Behavioral Inattention: From Micro to Macro

Xavier Gabaix
Harvard

NBER Behavioral Macro Bootcamp - March 12 2020

Overview for today

- Start from a fairly unified, general model of behavioral inattention
- Useful for micro
- Can do Arrow-Debreu with it
- Work up to macro

Ongoing work: unified BR in economics

1. Micro: "A sparsity-based model of bounded rationality" (QJE '14): Fairly general and simple device,

$$\operatorname*{smax}_{a} u(a,x) \text{ subject to } b(a,x) \geq 0$$

   Basic consumer theory: Walrasian demand, Hicksian demand, Slutsky matrix. Competitive equilibrium: Arrow-Debreu, Edgeworth boxes...
2. Partial survey: "Behavioral inattention" (Handbook of Behav. Ec. '19)
3. Macro: "Behavioral Macroeconomics via Sparse Dynamic programming": Life-cycle, RBC
4. "A Behavioral New Keynesian model": monetary and fiscal policy (cond. acc. AER '20)
5. Public economics: "Optimal taxation with behavioral agents" (with E. Farhi, AER '20)

   Ongoing work:
6. Finance: in the works. Merton problem...
7. Game theory: "Some game theory with sparsity-based

Attention: Anchoring and adjustment via Gaussian signal extraction

- True (but unknown value) $x \sim \mathcal{N}\left(x^d, \sigma_x^2\right)$
- Agent gets signal $s = x + \varepsilon$ with $\varepsilon \sim \mathcal{N}\left(0, \sigma_\varepsilon^2\right)$
- Agent wants $\max_a \mathbb{E}\left[-\frac{1}{2}(a-x)^2 \,|s\right]$, so

$$0 = \mathbb{E}\left[-(a-x)\,|s\right] = \mathbb{E}\left[x \mid s\right] - a$$

$$a = \hat{x}(s) := \mathbb{E}\left[x \mid s\right] = ms + (1-m)\,x^d$$

$$m = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_\varepsilon^2} \in [0,1].$$

- Average action:

$$\bar{a}(x) := \mathbb{E}\left[a(s)\,|x\right]$$

$$\bar{a}(x) = mx + (1-m)\,x^d$$

- "People make estimates by starting from an initial value that is adjusted to yield the final answer [...]. Adjustments are typically insufficient" (Kahneman-Tversky '74)

Pretty general framework

- Rational pb: $\max_a u(a, x)$
- Behavioral: with attention m

$$\max_a u(a, x, m)$$

so action is:

$$a(x, m) = \underset{a}{\operatorname{argmax}}\, u(a, x, m).$$

- Typical:

$$u(a, x, m) = u\left(a, m_1 x_1 + (1 - m_1) x_1^d, \ldots, m_n x_n + (1 - m_n) x_n^d\right).$$

- As if x_i is replaced by

$$x_i^s := m_i x_i + (1 - m_i) x_i^d,$$

- $m_i = 0$, the agent "does not think about x_i", i.e. replaces x_i by $x_i^s = x_i^d$
- When $m_i = 1$, rational agent.

Quadratic example

- *Quadratic example*: $\max_a u(a,x)$, with $u(a,x) = -\frac{1}{2}\left(a - \sum_i b_i x_i\right)^2$,

$$a^r = \sum_{i=1}^{10^6} b_i x_i : \qquad \text{Tradional Action}$$

$$a^s = \sum_{i=1}^{10^6} b_i m_i x_i : \qquad \text{Inattentive action}$$

Unification of biases

- Much of behavioral economics reflects a form of inattention
- **Inattention to true prices and shrouding of add-on costs**
- New price is p, perceived price is

$$p^s(p,m) = mp + (1-m)\,p^d.$$

- Demand $c^s(p) = c^r(p^s(p,m))$, so

$$c^s(p)' = mc^r(p^s)'$$

  Demand is muted by factor m
- Often better:

$$p^s = (p)^m \left(p^d\right)^{1-m}.$$

- Shrouding: (Gabaix and Laibson QJE '06)
  - Base price is $p^d = b$, add-on price is $\hat{p}$
  - Perceived price is $p^s = p^d + m\hat{p}$: people underestimate the add-on
- If $\hat{p}$ is a tax, underperception of the tax

Biases via inattention: Hyperbolic discounting

- $U_0 = \sum_{t=0}^{\infty} \delta^t u_t$, and call $U_1 = \sum_{t=1}^{\infty} \delta^{t-1} u_t$, so

$$U_0 = u_0 + \delta U_1$$

- Present-biased agent perceives

$$U_0^s = u_0 + m\delta U_1.$$

- So $m = \beta$ of hyperbolic literature (Laibson QJE '97)
- Normative implication is different here
  - Traditional: games between equally legitimate selves
  - Here: The rational utility U_0 is best criterion

Prospect theory: Inattention to true probability

- Take $p \in (0,1)$ a probability and log odds $q := \ln \frac{p}{1-p}$ (which is in $(-\infty, +\infty)$)
- Perception (Zhang and Maloney '12)

$$q^s = mq + (1-m)\, q^d.$$

- With $q^s = \ln \frac{p^s}{1-p^s}$, get:

$$p^s = \pi(p) = \frac{1}{1 + \left(\frac{1-p}{p}\right)^m \left(\frac{1-p^d}{p^d}\right)^{1-m}},$$

- Much like Kahneman-Tversky (ECTA '79)'s probability weighting function
- cf Steiner Stewart AER '16, Khaw, Li, Woodford '17, Enke Graeber '20, for more on that style.

Biases via inattention: When will we se overreaction vs underreaction?

- Suppose $y_{i,t+1} = \rho_i y_{it} + \varepsilon_{it}$, for many processes $i = 1 \dots n$
- Agent may simplify:

$$\rho_i^s = m\rho_i + (1-m)\,\rho^d.$$

- so $\mathbb{E}_t^s\left[y_{i,t+k}\right] = \left(\rho_i^s\right)^k y_{i,t}$, i.e.

$$\mathbb{E}_t^s\left[y_{i,t+k}\right] = \left(\frac{\rho_i^s}{\rho_i}\right)^k \mathbb{E}_t\left[y_{i,t+k}\right]$$

- So there's overreaction to non-persistent variables ($\rho_i < \rho^d$, so $\rho_i < \rho_i^s$), underreaction to very persistence variables
- Very non-persistent variables: stock growth, return $\rightarrow$overreaction (e.g. Greenwood Shleifer JFE '14)
- Very persistent variables (inflation) $\rightarrow$underreaction (Mankiw, Reis, Ricardo and Wolfers, MacroAnnual '03)
- Still, it would be nice to have systematic evidence on this.

Biases via inattention

- **Projection bias**
  - At time 0, I need to forecast x_t
  - I anchor on the current value: $x_t^d = x_0$. So there's projection bias (Loewenstein, O'Donoghue, Rabin QJE '03)
  $$x_t^s = m x_t + (1 - m) x_0$$
- **Overconfidence**: My default x^d is high, perhaps because I'm my own "lawyer" and "advocate"
- **Insensitivity to sample size** (KT '74). Replace true sample size by $N^s = \left(N^d\right)^{1-m} N^m$
- **Base-rate neglect** (KT '74). True base probability P is replaced by $P^s(y) = mP(y) + (1-m)P^d(y)$, where $P^d(y)$ is a uniform distribution
- **Cursedness** (Eyster Rabin ECTA '05): $\chi = 1 - m$

Biases via inattention

- **Neglected risks**
- True probability is p, but perceived probability is $p^s = mp$
- **Left-digit bias**
- Replace $x = a + \frac{b}{10}$, by

Measuring attention

- 5 main ways to measure attention:

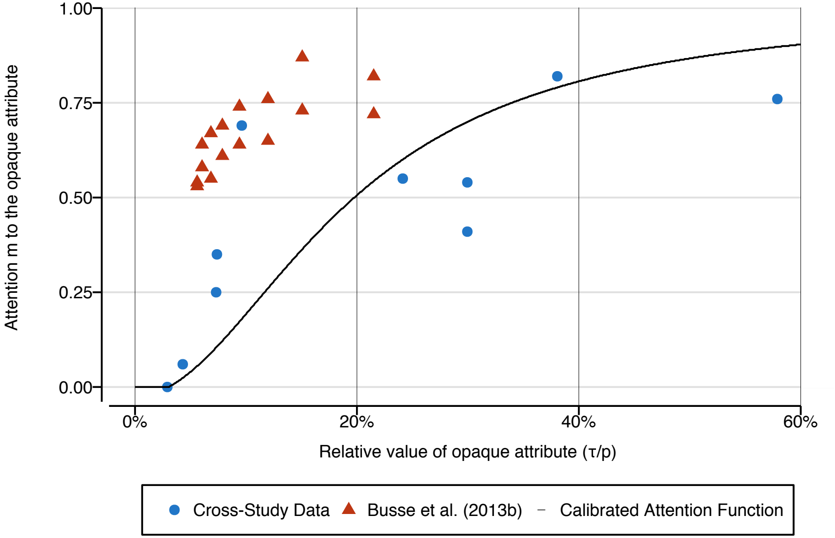
1. Deviations from an optimal action. (Requires knowing the correct action)
2. Deviations from normative cross-partials, e.g. from Slutsky symmetry. (does *not* require knowing the correct action) (We'll see that later)
3. Physical measurement, e.g. eye-tracking.
4. Surveys: eliciting people's beliefs.
   4.1 Simplest way to see if they don't know something
   4.2 But even if I know the interest rate, it doesn't mean that I'll take it into account
5. Qualitative measures: impact of reminders, of advice.

Measuring attention: deviation from optimal action

- Example: if demand is $D(p+\tau) = D^r(p+m\tau)$, then $D_\tau = mD_\tau^r$ and
$$m = \frac{D_\tau}{D_\tau^r}$$
- More generally, $a^r(x) := \text{argmax}_a\, u(a,x)$, $a^s(x) = a^r(mx)$, so $m = \frac{a_x^s}{a_x^r}$.
- How to know the "rational action" if people are confused?

1. This could be done in a "clear and understood" context, e.g. where all prices are very clear, perhaps with just a simple task (so that in this environment, $m = 1$), which allows us to measure a_x^r. Chetty et al AER '09, 15 Allcott and Taubinsky AER 'Taubinsky and Rees-Jones RES '17
2. Sometimes, see what experts do. E.g. do pharmacists buy generics or (more expensive) branded drugs. Bronnenberg, Dubé, et al. QJE '15

Studies put together



The Sparse BR Algorithm: Motivation

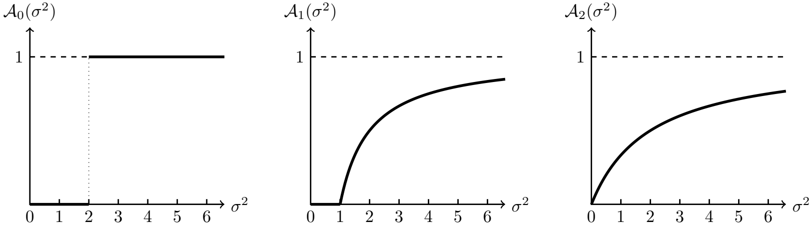
- Agent will take action $a(x,m) = \arg\max_a u(a,x,m)$ and experience utility
$$v(x,m) = u(a(x,m), x)$$
- So it is sensible to allocate attention m as:
$$\max_m \mathbb{E}[v(x,m)] - \mathcal{C}(m)$$
- **Definition**: smax agent simplifies the problem:
  - replaces utility by linear-quadratic approximation, and removes correlations in x
  - Chooses optimal attention m in that simplified model
  - Chooses optimal action with $a(x,m)$ action with correct utility.
- With $\iota = (1, ..., 1)$ =full attention, $\Lambda_{ij} = -\mathbb{E}\left[a_{m_i} u_{aa} a_{m_j}\right]$
$$\mathbb{E}[v(x,m)] = v(\iota) - \frac{1}{2}(m-\iota)' \Lambda (m-\iota) + o\left(\mathbb{E}\|x\|^2\right)$$
- Also, take $\mathcal{C}(m) = \kappa \sum_i m_i^\alpha$. Agent does:
$$\max_{m \in [0,1]^n} -\frac{1}{2}\sum_i (1-m_i)^2 \Lambda_{ii} - \kappa \sum_i m_i^\alpha$$

Sparse max: Quick version

- **Proposition**: *One solves* $\text{smax}_{a;m}\, u(a, x, m)$ *as follows.*
  Step 1: Choose attention:

  $$m_i^* = \mathcal{A}_\alpha\left(\mathbb{E}\left[a'_{m_i} u_{aa} a_{m_i}\right] / \kappa\right)$$

  Step 2: Choose action: $a^s = \arg\max_a u(a, x, m^*)$.
- $a_{m_i} = -u_{aa}^{-1} u_{am_i}$, and u_{aa} evaluated at $(a, m) = \left(a^d, m^d\right)$ with $a^d = \arg\max_a u\left(a, m^d, x\right)$
- $\mathcal{A}_\alpha(v) := \arg\min_m \frac{1}{2}(1-m)^2 |v| + m^\alpha$



Application

- **Prop**. The perceived x_i is $x_i^s = m_i x_i$ with attention is:

$$m_i = \mathcal{A}_\alpha\left(\frac{1}{\kappa} \times \left|a_{x_i} u_{aa} a_{x_i}\right| \sigma_{x_i}^2\right)$$
$$= \mathcal{A}_\alpha\left(\frac{2}{\kappa} \times \text{Utility gains from thinking about } x_i\right)$$

"Eliminate each feature of the world (i.e., $i = 1...n$) that would change the action by only a small amount" (i.e., eliminate the x_i such that $|\sigma_i \cdot \partial a / \partial x_i| \leq \sqrt{\kappa / |u_{aa}|}$).

- Quadratic example:

$$a^r = \sum_i b_i x_i : \text{Non-Sparse Action}$$

$$a^s = \sum_i b_i x_i^s, \qquad x_i^s = x_i \mathcal{A}_\alpha\left(b_i^2 \sigma_{x_i}^2 / \kappa\right) : \text{Sparse action}$$

Psychological underpinnings

- Limited attention and sparsity: it's people can't handle many elements in short-term memory ("7 ± 2" units).
- Reliance on default: Madrian and Shea, Carroll et al. '09.
- Anchoring and adjustment:

  "People make estimates by starting from an initial value that is adjusted to yield the final answer [...]. Adjustments are typically insufficient." (Tversky and Kahneman, 1974, p. 1129)

- Purposeful attention, towards seemingly important things. Chetty et al. '09, Caplin, Dean Martin '11, Bordalo, Gennaioli and Shleifer '12. Koszegi and Szeidl '12...

Comparison with Sims

- Compared to Sims (Entropy-based rational inattention):

$$
\begin{aligned}
a^r &= \sum_i b_i x_i \\
a^s &= \sum_i \mathcal{A}_\alpha \left(b_i^2 \sigma_{x_i}^2 / \kappa \right) \cdot b_i x_i \\
E\left[a^{\text{Sims}} \right] &= \lambda \sum_i b_i x_i
\end{aligned}
$$

  with $0 \leq \lambda < 1$: uniform dampening in Sims, feature by feature dampening in sparsity.
- In Sims, $a^{\text{Sims}}(x)$ is a random function
- With fixed costs, $a(x)$ is a discontinuous function
- With smax (with $\alpha \geq 1$), $a(x)$ is deterministic and continuous.

Consumer and equilibrium theory

- From this Marshallian demand, can re-do consumer theory, equilibrium theory

1. Slutsky matrices: non-symmetry, not negative-semi definite
2. Expenditure function, Hicksian demand
3. Roy's identity, Shephard's lemma, WARP
4. Nominal illusion
5. Producer theory
6. Equilibrium theory: Arrow-Debreu, 1st and 2nd welfare theorems
7. Edgeworth boxes: Phillips curve in each Edgeworth box

- Get a theory of what predictions of basic micro are robust to BR (e.g. sign predictions: "if the price goes up, demand goes down"), and what are not (like symmetry of Slutsky matrix; absence of nominal illusion).

A Behavioral New Keynesian Model

Xavier Gabaix

Harvard

March 2020

Introduction

- I write a theory of monetary and fiscal policy with behavioral agents.
- I use the workhorse framework of monetary policy, the New Keynesian model – and write a behavioral version of it.
- Agents are patient, but they're partially myopic to future disturbances.
- The rational model is a particular case.
- Motivation
    - If people aren't fully rational, our models and policies should incorporate that.
    - A number of empirical issues with New Keynesian framework.
    - The economy (e.g. Japan, US) looks stable at the ZLB, even though that contradicts the Taylor principle, so that in principle the economy could jump from one equilibrium to the next (Cochrane '15)?

Preview: Monetary and Fiscal

1. Behavioral version of the work-horse model used for policy.
2. Monetary policy is less powerful (esp. forward guidance).
3. Helicopter drops of money / Fiscal policy is more powerful.
4. Optimal joint fiscal+monetary policy.
5. Taylor principle strongly modified. Equilibrium is determinate (even with rigid monetary policy) at the ZLB.
6. The ZLB is much less costly.
7. Optimal policy
   - 7.1 Do "helicopter drops of money" at the ZLB $\rightarrow$ First Best.
   - 7.2 "Price-level targeting" is not optimal any more.
8. Resolution of neo-Fisherian paradoxes

- Empirical support for main features of model.

Related literature

- Inattention: Sims 03, Gabaix and Laibson 02, 06, Caballero 95, Mankiw Reis 02, Reis 06, Abel, Eberly and Panageas 09, Chetty, Kroft Looney 09, Angeletos La'O 10, Maćkowiak and Wiederholt 10, 16, Veldkamp 11, Matejka and Sims 11, Caplin, Dean and Martin 11, Caplin Dean Leahy 17, Woodford 12, Alvarez Lippi, Paciello 13, Ganong and Noel '17.
- Behavioral macro: Woodford '13, '18, Gabaix '14, '16, Garcia-Schmidt and Woodford '15, Angeletos and Lian '17a.
- Further explorations of behavioral NK: Farhi and Werning '17.
- New NK thinking: Ilut and Schneider '14, Auclert '15, Kaplan, Moll, Violante '17.
- Subjective expectations: Coibion and Gorodnichenko '15, Fuhrer '17, Afrouzi '17.
- Monetary policy and ZLB: many, including Eggertsson and Woodford 03, Evans, Fisher, Gourio, Krane 16, Werning 12.
- Forward Guidance puzzle: Del Negro, Giannoni, Patterson 15, McKay, Nakamura and Steinsson 15, Chung Herbst and Kiley 15, Caballero Farhi 15, Werning 15, Kiley 16, Campbell et al. 16, Angeletos and Lian 17b.

Traditional New Keynesian model

- Closed economy, no capital, Dixit-Stiglitz firms.
- With $\hat{r}_t = i_t - \mathbb{E}_t \pi_{t+1} - r_t^n$,

$$x_t = \mathbb{E}_t \left[x_{t+1} \right] - \sigma \hat{r}_t \text{ (IS curve)},$$
$$\pi_t = \beta E \left[\pi_{t+1} \right] + \kappa x_t \text{ (Phillips curve)}.$$

- $x_t = \ln y_t - \ln y_t^* =$ output gap.
- $\pi_t = \ln P_t - \ln P_{t-1} =$ inflation.
- Derivation in e.g. Woodford '03, Gali '15.

Individual problem

$$\max_{(c_t,N_t)_{t\geq 0}} U = \mathbb{E}\sum_{t=0}^{\infty}\beta^t u\left(c_t, N_t\right), \qquad u\left(c, N\right) = \frac{c^{1-\gamma}}{1-\gamma} - \frac{N^{1+\phi}}{1+\phi},$$

$$\begin{aligned} k_{t+1} &= \left(1+\bar{r}+\hat{r}\left(\boldsymbol{X}_t\right)\right)\left(k_t+\bar{y}+\hat{y}\left(N_t,\boldsymbol{X}_t\right)-c_t\right), \\ \boldsymbol{X}_{t+1} &= \boldsymbol{G}\left(\boldsymbol{X}_t,\epsilon_{t+1}\right), \\ \hat{y}\left(N_t,\boldsymbol{X}_t\right) &= \omega\left(\boldsymbol{X}_t\right)N_t + y^f\left(\boldsymbol{X}_t\right)-\bar{y}, \end{aligned}$$

with $\boldsymbol{X}_t$ = (de-meaned) state vector, and $y^f\left(\boldsymbol{X}_t\right)$ firms' profits.

- Behavioral agent maximizes in a subjective model with "cognitive discounting"

$$\boldsymbol{X}_{t+1} = \bar{m}\boldsymbol{G}\left(\boldsymbol{X}_t,\epsilon_{t+1}\right)$$

  with $\bar{m}\in[0,1]$. Rational case: $\bar{m}=1$.

- So behavioral agent maximizes U with perceived law of motions

Individual problem: Cognitive discounting

- Linearize: $\boldsymbol{X}_{t+1} = \bar{m}\left(\boldsymbol{\Gamma}\boldsymbol{X}_t + \boldsymbol{\epsilon}_{t+1}\right)$,

$$\begin{aligned}
\mathbb{E}_t^{BR}\left[\boldsymbol{X}_{t+1}\right] &= \bar{m}\boldsymbol{\Gamma}\boldsymbol{X}_t, \\
\mathbb{E}_t^{BR}\left[\boldsymbol{X}_{t+k}\right] &= \bar{m}^k\boldsymbol{\Gamma}^k\boldsymbol{X}_t, \\
\mathbb{E}_t^{BR}\left[\boldsymbol{X}_{t+k}\right] &= \bar{m}^k\mathbb{E}_t\left[\boldsymbol{X}_{t+k}\right].
\end{aligned}$$

- As linearizing, $\hat{r}\left(\boldsymbol{X}_t\right) = b^r\boldsymbol{X}_t$ for some coefficient b^r:

$$\mathbb{E}_t^{BR}\left[\hat{r}\left(\boldsymbol{X}_{t+k}\right)\right] = \bar{m}^k\mathbb{E}_t\left[\hat{r}\left(\boldsymbol{X}_{t+k}\right)\right].$$

- This is "cognitive discounting".

Behavioral IS curve: Derivation in Basic Case

- Traditional model: $\mathbb{E}_t \left[\beta R_t \left(\frac{c_{t+1}}{c_t} \right)^{-\gamma} \right] = 1$ gives

$$\hat{c}_t = \mathbb{E}_t \left[\hat{c}_{t+1} \right] - \frac{1}{\gamma R} \hat{r}_t$$

- Behavioral agent: $\mathbb{E}_t^{BR} [\beta R_t \left(\frac{c(\boldsymbol{X}_{t+1}, k_{t+1})}{c(\boldsymbol{X}_t, k_t)} \right)^{-\gamma}] = 1$. Also, agent correctly forecasts $k_{t+1} = 0$. So,

$$\mathbb{E}_t^{BR} \left[\beta R_t \left(\frac{c\left(\boldsymbol{X}_{t+1}\right)}{c\left(\boldsymbol{X}_t\right)} \right)^{-\gamma} \right] = 1.$$

i.e.

$$\begin{aligned} \hat{c}\left(\boldsymbol{X}_t\right) &= \mathbb{E}_t^{BR} \left[\hat{c}\left(\boldsymbol{X}_{t+1}\right) \right] - \frac{1}{\gamma R} \hat{r}_t. \\ &= \bar{m} \mathbb{E}_t \left[\hat{c}_t \left(\boldsymbol{X}_{t+1}\right) \right] - \frac{1}{\gamma R} \hat{r}_t, \end{aligned}$$

i.e., with $M = \bar{m}$,

$$\hat{c}_t = M \mathbb{E}_t \left[\hat{c}_{t+1} \right] - \sigma \hat{r}_t$$

Discounted Euler equation

- With $M := \frac{\bar{m}}{R - r m_Y}$,

$$x_t \quad = \quad M \mathbb{E}_t \left[x_{t+1} \right] - \sigma \hat{r}_t .$$

- Iterate:

$$x_t = -\sigma \mathbb{E}_t \sum_{\tau \geq t} M^{\tau - t} \hat{r}_\tau .$$

- In rational model, $M = 1$, so

$$x_t = -\sigma \mathbb{E}_t \sum_{\tau \geq t} \hat{r}_\tau .$$

- Hence, interest rate in 1000 periods has same impact as interest rate today: odd. "Forward guidance puzzle" (McKay, Nakamura Steinnson '15).

Firms' pricing problem: rational case

- Dixit-Stiglitz firms with Calvo pricing frictions.
- With $q_{i\tau} := \ln \frac{P_{i\tau}}{P_\tau} = p_{i\tau} - p_\tau$, $-\mu_t := w_t - \zeta_t$ social marginal cost, marginal cost is

$$MC_t = (1 - \tau_f)\, e^{w_t - \zeta_t} = (1 - \tau_f)\, e^{-\mu_t}.$$

- Profit is: $v_\tau = \left(\frac{P_{i\tau}}{P_\tau} - MC_\tau \right) \left(\frac{P_{i\tau}}{P_\tau} \right)^{-\varepsilon} c_\tau$,

$$v\left(q_{i\tau}, \mu_\tau, C_\tau\right) := \left(e^{q_{i\tau}} - (1 - \tau_f)\, e^{-\mu_\tau} \right) e^{-\varepsilon q_{i\tau}} c_\tau.$$

- Take firm simulating at time t. Call $\boldsymbol{X}_\tau$ the extended macro state vector $\boldsymbol{X}_\tau = \left(\boldsymbol{X}_\tau^1, \Pi_\tau \right)$ where $\Pi_\tau := \pi_{t+1} + \cdots + \pi_\tau$,

$$v^{\text{rat}}\left(q_{it}, \boldsymbol{X}_\tau\right) := v\left(q_{it} - \Pi\left(\boldsymbol{X}_\tau\right), \mu\left(\boldsymbol{X}_\tau\right), c\left(\boldsymbol{X}_\tau\right)\right).$$

- Traditional Calvo firm:

$$\max_{q_{it}} \mathbb{E}_t \sum_{\tau=t}^{\infty} (\beta\theta)^{\tau-t} \frac{c\left(\boldsymbol{X}_\tau\right)^{-\gamma}}{c\left(\boldsymbol{X}_t\right)^{-\gamma}} v^{\text{rat}}\left(q_{it}, \boldsymbol{X}_\tau\right).$$

Firms' problem: behavioral case

- Perceived profit:

$$v^{BR}(q_{it}, \boldsymbol{X}_\tau) := v\left(q_{it} - m^f_\pi \Pi(\boldsymbol{X}_\tau), m^f_x \mu(\boldsymbol{X}_\tau), c(\boldsymbol{X}_\tau)\right),$$

  with cognitive discounting: $\boldsymbol{X}_{\tau+1} = \bar{m} G^{\boldsymbol{X}}(\boldsymbol{X}_\tau, \varepsilon_{\tau+1})$
- Firms pay limited attention m^f_π to future inflation and attention m^f_x to real markups.
- Firm's problem:

$$\max_{q_{it}} \mathbb{E}^{BR}_t \sum_{\tau=t}^{\infty} (\beta\theta)^{\tau-t} \frac{c(\boldsymbol{X}_\tau)^{-\gamma}}{c(\boldsymbol{X}_t)^{-\gamma}} v^{BR}(q_{it}, \boldsymbol{X}_\tau).$$

- Hence, they choose price p^*_t:

$$p^*_t - p_t = (1-\beta\theta) \sum_{k\geqslant 0} (\beta\theta\bar{m})^k \mathbb{E}_t \left[m^f_\pi (\pi_{t+1} + \dots + \pi_{t+k}) - m^f_x \mu_{t+$$

- I put those firms and agents in general equilibrium, work out resulting dynamics.

Behavioral firms' problem (short version)

- Dixit-Stiglitz firms with Calvo pricing frictions.
- They pay limited attention m^f to future (macro) markup values.
- Hence, with $-\mu_t := w_t - \zeta_t$, the reset price p_t^* is:

$$p_t^* - p_t = (1-\beta\theta) \sum_{k \geqslant 0} (\beta\theta\bar{m})^k \mathbb{E}_t \left[m_\pi^f (\pi_{t+1} + \ldots + \pi_{t+k}) - m_x^f \mu_{t+\ldots} \right.$$

- I put those firms and agents in general equilibrium, work out resulting dynamics.

Behavioral NK model

- **Proposition:** with $\hat{r}_t = i_t - \mathbb{E}_t \pi_{t+1} - r_t^n$,

$$x_t = M \mathbb{E}_t \left[x_{t+1} \right] + b_d d_t - \sigma \hat{r}_t \text{ (IS curve)},$$

$$\pi_t = \beta M^f \mathbb{E}_t \left[\pi_{t+1} \right] + \kappa x_t \text{ (Phillips curve)},$$

with

$$M = \frac{\bar{m}}{R - r m_Y}, \quad \sigma = \frac{m_r \psi}{R \left(R - r m_Y \right)}, \quad b_d = \frac{\left(1 - \bar{m}\right) \phi m_y r R}{\left(\phi + \gamma\right) \left(R - m_Y r\right) \left(R - r\right.}$$

$$M^f = \bar{m} \left(\theta + \frac{1 - \beta\theta}{1 - \beta\theta\bar{m}} m_\pi^f \left(1 - \theta\right) \right), \qquad \kappa = \bar{\kappa} m_x^f.$$

- Rational model: $M = M^f = 1, b^d = 0$. Here: $M, M^f \in \left[0, 1\right], b^d \geq 0$.
- Empirical support for main features of model:
    1. Phillips curve: Gali and Gertler '99: $M^f \simeq 0.8$.
    2. Fuhrer Rudebusch 04, $M \simeq 0.6$. Fwd guidance puzzle lit.: Need $M < 1$.
    3. Ricardian equivalence doesn't fully hold: e.g. tax rebates etc. literature: $b^d > 0$.

Multiplicity of equilibria under the traditional model

- Consider a Taylor rule

$$i_t = \phi_\pi \pi_t + \phi_x x_t + j_t.$$

With $\boldsymbol{z}_t := (x_t, \pi_t)'$, $D = 1 + \sigma\phi_x + \kappa\sigma\phi_\pi$,
$b_t' = -\frac{\sigma}{D}(1, \kappa)(j_t - r_t^n)$,

$$\boldsymbol{z}_t = \boldsymbol{A}\mathbb{E}_t[\boldsymbol{z}_{t+1}] + \boldsymbol{b}_t,$$

$$\boldsymbol{A} = \frac{1}{D}\begin{pmatrix} M & \sigma(1 - \beta^f \phi_\pi) \\ \kappa M & \beta^f(1 + \sigma\phi_x) + \kappa\sigma \end{pmatrix}.$$

- We have equilibrium uniqueness ("Blanchard-Kahn determinacy") iff the eigenvalues of A are less than 1 in modulus.
- Then, we can write: $z_t = \mathbb{E}_t[\sum_{\tau \geq t} \boldsymbol{A}^{\tau - t} \boldsymbol{b}_\tau]$.
- Otherwise, there are other equilibria $z_{t+s} = \Lambda_2^{-s} v_1 \delta_t$ with $\mathbb{E}_{t-1}[\delta_t] = 0$.

Taylor principle reconsidered

- Consider a Taylor rule

$$i_t = \phi_\pi \pi_t + j_t.$$

- Traditional model: determinacy iff $\phi_\pi > 1$.
- **Proposition** We have equilibrium determinacy iff

$$\phi_\pi + \frac{\left(1 - \beta M^f\right)(1 - M)}{\kappa\sigma} > 1.$$

- In particular, if monetary policy is passive (e.g. stuck at ZLB, $\phi_\pi = 0$), uniqueness with strong enough BR:

$$\frac{\left(1 - \beta M^f\right)(1 - M)}{\kappa\sigma} > 1.$$

- Need enough BR (low M) and price stickiness (low κ; cf. Kocherlakota '16).
- Paper works out full Taylor rule $i_t = \phi_\pi \pi_t + \phi_x x_t + j_t$.

In the trad. model, the economy should be much more volatile at the ZLB

- With $z_t' = (x_t, \pi_t)$,$b_t' = -\frac{\sigma}{D}(1, \kappa)(i_t - r_t^n)$,

$$z_t = A_t \mathbb{E}_t [z_{t+1}] + b_t.$$

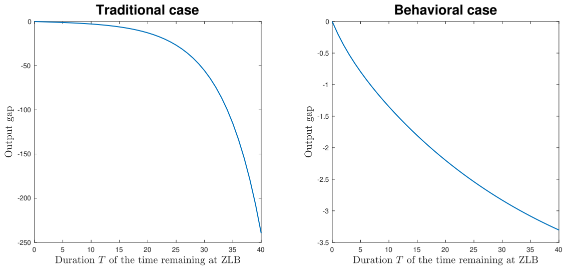
- Suppose we're at the ZLB for $t \leq T$: $A_t = A_{ZLT}$ for $t \leq T$, $A_t = A_{normal}$ for $t > T$. Then,

$$z_0 = \left(\boldsymbol{I} + \boldsymbol{A}_{ZLB} + ... + \boldsymbol{A}_{ZLB}^{T-1} \right) \underline{\boldsymbol{b}} + \boldsymbol{A}_{ZLB}^{T} \mathbb{E}_0 [\boldsymbol{z}_T].$$

- As (in the trad. model) $||A_{ZLB}|| > 1$, so $||A_{ZLB}^T|| \gg 1$ the economy should be extremely sensitive to forecasts about the future, so very erratic.
- In this behavioral model$||A_{ZLB}|| < 1$, so no high volatility.

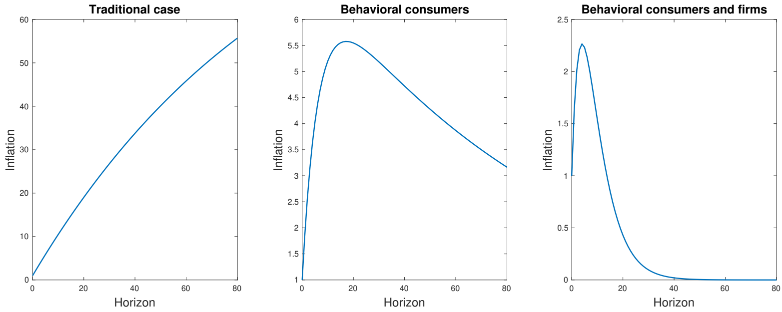
Impact of a ZLB for T periods

- Werning (2012), but with behavioral agents. Take $r_t^n = \underline{r}$ for $t \leq T$, and $r_t^n = \bar{r} < 0$ for $t > T$. For $t > T$, the CB sets $i_t = \bar{r} > 0$, $x_t = \pi_t = 0$.
- **Proposition**. When $\frac{\left(1-\beta M^f\right)(1-M)}{\kappa\sigma} < 1$, the recession is unbounded. But if $\frac{\left(1-\beta M^f\right)(1-M)}{\kappa\sigma} > 1$, then the recession is bounded).



Announcement of future rate cut

- Central bank announces today that it will cut the real rate for one period, in T periods (as McKay-Nakamura-Steinsson (2015)).
- What's the impact today on inflation today?



Optimal Monetary and Fiscal Policy

- Welfare: $\widetilde{W} = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t u(c_t, N_t)$.
- $\widetilde{W} = W^{FB} + W$,

$$W = -K\mathbb{E}_0 \sum_{t=0}^{\infty} \frac{1}{2}\beta^t \left(\pi_t^2 + \vartheta x_t^2\right) + W_-,$$

  with

$$\vartheta = \frac{\bar{\kappa}}{\varepsilon} = \frac{\kappa}{m_x^f \varepsilon},$$

  and $K = u_c c(\gamma + \phi) \frac{\epsilon}{\kappa} m^f$, and W_- is a constant
- Controlling for κ, the relative weight on the output gap (ϑ) is higher when firms are more behavioral (when m^f is lower).
- Intuition: inflation creates less between-firm price dispersion, because firms react less today to future inflation.
- First best: zero output gap and inflation, $x_t = \pi_t = 0$.

First Best in Behavioral NK model

- Recall: with $\hat{r}_t = i_t - \mathbb{E}_t \pi_{t+1} - r_t^n$,

$$x_t = M \mathbb{E}_t [x_{t+1}] + b_d d_t - \sigma \hat{r}_t \text{ (IS curve)},$$
$$\pi_t = \beta M^f \mathbb{E}_t [\pi_{t+1}] + \kappa x_t \text{ (Phillips curve)}.$$

- First best: we want 0 output gap and inflation: $x_t = \pi_t = 0$.
- So, we want:

$$b_d d_t - \sigma (i_t - r_t^n) = 0,$$

i.e. $i_t - \frac{b_d}{\sigma} d_t = r_t^n$.

Optimal policy with supply and demand shocks

First best via "helicopter drops of money"

- You get the first best iff:

$$i_t - \frac{b_d}{\sigma} d_t = r_t^n.$$

- *When the ZLB doesn't bind*. To obtain the first best, set (with Taylor rule around it),

$$i_t = r_t^n \text{ and zero deficit: } d_t = 0,$$

  like in the traditional model.
- *If we hit the ZLB*. Rational agents: Second best, very complex (Eggertson and Woodford, Werning, Gali).
- Behavioral agents: the right "helicopter drops of money" give First Best;

$$i_t = 0 \text{ and deficit: } d_t = \frac{-\sigma}{b_d} r_t^n.$$

- Because agents are not Ricardian, they spend the transfers you give them, hence output goes up.

Optimal policy with complex tradeoffs

- Cost-push shocks: with ν_t AR(1),

$$\pi_t = \beta M^f \mathbb{E}_t \left[\pi_{t+1}\right] + \kappa x_t + \nu_t.$$

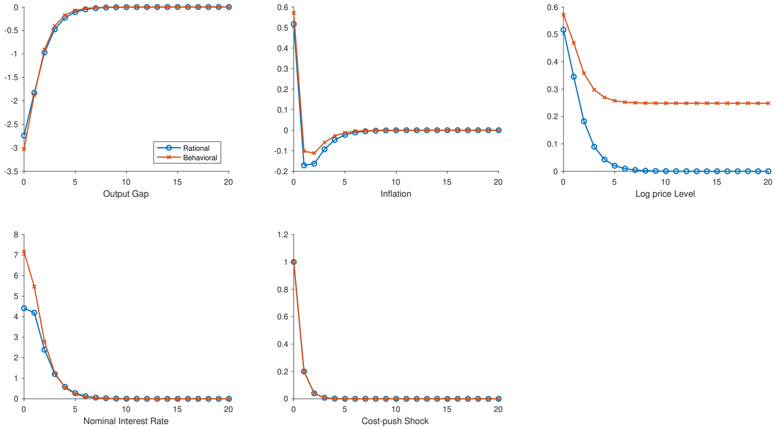
- Then first best cannot be achieved.
- Optimal policy with commitment:
  - Traditional model: Optimal policy gives "price level targeting", "nominal GDP" targeting.
  - With behavioral firms: this is not true any more:

$$\pi_t = \frac{-\vartheta}{\kappa}\left(x_t - M^f x_{t-1}\right),$$

$$p_t = \frac{-\vartheta}{\kappa}\left(x_t + \left(1 - M^f\right)\sum_{\tau=0}^{t-1} x_\tau\right).$$

- Without commitment: the optimal policy under rational vs behavioral economy are close: $\pi_t = \frac{-\vartheta}{\kappa} x_t$.

Optimal policy with commitment



Quantitative Exploration

- Model with partially backward looking firm, useful.
- Why is inflation stable? "Because agents' expectations are anchored at 2% inflation".
- Firms have two noisy signals: noisy rational expectations, and "default inflation".
- π_t^d = default inflation, which comes "for free" to the mind.
- π_t^{CB} = central bank guidance.
- Default inflation:

$$\pi_t^d = (1-\zeta)\,\bar{\pi}_t + \zeta \bar{\pi}_t{}^{CB},$$

  where $\bar{\pi}_t$ and $\bar{\pi}_t{}^{CB}$ are moving averages of past inflation and inflation guidance π_τ^{CB}.
- When passive, firm increases its price by π_t^d (indexation, like CEE, Smets Wouters).
- Call $\hat{\pi}_t := \pi_t - \pi_t^d$.

Extending the model: partially backward-looking firms

- Resulting general equilibrium: with $\pi_t = \hat{\pi}_t + \pi_t^d$,

$$x_t = M\mathbb{E}_t\left[x_{t+1}\right] + b_d d_t - \sigma\left(i_t - \mathbb{E}_t\pi_{t+1} - r_t^n\right) \text{ (IS curve)},$$
$$\hat{\pi}_t = \beta M^f \mathbb{E}_t\left[\hat{\pi}_{t+1}\right] + \kappa x_t \text{ (Phillips curve)}.$$

- So it's exactly the previous model, with "add π_t^d to predicted NK inflation".
- This embeds trad. NK model ($M = M^f = 1$, $\pi^d = 0$), basic behavioral model ($\pi^d = 0$), old Keynesian model ($M = M^f = 0, \zeta = 0$).

Estimation

- WP version of this paper, $M \simeq 0.85$ (with quarterly units)
- Ilabaca, Francisco, Greta Meggiorini, and Fabio Milani, "Bounded Rationality, Monetary Policy, and Macroeconomic Stability," Ec. Letters, 2020.
- Andrade, Cordeiro, and Lambais, "Estimating a Behavioral New Keynesian Model," Working Paper

Other payoffs from this extended model

- The data wants some backward looking inflation in the Phillips curve (Gali and Gertler '99).
- Inflation dynamics are more inert, via "default" inflation π_t^d.
- Notion of "the central bank raises rates to combat the past inflation", like in the old Keynesian model.
- Get costly disinflation (Ball 1994).
- Get Fisher neutrality.
- Condition for equilibrium determinacy:

$$\phi_\pi + \zeta \frac{\left(1 - \beta M^f\right)}{\kappa\sigma} \left(1 - M\right) > 1.$$

- ... something impossible at the ZLB in the NK model ($M = 1$) and Old Keynesian model ($M = \zeta = 0$).

Permanent interest rate shock

"If the Fed raises nominal interest rates, the [NK] model predicts that inflation will smoothly rise, both in the short run and long run. This paper presents a series of failed attempts to escape this prediction. Sticky prices, money, backward-looking Phillips curves, alternative equilibrium selection rules, and active Taylor rules do not [work]" (Cochrane 2015)

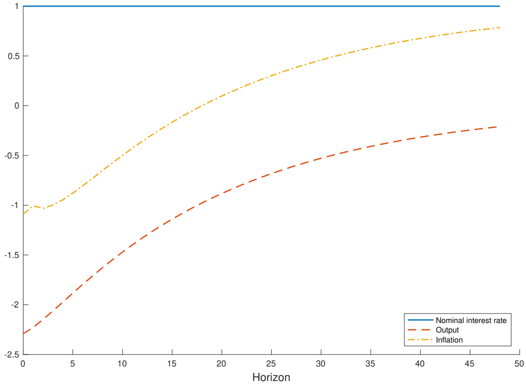
(Note: this depends on equilibrium selection)

Here, bounded rationality will work.

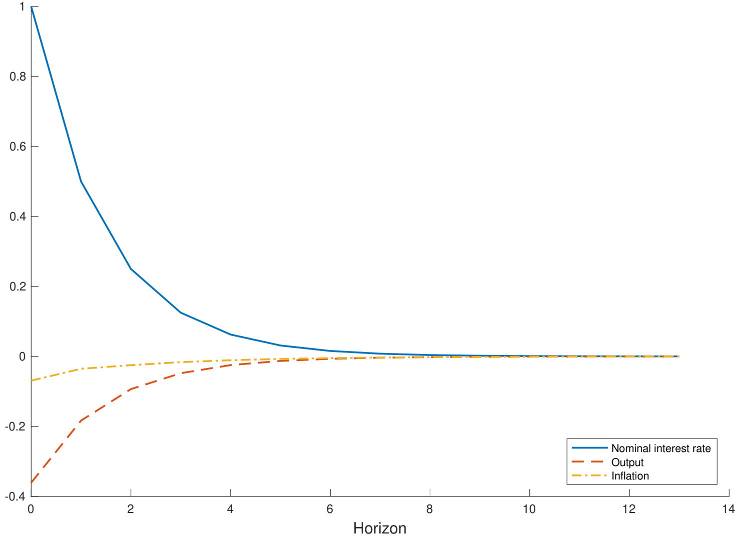
Permanent interest rate shock

Fed raises the nominal rate by 1%, permanently. No Taylor rule, but $\pi_t^{CB} = 1\%$.

Conclusion: the economy is Neo-Fisherian in long run, but Keynesian in run short. Solution to Cochrane's challenge.



Transitory interest rate shock



Extension: Term structure of attention

- For some purposes, we'll add a refinement: perceived law of motion for wealth:

$$k_{t+1} = \left(1 + \bar{r} + \hat{r}^{BR}\left(\boldsymbol{X}_t\right)\right)\left(k_t + \bar{y} + \hat{y}^{BR}\left(N_t, \boldsymbol{X}_t\right) - c_t\right),$$

$$\begin{aligned} \hat{r}^{BR}\left(\boldsymbol{X}_t\right) &= m_r \hat{r}\left(\boldsymbol{X}_t\right), \\ \hat{y}^{BR}\left(N_t, \boldsymbol{X}_t\right) &= m_y \hat{y}\left(\boldsymbol{X}_t\right) + \omega\left(\boldsymbol{X}_t\right)\left(N_t - N\left(\boldsymbol{X}_t\right)\right), \end{aligned}$$

with $m_r, m_y \in [0,1]$. Rational case: $m_r = m_y = 1$.

- So we get a "term structure of attention": intercept m_r, slope $\bar{m}$,

$$\begin{aligned} \mathbb{E}_t^{BR}\left[\hat{r}^{BR}\left(\mathbf{X}_{t+k}\right)\right] &= m_r \bar{m}^k \mathbb{E}_t\left[\hat{r}\left(\mathbf{X}_{t+k}\right)\right], \\ \mathbb{E}_t^{BR}\left[\hat{y}^{BR}\left(\mathbf{X}_{t+k}\right)\right] &= m_y \bar{m}^k \mathbb{E}_t\left[\hat{y}\left(\mathbf{X}_{t+k}\right)\right], \end{aligned}$$

with $\hat{y}^{BR}\left(\boldsymbol{X}_t\right) = m_y \hat{y}\left(\boldsymbol{X}_t\right)$.

- With $b_r(k_t) := \frac{-1}{\gamma R^2}$, $b_y := \frac{\bar{r}}{\bar{R}}$, $b_k := \frac{\bar{r}}{R}\frac{\phi}{\phi+\gamma}$, $m_Y := \frac{\phi m_y + \gamma}{\phi+\gamma}$
- **Proposition**: In this behavioral model (up to 2nd order terms)

$$c_t = \bar{y} + b_k k_t + \mathbb{E}_t\left[\sum_{\tau \geq t} \frac{\bar{m}^{\tau-t}}{R^{\tau-t}}\left(b_r m_r \hat{r}\left(\boldsymbol{X}_\tau\right) + b_y m_Y \hat{y}\left(\boldsymbol{X}_\tau\right)\right)\right].$$

- Then, I put these agents in general equilibrium, with $\hat{y}\left(\boldsymbol{X}_\tau\right) = \hat{c}\left(\boldsymbol{X}_\tau\right)$.
- Define $x_t = \ln y_t - \ln y_t^*$ the output gap.
- Define d_t=deficit after payment of interest rate on debt. I.e., active "transfers" by government to agents.

Behavioral IS curve: derivation (no fiscal policy yet)

- With $\tilde{b}_r = -\frac{\psi m_r}{R^2}$, consumption:

$$\hat{c}_t = \mathbb{E}_t\left[\sum_{\tau \geq t} \frac{\bar{m}^{\tau-t}}{R^{\tau-t}}\left(\frac{r}{R} m_Y \hat{y}_\tau + b_r m_r \hat{r}_\tau\right)\right].$$

- As $\hat{y}_\tau = \hat{c}_\tau$ in equilibrium, using output gap: $x_t = \hat{c}_t / c_t^*$ (here, just take case with $r_t^n = 0$ for simplicity)

$$x_t = \mathbb{E}_t\left[\sum_{\tau \geq t} \frac{\bar{m}^{\tau-t}}{R^{\tau-t}}\left(\frac{r}{R} m_Y x_\tau + b_r m_r \hat{r}_\tau\right)\right],$$

$$x_t = \frac{r}{R} m_Y x_t + \tilde{b}_r \hat{r}_t + \frac{\bar{m}}{R}\mathbb{E}_t\left[x_{t+1}\right],$$

so

$$x_t = M\mathbb{E}_t\left[x_{t+1}\right] - \sigma \hat{r}_t,$$

with $M := \frac{\bar{m}}{R - r m_Y} \in [0,1]$, $\sigma := \frac{-R\tilde{b}_r}{R - r m_Y} = \frac{m_r}{\gamma R(R - r m_Y)}$.

Partial vs General Equilibrium Effects

$$\hat{c}_0 = \mathbb{E}_0 \sum_{\tau \geq 0} \frac{\bar{m}^\tau}{R^\tau} \left(b_y m_Y \hat{c}_\tau - \frac{1}{\gamma} m_r \hat{r}_\tau \right)$$

- Rational agent

$$\hat{c}_0^{\text{direct}} = -\alpha \frac{\hat{r}_\tau}{R^\tau}, \;\; \hat{c}_0^{\text{GE}} = -\alpha \hat{r}_\tau,$$

$$\frac{\hat{c}_0^{\text{GE}}}{\hat{c}_0^{\text{direct}}} = R^\tau.$$

- Behavioral agent (cf Angeletos Lian 17)

$$\hat{c}_0^{\text{direct}} = -\alpha m_r \bar{m}^\tau \frac{\hat{r}_\tau}{R^\tau}, \hat{c}_0^{\text{GE}} = -\alpha m_r M^\tau \hat{r}_\tau,$$

with $M = \frac{\bar{m}}{R - r m_Y}$.

$$\frac{\hat{c}_0^{\text{GE}}}{\hat{c}_0^{\text{direct}}} = \left(\frac{R}{R - r m_Y} \right)^\tau \in \left[1, R^\tau\right].$$

Comparison with other models

$$x_t = M\mathbb{E}_t\left[x_{t+1}\right] + b_d d_t - \sigma \hat{r}_t \text{ (IS curve)}$$
$$\pi_t = \beta M^f \mathbb{E}_t\left[\pi_{t+1}\right] + \kappa x_t \text{ (Phillips curve)}$$

- Hand-to-mouth: keeps $M = M^f = 1$, though give (something like) $b_d > 0$.
- Sticky information (Mankiw-Reis): keeps $M = 1$, $b_d = 0$.
- Rational discounted Euler equation: Del Negro, Giannoni, Patterson ('15), McKay, Nakamura and Steinnsson ('15), Piergallini ('06), Nistico ('12), Caballero and Farhi ('15), Werning ('15): Keeps $M^f = 1$, and silent about b_d. In calibrations gives $M \simeq 1-$(small liquidity spread) or $M \simeq 1-$(small probability of death).
- Misperception of GE (Angeletos and Lian '17) without credit constraints: gives $b^d = 0$.
- Heterogeneity (McKay, Nakamura Steinsson '16, Farhi Werning '17): keeps $M^f = 1$; you lose rep. agent framework

Potential future work

- Theory: capital accumulation, unemployment, distortionary taxes, heterogeneity in wealth (Farhi Werning '16) and in rationality, noisy signal foundations (Angeletos Lian '17).
- Micro data: Estimate the $m_r, m_Y, m^f, \bar{m}$ for consumers and firms

$$c_t = \bar{y} + \frac{\bar{r}}{\bar{R}} k_t + \mathbb{E}_t \left[\sum_{\tau \geq t} \frac{\bar{m}^{\tau-t}}{R^{\tau-t}} \left(b_r\left(k_t\right) m_r \hat{r}_\tau + b_y m_Y \hat{y}_\tau \right) \right].$$

- Macro data: Estimate M, M^f, m^f.
- Estimate a Smets-Wouters model with bounded rationality.
- Link with expectations data (Coibion and Gorodnichenko '15).
- Learning for surveys
  - Most surveys: "where do you expect inflation to me in 2 years".
  - We need surveys eliciting the agents' subjective model (their "forecasting rule").
  - e.g. "Suppose that the central bank raises the interest rate now [or in a year, etc.], what do you think will happen in the economy? how will you change your consumption today"?

Conclusion: Monetary and Fiscal

1. Behavioral version of the work-horse model used for policy.
2. Monetary policy is less powerful (esp. forward guidance).
3. Helicopter drops of money / Fiscal policy is more powerful.
4. Optimal joint fiscal+monetary policy.
5. Taylor principle strongly modified. Equilibrium is determinate (even with rigid monetary policy) at the ZLB.
6. The ZLB is much less costly.
7. Optimal policy
   7.1 Do "helicopter drops of money" at the ZLB $\rightarrow$ First Best.
   7.2 "Price-level targeting" is not optimal any more.
8. Resolution of neo-Fisherian paradoxes

- Empirical support for main features of model.