Lecture 2: Macroeconomics with Mistakes

Joel P. Flynn
Yale

May 22, 2024
NBER Behavioral Macroeconomics Bootcamp

What Are The Macroeconomic Implications of Mistakes?

We have a bunch of tools now for modelling mistakes.

Goal: understand how mistakes matter for macro

Theme: Combining theory and data to answer macro questions

Focus on two main implications:

1. Monetary non-neutrality

2. Business cycle non-linearities implications

Outline

Monetary Non-Neutrality

Business Cycle Non-linearities

Takeaways

Inattention and Monetary Non-Neutrality

- Since Lucas (1972), well understood that imperfect information could lead to monetary non-neutrality

$$\frac{M}{P}$$

- The idea: if firms don't know that monetary shocks have happened, how could their prices perfectly adjust?

- But how informed should firms choose to actually be?

Seminal Contributions

1. Sims (1998), Stickiness
2. Mackowiak and Wiederholt (2009), Optimal Sticky Prices Under Rational Inattention (AER)
3. Stevens (2019), Coarse Pricing Policies (ReStud)
4. Gabaix (2020), A Behavioral New Keynesian Model (AER)

Mackowiak and Wiederholt (2009) – Micro vs. Macro

- Firms can acquire information about micro conditions and macro conditions
- Formally, they can acquire uncorrelated Gaussian signals about micro conditions and micro conditions at mutual information cost
- Main (quantitative) result: Firms should acquire very precise micro info and imprecise macro info

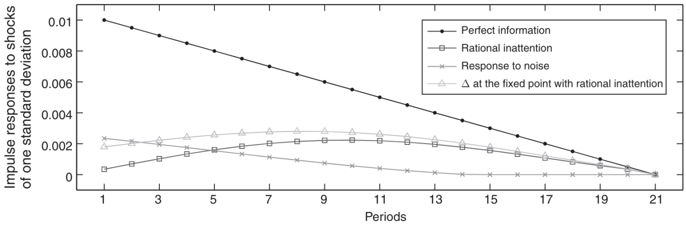


Figure 2. Impulse Responses of an Individual Price to an Innovation in Nominal Aggregate Demand, Benchmark Economy

Stevens (2019) – Coarse Pricing (I)

- Micro evidence: firms choose from a coarse set of prices and lumpily switch between

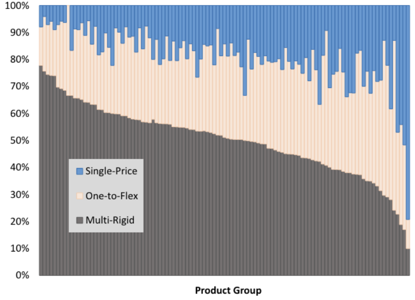


Figure 2: Classification of series by type of pricing policy, across product groups

Note: Nielsen Retail Scanner Data. Percent of series of each type in each product group.

- Estimates a model to match these micro-moments via SMM and shows coarseness matters for monetary non-neutrality

Promising Current Direction: Combining Theory and Data

I'll talk a bit about a recent paper (with Hassan Afrouzi and Choongryul Yang): "What Can Measured Beliefs Tell Us About Monetary Non-Neutrality?"

- Firms have optimal price $q_{i,t}$, which evolves according to a Brownian motion with instantaneous volatility σ
- Loss function given by:

$$\mathcal{L} = -\frac{B}{2}(p_{i,t} - q_{i,t})^2$$

- Pricing friction time-dependent with hazard rate $\theta(h)$
- Can acquire information about q at flow cost given by $\omega\,\mathrm{d}\mathbb{I}$, where $\mathrm{d}\mathbb{I}$ is the instantaneous change in mutual information

$$\sup_{\{\mu_{i,t}^{\mathcal{WS}},\hat{p}_{i,t}\}_{t\geq 0}} \mathbb{E}\left[\int_0^\infty e^{-rt}\left(-\frac{B}{2}\left(p_{i,t} - q_{i,t}\right)^2 \mathrm{d}t - \omega\,\mathrm{d}\mathbb{I}_t\right)\middle| S_i^0\right] \tag{1}$$

Optimal Dynamic Information Policy

Posterior uncertainty about its optimal reset price at time t, $U_{i,t} = \mathbb{V}[q_{i,t}|S_i^t]$

Theorem (Optimal Dynamic Information Policy)

The firm only acquires information when it changes its price. When the firm changes its price, there exists a threshold level of uncertainty U^* such that:

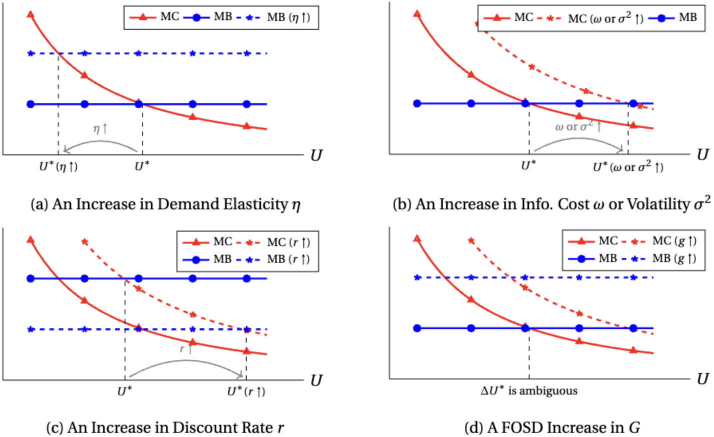
1. *If $U_{i,t-} \leq U^*$, then the firm acquires no information and $U_{i,t} = U_{i,t-}$.*
2. *If $U_{i,t-} > U^*$, then the firm acquires a Gaussian signal of its optimal price such that its posterior uncertainty is $U_{i,t} = U^*$.*

Moreover, U^* is the unique solution to:

$$\underbrace{\frac{\omega}{U^*} - \mathbb{E}^h\left[e^{-rh}\frac{\omega}{U^* + \sigma^2 h}\right]}_{\text{marginal cost of information}} = \underbrace{B\left(\frac{1 - \mathbb{E}^h[e^{-rh}]}{r}\right)}_{\text{marginal benefit of information}} \tag{2}$$

How The Economic Environment Determines Optimal Uncertainty

Figure 1: Comparative Statics of Optimal Reset Uncertainty in Model Parameters



(a) An Increase in Demand Elasticity η

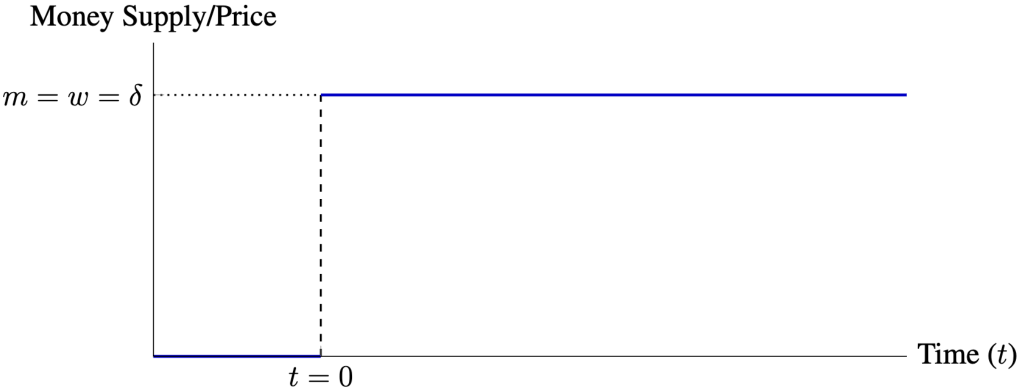
(b) An Increase in Info. Cost ω or Volatility σ^2

(c) An Increase in Discount Rate r

(d) A FOSD Increase in G

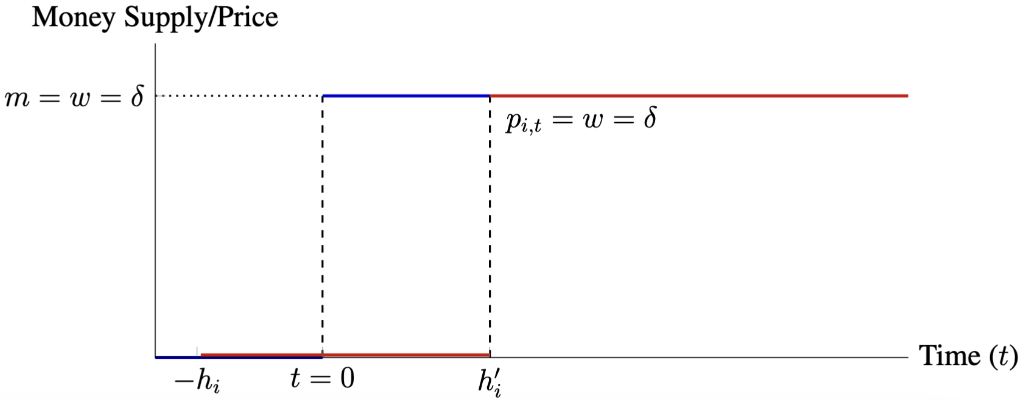
A Graphical Illustration of Monetary Non-Neutrality with Full Information

- Money supply increases δ percent at $t = 0$.
- Firms' nominal wage increase immediately to δ forever.



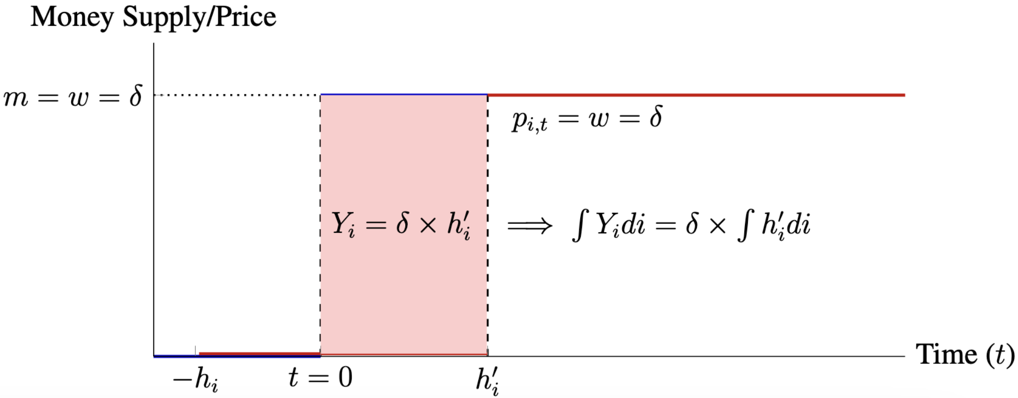
A Graphical Illustration of Monetary Non-Neutrality with Full Information

- Consider a firm i who last changed its price at $-h_i$ and gets to reset at h_i'
- With full information, price jumps at new $w = \delta$ at first opportunity



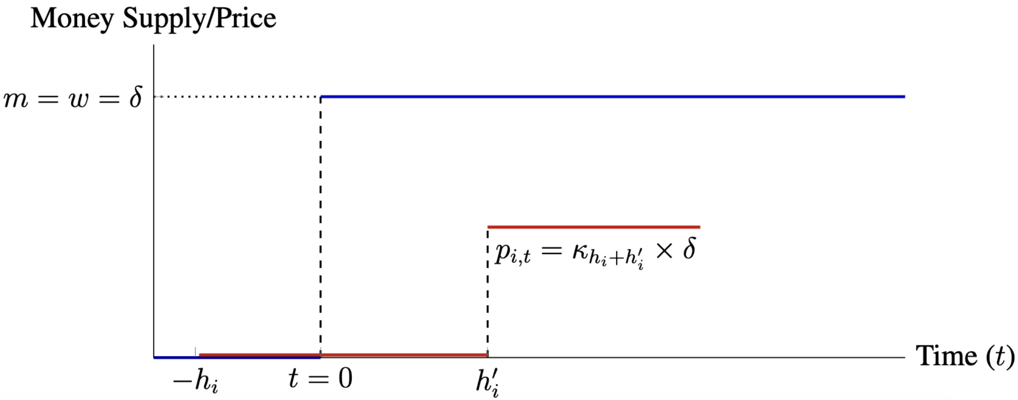
A Graphical Illustration of Monetary Non-Neutrality with Full Information

- Firm i's contribution to output is its duration since shock (h_i') times δ
- Aggregate contribution to output is average duration times δ



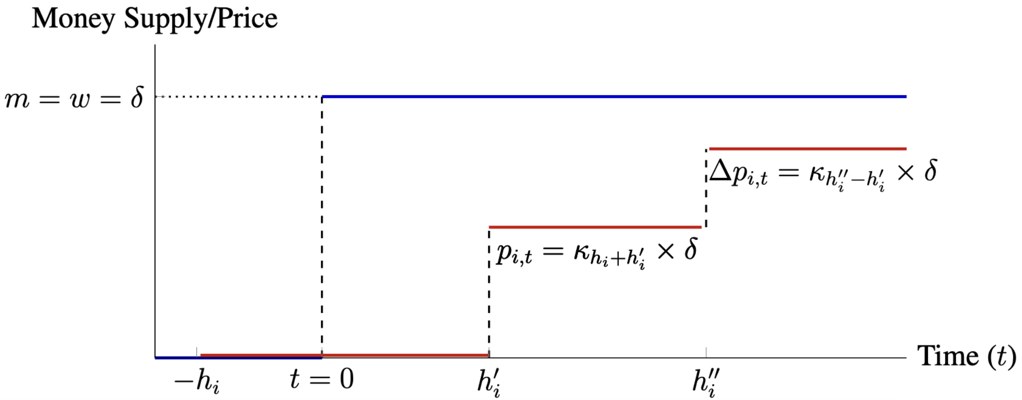
A Graphical Illustration of Monetary Non-Neutrality with Info. Frictions

- Firms' nominal wage increase immediately to δ forever.
- Firm i: price no longer jumps to $w = \delta$ at first price change (info. frictions)



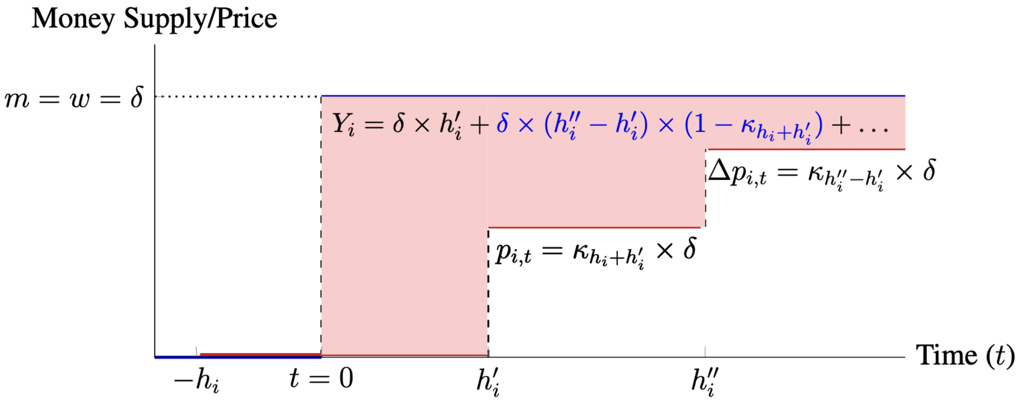
A Graphical Illustration of Monetary Non-Neutrality with Info. Frictions

- Instead, at every new price change, it gets closer to the new $w = \delta$
- At every price change, the size of the jump depends on the spell duration



A Graphical Illustration of Monetary Non-Neutrality with Info. Frictions

- Firm i's average contribution to output is now the sum of all these rectangles
- Aggregate non-neutrality is the sum over all firms



How Firms' Uncertainty Affects Monetary Non-Neutrality

The expected lifetime output gap of a firm who reset their price h periods ago and is y^b wrong about their optimal reset price is given by:

$$\bar{D}_h y^b + \sum_{k=0}^{\infty} \bar{D}_0 (1-\bar{\kappa}_0)^k (1-\bar{\kappa}_h) y^b = \bar{D}_h y^b + \bar{D}_0 y^b \frac{1-\bar{\kappa}_h}{\bar{\kappa}_0} \tag{3}$$

Theorem (Monetary Non-Neutrality)

The cumulative impulse response to an unobserved monetary shock $\mathcal{M}^b$ is:

$$\mathcal{M}^b = \bar{D} + \frac{U^*}{\sigma^2} \tag{4}$$

How Can We Identify The CIR in the Data?

Proposition (Characterization of the Distribution of Uncertainty)

The cross-sectional density of uncertainty about optimal reset prices $I \in \Delta(\mathbb{R}_+)$ is given by:

$$I(z) = \begin{cases} 0, & z < U^*, \\ \frac{1}{\sigma^2} f\left(\frac{z-U^*}{\sigma^2}\right), & z \geq U^*. \end{cases} \tag{5}$$

where $f(\cdot)$ is the density of ongoing spell lengths in the cross-section.

So, if we can measure (i) the empirical uncertainty distribution and (ii) the empirical distribution of spell lengths, we can back out σ^2 and pin down $\mathcal{M}^b$.

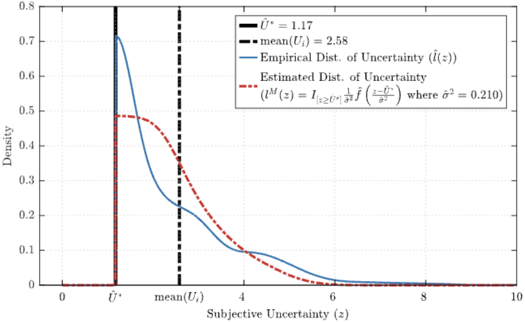
Eliciting the Information We Need In Survey Data

Survey question on distribution of beliefs about own price:
"If your firm was free to change its price (i.e. suppose there was no cost to renegotiating contracts with clients, no costs of reprinting catalogues, etcâ¦) today, what probability would you assign to each of the following categories of possible price changes the firm would make? Please provide a percentage answer."

Survey question on time since last price change:
"When did your firm last change its price (in months) and by how much (in % change)?"

Estimating the Model

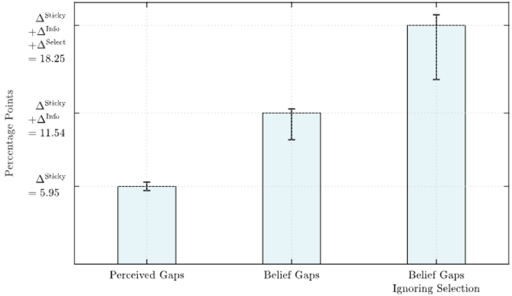
Figure 3: Distributions of Firms' Subjective Uncertainty in the Data and the Model



Notes: This figure shows the distribution of firms' subjective uncertainty about their ideal prices. The black vertical solid line shows the mode of the empirical distribution of subjective uncertainty ($\hat{U}^*$) and the black vertical dashed line shows the mean of the subjective uncertainty observed in the survey data. The blue solid line is the empirical distribution of uncertainty $\hat{l}(z)$. The red dashed line shows the estimated distribution of uncertainty ($l^M(z)$) from Equation (42) using the empirical distribution of time since the last price changes ($\hat{f}$) and the estimated uncertainty of shocks ($\hat{\sigma}^2$).

What Do Measured Beliefs Tell Us?

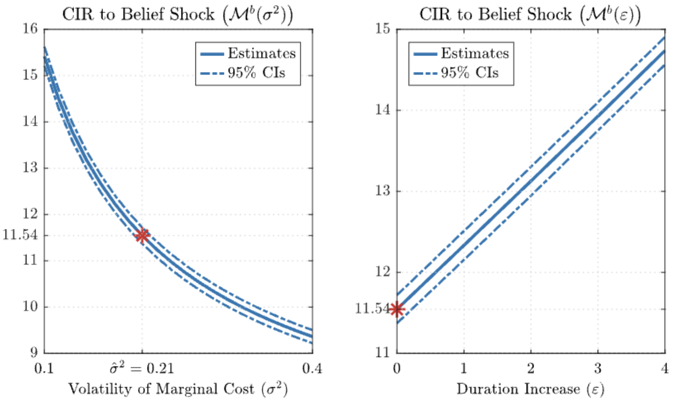
Figure 4: Estimated Monthly Cumulative Impulse Responses to an Initial 1 Percentage Point Output Gap under Different Scenarios



Notes: This figure shows the output effects of a 1 percentage point shock to perceived gaps (left bar), to belief gaps (middle bar), and belief gaps ignoring the selection effect (right bar). The output effect of a 1pp perceived gap is the average duration of firms' pricing spells $\Delta^{\text{Sticky}} = \bar{D}$, the effect of a 1pp belief gap is the effect of a perceived gap plus $\Delta^{\text{Info}} = \frac{U^*}{\sigma^2}$, and the effect of 1pp belief gap without selection effect is $\Delta^{\text{Sticky}} + \Delta^{\text{Info}}$ plus $\Delta^{\text{Select}} = \frac{\bar{U} - U^*}{\sigma^2}$. We present 95% confidence intervals as black vertical lines.

How Do Price Stickiness And Volatility Matter?

Figure 5: Microeconomic Volatility, Price Stickiness, and Monetary Non-Neutrality



Notes: This figure shows two counterfactual analyses on how micro uncertainty and price stickiness affect monetary non-neutrality. The left panel shows the effect of microeconomic uncertainty on monetary non-neutrality induced by information friction. The right panel shows the effect of price stickiness on monetary non-neutrality. Red stars show the estimates with the estimated $\hat{\sigma}^2 = 0.21$ and $\varepsilon = 0$. We present 95% confidence intervals as blue dashed lines.

Why Use Informational Models?

- We followed in the Lucas tradition of thinking about information
- But is that really essential?
- We care about firms' prices, not necessarily the beliefs that underlie those prices (while this can be informative)
- See Costain and Nakov (2019), "Logit Price Dynamics" for an analysis of monetary non-neutrality with logit stochastic choice

One Direction For Future Research

- Quite a lot of theoretical work on information frictions (reviewed today)
- Quite a lot of empirical work on expectations and surveys (reviewed by Chris and Karthik)
- Work that combines survey data and theories to speak to classic macro questions would be incredibly valuable
- Useful to do the theory and design surveys to measure exactly what is needed

Outline

Monetary Non-Neutrality

Business Cycle Non-linearities

Takeaways

The Macroeconomics of Managing "Mistakes"

- Firms, like the rest of us, optimize imperfectly
  see, e.g., Simon (1947, 1957) on attention constraints and "bounded rationality"

The Macroeconomics of Managing "Mistakes"

- Firms, like the rest of us, optimize imperfectly
  see, e.g., Simon (1947, 1957) on attention constraints and "bounded rationality"
- "Bounds of rationality" reflect choices and responses to economic conditions. The macroeconomy consists of many "mistake makers" responding to one another

The Macroeconomics of Managing "Mistakes"

- Firms, like the rest of us, optimize imperfectly
  see, e.g., Simon (1947, 1957) on attention constraints and "bounded rationality"
- "Bounds of rationality" reflect choices and responses to economic conditions. The macroeconomy consists of many "mistake makers" responding to one another

This paper ("Attention Cycles"): models a two-way interaction

Business Cycles	$\longleftrightarrow$	Attention Cycles
aggregate decisions		cognition, mistakes

Households, Final Goods, and Labor Supply

- Countably infinite time periods, indexed by $t \in \mathbb{N}$
- **Representative household** consumes C_t of final good and works L_t hours, with payoffs

$$\mathcal{U}\left((C_{t+j}, L_{t+j})_{j=0}^{\infty}\right) = \mathbb{E}_t\left[\sum_{j=0}^{\infty} \beta^j \left(\frac{C_{t+j}^{1-\gamma}}{1-\gamma} - v(L_{t+j})\right)\right]$$

for $\beta \in (0,1)$, $\gamma > 0$, and $v(\cdot)$ increasing + convex

Households, Final Goods, and Labor Supply

- Countably infinite time periods, indexed by $t \in \mathbb{N}$
- Representative household consumes C_t of final good and works L_t hours, with payoffs

$$\mathcal{U}\left((C_{t+j}, L_{t+j})_{j=0}^{\infty}\right) = \mathbb{E}_t\left[\sum_{j=0}^{\infty} \beta^j \left(\frac{C_{t+j}^{1-\gamma}}{1-\gamma} - v(L_{t+j})\right)\right]$$

  for $\beta \in (0,1)$, $\gamma > 0$, and $v(\cdot)$ increasing + convex
- **Final good** produced with CES ($\epsilon > 1$) technology, from intermediates $(x_{it})_{i\in[0,1]}$:

$$X_t = \left(\int_0^1 x_{it}^{1-\frac{1}{\epsilon}}\,\mathrm{d}i\right)^{\frac{\epsilon}{\epsilon-1}}$$

Households, Final Goods, and Labor Supply

- Countably infinite time periods, indexed by $t \in \mathbb{N}$
- Representative household consumes C_t of final good and works L_t hours, with payoffs

$$\mathcal{U}\left((C_{t+j}, L_{t+j})_{j=0}^{\infty}\right) = \mathbb{E}_t\left[\sum_{j=0}^{\infty} \beta^j \left(\frac{C_{t+j}^{1-\gamma}}{1-\gamma} - v(L_{t+j})\right)\right]$$

  for $\beta \in (0,1)$, $\gamma > 0$, and $v(\cdot)$ increasing + convex
- Final good produced with CES ($\epsilon > 1$) technology, from intermediates $(x_{it})_{i\in[0,1]}$:

$$X_t = \left(\int_0^1 x_{it}^{1-\frac{1}{\epsilon}} \, \mathrm{d}i\right)^{\frac{\epsilon}{\epsilon-1}}$$

- **Wage rule**, parameterized with slope $\chi > 0$ and constants $\bar{w}, \bar{X} > 0$:

$$w_t = \bar{w}\left(\frac{X_t}{\bar{X}}\right)^{\chi}$$

Intermediate Goods: Technology and Payoffs

Production function:

$$x_{it} = \theta_{it} \cdot L_{it}$$

- Productivity θ_{it}, with cross-sectional distribution G_t
- Single (labor) input + CRS, easily generalized to multiple flexible inputs + CRS

Firm's "flow payoff," risk-adjusted profits:

$$\Pi(x_{it}; \theta_{it}, w_t, X_t) = M(X_t) \cdot \pi\left(x_{it}; \theta_{it}, X_t, w_t\right)$$

Intermediate Goods: Technology and Payoffs

Production function:

$$x_{it} = \theta_{it} \cdot L_{it}$$

- Productivity θ_{it}, with cross-sectional distribution G_t
- Single (labor) input + CRS, easily generalized to multiple flexible inputs + CRS

Firm's "flow payoff," risk-adjusted profits:

$$\Pi(x_{it}; \theta_{it}, w_t, X_t) = M(X_t) \cdot \pi\left(x_{it}; \theta_{it}, X_t, w_t\right)$$

Intermediate Goods: Technology and Payoffs

Production function:

$$x_{it} = \theta_{it} \cdot L_{it}$$

- Productivity θ_{it}, with cross-sectional distribution G_t
- Single (labor) input + CRS, easily generalized to multiple flexible inputs + CRS

Firm's "flow payoff," risk-adjusted profits:

$$\Pi(x_{it}; \theta_{it}, w_t, X_t) = M(X_t) \cdot \pi\left(x_{it}; \theta_{it}, X_t, w_t\right)$$

Costly Control for Firms: Set-up

Premise: difficult for firms to digest "state" (macro and micro) and translate it into decisions

Model:

- Let state at t be $z_{it} := (\theta_{it}, X_t, w_t) \in \mathcal{Z}$
- Firm observes $z_{i,t-1}$ and conjectures transition density $f(z_{it} \mid z_{i,t-1})$
- Chooses conditional production distributions $p_t = (p(x \mid z_{it}))_{z_{it} \in \mathcal{Z}}$ to solve

$$\max_{p} \ \mathbb{E}_{f,p}\left[\Pi(x; z_{it})\right] - C_i(p)$$

We specialize to **entropy costs**, where $\lambda_i \sim H, \in \mathbb{R}_+$, is firm-level "inattentiveness" shifter:

$$C_i(p) = -\lambda_i \cdot \mathbb{E}_f\left[\text{Entropy}(p(x \mid z_i))\right]$$

Equilibrium

Aggregate productivity state θ_t

$$G_t = G(\theta_t), \qquad \theta' \geq \theta \implies G(\theta') \succsim_{\text{FOSD}} G(\theta)$$

and linear-quadratic approximation of profits, aggregator

Definition (Equilibrium)

Given a sequence of productivity shocks $(\theta_t)_{t=0}^{\infty}$, an equilibrium is a sequence for choices $((p_i^*(\theta_{t-1}))_{i\in[0,1]})_{t=1}^{\infty}$, output $(X(\theta_t))_{t=0}^{\infty}$, and wages $(w(\theta_t))_{t=0}^{\infty}$ such that

1. *Intermediate goods firms optimize given a correct conjecture for X.*
2. *Final output is consistent with the aggregator, and wages with the wage rule.*

Production Misoptimizations in Partial Equilibrium

Proposition (Production of Intermediate Goods Firms)

Each firm's production is described by the random variable

$$x_i = x^*(\theta_i, X, w) + \sqrt{\frac{\lambda_i}{|\pi_{xx}(\theta_i, X, w)| \cdot M(X)}} \cdot v_i, \qquad v_i \sim N(0,1), \ \textit{iid across } i$$

where x^* is the unconstrained optimal action, π_{xx} is the curvature of the dollar profit function, and M is the stochastic discount factor.

Production Misoptimizations in Partial Equilibrium

Proposition (Production of Intermediate Goods Firms)

Each firm's production is described by the random variable

$$x_i = x^*(\theta_i, X, w) + \boxed{\sqrt{\frac{\lambda_i}{|\pi_{xx}(\theta_i, X, w)| \cdot M(X)}} \cdot v_i}, \qquad v_i \sim N(0,1), \ \textit{iid across } i$$

where x^* is the unconstrained optimal action, π_{xx} is the curvature of the dollar profit function, and M is the stochastic discount factor.

Firms **make misoptimizations**

Production Misoptimizations in Partial Equilibrium

Proposition (Production of Intermediate Goods Firms)

Each firm's production is described by the random variable

$$x_i = x^*(\theta_i, X, w) + \sqrt{\frac{\lambda_i}{|\pi_{xx}(\theta_i, X, w)| \cdot M(X)}} \cdot v_i, \qquad v_i \sim N(0,1), \text{ iid across } i$$

where x^* is the unconstrained optimal action, π_{xx} is the curvature of the dollar profit function, and M is the stochastic discount factor.

Firms make misoptimizations, but rein them in based on incentives in

- Profit curvature: dollar cost of producing wrong level
- Stochastic discount factor: translation to utility cost

Production Misoptimizations in Partial Equilibrium

Proposition (Production of Intermediate Goods Firms)

Each firm's production is described by the random variable

$$x_i = x^*(\theta_i, X, w) + \sqrt{\frac{\lambda_i}{|\pi_{xx}(\theta_i, X, w)| \cdot M(X)}} \cdot v_i, \qquad v_i \sim N(0,1), \text{ iid across } i$$

where x^* is the unconstrained optimal action, π_{xx} is the curvature of the dollar profit function, and M is the stochastic discount factor.

Firms make misoptimizations, but rein them in based on incentives in

- Profit curvature: dollar cost of producing wrong level
- Stochastic discount factor: translation to utility cost

When Are Misoptimizations Highest? The Key Forces

Define *extent of misoptimization* $m(\lambda_i, \theta_i, X) := \mathbb{E}[(x_i - x_i^*)^2 \mid \theta_i, X]$

When Are Misoptimizations Highest? The Key Forces

Define *extent of misoptimization* $m(\lambda_i, \theta_i, X) := \mathbb{E}[(x_i - x_i^*)^2 \mid \theta_i, X]$

Corollary

Consider a type λ_i firm. Their extent of misoptimization

1. *Decreases in $|\pi_{xx}|$ (profit curvature), holding fixed M Profit sensitivity channel*
2. *Decreases in M (marginal utility), holding fixed $|\pi_{xx}|$ Risk-pricing channel*

When Are Misoptimizations Highest? The Key Forces

Define *extent of misoptimization* $m(\lambda_i, \theta_i, X) := \mathbb{E}[(x_i - x_i^*)^2 \mid \theta_i, X]$

Corollary

Consider a type λ_i firm. Their extent of misoptimization

1. *Decreases in $|\pi_{xx}|$ (profit curvature), holding fixed M Profit sensitivity channel*
2. *Decreases in M (marginal utility), holding fixed $|\pi_{xx}|$ Risk-pricing channel*

Corollary

Consider a type λ_i firm. Their extent of misoptimization

1. *Increases in productivity θ_i*
2. *Increases in output X if $\gamma > \chi(\epsilon + 1) - 1$ and decreases otherwise.*

Attention Cycles in Equilibrium

Assumption (Assumption ★)

$\gamma > \chi + 1$ and $\chi\epsilon < 1$ where γ is the coefficient of relative risk aversion, χ is the elasticity of real wages to real output, and ϵ is the elasticity of substitution between goods

Attention Cycles in Equilibrium

Assumption (Assumption ★)

$\gamma > \chi + 1$ and $\chi\epsilon < 1$ where γ is the coefficient of relative risk aversion, χ is the elasticity of real wages to real output, and ϵ is the elasticity of substitution between goods

Proposition (Proposition: Existence, Uniqueness, and Monotonicity)

For any $\chi > 0$, an equilibrium exists. Under ★, there is a unique such equilibrium with positive output X. Moreover, output is strictly increasing in productivity θ.

Attention Cycles in Equilibrium

Assumption (Assumption ★)

$\gamma > \chi + 1$ and $\chi\epsilon < 1$ where γ is the coefficient of relative risk aversion, χ is the elasticity of real wages to real output, and ϵ is the elasticity of substitution between goods

Proposition (Proposition: Existence, Uniqueness, and Monotonicity)

For any $\chi > 0$, an equilibrium exists. Under ★, there is a unique such equilibrium with positive output X. Moreover, output is strictly increasing in productivity θ.

Proposition (Proposition: Misoptimization Cycles)

Assume ★, or $\gamma > \chi + 1$ and $\chi\epsilon < 1$. In the unique linear-quadratic equilibrium, average misoptimization $m(\theta) := \mathbb{E}[(x_i - x_i^*)^2 \mid \theta]$ is lower when output $X(\theta)$ is lower.

An "Attention Wedge" Shapes Dynamics

Define sufficient statistics $\theta := \left(\mathbb{E}_G[\theta_i^{\epsilon-1}]\right)^{\frac{1}{\epsilon-1}}$ and $\lambda := \mathbb{E}_H[\lambda_i]$

Proposition (Consequences of Attention Cycles)

Output can be written in the following way:

$$\log X(\log\theta) = X_0 + \chi^{-1}\log\theta + \log W(\log\theta)$$

where $\log W(\log\theta) \leq 0$*, with equality iff* $\lambda = 0$*. Under* $\bigstar$*, the wedge satisfies:*

1. $\partial \log W / \partial\lambda < 0$ *Widens with larger cognitive costs*
2. $\partial \log W / \partial \log\theta < 0$ *for* $\lambda > 0$ *Is largest in productive, low-attention state*

Measuring Misoptimizations: Data

- **Dataset**: Compustat Annual Fundamentals, 1986-2017
  - *Strengths*: annual frequency, multi-sector coverage
  - *Acknowledged weaknesses*: only public firms
- Standard sample restrictions (e.g,. no financial or utility firms) Sample Restrictions
- Key variables: sales, total employees, total variable costs, value of capital stock

Measuring Misoptimizations: From Theory to Data

In the Theory

$$\log L_{it} = \log x^*_{it} - \log \theta_{it} + \log\left(1 + \frac{\sigma_{it}}{x^*_{it}} v_{it}\right), \quad v_{it} \sim N(0,1) \qquad \text{Proposition: Optimal Choices}$$

$$= \underbrace{\beta \log \theta_{it}}_{\text{Estimate}} + \underbrace{\tau \log X_t + \xi \log w_t}_{\substack{\text{Span by FE} \\ \text{(firm + sector-time)}}} + \underbrace{\log\left(1 + \frac{\sigma_{it}}{x^*_{it}} v_{it}\right)}_{\text{Treat as Residual}} \qquad \text{Log-linear } x^*$$

In The Data

$$\log L_{it} = \beta \log \hat{\theta}_{it} + \gamma_i + \chi_{j(i),t} + m_{it}$$

$$m_{it} = \rho m_{i,t-1} + \sqrt{1-\rho^2}\, u_{it} \qquad \mathbb{E}[u_{it}] = 0, \quad \mathbb{V}[u_{it}] = \tilde{\sigma}^2_{it} \approx \frac{\sigma^2_{it}}{(x^*_{it})^2}$$

Measuring Attention to the Macroeconomy: Methodology

Dataset: full text of all US-based public firms' 10-K and 10-Q

- Accounting summaries plus **discussions of risks and outlook**
- 1995 to 2018; 480,000 documents, or 5,000 per quarter

Measuring Attention to the Macroeconomy: Methodology

Dataset: full text of all US-based public firms' 10-K and 10-Q

- Accounting summaries plus discussions of risks and outlook
- 1995 to 2018; 480,000 documents, or 5,000 per quarter

1. Score words by their relative prominence in a macro reference R vs. 10K/Q

$$\text{tf-idf}(w; R) := \text{Frequency of } w \text{ in } R \times \log\left(\frac{1}{\text{Frequency of } w \text{ in 10K/Q}}\right)$$

Method: Calculating Macro Attention

References used: *Macroeconomics* by Mankiw, *Principles of Macroeconomics* by Mankiw, and *Macroeconomics: Principles and Policy* by Baumol and Blinder

Measuring Attention to the Macroeconomy: Methodology

Dataset: full text of all US-based public firms' 10-K and 10-Q

- Accounting summaries plus discussions of risks and outlook
- 1995 to 2018; 480,000 documents, or 5,000 per quarter

1. Score words by their relative prominence in a macro reference R vs. 10K/Q

$$\text{tf-idf}(w; R) := \text{Frequency of } w \text{ in } R \times \log\left(\frac{1}{\text{Frequency of } w \text{ in 10K/Q}}\right)$$

2. Generate "macro words" $=$ *intersection* of top 200 tf-idf for each reference

Method: Calculating Macro Attention

References used: *Macroeconomics* by Mankiw, *Principles of Macroeconomics* by Mankiw, and *Macroeconomics: Principles and Policy* by Baumol and Blinder

Measuring Attention to the Macroeconomy: Methodology

Dataset: full text of all US-based public firms' 10-K and 10-Q

- Accounting summaries plus discussions of risks and outlook
- 1995 to 2018; 480,000 documents, or 5,000 per quarter

1. Score words by their relative prominence in a macro reference R vs. 10K/Q

$$\text{tf-idf}(w; R) := \text{Frequency of } w \text{ in } R \times \log\left(\frac{1}{\text{Frequency of } w \text{ in 10K/Q}}\right)$$

2. Generate "macro words" = *intersection* of top 200 tf-idf for each reference
3. Define macro attention for firm i at time t as total IDF-weighted frequency of macro words, and time-series aggregate by averaging across firms

Method: Calculating Macro Attention

References used: *Macroeconomics* by Mankiw, *Principles of Macroeconomics* by Mankiw, and *Macroeconomics: Principles and Policy* by Baumol and Blinder

Fact 1: Misoptimizations Hurt Profitability and Returns

Are misoptimizations "bad" for firms, in both directions? (*not* mechanical from measurement)

Fact 1: Misoptimizations Hurt Profitability and Returns

Are misoptimizations "bad" for firms, in both directions? (*not* mechanical from measurement)

Binned scatter plots of

$$X_{it} = f(\hat{u}_{it}) + \chi_{j(i),t} + \epsilon_{it}$$

where X_{it} is stock return or firm profitability, $\chi_{j(i),t}$ are sector-by-time FE

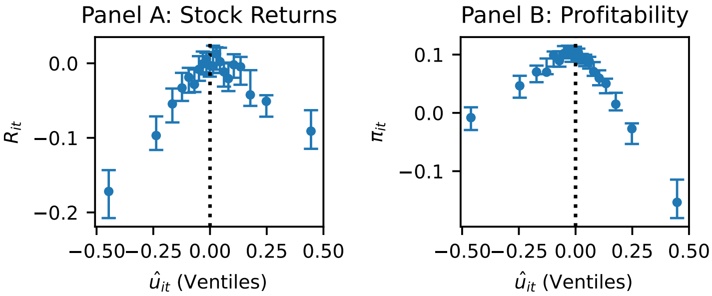
Fact 1: Misoptimizations Hurt Profitability and Returns

Are misoptimizations "bad" for firms, in both directions? (*not* mechanical from measurement)

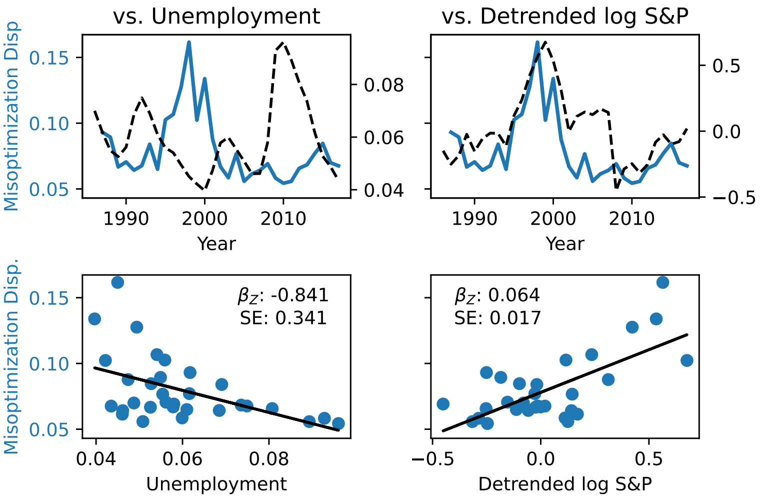
Binned scatter plots of

$$X_{it} = f(\hat{u}_{it}) + \chi_{j(i),t} + \epsilon_{it}$$

where X_{it} is stock return or firm profitability, $\chi_{j(i),t}$ are sector-by-time FE



Fact 2: Misoptimization Dispersion is Pro-Cyclical



Notes: SE are HAC-robust with two-year bandwidth.

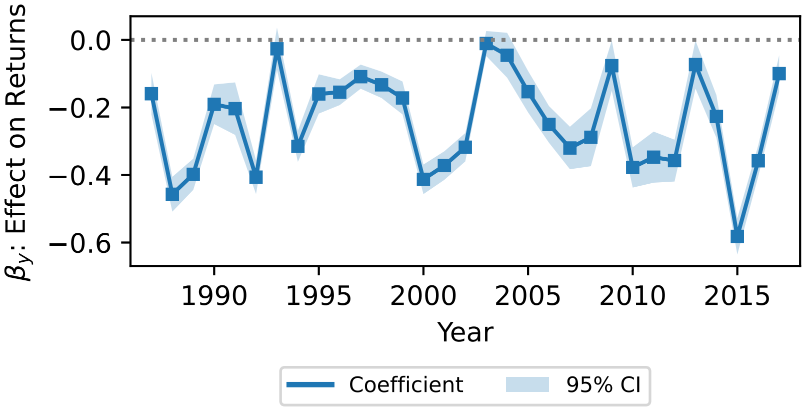
Fact 3: Misoptimizations Hurt Returns More in Bad Aggregate States

$$\Delta \log P_{it} = \sum_{y} \beta_y \cdot \hat{u}_{it}^2 \cdot \mathbb{I}[t = y] + \chi_{j(i),t} + \epsilon_{it}$$

- $\Delta \log P_{it}$: year-on-year stock return
- Industry-by-year fixed effects sweep out background trends
- Hypothesis from model: $|\beta_y|$ large in downturns, or economy experiences duress

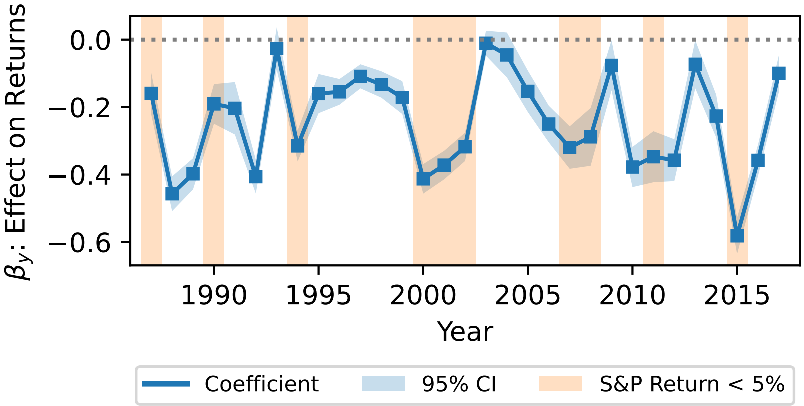
Fact 3: Misoptimizations Hurt Returns More in Bad Aggregate States

$$\Delta \log P_{it} = \sum_y \beta_y \cdot \hat{u}_{it}^2 \cdot \mathbb{I}[t = y] + \chi_{j(i),t} + \epsilon_{it}$$

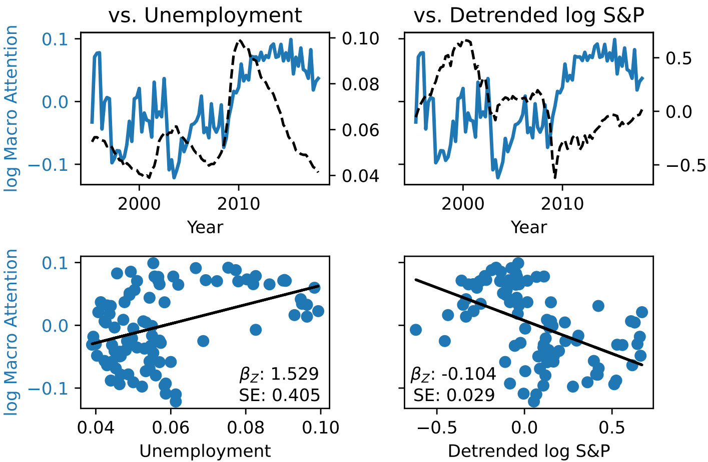


Fact 3: Misoptimizations Hurt Returns More in Bad Aggregate States

$$\Delta \log P_{it} = \sum_{y} \beta_y \cdot \hat{u}_{it}^2 \cdot \mathbb{I}[t = y] + \chi_{j(i),t} + \epsilon_{it}$$



Fact 4: Macro Attention in Language is Counter-Cyclical



Notes: standard errors are HAC-robust with two-year bandwith.

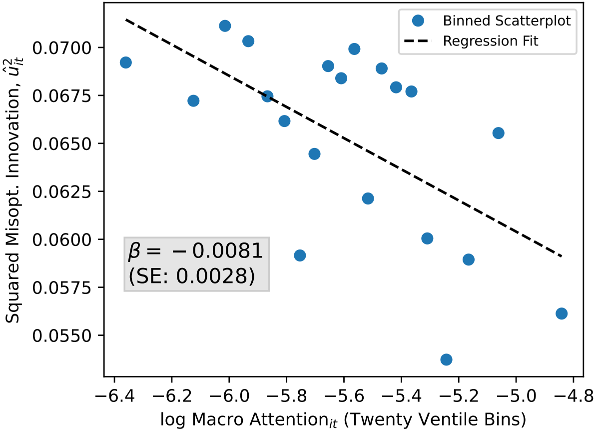
Fact 5: Macro-attentive Firms Make Smaller Misoptimizations

$$\hat{u}_{it}^2 = \beta \cdot \log \text{MacroAttention}_{it} + \chi_{j(i),t} + \Gamma' X_{it} + \epsilon_{it}$$

- $\log \text{MacroAttention}_{it}$: firm level Macro Attention in language
- Hypothesis: $\beta < 0$ implies that macro-attentive firms make more precise decisions, *sweeping out aggregate and industry-specific trends and cycles*

Fact 5: Macro-attentive Firms Make Smaller Misoptimizations

$$\hat{u}_{it}^2 = \beta \cdot \log \text{MacroAttention}_{it} + \chi_{j(i),t} + \Gamma' X_{it} + \epsilon_{it}$$



Notes: standard errors are double-clustered by firm and year.

Calibration of Model

Productivity sufficient statistic $\theta = \left(\mathbb{E}_G[\theta_i^{\epsilon-1}]\right)^{\frac{1}{\epsilon-1}}$ is Gaussian AR(1) in logs:

$$\log \theta_t = \rho \log \theta_{t-1} + \sigma u_t, u_t \sim N(0,1)$$

Calibration of Model

Productivity sufficient statistic $\theta = \left(\mathbb{E}_G[\theta_i^{\epsilon-1}]\right)^{\frac{1}{\epsilon-1}}$ is Gaussian AR(1) in logs:

$$\log\theta_t = \rho\log\theta_{t-1} + \sigma u_t, u_t \sim N(0,1)$$

	Parameter	Value	Strategy
χ	Wage Rule Slope	0.097	Direct (OLS) calculation, 1987-2018
ϵ	Elas of Substitution	4	1.33x average markup
ρ	Persistence of $\log\theta$	0.95	Standard

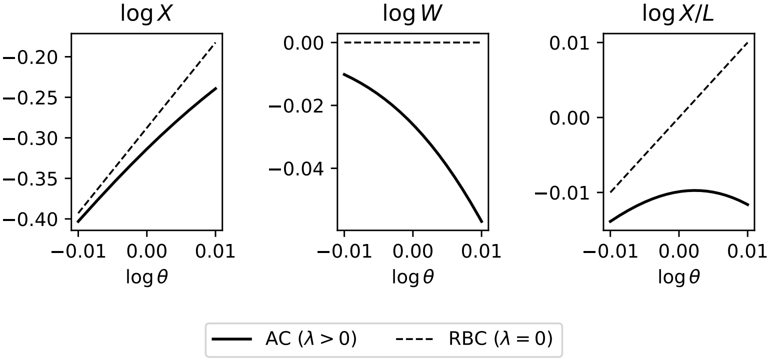
Calibration of Model

Productivity sufficient statistic $\theta = \left(\mathbb{E}_G[\theta_i^{\epsilon-1}]\right)^{\frac{1}{\epsilon-1}}$ is Gaussian AR(1) in logs:

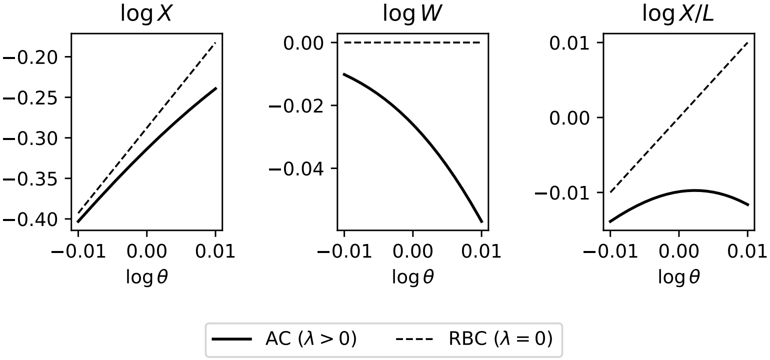
$$\log\theta_t = \rho\log\theta_{t-1} + \sigma u_t,\ u_t \sim N(0,1)$$

	Parameter	Value	Strategy
χ	Wage Rule Slope	0.097	Direct (OLS) calculation, 1987-2018
ϵ	Elas of Substitution	4	1.33x average markup
ρ	Persistence of $\log\theta$	0.95	Standard
γ	CRRA	11.5	Match: *Average level of Misopt. Disp.*
λ	Avg. Attention Cost	0.406	Match: *Slope of Misopt. Disp on* $-\frac{Unemp_t}{100}$
σ^2	Var. of $\log\theta$ Shock	4.8×10^{-7}	Match: *Variance of quarterly RGDP growth*

Output and the Attention Wedge in the Calibrated Model

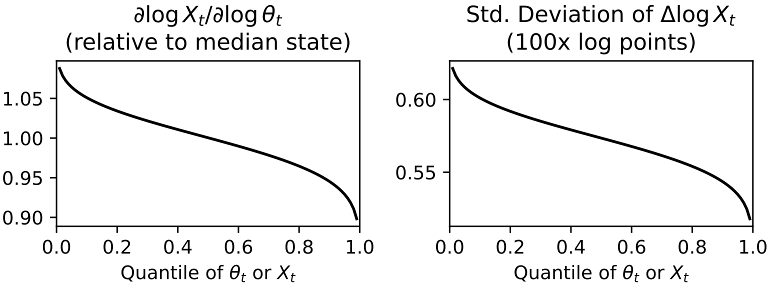


Output and the Attention Wedge in the Calibrated Model



- Median output cost of inattention $= 2.6\%$; productivity cost $= \chi \cdot \epsilon \cdot 2.6\% = 1.0\%$
- Non-monotone labor productivity
- Concave attention wedge $\rightarrow$ more shock response in low states

Results: Shock Responses and Stochastic Volatility



Signing the predictions from the theory,

- **Predictions 1 and 2**: More output effects of negative shocks, and of any shocks when productivity and output are low
- **Prediction 3**: Higher conditional volatility of output when productivity, output are low

Outline

Monetary Non-Neutrality

Business Cycle Non-linearities

Takeaways

Takeaways

- There is no cookie-cutter approach to studying macroeconomics with bounded rationality
- Bounded rationality is hard to measure, but theory helps
- Work that seriously combines theory and data will be immensely valuable in making behavioral macro impossible to ignore!