Attention Cycles

Joel P. Flynn and Karthik A. Sastry

Massachusetts Institute of Technology

August 25, 2022
European Summer Meeting of the Econometric Society

The Macroeconomics of "Mistakes"

- Firms, like the rest of us, optimize imperfectly

  see, e.g., Simon (1947, 1957) on attention constraints and "bounded rationality"

The Macroeconomics of “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 Macroeconomics of "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

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

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

This Paper (I): Motivation, Model, and Theoretical Results

Motivating Evidence (not today): textual analysis of US public firms' regulatory filings (Forms 10Q/K) suggests that firms speak more about macroeconomic topics in downturns

This Paper (I): Motivation, Model, and Theoretical Results

Motivating Evidence (not today): textual analysis of US public firms' regulatory filings (Forms 10Q/K) suggests that firms speak more about macroeconomic topics in downturns

Theoretical Framework:

- Neoclassical core
- *Cognitive friction*: costly control ("inattention") for firms

This Paper (I): Motivation, Model, and Theoretical Results

Motivating Evidence (not today): textual analysis of US public firms' regulatory filings (Forms 10Q/K) suggests that firms speak more about macroeconomic topics in downturns

Theoretical Framework:

- Neoclassical core
- *Cognitive friction*: costly control ("inattention") for firms

Two sets of **theoretical results**:

1. Characterization of how cognitive effort and misoptimization respond to business cycle
    - Sources of incentives: *profit curvature* ("dollar costs") and *risk pricing* ("utility costs")

This Paper (I): Motivation, Model, and Theoretical Results

Motivating Evidence (not today): textual analysis of US public firms' regulatory filings (Forms 10Q/K) suggests that firms speak more about macroeconomic topics in downturns

Theoretical Framework:

- Neoclassical core
- *Cognitive friction*: costly control ("inattention") for firms

Two sets of **theoretical results**:

1. Characterization of how cognitive effort and misoptimization respond to business cycle
    - Sources of incentives: *profit curvature* ("dollar costs") and *risk pricing* ("utility costs")
    - Prediction under standard calibration: low misoptimization in downturns

This Paper (I): Motivation, Model, and Theoretical Results

Motivating Evidence (not today): textual analysis of US public firms' regulatory filings (Forms 10Q/K) suggests that firms speak more about macroeconomic topics in downturns

Theoretical Framework:

- Neoclassical core
- *Cognitive friction*: costly control ("inattention") for firms

Two sets of **theoretical results**:

1. Characterization of how cognitive effort and misoptimization respond to business cycle
   - Sources of incentives: *profit curvature* ("dollar costs") and *risk pricing* ("utility costs")
   - Prediction under standard calibration: low misoptimization in downturns
2. Mapping from cyclical attention to macroeconomic dynamics, via *misallocation* channel

This Paper (II): Empirical Analysis and Quantification

Empirical Findings: using data on US public firms (1986-2017), we document that

This Paper (II): Empirical Analysis and Quantification

Empirical Findings: using data on US public firms (1986-2017), we document that

1. Average "Misoptimization Dispersion" in input choice, calculated consistently with the model, falls in downturns and rises in booms ("macro test")

This Paper (II): Empirical Analysis and Quantification

Empirical Findings: using data on US public firms (1986-2017), we document that

1. Average "Misoptimization Dispersion" in input choice, calculated consistently with the model, falls in downturns and rises in booms ("macro test")
2. The market punishes firm-level misoptimizations more in bad states ("micro test")

This Paper (II): Empirical Analysis and Quantification

Empirical Findings: using data on US public firms (1986-2017), we document that

1. Average "Misoptimization Dispersion" in input choice, calculated consistently with the model, falls in downturns and rises in booms ("macro test")
2. The market punishes firm-level misoptimizations more in bad states ("micro test")
3. Macro-attentive firms make smaller misoptimizations (connection to motivating evidence)

This Paper (II): Empirical Analysis and Quantification

Empirical Findings: using data on US public firms (1986-2017), we document that

1. Average "Misoptimization Dispersion" in input choice, calculated consistently with the model, falls in downturns and rises in booms ("macro test")
2. The market punishes firm-level misoptimizations more in bad states ("micro test")
3. Macro-attentive firms make smaller misoptimizations (connection to motivating evidence)

Quantification: model, calibrated to match facts and draw out GE consequences, generates

1. Larger aggregate responses to negative shocks vs. positive shocks
2. Larger aggregate shock responses in low states
3. Counter-cyclical volatility in output growth

   About 20% of what data ask for (Jurado, Ludvigson, and Ng, 2015)

Outline

Overview of Model

Households

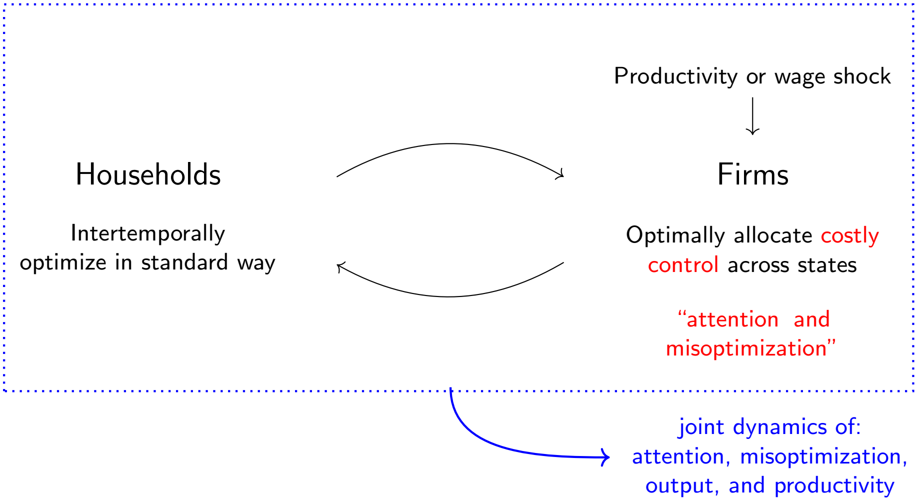
Intertemporally optimize in standard way

Firms

Optimally allocate costly control across states

"attention and misoptimization"

Overview of Model



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

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

  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 $\bar{w}, \bar{X} > 0$:

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

  Realistic and useful for analytical results (see also Blanchard and Galí, 2010)

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 Big Model

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 Big Model

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

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

Profit $x \cdot (P - MC)$
("dollars")

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 Big Model

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

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

Marginal utility of investor ("utils per dollar") → $M(X_t)$

Profit $x \cdot (P - MC)$ ("dollars") → $\pi\left(x_{it}; \theta_{it}, X_t, w_t\right)$

Costly Control for Firms: Set-up

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

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 $f(z_{it} \mid z_{i,t-1})$

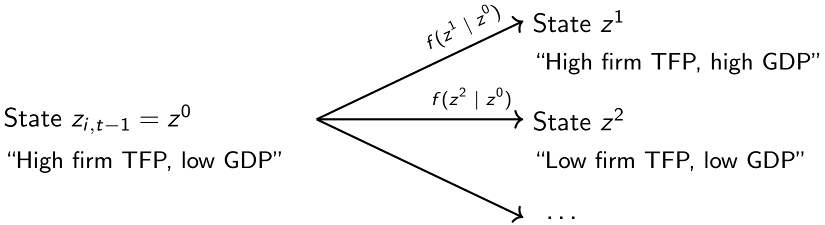
Costly Control for Firms: Set-up

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 $f(z_{it} \mid z_{i,t-1})$

Illustration: decision at time t



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 $p_t = (p(x \mid z_{it}))_{z_{it} \in \mathcal{Z}}$

Illustration: decision at time t

State $z_{i,t-1} = z^0$

Choose $(p(x \mid z_{it}))_{z_{it} \in \mathcal{Z}}$

State z^1

Plan: $p(x \mid z^1)$

State z^2

Plan: $p(x \mid z^2)$

...

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

$$\max_{p} \; \mathbb{E}_{f,p}\left[\Pi(x; z_{it})\right] - \left(-\lambda_i \, \mathbb{E}_f\left[\text{Entropy}[p(x \mid z_i)]\right]\right)$$

Illustration: decision at time t

State $z_{i,t-1} = z^0$

Choose $(p(x \mid z_{it}))_{z_{it} \in \mathcal{Z}}$

State z^1

Plan: $p(x \mid z^1)$

State z^2

Plan: $p(x \mid z^2)$

...

Friction embedded in cost: *precise planning is difficult*

Equilibrium

Aggregate productivity state θ_t Wage Rule Shock

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

and linear-quadratic approximation of profits, aggregator

Equilibrium

Aggregate productivity state θ_t Wage Rule Shock

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

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.

Extended Definition

Outline

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.

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**

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

Equilibrium Analysis: Existence, Uniqueness, Monotonicity

Assumption ★

1. $\gamma > \chi + 1$
2. $\chi\epsilon < 1$

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

Remark: if $\chi = 0.1$ (realistic wage rigidity), need $\epsilon < 10$ and $\gamma > 1.1$

Equilibrium Analysis: Existence, Uniqueness, Monotonicity

Assumption ★

1. $\gamma > \chi + 1$
2. $\chi\epsilon < 1$

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

Remark: if $\chi = 0.1$ (realistic wage rigidity), need $\epsilon < 10$ and $\gamma > 1.1$

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.

Misoptimization Cycles

Define “average misoptimization” m as average firm’s mean-squared-error in output choice:

$$m(\theta) := \mathbb{E}[m(\lambda_i, \theta_i, X) \mid \theta]$$

Misoptimization Cycles

Define "average misoptimization" m as average firm's mean-squared-error in output choice:

$$m(\theta) := \mathbb{E}[m(\lambda_i, \theta_i, X) \mid \theta]$$

Proposition: Misoptimization Cycles

Assume ★, or $\gamma > \chi + 1$ and $\chi\epsilon < 1$. In the unique linear-quadratic equilibrium, average misoptimization $m(\theta)$ is lower when output $X(\theta)$ is lower.

Misoptimization Cycles

Define "average misoptimization" m as average firm's mean-squared-error in output choice:

$$m(\theta) := \mathbb{E}[m(\lambda_i, \theta_i, X) \mid \theta]$$

Proposition: Misoptimization Cycles

Assume ★, or $\gamma > \chi + 1$ and $\chi\epsilon < 1$. In the unique linear-quadratic equilibrium, average misoptimization $m(\theta)$ is lower when output $X(\theta)$ is lower.

In words: under ★, key mechanism is:

lower aggregate productivity → hungrier investors → more precise firms

The Attention Wedge and Output 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.
2.

The Attention Wedge and Output 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

The Attention Wedge and Output 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 ★, 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

Corollary: labor productivity can be written as

$$\log A(\log\theta) = \log X(\log\theta) - \log L(\log\theta) = \log\theta + \chi\epsilon\log W(\log\theta)$$

Outline

Roadmap for Empirical Analysis

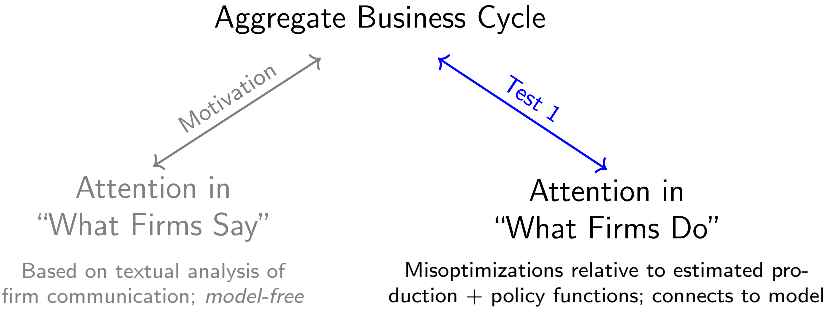
Aggregate Business Cycle

Motivation

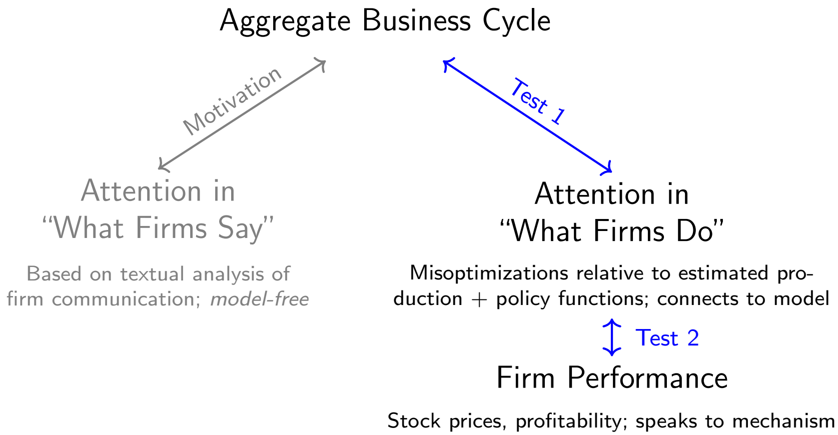
Attention in
"What Firms Say"

Based on textual analysis of
firm communication; *model-free*

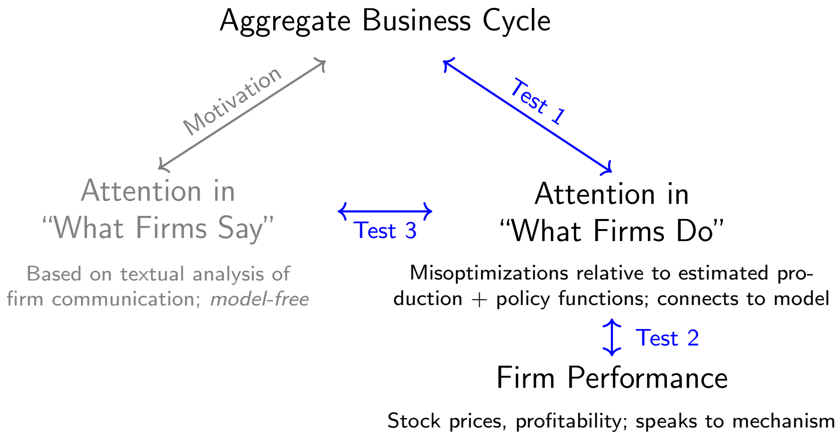
Roadmap for Empirical Analysis



Roadmap for Empirical Analysis



Roadmap for Empirical Analysis



Measuring Misoptimizations: Data

- **Dataset**: Compustat Annual Fundamentals, 1986-2017 Sample Restrictions
  - *Strengths*: annual frequency, multi-sector coverage
  - *Acknowledged weaknesses*: only public firms
- Key variables: sales, total employees, total variable costs, value of capital stock

Measuring Misoptimizations: From Theory to Data

In the Theory

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

Proposition: Optimal Choices

Measuring Misoptimizations: From Theory to Data

In the Theory

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

$$= \beta \log \theta_{it} + \tau \log X_t + \xi \log w_t + \log\left(1 + \frac{\sigma_{it}}{x_{it}^{*}} v_{it}\right) \qquad \text{Log-linear } x^{*}$$

Measuring Misoptimizations: From Theory to Data

In the Theory

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

$$= \underbrace{\beta \log \theta_{it}}_{\text{Estimate}} + \tau \log X_t + \xi \log w_t + \log\left(1 + \frac{\sigma_{it}}{x^*_{it}} v_{it}\right) \qquad \text{Log-linear } x^*$$

Measuring Misoptimizations: From Theory to Data

In the Theory

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

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

Measuring Misoptimizations: From Theory to Data

In the Theory

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

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

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 Misoptimizations: From Theory to Data

More Details

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

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

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

Measuring Misoptimizations: Empirical Implementation

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

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

Method: Calculating Misoptimizations and Misoptimization Dispersion

1. Estimate firm-level TFP as Solow residual, using industry-level revenue shares to estimate revenue elasticities Details

Measuring Misoptimizations: Empirical Implementation

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

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

Method: Calculating Misoptimizations and Misoptimization Dispersion

1. Estimate firm-level TFP as Solow residual, using industry-level revenue shares to estimate revenue elasticities Details
2. Estimate firm policy function with two-stage OLS procedure Details

Measuring Misoptimizations: Empirical Implementation

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

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

Method: Calculating Misoptimizations and Misoptimization Dispersion

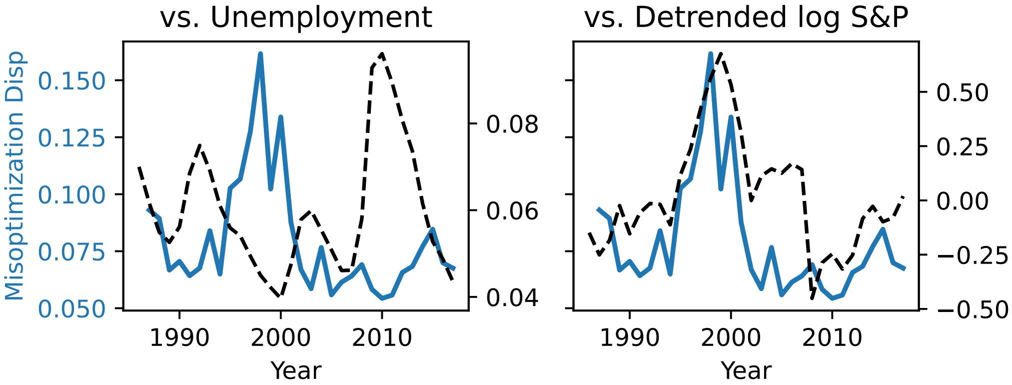
1. Estimate firm-level TFP as Solow residual, using industry-level revenue shares to estimate revenue elasticities Details
2. Estimate firm policy function with two-stage OLS procedure Details
3. Calculate $(\hat{m}_{it}, \hat{u}_{it})$ as residuals, and aggregate dispersion over set of firms $\mathcal{I}_t$ as

$$\text{MisoptimizationDisp}_t = \frac{\sum_{i \in \mathcal{I}_t} s^*_{it} \cdot \hat{u}^2_{it}}{\sum_{i \in \mathcal{I}_t} s^*_{it}}$$

where s^*_{it} are fitted values of predicting sales with TFP. Equation

Validation

Finding 1: Misoptimization Dispersion is Pro-Cyclical



Regression coefficient: -0.84
t-stat: -2.47

Regression coefficient: 0.06
t-stat: 3.76

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

IQR | Unwt. | Composition | Alt. Inputs | Alt. Meas. | Industry

Test: When Do Misoptimizations Hurt Performance More?

Hypothesis: cost of misoptimizations is highest in bad macro states

Test: When Do Misoptimizations Hurt Performance More?

Hypothesis: cost of misoptimizations is highest in bad macro states

Firm-by-year level regression model

$$\Delta P_{it} = \beta \cdot \hat{u}_{it}^2 + \phi \cdot \left(\hat{u}_{it}^2 \times \Delta \log \text{SP500}_t\right) + \chi_{j(i),t} + \Gamma' X_{it} + \epsilon_{it}$$

- **Outcome**: ΔP_{it} = annual stock return ("risk-adjusted profits")
- **Absorbed effects**: sectoral and macro trends
- **Possible additional controls**: TFP growth and Firm FE

Test: When Do Misoptimizations Hurt Performance More?

Hypothesis: cost of misoptimizations is highest in bad macro states ($\phi > 0$)

Firm-by-year level regression model

$$\Delta P_{it} = \beta \cdot \hat{u}_{it}^2 + \phi \cdot \left(\hat{u}_{it}^2 \times \Delta \log \mathrm{SP500}_t\right) + \chi_{j(i),t} + \Gamma' X_{it} + \epsilon_{it}$$

- Outcome: ΔP_{it} = annual stock return ("risk-adjusted profits")
- Absorbed effects: sectoral and macro trends
- Possible additional controls: TFP growth and Firm FE
- **Interpretation**: $\phi > 0$ indicates larger effects when investors are distressed
  - Formal articulation of this in model variant with CAPM structure in Appendix C

Finding 2: Greater Financial Punishment in Bad States

$$\Delta \log P_{it} = \beta \cdot \hat{u}_{it}^2 + \phi \cdot \left(\hat{u}_{it}^2 \times \Delta \log \text{SP500}_t\right) + \chi_{j(i),t} + \Gamma' X_{it} + \epsilon_{it}$$

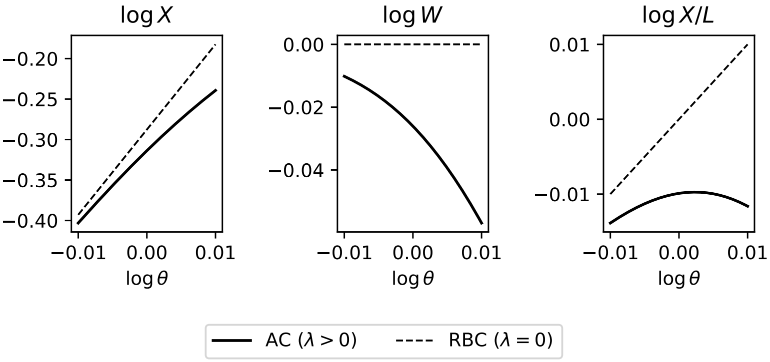
	(1)	(2)	(3)	(4)
	Outcome: $\Delta \log P_{it}$			
$\hat{u}_{it}^2$	-0.268	-0.262	-0.097	-0.087
	(0.025)	(0.023)	(0.034)	(0.033)
$\hat{u}_{it}^2 \times \Delta \log P_t$	0.376	0.376	0.443	0.431
	(0.123)	(0.124)	(0.171)	(0.167)
Sector x Time FE	✓	✓	✓	✓
Firm FE			✓	✓
TFP Control		✓		✓
N	41,578	41,578	41,206	41,206
R^2	0.239	0.261	0.385	0.403

Notes: standard errors are clustered by firm and year. Error bars are 95% confidence intervals.

Profitability | Binscatter and Symmetry | Dynamic Effects | Average Effect | Coefficient Time Series | Leverage and Betas | Alternative Measurement

Outline

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 Wedge Concavity

Conclusion

Introduced a theory of *Attention Cycles*

- Model: attention and misoptimization cycles are firms' rational reaction to higher stakes in downturns, and translate into mechanism for cyclical misallocation
- Data: consistent with the theory, we find direct evidence of (i) pro-cyclical "misoptimization" and (ii) state-dependent market punishment of misoptimizations
- Quantification: in model that matches the facts, endogenous attention magnifies aggregate shock response in low states

Conclusion

Introduced a theory of *Attention Cycles*

- Model: attention and misoptimization cycles are firms' rational reaction to higher stakes in downturns, and translate into mechanism for cyclical misallocation
- Data: consistent with the theory, we find direct evidence of (i) pro-cyclical "misoptimization" and (ii) state-dependent market punishment of misoptimizations
- Quantification: in model that matches the facts, endogenous attention magnifies aggregate shock response in low states

Companion work: Flynn and Sastry (2021) on "Strategic Mistakes" in abstract games, and additional macro and financial applications (speculative investment, price setting) Summary

New work: how do *macro narratives* arise, spread virally, and affect macro dynamics?

A Theory of *Strategic Mistakes* (Companion Paper)

Back

Model set-up:

- Continuum of agents care about own actions $x \in \mathcal{X}$, the state $\theta \in \Theta$ and an aggregate of actions of others $X : \Delta(\mathcal{X}) \to \mathbb{R}$. Payoff function: $u(x, X, \theta)$

A Theory of *Strategic Mistakes* (Companion Paper)

Back

Model set-up:

- Continuum of agents care about own actions $x \in \mathcal{X}$, the state $\theta \in \Theta$ and an aggregate of actions of others $X : \Delta(\mathcal{X}) \rightarrow \mathbb{R}$. Payoff function: $u(x, X, \theta)$
- Agents choose a non-parametric *stochastic choice* rule $p : \Theta \rightarrow \Delta(\mathcal{X})$, $p(x|\theta)$ PDF of actions $x \in \mathcal{X}$ in state $\theta \in \Theta$

A Theory of *Strategic Mistakes* (Companion Paper)

Back

Model set-up:

- Continuum of agents care about own actions $x \in \mathcal{X}$, the state $\theta \in \Theta$ and an aggregate of actions of others $X : \Delta(\mathcal{X}) \to \mathbb{R}$. Payoff function: $u(x, X, \theta)$
- Agents choose a non-parametric *stochastic choice* rule $p : \Theta \to \Delta(\mathcal{X})$, $p(x|\theta)$ PDF of actions $x \in \mathcal{X}$ in state $\theta \in \Theta$
- Cost of playing more precise actions mediated by a cost functional ($\pi \in \Delta(\Theta)$, $\phi : \mathbb{R}_+ \to \mathbb{R}$ strictly convex, *e.g.* entropy $\phi(x) = x \log x$)

$$c(P) = \sum_{\Theta} \int_{\mathcal{X}} \phi(p(x|\theta))\, \mathrm{d}x\, \pi(\theta)$$

A Theory of *Strategic Mistakes* (Companion Paper)

Back

Theoretical results: find conditions on (u, ϕ, X) such that:

1. Equilibria exist and are unique
2. Equilibrium action distribution $p(x|\theta)$ is FOSD-monotone in θ, aggregate $X(\theta)$ is monotone in θ
3. Equilibrium action distribution features dispersion or extent of mistakes that is monotone in the state
4. Equilibria are efficient

Technique: contraction-mapping arguments, which are essentially impossible under unrestricted information acquisition

Applications: financial speculation, price-setting, Bertrand competition. . .

An Example: General Motors in the Crisis

Back

[The] deteriorating economic and market conditions that have driven the drop in vehicle sales, including declines in real estate and equity values, rising unemployment, tightened credit markets, depressed consumer confidence and weak housing markets, may not improve significantly during 2010 and may continue past 2010 and could deteriorate further.

General Motors, 2009

An Example: General Motors in the Crisis

Back

[The] deteriorating economic and market conditions that have driven the drop in vehicle sales, including declines in real estate and equity values, rising unemployment, tightened credit markets, depressed consumer confidence and weak housing markets, may not improve significantly during 2010 and may continue past 2010 and could deteriorate further.

General Motors, 2009

Macro words: not that common in 10Q/K, but common in references

An Example: General Motors in the Crisis

Back

[The] deteriorating economic and market conditions that have driven the drop in vehicle sales, including declines in real estate and equity values, rising unemployment, tightened credit markets, depressed consumer confidence and weak housing markets, may not improve significantly during 2010 and may continue past 2010 and could deteriorate further.

General Motors, 2009

Macro words: not that common in 10Q/K, but common in references

Not macro words: too common in both!

Examples of Macro Words

Seem reasonable	False positives
unemployment	equilibrium
nominal	equation
productivity	theory
economists	question
macroeconomics	determinants
Fed	
inflation	
recession	

All Macro Words

Back

Word	Corr[tf,Unemp]	Word	Corr[tf,Unemp]	Word	Corr[tf,Unemp]	Word	Corr[tf,Unemp]	Word	Corr[tf,Unemp]
RECESSION	0.874	MUCH	0.397	SHIFTS	0.197	WORKERS	0.006	ANSWER	-0.221
ECONOMY	0.843	CURVE	0.393	LEADS	0.167	COUNTRY	-0.003	IMAGINE	-0.228
UNEMPLOYMENT	0.816	MACROECONOMICS	0.392	COUNTRIES	0.155	PRODUCTIVITY	-0.014	CON	-0.236
SAVING	0.641	BUY	0.388	DETERMINANTS	0.154	LABOR	-0.014	AGGREGATE	-0.314
PRICES	0.633	PRODUCTION	0.380	MONETARY	0.147	INFLATION	-0.030	LET	-0.363
FALL	0.631	ECONOMICS	0.355	SUP	0.130	RUN	-0.031	CALLED	-0.372
SPENDING	0.606	POINT	0.354	WAGE	0.127	THEORY	-0.043	DEMANDED	-0.379
UNEMPLOYED	0.589	FALLS	0.350	CUT	0.123	POLICY	-0.051	QUESTION	-0.438
ECONOMISTS	0.582	CONSUMER	0.345	FED	0.121	WAGES	-0.056	RISE	-0.481
MACROECONOMIC	0.570	PLY	0.341	BUDGET	0.104	MONEY	-0.078		
GOVERNMENT	0.549	HAPPENS	0.305	CONSUMERS	0.094	EXPLAIN	-0.107		
DEBATE	0.516	REAL	0.303	MULTIPLIER	0.066	EXPORTS	-0.125		
THUS	0.515	SUPPLY	0.299	WORLD	0.066	ARGUE	-0.130		
CONSUMPTION	0.505	DEMAND	0.286	FIRMS	0.065	TRADE	-0.135		
CHAPTER	0.476	EXAMPLE	0.286	EQUATION	0.059	FIGURE	-0.145		
PEOPLE	0.468	OUTPUT	0.274	WANT	0.054	THINK	-0.184		
WEALTH	0.441	GET	0.265	EQUILIBRIUM	0.045	SUPPOSE	-0.193		
SHOWS	0.437	NOMINAL	0.263	ECONOMIST	0.016	PROBLEM	-0.193		
SLOPE	0.423	RISES	0.262	CHAPTERS	0.011	PERCENT	-0.205		
NATION	0.398	QUANTITY	0.227	SHIFT	0.011	GOODS	-0.211		

Measuring Sector-Specific Production Functions Back

$$\log \text{Sales}_{it} = \mu^{-1} \left(\alpha_{L,j(i)} L_{it} + \alpha_{M,j(i)} M_{it} + \alpha_{K,j(i)} K_{it} \right)$$

where $\alpha_{L,j(i)} + \alpha_{M,j(i)} + \alpha_{K,j(i)} = 1$, for all j, and $\mu^{-1} = 0.75$ is one over markup.

	Quantity	Expenditure
Production, x_{it}	—	`sale`
Employment, L_{it}	`emp`	`emp` $\times$ industry wage (from CBP)
Materials, M_{it}	—	`cogs` + `xsga` − `dp` − wage bill
Capital, K_{it}	—	`ppegt` plus net investment (value of stock)

1. In industry j, calculate the estimated materials and labor shares over entire sample
2. If $\text{Share}_{M,j'} + \text{Share}_{L,j'} \leq \mu^{-1}$, where μ is externally identified markup, then set

$$\alpha_{M,j'} = \mu \cdot \text{Share}_{M,j'}$$
$$\alpha_{L,j'} = \mu \cdot \text{Share}_{L,j'}$$
$$\alpha_{K,j'} = 1 - \alpha_{M,j'} - \alpha_{L,j'}$$

3. Otherwise, adjust shares to match the assumed returns to scale

Estimating the Policy Function

Back

1. Estimate via OLS
$$\log L_{it} = \gamma_i + \chi_{j(i),t} + \beta \log \theta_{it} + m_{it}$$
2. Use residuals m^0_{it} to estimate AR(1) persistence, ρ, via
$$m^0_{it} = \rho m^0_{i,t-1} + \hat{u}_{it}$$
3. Estimate the "quasi-differenced" equation for labor choice,
$$\log L_{it} - \hat{\rho} \log L_{i,t-1} = \tilde{\eta}_i + \tilde{\chi}_{j(i),t} + \beta_0 \log \hat{\theta}_{it} + \beta_1 \log \hat{\theta}_{i,t-1} + (m_{it} - \hat{\rho} m_{i,t-1})$$
4. Translate into estimates of the residual,
$$\hat{u}_{it} = m_{it} - \hat{\rho} m_{i,t-1}$$

Weights for Averaging Misoptimization Dispersion

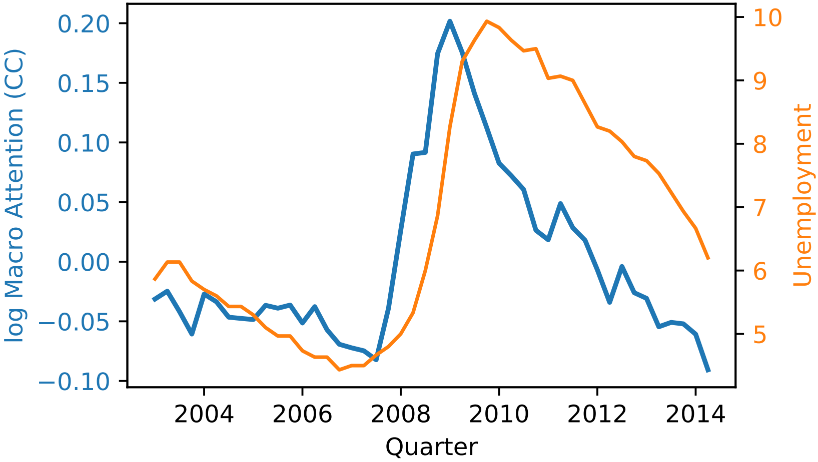
Back

$s_{it}^* = \exp\left(\hat{\beta} \log \theta_{it}\right)$ from the regression:

$$\log \text{Sales}_{it} = \beta \log \theta_{it} + \eta_i + \chi_{j(i),t} + \epsilon_{it}$$

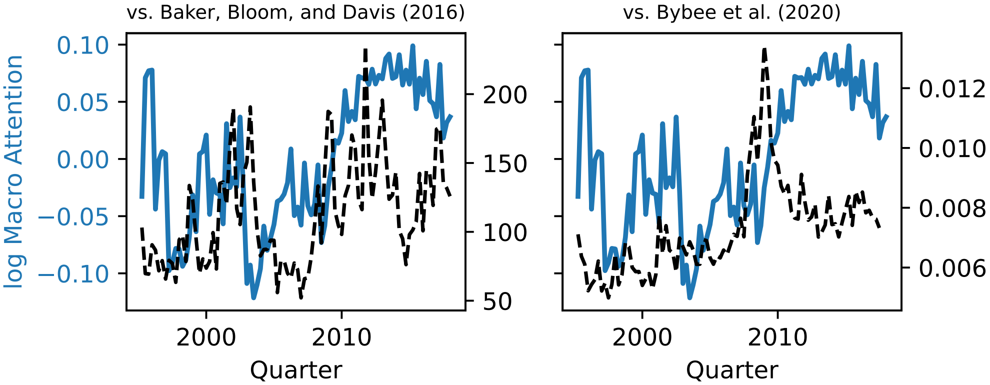
Conference-Call Measure Also Counter-cyclical and Persistent

Back 1 2



Macro Attention More Persistent Than News Attention

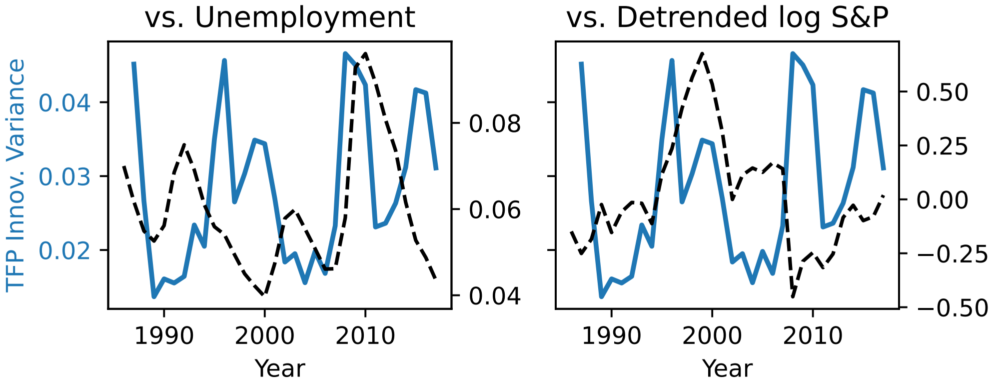
Back



Fundamental Dispersion is Pro-Cyclical Back Back

"TFP Innov. Variance" = (weighted) variance of ϵ_{it}:

$$\theta_{it} = \rho_\theta \theta_{i,t-1} + \chi_{j(i),t} + \gamma_i + \epsilon_{it}$$



Conference-Call Measure Has Similar Empirical Patterns

Back

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome:					
	$\log \text{MacroAttnCC}_t$			$\log \text{MacroAttnCC}_{it}$		
$\frac{\text{Unemployment}_t}{100}$	2.481 (0.596)					
$\log \text{SPDetrend}_t$		-0.270 (0.056)				
$\log \text{MacroAttnCC}_{t-1}$			0.949 (0.068)			
$\log \text{MacroAttn10K}_{it}$				0.463 (0.034)	0.372 (0.036)	0.121 (0.028)
Firm FE?						✓
Sector x Time FE?					✓	✓
N	46	46	45	8,023	7,994	7,670
R^2	0.376	0.593	0.873	0.123	0.308	0.804

Note: In the first three columns, standard errors are HAC robust with a bandwidth (Bartlett kernel) of four quarters. In the second three columns, standard errors are double-clustered by year and firm ID.

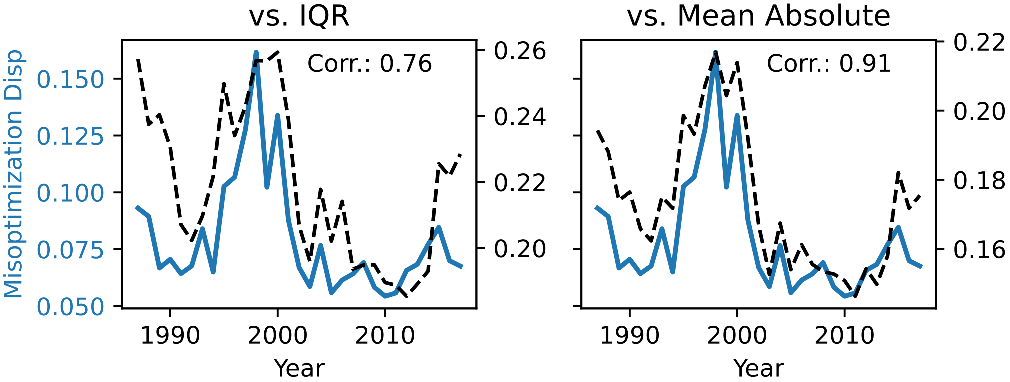
Robustness: Misoptimization Lowers Returns and Profit

Back

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Outcome: $\Delta \log P_{it}$				Outcome: π_{it}			
$\hat{u}_{it}^2$	-0.236	-0.230	-0.060	-0.051	-0.316	-0.316	-0.106	-0.105
	(0.026)	(0.026)	(0.032)	(0.032)	(0.024)	(0.024)	(0.018)	(0.017)
Sector x Time FE	✓	✓	✓	✓	✓	✓	✓	✓
Firm FE			✓	✓			✓	✓
TFP Control		✓		✓		✓		✓
N	41,578	41,578	41,206	41,206	51,015	51,015	50,966	50,996
R^2	0.238	0.261	0.384	0.403	0.117	0.131	0.663	0.681

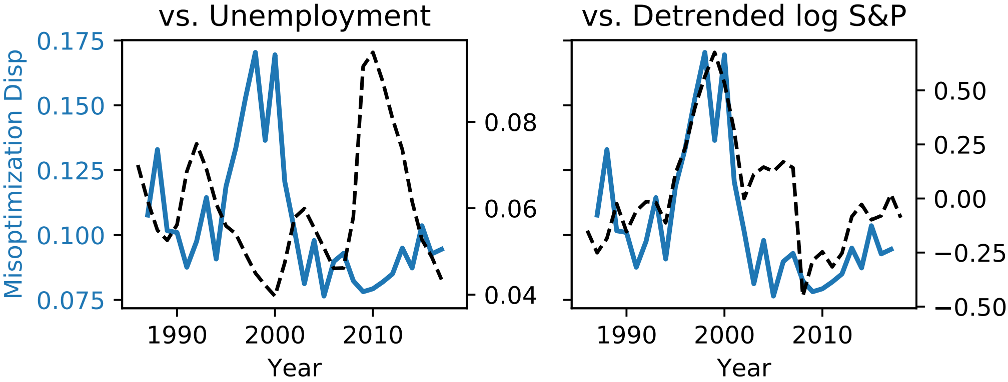
Note: Standard errors are double-clustered at the year and firm level. π_{it} = measured profitability of firm, or EBIT over lagged variable costs.

Fact 2: Misoptimization is Pro-Cyclical (IQR and Mean Absolute)



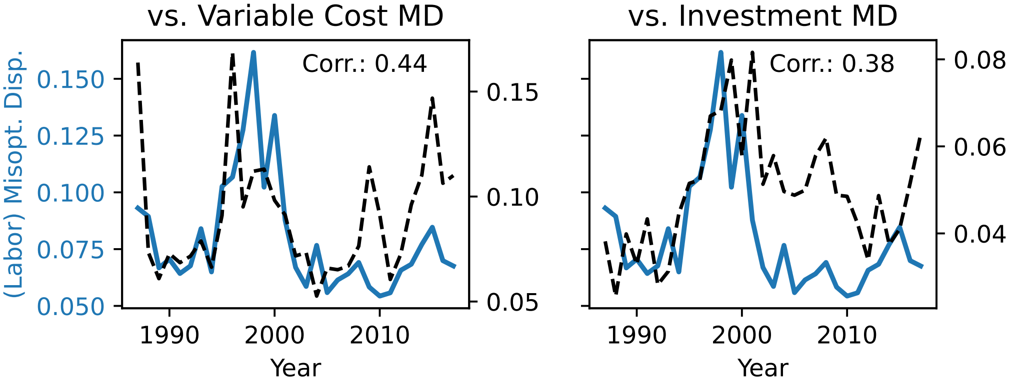
Back

Fact 2: Misoptimization is Pro-Cyclical (Unweighted Average)



Back

Fact 2: Misoptimization is Pro-Cyclical (Materials and Investment)



Back

Robustness: Pro-Cyclical Misoptimization

Back

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Outcome: MisoptimizationDispersion$_t$						
Unemployment$_t$	-0.810 (0.265)	-0.580 (0.214)	-0.895 (0.299)	-0.920 (0.309)	-0.857 (0.282)	-0.692 (0.272)	-1.256 (0.450)
Period	1986-2018	1986-2018	1986-2018	1986-2018	1986-2018	1998-2018	1986-2018
t, t^2 Control?	✓						
Manufacturing?		✓					
Sector Policy Fn.?			✓				
t-varying Policy Fn.?				✓			
Quadratic Policy Fn?					✓		
Pre-Period TFP?						✓	
OP (96) TFP?							✓
N	31	31	31	31	31	20	31
R^2	0.420	0.281	0.284	0.277	0.278	0.293	0.215

Note: Standard errors are HAC-robust with a 2-year Bartlett Kernel. The baseline estimate is a coefficient of -0.841 with a standard error of 0.341.

Robustness: Market Punishes Inattentiveness Harder in Downturns

Back

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Outcome: $\Delta \log P_{it}$						
$\hat{u}_{it}^2$	-0.097	-0.239	-0.101	-0.168	-0.090	-0.099	-0.109
	(0.034)	(0.941)	(0.035)	(0.045)	(0.036)	(0.035)	(0.034)
$\hat{u}_{it}^2 \times \Delta \log P_t$	0.443	0.941	0.415	0.680	0.420	0.330	0.447
	(0.171)	(0.370)	(0.169)	(0.182)	(0.156)	(0.163)	(0.163)
Sector x Time FE	✓	✓	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓	✓
Baseline	✓						
Adj. Control		✓					
Leverage Control			✓				
Manufacturing				✓			
Sector Policy Fn.					✓		
t-varying Policy Fn.						✓	
Quadratic Policy Fn.							✓
N	41,206	35,388	41,016	22,902	41,197	41,203	41,203
R^2	0.385	0.387	0.385	0.367	0.385	0.384	0.385

Note: Standard errors are double-clustered at the year and firm level. The baseline coefficient estimates are -0.097 (SE: 0.034) and 0.443 (SE: 0.171).

Robustness: Attentive Firms Make Smaller Mistakes

Back 1 2

	(1)	(2)	(3)	(4)	(5)	(6)
			Outcome: $\hat{u}_{it}^2$			
log $\text{MacroAttention}_{it}$	-0.0163	-0.0035	-0.0076	-0.0127	-0.0107	-0.0084
	(0.0066)	(0.0015)	(0.0028)	(0.0037)	(0.0028)	(0.0028)
Sector x Time FE	✓	✓	✓	✓	✓	✓
Conference Call Measure	✓					
Adj. Control		✓				
Leverage Control			✓			
Manufacturing				✓		
Sector Policy Fn.					✓	
t-varying Policy Fn.						✓
N	5,997	24,024	28,133	14,891	28,283	28,275
R^2	0.053	0.072	0.060	0.053	0.041	0.054

Note: Standard errors are double-clustered at the year and firm level. The baseline estimate, is -0.0081 with a standard error of 0.0028.

Advantage in Adversity: Winning the Next Downturn

FEBRUARY 04, 2019

By Martin Reeves, David Rhodes, Christian Ketels, and Kevin Whitaker

Back

Prediction: *Any* Shifter of Profit Curvature Should Predict Attention

Back

- Say we have *any other* variation in curvature of firm's objective function (e.g., firm-level variation in slope of demand curve)
- *Directly elicited* in the Coibion, Gorodnichenko, and Kumar (2018) survey of firms in New Zealand:

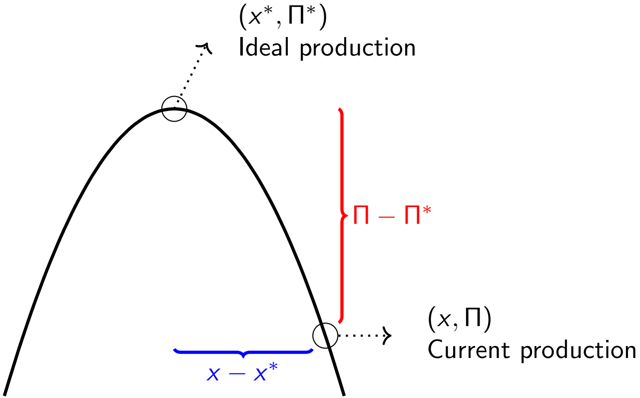
> If this firm was free to change its price (i.e. suppose there was no cost to renegotiating contracts with clients, no costs of reprinting catalogues, etc. . .) right now or in three months, by how much would it change its price in either case? Please provide a percentage answer. By how much do you think profits would change as a share of revenues in either case? Please provide a numerical answer in percent.

Recovering the Curvature of Profits from the Data Back

$$\Pi - \Pi^* = \frac{1}{2} u_{xx}(x^*(\theta), \theta) \cdot (x - x^*)^2$$

$$\frac{\Pi - \Pi^*}{(x - x^*)^2} = \frac{1}{2} u_{xx}(x^*(\theta), \theta)$$

$$= \text{ProfCurv}_{it}$$



High-curvature Firms Pay More Attention

Back

$$Y_{it} = \alpha + \beta \cdot \text{ProfCurv}_{it} + \gamma' X_{it} + \epsilon_{it}$$

Outcomes: back-cast error for recalling macro stat over last 12 months; indicator of whether you "keep track of" this variable (self-reported)

Variable	Inflation		GDP Growth		Unemployment	
Outcome	$\lvert BCE \rvert_{it}$	Track?	$\lvert BCE \rvert_{it}$	Track?	$\lvert BCE \rvert_{it}$	Track?
ProfCurv_{it}	-0.0328	0.050	-0.072	0.019	0.121	-0.022
	(0.091)	(0.029)	(0.041)	(0.028)	(0.077)	(0.081)
Controls?	✓	✓	✓	✓	✓	✓
R^2	0.457	0.332	0.006	0.074	0.032	0.065
N	3,153	1,254	1,256	1,254	716	1,254

Controlling for: 3-digit industries; bins in total value of output

CGK Survey: Hypothetical News Demand Back

Suppose that you hear on TV that the economy is doing well [or poorly]. Would it make you more likely to look for more information?

Response	Poorly	Well
Much more likely	44.96	9.77
Somewhat more likely	30.91	19.42
No change	12.56	8.67
Somewhat less likely	7.16	53.35
Much less likely	4.40	8.79
Total	100.00	100.00

Labor Wedge Shocks Propagate Like TFP Shocks

Back

Corollary: Dynamics with Labor Wedge Shocks

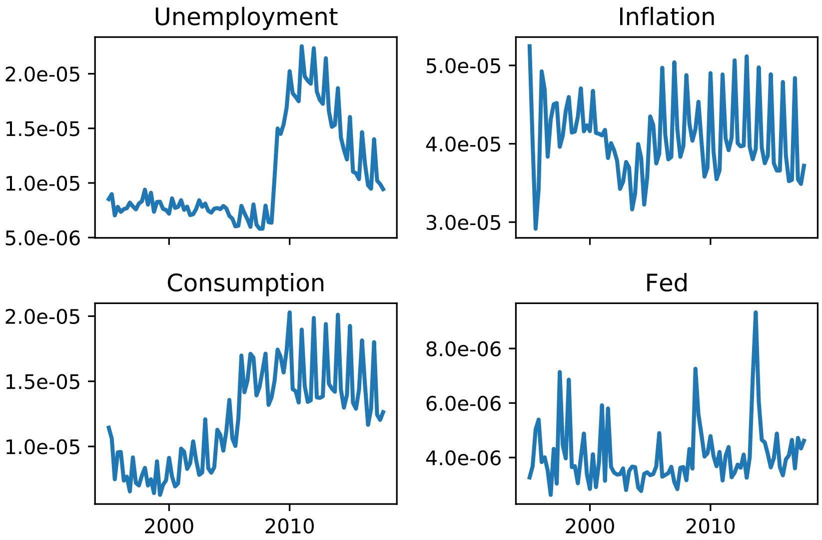
Consider a variant economy in which $\theta_t \equiv 0$, but $\bar{w}$, the constant in the wage rule, is time-varying. In particular, define $\epsilon_w = -(\log \bar{w}_t - \log \bar{w})$, where $\log \bar{w}$ is a long-run average. Then output is given as in the previous proposition, with ϵ_w replacing $\log \theta_t$.

Key observation in proof: θ_i and $1/\bar{w}$ enter in exactly the same way in the firm's problem.

Interpretation: no reason for cycles in our model to be driven by "technological fundamentals"—same propagation mechanisms work in a demand-driven economy.

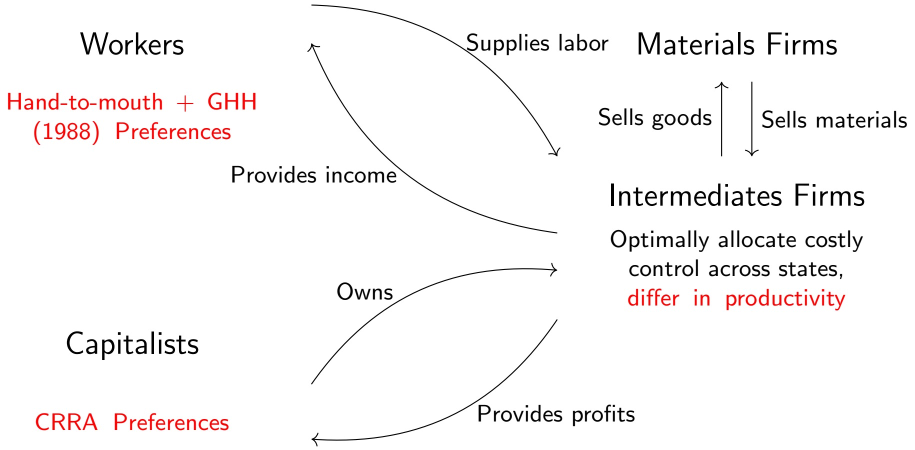
Pattern Dominated by "Real Conditions," Not Inflation or Policy

Back



Extended Model: Schematic Overview

Back



Summary of Results: Interpreting Measurement

Back

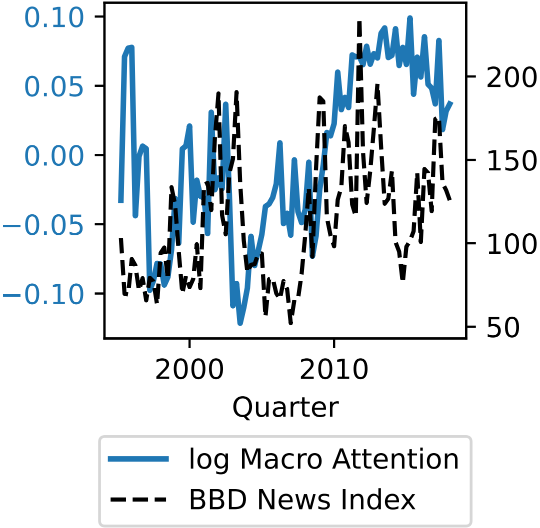
- **Lemma 1**: log-linear policy function for sales, in physical or revenue terms
  - Proof: straightforward algebra, given Cobb-Douglas and isoelastic structure
- **Lemma 2**: log-linear policy function for any input choice
  - Proof: similar to previous
- **Lemma 3**: cost shares identify output elasticities, up to (empirically small) Jensen's inequality correction
  - Proof: follows from fact that input choices are "right on average"
  - Jensen's inequality term arises *only* because of downward sloping demand; otherwise, input and sales mistakes cancel and cost shares exactly equal output elasticities

Sample Restrictions for Measuring Misoptimizations Back

1. Sales, material expenditures, and capital stock are strictly positive
   - Necessary for meaningful production, policy function estimation
2. Employees exceed 10
   - Screens out excessively young or poorly reported firms
3. 2-digit NAICS is not 52 (Finance and Insurance) or 22 (Utilities)
   - Sectors have drastically different production technology and market structures
4. Acquisitions as a proportion of assets (aqc over at) does not exceed 0.05.
   - Simple way to screen out large acquisition events
5. Fiscal year ends in December
   - Streamlines calculation of aggregates, comparison to business cycle

Macro Attention: Persistence and Comparison to News Indices

Back

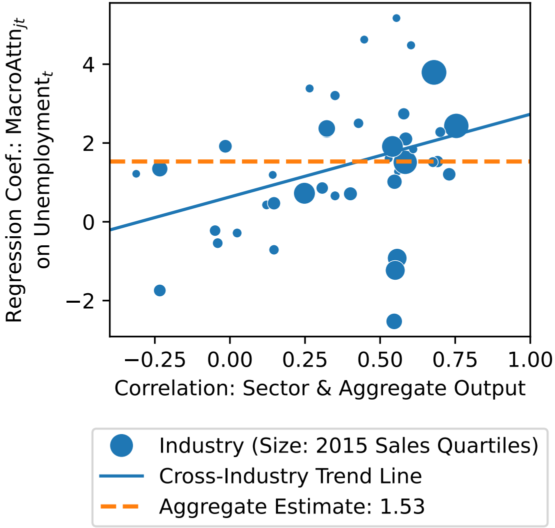


- Macro Attention nearly as persistent as business cycle (AR(1) coefficient: 0.820)
- *Comparison to literature*: Macro Attention more persistent than news-based Policy Uncertainty (Baker, Bloom, and Davis, 2016)

Economic News: BKMX

Macro Attention: Industry-Level Patterns in Cyclicality

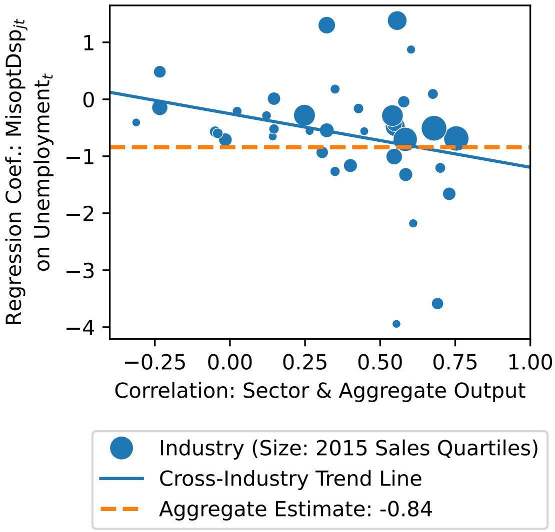
Back 1 Back 2



- Each dot is one of 42 industries (hybrid of NAICS2 and NAICS3)
- Attention counter-cyclical in most industries, but especially in pro-cyclical ones

Misoptimization Dispersion: Industry-Level Patterns

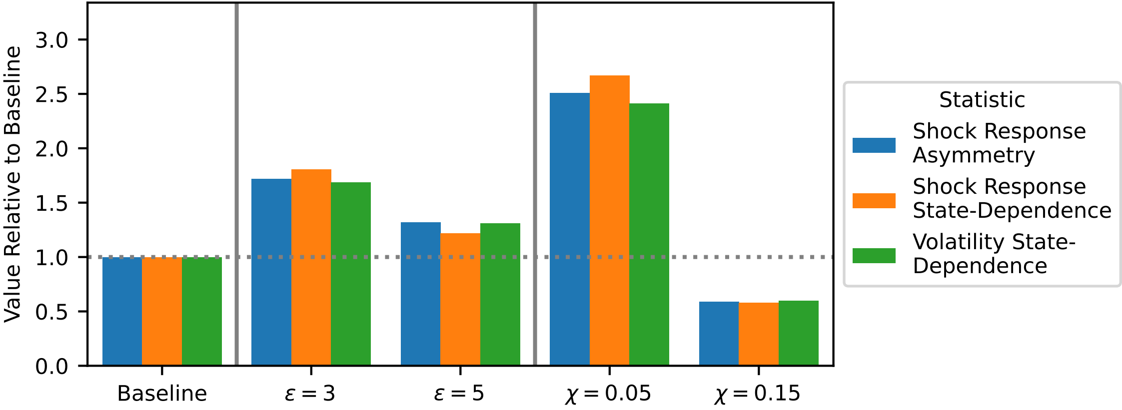
Back 1 Back 2



- *Industry-level analysis*: like with verbal attention, stronger patterns for more cyclical sectors

Robustness of Macro Predictions to Parameter Choices

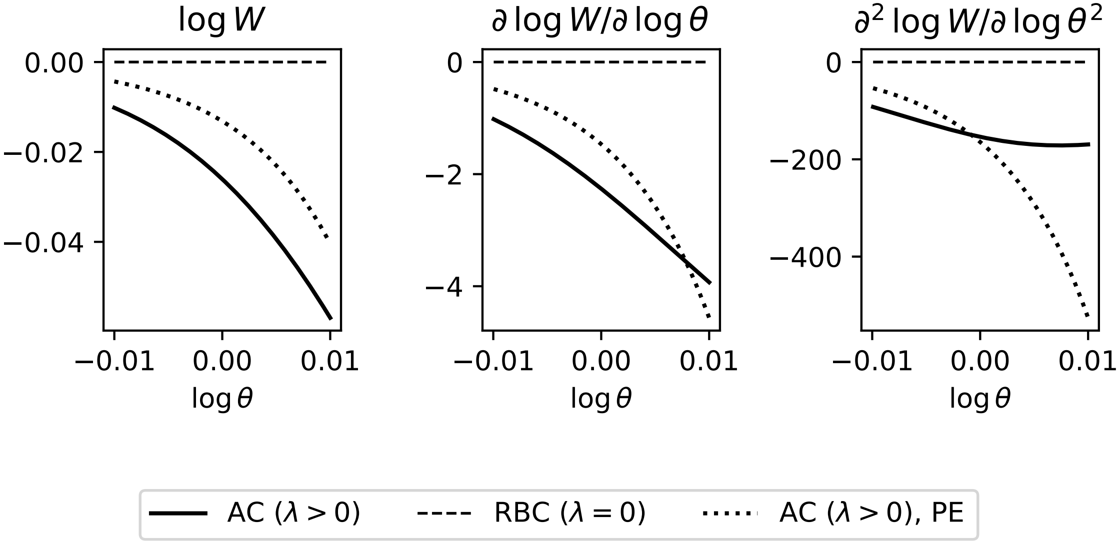
Exercise: re-calibrate model for different values of (ϵ, γ, χ), and re-do predictions for dynamics



Back

The Attention Wedge and GE Decomposition

Back



Translation of Results: Wage Rule Shocks Back

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

Symmetry of θ and $\bar{w}^{-1}$ in firm's problem $\rightarrow$ same implications for production + attention

	Model 1: θ Shock	Model 2: $\bar{w}^{-1}$ Shock
Proposition 1 equilibrium basics	Output monotone in θ	Output monotone in $\bar{w}^{-1}$
Proposition 2 firm choices	unchanged	unchanged
Proposition 3 dynamics and attention wedge	Wedge depends on θ	Wedge depends on $\bar{w}^{-1}$
Corollary 1 productivity and attention wedge	Ambiguous cyclicality of TFP	Counter-cyclical TFP
Corollary 2 dynamics properties	Describes dynamics in θ	Describes dynamics in $\bar{w}^{-1}$

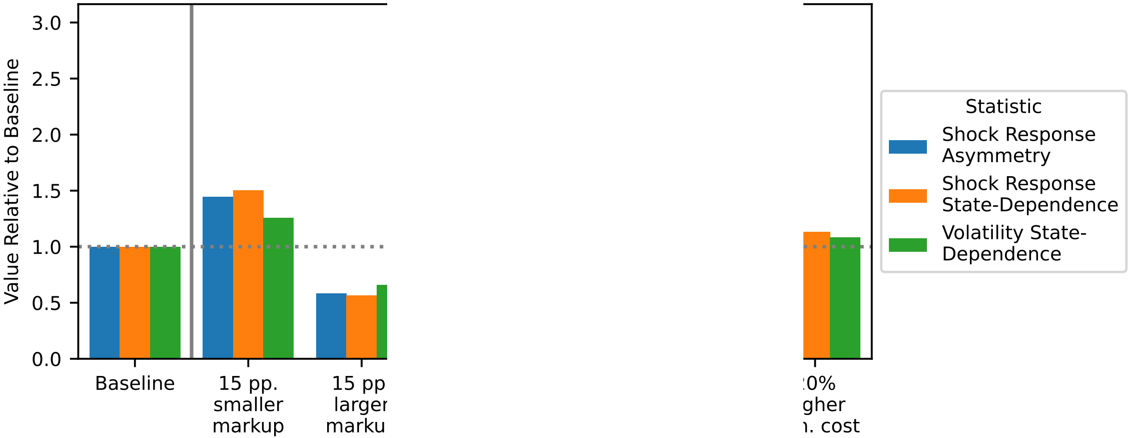
Counterfactuals: Attention Cycles Under Structural Changes Back

Comparative statics: what happens to main dynamic predictions if parameters change?

Model is *not* recalibrated

Counterfactuals: Attention Cycles Under Structural Changes

Back



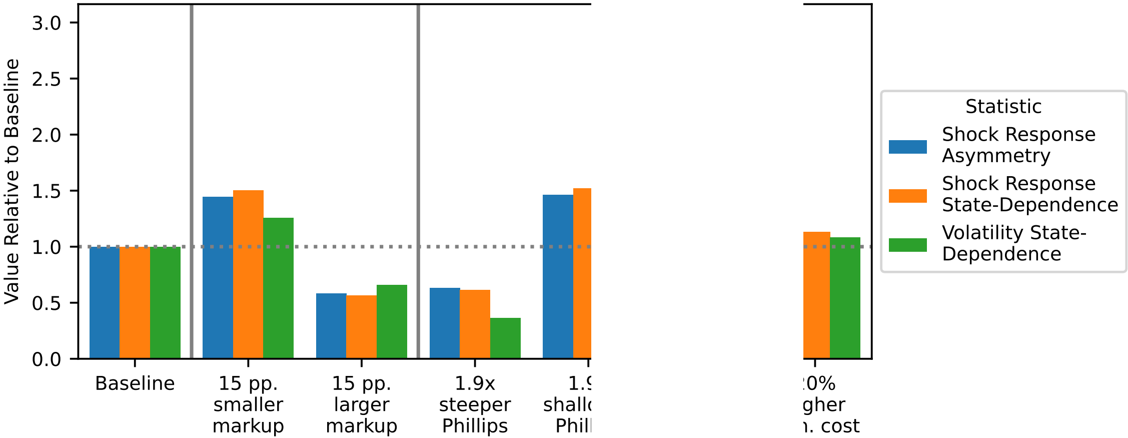
Rise in effective market power

Demirer (2020): 15 pp increase in markups over last half-century, translated to ϵ

Idea in model: more "competition" raises stakes for misoptimizations

Counterfactuals: Attention Cycles Under Structural Changes

Back



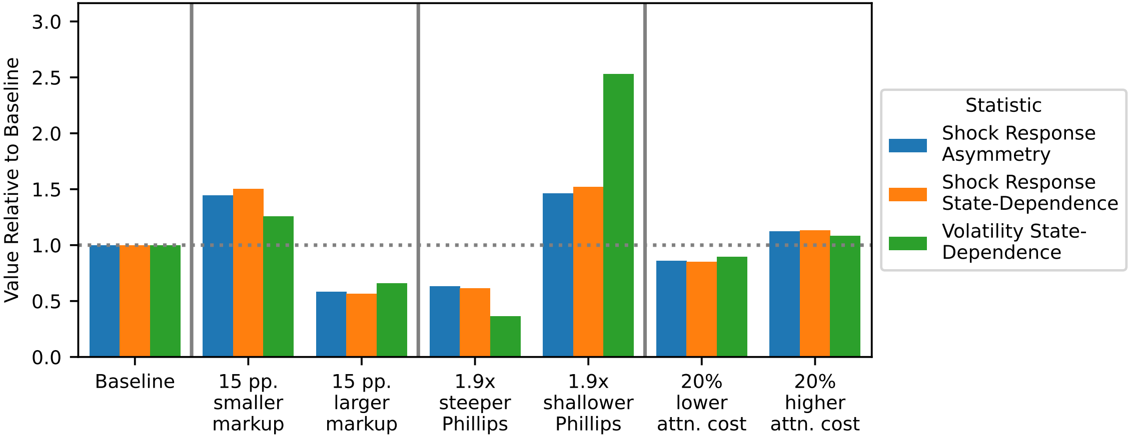
Flattening Wage Phillips Curve

Galí and Gambetti (2020): wage Phillips curve has flattened by factor of 1.9 since the 1980s

Idea in model: lower wage pressure → more cyclical attention

Counterfactuals: Attention Cycles Under Structural Changes

Back



Higher-uncertainty regime

Some economic/political/social shocks make optimal choices less "knowable"

Idea in model: higher attention costs increase sensitivity of misoptimization to incentives

Macro-attentive Firms Make Smaller Misoptimizations

Back

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

	Outcome: $\hat{u}_{it}^2$			
	(1)	(2)	(3)	(4)
$\log \text{MacroAttention}_{it}$	-0.0081	-0.0052	-0.0058	-0.0056
	(0.0028)	(0.0029)	(0.0044)	(0.0038)
Sector x Time FE?	✓	✓	✓	✓
Firm FE?			✓	✓
Other Controls?		✓		✓
N	28,279	24,392	27,875	23,930
R^2	0.053	0.067	0.383	0.384

Note: Standard errors are double-clustered at the year and firm level.

Alternative Measurement

Test: Does Punishment of Misoptimization Line Up With the Cycle? Back

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

$\Delta \log P_{it}$: year-on-year stock return

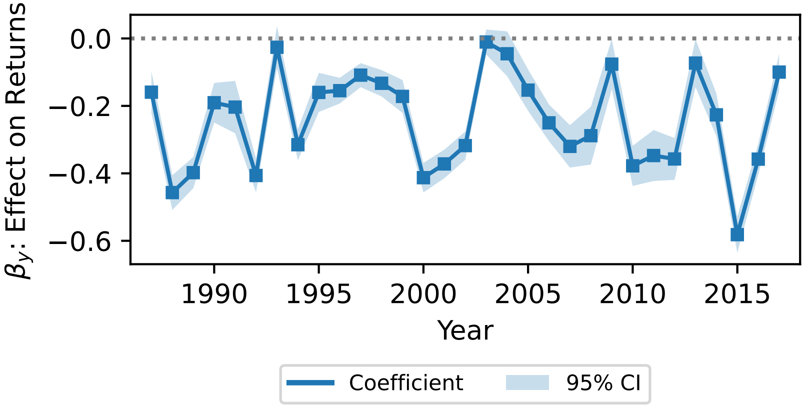
Industry-by-year fixed effects sweep out background trends

Hypothesis from model: $|\beta_y|$ large when stakes increase, or economy experiences duress

Test: Does Punishment of Misoptimization Line Up With the Cycle?

Back

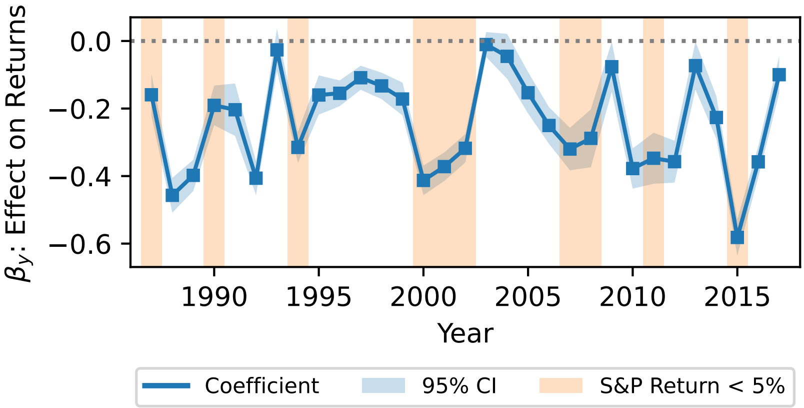
$$\Delta \log P_{it} = \sum_y \beta_y \cdot \hat{u}_{it}^2 \cdot \mathbb{I}[t = y] + \chi_{j(i),t} + \epsilon_{it}$$



Average Effect

Test: Does Punishment of Misoptimization Line Up With the Cycle? Back

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



Average Effect

State-Dependent Effect Driven by Returns, Not Profitability

Back

$$\pi_{it} = \beta \cdot \hat{u}^2_{it} + \phi \cdot \left(\hat{u}^2_{it} \times \Delta \log \text{SP500}_t\right) + \chi_{j(i),t} + \gamma_i + \epsilon_{it}$$

	(1) Outcome:
$\hat{u}^2_{it}$	-0.114
	(0.020)
$\hat{u}^2_{it} \times \Delta \log P_t$	0.112
	(0.089)
π_{it}	
$\pi_{it} \times \Delta \log P_t$	
Firm FE	✓
Sector x Time FE	✓
N	50,966
R^2	0.663
First-stage F	

State-Dependent Effect Driven by Returns, Not Profitability

Back

$$\Delta \log P_{it} = \beta \cdot \pi_{it} + \phi \cdot (\pi_{it} \times \Delta \log \text{SP500}_t) + \chi_{j(i),t} + \gamma_i + \epsilon_{it}$$

	(1) Outcome: π_{it}	(2) Outcome: $\Delta \log P_{it}$	(3)
$\hat{u}^2_{it}$	-0.114		
	(0.020)		
$\hat{u}^2_{it} \times \Delta \log P_t$	0.112		
	(0.089[illegible]		
π_{it}		0.42[illegible]	
		(0.03[illegible]	[illegible])
$\pi_{it} \times \Delta \log P_t$		-0.30[illegible]	[illegible]
		(0.16[illegible]	[illegible])
Firm FE	✓	✓	[illegible]
Sector x Time FE	✓	✓	[illegible]
N	50,96[illegible]	40,87[illegible]	[illegible]
R^2	0.663	0.40[illegible]	[illegible]
First-stage F			[illegible]

State-Dependent Effect Driven by Returns, Not Profitability

Back

$$\Delta \log P_{it} = \beta \cdot \pi_{it} + \phi \cdot (\pi_{it} \times \Delta \log \text{SP500}_t) + \chi_{j(i),t} + \gamma_i + \epsilon_{it}$$

First stage: $\hat{u}_{it}^2, \hat{u}_{it}^2 \times \Delta \log \text{SP500}_t$ as instruments

	(1) Outcome: π_{it}	(2) Outcome: $\Delta \log P_{it}$	(3) Outcome: $\Delta \log P_{it}$
$\hat{u}_{it}^2$	-0.114		
	(0.020)		
$\hat{u}_{it}^2 \times \Delta \log P_t$	0.112		
	(0.089)		
π_{it}		0.421	0.690
		(0.034)	(0.305)
$\pi_{it} \times \Delta \log P_t$		-0.303	-1.642
		(0.165)	(0.632)
Firm FE	✓	✓	✓
Sector x Time FE	✓	✓	✓
N	50,966	40,879	40,879
R^2	0.663	0.402	
First-stage F			17.80

The Attention Wedge in Productivity

Back

Corollary: The Attention Wedge and Measured Productivity

Productivity $A := X/L$ can be written as

$$\log A(\log \theta, \lambda) = \log \theta + \chi \epsilon \log W(\log \theta, \lambda)$$

Finding 2 (ctd.): Similar Profit Effects in All States Back

$$\pi_{it} = \beta \cdot \hat{u}_{it}^2 + \phi \cdot \left(\hat{u}_{it}^2 \times \Delta \log \mathrm{SP500}_t\right) + \chi_{j(i),t} + \gamma_i + \epsilon_{it}$$

	β	ϕ
estimate :	-0.11	0.11
SE :	(0.02)	(0.09)

Interpretation: profit curvature channel goes right way in data, but is smaller and more imprecise than the risk-pricing channel

Finding 2 (ctd.): Similar Profit Effects in All States Back

$$\pi_{it} = \beta \cdot \hat{u}_{it}^2 + \phi \cdot \left(\hat{u}_{it}^2 \times \Delta \log \text{SP500}_t\right) + \chi_{j(i),t} + \gamma_i + \epsilon_{it}$$

	β	ϕ
estimate :	-0.11	0.11
SE :	(0.02)	(0.09)

Interpretation: profit curvature channel goes right way in data, but is smaller and more imprecise than the risk-pricing channel

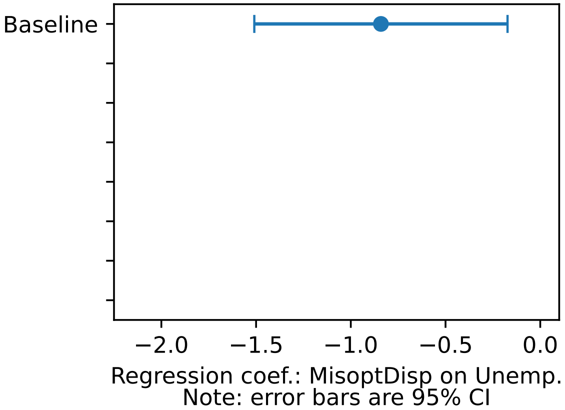
Also in the paper: Link

- Profitability has magnified effect on returns when S&P is doing poorly
- IV model: "misoptimization-caused variation" in profitability has (more pronounced) state-dependent effects

Robustness: Alternative Measurement of Misoptimizations

Back

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

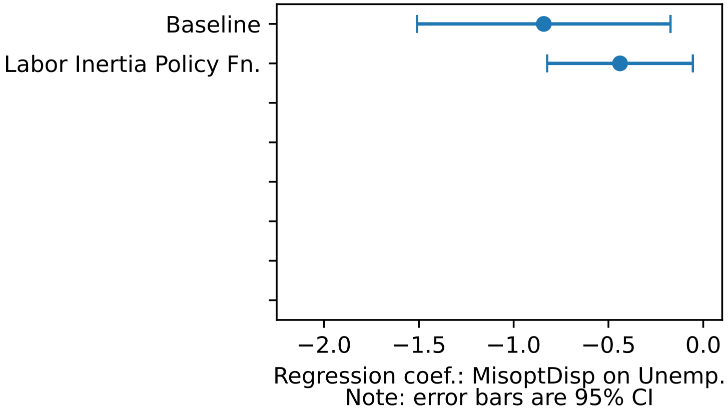


Robustness: Alternative Measurement of Misoptimizations

Back

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

Robustness: $\log L_{it} = \gamma_i + \chi_{j(i),t} + \beta \log \hat{\theta}_{it} + \tau \log L_{it-1} + m_{it}$



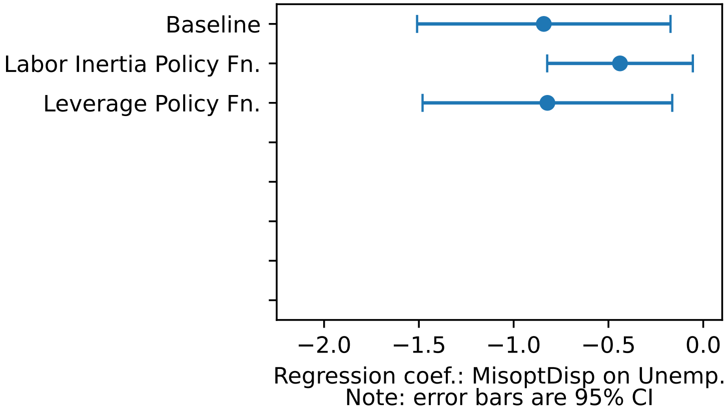
Allow for policy function (TFP responsiveness) to vary by sector

Robustness: Alternative Measurement of Misoptimizations

Back

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

Robustness: $\log L_{it} = \gamma_i + \chi_{j(i),t} + \beta \log \hat{\theta}_{it} + m_{it} + \xi \text{Lev}_{it} + \phi \cdot \text{Lev}_{it} \times \log \hat{\theta}_{it}$



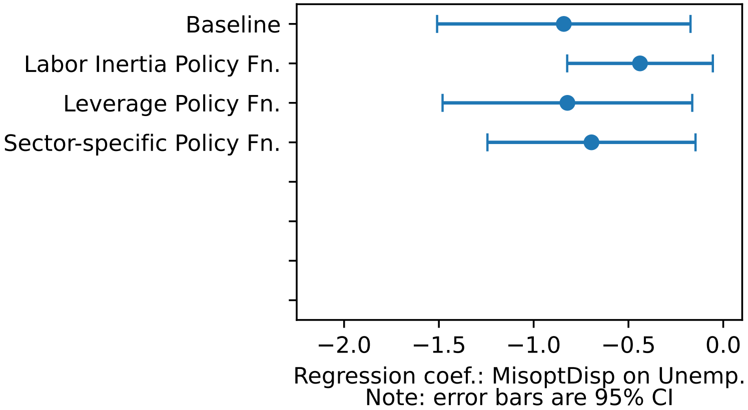
Allow for policy function (TFP responsiveness) to change over time (Decker et. al, 2020)

Robustness: Alternative Measurement of Misoptimizations

Back

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

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



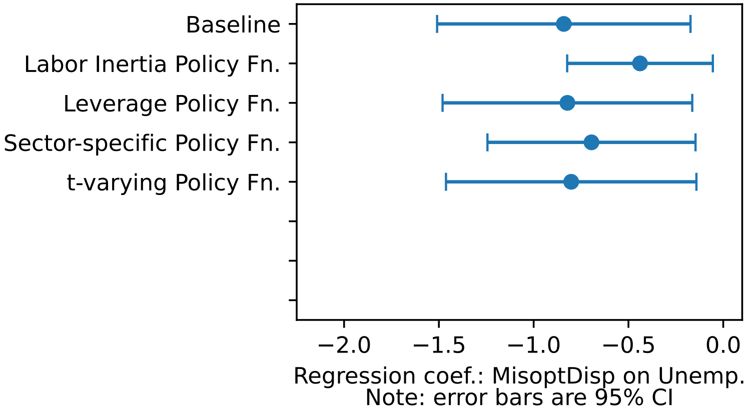
Allow for policy function to depend non-linearly on TFP, or have state-dependent elasticity (can capture asymmetries, as in Ilut et. al, 2018)

Robustness: Alternative Measurement of Misoptimizations

Back

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

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



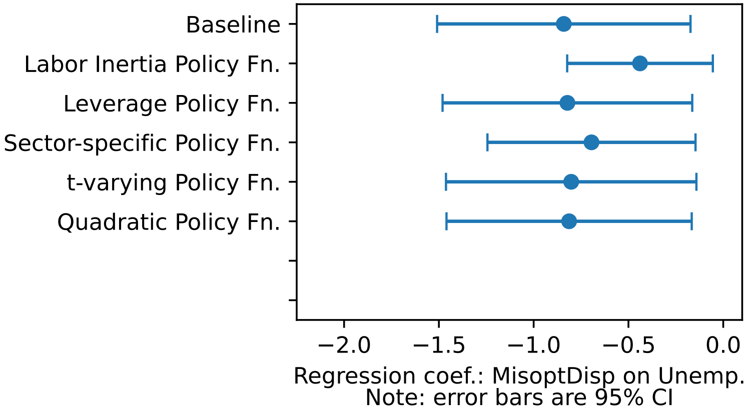
Absorb inertia from physical adjustment costs like hiring and firing costs

Robustness: Alternative Measurement of Misoptimizations

Back

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

Robustness: $\log L_{it} = \gamma_i + \chi_{j(i),t} + \beta_0 \log \hat{\theta}_{it} + \beta_1 (\log \hat{\theta}_{it})^2 + m_{it}$



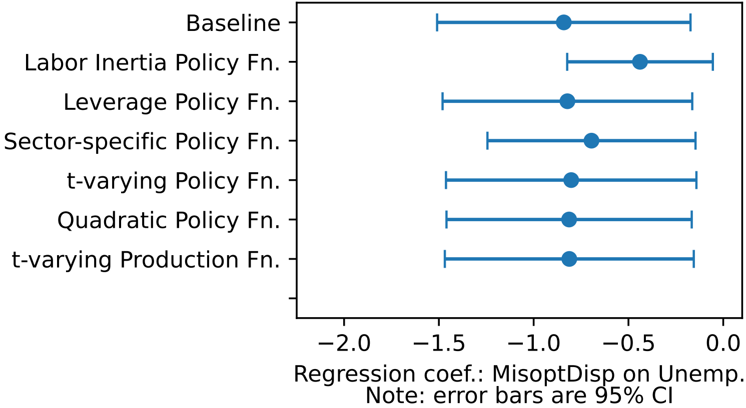
Allow for direct effect of leverage + interaction with productivity to proxy for credit frictions (Otonello and Winberry, 2020)

Robustness: Alternative Measurement of Misoptimizations

Back

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

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



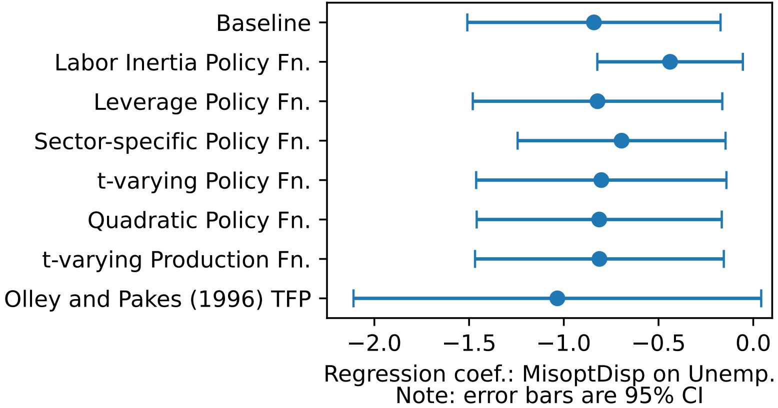
Allow production function (output elasticities) to change over time (e.g., due to automation)

Robustness: Alternative Measurement of Misoptimizations

Back

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

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



Calculate TFP using the structural method of Olley and Pakes (1996)

Extended Definition of Equilibrium Back

Let $\mathcal{Z}$ be the domain of (θ_i, X, w) and $f : \mathcal{Z} \to \Delta(\mathcal{Z})$ be each firm's conjectured transition density for this state variable. Let G_t be the productivity distribution at time t.

Definition: Equilibrium

An equilibrium is a stochastic choice rule $p \in \mathcal{P}$ and a transition density $f \in \mathcal{F}$ s.t.:

1. Intermediate goods firms' stochastic choice rules p solve their maximization program given f:

$$\max_{p \in \mathcal{P}} \int_{\mathcal{Z}} \int_{\mathcal{X}} \tilde{\Pi}(x, z_{it})\, p(x \mid z_{it})\, \mathrm{d}x\, f(z_{it} \mid z_{i,t-1})\, \mathrm{d}z - c\left(p, \lambda_i, z_{i,t-1}, f\right)$$

2. The transition density f is consistent with p in the sense that: (i) the marginal distribution of firm-level productivity is given by G; (ii) aggregate output is given by the aggregator evaluated in the cross-sectional distribution of production implied by p and G; and (iii) the wage is derived from the wage rule evaluated in aggregate output.

Strategic Mistakes vs. Mutual Information

Back

$$c^{MI}(p) = \underbrace{\int p(x|\theta) \log p(x|\theta) \, \mathrm{d}x \;\; f(\theta) \, \mathrm{d}\theta}_{\text{Entropy Term, or our } c(P)} - \underbrace{\int p(x) \log p(x) \, \mathrm{d}x}_{\text{Cross-State Interactions}}$$

Heuristic interpretation: "MI is entropy cost plus endogenous anchoring point"

In Appendix C of the paper,

- *Proposition 11*: entropy costs, and hence all our PE and GE results, are obtained in MI model if all agents believe that every production level x is *ex ante* equally likely
  - Intuition: "anchored to uninformative prior," like in Matějka and McKay (2015)
- *Proposition 12*: for any posterior separable cost (incl. MI), actions become more "precise" around the optimum when curvature of payoffs increases "locally" in state space
  - Intuition: agents respond to stakes for being precise

Dynamic Negative Effects of Misoptimization

Back 1 Back 2 Back 3

Outcome X is stock return, firm profitability, or TFP growth; and horizon j is varied

$$X_{i,t+j} = \beta_{j,X} \cdot \hat{u}_{it}^2 + \chi_{j(i),t} + \epsilon_{it}$$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Outcome: $\Delta \log P_{it}$			Outcome: π_{it}			Outcome: $\Delta \log \hat{\theta}_{it}$		
Horizon j	0	1	2	0	1	2	0	1	2
$\hat{u}_{it}^2$	-0.236	-0.252	-0.251	-0.316	-0.286	-0.265	-0.009	0.014	-0.007
	(0.026)	(0.027)	(0.038)	(0.024)	(0.018)	(0.019)	(0.007)	(0.008)	(0.010)
N	41,578	34,643	28,103	51,015	42,014	33,934	50,455	40,671	32,362
R^2	0.238	0.241	0.248	0.117	0.123	0.126	0.231	0.245	0.263

Note: Standard errors are double-clustered at the year and firm level. π_{it} = measured profitability of firm, or EBIT over lagged variable costs.

Robustness: Leverage Effect and Heterogeneous Betas

Back

$$\Delta \log P_{it} = \beta \cdot \hat{u}_{it}^2 + \phi \cdot \left(\hat{u}_{it}^2 \times \Delta \log \text{SP500}_t\right) + \tau \cdot Y_{it} + \eta \cdot \left(\hat{u}_{it}^2 \times Y_{it}\right) + \chi_{j(i),t} + \Gamma' X_{it} + \epsilon_{it}$$

	(1)	(2)	(3)	(4)	(5)	(6)
			Outcome: $\Delta \log P_{it}$			
$\hat{u}_{it}^2 \times \Delta \log P_t$	0.376	0.378	0.345	0.321	0.330	0.552
	(0.123)	(0.109)	(0.118)	(0.173)	(0.094)	(0.275)
Sector x Time FE	✓	✓	✓	✓	✓	✓
TFP and Interaction		✓				
Leverage and Interaction			✓			
Lag Return and Interaction				✓		
Firm FE and Interaction					✓	
Industry FE and Interaction						✓
N	41,578	41,578	41,429	34,805	41,578	41,206
R^2	0.239	0.261	0.246	0.239	0.240	0.379

Misoptimization Dispersion with Composition Adjustments

Back

Well-known issue: public firms skew younger, riskier over time

Davis, Haltiwanger, Jarmin, and Miranda (2006); Fama and French (2004); Brown and Kapadia (2007)

$$\text{MisoptDisp}_t = \alpha + \beta \cdot \frac{\text{Unemployment}_t}{100} + u_t$$

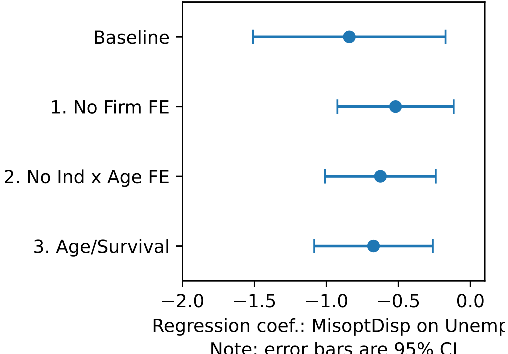
1. *Remove firm fixed effects*

$$\hat{u}_{it}^2 = \tau_i + \check{u}_{it}^2$$

2. *Remove industry-specific age trends*

$$\hat{u}_{it}^2 = \tau_{j(i),t-f(i)} + \check{u}_{it}^2$$

3. *Condition on survival* (for next 4 years) *and age* (older than 4 years)



Validation: Misoptimizations Hurt Performance

Are misoptimizations verifiably "bad" for firms, in both directions?

Validation: Misoptimizations Hurt Performance

Back 1 Back 2 Back 3

Are misoptimizations verifiably "bad" for firms, in both directions?

Binned scatter plots of

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

- X_{it} is stock return or firm profitability (= EBIT over lagged variable costs)
- $\chi_{j(i),t}$ are sector-by-time fixed effects

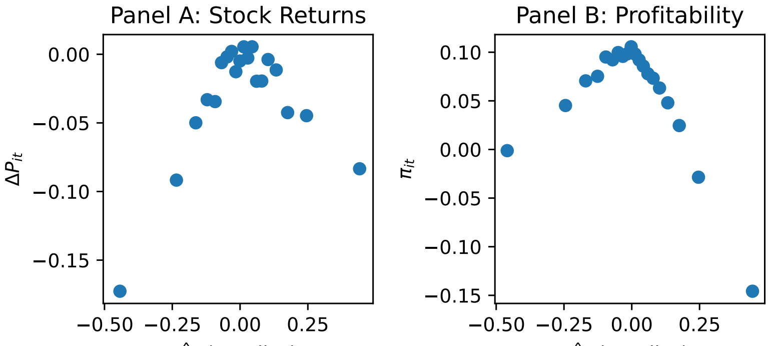
Validation: Misoptimizations Hurt Performance

Are misoptimizations verifiably "bad" for firms, in both directions?

Binned scatter plots of

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

- X_{it} is stock return or firm profitability (= EBIT over lagged variable costs)
- $\chi_{j(i),t}$ are sector-by-time fixed effects



Empirically Benchmarking Macro Predictions

Back

- **Asymmetric shock response**: model gets 25% of empirical benchmark
  - *Model*: 7% bigger output (5% bigger labor) response to negative vs. positive shock starting from steady state
  - *Data*: 20% bigger labor response to negative than positive productivity shock (Ilut, Kehrig, and Schneider, 2019: Table 9)

Empirically Benchmarking Macro Predictions

Back

- **Asymmetric shock response**: model gets 25% of empirical benchmark
  - *Model*: 7% bigger output (5% bigger labor) response to negative vs. positive shock starting from steady state
  - *Data*: 20% bigger labor response to negative than positive productivity shock (Ilut, Kehrig, and Schneider, 2019: Table 9)
- **Stochastic volatility**: model gets 20% of empirical benchmark
  - *Model*: 90th/10th percentile drop of productivity $\rightarrow$ 4.6% drop in GDP, 11% increase in conditional volatility of growth
  - *Data*: 57% increase of one-quarter ahead output (IP) uncertainty in Great Recession (Jurado, Ludvigson, and Ng, 2015: Figure 1)

Robustness: The Misoptimization Cycle With Different Measurement

Details

Back

$$\text{MisoptDisp}_t = \alpha + \beta \cdot \frac{\text{Unemployment}_t}{100} + u_t$$