The Effects of Labor and Industrial Regulations in India: Evidence from the Plant Size Distribution

Amrit Amirapu and Michael Gechter

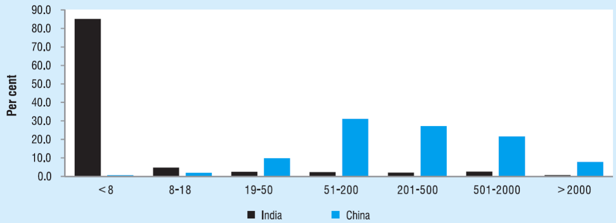
January 15th 2014

Motivation

- Conventional wisdom: “missing middle” among Indian firms

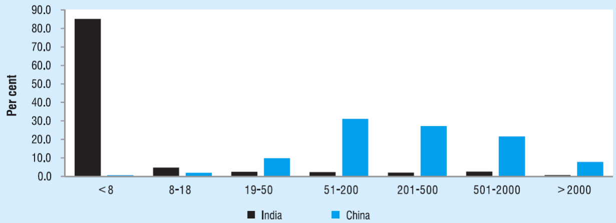
Motivation

- Conventional wisdom: "missing middle" among Indian firms
- India seems to have an unusually right-skewed distribution of firms:



Motivation

- Conventional wisdom: "missing middle" among Indian firms
- India seems to have an unusually right-skewed distribution of firms:



- This could have (ambiguous) implications for:
  - productivity
    - small firms are less productive and pay lower wages (e.g. Hasan and Jandoc, 2012; ADB, 2009)
    - learning/technological growth (Stiglitz)
  - inequality (in either direction)

Our Research Question and Contributions

- Qn: Could labor/industrial regulations play a role in producing the observed distribution (by constraining the size of enterprises)?

Our Research Question and Contributions

- Qn: Could labor/industrial regulations play a role in producing the observed distribution (by constraining the size of enterprises)?
- Not a new question...

Our Research Question and Contributions

- Qn: Could labor/industrial regulations play a role in producing the observed distribution (by constraining the size of enterprises)?
- Not a new question...
- Our contribution:
  - We use a methodology that has not been used on Indian data:
    - in particular we model these regulations as producing an increase in the labor costs of firms, and
    - we attempt to estimate the size of these costs - using distortions in the firm size distribution
  - We document a new phenomenon that may shed insight on the recent "casualisation" of the Indian work force

Our Research Question and Contributions

- Qn: Could labor/industrial regulations play a role in producing the observed distribution (by constraining the size of enterprises)?
- Not a new question...
- Our contribution:
  - We use a methodology that has not been used on Indian data:
    - in particular we model these regulations as producing an increase in the labor costs of firms, and
    - we attempt to estimate the size of these costs - using distortions in the firm size distribution
  - We document a new phenomenon that may shed insight on the recent "casualisation" of the Indian work force
- Note:
  - Limitations: data; not a welfare analysis
  - Preliminary... a lot to be done... feedback greatly desired...

Outline

Institutional Background

- Size-based regulations:
  - do not apply to firms below a certain size
  - Numerous relevant thresholds (**10**, 20, 50, 100, K = 5 crore)
  - **At 10:** Factories Act, ESI, Payment of Bonus/Gratuities, lots of paperwork, general sense of formality (eg: minimum wages)

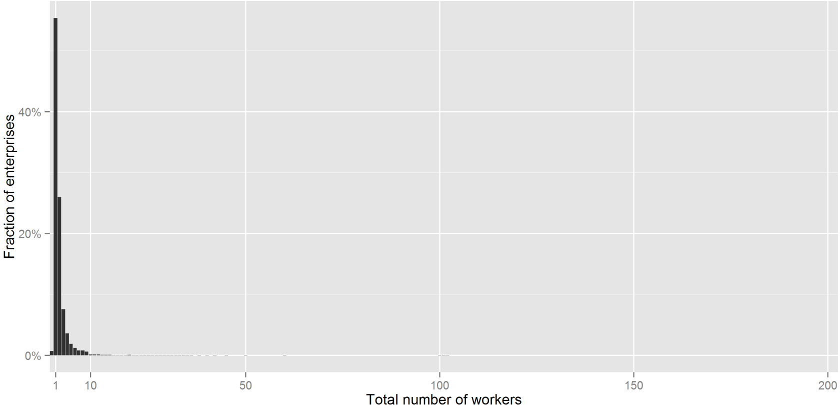
Literature Review

- Regulations and state-level outcomes in India: Besley and Burgess (2004), Bhattacharjea (2006 & 2009), and many many others
  - nearly all focused on IDA,
  - nearly all use (poor) variation in state laws
- Size-based regulations: Garicano, Lelarge and Van Reenen (2013), Gourio and Roys (2013)

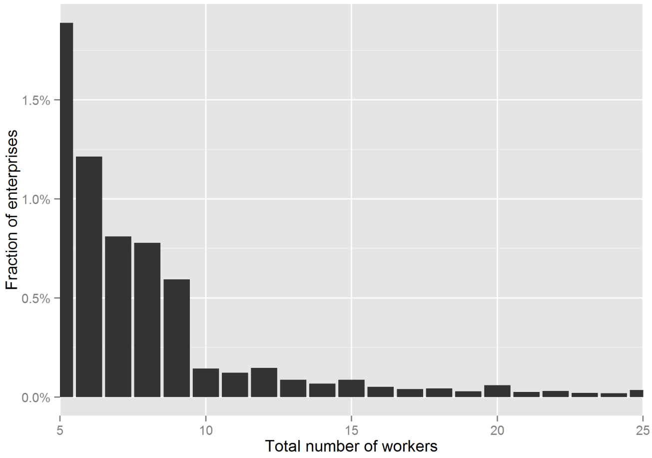
Data

- Economic Census of India (1990, 1998, 2005)
  - Intended to be an enumeration of all non-agricultural *enterprises* in India
  - Administered by state statistical offices
  - *Information is self-reported*, not tied to any other interaction with the government
  - very few variables

Distribution of Enterprises By Enterprise Size



Distribution of Enterprises By Enterprise Size



Distribution of Enterprises By Enterprise Size (log-scale)

Distribution of Enterprises By Enterprise Size (log-scale; linear fit)

Interstate Variation

Intertemporal Variation

Possible Explanations?

- One extreme:
  - Firms curb employment (and thus production) to avoid higher costs
- The other extreme:
  - Firms simply misreport employment (and are in fact relatively unconstrained)

Possible Explanations?

- One extreme:
  - Firms curb employment (and thus production) to avoid higher costs
- The other extreme:
  - Firms simply misreport employment (and are in fact relatively unconstrained)
- Alternative explanations:
  - Firms substitute permanent workers with non-permanent/contract/casual labor
  - Firms substitute labor for capital or higher-skill labor
  - Firms reduce employment through vertical disintegration/outsourcing of inputs

Theory

Based on Garicano, Lelarge and Van Reenen (2013)

- An individual firm's problem:

$$\pi(\alpha) = \max_{n} \alpha f(n) - wn$$

  - output, $y = \alpha f(n)$; $\quad n$: number of workers
  - Managerial ability: $\alpha \sim \phi(\alpha)$, defined on $[\underline{\alpha}, \infty)$

Theory

Based on Garicano, Lelarge and Van Reenen (2013)

- An individual firm's problem:

$$\pi(\alpha) = \max_{n} \alpha f(n) - wn$$

  - output, $y = \alpha f(n)$; n: number of workers
  - Managerial ability: $\alpha \sim \phi(\alpha)$, defined on $[\underline{\alpha}, \infty)$
- FOC:

$$\alpha = \frac{w}{f'(n)}$$

- If α has a power law distribution, so will $\chi(n)$.

Theory

Distribution with Size-Based Regulation

- With Size-Based Regulation:

$$\pi(\alpha) = \max_n \alpha f(n) - w\bar{\tau}n - \bar{k}$$

  - $\bar{\tau} = 1, \bar{k} = 0$ if $n \leq N$
  - $\bar{\tau} = \tau, \bar{k} = k$ if $n > N$; $\tau > 1$
- The result is that:
  - $\underline{\alpha} < \alpha < \alpha_1$: Unconstrained managers - choose low n ($<$ N)
  - $\alpha_1 < \alpha < \alpha_2$: Constrained managers - choose n = N and avoid regulation
  - $\alpha_2 < \alpha < \infty$: Taxed managers - choose high n ($>$ N) and bite the bullet

Theory

Distribution with Size-Based Regulation

- With Size-Based Regulation:

$$\pi(\alpha) = \max_{n} \alpha f(n) - w\bar{\tau}n - \bar{k}$$

  - $\bar{\tau} = 1, \bar{k} = 0$ if $n \leq N$
  - $\bar{\tau} = \tau, \bar{k} = k$ if $n > N$; $\tau > 1$
- The result is that:
  - $\underline{\alpha} < \alpha < \alpha_1$: Unconstrained managers - choose low n ($<$ N)
  - $\alpha_1 < \alpha < \alpha_2$: Constrained managers - choose n = N and avoid regulation
  - $\alpha_2 < \alpha < \infty$: Taxed managers - choose high n ($>$ N) and bite the bullet
- Now, adopting specific functional forms:
  - $f(n) = n^{\theta}$ & $\phi(\alpha) = c_{\alpha}\alpha^{-\beta_{\alpha}}$

Theoretical Densities

- And simplifying...

$$log\chi(n) = \begin{cases} logA - \beta log(n) & if \quad n \in [n_{\min}, N) \\ log(\delta_n) & if \quad n = N \\ 0 & if \quad n \in (N, n_u) \\ logA - \beta log(n) - \frac{\beta-1}{1-\theta} log(\tau) & if \quad n \geq n_u \end{cases}$$

- Key things are:
  - 1) power law in n 2) τ enters the log density additively

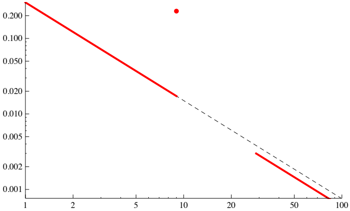
Theoretical Densities

- And simplifying...

$$
log\chi(n) = \begin{cases} logA - \beta log(n) & if \quad n \in [n_{\min}, N) \\ log(\delta_n) & if \quad n = N \\ 0 & if \quad n \in (N, n_u) \\ logA - \beta log(n) - \frac{\beta-1}{1-\theta} log(\tau) & if \quad n \geq n_u \end{cases}
$$

- Key things are:
  - 1) power law in n 2) τ enters the log density additively

Graphically:



Taking the model to the data

Our estimating equation:

$$\log(\chi(n)) = \alpha - \beta \log(n) + \delta * D,$$
$$D = 1\{n > 9\}$$
$$\hat{\tau} = \exp(\hat{\delta})^{-\frac{1-\theta}{\hat{\beta}-1}}$$

Taking the model to the data

Our estimating equation:

$$\log(\chi(n)) = \alpha - \beta \log(n) + \delta * D,$$
$$D = 1\{n > 9\}$$
$$\hat{\tau} = \exp(\hat{\delta})^{-\frac{1-\theta}{\hat{\beta}-1}}$$

Which we can take to the data:

ID Assumption & Preliminary "Results"

Main ID Assumptions:

- The distribution of firms is power law - *except* for the effect of the regulation.
- The downshift represents actual employment and is not due to misreporting (or any other explanations)

ID Assumption & Preliminary "Results"

Main ID Assumptions:

- The distribution of firms is power law - *except* for the effect of the regulation.
- The downshift represents actual employment and is not due to misreporting (or any other explanations)

All-India	α	β	δ	θ	τ
	-0.390 (0.0135)	2.155 (0.0404)	-1.147 (0.1338)	0.387 (0.0007)	**0.839** **(0.1725)**

Preliminary Results: Variation by State and Industry

State	τ	Industry	τ
Bihar	1.637	Construction	1.157
	(0.4629)		(0.3782)
Kerala	0.0441	Pub admin, etc	-0.224
	(0.0469)		(0.1247)
UP	1.465	Hotels, restaur.	0.966
	(0.4452)		(0.3610)
WB	2.017	***Manufacturing***	***2.501***
	(1.0039)		(1.0317)
TN	0.286	Wholesale, retail	1.569
	(0.0950)		(0.5851)

An Alternative Explanation: Misreporting (1)

- What if firms deliberately misreport their size?
  - Could this explain why we see fewer large firms than expected?
  - I.e. perhaps there is no ‘real’ effect at all…

An Alternative Explanation: Misreporting (1)

- What if firms deliberately misreport their size?
  - Could this explain why we see fewer large firms than expected?
  - I.e. perhaps there is no 'real' effect at all...
- In particular, let firms choose actual employment (n) *and* reported employment (l)
- Profits:

$$\pi(\alpha) = \max_{n,l} \alpha f(n) - wn - \tau l * 1(l > 9) - F(n,l) * p(n,l)$$

An Alternative Explanation: Misreporting (1)

- What if firms deliberately misreport their size?
  - Could this explain why we see fewer large firms than expected?
  - I.e. perhaps there is no 'real' effect at all...
- In particular, let firms choose actual employment (n) *and* reported employment (l)
- Profits:

$$\pi(\alpha) = \max_{n,l} \alpha f(n) - wn - \tau l * 1(l > 9) - F(n,l) * p(n,l)$$

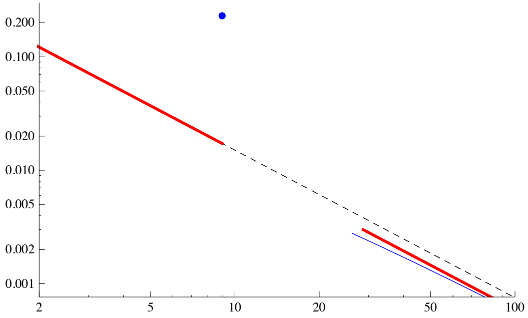
- May be reasonable to assume that the expected cost of misreporting [$F * p$] is an increasing, convex function of misreporting [$n - l$].
- I.e.: $F * \frac{(n-l)^2}{100}$, or $F(n-l) * \frac{(n-l)}{100}$

An Alternative Explanation: Misreporting (2)

- If the expected cost of misreporting $[F * p]$ is an increasing, strictly convex function of misreporting $[n - l]$, e.g.:

$$\pi(\alpha) = \max_{n,l} \alpha n^{\theta} - wn - \tau l * 1(l > 9) - F * \frac{(n-l)^2}{100}$$

- $\psi(l) \to \chi(n)$ for large l, n.
  - I.e.: This kind of misreporting cannot cause the observed 'downshift' *
  - more convex -> "faster" convergence -> lower bias



Misreporting (3)

The Takeaway

- For sufficiently convex misreporting penalties, the bias in estimating *tau* will be minimal.
- Under other assumptions, the bias may be significant.

Misreporting (3)

The Takeaway

- For sufficiently convex misreporting penalties, the bias in estimating *tau* will be minimal.
- Under other assumptions, the bias may be significant.
- Similar conclusions may hold if we amend the model to allow for:
  - substitution with capital, skilled labor or casual/contract labour

Further Issues: Explaining Interstate Variation

Interstate Variation

Table 2: Correlates of Tau table

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	tau	tau	tau	tau	tau	tau	tau
ln_strikes	-0.0936						
	(0.117)						
ln_workers_involved_strikes		0.0651					
		(0.0613)					
ln_mandays_lost_strikes			0.0760				
			(0.0752)				
ln_lockouts				0.216			
				(0.190)			
ln_workers_involved_lockouts					0.0908		
					(0.190)		
ln_mandays_lost_lockouts						0.169	
						(0.134)	
ln_lle_index_2005							-0.547
							(0.470)
Constant	0.883**	0.0391	-0.158	0.519	0.217	-1.168	3.819
	(0.269)	(0.633)	(0.857)	(0.455)	(1.567)	(1.697)	(2.685)
Observations	18	18	18	8	8	8	18

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Interstate Variation

Table 1: Correlates of Tau table

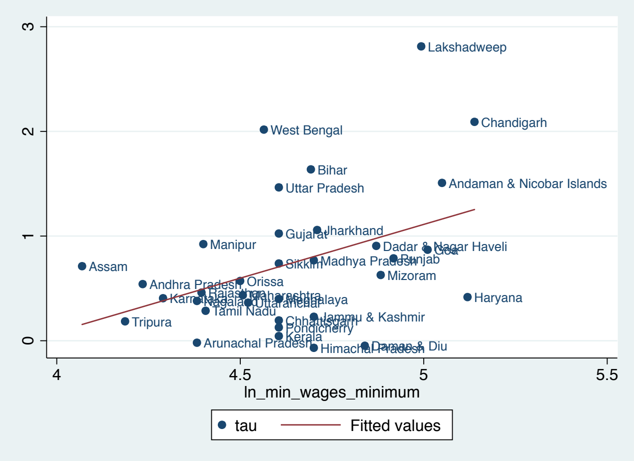
	(1) tau	(2) tau	(3) tau	(4) tau	(5) tau	(6) tau	(7) tau
ln_factory_inspections_fa	0.0430						
	(0.0661)						
ln_convictions_fa		0.0949					
		(0.0641)					
ln_inspections_mtwa			0.0999*				
			(0.0419)				
ln_cases_filed_mtwa				0.0796			
				(0.0950)			
ln_convictions_mtwa					0.0968		
					(0.0673)		
ln_min_wages_minimum						1.027*	
						(0.402)	
ln_min_wage_maximum							0.0811
							(0.338)
Constant	0.378	0.329	0.0182	0.447	0.310	-4.024*	0.330
	(0.468)	(0.364)	(0.280)	(0.377)	(0.268)	(1.862)	(1.677)
Observations	13	8	18	13	10	34	34

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Interstate Variation

Tau Against Lowest Minimum Wage

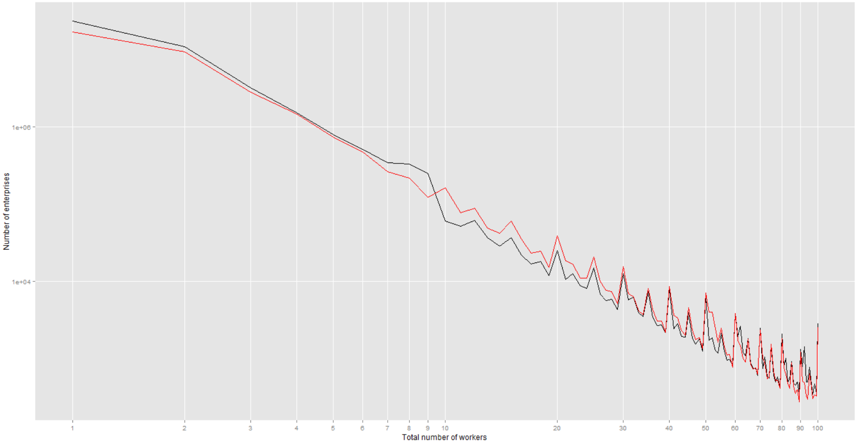


Further Issues: Explaining Intertemporal Variation

Further Issues: Explaining Intertemporal Variation

Frequency by enterprise size (Overlaying EC Waves)

Black = 2005 EC; Red = 1998 EC



Intertemporal Variation

Possible "economic" explanations

- Some suggestive facts:
  - significant decrease in organized employment over the time period.
    - 1997 - 2004, 1.8 million jobs lost (6.3% of org sector)[1]
  - significant increase in casual and contract labor (at least in registered manuf):
    - 1990 - 12%; 1998 - 15.5%; 2005 - 26.8%[2]
- But it could also be...
  - Change in organizational form (eg: more subcontracting/disintegration)?
  - Entry of small firms between 98 and 2005, coupled with exit of larger sized firms?
  - Anything else?

---

[1] 1.2 million in org manuf sector (18% of jobs) [Nagaraj, 2007]

[2] [Sundar, 2012; Maiti, 2013]

Intertemporal Variation

Possible "statistical" explanations

- Change in cost of misreporting or the benefit of reporting accurately?
  - dilution of powers of ministry of factories in 2000?
- In 2005 EC, an extra burden placed on enumerators for enterprises > 10 - the address slip
  - enumerators were paid extra in some states (Bihar, Tamil Nadu - not WB or MH)
  - some evidence that this is not an issue (from Post Enumeration Checks in WB and Tamil Nadu, anyway)
  - can't explain the downshift
  - similar downshift doesn't appear in other waves of the EC

Conclusion

- We observe a distortion in the size distribution of Indian enterprises arising between the period 1998-2005
  - *largest in the manufacturing sector.*
- We suspect that these distortions have something to do with size-based industrial regulations.
  - eg: perhaps they arise due to an interaction between these regulations and an increase in the use and availability of contract labor
- Following Garicano et al. (2013) we translate the size of the distortion into economic costs - under the assumption that the distortions are *not* caused by misreporting or other types of adjustment.
  - Under this assumption, employing 10 or more workers results in an 84% (average) increase in the cost of labor.
  - Given the caveats above, this is likely an upper bound.

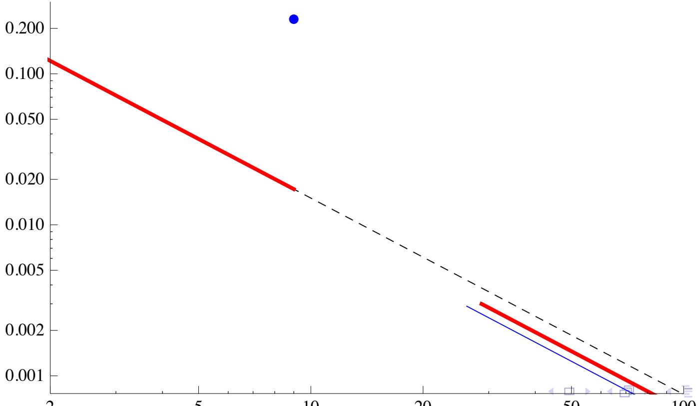
Thank you!

Graphically

- Case 2:

$$\pi(\alpha) = \max_{n,l} \alpha n^{\theta} - wn - \tau l * 1(l > 9) - F * (n-l) * \frac{(n-l)}{l}$$

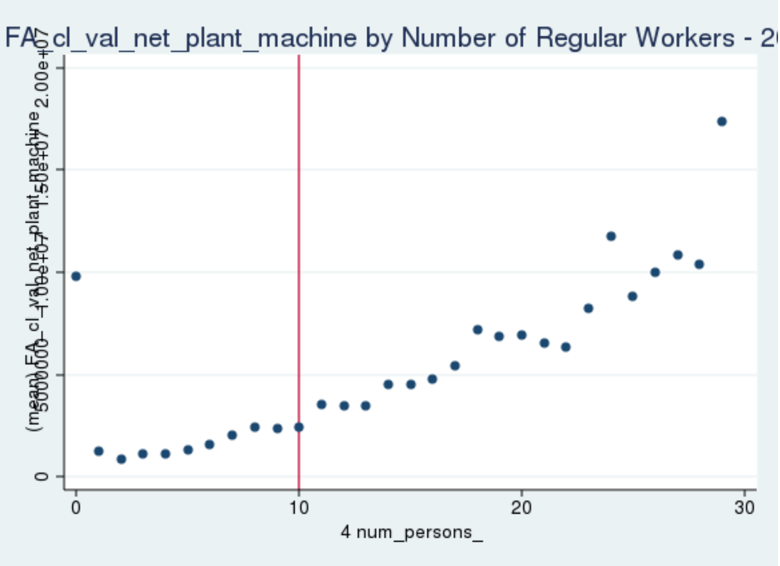
- Dashed: no tax;
- Red line: true distribution $\chi(n)$
- Blue line: reported distribution $\psi(l)$



Distribution of Enterprises By Enterprise Size (log-scale; nonparametric fit)

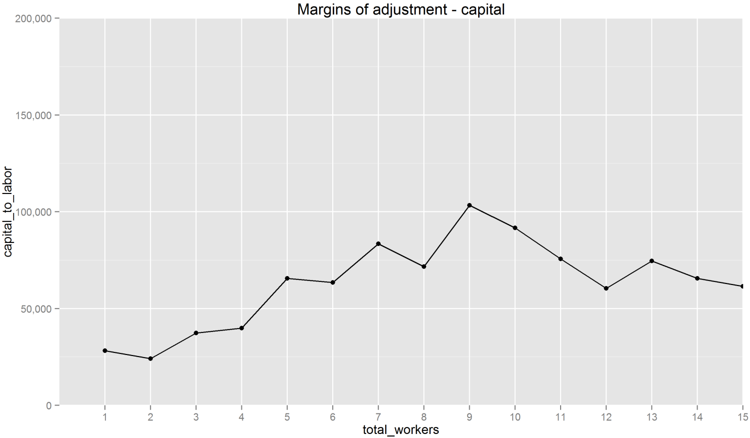
Other Margins of Adjustment

(i.e. other than employment) - ASI



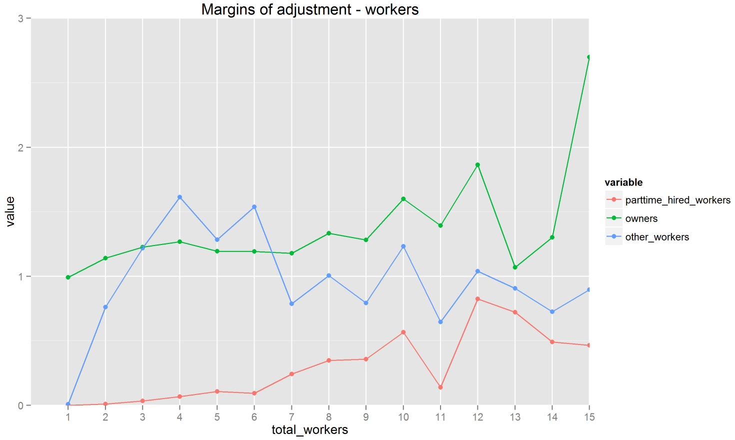
Other Margins of Adjustment

From the NSSO - by hired workers



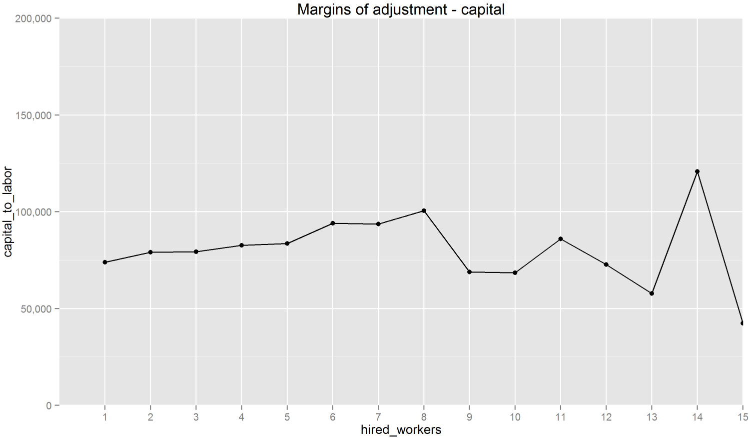
Other Margins of Adjustment

From the NSSO - by hired workers



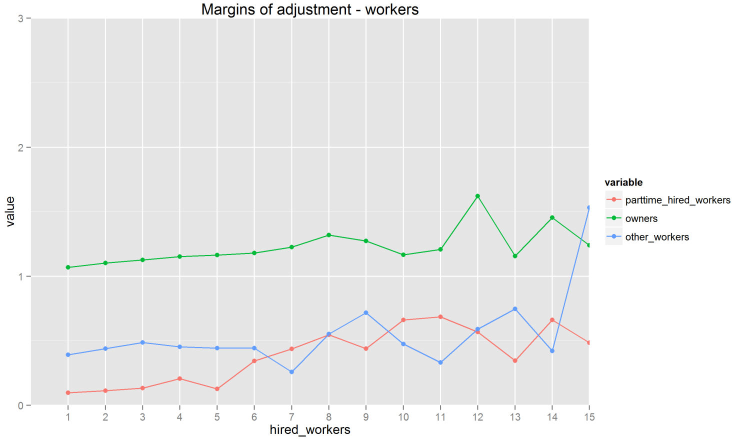
Other Margins of Adjustment

From the NSSO



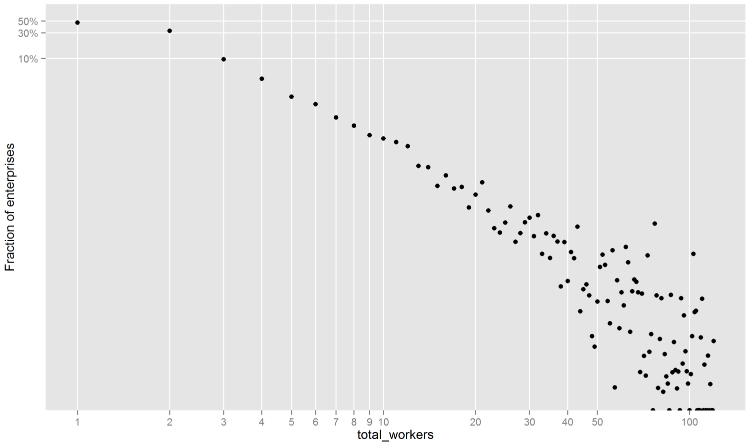
Other Margins of Adjustment

From the NSSO



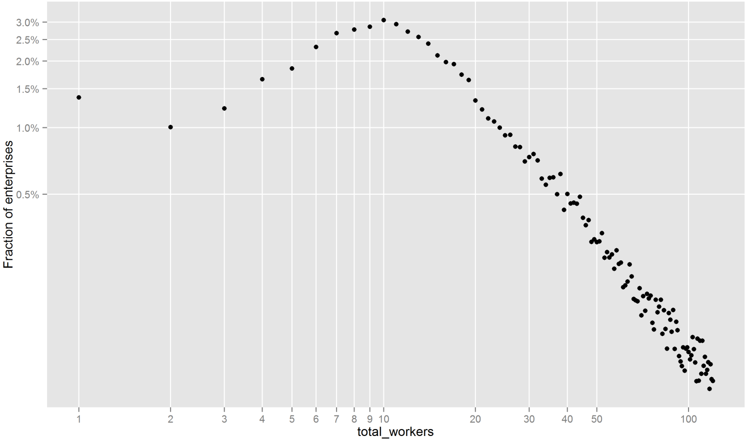
Firm Size Distribution in Other Datasets

2005 NSSO



Firm Size Distribution in Other Datasets

2005 ASI

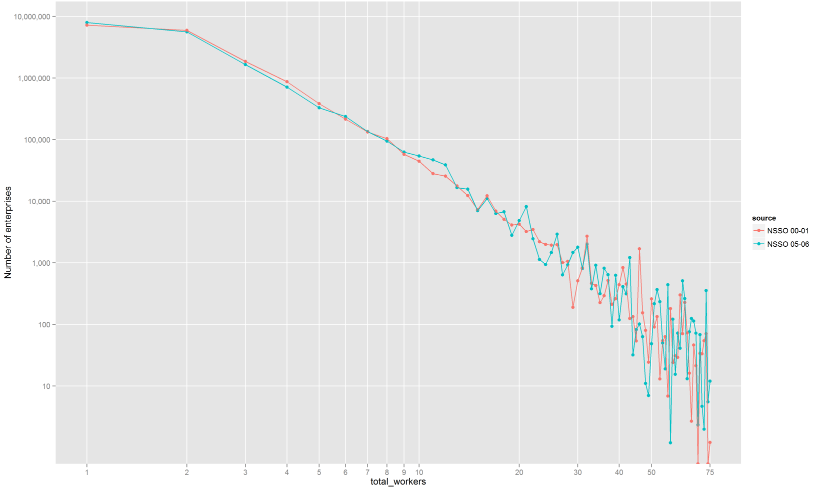


Firm Size Distribution in Other Datasets

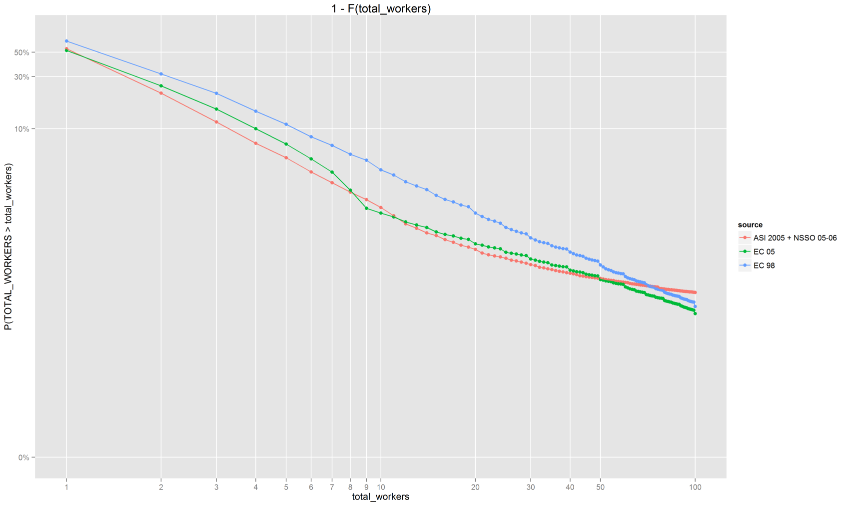
MSME coming soon...

Intertemporal Variation -Other Datasets?

NSSO Unorganized Manufacturing Surveys

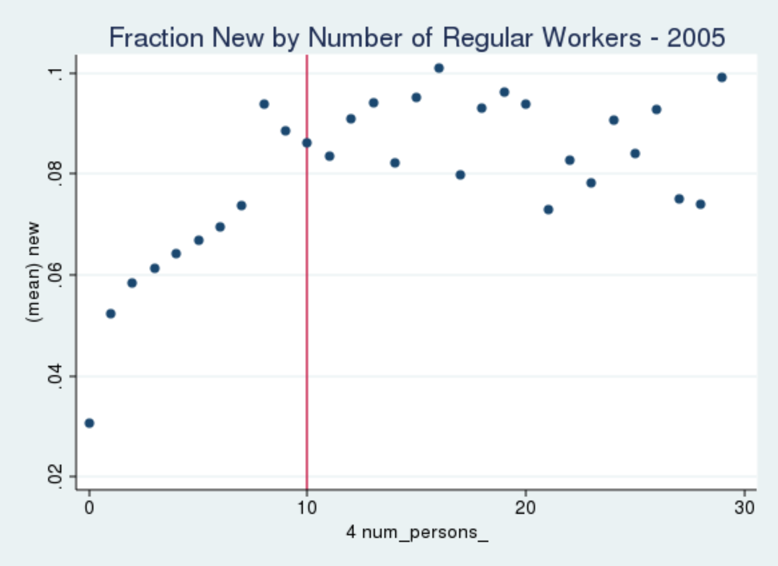


Intertemporal Variation



Intertemporal Variation

Entry - ASI



An Alternative Explanation: Misreporting

- What if firms deliberately misreport their size?
  - Could this explain why we see fewer large firms than expected?
  - I.e. perhaps there is no 'real' effect at all...

An Alternative Explanation: Misreporting

- What if firms deliberately misreport their size?
  - Could this explain why we see fewer large firms than expected?
  - I.e. perhaps there is no 'real' effect at all...

- In particular, let firms choose actual employment (n) *and* reported employment (l)
- Profits:

$$\pi(\alpha) = \max_{n,l} \alpha f(n) - wn - \tau l * 1(l > 9) - F(n,l) * p(n,l)$$

An Alternative Explanation: Misreporting

- What if firms deliberately misreport their size?
  - Could this explain why we see fewer large firms than expected?
  - I.e. perhaps there is no 'real' effect at all...

- In particular, let firms choose actual employment (n) *and* reported employment (l)
- Profits:

$$\pi(\alpha) = \max_{n,l} \alpha f(n) - wn - \tau l * 1(l > 9) - F(n,l) * p(n,l)$$

- With specific functional forms:

$$\pi(\alpha) = \max_{n,l} \alpha n^{\theta} - wn - \tau l * 1(l > 9) - F * \frac{(n-l)^2}{100}$$

$$\pi(\alpha) = \max_{n,l} \alpha n^{\theta} - wn - \tau l * 1(l > 9) - F * (n-l) * \frac{(n-l)}{l}$$

Modelling Misreporting

- The result:
  - $\underline{\alpha} < \alpha < \alpha_1$: Unconstrained managers - *choose* $n \leq N$ *&* $l = n$
  - $\alpha_1 < \alpha < \alpha_2$: Misreporting managers - *choose* $l = N$ *but* $n > N$ *(to skirt the regulation)*
  - $\alpha_2 < \alpha < \infty$: Taxed managers - *choose* $n > N$ *&* $l > N$ *(but* $l \neq n$)

Modelling Misreporting

- The result:
  - $\underline{\alpha} < \alpha < \alpha_1$: Unconstrained managers - *choose* $n \leq N$ *&* $l = n$
  - $\alpha_1 < \alpha < \alpha_2$: Misreporting managers - *choose* $l = N$ *but* $n > N$ *(to skirt the regulation)*
  - $\alpha_2 < \alpha < \infty$: Taxed managers - *choose* $n > N$ *&* $l > N$ *(but* $l \neq n$)
- In case 1: $l = n - \frac{50}{F}\tau$
- In case 2: $l = n * (\frac{1}{\sqrt{\frac{\tau}{F} - 1}})$

Model 1 Implications - Densities

- Then $log\chi(n) =$
$$\begin{cases} logA - \beta log(n) & if \quad n \in [n_{\min}, 9) \\ log[\xi(n)] & if \quad n \in [9, n_m(\alpha_2)] \\ 0 & if \quad n \in (n_m(\alpha_2), n_t(\alpha_2)) \\ logA'(\tau) - \beta log(n) & if \quad n \geq n_t(\alpha_2) \end{cases}$$

Model 1 Implications - Densities

- Then $log\chi(n) =$

$$\begin{cases} logA - \beta log(n) & if \quad n \in [n_{\min}, 9) \\ log[\xi(n)] & if \quad n \in [9, n_m(\alpha_2)] \\ 0 & if \quad n \in (n_m(\alpha_2), n_t(\alpha_2)) \\ logA'(\tau) - \beta log(n) & if \quad n \geq n_t(\alpha_2) \end{cases}$$

- Then $log\psi(l) =$

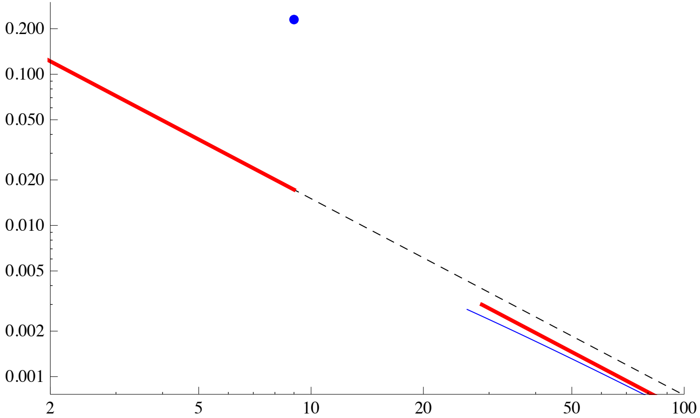
$$\begin{cases} logA - \beta log(l) & if \quad l \in [l_{\min}, 9) \\ log(\delta_l) & if \quad l = 9 \\ 0 & if \quad n \in (9, l_t(\alpha_2)) \\ logA'(\tau) - \beta log(l + \frac{50}{F}\tau) & if \quad l \geq l_t(\alpha_2) \end{cases}$$

Model 1 Implications - Densities

- Then $log\chi(n) =$
$$\begin{cases} logA - \beta log(n) & if \quad n \in [n_{\min}, 9) \\ log[\xi(n)] & if \quad n \in [9, n_m(\alpha_2)] \\ 0 & if \quad n \in (n_m(\alpha_2), n_t(\alpha_2)) \\ logA'(\tau) - \beta log(n) & if \quad n \geq n_t(\alpha_2) \end{cases}$$
- Then $log\psi(l) =$
$$\begin{cases} logA - \beta log(l) & if \quad l \in [l_{\min}, 9) \\ log(\delta_l) & if \quad l = 9 \\ 0 & if \quad n \in (9, l_t(\alpha_2)) \\ logA'(\tau) - \beta log(l + \frac{50}{F}\tau) & if \quad l \geq l_t(\alpha_2) \end{cases}$$
- Key things are:
  - The presence of 'bunching' and 'valleys' can be explained entirely by this kind of misreporting (δ_l)
  - The 'downshift' cannot: $\psi(l) \to \chi(n)$ for large l,n.

Graphically

- Dashed: no tax;
- Red line: true distribution $\chi(n)$
- Blue line: reported distribution $\psi(I)$



Model 2 Implications - Densities

Model 2 Implications - Densities

- Then $log\chi(n) =$
$$\begin{cases} logA * -\beta log(n) & if \quad n \in [n_{\min}, 9) \\ log[\xi(n)] & if \quad n \in [9, n_m(\alpha_2)] \\ 0 & if \quad n \in (n_m(\alpha_2), n_t(\alpha_2)) \\ logA'(\tau) - \beta log(n) & if \quad n \geq n_t(\alpha_2) \end{cases}$$

Model 2 Implications - Densities

- Then $log\chi(n) =$

$$\begin{cases} logA * -\beta log(n) & if \quad n \in [n_{\min}, 9) \\ log[\xi(n)] & if \quad n \in [9, n_m(\alpha_2)] \\ 0 & if \quad n \in (n_m(\alpha_2), n_t(\alpha_2)) \\ logA'(\tau) - \beta log(n) & if \quad n \geq n_t(\alpha_2) \end{cases}$$

- Then $log\psi(l) =$

$$\begin{cases} logA * -\beta log(l) & if \quad l \in [l_{\min}, 9) \\ log(\delta_l) & if \quad l = 9 \\ 0 & if \quad n \in (9, l_t(\alpha_2)) \\ logA'(\tau) - \beta log(l) + (\beta - 1) log[f(\tau)] & if \quad l \geq l_t(\alpha_2) \end{cases}$$

Model 2 Implications - Densities

- Then $log\chi(n) =$

$$\begin{cases} logA * -\beta log(n) & if \quad n \in [n_{\min}, 9) \\ log[\xi(n)] & if \quad n \in [9, n_m(\alpha_2)] \\ 0 & if \quad n \in (n_m(\alpha_2), n_t(\alpha_2)) \\ logA'(\tau) - \beta log(n) & if \quad n \geq n_t(\alpha_2) \end{cases}$$

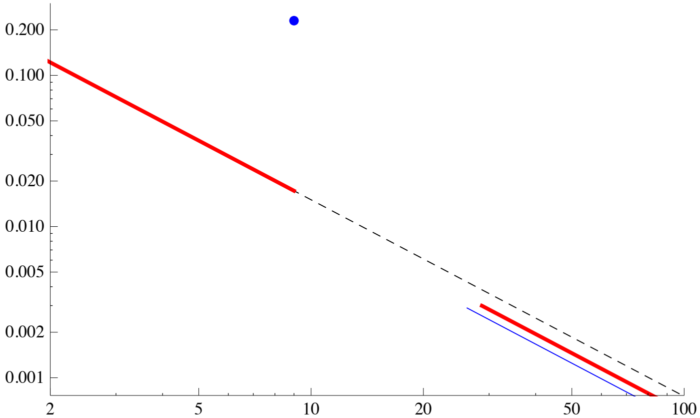
- Then $log\psi(l) =$

$$\begin{cases} logA * -\beta log(l) & if \quad l \in [l_{\min}, 9) \\ log(\delta_l) & if \quad l = 9 \\ 0 & if \quad n \in (9, l_t(\alpha_2)) \\ logA'(\tau) - \beta log(l) + (\beta - 1) log[f(\tau)] & if \quad l \geq l_t(\alpha_2) \end{cases}$$

- Key things are:
  - Again, there is a downshift in the distribution of $\chi(n)$ after 10,
  - But now there is an additional downshift in the distribution of $\psi(l)$, so that any estimate based on $\psi(l)$ will overestimate the true effect.

Graphically

- Dashed: no tax;
- Red line: true distribution $\chi(n)$
- Blue line: reported distribution $\psi(l)$



Theory: Last Words

- If there is the "right kind" of misreporting, the estimates above may be upwardly biased.
- Still, there are costs to misreporting (informalization, tax avoidance), but these may be difficult to measure.
- Same with the other stories (contracting, outsourcing, etc)...

Sources of Finance by size

No finance, govt finance and private finance