Diamond–Mirrlees meets Sims: Optimal Taxation with Rational Inattention

George-Marios Angeletos
Northwestern University

Matias Bayas-Erazo
University of Zürich

NBER SI Macro Public Finance, July 2026

Motivation

Q: How should we set commodity taxes when consumers are **inattentive**?

- Consumers misperceive, ignore, or underreact to taxes

  [Chetty et al., 2009; Rees-Jones and Taubinsky, 2020 ...]

- Existing (behavioral) literature: [Goldin, 2015; Gerritsen, 2016; Farhi and Gabaix, 2020 ...]

  1. *Main lessons:* we should use taxes to ***correct*** and/or ***exploit*** inattention
  2. *Underlying assumption:* inattention is a mistake

- **Today:** treat inattention as rational $\implies$ planner internalizes attention costs

Main takeaway: irrelevance, no need to pay special attention to inattention!

This paper

Introduce **general form** of rational inattention into multi-good Ramsey taxation problem

- Consumers choose based on $\underbrace{\textbf{imperfect signals}}_{\text{costly info acquisition, costly cognition, or costly optimization}}$ of fundamentals, prices, and taxes

- Non-paternalistic planner internalizes attention costs

- **Strongest irrelevance result** under two restrictions on attention costs:

    1. *Invariance* [Hébert and La'O, 2023; Angeletos and Sastry, 2025]
    2. *State separability* [Flynn and Sastry, 2023]

- Relax restrictions to identify when and why inattention *might* matter

Main results

1. Irrelevance: standard optimal tax formulas survive under **benchmark** conditions
   - just look at **effective elasticities**, ignore whatever is "under the hood"
   - inattention can *increase* elasticities and justify **lower** taxes

2. Away from benchmark, two motives for state-dependent taxes
   i. Ramsey: reshape uncertainty, relaxing usual tradeoff when costs are **not separable**
   ii. Pigou: stabilize or "simplify" prices to ease **non-invariant** cognitive costs

3. Salience can be interpreted as a violation of invariance, bridge to behavioral PF
   - it matters only when **producer prices are endogenous**
   - *still irrelevant* with the right (GE) elasticity

Plan for Today

1. A simple model of rational inattention
    - A first irrelevance result
2. How general is irrelevance?
3. When does rational inattention matter?
    - State dependence: relaxing separability
    - Salience: relaxing invariance
4. Conclusion

A simple model of rational inattention

Environment

- Two goods: taxed good x and untaxed numeraire y
- **Representative consumer** with quasilinear preferences over (x, y) and income w

$$u(x) + y, \qquad q\,x + y = w$$

- **Exogenous** state $\theta \in \Theta$ with distribution $\pi \equiv (\pi_\theta)_{\theta \in \Theta}$
- θ determines producer price p_θ, tax policy sets τ_θ, and *true* **consumer price** is

$$q_\theta = p_\theta + \tau_\theta, \qquad \boldsymbol{Q} = (q_\theta)_{\theta \in \Theta}$$

Rational inattention as stochastic demand

Two flavors of RI

- Consumer has prior ϕ over **possible prices** $q \in \mathcal{Q}_\phi$, where $\mathcal{Q}_\phi = \{q : \phi(q) > 0\}$

- She chooses a **stochastic** choice rule $F = \{F(\cdot \mid q)\}_{q \in \mathcal{Q}_\phi}$
  - $F(\cdot \mid q)$ is a distribution over consumption when consumer price is q

- **Individual problem:**

$$v(\phi) = \max_{F \in \mathcal{F}} \left\{ \sum_{q \in \mathcal{Q}_\phi} \phi(q) \int_{\mathcal{X}} [u(x) - qx] dF(x \mid q) - C(F, \phi) \right\}.$$

- $C(F, \phi)$ captures costly information, costly cognition, or costly control

Equilibrium and Ramsey problem

Definition

An equilibrium is a price system $\boldsymbol{Q}$, a stochastic choice rule F, and a prior ϕ such that

1. *Optimality.* F solves the consumer's problem given ϕ
2. *Consistency.* $\phi = \phi^*_{\boldsymbol{Q}}$, where $\phi^*_{\boldsymbol{Q}}(q) = \sum_{\{\theta:\, q_\theta = q\}} \pi_\theta$ for every q

Ramsey problem: Planner chooses after-tax prices $\boldsymbol{Q}$ to maximize:

– consumer's indirect utility $v(\cdot)$ plus shadow value of tax revenue

A simple irrelevance result

- Suppose attention costs are given by:

$$C(F, \phi) = \sum_{q \in \mathcal{Q}_\phi} \phi(q) \int_{\mathcal{X}} K(f(x \mid q))dx$$

- This imposes **two restrictions**:
  - costs are additively separable across states
  - costs are invariant $\approx$ independent of how underlying uncertainty is labeled

Proposition

In every state, optimal tax satisfies

$$\frac{\tau}{q} = \frac{\Lambda}{\epsilon(q)}$$

where $\epsilon(q) \equiv -\frac{\partial \log \bar{x}(q)}{\partial \log q}$ is **price elasticity** of aggregate demand and $\Lambda \equiv 1 - \lambda^{-1}$

Proof sketch: As-if attentive representation

- Inattention can be folded into modified utility function:

$$\hat{u}(\bar{x}) = \max_{f \geq 0} \left\{ \int_{\mathcal{X}} \left[u(x)f(x) - K(f(x)) \right] dx \quad \text{s.t.} \quad \int_{\mathcal{X}} xf(x)\, dx = \bar{x}, \quad \int_{\mathcal{X}} f(x)dx = 1 \right\}$$

- Indirect utility and demand associated with $\hat{u}$ are

$$\hat{v}(q) = \max_{\bar{x}} \{\hat{u}(\bar{x}) - q\bar{x}\}, \qquad \hat{x}(q) = \arg\max_{\bar{x}} \{\hat{u}(\bar{x}) - q\bar{x}\}$$

- The planner therefore faces a standard Ramsey problem:

$$\max_{\boldsymbol{Q}} \sum_{\theta \in \Theta} \pi_\theta \left[\hat{v}(q_\theta) + \lambda (q_\theta - p_\theta) \hat{x}(q_\theta) \right]$$

- *As-if* demand comes from **fully attentive consumer** whose indirect utility is $\hat{v}(q)$

$\implies$ elasticity of $\hat{x}$ is a *sufficient statistic* for optimal taxes!

How general is irrelevance?

General environment

- Arbitrary number of taxed goods $i \in \mathcal{I}$ and household types $h \in \mathcal{H}$ with masses μ^h
- **Production:** $H(\boldsymbol{x}^S, y^S, \theta) \leq 0$ operated by fully attentive competitive firm
- Agents' subjective uncertainty captured by **cognitive state** $z = (\theta, \boldsymbol{p}, \boldsymbol{q})$ with **prior** ϕ
- Type h chooses $\{F^h(\cdot \mid z)\}_{z \in Z_\phi}$ where $Z_\phi = \{z : \phi(z) > 0\}$
  - $F^h(\cdot \mid z)$ is distribution over *actions* $a = (\boldsymbol{x}, y)$ when cognitive state is z

Individual problem:

$$V^h(\phi) = \max_{F^h} \sum_{z \in Z_\phi} \phi(z) \int u^h(a, \theta)\, dF^h(a \mid z) - C^h(F^h, \phi) \;\text{ s.t. } \int (\boldsymbol{q} \cdot \boldsymbol{x} + y)\, dF^h(a \mid z) \leq w^h_\theta, \; \forall z$$

Note: indexing budget constraints by z relies on complete markets

The two benchmark restrictions, in words

▸ Formal definitions

1. Invariance: *same uncertainty, same behavior* $\Longrightarrow$ *same cost*

   [Hébert and La'O, 2023; Angeletos and Sastry, 2025]

   – satisfied by mutual information: agents "see through" the price system

   $\Rightarrow$ no **Pigouvian correction**

2. State separability: fine-tuning choices in one state doesn't affect costs in others

   [Flynn and Sastry, 2023]

   $\Rightarrow$ indirect utility **separable across states**

$\Longrightarrow$ **Together:** planner faces a *standard, state-separable* Ramsey problem

General irrelevance result

Define marginal social value of income for type h in state θ [Diamond, 1975]

$$\gamma_\theta^h = \beta^h \partial_w v_\theta^h + \lambda_\theta \boldsymbol{\tau}_\theta \cdot \partial_w \bar{\mathbf{x}}_\theta^h$$

Theorem

If attention costs are invariant and state separable, optimal taxes satisfy

$$\sum_{h \in \mathcal{H}} \mu^h \left[(\lambda_\theta - \gamma_\theta^h) \bar{\mathbf{x}}_\theta^h + \lambda_\theta \mathbf{S}_\theta^h \boldsymbol{\tau}_\theta \right] = \mathbf{0} \qquad \forall \theta.$$

- *As-if* demand $\bar{\mathbf{x}}_\theta^h = \hat{\mathbf{x}}^h(\theta, \mathbf{q}_\theta, w_\theta^h)$ with associated Slutsky matrix $\mathbf{S}_\theta^h$

When might rational inattention matter?

Relaxing separability calls for state dependence

Mutual-information example

- Attention in state θ may depend on incentives in other states (e.g. mutual info)
- As-if demand now depends on full price system: $\bar{\boldsymbol{x}}^h_{\theta'} = \tilde{\boldsymbol{x}}^h(\theta', \boldsymbol{Q}, \boldsymbol{W}^h)$
  - cross-state Slutsky $\boldsymbol{S}^h_{\theta,\theta'}$ defined analogously: response of demand in θ' to prices in θ

Theorem *(weak irrelevance)* Moment formula

With invariant but non-separable attention costs,

$$\sum_{h\in\mathcal{H}} \mu^h \left[(\pi_\theta \lambda_\theta - \gamma^h_\theta) \bar{\boldsymbol{x}}^h_\theta + \sum_{\theta'} \pi_{\theta'} \lambda_{\theta'} (\boldsymbol{S}^h_{\theta,\theta'})^\top \boldsymbol{\tau}_{\theta'} \right] = \mathbf{0} \qquad \forall \theta.$$

- Same formula as in fully attentive economy w/ non-separable preferences

Relaxing invariance: when does inattention become a target?

- Without invariance, agents cannot "see through" relabelings of $z = (\theta, \boldsymbol{p}, \boldsymbol{q})$
- Consistent prior $\phi^*_{\boldsymbol{Q},\boldsymbol{P}}$ no longer drops out of as-if utility and demand:

$$V^h(\boldsymbol{Q}, \boldsymbol{W}^h, \phi^*_{\boldsymbol{Q},\boldsymbol{P}}), \qquad \boldsymbol{X}^h(\boldsymbol{Q}, \boldsymbol{W}^h, \phi^*_{\boldsymbol{Q},\boldsymbol{P}})$$

- Prior acts like an endogenous taste shifter
  - depends on price distribution, hence on taxes
- Split $q_\theta = p_\theta + \tau_\theta$ is no longer a mere relabeling:
  - taxes reshape cognitive burden of price system (**Pigou**)
  - demand can respond to split between producer prices and taxes (**salience**)

General formula with non-invariant costs

Theorem

With general attention costs and (possibly) endogenous production,

$$\sum_h \mu^h \left[(\pi_\theta \lambda_\theta - \gamma_\theta^h) \bar{\boldsymbol{x}}_\theta^h + \sum_{\theta'} \pi_{\theta'} \lambda_{\theta'} (\boldsymbol{S}_{\theta,\theta'}^h)^\top \boldsymbol{\tau}_{\theta'} \right] = \Xi_\theta \qquad \forall \theta.$$

- Ξ_θ is Pigouvian wedge from changing cognitive burden of prices

▸ Cognitive externality

- Slutsky matrices are **total equilibrium** responses: embed salience

▸ Unpacking salience

Zooming in on salience

- Specialize to one good, state-separable demand, and no income effects: $\bar{x}_\theta(p_\theta, q_\theta)$
- Two elasticities:

$$\underbrace{\epsilon_\theta^q = -\frac{\partial \log \bar{x}_\theta}{\partial \log q_\theta}\bigg|_{p_\theta \text{ fixed}}}_{\text{vary the tax}} \qquad \underbrace{\epsilon_\theta^p = -\frac{\partial \log \bar{x}_\theta}{\partial \log p_\theta}\bigg|_{q_\theta \text{ fixed}}}_{\text{vary the split}}$$

- Under invariance, $\epsilon_\theta^p = 0$
- Tax vs. posted-price experiments identify salience elasticity: [Chetty et al., 2009]

$$\epsilon_\theta^{posted} - \epsilon_\theta^{tax} = \frac{q_\theta}{p_\theta}\, \epsilon_\theta^p$$

$\Rightarrow$ measured salience gap maps to ϵ_θ^p

Simple optimal tax formula with salience

Corollary

$$\frac{\tau_\theta}{q_\theta} = \frac{1}{\epsilon_\theta^{GE}} \left(\Lambda_\theta - \frac{\Xi_\theta}{\pi_\theta \lambda_\theta \bar{x}_\theta} \right).$$

- Relevant elasticity after producer prices adjust:

$$\epsilon_\theta^{GE} \equiv \frac{\epsilon_\theta^q}{1 + \epsilon_\theta^p / \eta_\theta^S}$$

  where η_θ^S denotes supply elasticity

- Salience matters through $\epsilon_\theta^p / \eta_\theta^S$: no GE price response, no salience correction!

- ϵ_θ^{GE} can be identified *directly* by economy-wide tax-reform

Conclusion

- Rational inattention need not change optimal tax design
  - invariance $+$ separability $\Rightarrow$ classical formulas with **effective elasticities**
  - no license to correct or exploit inattention, optimal taxes may even be lower
- Away from benchmark, two motives for state-dependent taxes
  - Ramsey: reshape the uncertainty consumers face (non-separable costs)
  - Pigou: "simplify" prices to ease cognitive costs (non-invariant costs)
- Salience $=$ failure of invariance, and correction only matters in **general equilibrium**
  - enters optimal taxes only through $\epsilon^p_\theta/\eta^S_\theta$ and the cognitive wedge Ξ_θ
  - *still irrelevant* with the right (GE) elasticity: no separate stance on salience needed

Thank You!

matias.bayas-erazo@econ.uzh.ch

Backup: moment setup without separability

- Maintain invariance, but allow demand to depend on other price realizations
- Abstract from distributional concerns and treat states as price realizations
- Summarize the after-tax price distribution by moments

$$\sigma_k = \mathbb{E}[m_k(\boldsymbol{q})]$$

- Aggregate demand becomes $\hat{\boldsymbol{x}}(\boldsymbol{q}, \boldsymbol{\sigma})$

formula

Moment formula without separability

- Abstract from distributional concerns and treat states as price realizations
- Summarize after-tax price distribution by moments: $\sigma_k = \mathbb{E}[m_k(\boldsymbol{q})]$
- Aggregate demand becomes $\hat{\boldsymbol{x}}(\boldsymbol{q}, \boldsymbol{\sigma})$

Corollary

$$\boldsymbol{S}(\boldsymbol{q}, \boldsymbol{\sigma})^\top \boldsymbol{\tau} + \sum_k \mathbb{E}[\boldsymbol{\tau} \cdot \partial_{\sigma_k} \hat{\boldsymbol{x}}(\boldsymbol{q}, \boldsymbol{\sigma})] \partial_{\boldsymbol{q}} m_k(\boldsymbol{q}) = -\Lambda \hat{\boldsymbol{x}}(\boldsymbol{q}, \boldsymbol{\sigma})$$

- New term: how taxes move demand by changing those moments
- Still a revenue effect, not a Pigouvian wedge for inattention

back to theorem

Backup: mutual information and state-dependent taxes

- Let producer prices be log normal and restrict after-tax prices to

$$q = \chi p^{\varphi}, \qquad \mathbb{V}[\log q] = \varphi^2 v_\theta$$

- Consumers observe a noisy signal about $\log q$ and choose attention

$$\rho = \frac{\mathbb{V}[\log q]}{\mathbb{V}[\varepsilon] + \mathbb{V}[\log q]}$$

- More after-tax price volatility raises the benefit of attention
- The planner regulates attention through pass-through φ:

$$\varphi^* = 1 \text{ if } \eta = 1, \qquad \varphi^* < 1 \text{ if } \eta < 1, \qquad \varphi^* > 1 \text{ if } \eta > 1.$$

- State dependence due to Ramsey logic, not a corrective tax on inattention

back

Backup: moment setup without invariance

- Shut down producer-price feedback to isolate cognitive externality
- Price system enters attention costs through moments:

$$\sigma_k = \mathbb{E}[m_k(\boldsymbol{q})]$$

- Demand and attention costs become

$$\hat{\boldsymbol{x}}(\boldsymbol{q}, \boldsymbol{\sigma}), \qquad C^h(F^h, \boldsymbol{\sigma})$$

- Define aggregate cognitive wedge:

$$\Psi_k \equiv \sum_h \mu^h \beta^h \partial_{\sigma_k} C^h(F^h, \boldsymbol{\sigma})$$

▸ formula

Cognitive externalities and state dependence

Non-invariant moment formula

$$\boldsymbol{S}(\boldsymbol{q},\boldsymbol{\sigma})^\top \boldsymbol{\tau} + \sum_k \mathbb{E}[\boldsymbol{\tau} \cdot \partial_{\sigma_k} \hat{\boldsymbol{x}}] \partial_{\boldsymbol{q}} m_k(\boldsymbol{q}) = -\Lambda \hat{\boldsymbol{x}} + (1-\Lambda) \sum_k \Psi_k \partial_{\boldsymbol{q}} m_k(\boldsymbol{q})$$

- Left side: how taxes affect revenue through demand and attention
- Final term: Pigouvian wedge from cognitive costs Ψ_k
- Example with $K(\rho, v)$ and $v = \mathbb{V}[\log q]$:

$$K_v > 0 \Rightarrow \varphi^* < 1, \qquad K_v < 0 \Rightarrow \varphi^* > 1$$

- State dependence now simplifies or amplifies price variation for Pigouvian reasons

back to theorem

Backup: unpacking the general equilibrium salience response

- Taxes move demand, market clearing moves p, and salience moves demand again
- Initial salience response:

$$\left(\boldsymbol{K}^{h}_{\theta,\theta'}\right)^{\top} = \sum_{\vartheta} \left(\nabla_{\boldsymbol{q}_{\theta}} \bar{\boldsymbol{X}}_{\vartheta}\right)^{\top} \boldsymbol{A}_{\vartheta} \left(\nabla_{\boldsymbol{p}_{\vartheta}} \bar{\boldsymbol{x}}^{h}_{\theta'}\right)^{\top}$$

- Total response substitutes the GE feedback rounds:

$$\left(\mathcal{K}^{h}_{\theta,\theta'}\right)^{\top} = \left(\boldsymbol{K}^{h}_{\theta,\theta'}\right)^{\top} + \sum_{\kappa} \left(\boldsymbol{K}^{h}_{\theta,\kappa}\right)^{\top} \boldsymbol{A}_{\kappa} \left(\nabla_{\boldsymbol{p}_{\kappa}} \bar{\boldsymbol{X}}_{\theta'}\right)^{\top} + \cdots$$

- This channel disappears under invariance or fixed producer prices

back to theorem

Two flavors of rational inattention

back

1. Noisy information [Sims, 2003]

   - Agents condition demand on **noisy signal** $\tilde{q} = q + \sigma\varepsilon$, precision σ chosen at cost $K(\sigma)$

$$\max_{\sigma,\, x^*(\cdot)} \mathbb{E}\left[u(x^*(\tilde{q})) - q\, x^*(\tilde{q}) - K(\sigma)\right]$$

2. Costly control [Flynn and Sastry, 2023]

   - Consumer observes q but consumption **trembles**: $x = x^*(q) + \sigma(q)\varepsilon$

$$\max_{\sigma(\cdot),\, x^*(\cdot)} \mathbb{E}\left[u(x^* + \sigma\varepsilon) - q\,(x^* + \sigma\varepsilon) - K(\sigma)\right]$$

- Both reduce to choosing F with cost $C(F, \phi)$
- General model nests both plus sparsity, narrow bracketing, salience

Invariance

▸ State separability

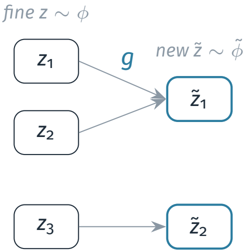
Same uncertainty, same behavior $\Rightarrow$ same attention cost, no matter how states are described

1. $(\tilde{F}, \tilde{\phi})$ is a transformation of (F, ϕ) under mapping g if

$$\tilde{\phi}(z) = \sum_{z' : g(z') = z} \phi(z'), \quad \tilde{f}(a \mid z) = \frac{\sum_{z' : g(z') = z} f(a \mid z')\phi(z')}{\tilde{\phi}(z)}$$

2. Transformation is sufficient if it preserves original behavior

$$F(a \mid z) = \tilde{F}(a \mid g(z)) \quad \forall a, z$$



Invariance

C is invariant if $C(F, \phi) = C(\tilde{F}, \tilde{\phi})$ for any sufficient transformation of (F, ϕ)

Mutual information is invariant: only depends on **joint distribution** of actions & states

◂ back

State separability

Definition

C is state separable if there exists a strictly convex K and a weighting function α such that, for any F with density f,

$$C(F, \phi) = \sum_{z \in Z_\phi} \alpha(z)\phi(z) \int_{\mathcal{A}} K(f(a \mid z))da.$$

- Cost of fine-tuning choices in one state is independent of mistakes in other states
- Some flexibility in choice of α to capture different types of decision frictions
  - ex-post optimization errors, consideration sets [cf. Flynn and Sastry, 2023]
- Key implication: gives a state-by-state *as-if* utility representation

back

References

Angeletos, G.-M. and Sastry, K. A. (2025). Inattentive economies. *Journal of Political Economy*, 133(7):000–000.

Chetty, R., Looney, A., and Kroft, K. (2009). Salience and Taxation: Theory and Evidence. *American Economic Review*, 99(4):1145–1177.

Diamond, P. A. (1975). A many-person Ramsey tax rule. *Journal of Public Economics*, 4(4):335–342.

Farhi, E. and Gabaix, X. (2020). Optimal taxation with behavioral agents. *American Economic Review*, 110(1):298–336.

Flynn, J. P. and Sastry, K. A. (2023). Strategic mistakes. *Journal of Economic Theory*, 212:105704.

Gerritsen, A. (2016). Optimal taxation when people do not maximize well-being. *Journal of Public Economics*, 144:122–139.

Goldin, J. (2015). Optimal tax salience. *Journal of Public Economics*, 131:115–123.

Hébert, B. and La'O, J. (2023). Information acquisition, efficiency, and nonfundamental volatility. *Journal of Political Economy*, 131(10):2666–2723.

Rees-Jones, A. and Taubinsky, D. (2020). Measuring "schmeduling". *The Review of Economic Studies*, 87(5):2399–2438.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690.