Trade and Inequality:
From Theory to Estimation

Elhanan Helpman
Harvard

Oleg Itskhoki
Princeton

Marc Muendler
UC San Diego

Stephen Redding
Princeton

August 2015

Motivation

- Neoclassical trade theory (H-O, SF, R)
  - sector-level comparative advantage
  - focus on "**between**" effects
- New trade theory
  - Krugman: intra-industry trade
  - Melitz: firm-level comparative advantage
  - focus on "**within**" effects
- Trade and inequality
  - Heavily influenced by H-O framework
  - Empirically has limited explanatory power
- "New view" of trade and inequality
  - link wages to firm performance
  - **within**-industry, **between**-firm

This Paper

- Uses linked employee-employer data for Brazil from 1986-98
  - Distribution of wages across workers and firms
  - Firm trade participation

- Establishes stylized facts about Brazilian wage inequality
  - within sector-occupations
  - for workers with similar observables (residual inequality)
  - between firms

- Develops a structural model to quantify the role of firm heterogeneity in wage inequality
  - extension of HIR (2010)
  - a model of *within*-sector, *between*-firm *residual* inequality
  - wages and employment vary with firm productivity and trade participation

Related Literature

- Long and large tradition in labor literature
- "New view" empirics:
  - Bernard and Jensen (1995)...
  - Verhoogen (2008)
  - Amity and Davis (2011)
  - AKM (1999) estimation used in trade context
- "New view" theory:
  - Feenstra and Hanson (1999)...
  - Yeaple (2005)...
  - Egger and Kreickemeier (2009)
  - HIR (2010)...

DATA

Brazilian RAIS Data

- Matched employer-employee data from 1986–1998
  - All workers employed in the formal sector
  - Focus on the manufacturing sector
  - Observe firm, industry and occupation
  - Observe worker education (high school, college degree), demographics (age, sex) and experience (employment history)
  - 5 aggregate and 350 disaggregate occupations
  - 13 aggregate and (from 1994) 250 disaggregate sectors

- Over the period 1986-1998 as a whole, our sample includes more than 7 million workers and 100,000 firms in every year

- Trade transactions data from 1986-1998
  - Merged with the matched employer-employee data
  - Observe firm exports and export products and destinations

STYLIZED FACTS

Within and Between Inequality

Sector-occupation bins

A. Main Period	Level (%) 1994	Change (%) 1986–95
Within occupation	82	92
Within sector	83	73
Within sector-occupation	68	66
Within detailed-occupation	61	60
Within sector–detailed-occupation	56	54
B. Late Period	**1994**	**1994–98**
Within detailed-sector –detailed-occupation	47	141

Fact 1

Within sector-occupation component of wage inequality accounts for over 2/3 of both level and growth of wage inequality

▸ show within regions

Residual Inequality

Conditional on worker observables

	Level (%) 1994	Change (%) 1986–95
Residual wage inequality	59	49
— within sector-occupation	89	91

Fact 2

(i) *Residual inequality is at least as important as worker observables for both level and growth of wage inequality*

(ii) *Almost all residual inequality is within sector-occupations*

Between-firm Inequality

- Mincer log-wage regression with firm fixed effect:

$$w_{it} = z_{it}'\vartheta_{\ell t} + \psi_{j\ell t} + \nu_{it}$$

  - i worker
  - j firm
  - ℓ sector-occupation bin

- $\psi_{j\ell t}$ firm fixed effect includes:
  - Returns to unobserved skill (workforce composition)
  - Worker rents (differences in wage for same workers)
  - Match effects

- Decomposition of within inequality:
  - Observables: $\operatorname{var}\left(z_{it}'\hat{\vartheta}_{\ell t}\right)$
  - Between-firm component: $\operatorname{var}\left(\hat{\psi}_{j\ell t}\right)$
  - Covariance: $\operatorname{cov}\left(z_{it}'\hat{\vartheta}_{\ell t}, \hat{\psi}_{j\ell t}\right)$
  - Within-firm component: $\operatorname{var}\left(\hat{\nu}_{it}\right)$

Between-firm Inequality

Within sector-occupation bins

	UNCONDITIONAL FIRM WAGE COMPONENT, $\psi^{U}_{j\ell t}$		CONDITIONAL FIRM WAGE COMPONENT, $\psi^{C}_{j\ell t}$	
	Level (%) 1994	Change (%) 1986–1995	Level (%) 1994	Change (%) 1986–1995
Between-firm wage inequality	55	115	39	86
Within-firm wage inequality	45	−15	37	−11
Worker observables			13	2
Covar observables–firm effects			11	24

Fact 3

Between-firm component account for about half of level and the majority of growth of within sector-occupation wage inequality

Between-firm Inequality

Size and exporter wage premia

	UNCONDITIONAL FIRM WAGE COMPONENT, $\hat{\psi}_{jt}^{U}$	CONDITIONAL FIRM WAGE COMPONENT, $\hat{\psi}_{jt}^{C}$
Firm Employment Size	0.122^{***}	0.104^{***}
	(0.010)	(0.009)
Firm Export Status	0.262^{***}	0.168^{***}
	(0.042)	(0.024)
Sector Fixed Effects	yes	yes
Within R-squared	0.17	0.13
Observations	91,410	91,410

Fact 4

Larger firms on average pay higher wages; exporters on average pay higher wages even after controlling for size. The remaining variation in wages is substantial.

STRUCTURAL MODEL

Model: Extension of HIR

1. Melitz (2003) product market:

$$R = \Upsilon A y^{\beta}, \qquad \Upsilon \in \{1, \Upsilon_x > 1\}$$

2. Heterogeneity in fixed cost of exports: $e^{\varepsilon} F_x$

3. Complementarity between productivity and worker ability:

$$y = e^{\theta} H^{\gamma} \bar{a}, \qquad \gamma < 1$$

4. Unobserved heterogeneity and costly screening:

$$e^{-\eta} C \frac{(a_c)^{\delta}}{\delta} \qquad \Rightarrow \qquad \bar{a} = \frac{k}{k-1} a_c$$

5. DMP search friction (cost b per worker) and wage bargaining:

$$W = \frac{\beta\gamma}{1+\beta\gamma} \frac{R}{H} = b \cdot (a_c)^k$$

Econometric Model

- Empirical model of $X_j = \{h_j, w_j, \iota_j\}_j$:

$$\begin{cases} h_j = \alpha_h + \mu_h \cdot \iota_j + u_j, \\ w_j = \alpha_w + \mu_w \cdot \iota_j + \zeta u_j + v_j, \\ \iota_j = \mathbb{I}\{z_j \geq f\} \end{cases}$$

- Distributional assumption:

$$(u_j, v_j, z_j)' \sim \mathcal{N}(\mathbf{0}, \Sigma), \qquad \Sigma = \begin{pmatrix} \sigma_u^2 & & \\ 0 & \sigma_v^2 & \\ \rho_u \cdot \sigma_u & \rho_v \cdot \sigma_v & 1 \end{pmatrix}$$

- **Selection** (ρ_u, ρ_v) versus **Market access** (μ_h, μ_w)

- Theoretical restriction: $\mu_h, \mu_w > 0$

Identification

1. Maximum Likelihood
   — under additional orthogonality assumption between structural productivity shocks θ and η:

$$\zeta \leq \frac{\mu_w}{\mu_h} \leq \zeta + \frac{\sigma_v^2}{(1+\zeta)\sigma_u^2}$$

2. GMM Bounds
   — based on a subset of moments

3. Semi-parametric estimation
   — using alternative instruments for export participation

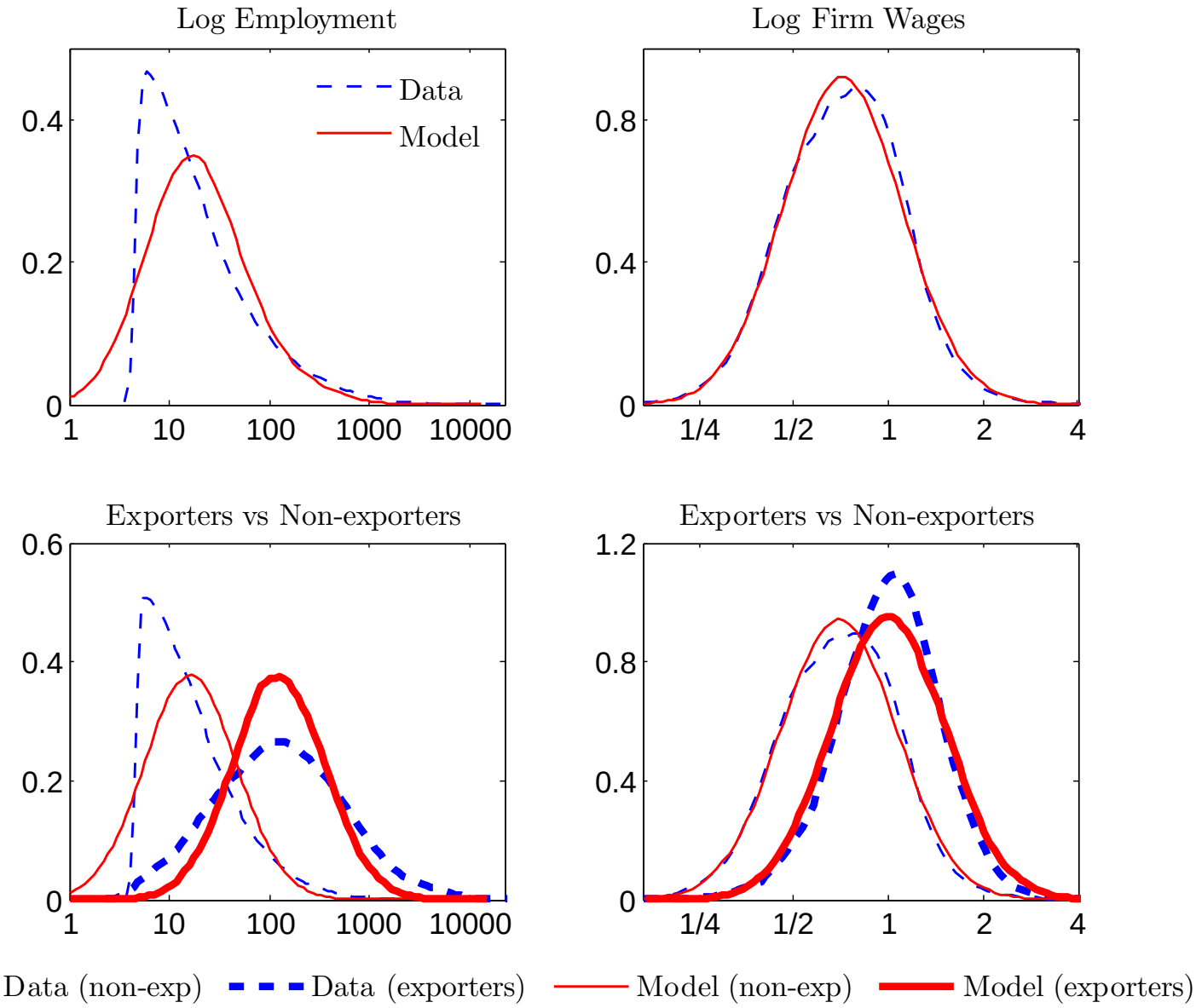
RESULTS

Coefficient Estimates 1994

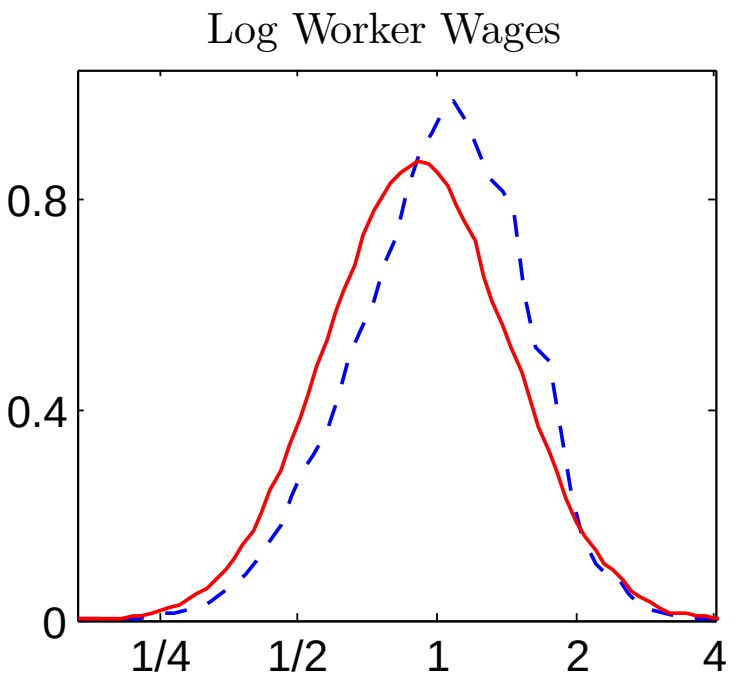
	Coefficient	Std Error
μ_h	1.992	0.019
μ_w	0.197	0.022
ρ_u	0.023	0.004
ρ_v	0.199	0.024
f	1.341	0.006

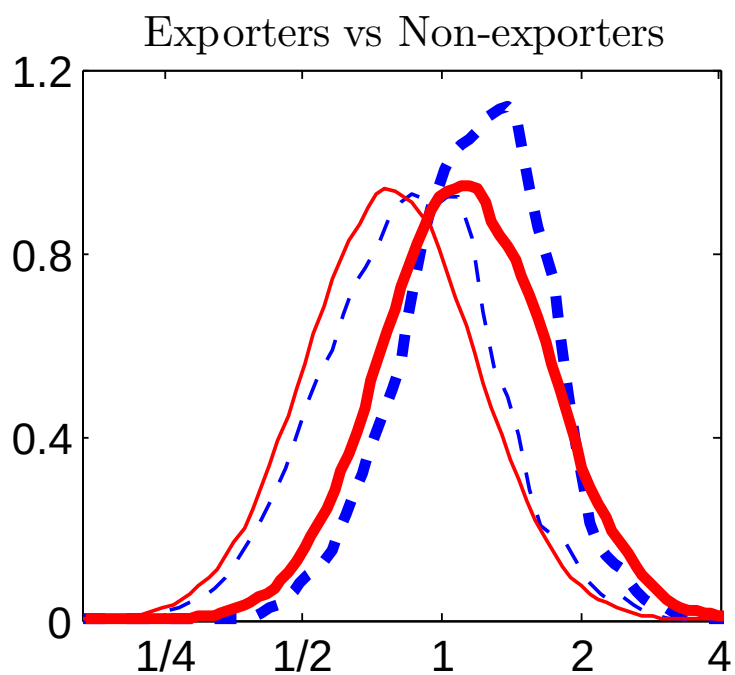
Note: Maximum likelihood estimates and robust (sandwich-form) asymptotic standard errors (see the online supplement) for 1994. Number of observations (firms): 91,410.

Employment and Wage Distributions



Worker Wage Distribution





Counterfactuals

- Estimated model:

$$\begin{cases} h_j = \alpha_h + \mu_h \cdot \iota_j + u_j, \\ w_j = \alpha_w + \mu_w \cdot \iota_j + \zeta u_j + v_j, \qquad (u_j, v_j, z_j)' \sim \mathcal{N}(\mathbf{0}, \Sigma) \\ \iota_j = \mathbb{I}\{z_j \geq f\} \end{cases}$$

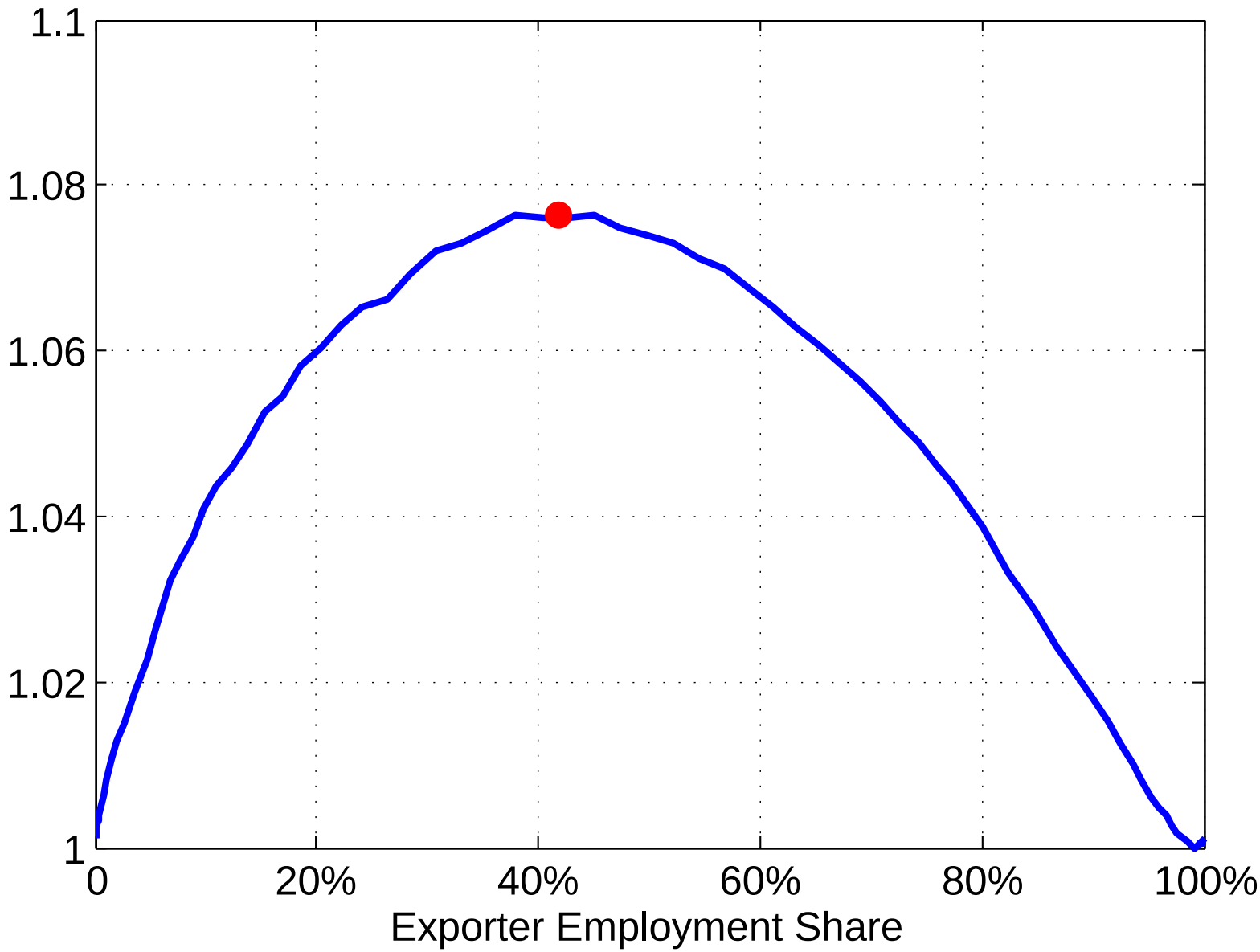
- Parameters (μ_h, μ_w, f) form a **sufficient statistic**:

$$f = \frac{1}{\sigma}\left[\alpha_f + \log F_x - \log\left(\Upsilon_x^{\frac{1-\beta}{\Gamma}} - 1\right)\right]$$

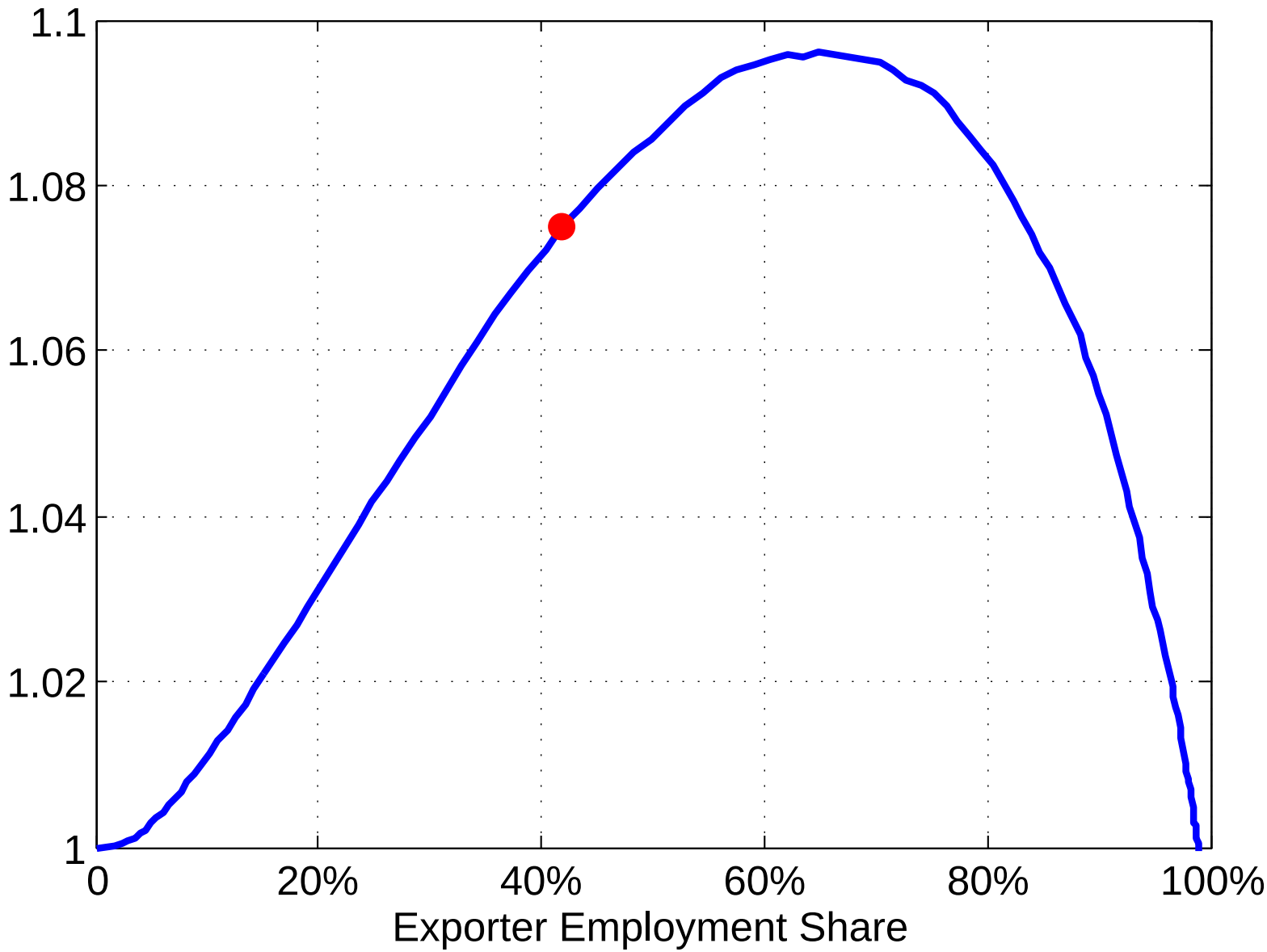
$$\mu_h + \mu_w = \Upsilon_x^{\frac{1-\beta}{\Gamma}}, \qquad \Upsilon_x = 1 + \tau^{-\frac{\beta}{1-\beta}} \frac{A_x}{A_d}$$

- Two counterfactuals: variation in F_x and τ

Variation in Fixed Export Cost



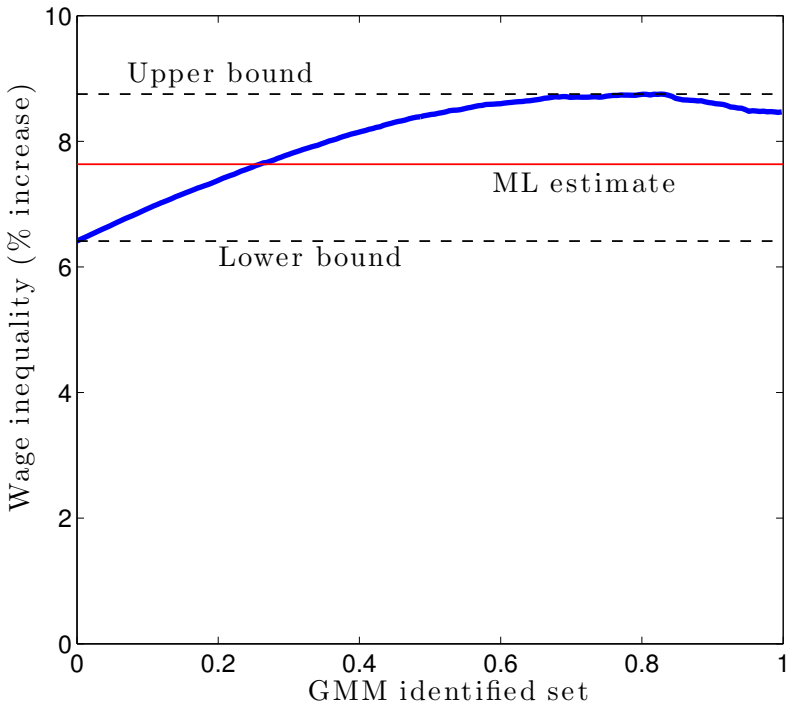
Variation in Variable Trade Cost



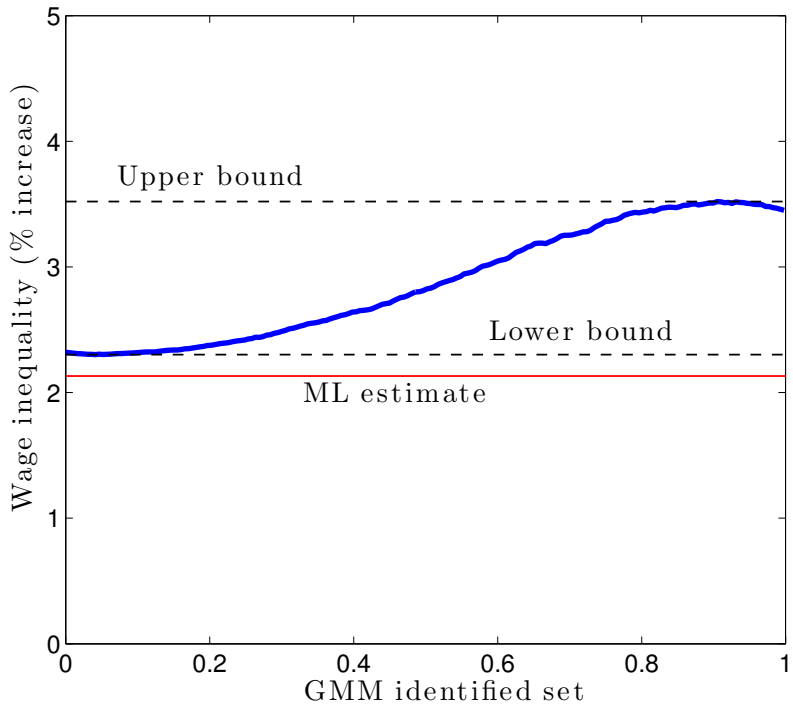
GMM BOUNDS

GMM Bounds

(a) Autarky counterfactual



(b) τ counterfactual



- Autarky bounds: $[6.6\%, 9.0\%]$ vs ML estimate 7.6%
- τ bounds: $[2.3\%, 3.5\%]$ vs ML estimate 2.2%

Semi-parametric

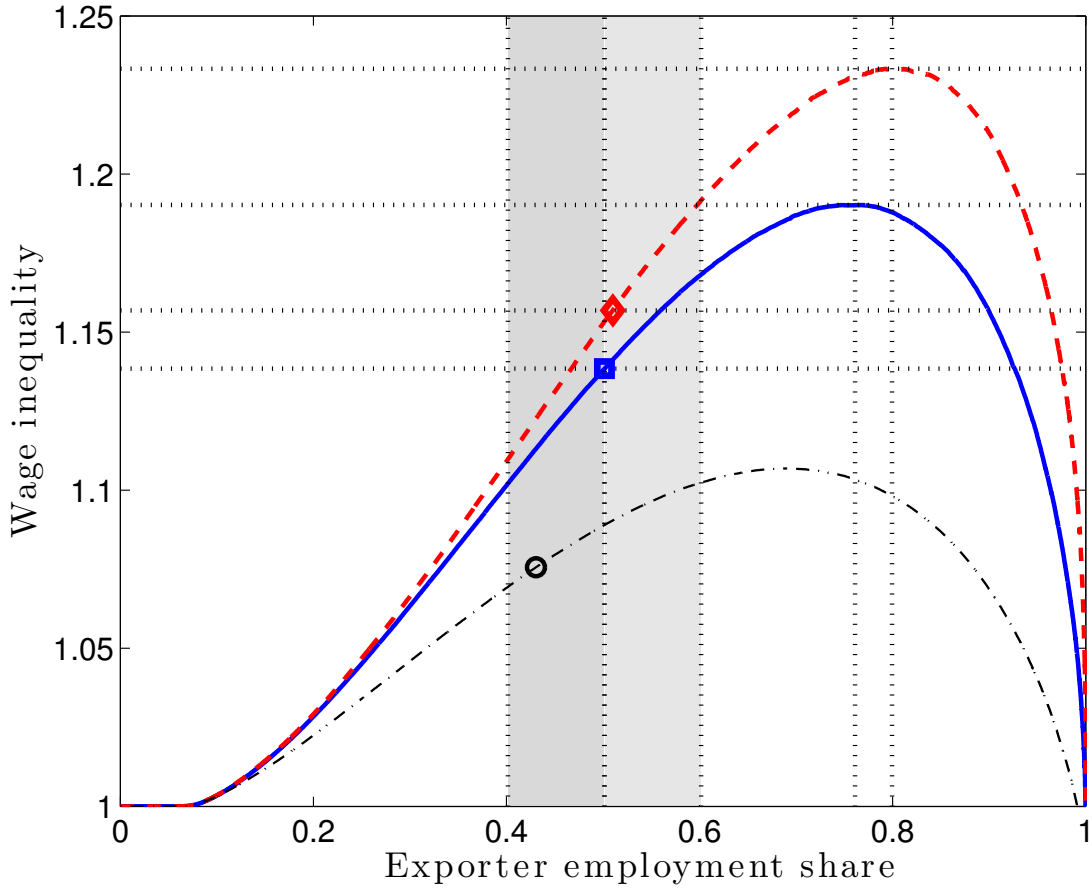
Two-stage estimation

	Business Procedures	Foreign Workers			Both Excluded Variables
		Firm	Meso	Layoff	
	(1)	(2)	(3)	(4)	(5)
Panel A: Selection					
Business Procedures	-0.139^{***}	—	—	—	-0.139^{***}
	(0.025)				(0.025)
Foreign Worker	—	0.070^{***}	0.129^{***}	0.022^{**}	0.019^{*}
		(0.008)	(0.034)	(0.010)	(0.010)
First-stage F-statistic	30.60	85.96	14.56	4.36	37.36
[p-value]	[0.000]	[0.000]	[0.000]	[0.037]	[0.000]
Panel B: Employment					
Employment premium (μ_h)	2.004^{***}	1.997^{***}	2.032^{***}	2.039^{***}	2.012^{***}
	(0.031)	(0.034)	(0.034)	(0.033)	(0.032)
Second-stage F-statistic	16.57	83.40	2.69	2.18	14.37
[p-value]	[0.000]	[0.000]	[0.045]	[0.088]	[0.000]
Panel C: Wages					
Wage premium (μ_w)	0.361^{***}	0.343^{***}	0.312^{***}	0.356^{***}	0.361^{***}
	(0.016)	(0.015)	(0.012)	(0.016)	(0.017)
Second-stage F-statistic	4.07	59.70	171.67	2.30	4.00
[p-value]	[0.007]	[0.000]	[0.000]	[0.075]	[0.007]

MULTIDESTINATION

Multidestination Model

Counterfactuals



▸ Show the model

Conclusions

- Neoclassical trade theory emphasizes wage inequality between occupations and industries
- In contrast, new theories of firm heterogeneity and trade point to wage dispersion within occupations and industries
- Using matched employer-employee data for Brazil, we show:
  - Much of the increase in wage inequality since the mid-1980s has occurred within sector-occupations
  - Increased within-group wage inequality
  - Increased wage dispersion between firms
  - Between-firm wage dispersion related to trade participation
- Develop a framework for the structural estimation of a model with firm heterogeneity and wage dispersion across firms
- Use this framework to quantify the effect of trade on wage dispersion

Model Predictions

- A firm with idiosyncratic shock $\{\theta, \eta, \varepsilon\}$:

$$R(\theta,\eta,\varepsilon) = \kappa_r \Upsilon^{\frac{1-\beta}{\Gamma}} \left(e^{\theta}\right)^{\frac{\beta}{\Gamma}} \left(e^{\eta}\right)^{\frac{\beta(1-\gamma k)}{\delta\Gamma}}$$

$$H(\theta,\eta,\varepsilon) = \kappa_h \Upsilon^{\frac{(1-\beta)(1-k/\delta)}{\Gamma}} \left(e^{\theta}\right)^{\frac{\beta(1-k/\delta)}{\Gamma}} \left(e^{\eta}\right)^{\frac{\beta(1-\gamma k)(1-k/\delta)}{\delta\Gamma} - \frac{k}{\delta}}$$

$$W(\theta,\eta,\varepsilon) = \kappa_w \Upsilon^{\frac{k(1-\beta)}{\delta\Gamma}} \left(e^{\theta}\right)^{\frac{\beta k}{\delta\Gamma}} \left(e^{\eta}\right)^{\frac{k}{\delta}\left(1+\frac{\beta(1-\gamma k)}{\delta\Gamma}\right)}$$

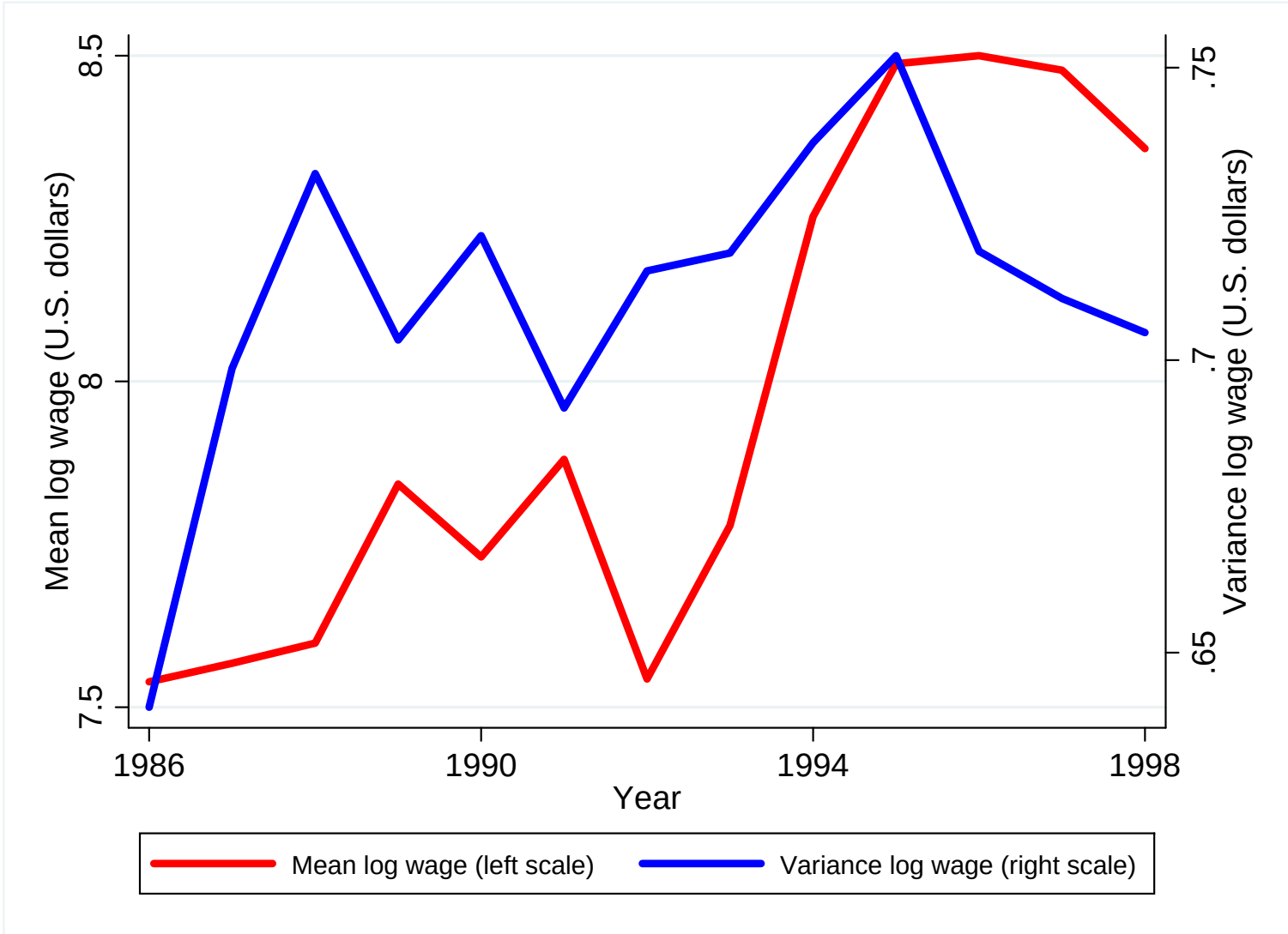
- Market access variable

$$\Upsilon = 1 + \iota \cdot \left(\Upsilon_x - 1\right), \qquad \Upsilon_x = 1 + \tau^{-\frac{\beta}{1-\beta}} \frac{A_x}{A_d}$$

- Selection into exporting

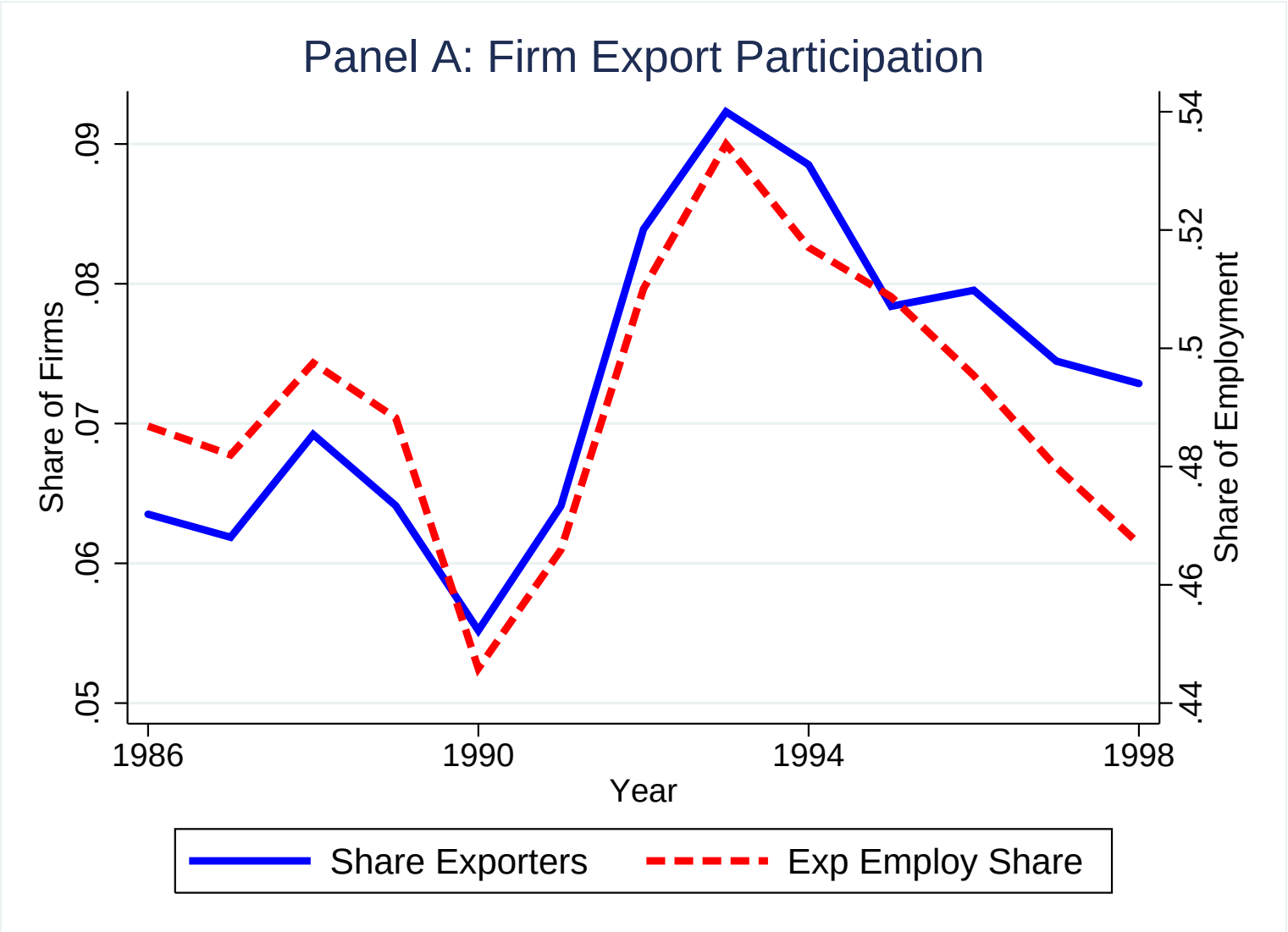
$$\iota = \iota(\theta,\eta,\varepsilon) = \mathbb{I}\left\{\kappa_\pi \left(\Upsilon_x^{\frac{1-\beta}{\Gamma}} - 1\right) \left(e^{\theta}\right)^{\frac{\beta}{\Gamma}} \left(e^{\eta}\right)^{\frac{\beta(1-\gamma k)}{\delta\Gamma}} \geq F_x e^{\varepsilon}\right\}$$

Wage Inequality



Back to slides

Trade Openness



Back to slides

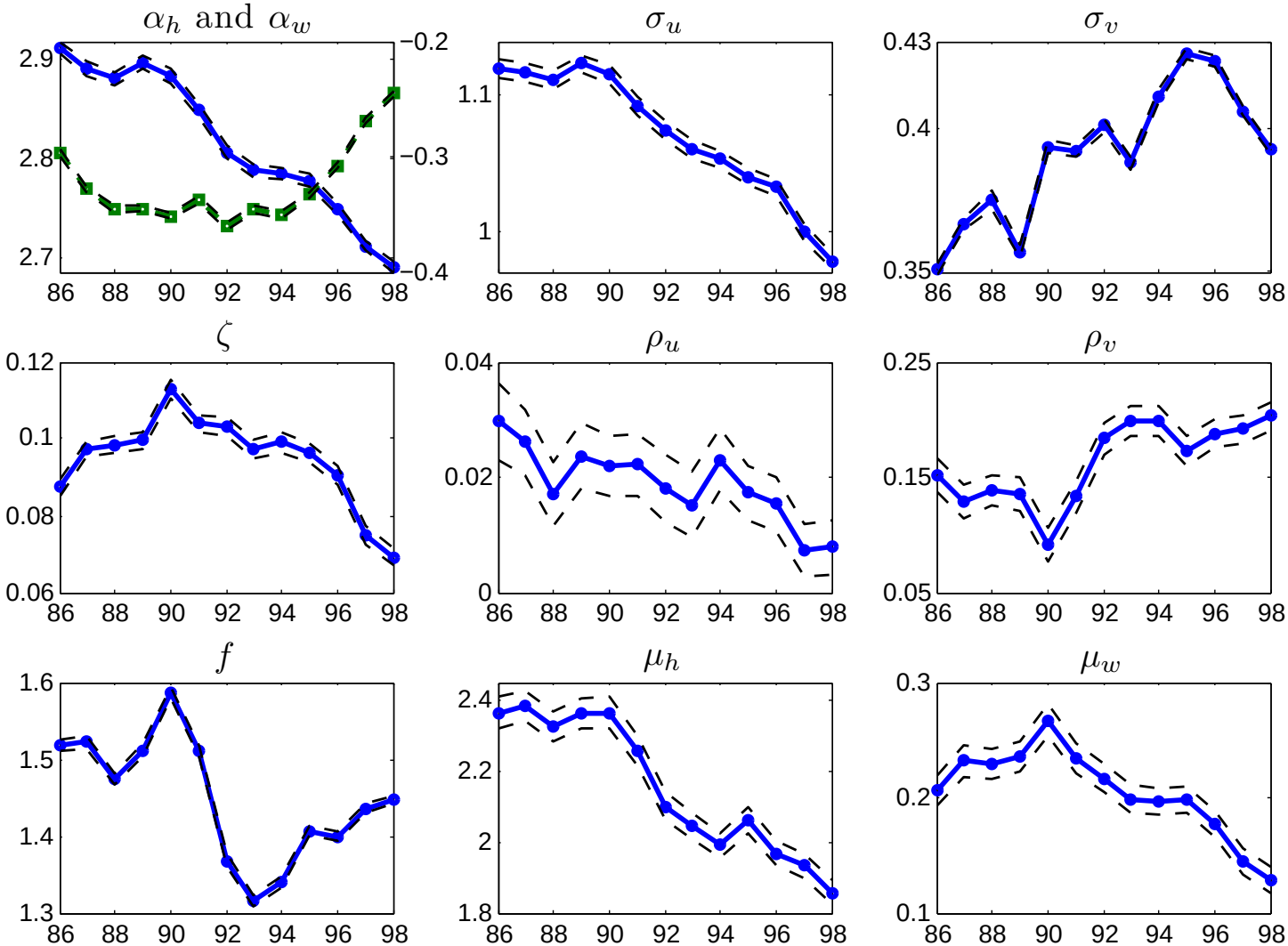
Regional Robustness

	OVERALL INEQUALITY		RESIDUAL INEQUALITY	
	Level 1994	Change 1986–95	Level 1994	Change 1986–95
Within sector-occupation	68	66	89	91
Within sector-occupation, São Paulo	64	49	89	71
Within sector-occupation-state	58	38	76	56
Within sector-occupation-meso	54	30	72	49

◂ Back to slides

Estimation Results

Parameters



Model Fit

Firm-level moments

	All firms	Non-exp.	Exporters
		DATA	
Mean h	2.96	2.78	4.82
Mean w	−0.33	−0.37	−0.01
Std deviation h	1.20	1.00	1.46
Std deviation w	0.43	0.43	0.38
Correlation h & w	0.33	0.24	0.32
Fraction of exporters	9.0%		
		MODEL	
Mean h	2.96	2.78	4.83
Mean w	−0.33	−0.37	0.00
Std deviation h	1.20	1.05	1.05
Std deviation w	0.43	0.42	0.42
Correlation h & w	0.32	0.25	0.24
Fraction of exporters	9.0%		

Model Fit

Worker wage dispersion

	Data	Model
Std deviation	0.42	0.46
— non-exporters	0.42	0.42
— exporters	0.35	0.42
Gini coefficient	0.23	0.25
90/10-ratio	2.95	3.23
— 90/50	1.63	1.80
— 50/10	1.81	1.80

Model Fit

Worker wage dispersion

	Data	Model
Std deviation	0.42	0.46
— non-exporters	0.42	0.42
— exporters	0.35	0.42
Gini coefficient	0.23	0.25
90/10-ratio	2.95	3.23
— 90/50	1.63	1.80
— 50/10	1.81	1.80

Size and exporter wage premia

	Data	Model
Employment premium	0.10	0.10
Exporter premium	0.16	0.16
R-squared	0.11	0.11

Likelihood Function

$$\mathcal{L}(\Theta|X_j) = \prod_j \mathbb{P}\{(h_j, w_j, \iota_j)|\Theta\}$$

$$\mathbb{P}\{(h_j, w_j, \iota_j)|\Theta\} =$$

$$\frac{1}{\sigma_u}\phi(\hat{u}_j)\frac{1}{\sigma_v}\phi(\hat{v}_j)\left[\Phi\left(\frac{f-\rho_u\hat{u}_j-\rho_v\hat{v}_j}{\sqrt{1-\rho_u^2-\rho_v^2}}\right)\right]^{1-\iota_j}\left[1-\Phi\left(\frac{f-\rho_u\hat{u}_j-\rho_v\hat{v}_j}{\sqrt{1-\rho_u^2-\rho_v^2}}\right)\right]^{\iota_j}$$

$$\hat{u}_j = \frac{h_j-\alpha_h-\mu_h\iota_j}{\sigma_u},$$

$$\hat{v}_j = \frac{(w_j-\alpha_w-\mu_w\iota_j)-\zeta(h_j-\alpha_h-\mu_h\iota_j)}{\sigma_v}$$

◂ back to slides

GMM Bounds

- We drop the orthogonality assumption and use the following set of moments:
  (a) conditional first moments: $\mathbb{E}\iota$, $\mathbb{E}\{h|\iota\}$ and $\mathbb{E}\{w|\iota\}$
  (b) unconditional second moments: $\mathrm{var}(h)$, $\mathrm{var}(w)$ and $\mathrm{cov}(h, w)$
  (c) size and exporter wage premia λ_s and λ_x, and R^2 from:

$$\mathbb{E}\{w|h, \iota\} = \lambda_0 + \lambda_s h + \lambda_x \iota$$

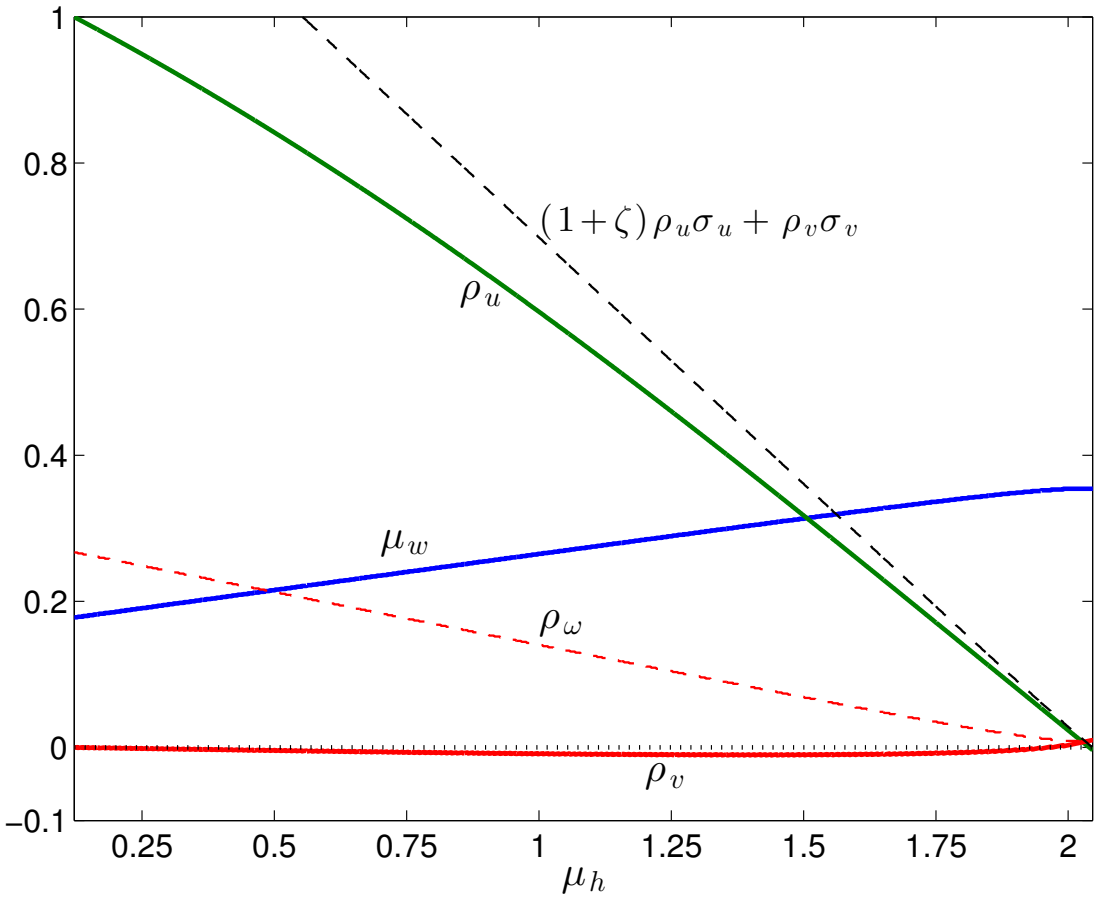
  — In addition, we impose $|\rho_u|, |\rho_v| < 1$ and $\sigma_u, \sigma_v > 0$
  — and $\mathrm{corr}\big((1+\zeta)u + v, z\big) = (1+\zeta)\rho_u\sigma_u + \rho_v\sigma_v > 0$
  — We check that $\mu_h, \mu_w > 0$

- This identifies a uni-dimensional interval in the 10-dimensional parameter space, the GMM identified set ▸ Show the iSet

- For each element of this set we conduct: (a) autarky and (b) variable trade cost counterfactual:
  — $\tau \uparrow$ to generate a 10p.p.$\downarrow$ in exporter employment share

GMM Identified Set

- Main idea: $\bar{h}_1 - \bar{h}_0 = \mu_h + \rho_u \sigma_u (\lambda_1 - \lambda_0)$



back to slides

Multidestination Model

$$\begin{cases} h &= \alpha_h + \mu_{h,1}\iota_1 + (\mu_{h,2} - \mu_{h,1})\iota_2 + (\mu_{h,3} - \mu_{h,2})\iota_3 + u, \\ w &= \alpha_w + \mu_{w,1}\iota_1 + (\mu_{w,2} - \mu_{w,1})\iota_2 + (\mu_{w,3} - \mu_{w,2})\iota_3 + \zeta u + v, \\ \iota_\ell &= \mathbb{I}\{f_{\ell-1} \le z \le f_\ell\}, \qquad \ell = 1,2,3, \end{cases}$$

$$\mu_{h,\ell} = \frac{\delta - k}{\delta} \log \Upsilon_{x,\ell}^{\frac{1-\beta}{\Gamma}}, \qquad \mu_{w,\ell} = \frac{k}{\delta - k}\mu_{h,\ell},$$

$$f_\ell = \frac{1}{\sigma}\Big[-\alpha_\pi + \log F_{x,\ell} - \log\left(\Upsilon_{x,\ell}^{\frac{1-\beta}{\Gamma}} - \Upsilon_{x,\ell-1}^{\frac{1-\beta}{\Gamma}}\right)\Big],$$

$$\Upsilon_x = 1 + \tau^{-\frac{\beta}{1-\beta}} \sum_{\ell=1,2,3} \iota_\ell \left(\frac{A_{x,\ell}}{A_d}\right)^{\frac{1}{1-\beta}}$$

$$\iota_\ell = \mathbb{I}\left\{ \kappa_\pi \left[\Upsilon_{x,\ell}^{\frac{1-\beta}{\Gamma}} - \Upsilon_{x,\ell-1}^{\frac{1-\beta}{\Gamma}}\right] \left(e^\theta\right)^{\frac{\beta}{\Gamma}} \left(e^\eta\right)^{\frac{\beta(1-\gamma k)}{\delta\Gamma}} \ge e^\varepsilon F_{x,\ell} \right\}, \qquad \ell = 1,2,3$$

◂ back to slides

Sectors

- Twelve aggregate sectors (IBGE) 1986-1998

	Industry	Empl't share (%)	Rel. mean log wage	Fraction Exporters	Exporter Empl't %
2	Non-metallic mineral products	5.5	−0.12	2.3	32.3
3	Metallic products	9.8	0.27	6.1	49.9
4	Machinery, equipment & instr.	6.6	0.38	12.3	54.1
5	Electrical & telecomm. equip.	6.0	0.37	11.8	56.3
6	Transport equipment	6.3	0.61	11.2	70.6
7	Wood products & furniture	6.5	−0.48	3.2	23.5
8	Paper, publishing & printing	5.4	0.14	2.5	30.6
9	Rubber, tobacco, leather & fur	7.0	−0.04	8.6	50.8
10	Chemical & pharma. products	9.9	0.40	11.2	50.6
11	Apparel & textiles	15.7	−0.32	2.5	34.8
12	Footwear	4.4	−0.44	12.2	65.7
13	Food, beverages & alcohol	16.9	−0.30	3.9	38.0

- More than 250 disaggregated industries (CNAE) 1994-1998

Occupations

- Five aggregate occupations 1986-1998

	Occupation	Employment share (%)	Relative mean log wage
1	Professional and Managerial	7.8	1.08
2	Skilled White Collar	11.1	0.40
3	Unskilled White Collar	8.4	0.13
4	Skilled Blue Collar	57.4	−0.15
5	Unskilled Blue Collar	15.2	−0.35

- More than 300 disaggregated occupations (CBO) 1986-1998