Do larger firms exert more market power?

Markups and markdowns along the size distribution

M. Mertens[1] B. Mottironi[2]

[1]Halle Institute for Economic Research (IWH)

[2]London School of Economics (LSE)

CompNet Annual Conference - October 19, 2023

Table of Contents

Motivation

- Extensive debate on rising concentration
  - Potentially good for aggregate productivity
  - Potentially harmful to competition
- A critical Q: *Do larger firms exert more market power?*
  - Most theories assume they do (e.g. Cournot, Bertrand, linear demand)
  - Consistently with Marshall's second law of demand
    - Moving towards lower prices, the demand becomes less elastic
  - Implications for several economic outcomes
    - Resource allocation, factor shares, gains from trade, etc

Preview of the results

- Standard definitions
  - (Price) Markup: price over marginal cost ($\equiv \mu_{it}$)
  - (Wage) Markdown: MRPL over wage ($\equiv \gamma_{it}$)
- Empirical findings:
  1. Unexpectedly, larger firms have lower markups
  2. Their markdowns are instead (much) higher
  3. *Controlling for markdowns*, markups grow in size
  4. Markups and markdowns are negatively correlated
- We discuss a potential mechanism
  - Consistent with the Marshall's second law of demand

Literature

- 'Production approach' to markup (and markdown) estimation
  - Hall (1986); Dobbelaere & Mairesse, (2013); De Loecker & Warzynski, (2012)
- Application to recent trends
  - De Loecker et al. (2020); Autor et al. (2020); Mertens (2021); Yeh et al. (2022); Deb et al. (2022); Bighelli et al. (2023)
- Discussion of the related issues:
  - Unobserved monopsony (Hashemi et al., 2022)
  - Price bias (Bond et al., 2021; De Ridder et al., 2022)
  - Non-neutral technology (Demirer, 2022; Raval, 2023; Raval, 2023b)

Table of Contents

The 'production approach': intuition

- Ratio estimators are derived from firms' optimality conditions
  - In perfect competition, expenditure shares equal output elasticities
  - Both markups and markdowns generate a wedge between the two
  - Take a flexible input with no markdown (e.g. raw materials):
    - Markup = its output elasticity over its expenditure share
  - With monopsony, the formula above includes the markdown, too
    - Markdown = 'labour ratio' over 'materials ratio'

The 'production approach': pros and cons

- Main advantages
  - No need to estimate firm-specific product demand and labour supply
  - No need to make assumptions on the 'source' of market power
  - A convenient method to apply in a variety of countries and industries
- Main critical assumptions (Bond et al., 2021)
  1. No monopsony over the flexible input
  2. No adjustment costs in the flexible input
  3. Flexible input is not used to affect the demand

Estimators from cost minimization

- $\text{Min}\{r_{it}K_{it} + w_{it}L_{it} + z_{it}M_{it}\}$ s.t. $Q_{it} > Q^*$
  - FOC(M_{it}): $z_{it} = \lambda_{it}\frac{\partial Q_{it}}{\partial M_{it}}$
  - Combine with $\mu_{it} = \frac{P_{it}}{\lambda_{it}}$ and $\theta_{it}^M \equiv \frac{\partial Q_{it}}{\partial M_{it}}\frac{M_{it}}{Q_{it}}$ to get the markup:

$$\mu_{it} = \theta_{it}^M \frac{P_{it}Q_{it}}{z_{it}M_{it}}$$

  - FOC(L_{it}): $w_{it}\left(1 + \frac{\partial w_{it}}{\partial L_{it}}\frac{L_{it}}{w_{it}}\right) = \lambda_{it}\frac{\partial Q_{it}}{\partial L_{it}} = \frac{P_{it}}{\mu_{it}}\theta_{it}^L\frac{Q_{it}}{L_{it}}$
  - Substitute the equation from markup to get $\left(1 + \frac{\partial w_{it}}{\partial L_{it}}\frac{L_{it}}{w_{it}}\right) = \frac{\theta_{it}^L z_{it}M_{it}}{\theta_{it}^M w_{it}L_{it}}$
  - Profit maximization implies $\gamma_{it} = \left(1 + \frac{\partial w_{it}}{\partial L_{it}}\frac{L_{it}}{w_{it}}\right)$, hence the markdown:

$$\gamma_{it} = \frac{\theta_{it}^L}{\theta_{it}^M}\frac{z_{it}M_{it}}{w_{it}L_{it}}$$

Estimators from 'efficient bargaining'

- Consider the following problem: $\text{Max}\{(w_{it}L_{it})^{\phi_{it}}\Pi_{it}^{1-\phi_{it}}\}$
- The estimators of μ_{it} and γ_{it} remain the same
  - Dobbelaere & Mairesse (2013), Mertens (2021), Chern Wong (2023)
- The markdown is now the outcome of two opposite forces
  - Monopsony power of firms vs bargaining power of workers
  - It lies below unity if worker power is relatively higher

Cobb-Douglas vs Translog

- What if we assume a Cobb-Douglas production function?

$$\mu_{it} = \theta^M \frac{P_{it} Q_{it}}{z_{it} M_{it}} \qquad \gamma_{it} = \frac{\theta^L}{\theta^M} \frac{z_{it} M_{it}}{w_{it} L_{it}}$$

- If output elasticities are constant across firms of the same industry:
  - *Levels* of markups and markdowns need production function estimates
  - But their *variation* is fully captured by the expenditure shares alone
- What if output elasticities vary across firms?
  - The estimation of (translog) production functions becomes necessary
  - We need prices

A dual approach

1. Cobb-Douglas assumption for a large database
   - Market power vs firm size across 19 European countries
2. Richer specification for a sub-sample with firm-level prices
   - Same analysis, but using quantity-based translog output elasticities
   - In both cases: *no price bias*

- Our procedure for #2 follows the literature:
  - Control function for productivity (Olley & Pakes, 1996)
  - Material as flexible input (Levinsohn & Petrin, 2003)
  - Control function for unobserved input prices (De Loecker et al., 2016)
  - Firm-level price indices for multi-product firms (Eslava et al., 2004)

Table of Contents

CompNet Data

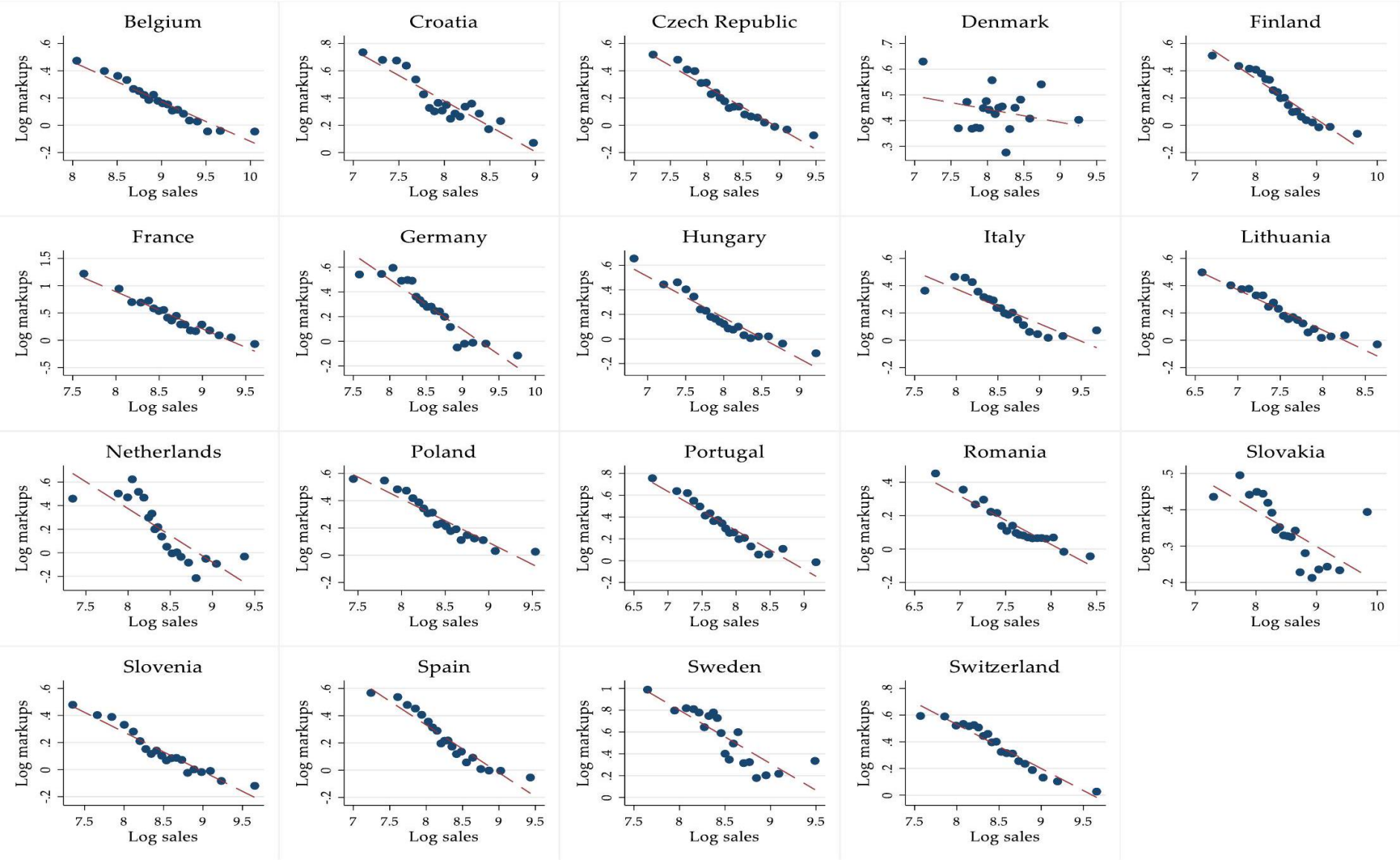
- Micro-based indicators from firm-level administrative data
  - Aggregated at the 2-digit industry level for confidentiality issues
- It includes:
  - Percentiles and dispersion measures of firm-level variables
  - Parametric indicators (TFP, markups, markdowns, MRP, OP gap)
  - Joint distributions of two variables
    - E.g. Average markups by quintile of sales
    - Or average markdown by quintile of markup
- Firm-level prices are not available
  - Our CompNet-based analysis relies on the Cobb-Douglas assumption

AFiD Data

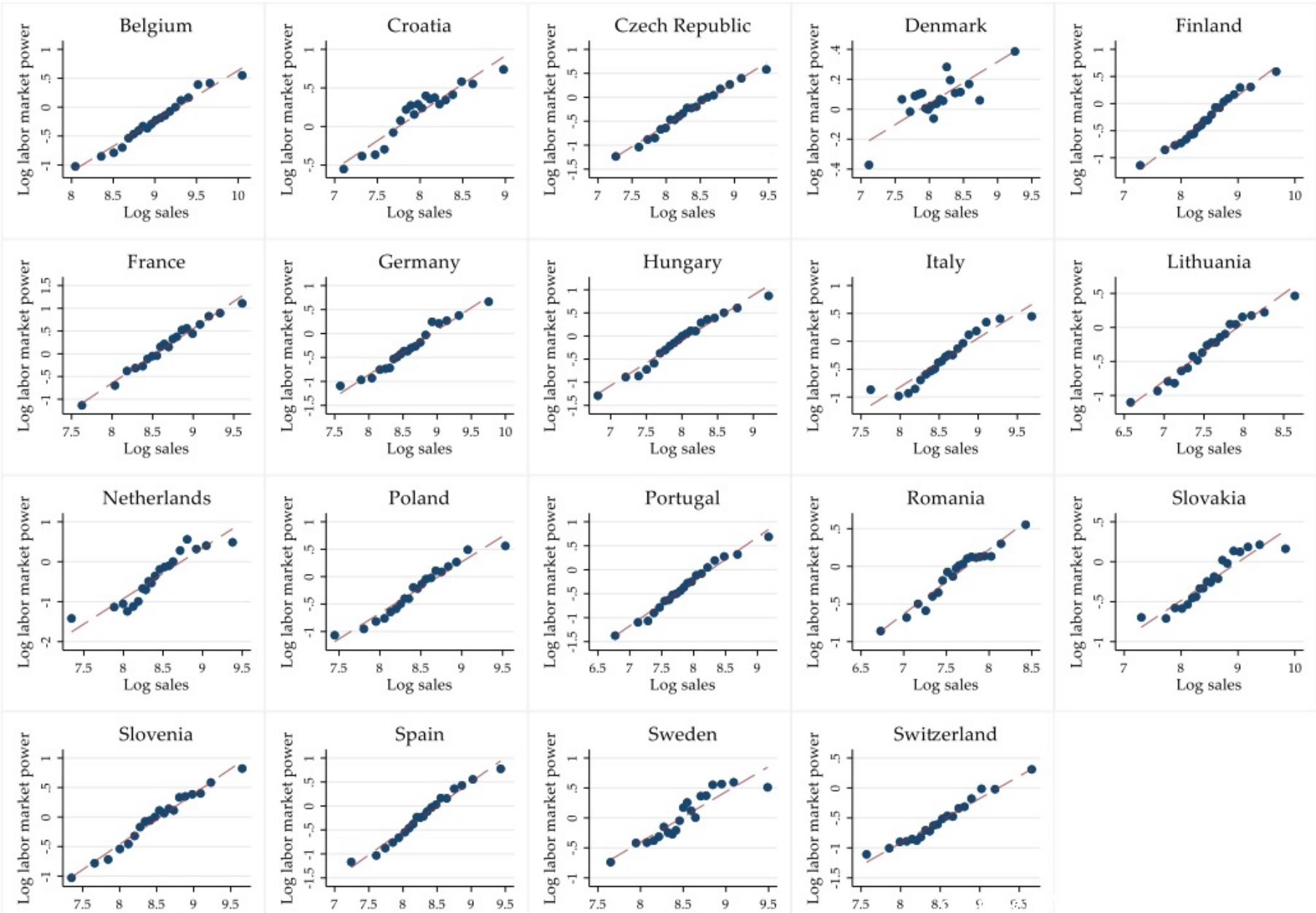
- Firm-level data from German administrative sources
  - Includes manufacturing firms with > 20 employees
  - Rotating sample, covers 40% of the underlying population
  - Time period observed: 1995-2016
- All necessary information for production function estimation
  - Balance-sheet data
  - Income statements
  - NACE classifications
- And, remarkably, product-level prices and quantities, up to 10-digit
  - E.g. "Workwear - Long trousers for men, cotton"

Table of Contents

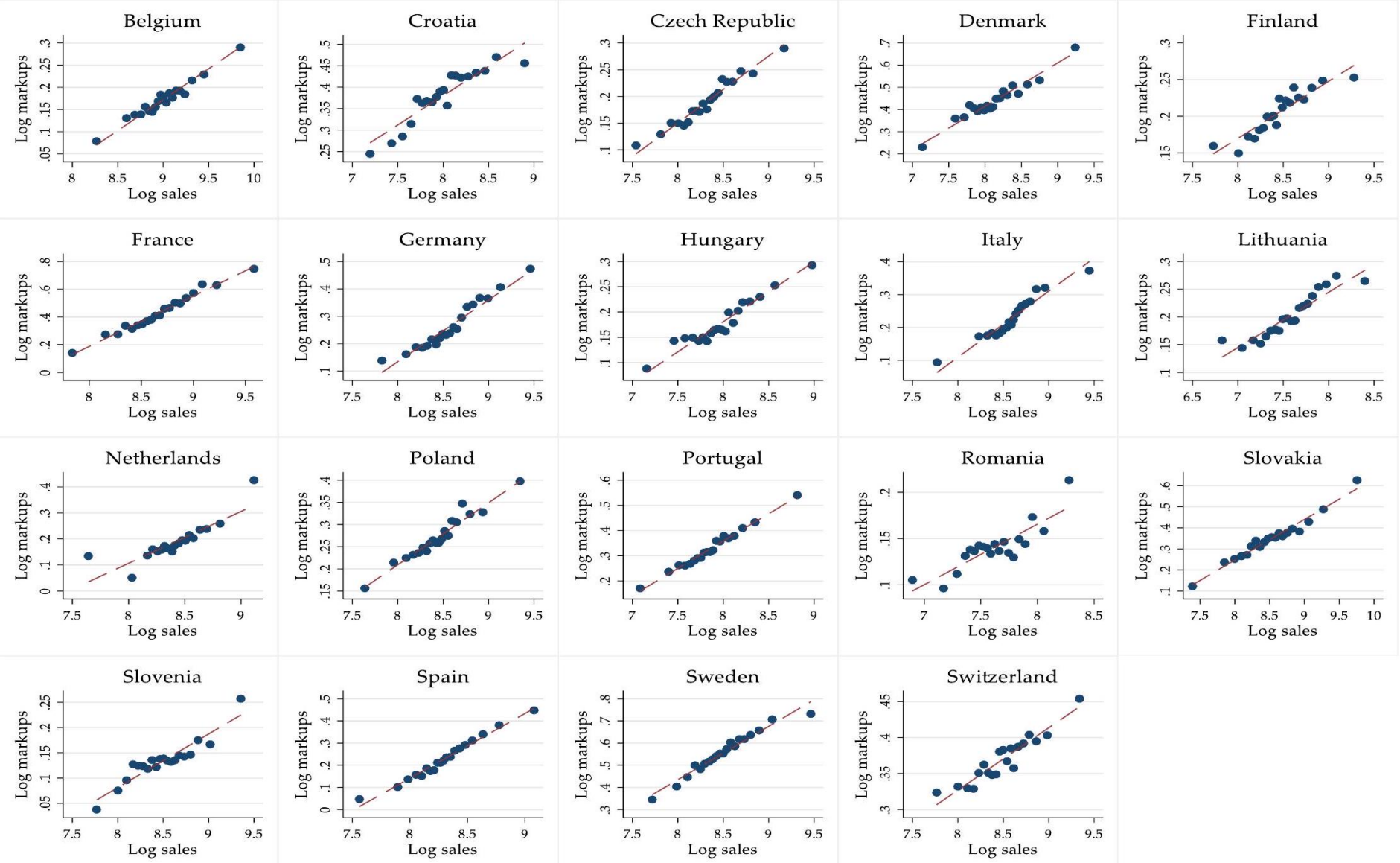
CompNet data: markups vs sales



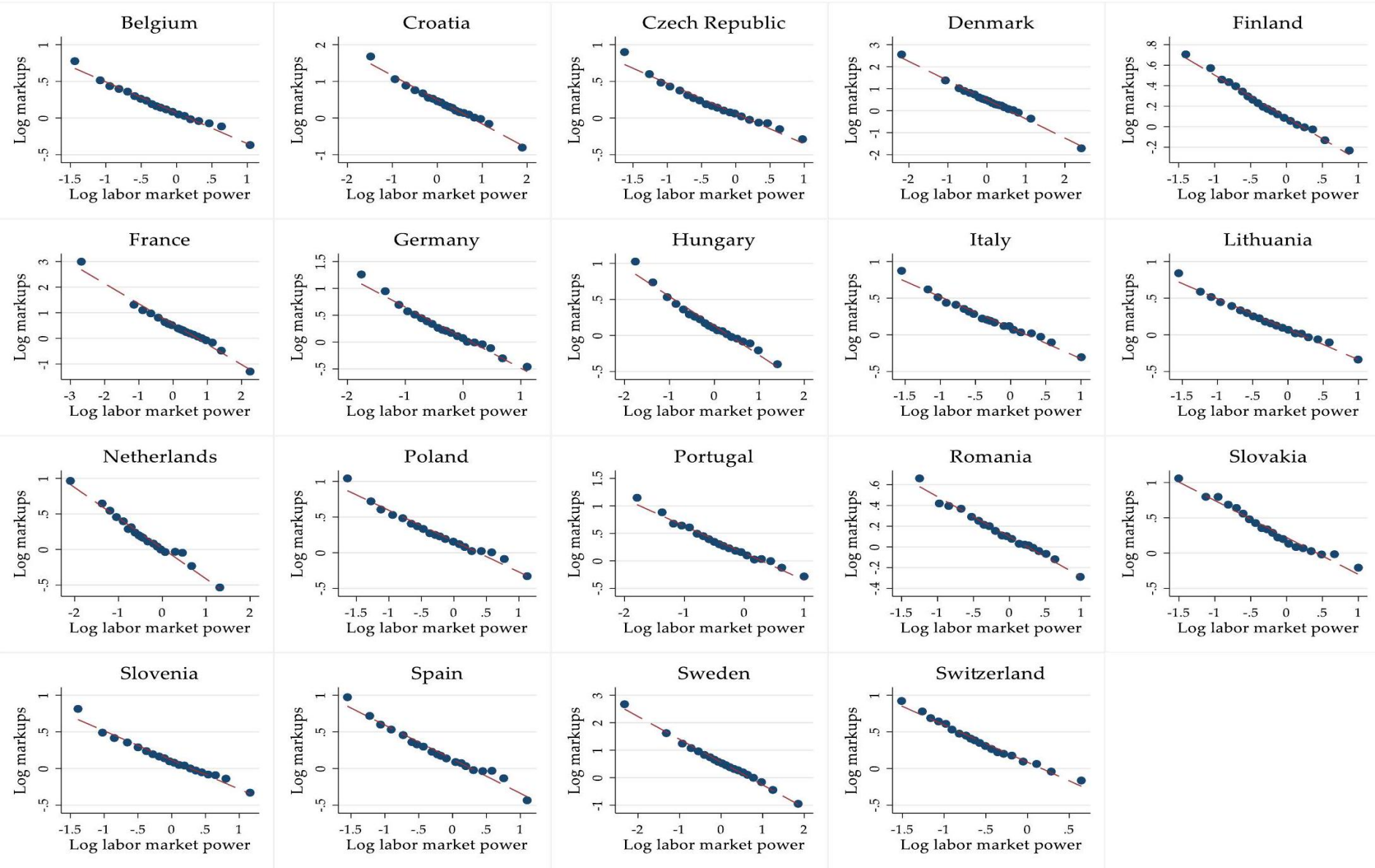
CompNet data: markdowns vs sales



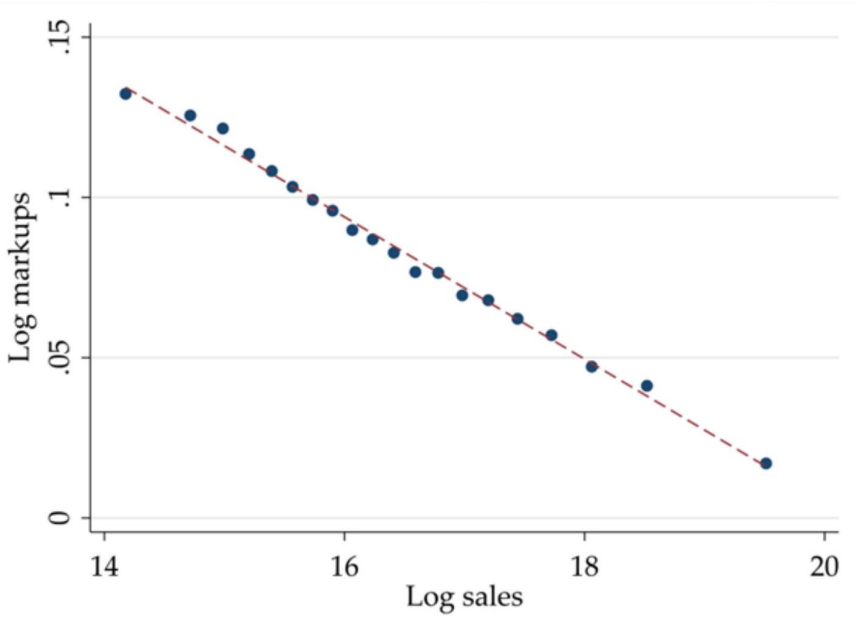
CompNet data: markups vs sales controlling for markdowns

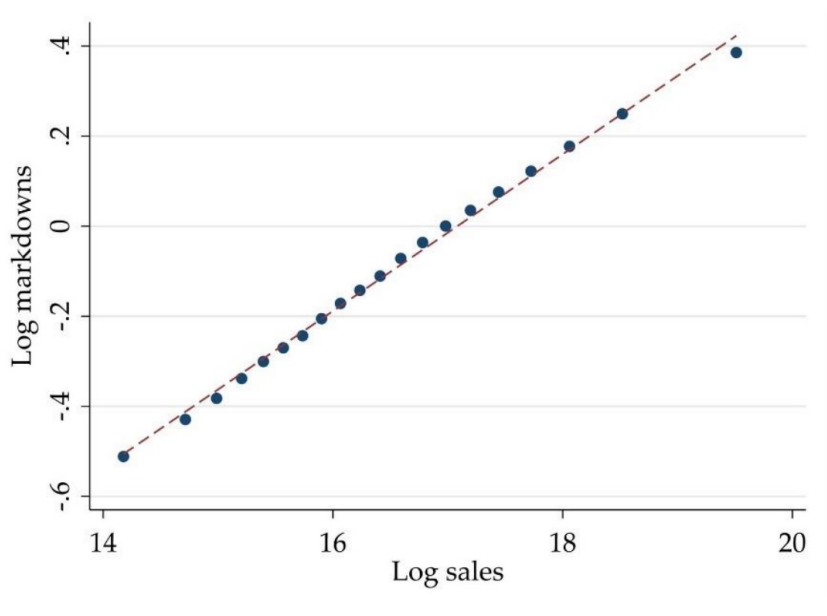


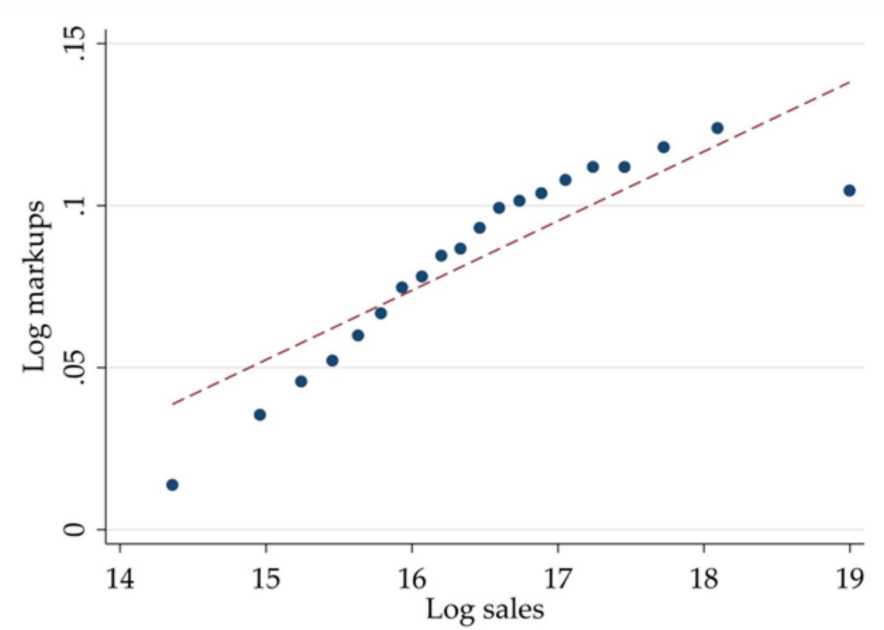
CompNet data: markups vs markdowns



AFiD data: same 4 charts







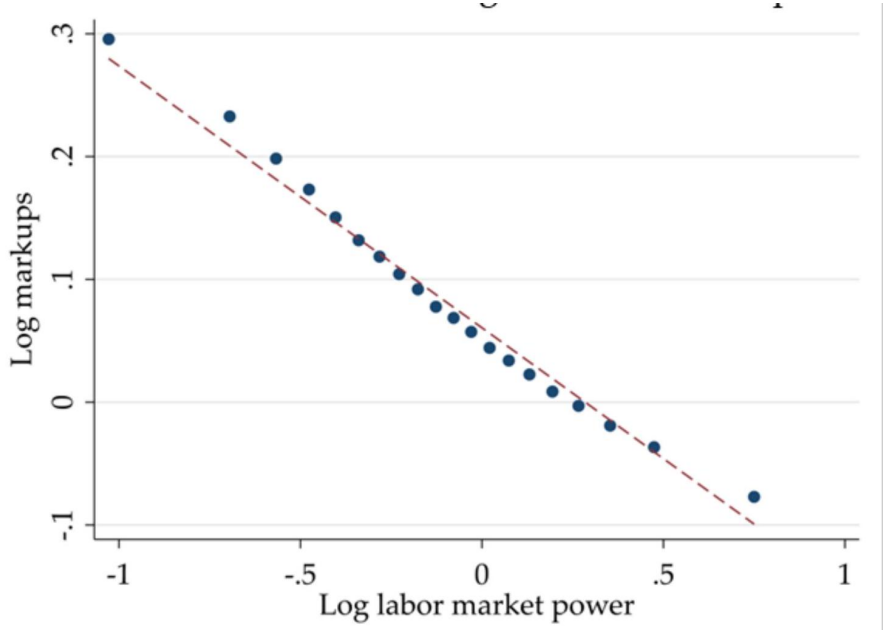


Table of Contents

A possible mechanism

- Efficient bargaining, with worker power depending on firm size
  - $\text{Max}\{(w_{it}L_{it})^{\phi_{it}(Q_{it})}\Pi_{it}^{1-\phi_{it}(Q_{it})}\}$
- Taking the FOC wrt quantities, and rearranging, we get:
  - $\left(\frac{\partial P_{it}}{\partial Q_{it}}\frac{Q_{it}}{P_{it}}+1\right)-\frac{1}{\mu_{it}}=\frac{\phi'(Q_{it})}{1-\phi(Q_{it})}\log\left(\frac{\Pi_{it}}{w_{it}L_{it}}\right)\frac{\Pi_{it}}{P_{it}}$

$$\phi'_{it}(Q_{it})<0\rightarrow\mu_{it}<\frac{1}{\frac{\partial P_{it}}{\partial Q_{it}}\frac{Q_{it}}{P_{it}}+1}$$

- Larger inequality for large firms with high wage markdowns
- *Intuition*: profitable firms cut prices to expand and reduce rent-sharing

Implications

1. Misallocation and optimal policies
2. Concentration and factor shares
3. Gains from trade
4. Demand-based estimations
5. Pass-through and rent-sharing

Table of Contents

Summary

- Novel empirical results on firm-level markups
  - Robust to typical issues:
    - Unobserved monopsony, price bias, non-neutral technology, etc.
  - Consistent across 19 European countries (all available)
- Key findings:
  1. Larger firms appear to have lower markups and higher markdowns
  2. Evidence suggests a trade-off between product and labour market power

Thank you!

Table of Contents

Assumption (1/3): Unobserved monopsony

- A relevant threat for studies that use labour or COGS
  - From Hall (1986): $\mu_{it}^{L} = \theta_{it}^{L} \frac{P_{it} Q_{it}}{w_{it} L_{it}} = \mu_{it} \gamma_{it}$
  - This can explain different results on the markups-size correlation
  - Some claims might be reinterpreted in terms of *labour* market power
- Intermediate input markets are typically assumed to be competitive
  - If not, upward bias in markups, downward bias in markdowns
- Evidence from Morlacco (2020) on imported inputs:
  - Large firms have more monopsony over materials than small firms
  - Correcting this bias would reinforce our results

Assumption (2/3): Adjustment costs

- Intermediate inputs are typically assumed to be flexible
  - E.g. Hall (2004) use them to estimate adjustment costs in K_{it} and L_{it}
- Labour adjustment costs lead to an upward bias in markdowns
  - But Yeh et al. (2022) show that the impact is negligible (+3%)
  - Much smaller than the gap in markdowns between small and large firms

Assumption (3/3): Marketing

- Downward bias in markups if the flexible input is used to affect the demand
- For robustness, we test our correlations separately for firms producing:
  1. Consumption goods and services
  2. Capital goods and services
  3. Intermediate goods and services
- No significant difference

Non-neutral technology and expenditure shares

- Does a labour-augmenting productivity shock affect the output elasticities?
  - It depends on the elasticity of substitution
  - With Cobb-Douglas ($\sigma = 1$), no effect
  - If inputs are gross complements ($\sigma < 1$), then θ_{it}^{M} goes up and θ_{it}^{L} down
- Heterogeneous expenditure shares may depend on this:
  - $\sigma < 1$, while larger firms have higher levels of labour-augmenting tech
- Are translog elasticities flexible enough to capture this?

$$\frac{\theta_{it}^{L}}{\theta_{it}^{M}} = \frac{\beta^{L} + 2\beta^{LL}\ l_{it} + \beta^{KL}\ k_{it} + \beta^{ML}\ m_{it}}{\beta^{M} + 2\beta^{MM}\ m_{it} + \beta^{KM}\ k_{it} + \beta^{LM}\ l_{it}}$$

Translog estimates with non-neutral technology

- Raval (2023) suggests splitting firms in quintiles by the ratio $\frac{z_{it} M_{it}}{w_{it} L_{it}}$
  - Absent labour market power, the ratio is proportional to labour-augmenting productivity
- ...and then measuring output elasticities as cost shares, by quintile
- But cost shares are not informative with imperfect labour markets
- We can run separate translog estimates in the same quintiles
  - If non-neutral technology is the major force behind our results, this solution should generate sufficient heterogeneity in output elasticities

Non-neutral technology: a robustness check