The Incidence of Distortions*

David Atkin[1] Baptiste Bernadac[2] Dave Donaldson[1]
Tishara Garg[1] Federico Huneeus[3]

MIT UCLA Duke & Central Bank of Chile

July 2025

*This research is part of the research agenda of the Bank to track the Chilean economy, under which the Bank has an agreement to access data from the tax authority. The views and opinions expressed are those of the authors alone and do not necessarily reflect those of the Central Bank of Chile.

The Incidence of Distortions

- Economic distortions $\implies$ Misallocation + Low aggregate output
  - Large body of empirical work suggests magnitudes may be large
  - Especially for developing economies (e.g. Banerjee Duflo 2005, Restuccia Rogerson 2008, Hsieh Klenow 2009)

The Incidence of Distortions

- Economic distortions $\implies$ Misallocation + Low aggregate output
  - Large body of empirical work suggests magnitudes may be large
  - Especially for developing economies (e.g. Banerjee Duflo 2005, Restuccia Rogerson 2008, Hsieh Klenow 2009)
- But who bears the burden of distortions?
  - If poor disproportionately harmed, distortions may be even bigger issue
  - Knowledge of equity-efficiency trade-off is vital for policymakers

Existing Work

- Typically documents variation in one distortion, in one sector, across one dimension
  - E.g. retail markups by income (Faber Fally 2021, Gupta 2022, Sangani 2022), taxes by income (Conlon et al. 2022), tariffs by income (Faber 2014, Acosta Cox 2024), textile worker markdowns by gender (Sharma 2022), input markups by buyer (Burstein Cravino Rojas 2024)

Existing Work

- Typically documents variation in one distortion, in one sector, across one dimension
  - E.g. retail markups by income (Faber Fally 2021, Gupta 2022, Sangani 2022), taxes by income (Conlon et al. 2022), tariffs by income (Faber 2014, Acosta Cox 2024), textile worker markdowns by gender (Sharma 2022), input markups by buyer (Burstein Cravino Rojas 2024)
- But...
  - Distortions on final goods may hide exposure to distortions higher up supply chain
  - Revenues from distortions need accounting for
  - What appears harmful focusing on single distortion or sector may be beneficial if correcting other distortions (second best argument)
  - Equity along one dimension may hide inequalities among others

What We Do

1. **Theory** Advances in heterogeneous agent GE models of distorted economies
   - Building on Baqaee and Farhi (2020), derive incidence formula that depend on:

   a. Full matrices of individual-level exposure to (static) distortions

   b. The size of the distortions

What We Do

1. **Theory** Advances in heterogeneous agent GE models of distorted economies
   - Building on Baqaee and Farhi (2020), derive incidence formula that depend on:

   a. Full matrices of individual-level exposure to (static) distortions
   b. The size of the distortions

2. **Measurement** Most granular admin & survey microdata we are aware of: Chile
   - Administrative: Consumer-to-firm, firm-to-firm, firm-to-employee, firm-to-owner registries, pension holdings, tax and transfer records
   - Surveys: Large-sample household, employment and microenterprise $\Rightarrow$ Informality
   - Combine Admin+Surveys: Statistical matching on income and consumption side

More Related Work

- Building national accounts from microdata (Adao et al. 2022 for Ecuador, Andersen et al. 2023 for Denmark)
  - No distortions; Either no consumers, or no data on what is bought/owned/F2F links
- Impact of globalization in presence of multiple distortions (Atkin Donaldson 2022 for trade shocks; Manelici, Ulate, Vasquez, Zarate 2024 for FDI)
  - No focus on incidence, limited ability to track linkages between agents in economy

Today

1. Introduction
2. Theoretical Framework and Incidence Formula
3. Distortions Estimation Strategy
4. Data
5. Exposure Matrices and Wedges
6. The Incidence of Distortions
7. Next Steps

Theoretical Framework: Set-up

- $\mathcal{C}$ consumers c, $\mathcal{N}$ firm-products i, and $\mathcal{F}$ factors f (in fixed supply, L_f)
  - Consumers have heterogeneous (homothetic) prefs, endowments, ownership, claims
  - Firms have (CRS) technologies

- What is welfare impact of reducing distortions, how does it vary across individuals?
  - For each individual, calculate change in real income from going from current, exogenous wedge distribution to $p = mc$ and $w = vmp$
  - To do so, we start with a small change in wedges:

Theoretical Framework: Set-up

- $\mathcal{C}$ consumers c, $\mathcal{N}$ firm-products i, and $\mathcal{F}$ factors f (in fixed supply, L_f)
  - Consumers have heterogeneous (homothetic) prefs, endowments, ownership, claims
  - Firms have (CRS) technologies

- What is welfare impact of reducing distortions, how does it vary across individuals?
  - For each individual, calculate change in real income from going from current, exogenous wedge distribution to $p = mc$ and $w = vmp$
  - To do so, we start with a small change in wedges:

$$d \ln \mathcal{Y}_c = -\sum_{i \in \mathcal{N}} b_{ci} \, d \ln p_i + \underbrace{\sum_{f \in \mathcal{F}} \Theta_{cf} \, d \ln w_f + \sum_{i \in \mathcal{N}} \Theta_{ci} \, d \ln \pi_i + \Theta_{ct} \, d \ln T}_{d \ln \chi_c}$$

  - $\mathcal{Y}_c$ is $c's$ real income, b_{ci} is c's expenditure share on goods from firm i
  - Θ_{cf}, Θ_{ci} and Θ_{ct} are shares of income χ_c from factor f's wages w_f, firm i's after-tax profits π_i, and government transfers T

Theoretical Framework: Consumption-Based Exposure

- Price changes due to changes in distortions

$$d \ln p_i = \sum_{j \in \mathcal{N}} \tilde{\Psi}_{ij} \left(d \ln \mu_j + \sum_{k \in \mathcal{N}, \mathcal{F}} \tilde{\Omega}_{jk} d \ln \tau_{jk} \right) + \sum_{f \in \mathcal{F}} \tilde{\Psi}_{if} d \ln w_f$$

  - where $\Omega_{jk} = \frac{p_k^I x_{jk}}{p_j y_j}$ is share of input k in revenues of j, $\tilde{\Omega}_{jk} \equiv \tau_{jk} \mu_j \Omega_{jk}$ is share of the cost of input k in costs of j, and $\tilde{\Psi} = (I - \tilde{\Omega})^{-1}$

- μ_j is output wedge for firm-product j ($p_j = \mu_j c_j$, e.g. markup, sales tax)
- τ_{jk} is the k^{th} input wedge for firm j (as-if mc is $\tau_{jk} p_k^I$, e.g. monopsony, payroll tax)
- Do not need to separate $\mu_j \tau_{jk}$ beyond allocating wedge revenue (e.g. via tax data)

Theoretical Framework: Income-Based Exposure

- Factor income changes due to changes in distortions

$$d \ln w_f = -\underbrace{\sum_{l,m \in \mathcal{F},\mathcal{N}} (d \ln \mu_l + d \ln \tau_{lm}) \lambda_l \Omega_{lm} \Psi_{mf}}_{\text{Direct effect of } p \text{ change}}$$

$$+ \underbrace{\sum_{k \in \mathcal{C},\mathcal{N}} \mu_k^{-1} \lambda_k \text{Cov}_{\tilde{\Omega}^{(k)}}(d \ln \tilde{\Omega}^{(k)}, diag(\tau^{(k)})^{-1} \Psi_{(f)})}_{\text{Substitution within } k} + \underbrace{\sum_{c \in \mathcal{C}} d\chi_c \sum_{k \in \mathcal{N}} b_{ck} \Psi_{kf}}_{\text{Income changes for consumers}}$$

  - where λ_i is sales share of firm or income share of factor/consumer i

- Related expressions for changes in distortion revenue via ownership ($d \ln \pi_i$) and via transfers ($d \ln T$)

Theoretical Framework: Solving for Incidence

- Hsieh Klenow (2009) benchmark:
  - (Nested) CES demand $\Rightarrow$ Change in consumer's budget shares of firm-product i

$$d \ln b_{ci} = \underbrace{(1-\theta_{g(i)})(d \ln p_i - \sum_{i' \in g(i)} \alpha_{i' g(i)} d \ln p_{i'})}_{\text{Within HS3 industry}} + \underbrace{(1-\theta_G)(d \ln p_{g(i)} - \sum_{g \in G} \alpha_{gG} d \ln p_g)}_{\text{Across HS3 industries}}$$

  - Sector-specific Cobb-Douglas production function

$$Y_i = A_i K_i^{\beta_K^s} L_i^{\beta_L^s} \prod_{s'} M_{s'i}^{\beta_{Ms'}^s} \text{ with } \sum_{s'} \beta_{Ms'}^s = 1 - \beta_K^s - \beta_L^s$$

Theoretical Framework: Solving for Incidence

- Hsieh Klenow (2009) benchmark:
  - (Nested) CES demand $\Rightarrow$ Change in consumer's budget shares of firm-product i

$$d \ln b_{ci} = \underbrace{(1-\theta_{g(i)})(d \ln p_i - \sum_{i' \in g(i)} \alpha_{i'g(i)} d \ln p_{i'})}_{\text{Within HS3 industry}} + \underbrace{(1-\theta_G)(d \ln p_{g(i)} - \sum_{g \in G} \alpha_{gG} d \ln p_g)}_{\text{Across HS3 industries}}$$

  - Sector-specific Cobb-Douglas production function

$$Y_i \ = \ A_i K_i^{\beta_K^s} L_i^{\beta_L^s} \prod_{s'} M_{s'i}^{\beta_{Ms'}^s} \text{ with } \sum_{s'} \beta_{Ms'}^s = 1 - \beta_K^s - \beta_L^s$$

- Linear system can be solved for any values of elasticities θ and β, wedges μ_i and τ_{ij}, shares Ω_{ij}, $b_c i$, ϕ's, and χ_c (for small changes, or iterating for large changes)

Today

1. Introduction
2. Theoretical Framework and Incidence Formula
3. Distortions Estimation Strategy
4. Data
5. Exposure Matrices and Wedges
6. The Incidence of Distortions
7. Next Steps

Distortions Estimation Strategy

- Wedge: Gap between efficient resource allocation and actual allocation

$$\mu_j \tau_{jk} = \frac{\eta_{jk}}{\Omega_{jk}}$$

  - Ω_{jk}: observed share of input k in j's revenues
  - η_{jk}: output elasticity of firm j with respect to input k
    - Hsieh Klenow (2009) benchmark:
      Sector-specific Cobb-Douglas + US frictionless benchmark (no wedges on average)
      $\Rightarrow$ Calibrate η_{jk} using US sector-specific cost shares

Today

Data: What We Need, Where We Get It From

$$d \ln \mathcal{Y}_c = -\sum_{i\in\mathcal{N}} b_{ci} d \ln p_i + \sum_{f\in\mathcal{F}} \Theta_{cf} d \ln w_f + \sum_{i\in\mathcal{N}} \Theta_{ci} d \ln \pi_i + \Theta_{ct} d \ln T$$

1. Data on Individuals
   - Individual c's consumption shares on firm-product i (b_{ci}) $\Rightarrow$ Direct exposure
     - Product-level firm-to-individual transaction data pushes beyond home/retail scanner data: links individuals to firms, wider coverage
     - Individual-level consumption surveys at store brand-product level to fill in gaps due to no tax ids/informality by statistical matching on F2I consumption + demographics

Data: What We Need, Where We Get It From

$$d \ln \mathcal{Y}_c = -\sum_{i\in\mathcal{N}} b_{ci} d \ln p_i + \sum_{f\in\mathcal{F}} \Theta_{cf} d \ln w_f + \sum_{i\in\mathcal{N}} \Theta_{ci} d \ln \pi_i + \Theta_{ct} d \ln T$$

1. Data on Individuals
   - Consumption shares by product and firm (b_{ci})
   - Labor income by source (Θ_{cf}): Employer-employee data+consumption and labor survey. Define factor as location, including the ROW (soon: education$\times$location)
   - Ownership shares of firms to allocate capital income (Θ_{cf}), profits (Θ_{ci}): Ownership registry, pension holdings by fund, microenterprise survey informal ownership
   - Taxes and transfers for tax revenue rebated (Θ_{ct}): Tax forms, transfer records
   - Grouping variables are age, gender, geography, country of birth, race, education: Civil registry data, voting registry, unemployment insurance records

Data: What We Need, Where We Get It From

$$d \ln \mathcal{Y}_c = -\sum_{i\in\mathcal{N}} b_{ci} \sum_{j\in\mathcal{N}} \tilde{\Psi}_{ij}(d \ln \mu_j + \sum_k \tilde{\Omega}_{jk} d \ln \tau_{jk}) + \sum_f \tilde{\Psi}_{if} d \ln w_f + \ldots$$

2. Data on Firms
   - Input shares by firm-product for indirect exposure ($\tilde{\Omega}_{ij}$): Firm-to-firm transactions
     - Split retailers into single-product i, using products coming into multiproduct firm
   - Revenue shares of labor, capital and materials + total output + tax payments: Tax forms, employer-employee data, firm-to-firm transactions, microenterprise survey
3. Wedge estimation: $\mu_j \tau_{jk} = \eta_{jk}/\Omega_{jk}$
   - Output elasticities (η_{ij}): 3-digit US industry cost shares
   - Input shares of revenue (Ω_{ij}): Firm-to-firm transactions, employer-employee, tax forms

Administrative Data: Summary Statistics for 2022

1	Firm-to-Individual Consumption	Consumers 13,453,311	Suppliers 2,124	Pairs 43,626,887	Transactions 6 Billion
2	Firm-to-Firm Domestic Trade	Buyers 1,354,408	Suppliers 624,073	Pairs 35,993,564	Transactions 2.1 Billion
3	Firm-to-Firm International Trade	Buyers 93,423	Suppliers 155,283	Pairs 273,110	Transactions 5,298,769
4	Firm-to-Employee Employment/Wages	Firms 702,729	Workers 8,242,191	Pairs 13,138,247	Jobs per Worker 1.6
5	Firm-to-Individual Ownership	Owners 1,781,539	Owned 1,445,504	Pairs 3,172,853	Med. Owner Share 34%
6	Gov-to-Individual Net Transfers	Individuals 8,021,862	Policies 10	Pairs 16,262,917	Policy Transactions 16,495,680

Data Coverage Implementation Challenges

Data: What We Need, Where We Get It From

$$d \ln \mathcal{Y}_c = -\sum_{i\in\mathcal{N}} b_{ci}\, d\ln p_i + \sum_{f\in\mathcal{F}} \Theta_{cf}\, d\ln w_f + \sum_{i\in\mathcal{N}} \Theta_{ci}\, d\ln \pi_i + \Theta_{ct}\, d\ln T$$

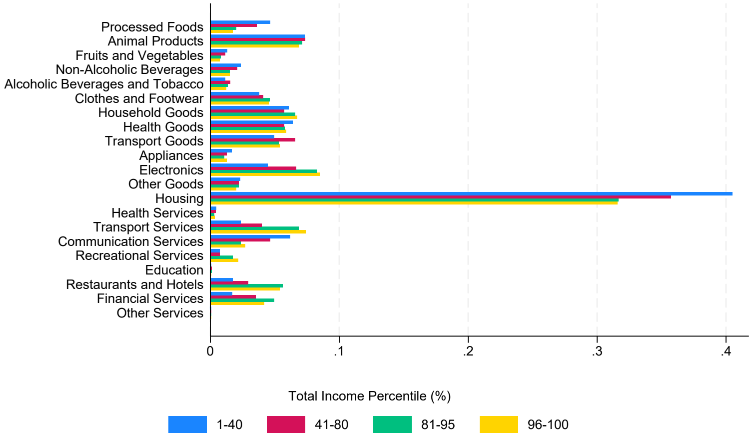
- Taking stock: Framework allows for

1. Interacting distortions across sectors and markets: Y,L,K,M
2. Indirect exposure through entire supply chain
3. Accrual of distortion revenue $\Rightarrow$ Highlight income channel, on top of consumption
4. *ij*-specific wedges
5. General equilibrium

Today

Exposure Matrices: Expenditure Shares From F2H + Survey

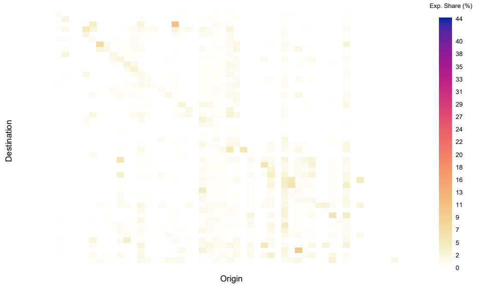
$$d\ \ln \mathcal{Y}_c \approx -\sum_{i \in \mathcal{N}} b_{ci}\, d \ln p_i + \sum_{f \in \mathcal{F}} \Theta_{cf}\, d \ln w_f + \sum_{i \in \mathcal{N}} \Theta_{ci}\, d \ln \pi_i + \Theta_{ct}\, d \ln T$$



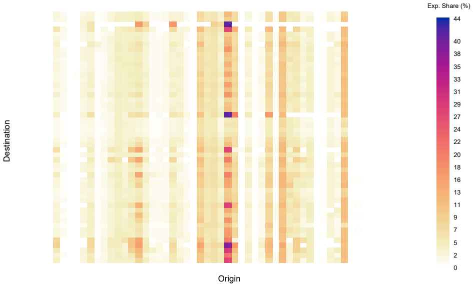
(a) Across Products

Exposure Matrices: Firm-to-Firm Data Across Sectors

$$d\ \ln \mathcal{Y}_c \approx -\sum_{i \in \mathcal{N}} b_{ci} \left(\sum_{j \in \mathcal{N}} \tilde{\Psi}_{ij} \left(d \ln \mu_j + \sum_{k \in \mathcal{N}, \mathcal{F}} \tilde{\Omega}_{jk}\, d \ln \tau_{jk} \right) + \sum_{f \in \mathcal{F}} \tilde{\Psi}_{if}\, d \ln w_f \right) + \ldots$$



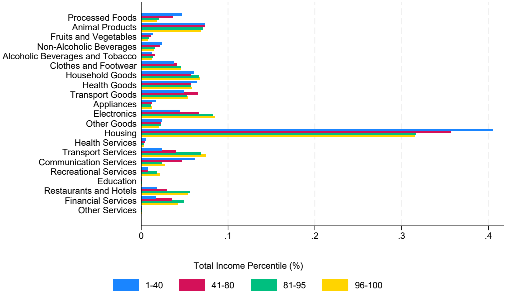
(a) Direct Linkages: $\tilde{\Omega}_{jk}$



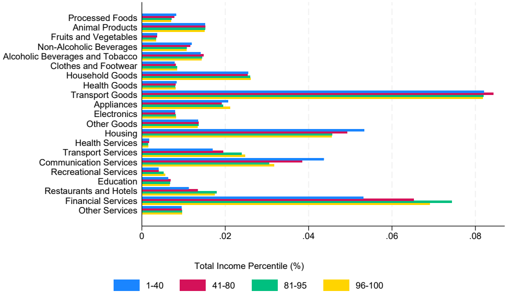
(b) Indirect Linkages: $\tilde{\Psi}_{jk} - \tilde{\Omega}_{jk} - I$

Exposure Matrices: Expenditure Shares, Direct and Indirect

$$d \ln \mathcal{Y}_c \approx -\sum_{i\in\mathcal{N}} b_{ci}\left(\sum_{j\in\mathcal{N}} \tilde{\Psi}_{ij}\left(d\ln\mu_j + \sum_{k\in\mathcal{N},\mathcal{F}} \tilde{\Omega}_{jk}\, d\ln\tau_{jk}\right) + \sum_{f\in\mathcal{F}} \tilde{\Psi}_{if}\, d\ln w_f\right) + \ldots$$



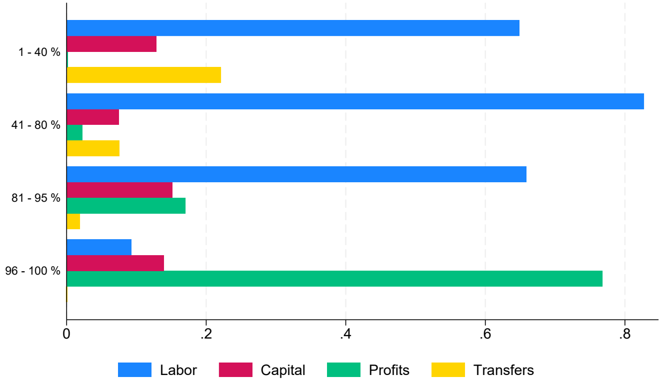
(a) Direct: b_{ci}



(b) Direct+Indirect: $\sum_i b_{ci}\tilde{\Psi}_{ij}$

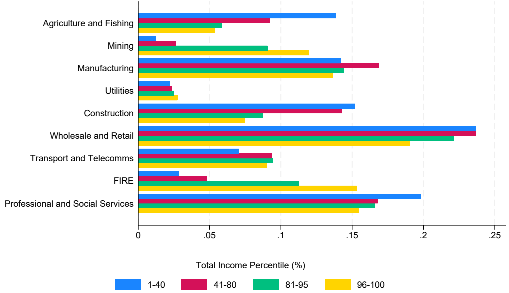
Exposure Matrices: Income by Source

$$d\ \ln \mathcal{Y}_c \approx -\sum_{i\in\mathcal{N}} b_{ci} d\ln p_i + \sum_{f\in\mathcal{F}} \Theta_{cf} d\ln w_f + \sum_{i\in\mathcal{N}} \Theta_{ci} d\ln \pi_i + \Theta_{ct} d\ln T$$

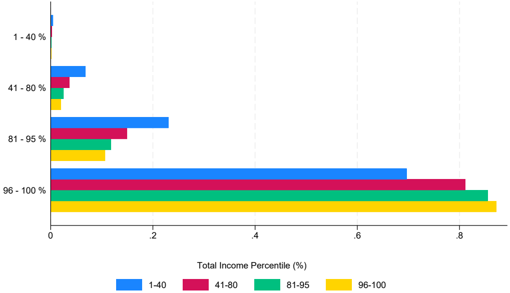


Exposure Matrices: Labor Shares from Employer-Employee

$$\text{d}\ \ln \mathcal{Y}_c \approx -\sum_{i\in\mathcal{N}} b_{ci}\, d\ln p_i + \sum_{f\in\mathcal{F}} \Theta_{cf}\, d\ln w_f + \sum_{i\in\mathcal{N}} \Theta_{ci}\, d\ln \pi_i + \Theta_{ct}\, d\ln T$$



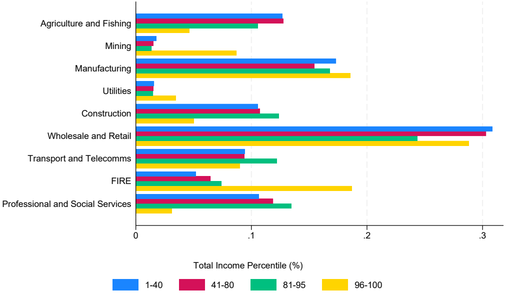
(a) Across Sectors



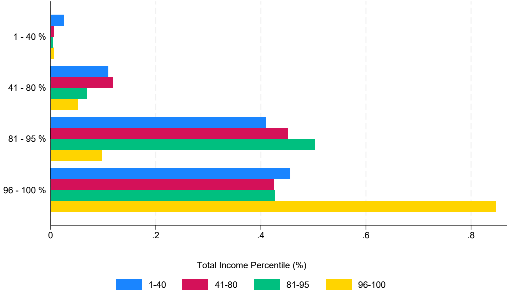
(b) Across Firm Size Bins

Exposure Matrices: Profit Shares from Ownership Registry

$$d \ln \mathcal{Y}_c \approx -\sum_{i\in\mathcal{N}} b_{ci}\, d\ln p_i + \sum_{f\in\mathcal{F}} \Theta_{cf}\, d\ln w_f + \sum_{i\in\mathcal{N}} \Theta_{ci}\, d\ln \pi_i + \Theta_{ct}\, d\ln T$$



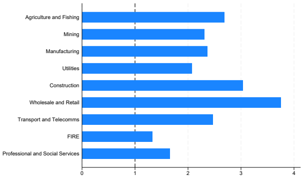
(a) Across Sectors



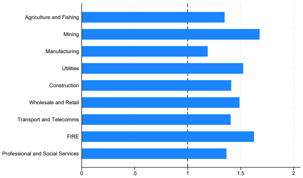
(b) Across Firm Size Bins

Wedge Estimates: Mean Across Sectors and Firm Size

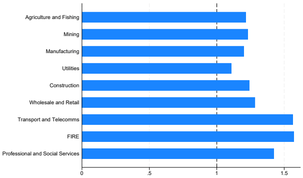
Sector



(a) Labor

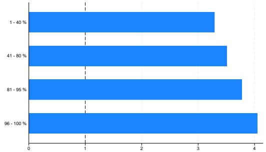


(b) Capital

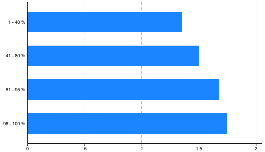


(c) Materials

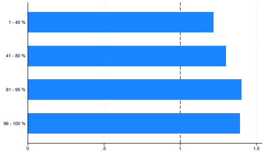
Firm Size



(d) Labor



(e) Capital



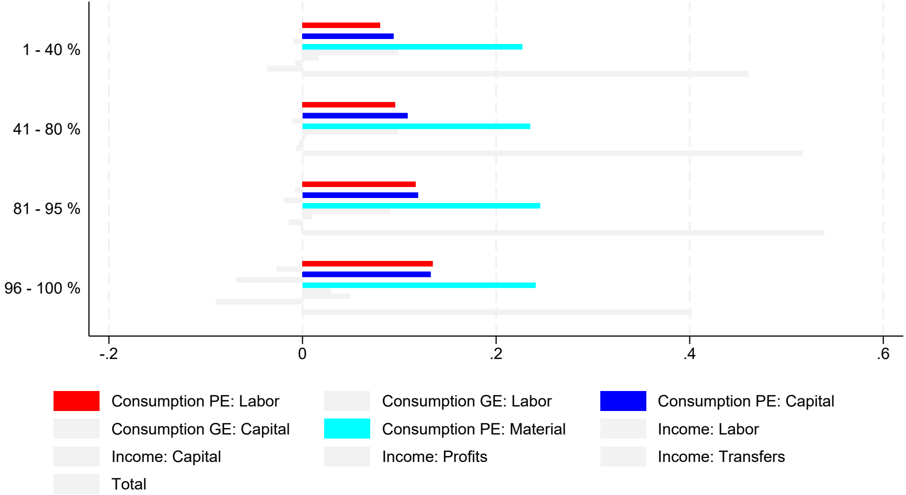
(f) Materials

Today

1. Introduction
2. Theoretical Framework and Incidence Formula
3. Distortions Estimation Strategy
4. Data
5. Exposure Matrices and Wedges
6. **The Incidence of Distortions**
7. Next Steps

PE Consumption Price Gains: Somewhat Unevenly Distributed

$$-\sum_{i\in\mathcal{N}} b_{ci}\, d\ln p_i = -\sum_{i\in\mathcal{N}} b_{ci}\left[\sum_{j\in\mathcal{N}} \tilde{\Psi}_{ij}\left(d\ln\mu_j + \sum_{k\in\mathcal{N},\mathcal{F}} \tilde{\Omega}_{jk}\, d\ln\tau_{jk}\right) + \sum_{f\in\mathcal{F}} \tilde{\Psi}_{if}\, d\ln w_f\right]$$

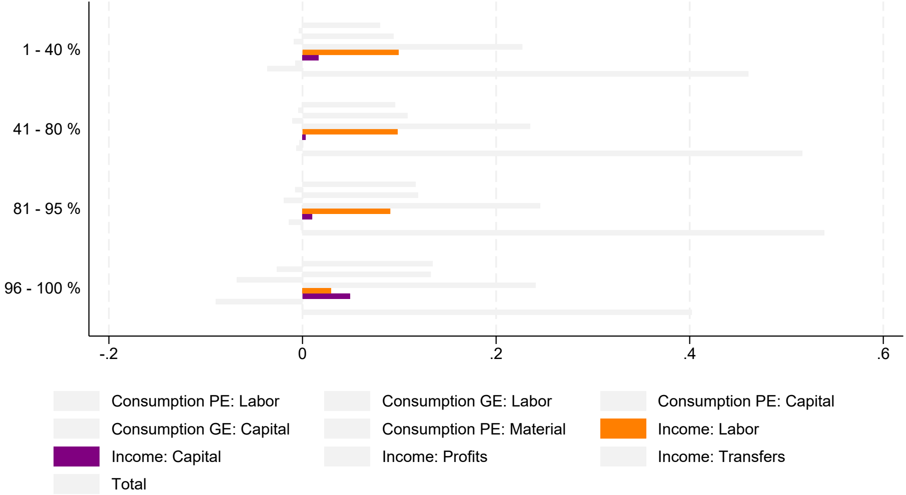


GE Consumption Price Losses: Harms the Richest 4%

$$- \sum_{i\in\mathcal{N}} b_{ci}\, d\ln p_i = -\sum_{i\in\mathcal{N}} b_{ci} \left[\sum_{j\in\mathcal{N}} \tilde{\Psi}_{ij} \left(d\ln\mu_j + \sum_{k\in\mathcal{N},\mathcal{F}} \tilde{\Omega}_{jk}\, d\ln\tau_{jk} \right) + \sum_{f\in\mathcal{F}} \tilde{\Psi}_{if}\, d\ln w_f \right]$$

Factor Income Gains: Labor (Capital) Benefits the Poor (Rich)

$$\sum_{f\in\mathcal{F}}\Theta_{cf}\,d\ln w_f=\sum_{f\in\mathcal{F}}\Theta_{cf}\left[-\sum_{l,m\in\mathcal{F},\mathcal{N}}\lambda_l\Omega_{lm}(d\ln\mu_l+d\ln\tau_{lm})\Psi_{mf}+\sum_{c\in\mathcal{C}}d\chi_c\sum_{k\in\mathcal{N}}b_{ck}\Psi_{kf}\right]$$



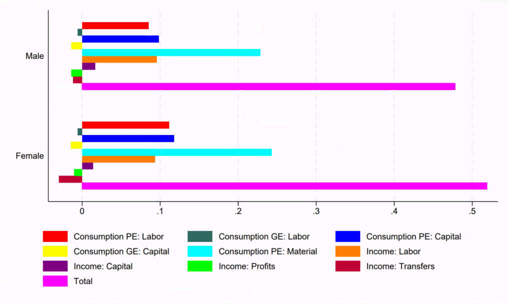
Non-Factor Income: Transfers (Profits) Harm the Poor (Rich)

$$\sum_{i\in\mathcal{N}} \Theta_{ci} d\ln \pi_i = \sum_{i\in\mathcal{N}} \Theta_{ci}\left[d\ln\lambda_i + \frac{\lambda_i}{\pi_i}\sum_{j\in\mathcal{F},\mathcal{N}} \Omega_{ij}(d\ln\mu_i + d\ln\tau_{ij} - d\ln\tilde{\Omega}_{ij}) - \frac{\lambda_i}{\pi_i} dT_i\right]$$

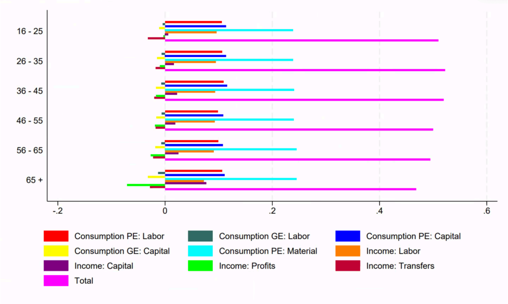
Distortion Reductions Benefit The Richest Relatively Less

$$d\ \ln \mathcal{Y}_c = -\sum_{i \in \mathcal{N}} b_{ci} d \ln p_i + \sum_{f \in \mathcal{F}} \Theta_{cf} d \ln w_f + \sum_{i \in \mathcal{N}} \Theta_{ci} d \ln \pi_i + \Theta_{ct} d \ln T$$

Distortions Harm Women and Young More



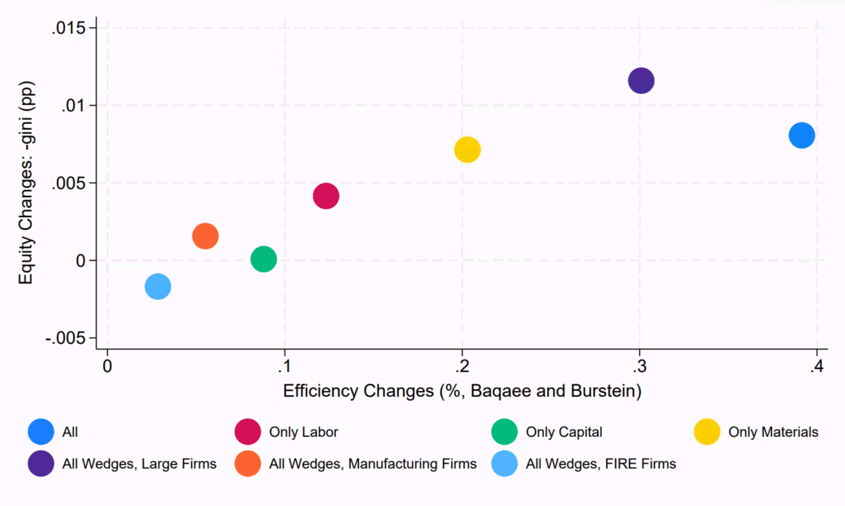
(a) Across Gender



(b) Across Age

- Pro-Rich: Driven by income channels (factor + non-factor) and consumption GE
- Pro-Male: Driven mostly by non-factor income channel
- Pro-Old: Driven mostly by non-factor income channel

Equity Efficiency Tradeoffs of Reducing Different Wedges?



Next Steps

- Further enrich construction of micro-level national accounts
    - Household-level analysis, richer demographics/race/immigration status
    - Add pension data
    - Worker skill levels by education
    - More granular CES nests and CD materials breakdown
- Bring wedges to zero via full iterative solution
- Richer counterfactuals and takeaways
    - Tax versus non-tax wedges
    - Which wedges reinforce each other and which are countervailing?
    - How misleading would results be if we only had access to typically-available data?
    - What drives incidence heterogeneity?

Theoretical Framework: (Non-Factor) Income-Based Exposure ▸ Return

- Changes in distortion revenue via ownership

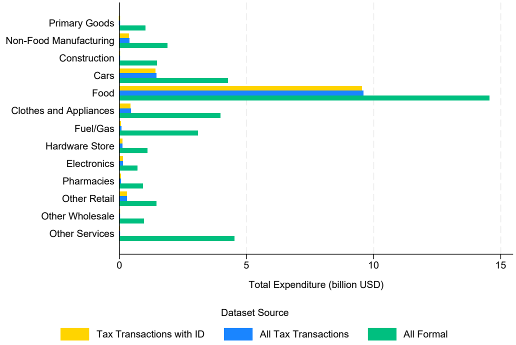
$$d \ln \pi_i = d \ln \lambda_i + \frac{\lambda_i}{\pi_i} \sum_{j \in \mathcal{F}, \mathcal{N}} \Omega_{ij} (d \ln \mu_i + d \ln \tau_{ij} - d \ln \tilde{\Omega}_{ij}) - \frac{\lambda_i}{\pi_i} dT_i$$

  - Change in Domar weights ($d \ln \lambda_i$): Similar to $d \ln w_f$, as factor is firm with no inputs

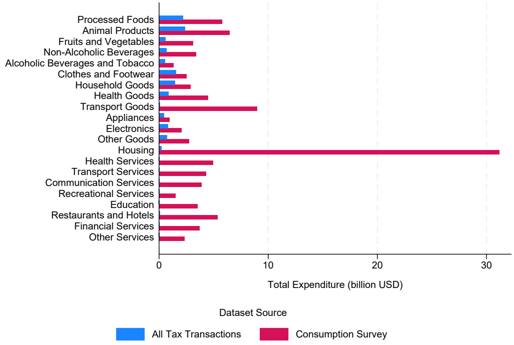
- Changes in distortion revenue via transfers

$$d \ln T = \sum_{i \in \mathcal{C}, \mathcal{N}} \frac{\lambda_i T_i}{T} (d \ln \lambda_i + d \ln T_i)$$

Firm-to-Individual Consumption: Data Coverage Return



(a) Across Retailer Types



(b) Across Products

Implementation Challenges Return

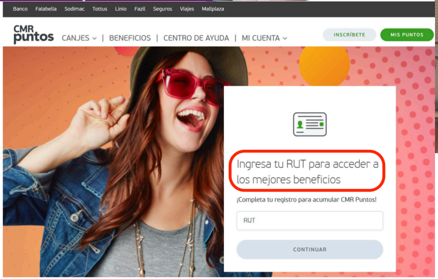
1. Filling in Administrative Data Gaps
   - Predict non-F2I consumption, informal income, government transfers from consumption (EPF 21/22, 44k) and employment (ESI 20–22, 300k) surveys
   - Statistically match EPF/ESI individuals to admin data using F2I consumption /formal income patterns within characteristic bins (Blanchet et al. 2023)
   - Carry across match values, residual after hitting total formal expenditure = informal

Implementation Challenges Return

1. Filling in Administrative Data Gaps
   - Predict non-F2I consumption, informal income, government transfers from consumption (EPF 21/22, 44k) and employment (ESI 20–22, 300k) surveys
   - Statistically match EPF/ESI individuals to admin data using F2I consumption /formal income patterns within characteristic bins (Blanchet et al. 2023)
   - Carry across match values, residual after hitting total formal expenditure = informal

2. Computational Burden of Large Matrices
   - E.g. Leontief inverse of I-O matrix $\Psi = (I - \Omega)^{-1}$ at firm product level (1.3 million)
   - Matrices are sparse $\Rightarrow$ Leverage computational advancements on sparse matrices
   - Approximate inverse with power series, $\Psi = I + \Omega + \Omega^2 + \Omega^3 + \ldots$, computing smaller (not $\Omega \times \Omega$) operations iteratively, $\Psi b = b + \Omega b + \Omega(\Omega b) + \Omega(\Omega(\Omega b)) + \ldots$
   - $\Rightarrow$ Compute the Leontief inverse of the firm-product IO matrix in 7 seconds

Source of Firm-to-Individual Consumption Data: Examples



Computational Burden of Large Matrices

- These matrices need inverting and they are big!
  - E.g., need compute Leontief inverse of input-output matrices $\Psi = (I - \Omega)^{-1}$
  - Input-output matrix Ω is defined at firm-product level, i.e., around 1.3 million

- We invert and multiply these massive matrices by exploiting the following:
  1. Matrices are sparse $\Rightarrow$ Leverage computational advancements on sparse matrices
  2. Approximate Leontief inverse with power series, e.g., $\Psi = I + \Omega + \Omega^2 + \Omega^3 + \dots$
  3. Leverage linear algebra by computing smaller operations iteratively
     - Consider a column vector b of dimension $N \times 1$, with N=number of firm-products
     - Use b to reduce the burden of multiplying Ω times Ω
     - E.g., $\Psi b = b + \Omega b + \Omega(\Omega b) + \Omega(\Omega(\Omega b)) + \dots$

- $\Rightarrow$ Compute the Leontief inverse of the firm-product IO matrix in 7 seconds