Measuring Productivity:

Lessons from Researcher-Designed Surveys & Productivity Benchmarking

David Atkin, MIT
Amit Khandelwal, Columbia GSB
Adam Osman, UIUC

Motivation

- Large differences in productivity between developed and developing world (Hall & Jones 1999, Bloom & Van Reenan 2007)

- Motivates some obvious questions:
  - How large are productivity differences across firms?
  - What drives these productivity differences?
  - Which policies could raise productivity and potentially reduce dispersion?

- Central challenge: *productivity is never observed!*
  - Detailed data rarely available, many measurement issues
  - Hundreds of different products often reside within a single administrative code

What We Do

- We focus on a specific industry—flat-weave rugs—and design surveys to compare alternative measures of productivity

- Collect survey data that allow us to calculate:
  1. **TFPQ**: Quantity productivity
     - ability to produce quantity with a given set of inputs

  2. **TFPZ**: Quality productivity
     - ability to produce quality with a given set of inputs

  3. **TFPC**: Capability
     - combine TFPQ & TFPZ using consumers' quality-quantity tradeoff to estimate firm's overall capability

  4. **TFPR**: Revenue productivity

- Measure "**true**" productivity in a lab benchmarking setting

Main Takeways

1. Normal estimates overstate the amount of productivity dispersion across firms because they fail to account for variation in product specifications

2. Standard TFPQ isn't great:
   - Weakly correlated with true productivity
   - Inversely correlated with quality productivity (TFPZ)!

3. But adjusting for specifications provides a measure that strongly correlates with true productivity, TFPZ

4. Large variation in firm capabilities (TPFC)
   - TFPC requires production data, product specifications and quality data
   - In the absence of such detailed data, TFPR may be a reasonable proxy for capability

Setting - Handmade Rugs in Egypt

- Working in Fowa, Egypt
  - About 2hrs from Alexandria
  - Well-known carpet cluster

- Conducted recruitment drive for firms in Fowa in early 2011

- Found 219 firms who:
  - Worked on own account (e.g., bought own inputs)
  - Less than 5 employees
  - Same sample as Atkin et al (2017)
  - Simple technology allows us to collect detailed data on every aspect of production

1- Unadjusted TFPQ: Quantity Productivity

Let's start by discussing & comparing 3 different ways to estimate TFPQ

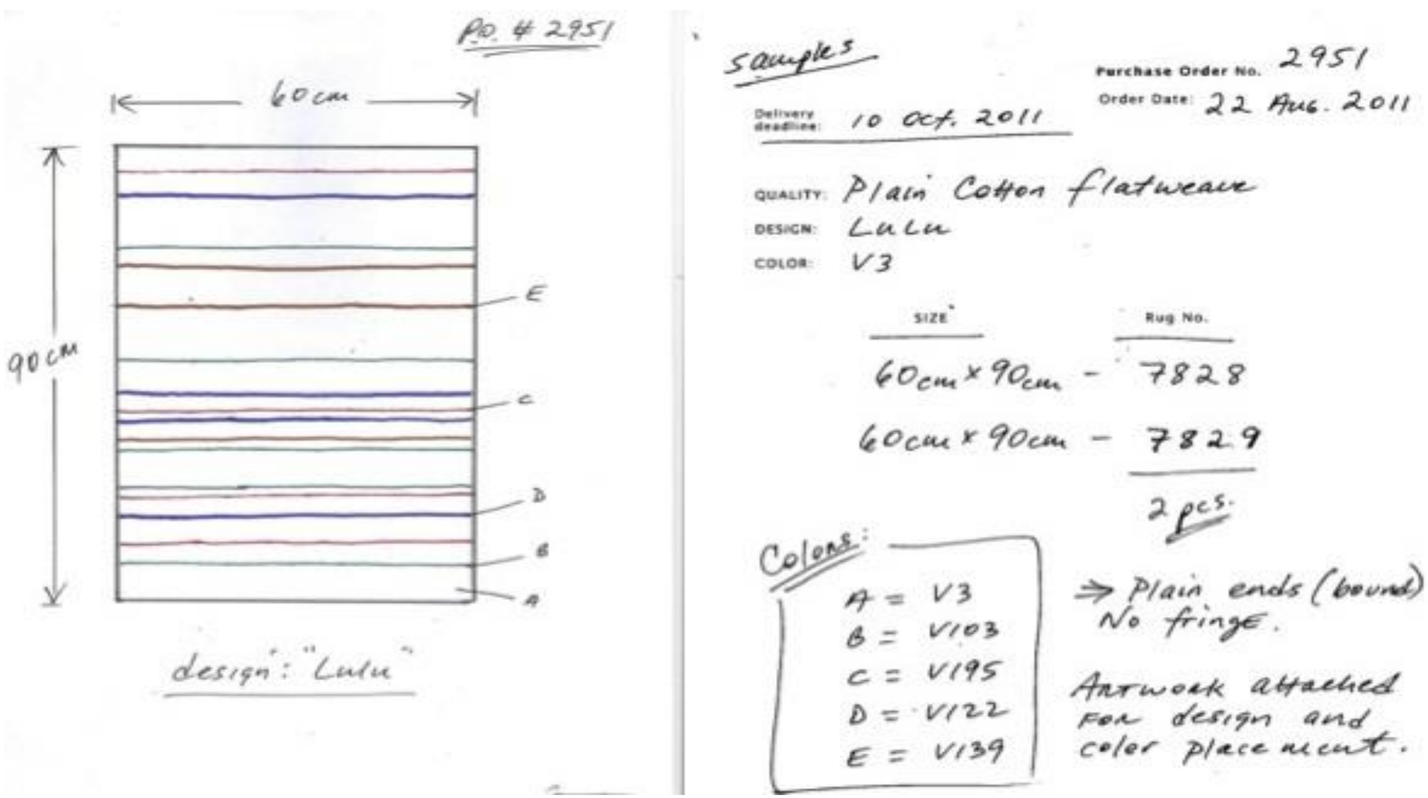
We can estimate TFPQ using Cobb-Douglas production function:

$$x = \phi_u l^{\alpha_l} k^{\alpha_k} e^{\epsilon}$$

- Output (*x*): number of square meters of rugs produced that month
- Labor (*l*): number of man-hours used that month
- Looms (*k*): number of looms used that month
- Estimate via OLS (similar results using Olley Pakes control function)
- **Unadjusted TFPQ** (ϕ_u)

Adjusting TFPQ: Controlling for Specifications

- What is a **specification**?
  - Codifiable attributes of the rug that are typically chosen by the buyer
  - E.g. design, thread type, thread count, colors, rug subcategory
  - All rugs fall within a a single HS10 code
  - We observe 435 combinations of specifications



2- Specification-Adjusted TFPQ

Our Original Cobb-Douglas production function: $x = \phi_u l^{\alpha_l} k^{\alpha_k} e^{\epsilon}$

- Output (x), Labor hours (l), Looms (k)

But products have different specifications which will affect productivity:

- e.g. a high thread count rug requires more inputs

Adjusting TFPQ for product specifications

$$\phi_u = \phi_a e^{\lambda \gamma}$$

- Where unit input requirements vary with vector of specifications (λ)

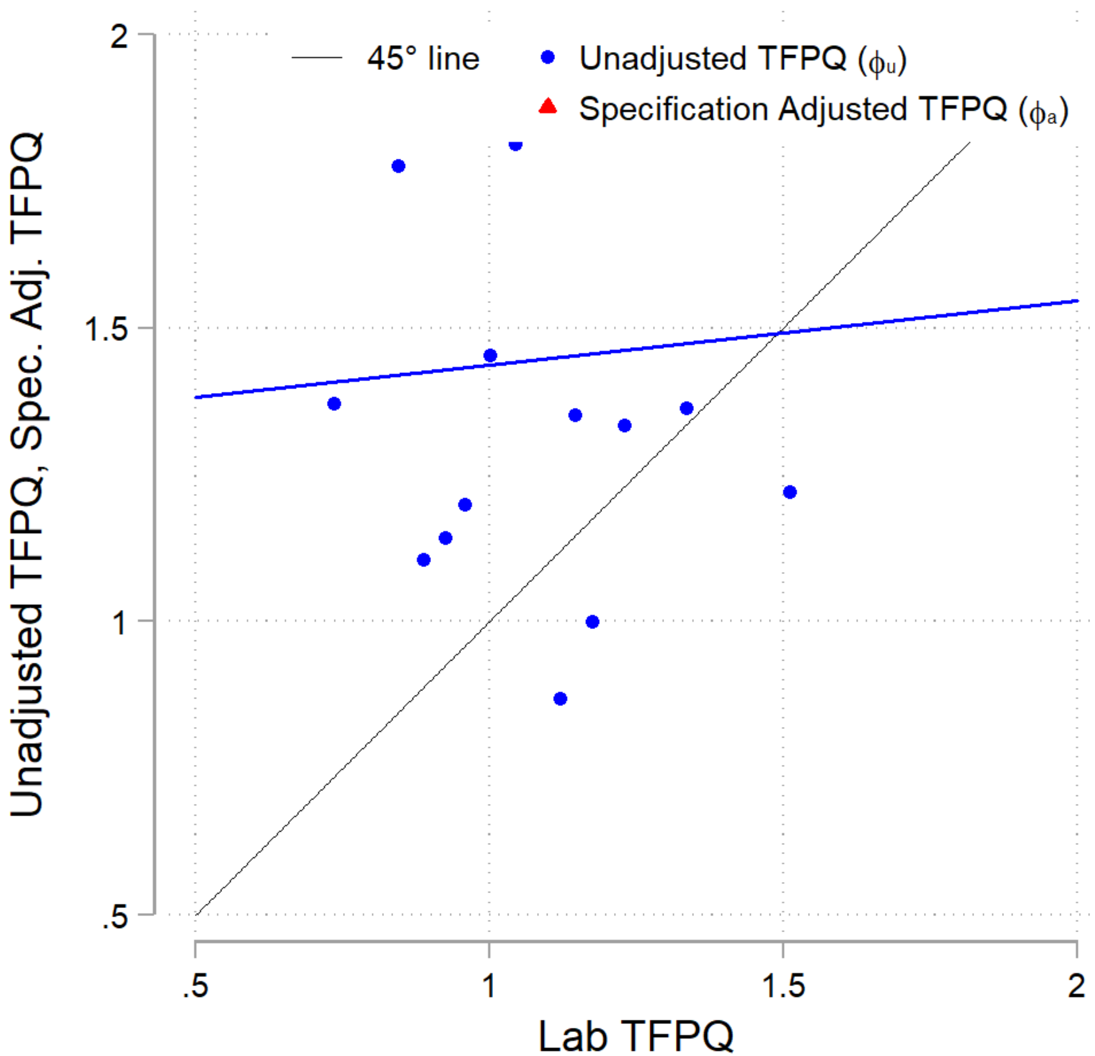
- **Specification-Adjusted TFPQ** (ϕ_a)

3- Laboratory TFPQ

Set up a controlled laboratory:

- Firms paid a flat fee to produce an identical rug
  - domestic design, 140cm by 70cm, should be 1750g
- Provided all firms with identical material inputs and loom in rented facility
- Recorded dimensions, weight, time to produce the rug
  - **Lab TFPQ**: Direct measure of quantity productivity: m^2 per labor hour
- "Benchmark" since inputs, product specifications identical across firms

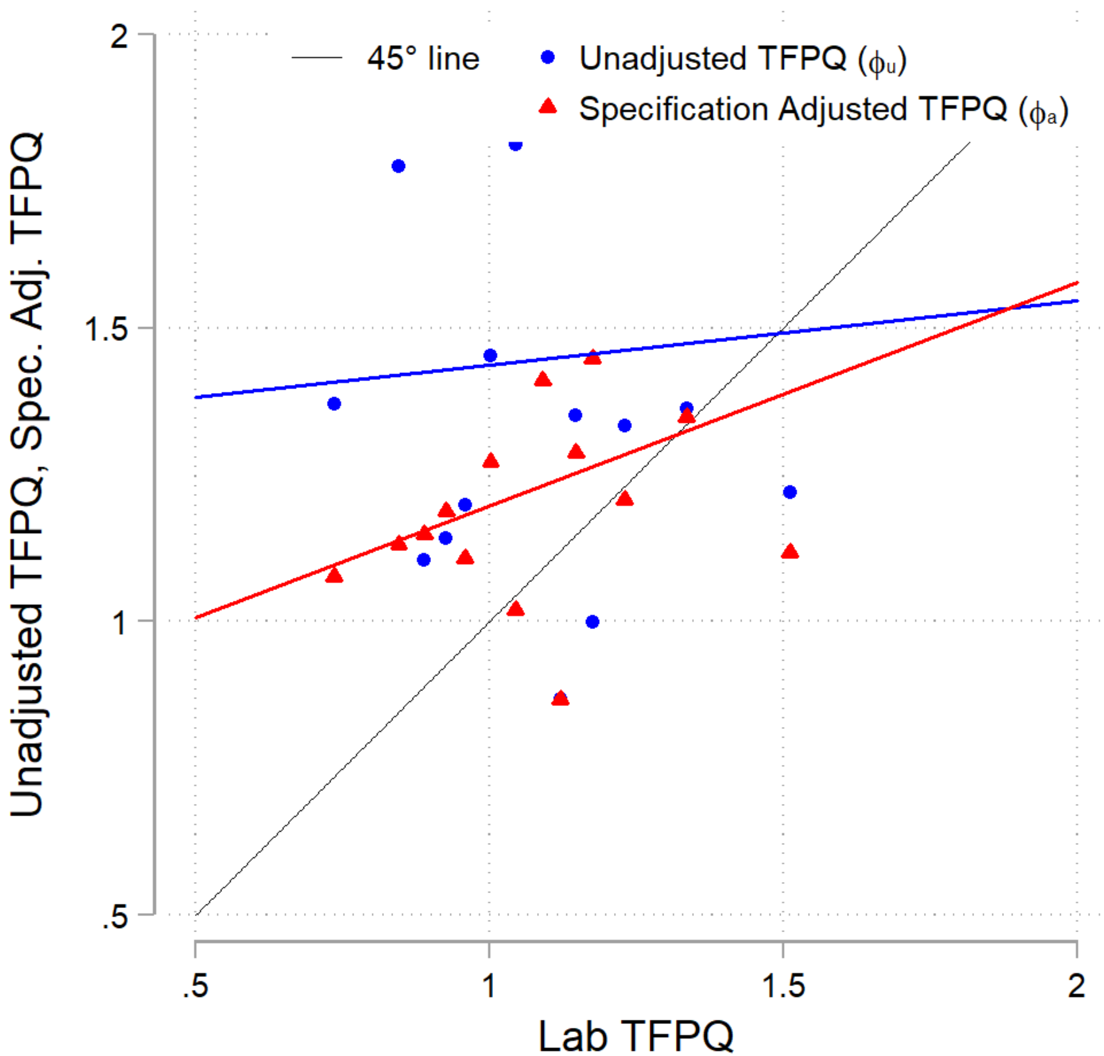
Result 1: Adjusting for Product Specifications is Important



ϕ_u: β=0.11 (se=0.62) (N=186); ϕ_a: β=0.38 (se=0.21) (N=186)

- Unadjusted TFPQ is weakly correlated with Lab TFPQ

Result 1: Adjusting for Product Specifications is Important



ϕ_u: β=0.11 (se=0.62) (N=186); ϕ_a: β=0.38 (se=0.21) (N=186)

- Unadjusted TFPQ is weakly correlated with Lab TFPQ
- Controlling for specifications makes TFPQ strengthens the correlation significantly
- Considering Dispersion 90/10 percentile ratios:
  - Unadjusted TFPQ: 4.7
  - Spec-Adjusted TFPQ: 3.1
  - Lab TFPQ: 1.3

TFPZ – Quality Productivity

Even if two firms are producing the same product they may differ in quality

We collected data on 11 different dimensions of quality:

- E.g., corners, waviness, packedness, design accuracy, etc.

We estimate TFPZ by replacing quantity produced by quality produced in our production function

We convert our quality metrics to consumer valuation using a simple CES demand system to create a theory-based quality index

$$\ln x = (\sigma - 1) \sum_j \theta_j \ln q_j - \sigma \ln p + c$$

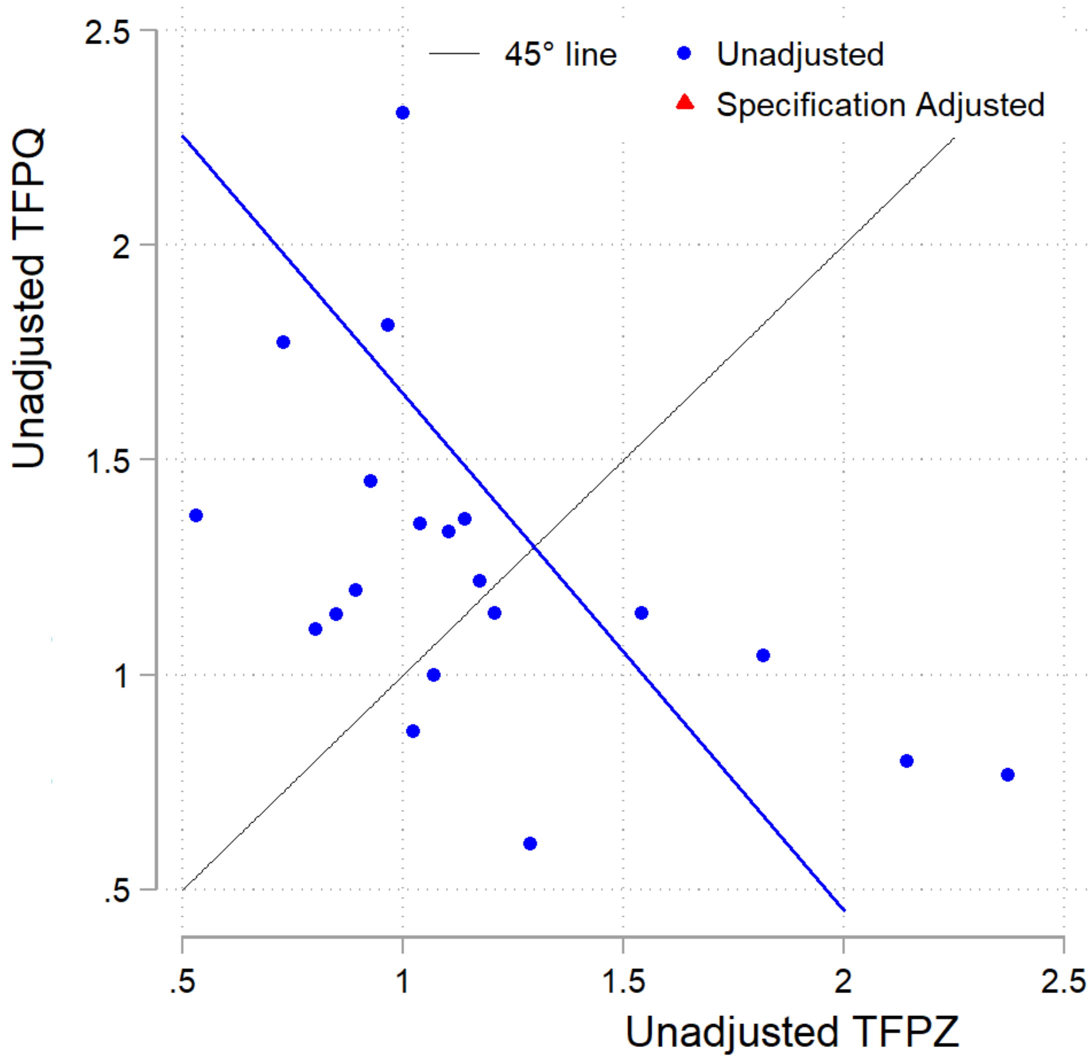
We then estimate:

- **Unadjusted TFPZ** (ζ_u)

-and-

- **Specification-Adjusted TFPZ** (ζ_a)

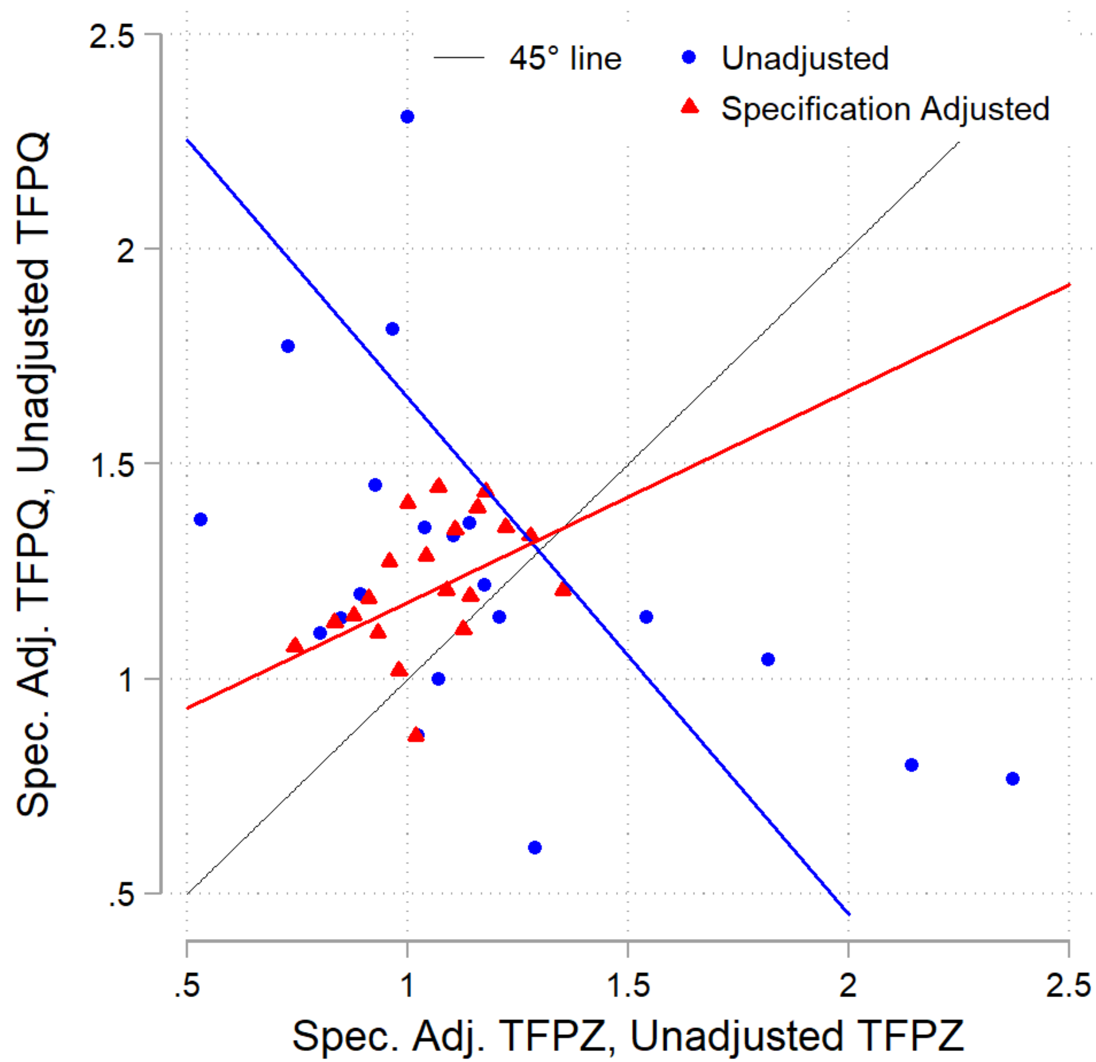
Result 2: Quantity versus Quality Productivity



ϕ_u: β=-1.20 (0.19) (N=209); ϕ_a: β=0.49 (0.22) (N=209)

- Unadjusted TFPQ and TFPZ are negatively correlated!

Result 2: Quantity versus Quality Productivity



ϕ_u: β=-1.20 (0.19) (N=209); ϕ_a: β=0.49 (0.22) (N=209)

- Unadjusted TFPQ and TFPZ are negatively correlated!
- But positively correlated after spec adjusting
- More capable firms make varieties with more demanding specs

TFPC: Firm Capabilities

We aggregate quality and quantity production functions to form **firm capability**

We multiply TFPQ and TFPZ to get TFPC (as implied by CES utility)

$$\Pi_j q_j^{\theta_j} x = \zeta_a \phi_a e^{\lambda(\gamma+\delta)} l^{\alpha_l+\beta_l} k^{\alpha_k+\beta_k} e^{\epsilon+\varepsilon}$$

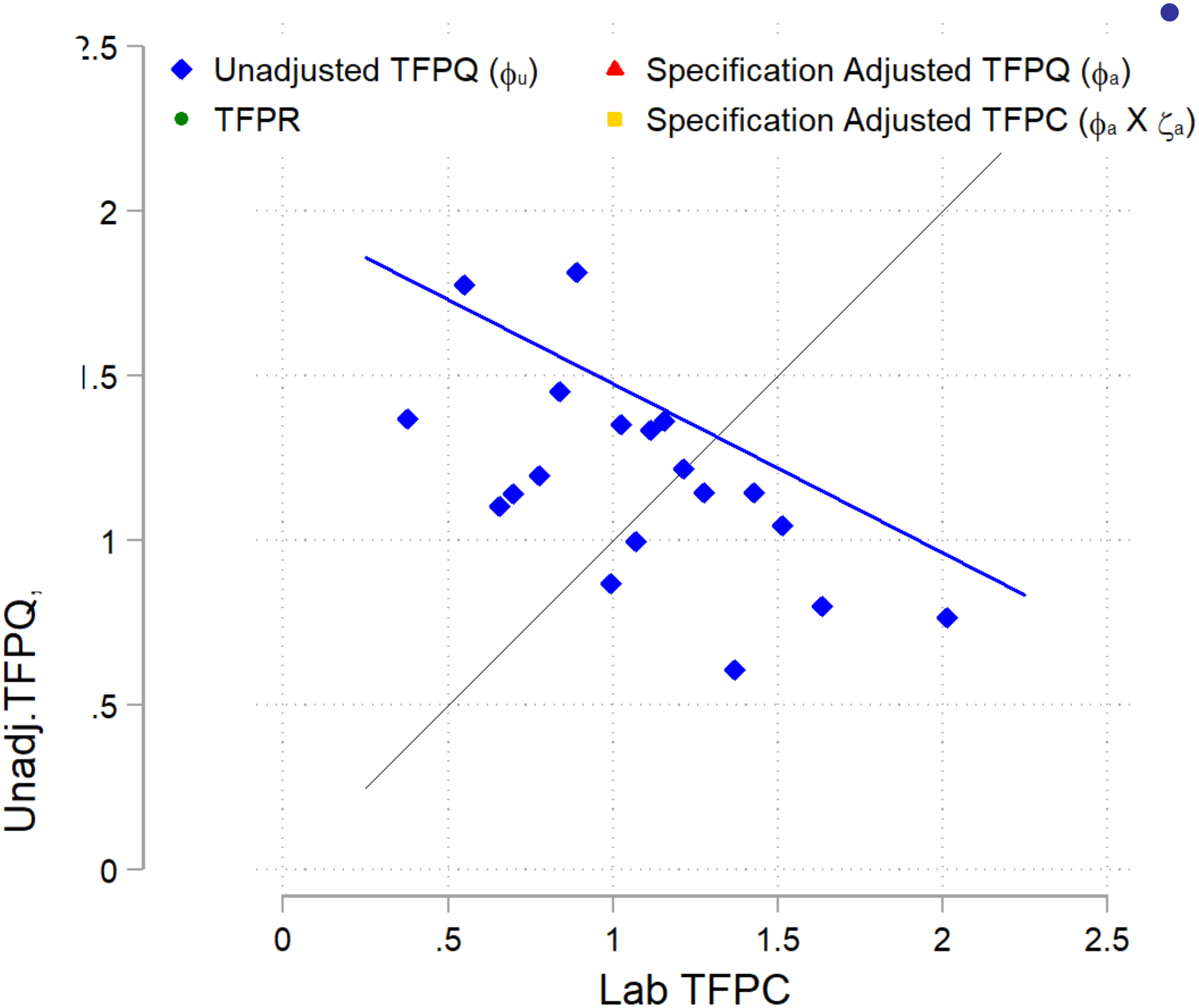
And then we estimate:

- **Unadjusted TFPC** ($\phi_u \zeta_u$)
- **Specification-Adjusted TFPC** ($\phi_a \zeta_a$)

We do the same for the quality of the rug produced in the lab and estimate

Lab TFPC = **Lab TFPQ x Lab TFPZ**

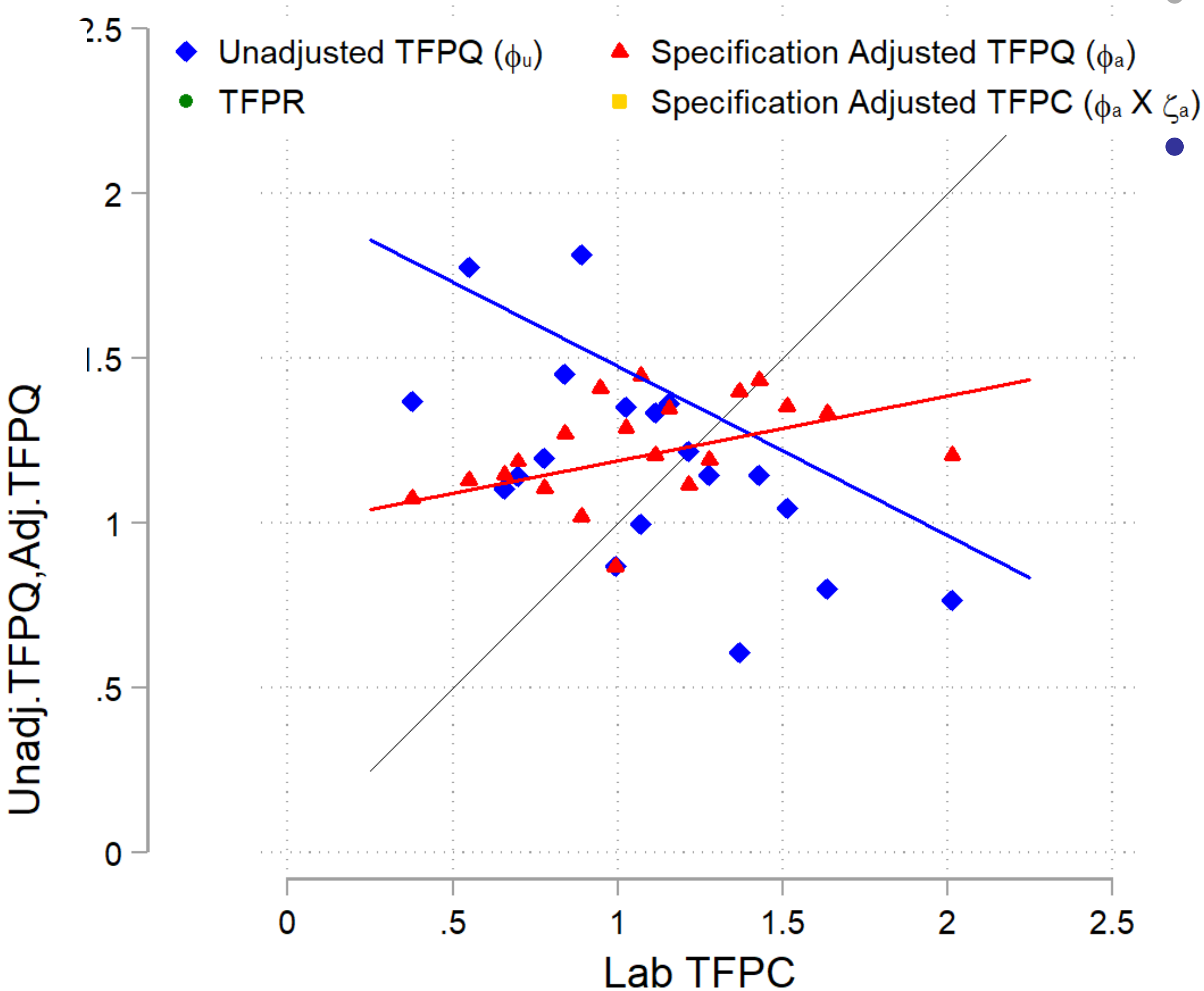
Result 3: Proxies for "True" Firm Capability



ϕ_u:β= -0.51 (0.27) (N=186); ϕ_a:β= 0.20 (0.09) (N=186)
TFPR:β= 0.05 (0.10) (N=180); ϕ_a X ζ_a:β= 0.26 (0.10) (N=186)

- Unadjusted TFPQ is a misleading measure for firm capability

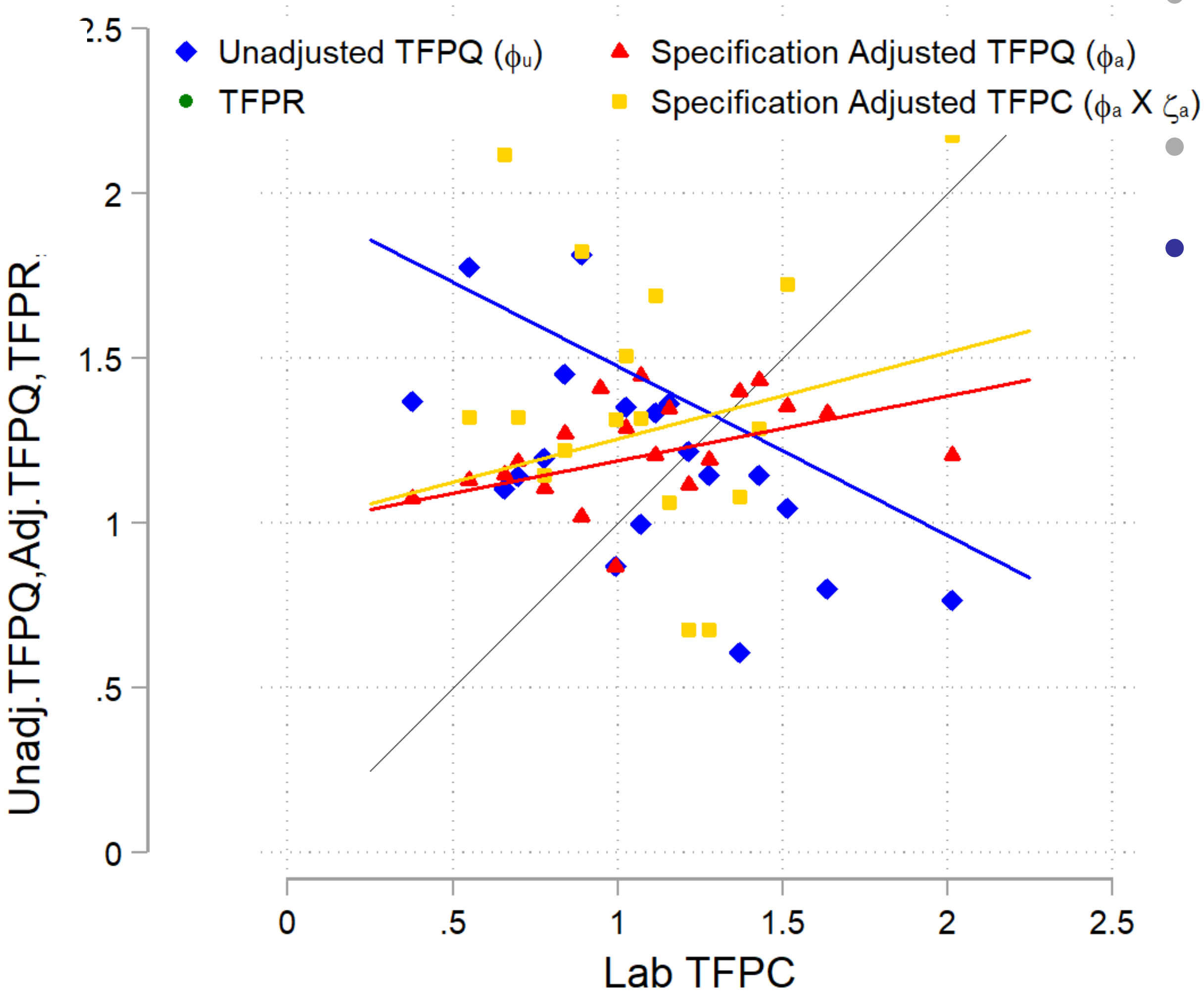
Result 3: Proxies for "True" Firm Capability



ϕ_u:β= -0.51 (0.27) (N=186); ϕ_a:β= 0.20 (0.09) (N=186)
TFPR:β= 0.05 (0.10) (N=180); ϕ_a X ζ_a:β= 0.26 (0.10) (N=186)

- Unadjusted TFPQ is a misleading measure for firm capability
- Spec-adjusting is a better proxy

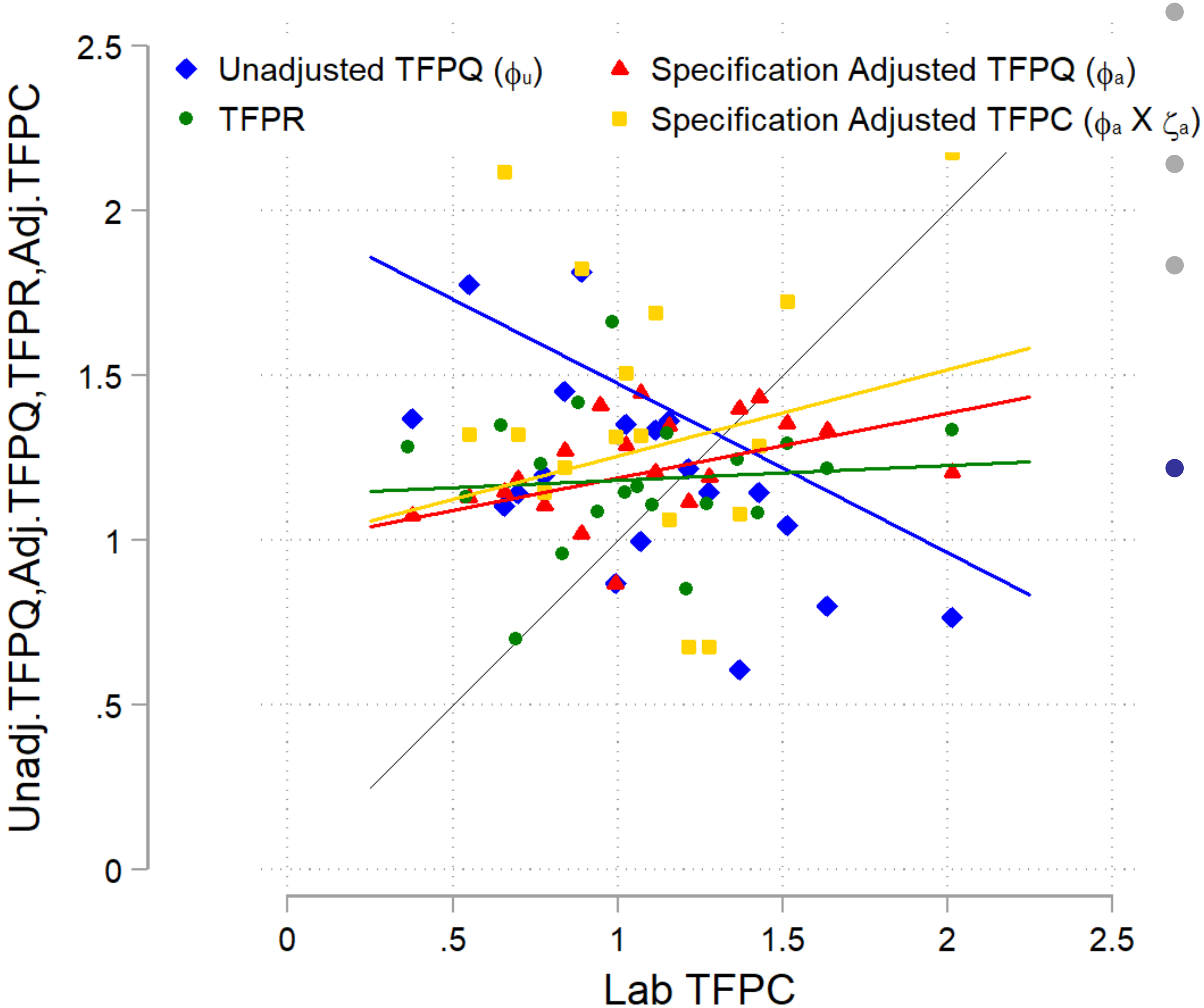
Result 3: Proxies for "True" Firm Capability



ϕ_u:β= -0.51 (0.27) (N=186); ϕ_a:β= 0.20 (0.09) (N=186)
TFPR:β= 0.05 (0.10) (N=180); ϕ_a X ζ_a:β= 0.26 (0.10) (N=186)

- Unadjusted TFPQ is a misleading measure for firm capability
- Spec-adjusting is a better proxy
- Re-assuringly, spec-adjust capabilities (from survey data) matches lab capability best

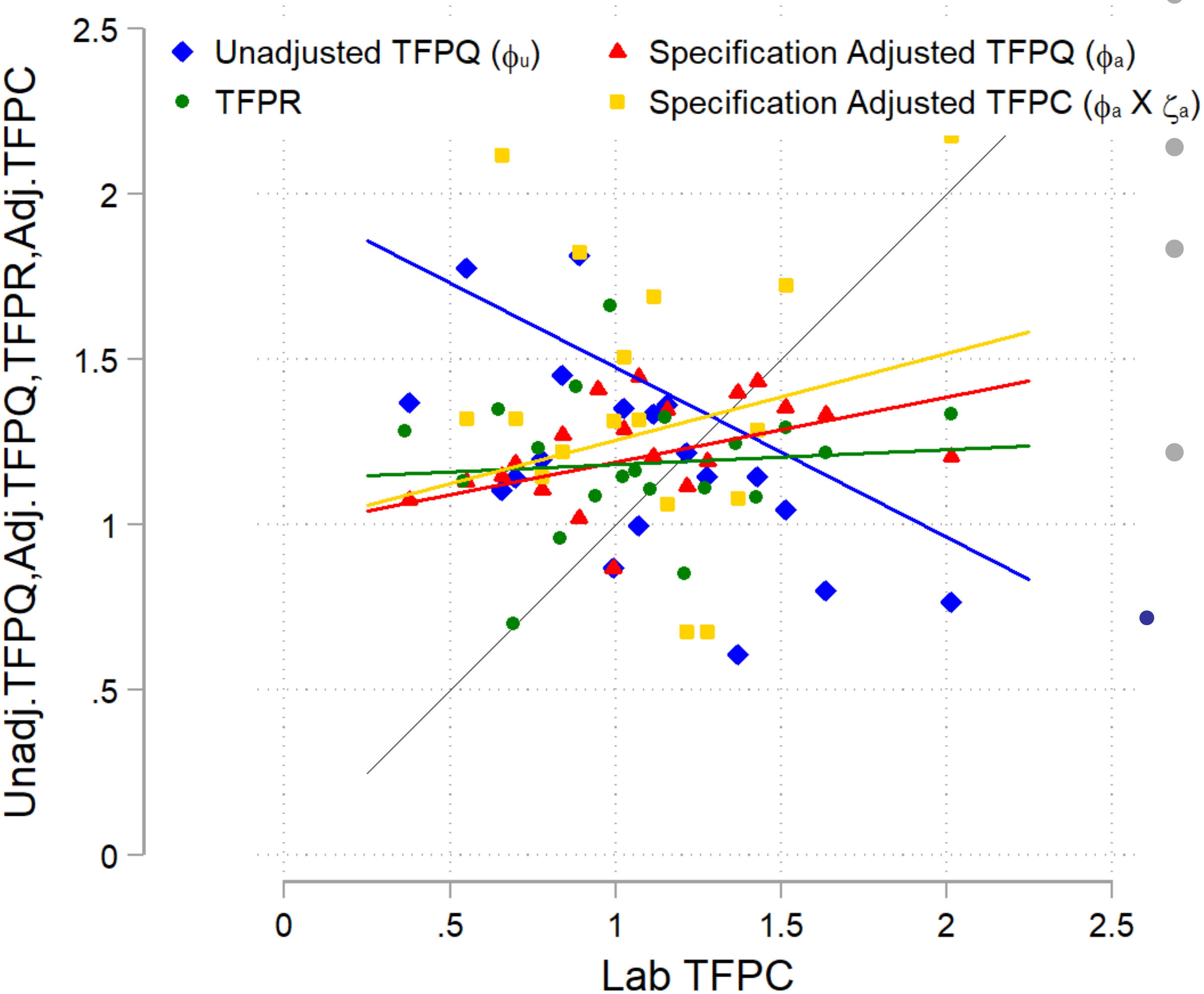
Result 3: Proxies for "True" Firm Capability



ϕ_u:β= -0.51 (0.27) (N=186); ϕ_a:β= 0.20 (0.09) (N=186)
TFPR:β= 0.05 (0.10) (N=180); ϕ_a X ζ_a:β= 0.26 (0.10) (N=186)

- Unadjusted TFPQ is a misleading measure for firm capability
- Spec-adjusting is a better proxy
- Re-assuringly, spec-adjust capabilities (from survey data) matches lab capability best
- TFPR is positively correlates (but weak)

Result 3: Proxies for "True" Firm Capability



ϕ_u:β= -0.51 (0.27) (N=186); ϕ_a:β= 0.20 (0.09) (N=186)
TFPR:β= 0.05 (0.10) (N=180); ϕ_a X ζ_a:β= 0.26 (0.10) (N=186)

- Unadjusted TFPQ is a misleading measure for firm capability
- Spec-adjusting is a better proxy
- Re-assuringly, spec-adjust capabilities (from survey data) matches lab capability best
- TFPR is positively correlates (but weak)

- Considering Dispersion 90/10 percentile ratios:
  - Unadjusted TFPC: 4.3
  - Spec-Adjusted TFPC: 3.5
  - **Lab TFPC: 2.3**
  - **TFPR: 2.7**

Conclusions

- Adjusting survey-based productivity measures using specifications reduces measured dispersion by 1/3

- Large variation in firm capabilities
  - Efficient firms are also those that produce high quality, conditional on product specs

- If researchers are interested in broader capabilities of firms:

  - TFPR may be better proxy than (unadjusted) TFPQ

  - But tailored surveys and/or benchmarking may be best way to understand performance differences across firms