_i} y_i \]

where \( \psi_i \) rational demand elasticity, \( y_i \) expenditure with no tax
Ramsey: Inverse Elasticity Rule

- Behavioral elasticity $m_i \psi_i$

- Behavioral Ramsey formula

\[
\frac{\tau_i}{p_i} = \frac{\Lambda}{m_i^2 \psi_i}
\]

- Contrast with traditional Ramsey formula

\[
\frac{\tau_i^R}{p_i} = \frac{\Lambda}{\psi_i}
\]

- Taxation and salience: $\frac{1}{m_i^2}$
Pigou: Dollar for Dollar Principle

- Representative agent with quasilinear utility
- One taxed good with price $p$ and externality $-\xi c$
- Inattention to tax $\tau^s = m\tau$
- Behavioral Pigou formula
  \[ \tau = \frac{\xi}{m} \]
- Contrast with traditional Pigou formula
  \[ \tau^R = \xi \]
- Taxation and salience: Pigou $\frac{1}{m}$ vs. Ramsey $\frac{1}{m^2}$
Ramsey and Pigou: Heterogeneous Attention

- Heterogeneous attention $m_i^h$
- Additional deadweight loss from misallocation
- Behavioral Ramsey and Pigou formula become

$$\frac{\tau_i}{p_i} = \frac{\Lambda}{\psi_i \mathbb{E}[m_i^h]^2} = \frac{\Lambda}{\psi_i \left( \mathbb{E}[m_i^h]^2 + \text{var}[m_i^h] \right)}$$

$$\tau^* = \frac{\mathbb{E}[\xi^h m^h]}{\mathbb{E}[m^h]^2} = \frac{\mathbb{E}[\xi^h] \mathbb{E}[m^h] + \text{cov}(\xi^h, m^h)}{\mathbb{E}[m^h]^2 + \text{var}[m^h]}$$
Pigou: Taxes vs. Quantity Restrictions

- Revisit traditional presumption:
  Pigouvian taxes $>\,\text{quantity restrictions}$

- Heterogeneity:
  - externality $\xi_h$
  - misperception $m_h$

- Quasilinear + quadratic utility:
  - social bliss point $c_h^*$
  - “elasticity” (slope) of demand $\Psi$
Pigou: Taxes vs. Quantity Restrictions

- Quantity restrictions better than taxation iff

\[
\frac{1}{\Psi} \text{var}(c_h^*) \leq \Psi \frac{E[\xi_h^2] E[m_h^2] - (E[\xi_h m_h])^2}{E[m_h^2]}
\]

1. enough heterogeneity in attention \((m_h)\) or externality \((\xi_h)\)
2. not too much heterogeneity in preferences \((c_h^*)\)
3. high demand elasticity \((\Psi \text{ high})\)
Useful Simple Parametrization

- Experienced utility $u^h(c_0, C) = c_0 + U^h(C) - \xi$

- Decision utility $u^{s,h}(c_0, C) = c_0 + U^{s,h}(C) - \xi$

- Misperception $\tau^{s,h} = \tau M^h$

- Internality wedge $\tau^{l,h} = U^{s,h}_C(C) - U^h_C(C)$

- Internality/externality wedge $\tau^{X,h} = \frac{\beta^h}{\lambda} \tau^{l,h} + \tau^{\xi,h}$

- Misoptimization wedge $\tau^{b,h} = \tau^{l,h} + \tau - \tau^{s,h}$

- Optimal tax

$$\tau = \left( \sum_h M^{h'} S^{h',r} (I - \left( I - M^h \right) \frac{\gamma^{\xi,h}}{\lambda}) \right)^{-1} \cdot \sum_h [M^{h'} S^{h',r} \tau^{X,h} - (1 - \frac{\gamma^{h',\xi}}{\lambda}) c^h]$$
Pigou: Principle of Targeting

- Traditional principle of targeting:
  - tax eternality good
  - do not tax complements
  - do not subsidize substitutes

- Behavioral (heterogeneous attention):
  - tax complements
  - subsidize substitutes

- cf Allcott, Mullainathan, Taubinsky ('14): if consumers partly “forget” about cost of gas when purchasing car, subsidize fuel efficiency, or mandate fuel-efficiency standards
Pigou: Principle of Targeting

- Use simple parametrization
- Two goods, negative externality from good 1
  \[ \tau_1^X = \xi > 0 \text{ and } \tau_2^X = 0 \]

- Homogenous preferences, decision=experienced, heterogenous misperceptions, no redistributive or revenue raising motive

- Optimal tax on good 2
  \[ \tau_2 = \frac{S_{11}^r S_{12}^r E \left[ m_{1,h} \right] \left[ E \left[ m_{1h}^2 \right] E \left[ m_{2h} \right] - E \left[ m_{1h} m_{2h} \right] E \left[ m_{1h} \right] \right]}{\det E \left[ M^h S^r M^h \right]} \tau_1^X \]
  - \( \tau_2 = 0 \) with homogenous misperceptions
  - \( \tau_2 > 0 \) iff \( S_{12}^r > 0 \) with heterogenous misperceptions
    (if not too correlated)
Vouchers and Mental Accounts

- Two goods, food (1) and non-food (2)
- Internality from food (decisions vs. experienced utility)
  \[ u^s(c_1, c_2) = \frac{c_1^{\alpha_1^s} c_2^{\alpha_2^s}}{\alpha_1^1 \alpha_2^2} \quad \text{vs.} \quad u(c_1, c_2) = \frac{c_1^{\alpha_1} c_2^{\alpha_2}}{\alpha_1^1 \alpha_2^2} \]
  with \( \alpha_1^s + \alpha_2^s = \alpha_1 + \alpha_2 = 1 \) and \( \alpha_1^s < \alpha_1 \)
- Mental accounting (perceived vs. actual budget constraint)
  \[ c_1 + c_2 + \kappa_1 |c_1 - \omega_1^d| = w \quad \text{vs.} \quad c_1 + c_2 = w \]
- Transfers \( t \) and food voucher \( b \)
  \[ w = w^* + t + b \quad \text{and} \quad \omega_1^d = \alpha_1^s w + \beta b \]
- Government objective function
  \[ \frac{[u(c(t, b))]^{1-\sigma}}{1-\sigma} - \lambda (t + b) \]
Vouchers and Mental Accounts

- MPCF from voucher \((\alpha_1^s + \beta)\) > MPCF from transfer \((\alpha_1^s)\), even if voucher inframarginal \((c_1 > b)\)

- Given \(T = t + b\), optimal voucher

\[
\frac{b}{w} = \frac{\alpha_1 - \alpha_1^s}{\beta}
\]

- Higher overall transfers iff weak taste for redistribution \((\sigma < 1)\)

- Higher welfare with vouchers.
Assume 

\[ u^{s,h}(c_1, c_2) = \frac{c_1^{\alpha_{1}^{h,s}} c_2^{\alpha_{2}^{h,s}}}{\alpha_{1}^{\alpha_{1}} \alpha_{2}^{\alpha_{2}}} \quad \text{and} \quad u^{h}(c_1, c_2) = \frac{c_1^{\alpha_{1}} c_2^{\alpha_{2}}}{\alpha_{1}^{\alpha_{1}} \alpha_{2}^{\alpha_{2}}} \]

with \( \alpha_{1}^{h,s} + \alpha_{2}^{h,s} = \alpha_{1} + \alpha_{2} = 1 \)

Samuelsonian welfare function \( \sum_{h} \left[ u^{h,s}(c_{1}^{h}, c_{2}^{h}) \right]^{1-\sigma} \)

Linear income tax \( \tau_{z} \) and a lump sum rebate
Mistakes and Redistribution

- Strong preference for redistribution ($\sigma > 1$): larger behavioral biases (reductions in $A^h$) for poor lead to more redistribution (higher $\tau_z$)

- Reverse if weak preference for redistribution ($\sigma < 1$)

- Mistakes lower utility and marginal utility of wealth, ambiguous effect on social marginal utility of income $\gamma^h$:

$$\nu^h(z) = A^h z, \quad A^h = \left(\frac{\alpha_1^{h,b}}{\alpha_1}\right)^{\alpha_1} \left(\frac{\alpha_2^{h,b}}{\alpha_2}\right)^{\alpha_2} \leq 1$$

$$\gamma^h = \left(A^h z\right)^{-\sigma} \quad A^h = z^{-\sigma} \left(A^h\right)^{1-\sigma}$$
Internalities and Redistribution

- Use simple parametrization
- No externalities, no misperceptions, decision=experienced except...
- ...good 1 only consumed by type $h^*$ with internality $\tau^l,h^* > 0$
- Optimal tax

$$
\frac{\tau_1}{q_1} = \frac{1 - \frac{\gamma h^*}{\lambda}}{\psi_1} + \frac{\gamma h^*}{\lambda} \frac{\tau^l,h^*}{q_1}
$$

- Sign ambiguous, internality correction vs. redistribution
- Ex. “sugary sodas” (cf. also Lockwood and Taubinsky ’15)
Aversive Nudges vs. Taxes

- Allow for misperceptions
- Use \( U^h(c) = \frac{a^h c - \frac{1}{2} c^2}{\Psi} \)
- Nudge as a tax \( c^h^* (\tau, \chi) = c^h^*_0 - \Psi (m^h^* \tau + \chi \eta^h^*) \)
- Aversive nudge \( u^h^* (c, \chi) = u^h^* (c) - \iota^h^* \chi c_1 \)
- Tax dominates nudge iff
  \[
  \frac{\lambda - \gamma^h^*}{m^*_h} > \frac{-\iota^h^* \gamma^h^*}{\eta^h^*}
  \]
- “Nudge the poor, tax the rich”
Mirrlees (1971)

- General behavioral biases with non-linear income tax $T(z)$
- Behavioral Saez formula (Saez 2001)

Sufficient statistics:
- traditional: elasticity of labor supply, welfare weights, hazard...
- behavioral: misoptimization wedge, behavioral cross-influence
Behavioral Saez Formula

\[
\frac{T'(z^*) - \tilde{\tau}^b(z^*)}{1 - T'(z^*)} + \int_0^\infty \omega(z^*, z) \frac{T'(z) - \tilde{\tau}^b(z)}{1 - T'(z)} dz
\]

\[
= \frac{1}{\zeta^c(z^*)} \frac{1 - H(z^*)}{z^* h^*(z^*)} \int_{z^*}^\infty e^{-\int_{z^*}^{\infty} \rho(s) ds} \left(1 - g(z) - \frac{\eta(z) \tilde{\tau}^b(z)}{1 - T'(z)}\right) \frac{h(z)}{1 - H(z^*)},
\]

where

\[
\rho(z) = \frac{\eta(z)}{\zeta^c(z)} \frac{1}{z},
\]

\[
\omega(z^*, z) = \frac{\zeta^c_{Q_{z^*}}(z) - \int_{z^*}^\infty e^{-\int_{z^*}^{z'} \rho(s) ds} \rho(z') \zeta^c_{Q_{z^*}}(z) dz'}{\zeta^c(z^*)} \frac{zh^*(z)}{z^* h^*(z^*)},
\]

and traditional Saez formula obtains with \(\tilde{\tau}^b(z) = \zeta^c_{Q_{z^*}} = 0\).
Some Applications (See Paper)

- Nonzero taxes at top and bottom (bounded skills)
- Behavioral Saez top tax formula (unbounded skills)
- Possibility of negative marginal income tax rates
  - rationalization of EITC if poor undervalue benefits of work
  - see also Lockwood (JMP, in progress)
- Schmeduling (Liebman and Zeckhauser 2004): confusion of average for marginal tax rates
Additional General Results (See Paper)

- **Endogenous attention:**
  - attention as a good, optimal/suboptimal attention
  - typically lower taxes with endogenous attention

- **Salience as policy choice:**
  - low salience to raise taxes
  - high salience to correct for internalities or externalities
Additional General Results (See Paper)

- **Diamond-Mirrlees (1971):**
  - traditional $\rightarrow$ productive efficiency (ex. no taxes on intermediate goods) if complete set of taxes on final goods
  - behavioral $\rightarrow$ productive efficiency if complete set of salient taxes on final goods
  - in both cases, no productive efficiency $\rightarrow$ supply elasticities and incidence enter tax formulas

- **Atkinson-Stiglitz (1976):**
  - traditional $\rightarrow$ uniform commodity taxation if separable preferences
  - behavioral $\rightarrow$ not true anymore in general, e.g. tax more non-salient goods and high internality goods
Conclusion

- Traditional optimal taxation theory:
  - general using traditional price theory
  - unification $\rightarrow$ tax formulas with sufficient statistics
  - concrete lessons

- Behavioral optimal taxation theory:
  - general using behavioral price theory
  - unification $\rightarrow$ tax formulas with old and new sufficient statistics
  - new concrete lessons