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-Stigltiz...)

Example: Decision vs. Experienced Utility

- Decision utility u^s and experience utility u
- Agent behavior

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

- Ex. internalities from temptation, hyperbolic discounting...

Example: Misperception

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

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

i.e.

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

- 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 $\boldsymbol{q}\cdot\boldsymbol{c} = w$

General Model: Behavioral Price Theory

- Two primitives:
  - Marshallian demand function $\boldsymbol{c}(\boldsymbol{q}, w)$ with $\boldsymbol{q} \cdot \boldsymbol{c}(\boldsymbol{q}, w) = w$
  - "experienced" utility function $u(\boldsymbol{c})$
- Indirect utility function $v(\boldsymbol{q}, w) = u(\boldsymbol{c}(\boldsymbol{q}, w))$
- Misoptimization wedge $\boldsymbol{\tau}^b = \boldsymbol{q} - \frac{u_{\boldsymbol{c}}(\boldsymbol{c}(\boldsymbol{q},w))}{v_w(\boldsymbol{q},w)}$
- Slutsky matrix $\boldsymbol{S}_j^C(\boldsymbol{q}, w) = \boldsymbol{c}_{q_j}(\boldsymbol{q}, w) + \boldsymbol{c}_w(\boldsymbol{q}, w) c_j(\boldsymbol{q}, w)$
- Behavioral Roy identity $\frac{v_{q_j}(\boldsymbol{q},w)}{v_w(\boldsymbol{q},w)} = -c_j - \boldsymbol{\tau}^b \cdot \boldsymbol{S}_j^C$

Mapping to the General Model: Concrete Examples

- Decision vs. experienced utility model:
  - misoptimization wedge $\boldsymbol{\tau}^b = \frac{u_c^s}{v_w^s} - \frac{u_c}{v_w}$
  - $\tau_i^b > 0$ for "tempting" goods
  - Slutsky $S_{ij} = S_{ij}^s$
- Misperception model:
  - misoptimization wedge $\boldsymbol{\tau}^b = \boldsymbol{q} - \boldsymbol{q}^s$
  - $\tau_i^b > 0$ for goods with non-salient taxes
  - Slutsky $S_{ij}^H = \sum_k S_{ik}^r \frac{\partial q_k^s(\boldsymbol{q}, w)}{\partial q_j}$

Many-Person Ramsey (Diamond 1975)

- Social objective function

$$L(\boldsymbol{\tau}) = W(v^h(\boldsymbol{p}+\boldsymbol{\tau}, w)) + \lambda \sum_h [\boldsymbol{\tau} \cdot \boldsymbol{c}^h(\boldsymbol{p}+\boldsymbol{\tau}, w) - w]$$

- Optimal tax formula

$$0 = \frac{\partial L(\boldsymbol{\tau})}{\partial \tau_i} = \sum_h [(\lambda - \gamma^h) c_i^h + \lambda (\boldsymbol{\tau} - \widetilde{\boldsymbol{\tau}}^{b,h}) \cdot \boldsymbol{S}_i^{C,h}]$$

- Sufficient statistics:
  - social marginal welfare weight $\beta^h = W_{v^h} v_w^h$
  - social marginal utility of income $\gamma^h = W_{v^h} v_w^h + \lambda \boldsymbol{\tau} \cdot \boldsymbol{c}_w^h$
  - substitution elasticities $\boldsymbol{S}_i^{C,h}$
  - weighted misoptimization wedge $\widetilde{\boldsymbol{\tau}}^{b,h} = \frac{\beta^h}{\lambda} \boldsymbol{\tau}^{b,h}$

Many-Person Ramsey (Diamond 1975)

- Optimal tax formula

$$0 = \frac{\partial L(\boldsymbol{\tau})}{\partial \tau_i} = \sum_h [(\lambda - \gamma^h) c_i^h + \lambda(\boldsymbol{\tau} - \widetilde{\boldsymbol{\tau}}^{b,h}) \cdot \boldsymbol{S}_i^{C,h}]$$

- Three terms:
  - mechanical $(\lambda - \gamma^h) c_i^h$
  - substitution $\lambda \boldsymbol{\tau} \cdot \boldsymbol{S}_i^{C,h}$
  - misoptimization $-\lambda \widetilde{\boldsymbol{\tau}}^{b,h} \cdot \boldsymbol{S}_i^{C,h}$
- Additional condition if lump sum taxes $\sum_h (\lambda - \gamma_h) = 0$

Many-Person Ramsey (Diamond 1975)

- Assume symmetric Slutsky matrices $\boldsymbol{S}_{ij}^{C,h} = \boldsymbol{S}_{ji}^{C,h}$
- Then tax formula expressible in "discouragement" form

$$\frac{-\sum_{h,j} \tau_j \boldsymbol{S}_{ij}^{C,h}}{c_i} = 1 - \frac{\bar{\gamma}}{\lambda} - cov\left(\frac{\gamma^h}{\lambda}, \frac{Hc_i^h}{c_i}\right) - \frac{\sum_{h,j} \tilde{\tau}_j^{b,h} \boldsymbol{S}_{ij}^{C,h}}{c_i}$$

Pigou (Sandmo 1975)

- Externality $\xi = \xi((\boldsymbol{c}^h))_{h=1\ldots H}$, indirect utility $v^h(q, w, \xi)$
- Optimal tax formula

$$0 = \frac{\partial L(\boldsymbol{\tau})}{\partial \tau_i} = \sum_h [(\lambda - \gamma^{\xi,h})c_i^h + \lambda(\boldsymbol{\tau} - \boldsymbol{\tau}^{\xi,h} - \widetilde{\boldsymbol{\tau}}^{b,h}) \cdot \boldsymbol{S}_i^{C,h}]$$

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

- General model NOT subsumed by traditional theory of externalities

Nudges

- Nudge χ: influences demand $\boldsymbol{c}(\boldsymbol{q}, w, \chi)$, possibly utility $u(\boldsymbol{c}, \chi)$, but not budget $\boldsymbol{q} \cdot \boldsymbol{c} = w$
- Ex. decision utility $u^s(\boldsymbol{c})$, perceived price $\boldsymbol{q}^{s,*}(\boldsymbol{q}, w)$, nudgeability $\eta \geq 0$
- Agent behavior

$$\boldsymbol{c}(\boldsymbol{q}, w, \chi) = \arg smax_{\boldsymbol{c}|u^s, B^s} u^s(\boldsymbol{c}) \text{ s.t. } \boldsymbol{q} \cdot \boldsymbol{c} \leq w$$

i.e.

$$u^{s\prime}(\boldsymbol{c}) = \Lambda B^s_{\boldsymbol{c}}(\boldsymbol{q}^s, \boldsymbol{c}, \chi) \text{ with } \Lambda \text{ such that } \boldsymbol{q} \cdot \boldsymbol{c}(\boldsymbol{q}, w, \chi) = w$$

- Nudge as a tax $B^s(\boldsymbol{q}, \boldsymbol{c}, \chi) = \boldsymbol{q}^{s,*}(\boldsymbol{q}, w) \cdot \boldsymbol{c} + \chi \eta c_i$
- Nudge as an anchor $B^s(\boldsymbol{q}, \boldsymbol{c}, \chi) = \boldsymbol{q}^{s,*}(\boldsymbol{q}, w) \cdot \boldsymbol{c} + \eta |c_i - \chi|$

Optimal Nudges

- Optimal nudge formula

$$0 = \frac{\partial L}{\partial \chi} = \sum_h [\lambda(\boldsymbol{\tau} - \boldsymbol{\tau}^{\xi,h} - \widetilde{\boldsymbol{\tau}}^{b,h}) \cdot \boldsymbol{c}_\chi^h + \beta^h \frac{u_\chi^h}{v_w^h}]$$

- 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

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

- Misperception of taxes $\tau_i^s = m_i \tau_i$ (salience)
- Social objective, limit of small taxes ($\Lambda = \lambda - 1$ small)

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

where ψ_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

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

$$\tau^* = \frac{\mathbb{E}\left[\xi^h m^h\right]}{\mathbb{E}\left[{m^h}^2\right]} = \frac{\mathbb{E}\left[\xi^h\right]\mathbb{E}\left[m^h\right] + cov\left(\xi^h, m^h\right)}{\mathbb{E}\left[m^h\right]^2 + var\left[m^h\right]}$$

Pigou: Taxes vs. Quantity Restrictions

- Revisit traditonal presumption:

  $$\text{Pigouvian taxes} > \text{quantity restrictions}$$

- Heterogeneity:
  - externality ξ_h
  - mispereception m_h
- Quasilinear + quadratic utility:
  - social bliss point c_h^*
  - "elasticity" (slope) of demand Ψ

Pigou: Taxes vs. Quantity Restrictions

- Quantity restrictions better than taxation iff

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

1. enough heterogeneity in attention (m_h) or externality (ξ_h)
2. not too much heterogeneity in preferences (c_h^*)
3. high demand elasticity (Ψ high)

Useful Simple Parametrization

- Experienced utility $u^h(c_0, \boldsymbol{C}) = c_0 + U^h(\boldsymbol{C}) - \xi$
- Decision utility $u^{s,h}(c_0, \boldsymbol{C}) = c_0 + U^{s,h}(\boldsymbol{C}) - \xi$
- Misperception $\boldsymbol{\tau}^{s,h} = \boldsymbol{\tau} M^h$
- Internality wedge $\boldsymbol{\tau}^{I,h} = U^{s,h}_{\boldsymbol{C}}(C) - U^h_{\boldsymbol{C}}(\boldsymbol{C})$
- Internality/externality wedge $\boldsymbol{\tau}^{X,h} = \frac{\beta^h}{\lambda}\boldsymbol{\tau}^{I,h} + \boldsymbol{\tau}^{\xi,h}$
- Misoptimization wedge $\boldsymbol{\tau}^{b,h} = \boldsymbol{\tau}^{I,h} + \boldsymbol{\tau} - \boldsymbol{\tau}^{s,h}$
- Optimal tax

$$
\begin{aligned}
\boldsymbol{\tau} \quad = \quad & (\sum_h \boldsymbol{M}^{h\prime} \boldsymbol{S}^{h,r} (I - \left(I - \boldsymbol{M}^h\right) \frac{\gamma^{\xi,h}}{\lambda}))^{-1} \\
& \cdot \sum_h [\boldsymbol{M}^{h\prime} \boldsymbol{S}^{h,r} \boldsymbol{\tau}^{X,h} - (1 - \frac{\gamma^{h,\xi}}{\lambda}) \boldsymbol{c}^h]
\end{aligned}
$$

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

$$\tau_1^X = \xi > 0 \quad \text{and} \quad \tau_2^X = 0$$

- 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)

$$u^s(c_1, c_2) = \frac{c_1^{\alpha_1^s} c_2^{\alpha_2^s}}{\alpha_1^{\alpha_1} \alpha_2^{\alpha_2}} \quad \text{vs.} \quad u(c_1, c_2) = \frac{c_1^{\alpha_1} c_2^{\alpha_2}}{\alpha_1^{\alpha_1} \alpha_2^{\alpha_2}}$$

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)

$$c_1 + c_2 + \kappa_1 \left| c_1 - \omega_1^d \right| = w \quad \text{vs.} \quad c_1 + c_2 = w$$

- 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

$$\frac{\left[u(\boldsymbol{c}(t, b))\right]^{1-\sigma}}{1-\sigma} - \lambda(t + b)$$

Vouchers and Mental Accounts

- MPCF from voucher $(\alpha_1^s + \beta)$ > MPCF from transfer (α_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.

Mistakes and Redistribution

- 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 $\sum_h \frac{\left[u^{h,s}(c_1^h,c_2^h)\right]^{1-\sigma}}{1-\sigma}$
- Linear income tax τ_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 τ_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 γ^h:

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

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

Internalities and Redistribution

- Use simple parametrization
- No externalities, mo misperceptions, decision=experienced except...
- ...good 1 only consumed by type h^* with internality $\tau_1^{I,h^*} > 0$
- Optimal tax

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

- 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 $c^{h^*}(\tau, \chi) = c_0^{h^*} - \Psi\left(m^{h^*}\tau + \chi\eta^{h^*}\right)$
- Aversive nudge $u^{h^*}(\boldsymbol{c}, \chi) = u^{h^*}(\boldsymbol{c}) - \iota^{h^*}\chi c_1$
- 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

$$\frac{T'(z^*)-\widetilde{\tau}^b(z^*)}{1-T'(z^*)}+\int_0^\infty \omega(z^*,z)\frac{T'(z)-\tilde{\tau}^b(z)}{1-T'(z)}dz$$
$$=\frac{1}{\zeta^c(z^*)}\frac{1-H(z^*)}{z^*h^*(z^*)}\int_{z^*}^\infty e^{-\int_{z^*}^z \rho(s)ds}\left(1-g(z)-\frac{\eta(z)\tilde{\tau}^b(z)}{1-T'(z)}\right)\frac{h(z)}{1-H(z^*)},$$

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