From Micro to Macro in an Equilibrium Diffusion Model

Wyatt Brooks — Arizona State

Kevin Donovan — Yale SOM

Terence Johnson — Virginia

Motivation

Goal: Study cost of distortions that limit intra-firm learning opportunities

1. **Micro evidence suggest these frictions are important** (Atkin, et al, 2017; Brooks et al, 2018; Cai and Szeidl, 2018)
   - ▷ Randomly create new opportunities for firm-to-firm interaction
     $\Rightarrow$ higher profit, tech adoption, management practices

2. **. . . but likely incomplete accounting at scale**
   - ▷ If that learned ability diffuses to others (Alverez, et al., 2008; Perla and Tonetti, 2014; Buera and Lucas, 2018)

Motivation

Goal: Study cost of distortions that limit intra-firm learning opportunities

1. **Micro evidence suggest these frictions are important** (Atkin, et al, 2017; Brooks et al, 2018; Cai and Szeidl, 2018)
   - Randomly create new opportunities for firm-to-firm interaction
     $\Rightarrow$ higher profit, tech adoption, management practices
2. **. . . but likely incomplete accounting at scale**
   - If that learned ability diffuses to others (Alverez, et al., 2008; Perla and Tonetti, 2014; Buera and Lucas, 2018)
3. **Big picture: increase managerial skill at scale**
   - Infer aggregate implications from same micro evidence (World Bank, 2020)
   - Warranted? If not, how do we link the two?

What We Do

- **GE model of intra-firm learning + diffusion**
  - **Micro-foundation:** Interaction $\Rightarrow$ exchange "ideas" (skills, info, etc.)
  - **Link to aggregates with diffusion:** Distribution of ideas $\Rightarrow$ learning tomorrow, prices, ...
- **Question: What is the cost of distortions that limit interactions?**
  - **Problem:** depends on hard-to-measure elasticities (who/how often do I meet?)
- **Derive relationship between key model parameters and micro evidence**
  - Holds for broad class of recent experiments + diffusion models
  - Links micro evidence with models that motivate it

What We Find

- **Use to re-interpret smaller scale experiments that making learning easier**
  - Average treatment effect has no (direct) relation to at-scale gains
- **Highlight alternative covariance moment**
  - Better summarizes key model forces
  - Simple OLS interpretation (non-parametric extensions to more complicated settings)

What We Find

- **Use to re-interpret smaller scale experiments that making learning easier**
  - Average treatment effect has no (direct) relation to at-scale gains
- **Highlight alternative covariance moment**
  - Better summarizes key model forces
  - Simple OLS interpretation (non-parametric extensions to more complicated settings)
- **Quantify in specific Kenyan RCT** (Brooks, et al., 2018)
  - **Treatment:** random matches between high- and low-profit firm owners
  - **ATE** $= +19\%$
- **Continuum of economist deliver same ATE, but aggregate gains** $\in (0.6\%, 38\%)$
  - ATE + Covariance $\Rightarrow +11\%$

Outline

1. **Lay out (part of) diffusion model**
   - Highlight importance of various parameters
2. **Link parameters to promising micro evidence**
   - Highlight why standard empirical moments provide little help
3. **Quantify importance with RCT in Kenya**

Economic Environment

- **Discrete time, infinite horizon economy**
  - Measure one agents ("firms") with ability z
  - Each period, get two shocks:
    - Imitation shock $\hat{z}$, adopt if profitable
    - Random innovation in ability ε
  - ε uncorrelated with $z, \hat{z}$, not i.i.d.

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Economic Environment

- **Discrete time, infinite horizon economy**
  - Measure one agents ("firms") with ability z
  - Each period, get two shocks:
    - Imitation shock $\hat{z}$, adopt if profitable
    - Random innovation in ability ε
  - ε uncorrelated with $z, \hat{z}$, not i.i.d.

1. **How does z evolve?** Law of motion for ability $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Economic Environment

- **Discrete time, infinite horizon economy**
  - Measure one agents ("firms") with ability z
  - Each period, get two shocks:
    - Imitation shock $\hat{z}$, adopt if profitable
    - Random innovation in ability ε
  - ε uncorrelated with $z, \hat{z}$, not i.i.d.

1. **How does z evolve?** Law of motion for ability $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$
2. **What gets transmitted?** In equilibrium, profit $\pi \propto z$

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Economic Environment

- **Discrete time, infinite horizon economy**
  - Measure one agents ("firms") with ability z
  - Each period, get two shocks:
    - Imitation shock $\hat{z}$, adopt if profitable
    - Random innovation in ability ε
  - ε uncorrelated with $z, \hat{z}$, not i.i.d.

1. **How does z evolve?** Law of motion for ability $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$
2. **What gets transmitted?** In equilibrium, profit $\pi \propto z$
3. **Who interacts?** Equilibrium source distribution $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$, θ orders via FOSD

Special cases: Jovanovic and Rob (1989), Alvarez, et al. (2008), Lucas (2009), Lucas and Moll (2014), Perla and Tonetti (2014), Buera and Lucas (2018), Buera and Oberfield (2020)

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Measuring Cost of Distortions that Limit Interactions

- **Question: how large are benefits from better matching at scale?**
  - i.e., a permanent increase in θ

- **Many experiments do something like this in partial equilibrium . . .**
  - buyer/supplier links (Atkin, et al., 2017)
  - groups meetings of firm managers (Cai and Szeidl, 2018)
  - 1-1 meetings of high- and low-profit SMEs (Brooks et al., 2018)
  - 1-1 meeting with "role model" owner (Lafortune, et al., 2018)
  - business plan competition interactions (Fafchamps and Quinn, 2018)

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Measuring Cost of Distortions that Limit Interactions

- **Question: how large are benefits from better matching at scale?**
  - i.e., a permanent increase in θ

- **Many experiments do something like this in partial equilibrium ...**
  - buyer/supplier links (Atkin, et al., 2017)
  - groups meetings of firm managers (Cai and Szeidl, 2018)
  - 1-1 meetings of high- and low-profit SMEs (Brooks et al., 2018)
  - 1-1 meeting with "role model" owner (Lafortune, et al., 2018)
  - business plan competition interactions (Fafchamps and Quinn, 2018)

- **All of these are the same shock at this level of abstraction**
  - (Weakly) Better set of draws for treatment group.
  - **In model:** Replace $\widehat{M}$ with better, exogenous $\widehat{H}_T$

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Linking experiments with model parameters

$$\text{ATE}^{\text{data}} = \frac{\mathbb{E}[\pi'|i \in \mathbf{T}]}{\mathbb{E}[\pi'|i \in \mathbf{C}]} \qquad \text{ATE}^{\text{model}} = \frac{\int\int \pi^{\rho} \max\{1, \hat{\pi}/\pi\}^{\beta} \; d\widehat{H}_T(\hat{\pi}) \; dH(\pi)}{\int\int \pi^{\rho} \max\{1, \hat{\pi}/\pi\}^{\beta} \; d\widehat{M}(\hat{\pi}, \pi, \theta) \; dH(\pi)}$$

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Linking experiments with model parameters

$$\text{ATE}^{\text{data}} = \frac{\mathbb{E}[\pi' | i \in \mathbf{T}]}{\mathbb{E}[\pi' | i \in \mathbf{C}]} \qquad \text{ATE}^{\text{model}} = \frac{\int \int \pi^{\rho} \max\{1, \hat{\pi}/\pi\}^{\beta} \;\; d\widehat{H}_T(\hat{\pi}) \; dH(\pi)}{\int \int \pi^{\rho} \max\{1, \hat{\pi}/\pi\}^{\beta} \;\; d\widehat{M}(\hat{\pi}, \pi, \theta) \; dH(\pi)}$$

Proposition

For any (β, ρ), there is a unique θ that solves $ATE^{model} = ATE^{data}$*

(if ATE^{data} is within some computable range, otherwise no solution)*

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Linking experiments with model parameters

$$\text{ATE}^{\text{data}} = \frac{\mathbb{E}[\pi' | i \in \mathbf{T}]}{\mathbb{E}[\pi' | i \in \mathbf{C}]} \qquad \text{ATE}^{\text{model}} = \frac{\int\int \pi^{\rho} \max\{1, \hat{\pi}/\pi\}^{\beta} \; d\widehat{H}_T(\hat{\pi}) \; dH(\pi)}{\int\int \pi^{\rho} \max\{1, \hat{\pi}/\pi\}^{\beta} \; d\widehat{M}(\hat{\pi}, \pi, \theta) \; dH(\pi)}$$

Proposition

For any (β, ρ), there is a unique θ that solves $ATE^{model} = ATE^{data}$*

(if ATE^{data} is within some computable range, otherwise no solution)*

- **Observe small ATE^{data}. How to rationalize in model?**

#1	**Direct effect:**	(β, ρ) are low	No one learns from a good match
#2	**Extensive margin:**	θ is high	Everyone already meets smart firms

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Linking experiments with model parameters

$$\text{ATE}^{\text{data}} = \frac{\mathbb{E}[\pi'|i \in \mathbf{T}]}{\mathbb{E}[\pi'|i \in \mathbf{C}]} \qquad \text{ATE}^{\text{model}} = \frac{\int\int \pi^{\rho} \max\{1, \hat{\pi}/\pi\}^{\beta} \;\; d\widehat{H}_T(\hat{\pi}) \; dH(\pi)}{\int\int \pi^{\rho} \max\{1, \hat{\pi}/\pi\}^{\beta} \;\; d\widehat{M}(\hat{\pi}, \pi, \theta) \; dH(\pi)}$$

Proposition

For any (β, ρ), there is a unique θ that solves $ATE^{model} = ATE^{data}$*

(if ATE^{data} is within some computable range, otherwise no solution)*

- **Observe small ATE^{data}. How to rationalize in model?**

#1	**Direct effect:**	(β, ρ) are low	No one learns from a good match
#2	**Extensive margin:**	θ is high	Everyone already meets smart firms

- **(β, ρ, θ) combinations do not have same aggregate implications**
  - Need to pin down relative importance

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Measuring the direct effect (β, ρ)

Proposition

(β, ρ) *are identified by coefficients from the following regression run only on treated firms*

$$\log(\pi_i') = \tilde{c} + \rho \log(\pi_i) + \beta \log\left(\max\left\{1, \frac{\hat{\pi}_i}{\pi_i}\right\}\right) + \varepsilon_i$$

- **Interpretation:** impact of a better match, controlling for initial profit

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Measuring the direct effect (β, ρ)

Proposition

(β, ρ) *are identified by coefficients from the following regression run only on treated firms*

$$\log(\pi_i') = \tilde{c} + \rho \log(\pi_i) + \beta \log\left(\max\left\{1, \frac{\hat{\pi}_i}{\pi_i}\right\}\right) + \varepsilon_i$$

- **Interpretation:** impact of a better match, controlling for initial profit
- **(β, ρ) are unbiased:** random matching creates exploitable heterogeneity

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Measuring the direct effect (β, ρ)

Proposition

(β, ρ) *are identified by coefficients from the following regression run only on treated firms*

$$\log(\pi_i') = \tilde{c} + \rho \log(\pi_i) + \beta \log\left(\max\left\{1, \frac{\hat{\pi}_i}{\pi_i}\right\}\right) + \varepsilon_i$$

- **Interpretation:** impact of a better match, controlling for initial profit
- **(β, ρ) are unbiased:** random matching creates exploitable heterogeneity
- **Why this moment matters:** aggregates driven by mass in right tail

$$\text{If } \hat{\pi} > \pi \text{ for everyone} \ldots \hat{\beta} = \frac{cov\left(\log(\pi_i'), \log(\hat{\pi}_i)\right)}{\sigma^2_{\log(\hat{\pi}_i)}}$$

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

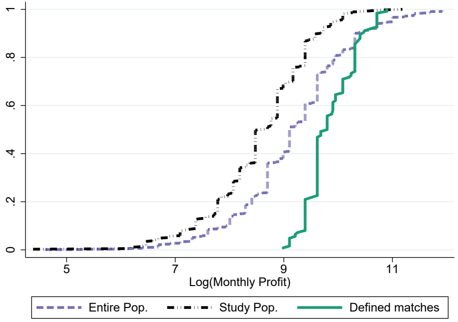
Outline

1. **Lay out (part of) diffusion model**
    - Highlight importance of various parameters
2. **Link parameters to promising micro evidence**
    - Highlight why standard empirical moments provide little help
3. **Quantify importance with RCT in Kenya**

RCT: Dandora, Kenya, 2014-2015

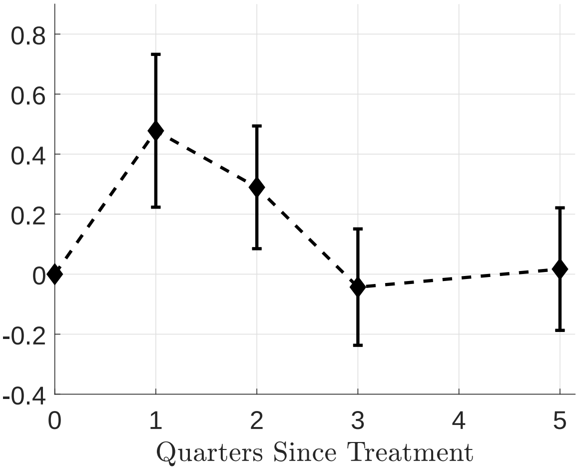
(Brooks et al., 2018)

- **Treatment: Random match to high profit business owner**
  - 2x as profitable, 10 years more experience
- **To more productive member of the match:**
  - Help less profitable firm learn about business
  - One meeting during November 2014
  - No topics, meeting length, cost to not meeting
- **To treatment firm: phone number of the match**
- **Track outcomes over 5 quarters**

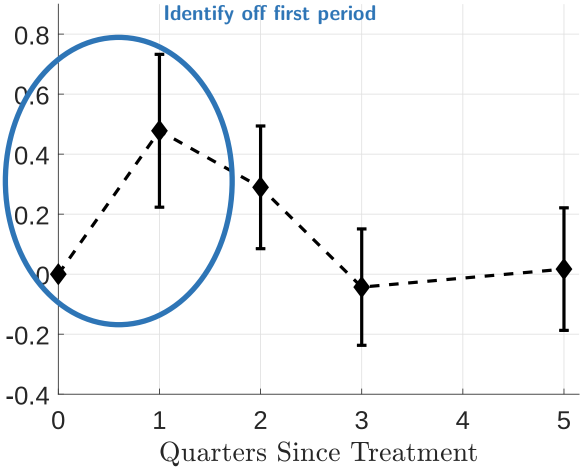


Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Time Series of Profit Average Treatment Effect



Time Series of Profit Average Treatment Effect



Underlying Channels

- **Key mechanism: primarily costs**
  - 33% more likely to switch suppliers
  - 45% in inventory costs
- **Massive supplier turnover: 2/3 of control firms switch suppliers**
- **Diffusion model seems reasonable, despite quick fade-out**
  1. Profit gains are surplus, not redistribution
  2. **Second RCT: after being in treatment, go mentor a control firm**
     - Original mentor profit strongly predicts treatment
     - Inconsistent with span-of-control theory

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Model Overview

- **Measure one of agents with heterogeneous ability z**
  - Aggregate state: $M(z)$
  - Die at rate δ, replaced with new agents who draw initial ability $z_0 \sim G(z)$
- **Occupational choice each period**
  - **Worker:** paid market clearing wage w
  - **Firm:** earn profit $\pi(z) = x^{\alpha} n^{\eta} - p_x x - wn$

Utility flow: $u = \omega \log(y) + (1-\omega)\log(1-s)$

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Model Overview

- **Measure one of agents with heterogeneous ability z**
  - Aggregate state: $M(z)$
  - Die at rate δ, replaced with new agents who draw initial ability $z_0 \sim G(z)$
- **Occupational choice each period**
  - **Worker:** paid market clearing wage w
  - **Firm:** earn profit $\pi(z) = x^\alpha n^\eta - p_x x - wn$ Utility flow: $u = \omega \log(y) + (1-\omega)\log(1-s)$
- **Firms required to find supplier by exerting effort s**
  - Continuum of suppliers with different marginal cost m: $\pi^s = (p_x - m)x$
  - Suppliers source from some outside entity, remove profit
  - Nash bargain over price p_x

$$p_x^* = argmax_{p_x}\ (\pi)^\nu (\pi^s)^{1-\nu}$$

[value functions]

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^\rho \max\{1, (\hat{z}/z)\}^\beta$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Diffusion in the Model

- **Ability z + effort s helps agents find a good supplier**

$$m = exp(-s)z^{\frac{\alpha+\eta-1}{\alpha}}$$

- **Can be diffused across agents**

$$z' = e^{c+\varepsilon} z^{\rho} \max\left\{1, \frac{\hat{z}}{z}\right\}^{\beta}$$

- **Learn from operating firms**

$$\hat{z} \sim \widehat{M}(\hat{z}; \theta, M) = M^f(\hat{z}; M)^{\frac{1}{1-\theta}}$$

[value functions]

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Stationary Equilibrium

- **Stationary equilibrium is:**
  - Value functions and decision rules
  - Bargaining outcomes
  - Distribution M^* is consistent with the decision rules and evolves according to

$$\Lambda(M(z')) = \delta G(z') + \int\int F(\log(z') - \rho\log(z) - \beta\log(\max\{1, \hat{z}/z\}) - c)\, d\widehat{M}(\hat{z}; M)\, dM(z)$$

  and $\Lambda(M^*(z')) = M^*(z)$

- **Model satisfies all the relevant assumptions to use previous results (details in paper)**

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Estimating Diffusion Parameters

- **Step 1: Use law of motion to estimate (β, ρ) within treated firms**

$$\begin{aligned} \log(\pi_i') &= \tilde{c} + \rho \log(\pi_i) + \beta \log(\max\{1, \hat{\pi}_i/\pi_i\}) + \varepsilon_i \\ &= \underset{(2.24)}{2.41} + \underset{(0.27)^{**}}{0.59} \log(\pi_i) + \underset{(0.24)^{**}}{0.54} \log(\max\{1, \hat{\pi}_i/\pi_i\}) \end{aligned}$$

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Estimating Diffusion Parameters

- **Step 1: Use law of motion to estimate (β, ρ) within treated firms**

$$
\begin{aligned}
\log(\pi_i') &= \tilde{c} + \rho \log(\pi_i) + \beta \log(\max\{1, \hat{\pi}_i/\pi_i\}) + \varepsilon_i \\
&= \underset{(2.24)}{2.41} + \underset{(0.27)^{**}}{0.59} \log(\pi_i) + \underset{(0.24)^{**}}{0.54} \log(\max\{1, \hat{\pi}_i/\pi_i\})
\end{aligned}
$$

- **Step 2: Estimate extensive margin θ from ATE, under assumption $\widehat{M}(\hat{z}, z, \theta) = M^f(\hat{z})^{\frac{1}{1-\theta}}$**

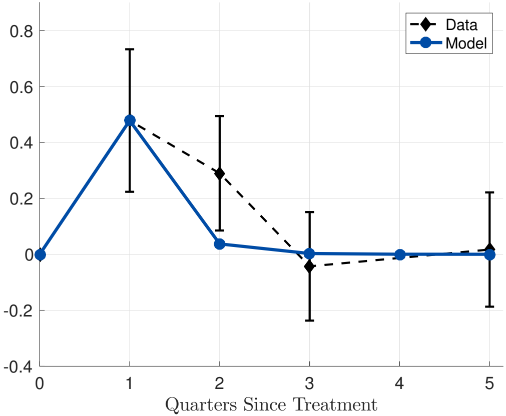
$$
\min_{\theta} \; abs\left(\frac{\mathbb{E}[\pi_T']}{\mathbb{E}[\pi_C']} - \frac{\int\int \pi^\rho \max\{1, \hat{\pi}/\pi\}^\beta \, d\widehat{H}_T(\hat{\pi}) dH_T(\pi)}{\int\int \pi^\rho \max\{1, \hat{\pi}/\pi\}^\beta \, d\widehat{M}(\hat{\pi}, \pi, \theta) dH_C(\pi)} \right) \quad \Rightarrow \quad \theta = -0.41
$$

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^\rho \max\{1, (\hat{z}/z)\}^\beta$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Implied RCT Dynamics

Replicate RCT in the model

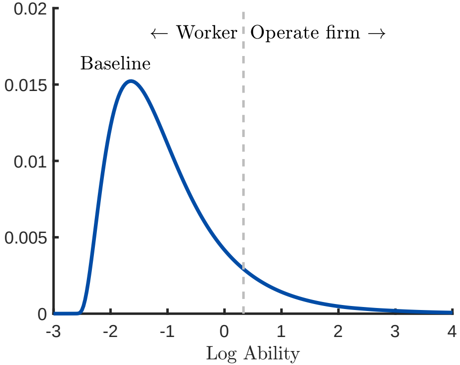
- Replicate empirical profit distribution, matches
- Trace impulse response
- Hold fixed distribution $M^*(z)$



Aggregate Implications

- **Aggregate experiment: permanently increase θ by 25%**
  - **Main channel:** learning shifts ability dist (64% of total)
  - **Amplification:** wage $\nearrow$ causes marginal firms to exit (36%)

	(1) Fixed Wage	(2) Total
Income	1.07	1.11
Ability	1.08	1.12
Labor Supply	0.92	0.98
Wage	1.00	1.13

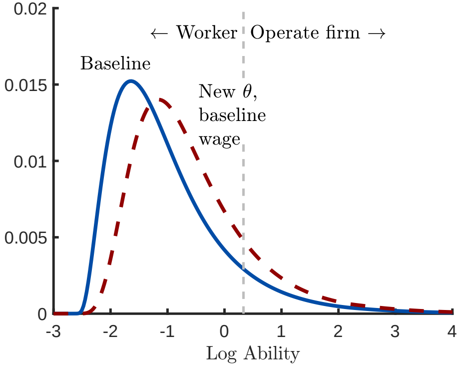


Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Aggregate Implications

- **Aggregate experiment: permanently increase θ by 25%**
  - **Main channel:** learning shifts ability dist (64% of total)
  - **Amplification:** wage $\nearrow$ causes marginal firms to exit (36%)

	(1) Fixed Wage	(2) Total
Income	1.07	1.11
Ability	1.08	1.12
Labor Supply	0.92	0.98
Wage	1.00	1.13

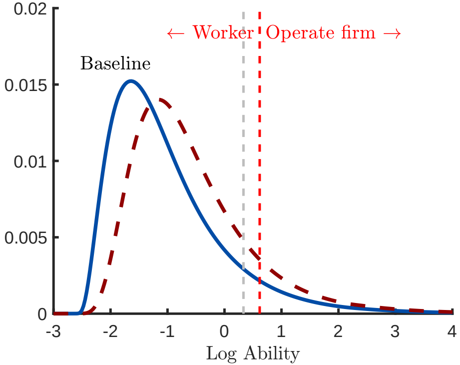


Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Aggregate Implications

- **Aggregate experiment: permanently increase θ by 25%**
  - **Main channel:** learning shifts ability dist (64% of total)
  - **Amplification:** wage $\nearrow$ causes marginal firms to exit (36%)

	(1) Fixed Wage	(2) Total
Income	1.07	1.11
Ability	1.08	1.12
Labor Supply	0.92	0.98
Wage	1.00	1.13

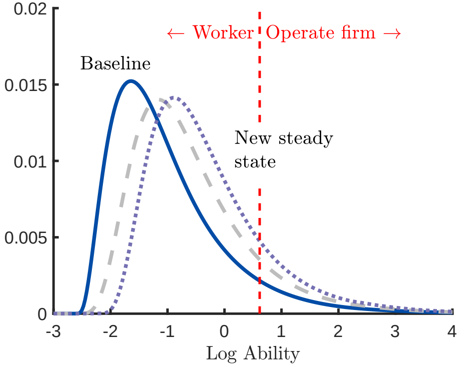


Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Aggregate Implications

- **Aggregate experiment: permanently increase θ by 25%**
  - **Main channel:** learning shifts ability dist (64% of total)
  - **Amplification:** wage $\nearrow$ causes marginal firms to exit (36%)

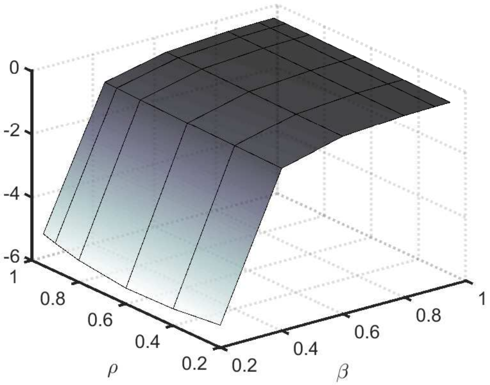
	(1) Fixed Wage	(2) Total
Income	1.07	1.11
Ability	1.08	1.12
Labor Supply	0.92	0.98
Wage	1.00	1.13



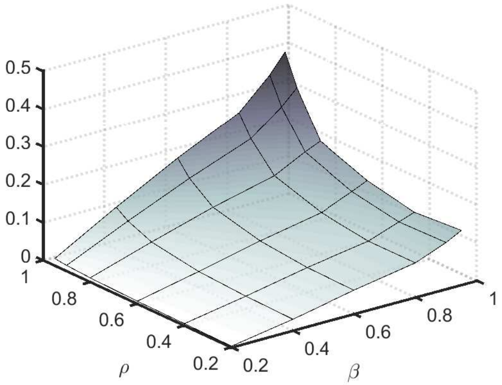
Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Getting diffusion parameters right is critical for at-scale gains . . .

- **For each (β, ρ) re-estimate θ to match same 1-period ATE**
  - At-scale gains vary between 0.6% and 38%



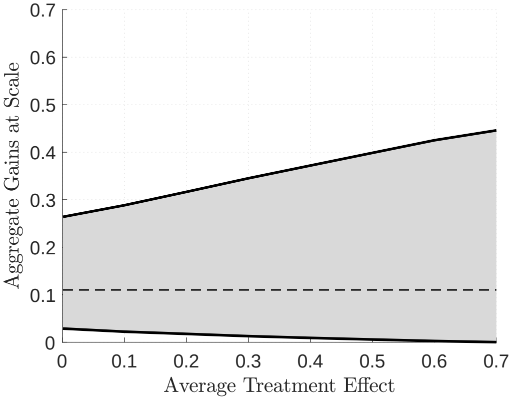
Implied θ



Aggregate Gains

. . . and for extracting policy recommendations

- **Replicate same exercise for every ATE**
- **Perverse policy recommendations**
  - Easy to come up with negative relationship
- **Implications:**
  - Maximize covariance given ATE
  - Locally, covariance is better predictor



[More details on result] [ATE Dynamics] [add mis-measurement] [Cai and Szeidl (2018)]

Conclusion

- **Relationship between reduced-form RCT results and at-scale gains from learning**
  - Many parameter combinations for the same treatment effect (intensive/extensive margins)
  - New covariance moment can help disentangle
- **Kenya implementation: cost of not doing so can be large**
  - At-scale gains vary $(0.6\%, 38\%)$ for same ATE
  - **Why:** covariance closely connected to key aggregate channel
- **Leaves open a number of issues:**
  - Increased competition may limit sharing skills
  - Method assumes matching function

Appendix Slides

- Different matching processes under $\widehat{M}(\hat{z}, z, \theta)$
- Extensions of identification procedure
- Value Functions
- Quantitative investigation of parameter importance
- Dynamics of the treatment effect
- Quantitative evaluation of mis-measurement
- Same procedure in Cai and Szeidl (2018)

Different Matching Processes

- `Adding exogenous innovations/noise`
- `Effort and bargaining`

Noise in Diffusion Process

- **Ability z receives idea that has two components: $\hat{z} = \gamma^{1/\theta} z_m$**
  - Random match z_m from another agent
  - Random exogenous innovation on that idea γ
- **Distribution of draws:**

$$
\begin{aligned}
\widehat{M}(c) &= Prob(\hat{z} \leq c) \\
&= Prob(z_m \leq c\gamma^{-1/\theta}) \\
&= \int M(c\gamma^{-1/\theta}) d\Gamma(\gamma)
\end{aligned}
$$

[back to matching options] [back to appendix] [back to model]

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Effort Choice and Bargaining

- **Ability z receives match z_m**
  - Exert effort x implies draw $\hat{z} = z_m^x z^{1-x}$
  - Match m gets benefit $b(x)$
- **Nash bargain over effort, bargaining weight θ:** $\max_{x\in[0,1]} \left(z_m^x z^{1-x}\right)^{\theta} b(x)^{1-\theta}$
  - Realized draw $\hat{z} = \max\left\{z, z_m e^{1-1/\theta}\right\}$
- **Distribution of draws:**

$$\begin{aligned}\widehat{M}(c) &= Prob(\hat{z} \le c) = Prob(z_m e^{1-1/\theta} \le c) \\ &= Prob(z_m \le c e^{1/\theta-1}) \\ &= M(c e^{1/\theta-1})\end{aligned}$$

[back to matching options] [back to appendix] [back to model]

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Extensions of Identification Procedure

- Semi-parametric identification
- Relationship between observables and z
- Mis-measurement
- Additional characteristics

Semi-Parametric Identification

- **Estimate (ρ, f) in**

$$\log(z') = c + \rho \log(z) + f\left(\frac{\hat{z}}{z}\right) + \varepsilon$$

- **Follows directly from assumptions + literature on non-linear error-in-variable regressions**
  - Yatchew (1997), Härdle et al. (2000)

1. **Estimate ρ**
   - Order data $\hat{\pi}_1/\pi_1 < \hat{\pi}_2/\pi_2 \ldots < \hat{\pi}_N/\pi_N$
   - Difference out the nonlinear f in limit as gap between i, $i+1$ goes to zero
2. **Now estimate f non-parametrically given remaining variation**

[back to id extensions] [back to appendix] [back to id procedures]

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

General relationship between observables and z

- **There exists a known function $g(\mathbf{x}) = Cz$ for some potentially unknown constant C and observables $\mathbf{x}$.**
  - Estimate production function using data from control group, $g(y, n, k) = Cz$
- **Similar regression with a data transform**

$$\log(g(\mathbf{x}')) = c + \rho \log(g(\mathbf{x})) + \beta \log\left(\max\left\{1, \frac{g(\widehat{\mathbf{x}})}{g(\mathbf{x})}\right\}\right) + \varepsilon,$$

- **Key: not $\pi \propto z$, but any set of observables and z**

[back to id extensions] [back to appendix] [back to id procedures]

Mis-measurement

- **Will do so with more general law of motion**

$$log(\pi') = \sum_{j=1}^{M} \beta_j g_j(\vec{\pi}) + \varepsilon$$

Assumption

We observe two outcomes that are mis-measured versions of the true value, $\vec{\pi}^* = (\pi^*, \hat{\pi}^*)$,

$$\begin{aligned} \vec{\pi}_{1i}^k &= \vec{\pi}_i^{*k} + \nu_{1i}^k, \quad k = 1, 2 \\ \vec{\pi}_{2i}^k &= \vec{\pi}_i^{*k} + \nu_{2i}^k, \quad k = 1, 2 \end{aligned}$$

We assume the following relationships between the measurement error and true values:

$$\mathbb{E}[\nu_1^k | \pi^{*k}, \nu_2^k] = 0, \quad k = 1, 2$$

$$\nu_2^k \text{ is independent from } \vec{\pi}^*, \nu_2^{-k}, \text{ where } -k \neq k$$

Basic Idea

- **Basic idea from Kotlarki's lemma, in $\mathbb{R}^1$**

$$\phi_{\pi^*}(t) = \exp\left(\int_0^t \frac{\mathbb{E}[i\pi_1 e^{it\pi_2}]}{\mathbb{E}[e^{it\pi_2}]}\right)$$

- **Inverse Fourier transform gives distribution of true values $f(\pi^*)$**
- **Our model is $\vec{\pi}^* = (\pi^*, \hat{\pi}^*) \in \mathbb{R}^2$. Schennach (2004):**

Proposition

If $\mathbb{E}[|\vec{\pi}^k|]$ and $\mathbb{E}[|\eta_1^k|]$ are finite, then there exists a closed form for any function $\mathbb{E}[u(\vec{\pi}^*, \beta)]$ whenever it exists.

[back to id extensions]   [back to appendix]   [back to id procedures]

Additional Characteristics

- **Let β depend on own and match characteristics, $\mathbf{x}$ and $\widehat{\mathbf{x}}$**
- **Law of motion:**

$$\log(\pi_i') = c + \rho \log(\pi_i) + \beta(\mathbf{x}, \widehat{\mathbf{x}}) \log\left(\max\left\{1, \frac{\hat{\pi}_i}{\pi_i}\right\}\right)$$

Bin the characteristics in some way: $$\log(\pi_{ib}') = c + \rho \log(\pi_{ib}) + \sum_{b=1}^{B} \beta_b \log\left(\max\left\{1, \frac{\hat{\pi}_{ib}}{\pi_{ib}}\right\}\right)$$

- **Identifies $(\rho, \beta_1, \ldots, \beta_B)$ with A3 adjustment**

$$\widehat{M}(\hat{z}, b; z, \theta) = \widehat{M}_b(\hat{z}; z, \theta)\Gamma_b$$

[back to id extensions] [back to appendix] [back to id procedures]

Individual's Problem

Static problem for each occupation

$$u^f(z,M) = \max_{s,x,n\geq 0} \; \omega \log(\pi) + (1-\omega)\log(1-s) \qquad u^w(z,M) = \omega \log(w)$$

$$\begin{aligned} s.t. \quad & \pi = x^{\alpha} n^{\eta} - p_x x - wn \\ & m = f(s,z) \\ & p_x = \text{argmax}_{p_x} \; [\pi]^{\nu}[\pi^s(m)]^{1-\nu} \end{aligned}$$

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Individual's Problem

Static problem for each occupation

$$u^f(z,M) = \max_{s,x,n\geq 0} \; \omega \log(\pi) + (1-\omega)\log(1-s) \qquad u^w(z,M) = \omega \log(w)$$

$$\begin{aligned} s.t. \quad & \pi = x^\alpha n^\eta - p_x x - wn \\ & m = f(s,z) \\ & p_x = \text{argmax}_{p_x} \; [\pi]^\nu [\pi^s(m)]^{1-\nu} \end{aligned}$$

$\Big\downarrow$ **+ diffusion**

$$v(z,M) = \max\{u^f(z,M), u^w(z,M)\} + (1-\delta)\int_\varepsilon \int_{\hat{z}} v(z'(\hat{z},\varepsilon;z), M')\widehat{M}(d\hat{z},M)dF(\varepsilon)$$

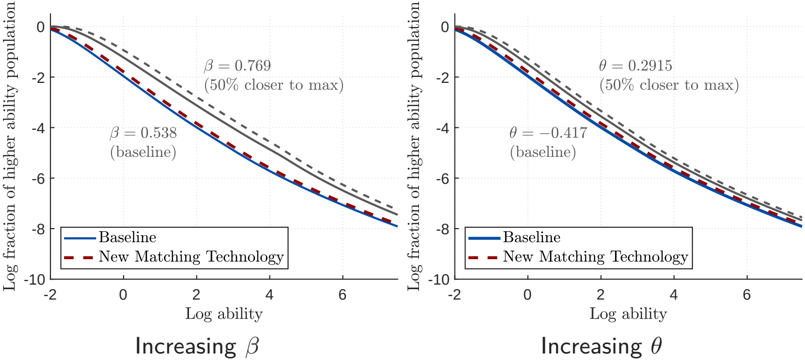
$$s.t. \quad z'(\hat{z},\varepsilon;z) = e^{c+\varepsilon} z^\rho \max\left\{1, \frac{\hat{z}}{z}\right\}^\beta$$

[back]

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^\rho \max\{1, (\hat{z}/z)\}^\beta$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

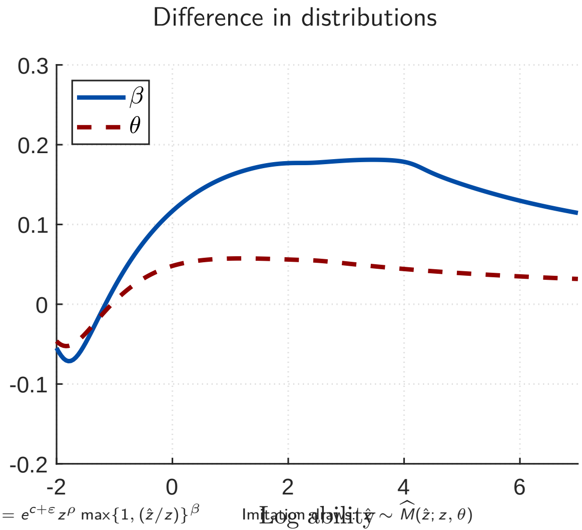
The Importance of Heterogeneity

Fraction with ability above z, $\log(1 - M(z))$



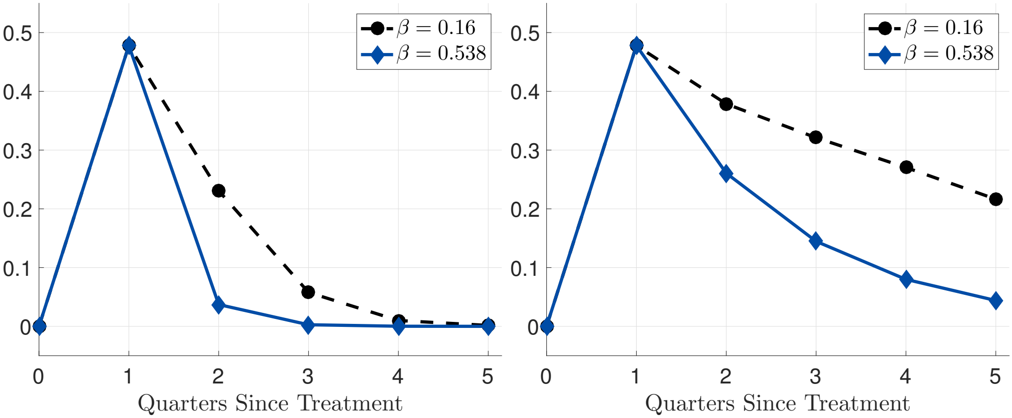
Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

The Importance of Heterogeneity



Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

The Dynamics of the ATE



At Estimated Persistence $\rho = 0.595$ | At High Persistence $\rho = 0.99$

Profit: $\pi \propto z$ [back to results] Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Quantitative Evaluation of Mis-Measurement

- **We observe $\pi = \tau\pi^*$, where $\tau \sim N(0, \sigma_\tau)$ is classical measurement error**
  - $\tau \sim N(0, \sigma_\tau)$, where σ_τ is known but the individual realizations are not
  - Can extend to unknown σ_τ with more machinery
- **Need a little notation for simplicity**
  - $\tilde{x} = \log(x)$
  - $f_x(x)$ is pdf
  - $\phi_x(t) = \int_{\mathbb{R}} e^{itx} f_x(x) dx$ as its characteristic function.

Quantitative Evaluation of Mis-Measurement

- **Estimate characteristic functions of the observed π and $\hat{\pi}$**

$$\hat{\phi}_{\tilde{\pi}}(t) = \left(\frac{1}{n} \sum_{j=1}^{n} e^{it \log(\pi_j)} \right) \phi_{k,\pi}(h_\pi t) \qquad \hat{\phi}_{\tilde{\hat{\pi}}}(t) = \left(\frac{1}{n} \sum_{j=1}^{n} e^{it \log(\hat{\pi}_j)} \right) \phi_{k,\hat{\pi}}(h_{\hat{\pi}} t)$$

- **This gives us true characteristic functions:** $\phi_{\tilde{\pi}*}(t) = \hat{\phi}_{\tilde{\pi}}(t) / \phi_{\tilde{\tau}}(t)$

- **Then recover densities from inverse Fourier transform**

$$f_{\pi*}(\pi^*) = \frac{1}{2\pi} \int \hat{\phi}_{\pi*}(t) e^{-it\pi^*} dt \qquad f_{\hat{\pi}*}(\hat{\pi}^*) = \frac{1}{2\pi} \int \hat{\phi}_{\hat{\pi}*}(t) e^{-it\hat{\pi}^*} dt$$

Quantitative Evaluation of Mis-Measurement

- **Minimum distance estimator to estimate**
- **Choose $\Gamma = (c, \rho, \beta)$ to solve**

$$\min_{\Gamma} \sum_{i=1}^{n} \left(\pi_i' - G(\pi_i, \hat{\pi}_i; \Gamma)\right)^2$$

where

$$G(\pi, \hat{\pi}; \Gamma) = \int \int g(\pi^*, \hat{\pi}^*) f_{\pi^*|\pi}(\pi^*|\pi, \Gamma) f_{\hat{\pi}^*|\hat{\pi}}(\hat{\pi}^*|\hat{\pi}, \Gamma) d\pi^* d\hat{\pi}^*$$

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Aggregate Gains

Table: Equilibrium Moments

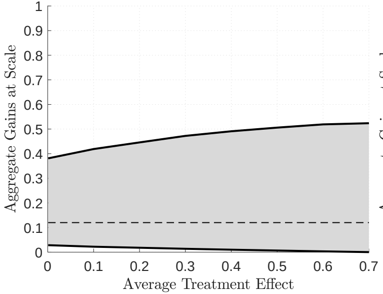
	$\sigma_\tau = 0.3$		$\sigma_\tau = 1$	
	(1) Fixed Wage	(2) At-Scale	(3) Fixed Wage	(4) At-Scale
Income	1.08	1.12	1.20	1.38
Ability	1.08	1.14	1.20	1.42
Aggregate Labor Supply	0.92	0.99	0.89	1.00
Wage	1.00	1.14	1.00	1.39

Table notes: All are measured relative to the baseline equilibrium at the give value of σ_τ. Each column reports the new stationary equilibrium after shocking the matching technology, where the first (columns 1 and 3) holds the wage fixed at its baseline level and the second allows it to adjust.

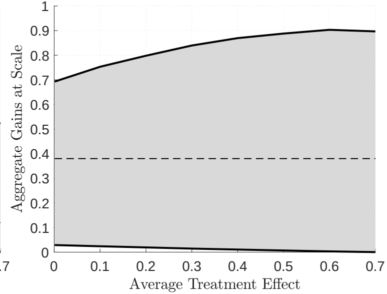
Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Range of Aggregate Gains

Figure: Range of Aggregate Gains for each ATE



$\sigma_\tau = 0.3$



$\sigma_\tau = 1$

[back to results]

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^\rho \max\{1, (\hat{z}/z)\}^\beta$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Cai and Szeidl (2018): Basics

- **RCT among 2,820 Chinese firms**
  - Treated firms (1,500 of 2,820) are randomly placed into a group of approximately 10 other firms
- **Different economy than our Kenya example**
  - Group meetings, instead of individual
  - More intense: monthly for one year
  - Larger firms: average size is 35 workers
  - Cross-randomize info about new financial products
- **Survey waves:**
  - Pre-treatment
  - 1 year later (end of treatment period)
  - 2 years later (1 year post-treatment)

[back to appendix] [back to results]

Cai and Szeidl (2018): Results Overview

- **Large and persistent effects on sales, profit, employment, productivity, management practices**
  - Hold in both post-treatment waves
- **Information about financial products flows within groups**
- **Group-level scale predicts treatment effect**
  - Firms in groups with larger average size see larger treatment effect
  - "Internal consistency check" of results

[back to appendix] [back to results]

Cai and Szeidl (2018): Model

- **Representative household:**

$$\max_{\{C_t, K_{t+1}\}\geq 0} \sum_{t=0}^{\infty} (1-\delta)^t u(C_t)$$

$$s.t. \quad C_t + K_{t+1} - (1-\lambda)K_t = w_t + r_t K_t + \Pi_t$$

$$K_0 \text{ given}$$

- **Firm profit:** $z^{1-\alpha-\eta} n^{\alpha} k^{\eta} - wn - rk$
  - Productivity evolves $z_{t+1} = e^{c+\varepsilon_t} z_t^{\rho} \left(1 + \frac{\hat{z}_t}{z_t}\right)^{\beta}$

[back to appendix] [back to results]

Cai and Szeidl (2018): Diffusion

- **Pr $= \theta$, join random group of exogenous size N**
- **Potential gains from match depends on average productivity of group**

$$\hat{z} = \sum_{i=1}^{N} \hat{z}_i / N$$

- **Distribution of draws:**

$$\widehat{M}(\hat{z}) = 1 - \theta + \theta Q(\hat{z}),$$

[back to appendix] [back to results]

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Cai and Szeidl (2018): Estimating Diffusion Parameters

- **Focus on firm size, given available results**
  - Estimate off $t = 0, 1$ data, check if we can match $t = 2$
- **First step within treatment group:**

$$\log(n_i') = c + \rho \log(n_i) + \beta \log\left(1 + \frac{\hat{n}_i}{n_i}\right) + \varepsilon \qquad \text{for all } i \text{ in treatment}$$

- **Then estimate θ**

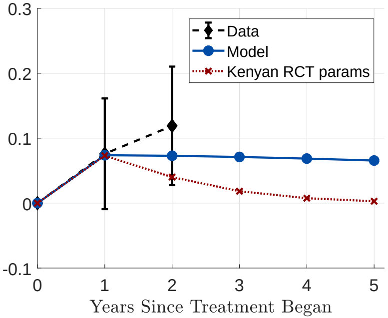
$$\min_{\theta} \quad abs\left(\frac{\mathbb{E}[n_T']}{\mathbb{E}[n_C']} - \frac{\int\int \pi^{\rho}\left(1 + \hat{n}/n\right)^{\beta} d\widehat{H}_T(\hat{n}) dH_T(n)}{\int\int n^{\rho}\left(1 + \hat{n}/n\right)^{\beta} d\widehat{M}(\hat{n};\theta) dH_C(n)}\right)$$

[back to appendix] [back to results]

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Cai and Szeidl (2018): Persistence of ATE

Figure: Dynamics of Average Treatment Effect (Firm Size)



[back to appendix] [back to results]

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$

Cai and Szeidl (2018): Range of At-Scale Gains

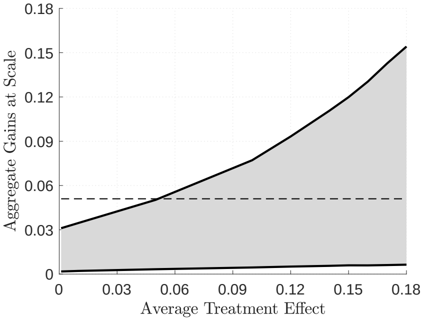


Figure: Range of Aggregate Gains for each ATE (Firm Size)

[back to appendix] [back to results]

Profit: $\pi \propto z$ Ability: $z' = e^{c+\varepsilon} z^{\rho} \max\{1, (\hat{z}/z)\}^{\beta}$ Imitation draws: $\hat{z} \sim \widehat{M}(\hat{z}; z, \theta)$