Title: Productivity and Entrepreneurship

Subtitle: Mice, Gazelles and Elephants: The Menagerie of U.S. Business Dynamics

Lecture Notes for NBER Entrepreneurship Bootcamp
July 2010
By John Haltiwanger
University of Maryland and NBER

CHRYSLER

Ford

GM

kmart

USS

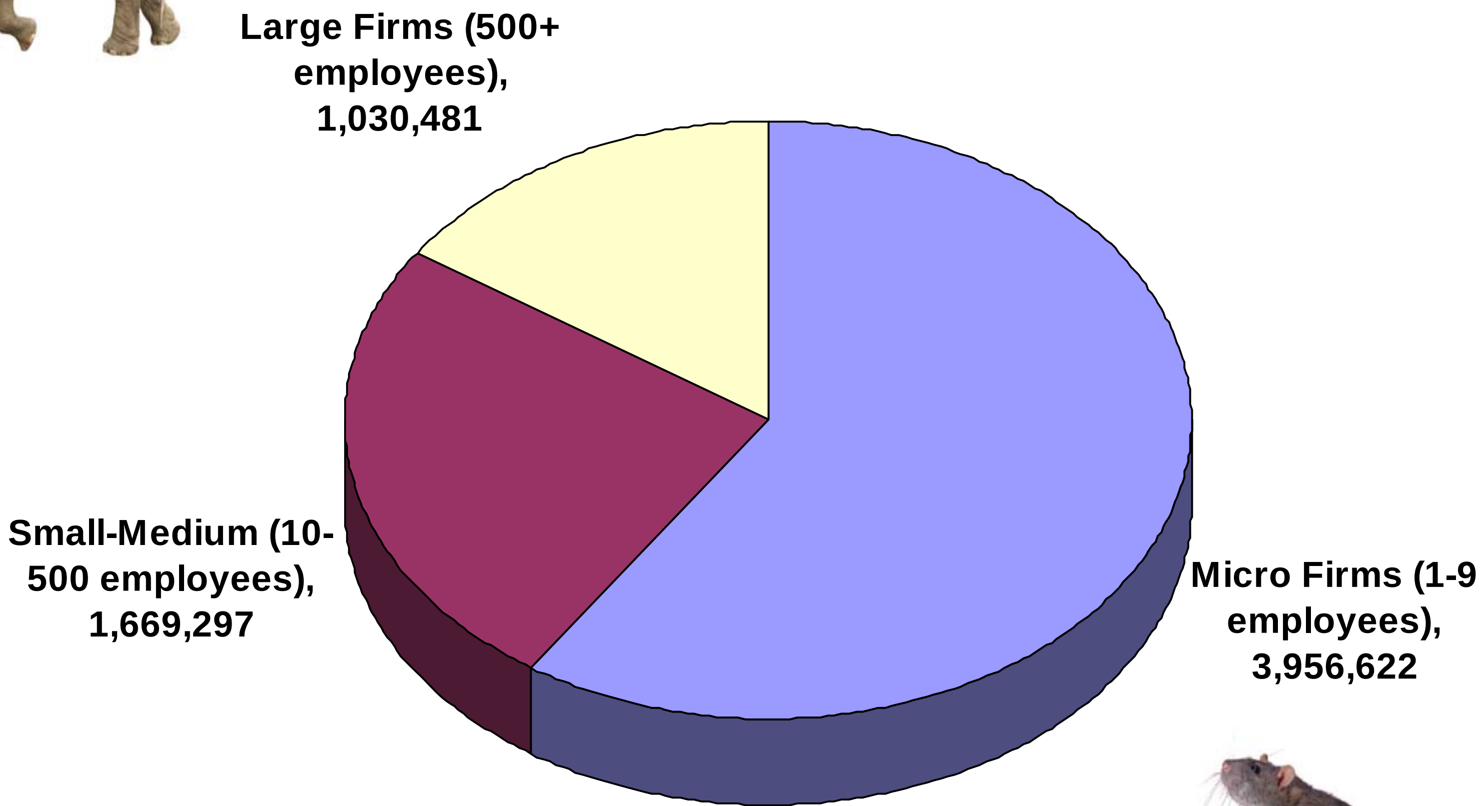
TOYOTA

Google

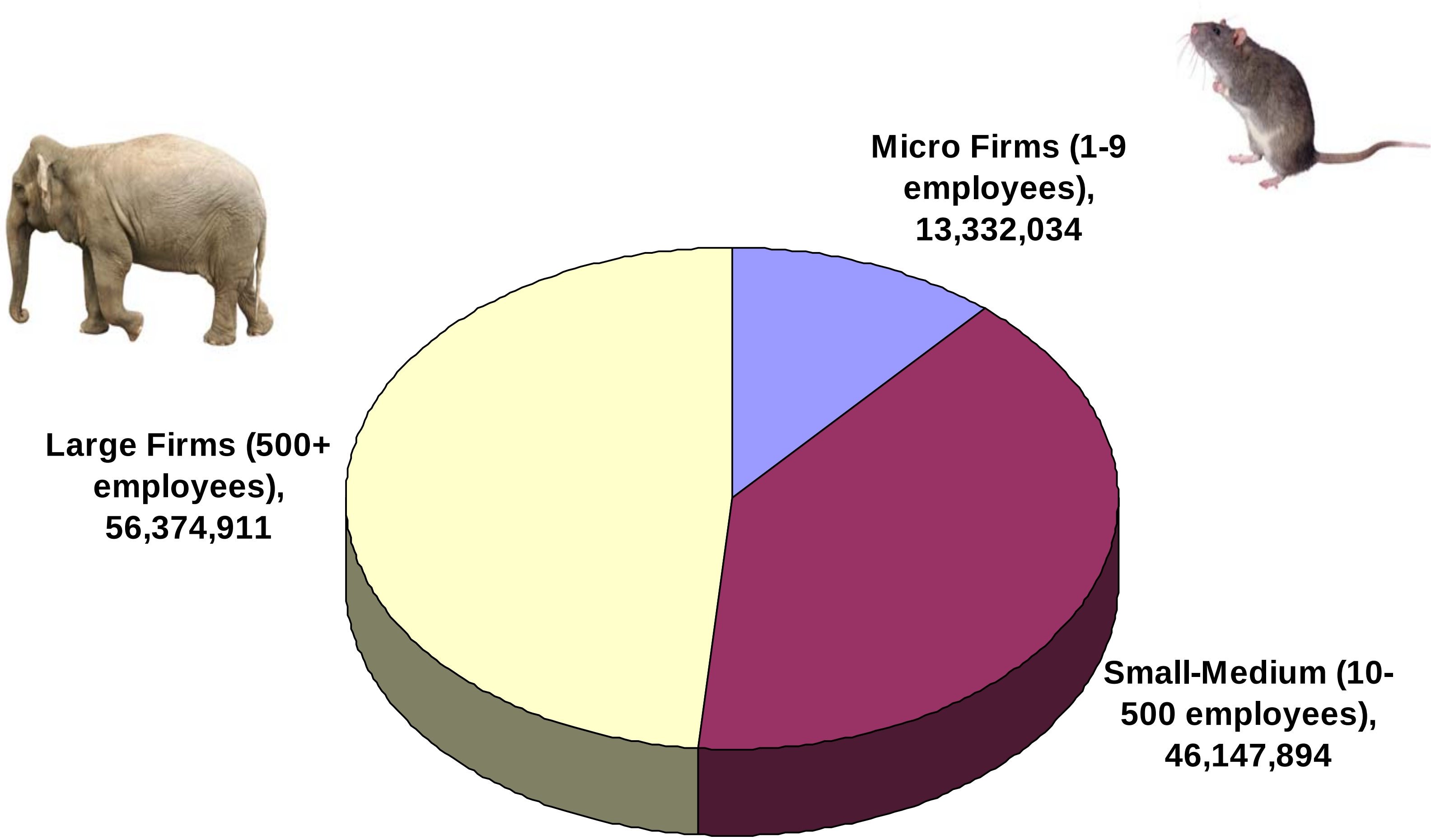
HONDA

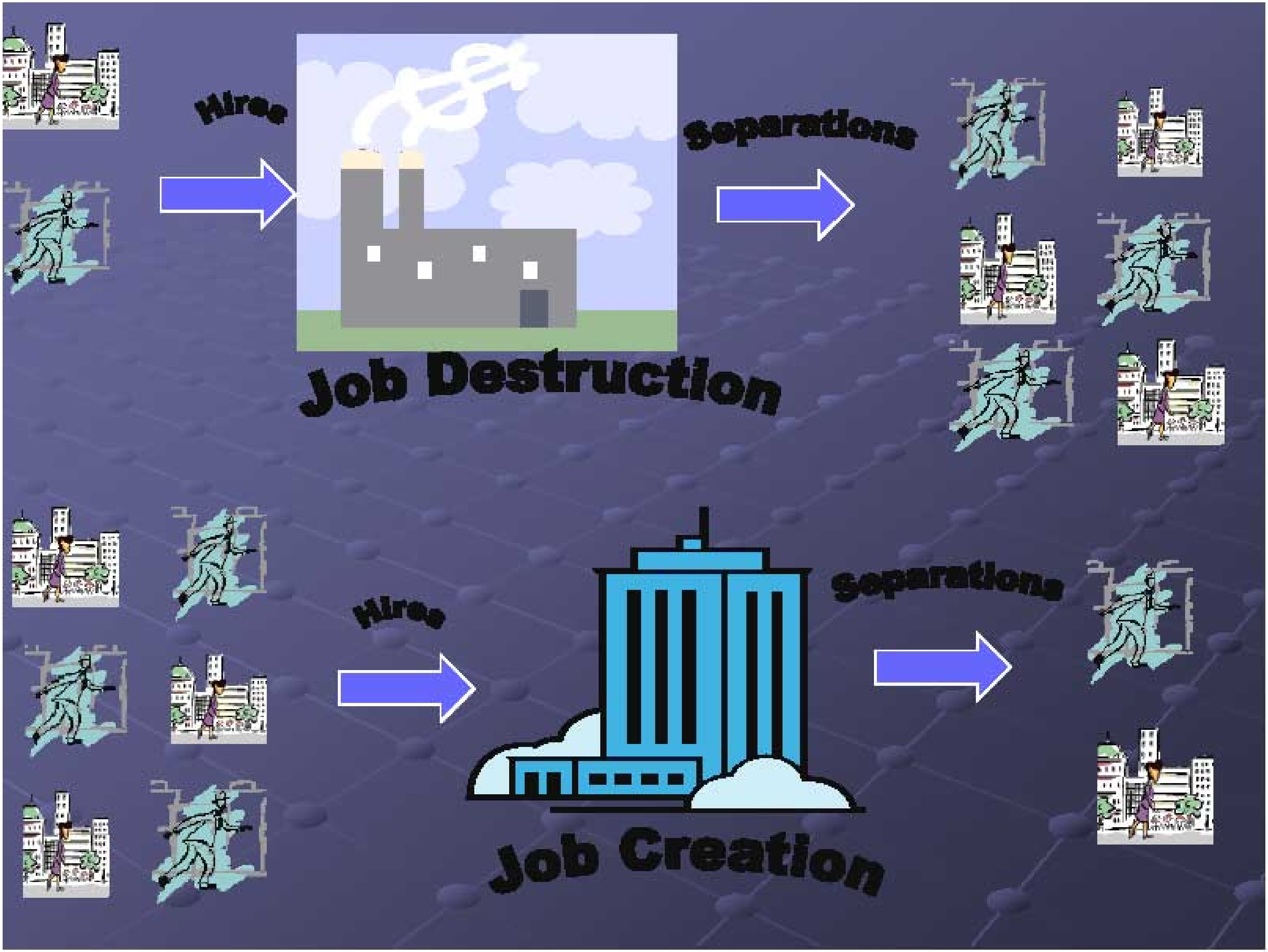
STARBUCKS COFFEE

Share of Establishments by Firm Size, 2005

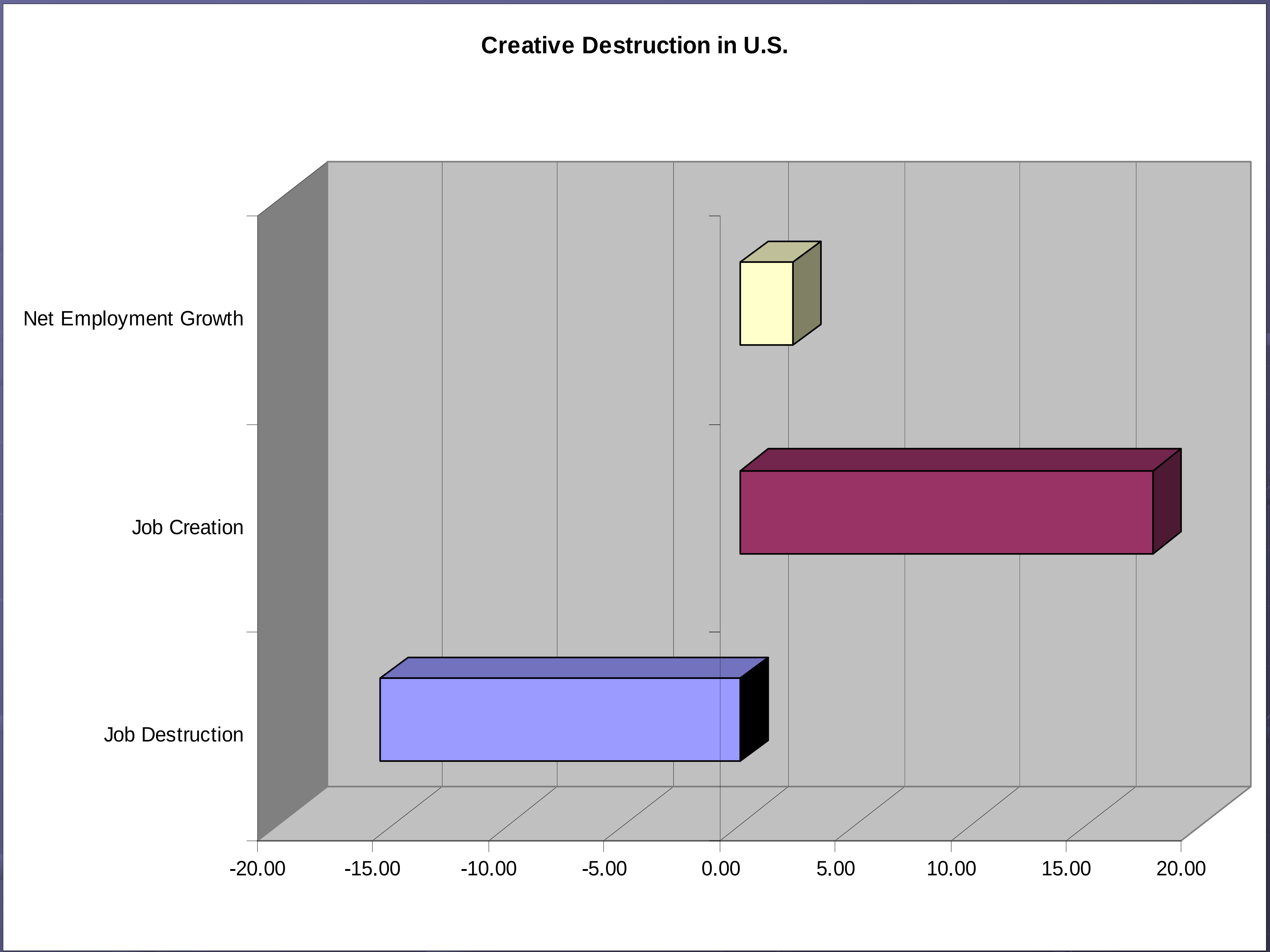


Share of Employment by Firm Size, 2005

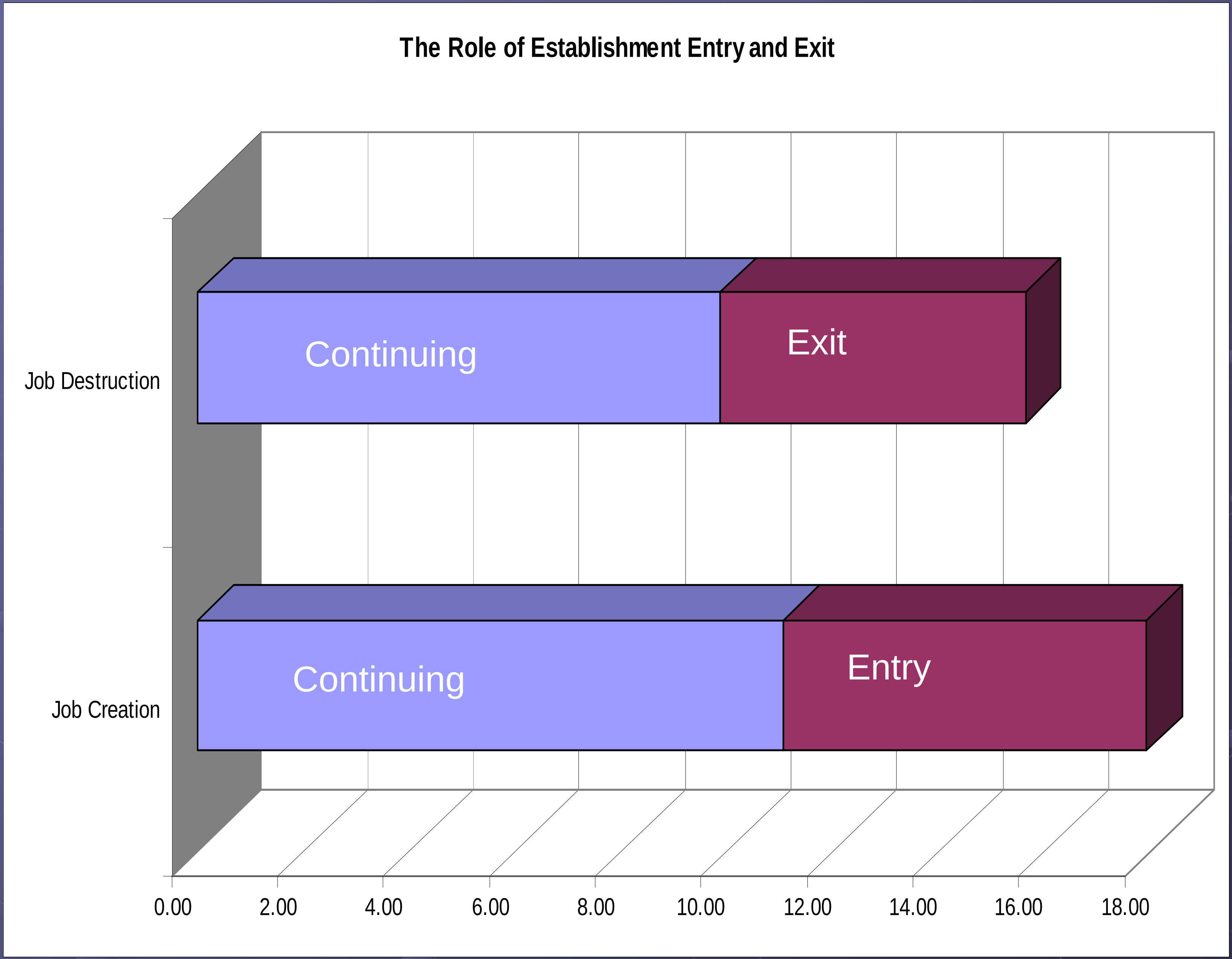




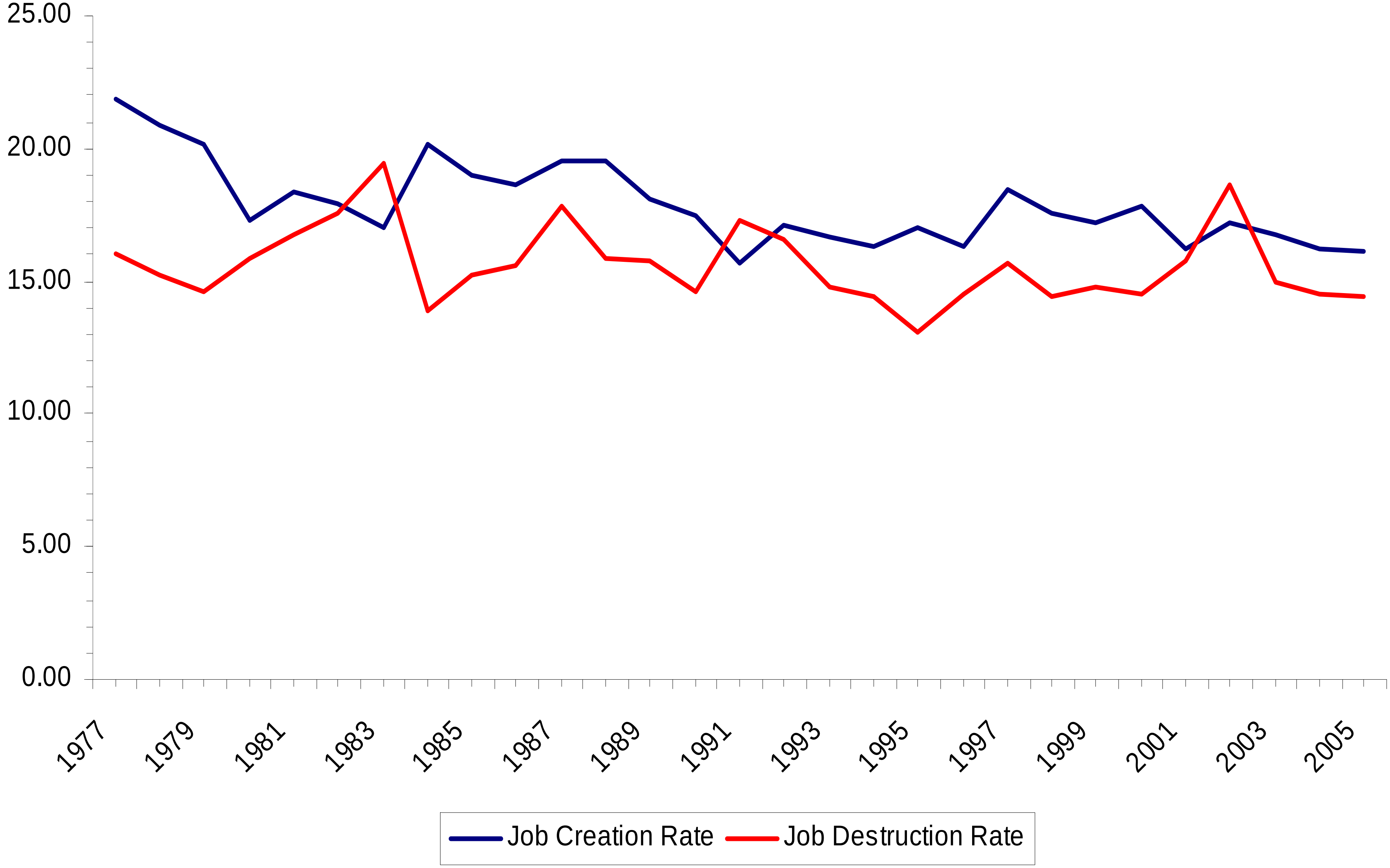
Hires
Separations
Job Destruction
Hires
Separations
Job Creation



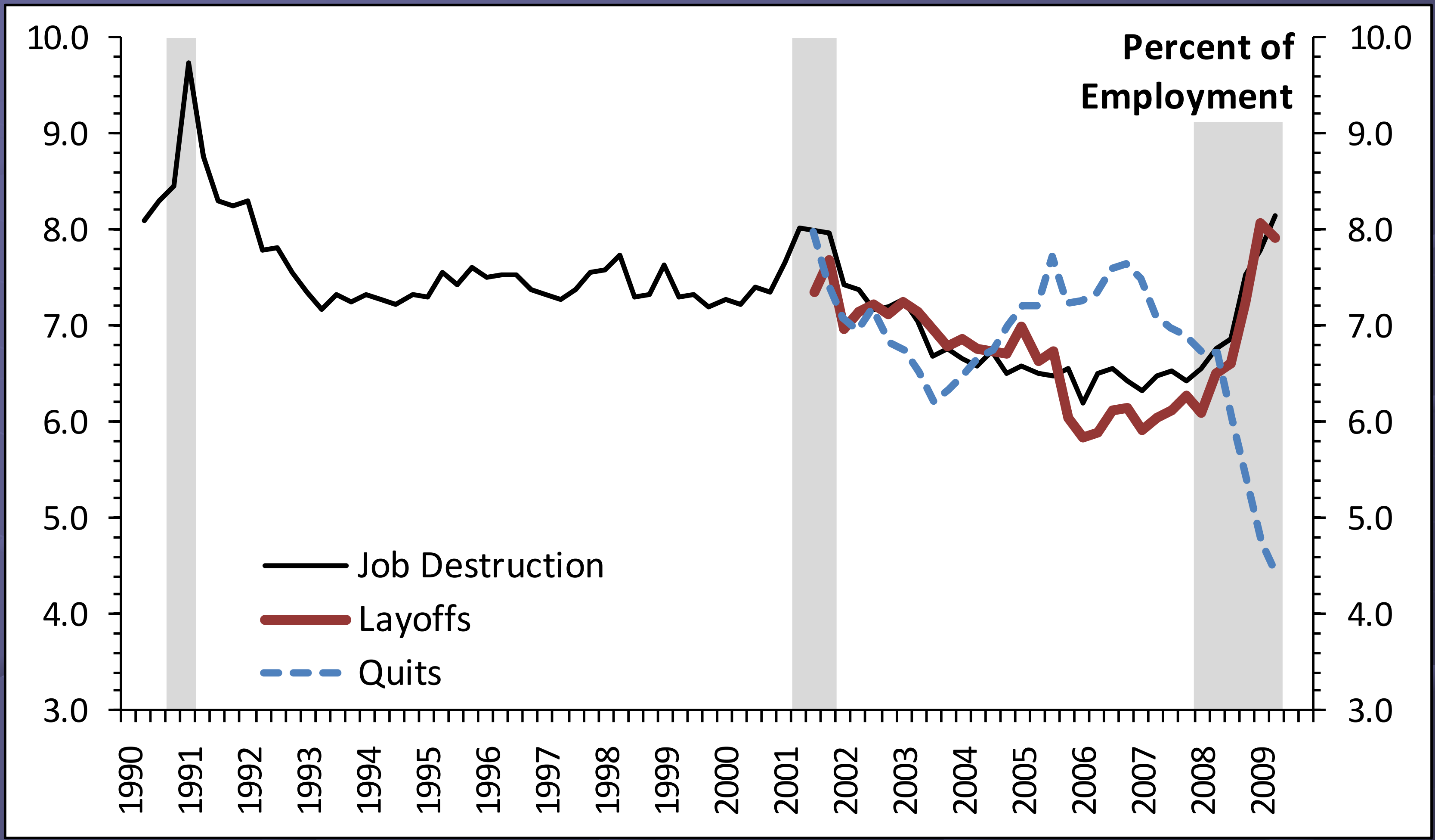
Creative Destruction in U.S.
Net Employment Growth
Job Creation
Job Destruction
-20.00
-15.00
-10.00
-5.00
0.00
5.00
10.00
15.00
20.00


The Role of Establishment Entry and Exit
Job Destruction
Job Creation
Continuing
Exit
Continuing
Entry
0.00
2.00
4.00
6.00
8.00
10.00
12.00
14.00
16.00
18.00

Job Creation/Destruction Rates: Economy by Year

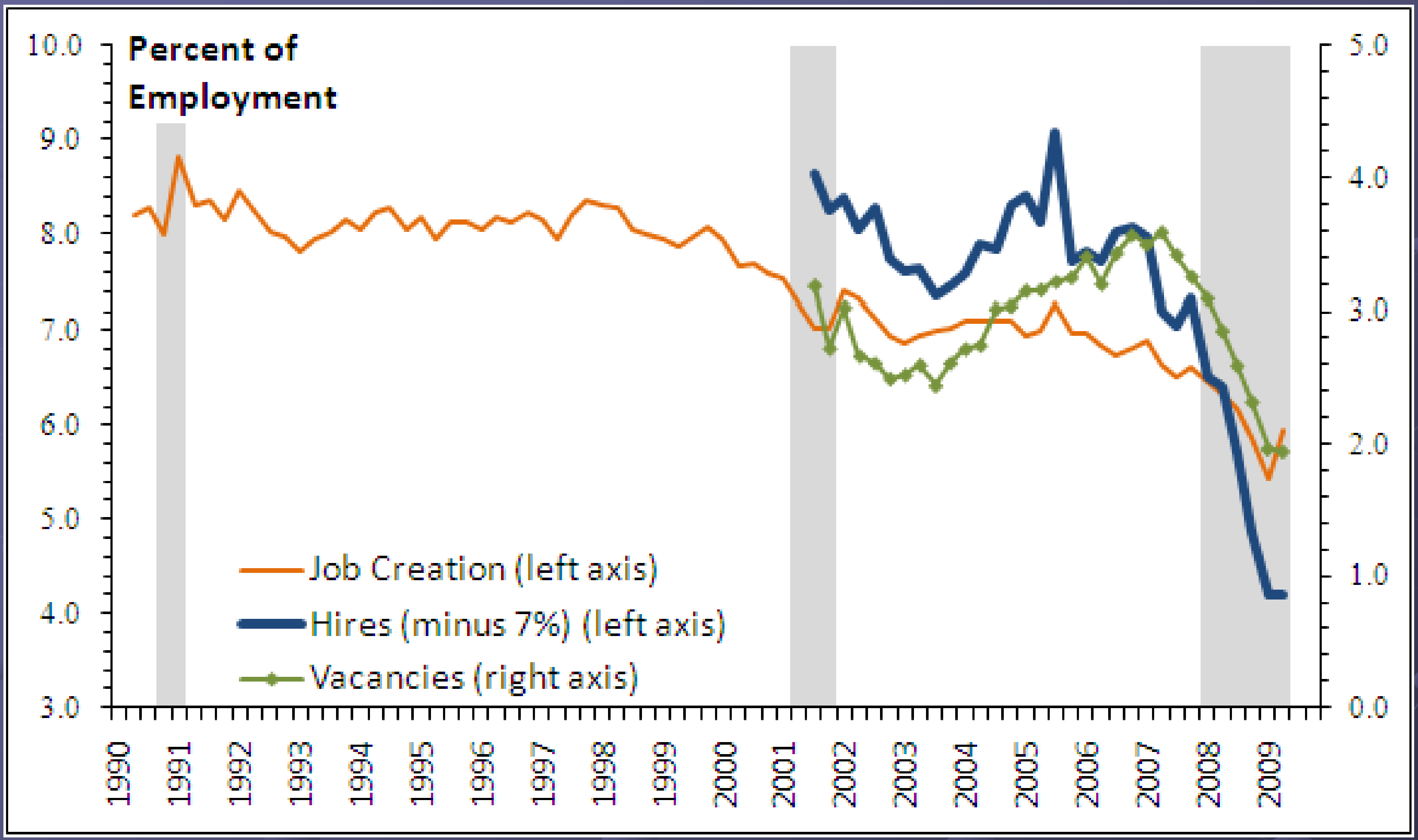


Aggregate Worker and Job Flows (Quarterly)



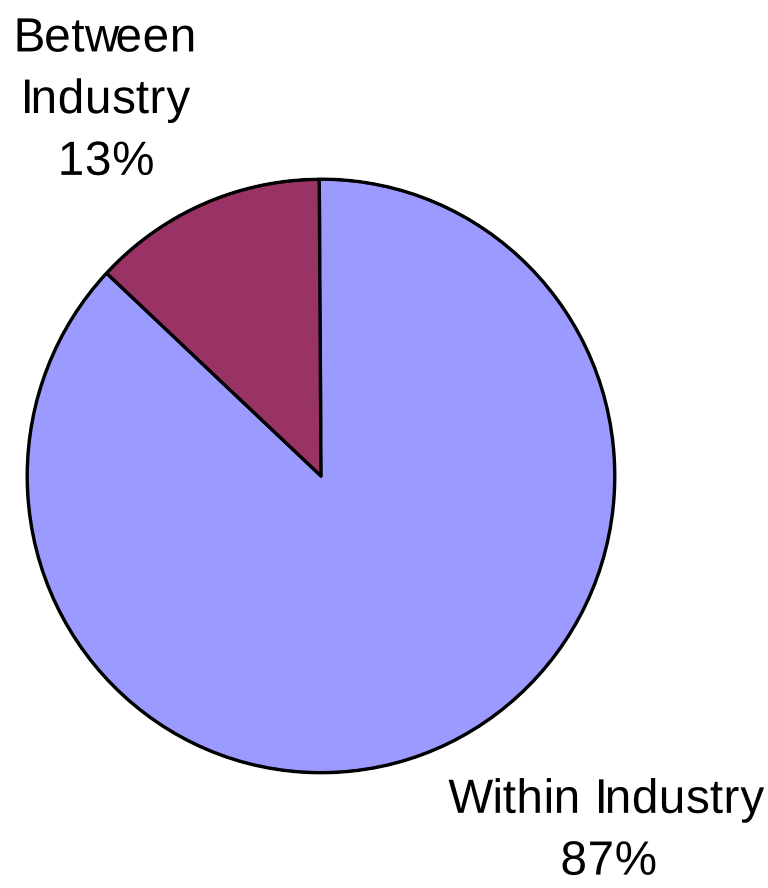
- Layoffs (JOLTS) move with job destruction (BED), and quits (JOLTS) moves opposite to both.

Aggregate Worker and Job Flows

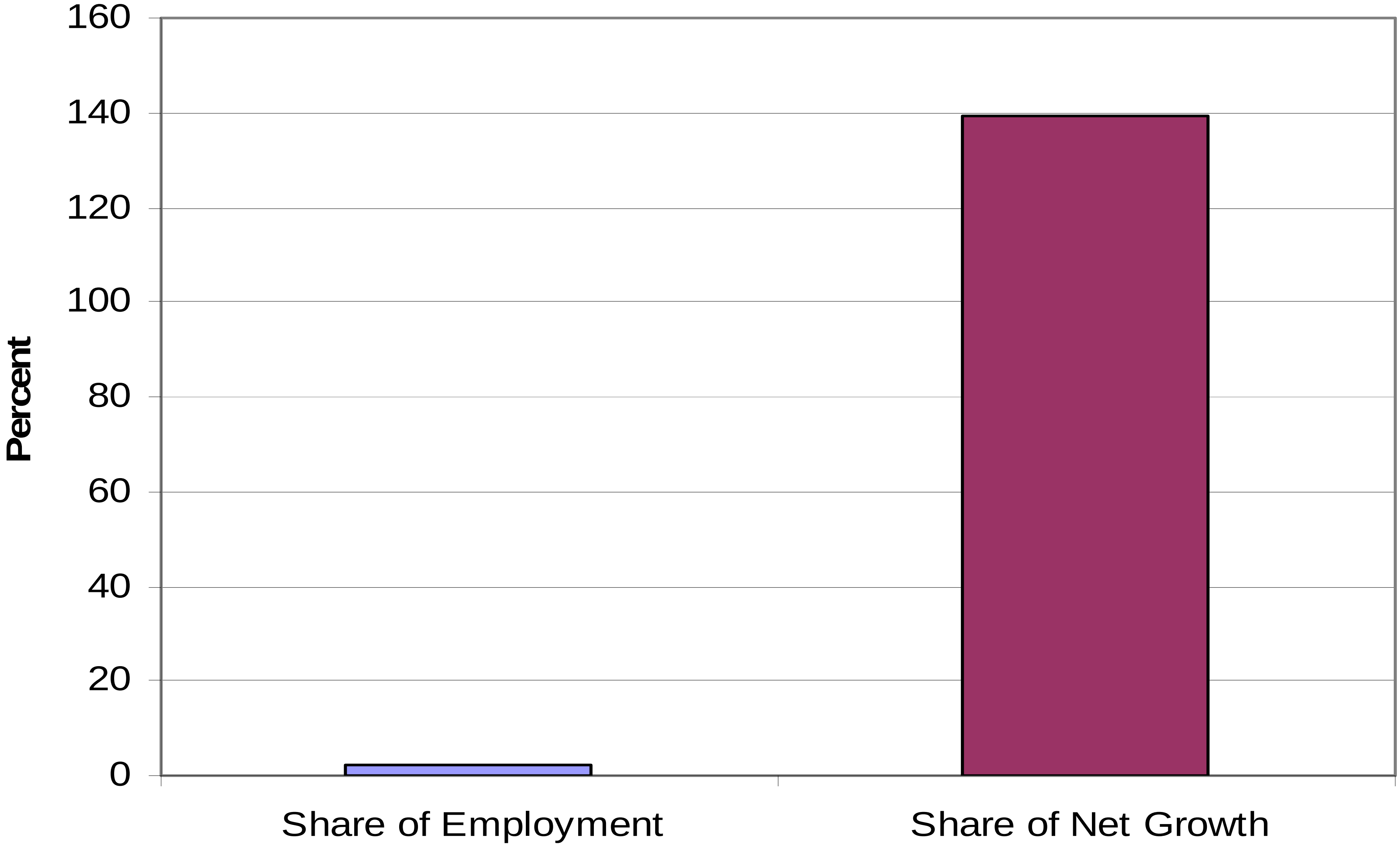


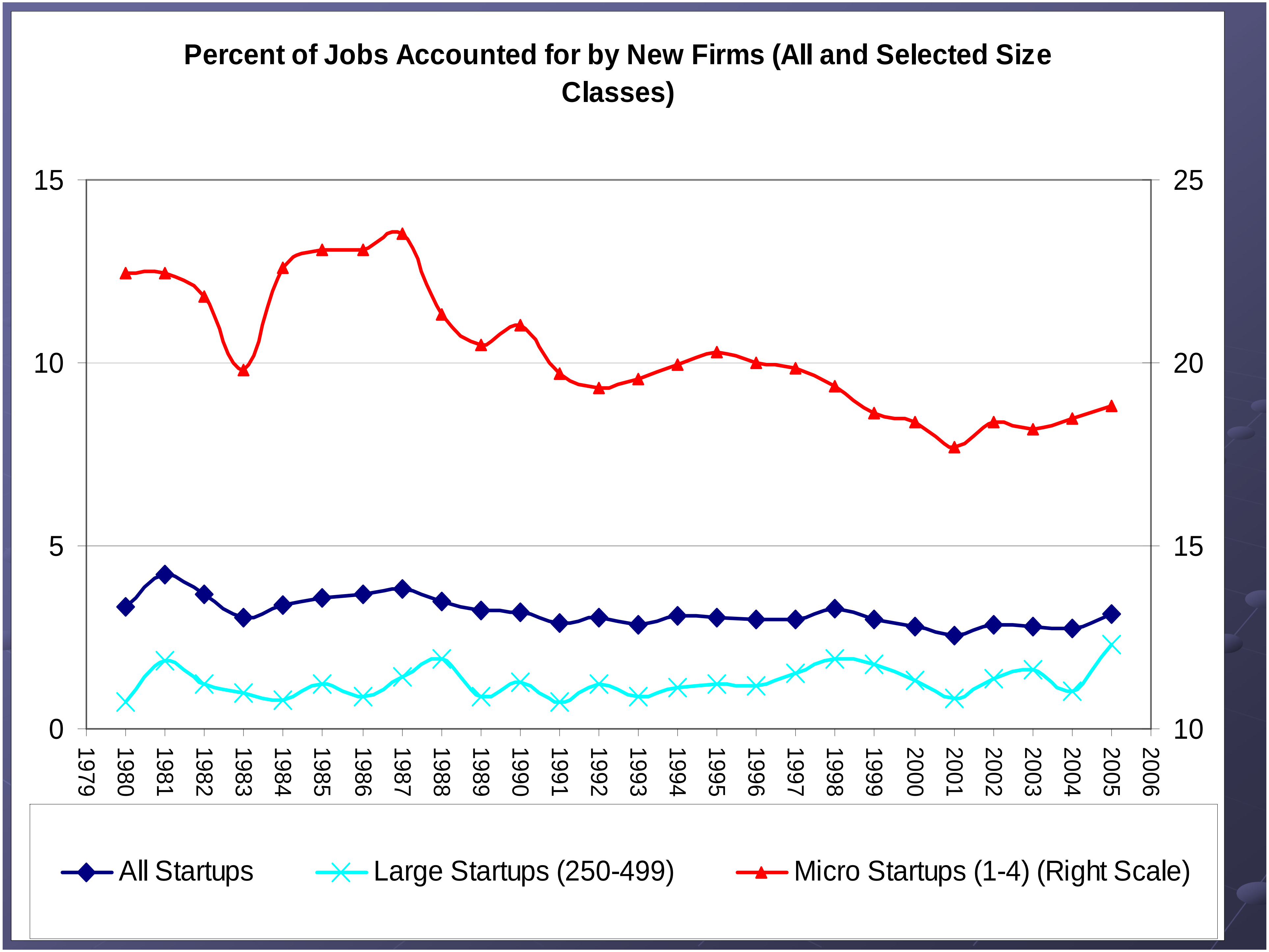
- Hires and vacancies (JOLTS) tend to move with job creation (BED).

Share of Reallocation Between and Within Detailed Industries

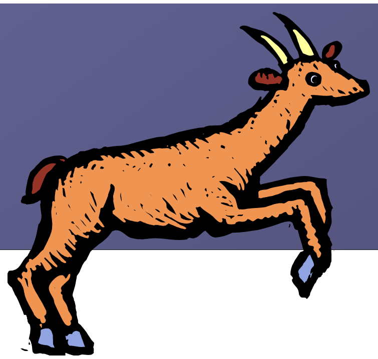


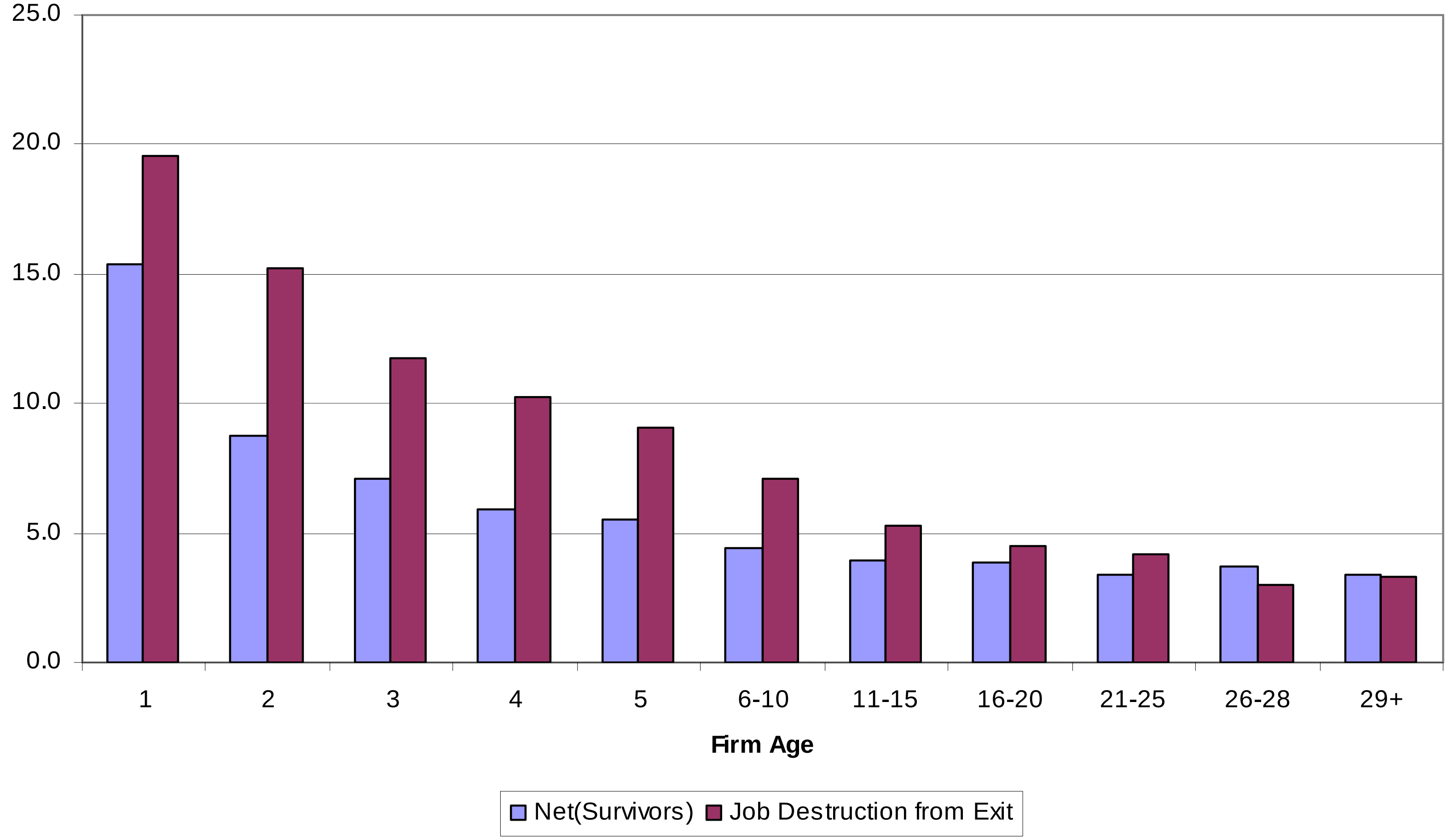
Business Startups as Percentage of Employment and Net Growth



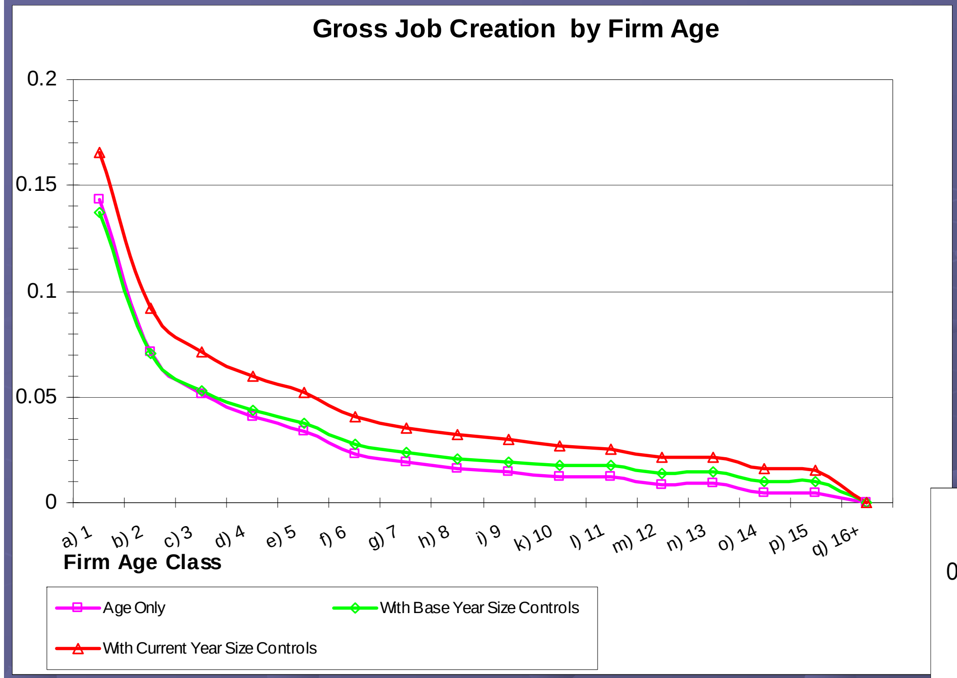

Percent of Jobs Accounted for by New Firms (All and Selected Size Classes)
15
10
5
0
25
20
15
10
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
All Startups
Large Startups (250-499)
Micro Startups (1-4) (Right Scale)

Are there Gazelles?

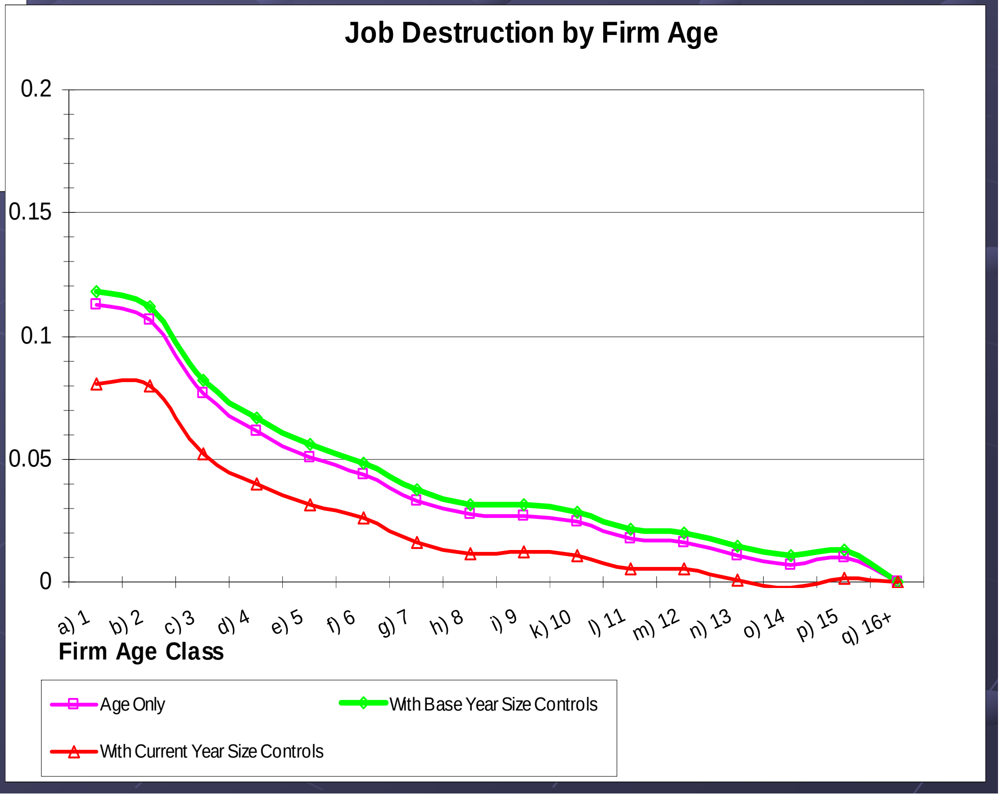




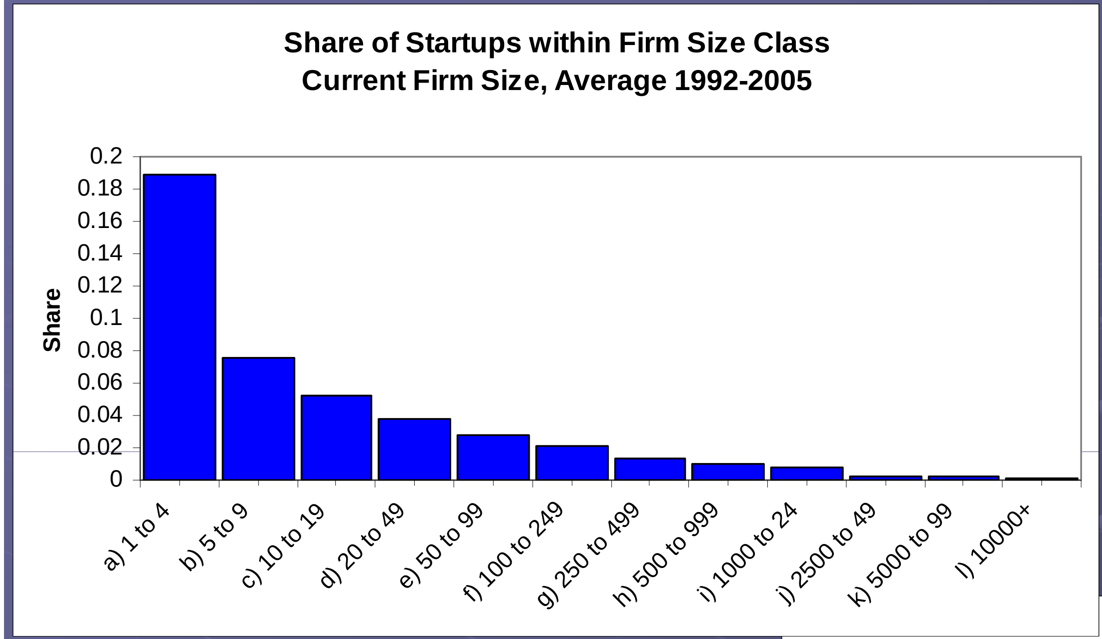
Even excluding startups, young businesses disproportionately create and destroy jobs



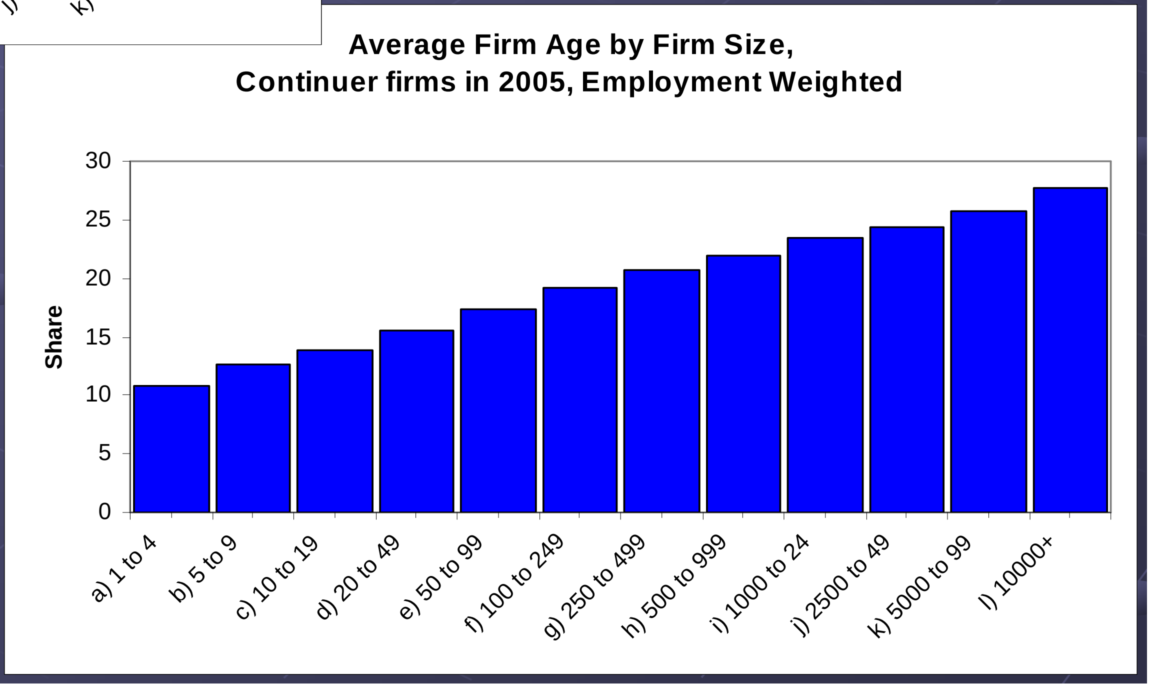
These patterns robust to size controls



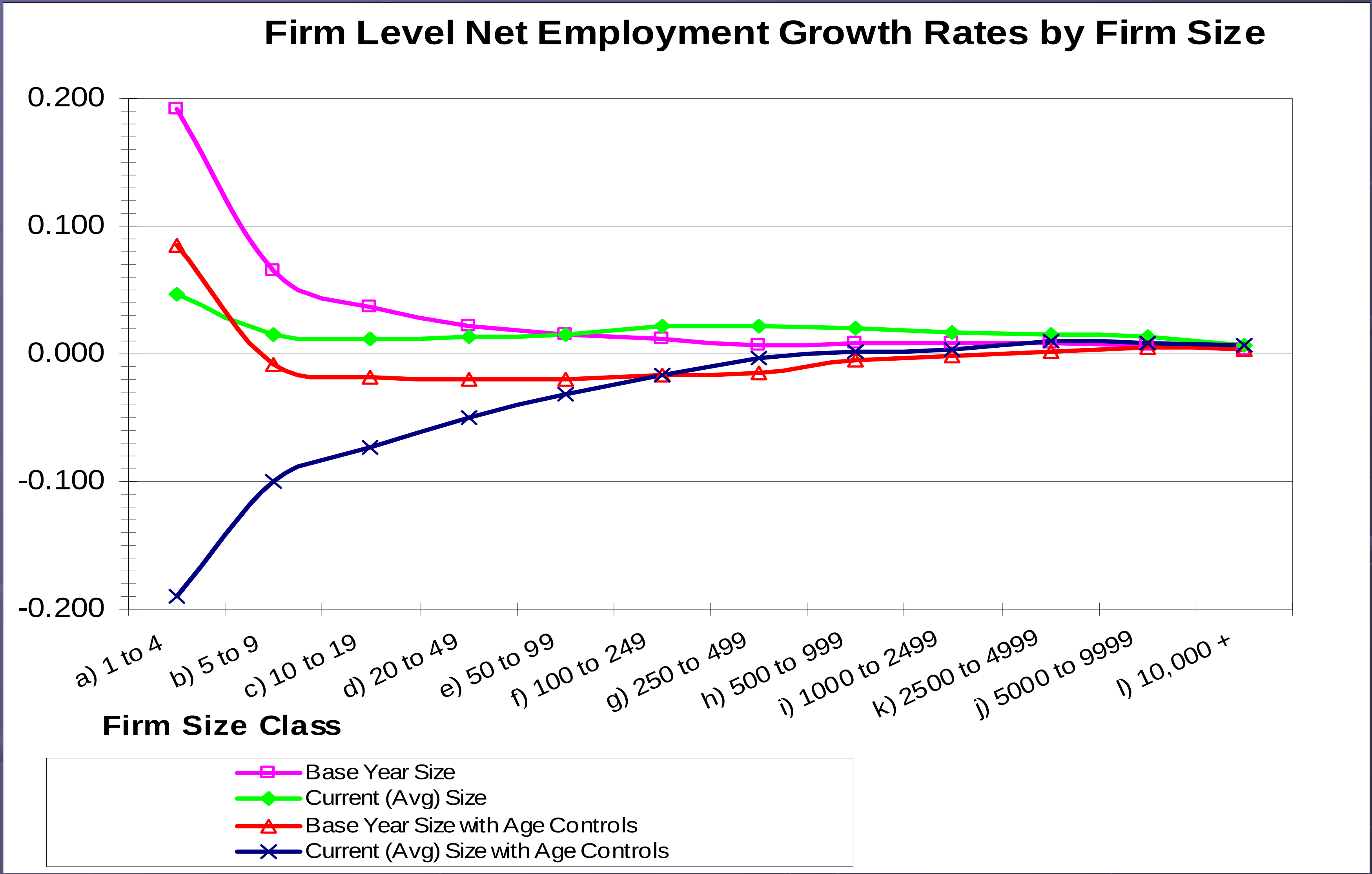
The Size/Age Relation



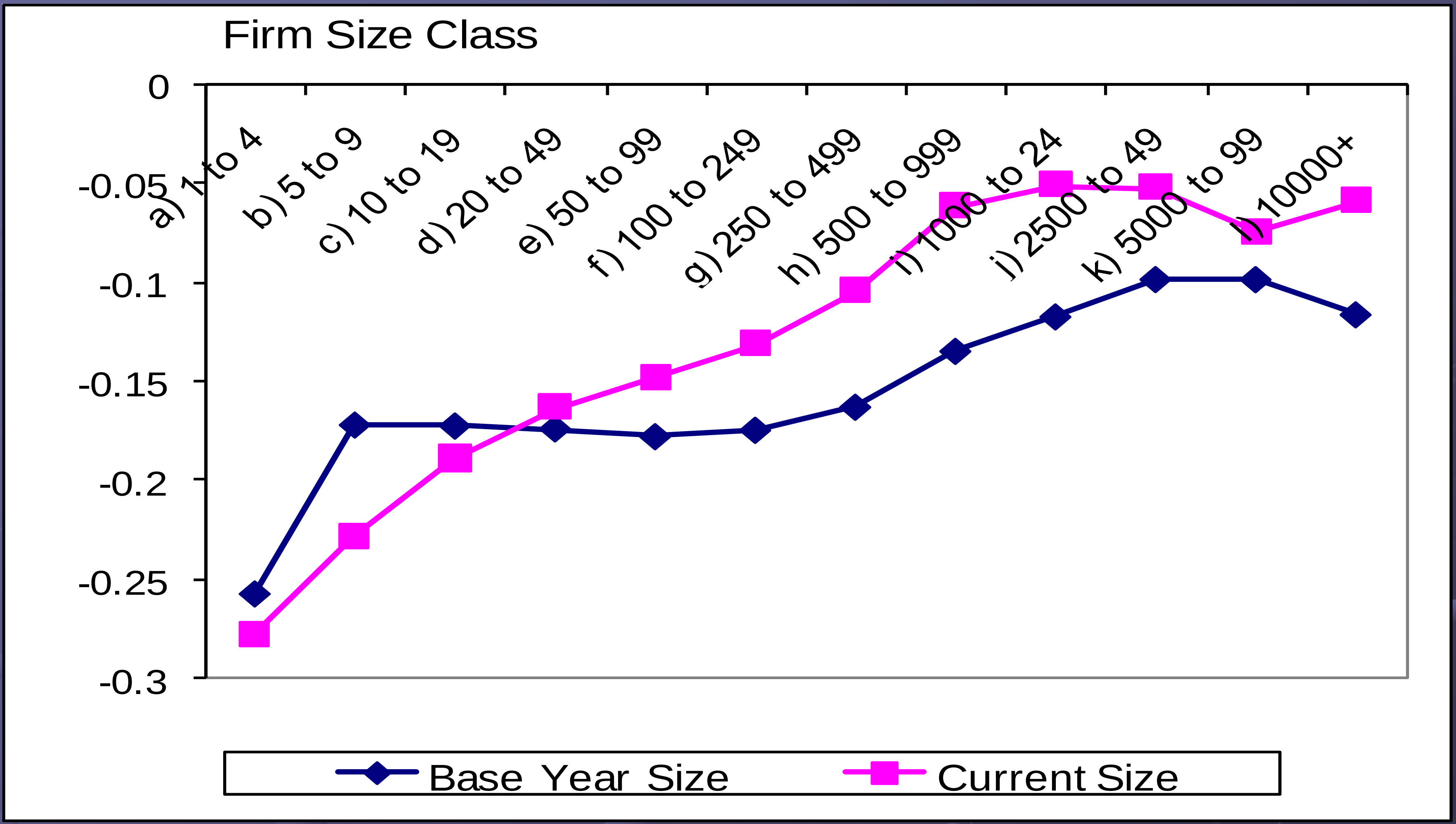
Firm startups are small.
Larger firms are older



Firm Size: Sensitivity to controlling for age and size methodology

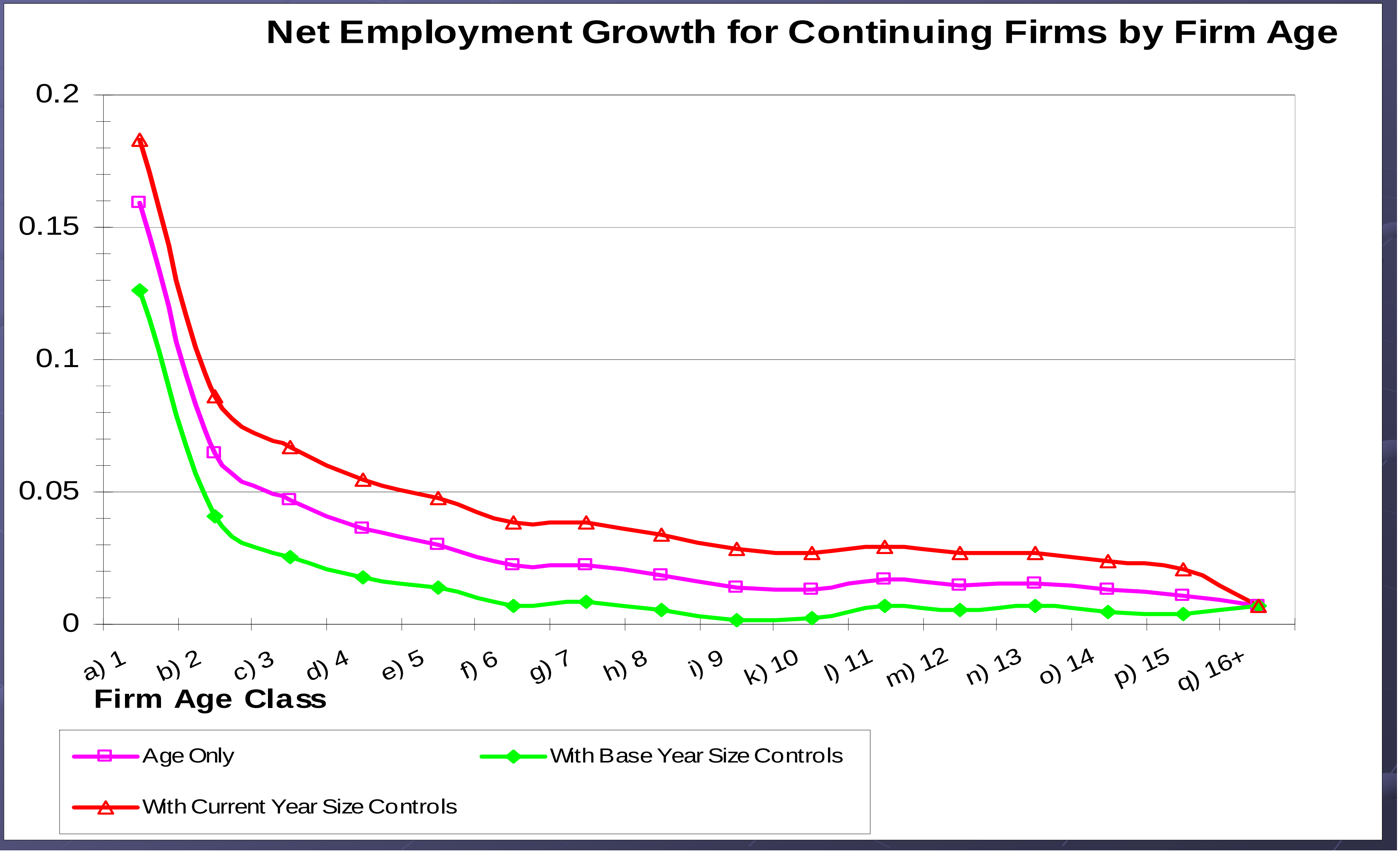


Regression to the Mean

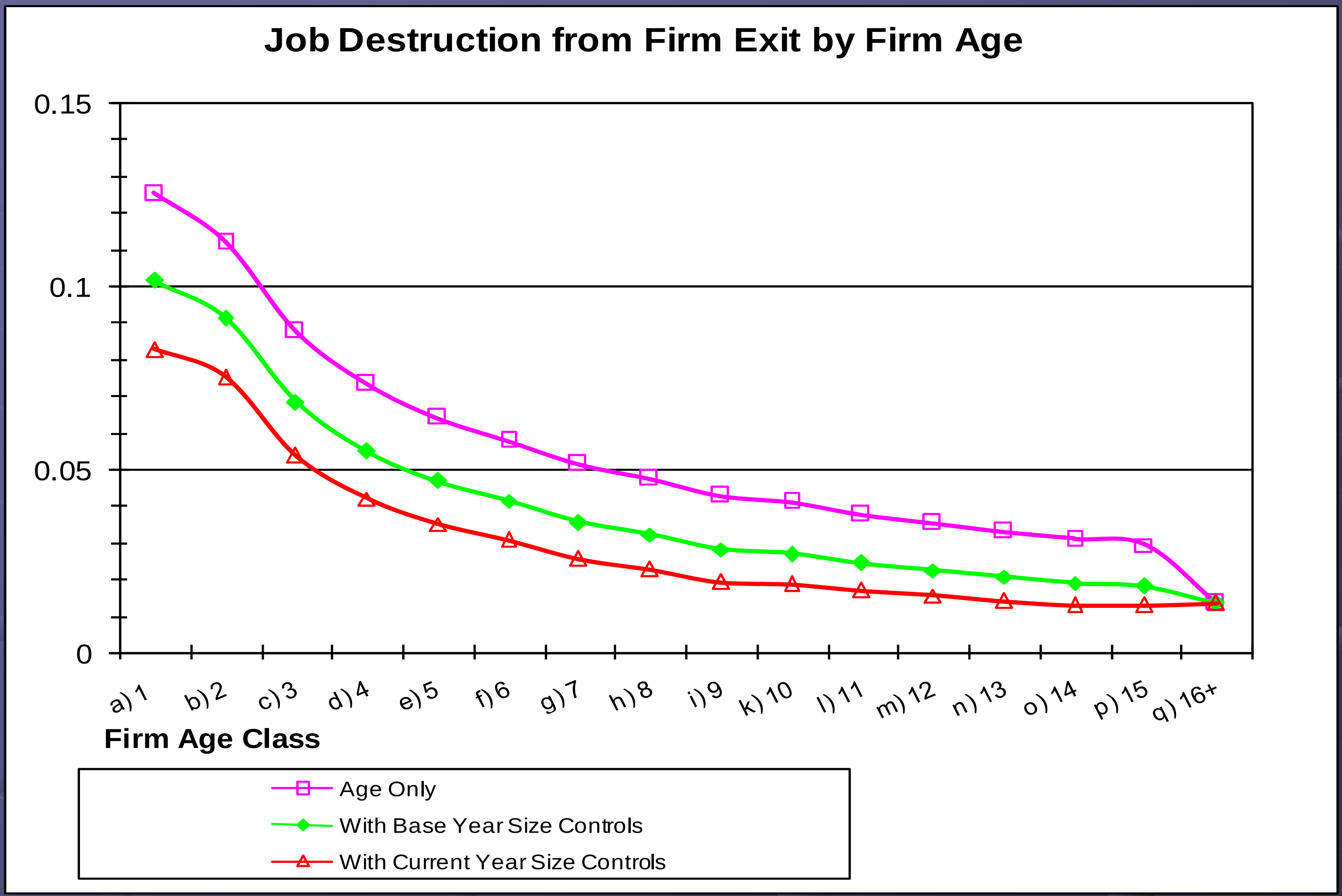


Negative autocorrelation greater for smaller size classes suggests transitory shocks more important for smaller firms

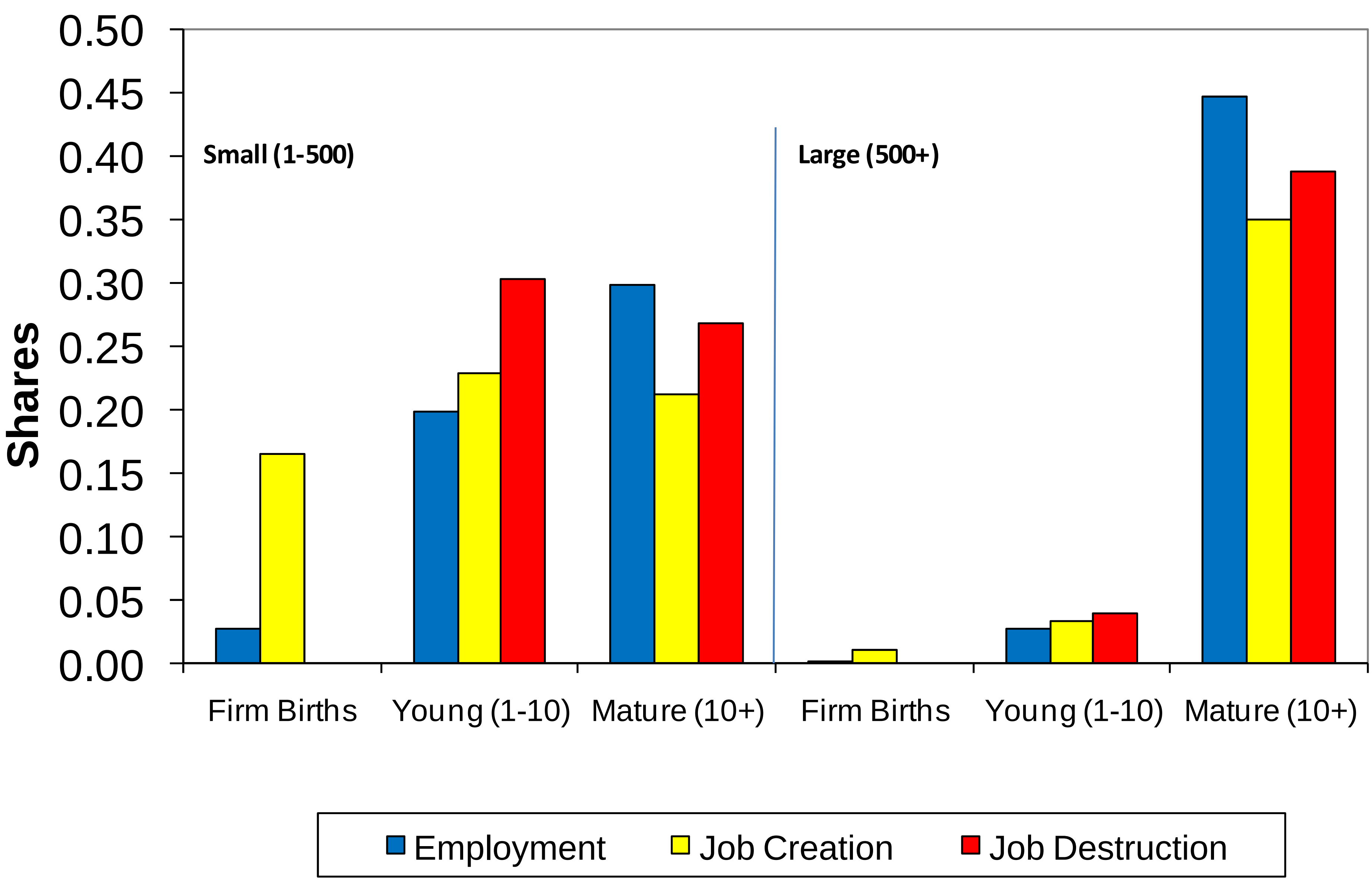
“Up or Out” Dynamics of Young Businesses (robust to size controls)



“Out” component…



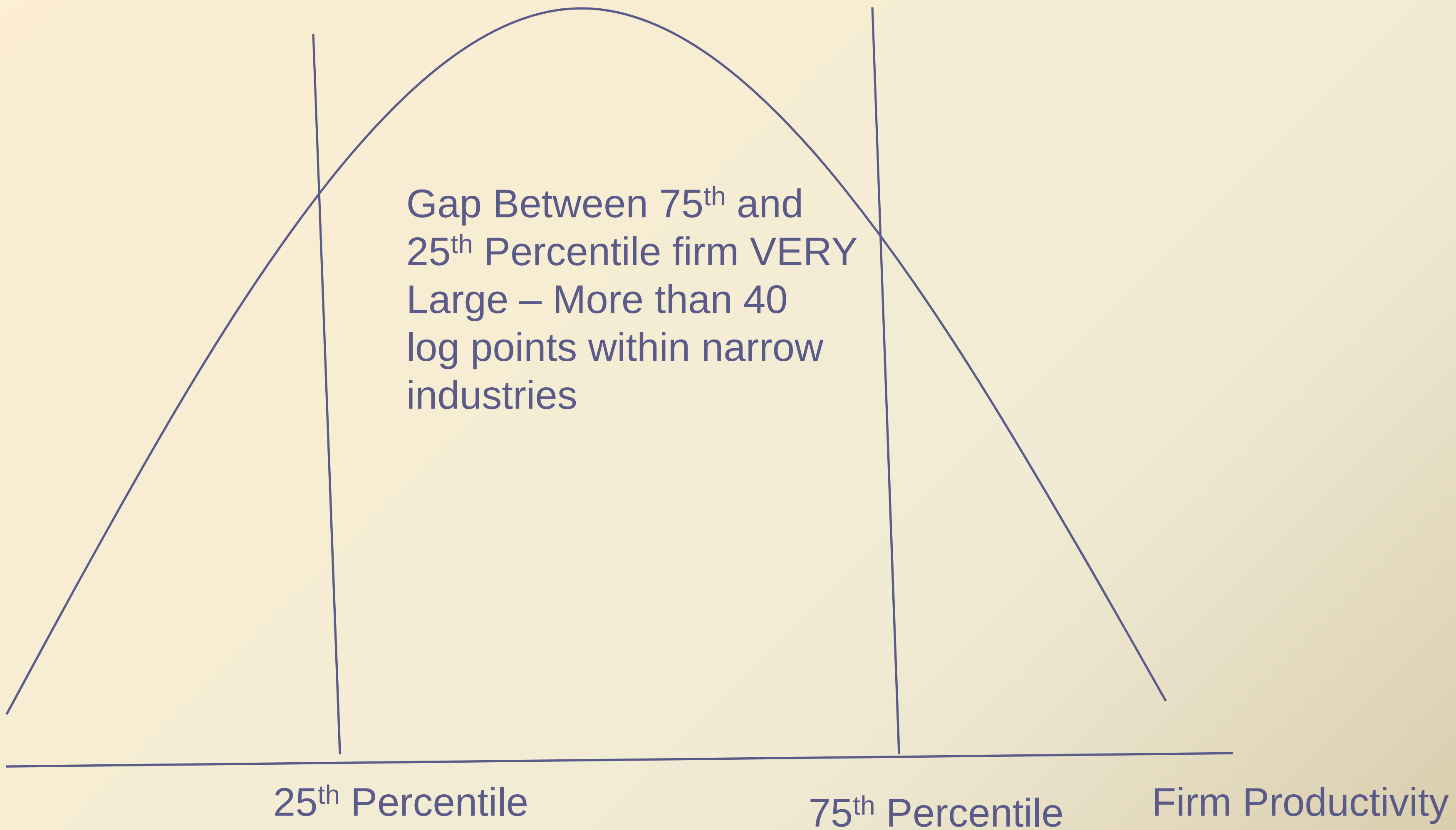
Shares of Creation, Destruction and Employment



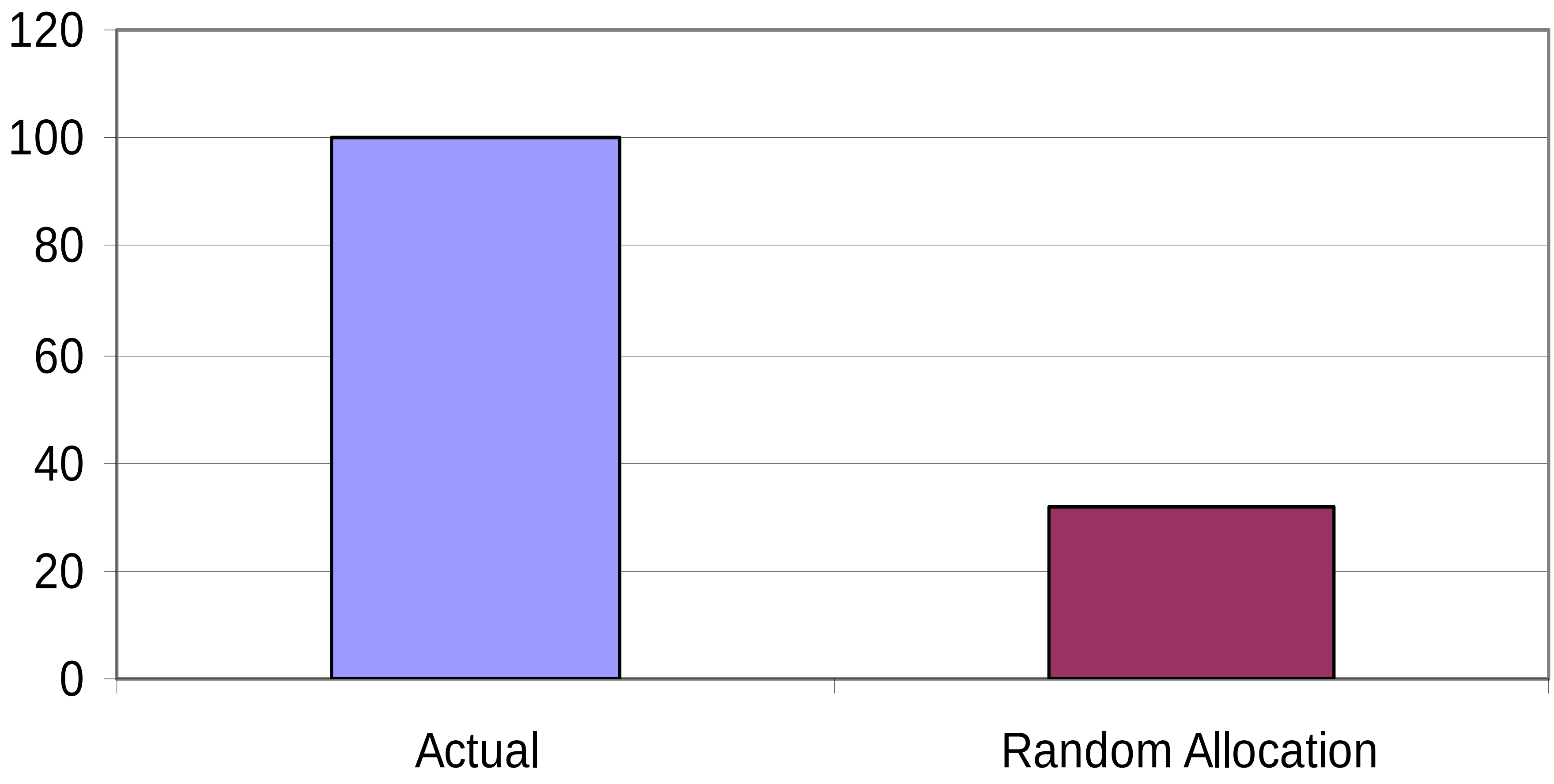
What accounts for cross sectional and dynamic patterns?

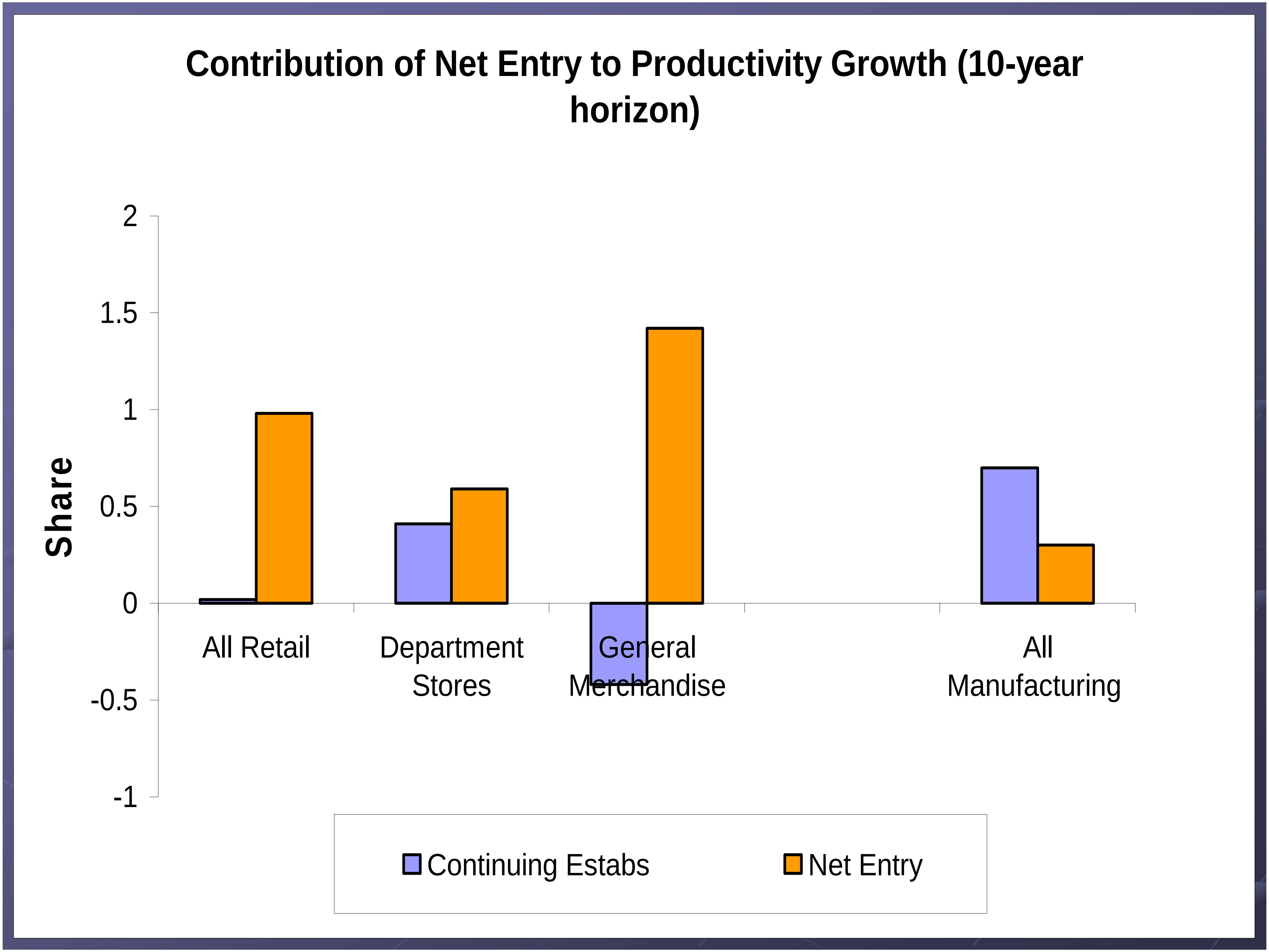
- Very skewed size distribution
- Constant state of churning
  - Wave of entering firms contributes substantially to job creation each year
  - Most exit
  - Conditional on survival, young businesses grow quickly
  - Even amongst large, mature businesses high pace of churning of jobs and businesses

Firm's subject to LARGE and persistent productivity shocks…in healthy economy constantly reallocating outputs and inputs away from less productive to more productive businesses



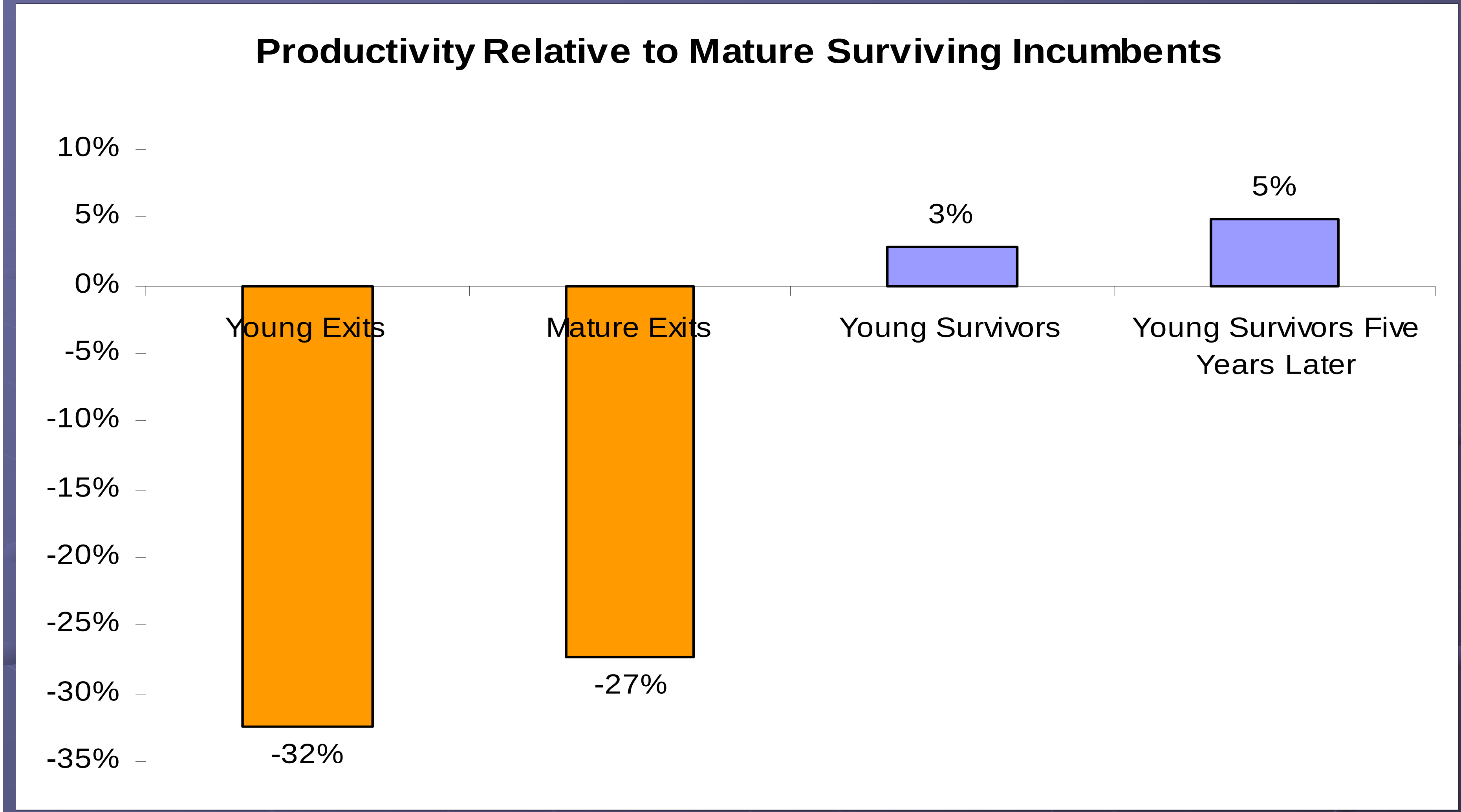
U.S. Labor Productivity: Comparison Between Actual and Random Allocation of Size of Businesses





Contribution of Net Entry to Productivity Growth (10-year horizon)
Share
2
1.5
1
0.5
0
-0.5
-1
All Retail
Department Stores
General Merchandise
All Manufacturing
Continuing Estabs
Net Entry

In Retail Trade, selection and learning effects play critical roles….



Based on regression on establishment-level data for U.S. Retail Trade (FHK(2006))

Why is there so much dispersion in productivity across businesses in narrowly defined sectors?

- Background facts:
  - Interquartile range of log of Revenue TFP (TFPR) is 0.29
  - Interquartile range of log of Revenue Labor Productivity (RLP) is 0.65
  - Dispersion in TFPQ, TFPR, and output price within narrow product classes (7-digit) in U.S. (Source: FHS (2008)):
    - Std. Dev of log(TFPQ) is: 0.26
    - Std. Dev of log(TFPR) is: 0.22
    - Std. Dev of log(RLP) is: 0.65
    - Std. Dev of log(P) is: 0.18
    - Std. Dev of log(Q) is: 1.05
    - Corr(log(TFPQ),log(P)) is: -0.54
    - Corr(log(TFPQ),log(Q)) is: 0.28
    - Corr(log(TFPQ),log(TFPR)) is: 0.75
    - Corr(log(TFPQ),log(RLP)) is: 0.56

Frictions + Distortions

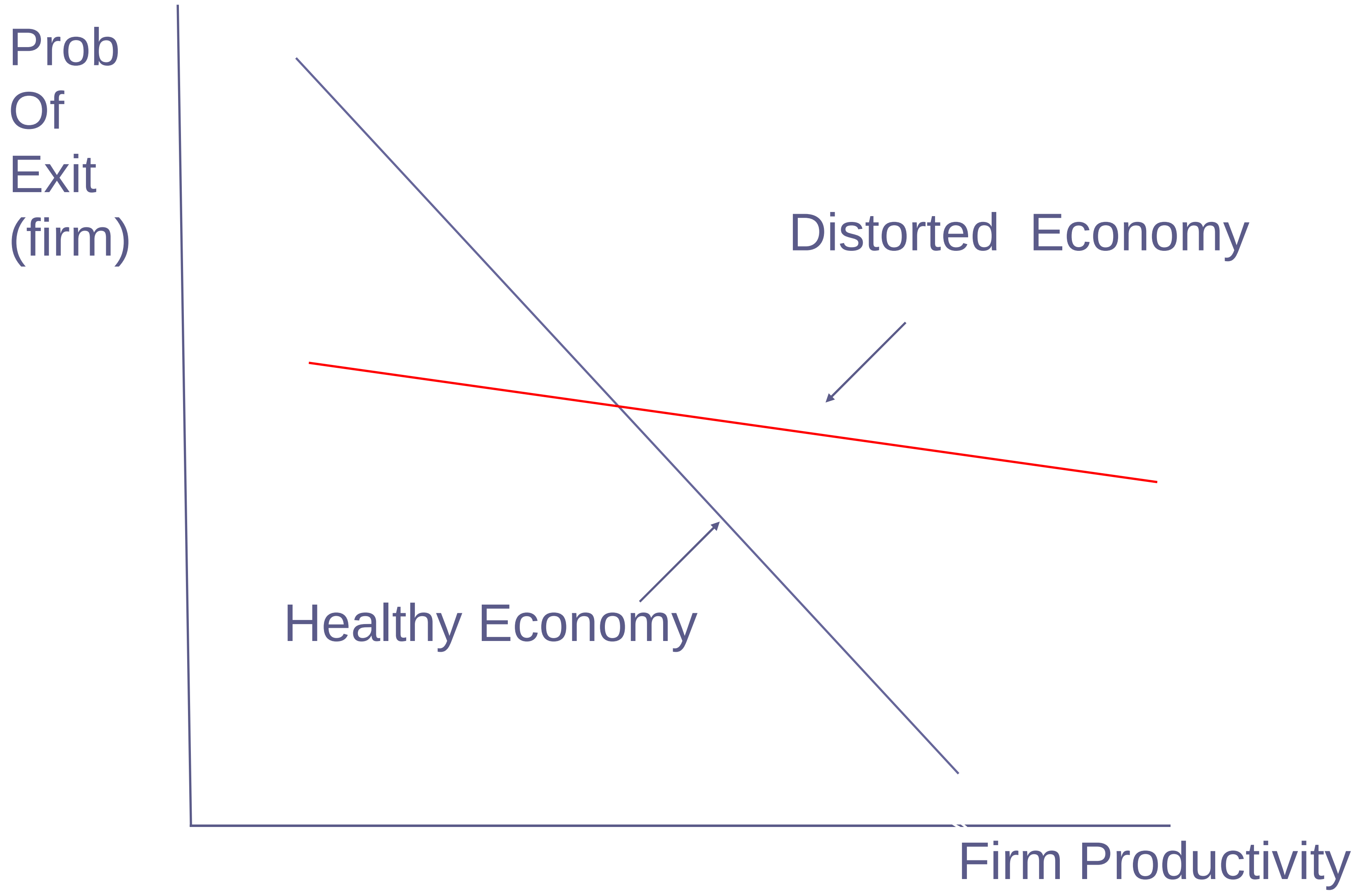
- Costs of Entry (and exit)
  - Including costs of entering new markets
  - Hopenhayn (1992), Melitz (2003), Melitz and Ottaviano (2005)
- Learning (initial conditions and after changing products/processes)
  - Jovanovic (1982) and Ericson and Pakes (1998)
  - Experimentation
- Adjustment costs for factors of production (capital, labor, intangible capital)
  - Convex vs. Nonconvex
- Economies of scope and control
- Product Differentiation:
  - Horizontal (e.g., spatial) vs. Vertical
- Output and input price dispersion and determination
- Imperfections in product, labor, capital, credit markets
- Distortions to all of the above + market institutions
  - Idiosyncratic distortions as in Banerjee and Duflo (2003), Restuccia and Rogerson (2007), Hsieh and Klenow (2007)

What frictions matter the most?

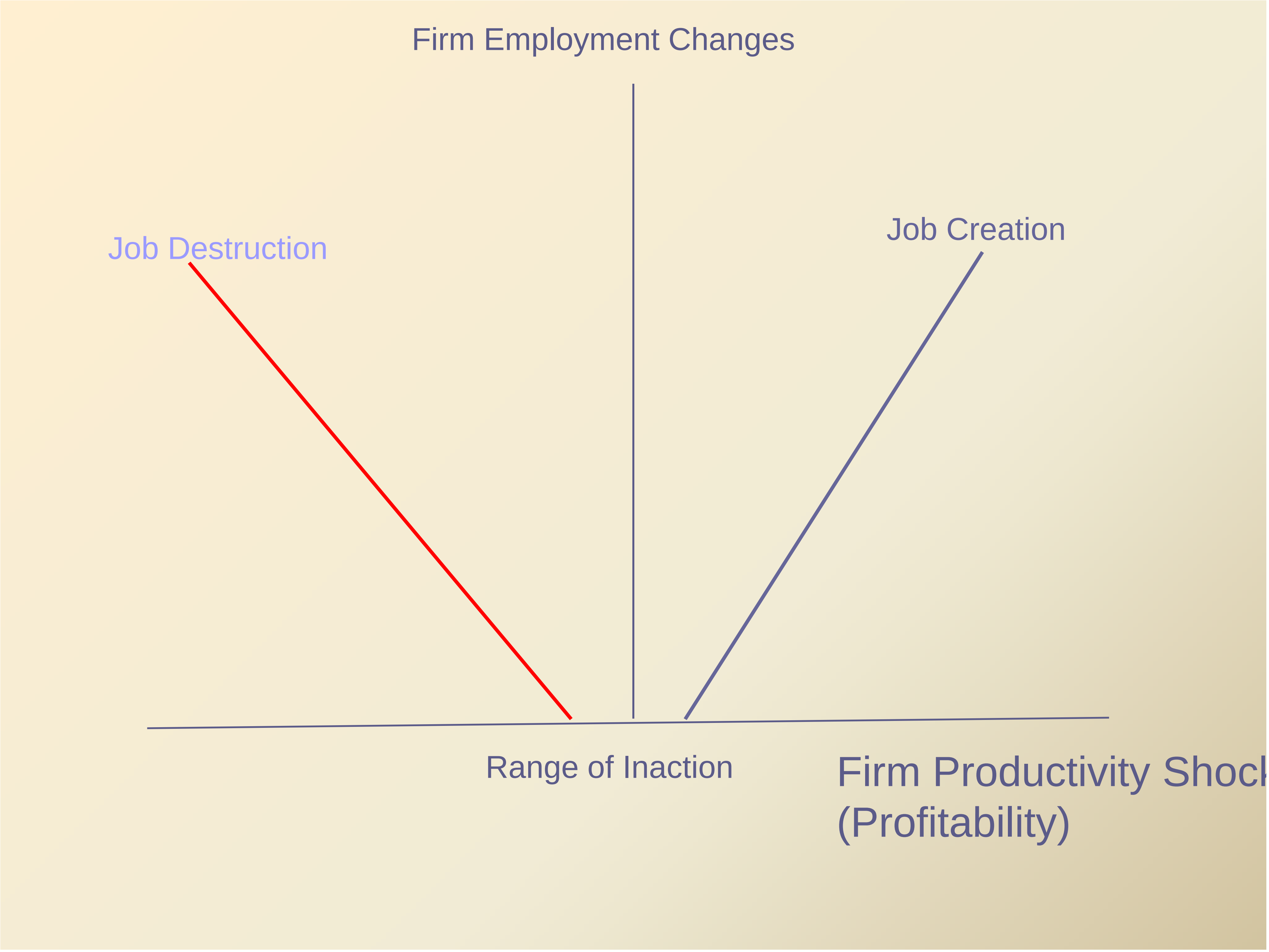
- Many studies showing evidence of entry costs, labor adjustment costs, capital adjustment costs, trade costs, product differentiation, and so on.
- Many open questions and issues:
  - Not practical to include all frictions in all models – but caution about identification since we are all using same data
  - How do frictions vary across advanced vs. emerging vs. transition?
- Important to distinguish between those frictions that yield some plants persistently higher productivity than others as opposed to adjustment dynamics

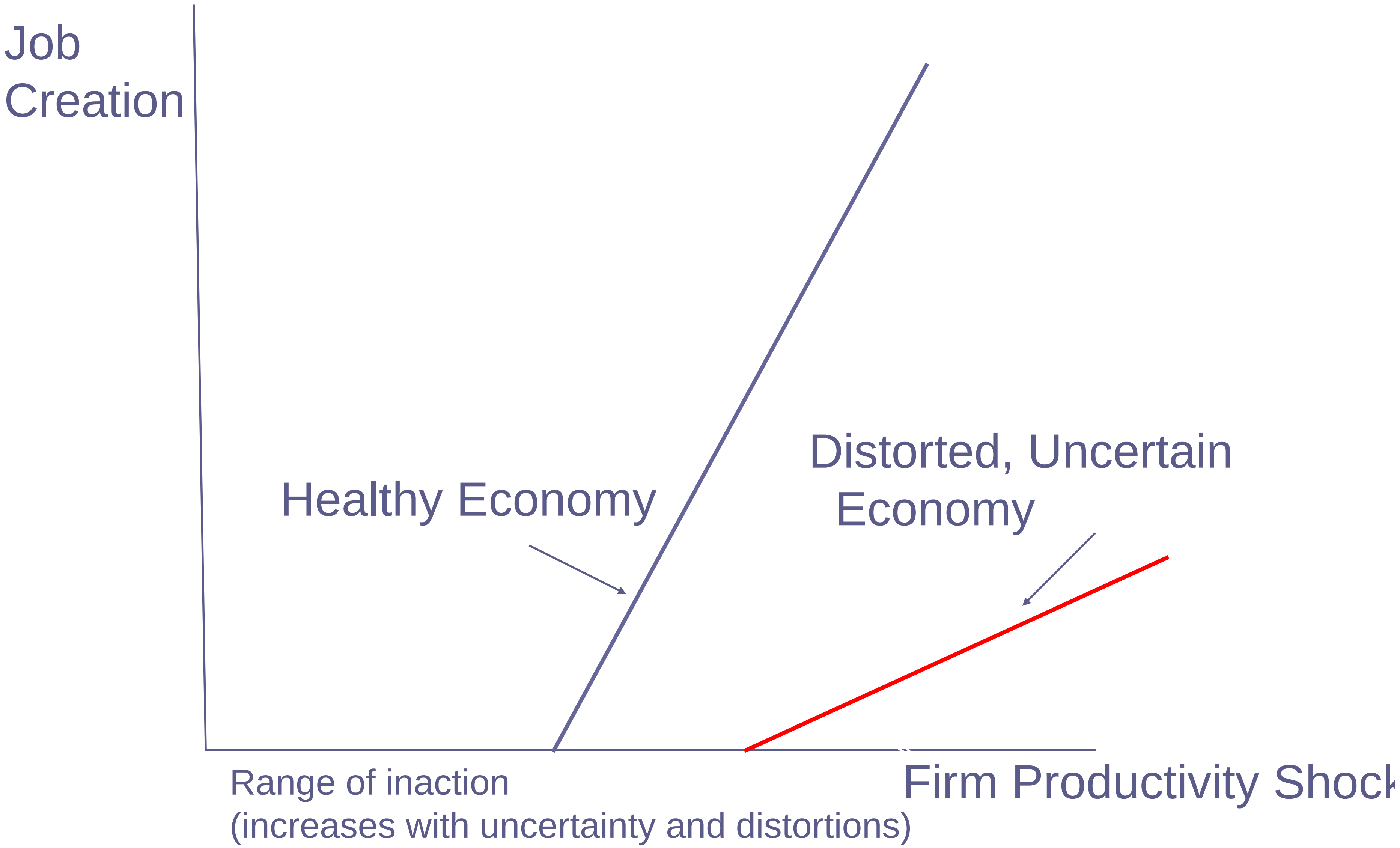
Lots of margins for distortions…

- Cross sectional misallocation
- Dynamic distortions:
  - Startups
  - Post-entry up or out dynamic
  - Creative Destruction
- Secular vs. Cyclical Distortions



Prob
Of
Exit
(firm)
Distorted Economy
Healthy Economy
Firm Productivity


Firm Employment Changes
Job Destruction
Job Creation
Range of Inaction
Firm Productivity Shock
(Profitability)


Job
Creation
Healthy Economy
Distorted, Uncertain
Economy
Range of inaction
(increases with uncertainty and distortions)
Firm Productivity Shock

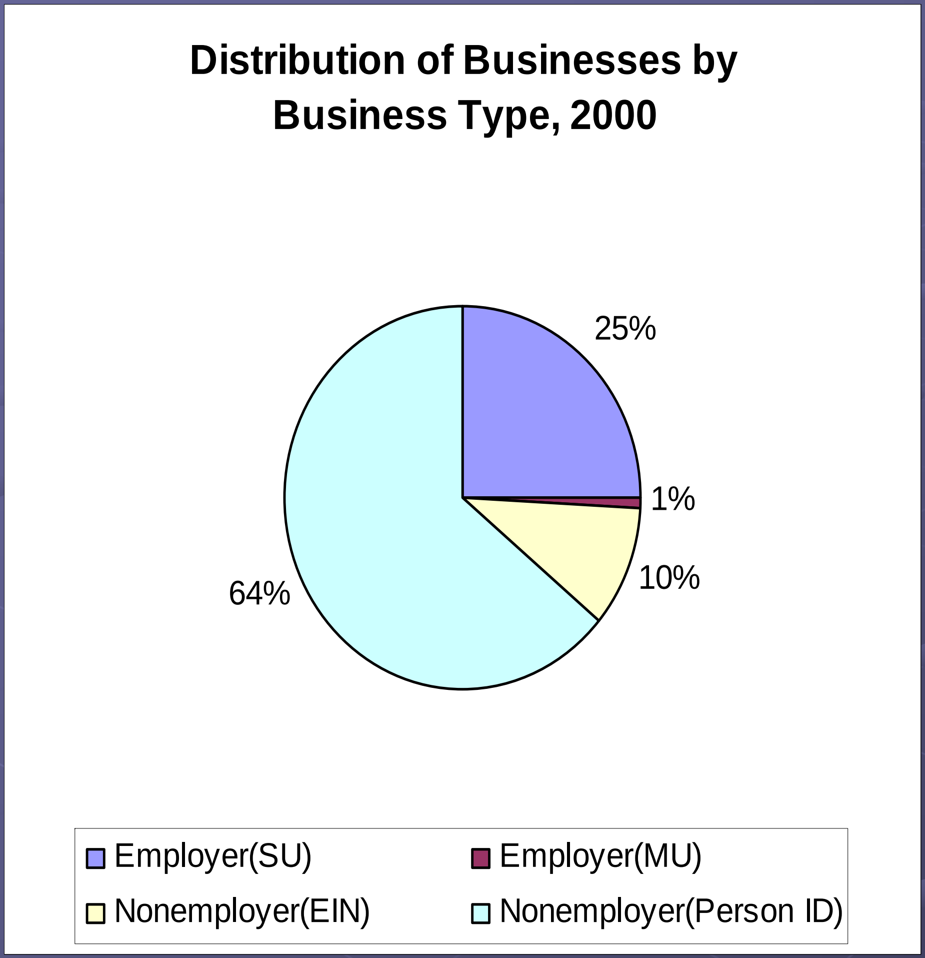
Taking Stock

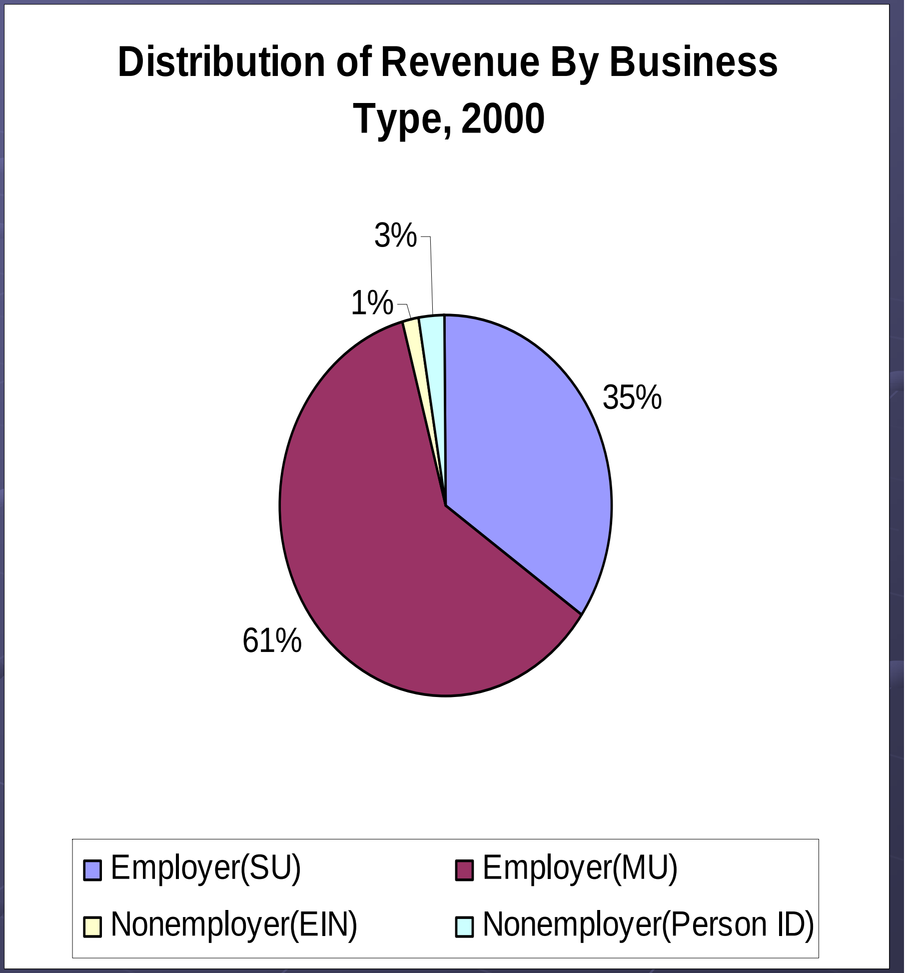
- High pace of churning of businesses within narrowly defined industries
- Startups and young businesses play an important role in these dynamics
- Up or out dynamics
- These dynamics connected to productivity (and demand) dynamics at the micro level
- Identifying the frictions and how they vary across industry, time, and country ongoing activity
- But what about before entry?

“Before” Entry….

- Entrepreneurial dynamics starts at micro business level
- Entrepreneurs start with an idea – often while employed elsewhere
- New longitudinal databases at U.S. Census Bureau tracking this process
  - ILBD: Nonemployers (e.g., sole props without employees) + Employers
  - LEHD/SED: Tracking transitions from W&S jobs to self-employed jobs

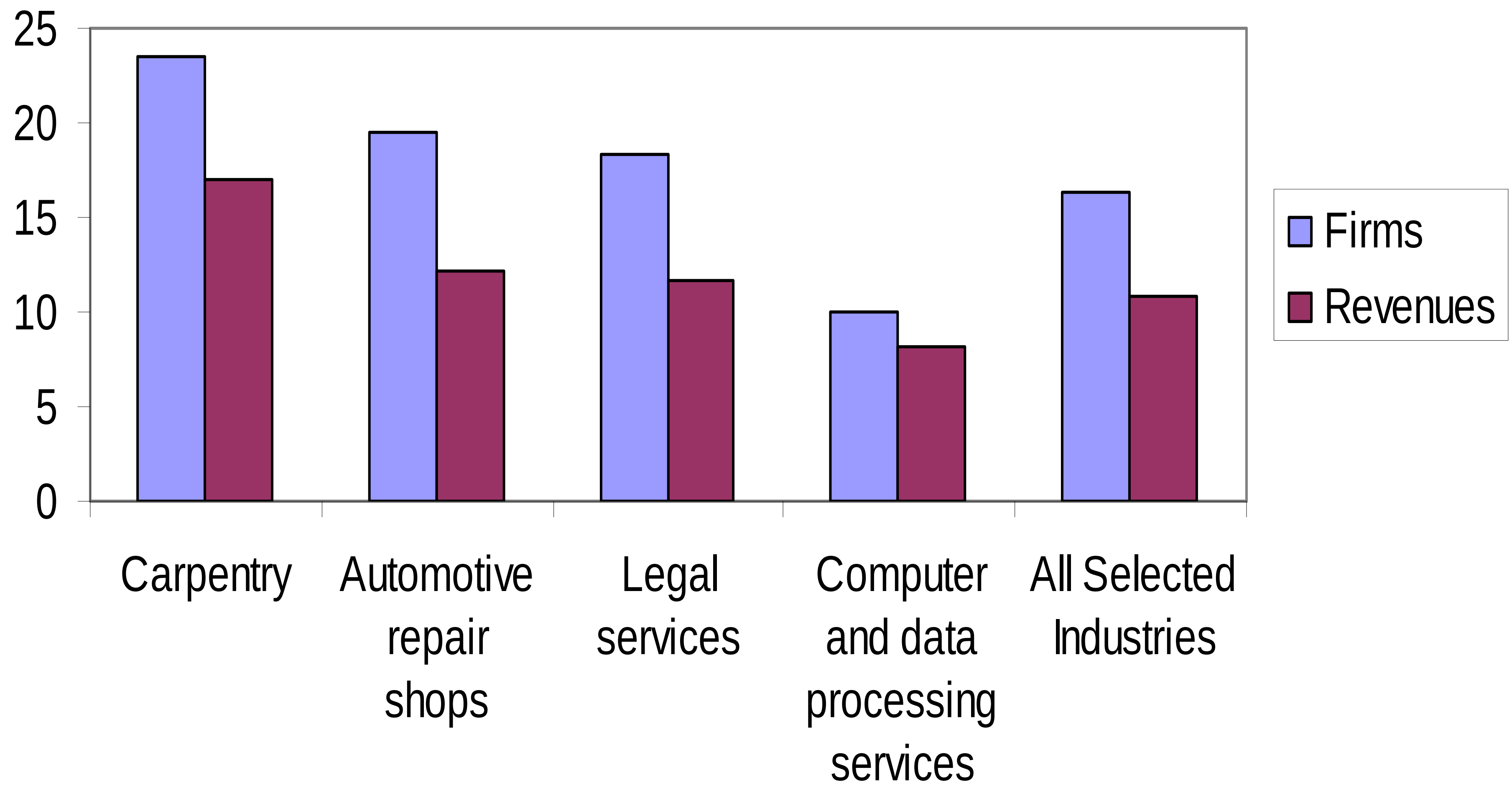
Micro Businesses constitute a large share of businesses and a small share of revenue…





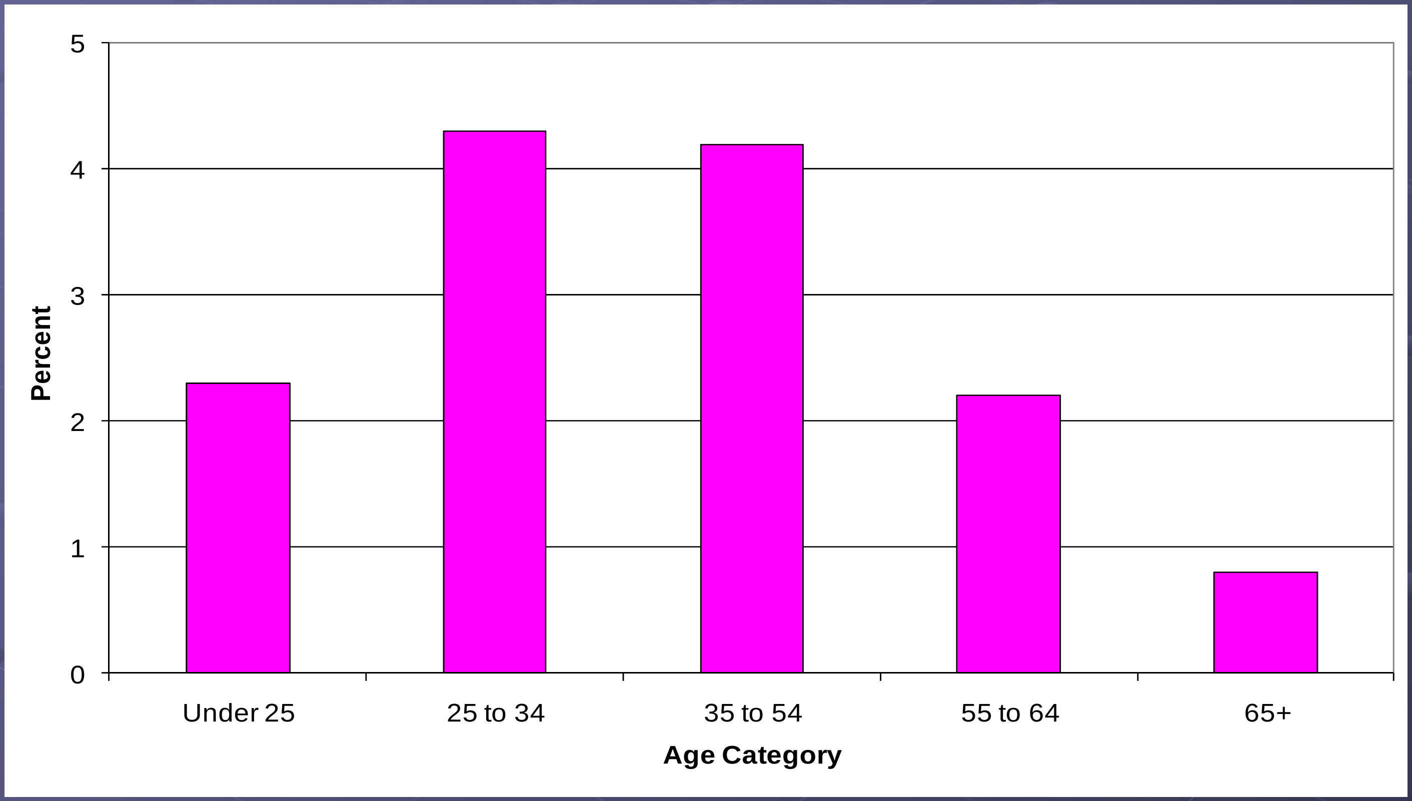
Source: Davis et. al. (2008)

Shares of New Employer Businesses in 1997 with Pre-History as Nonemployer Businesses



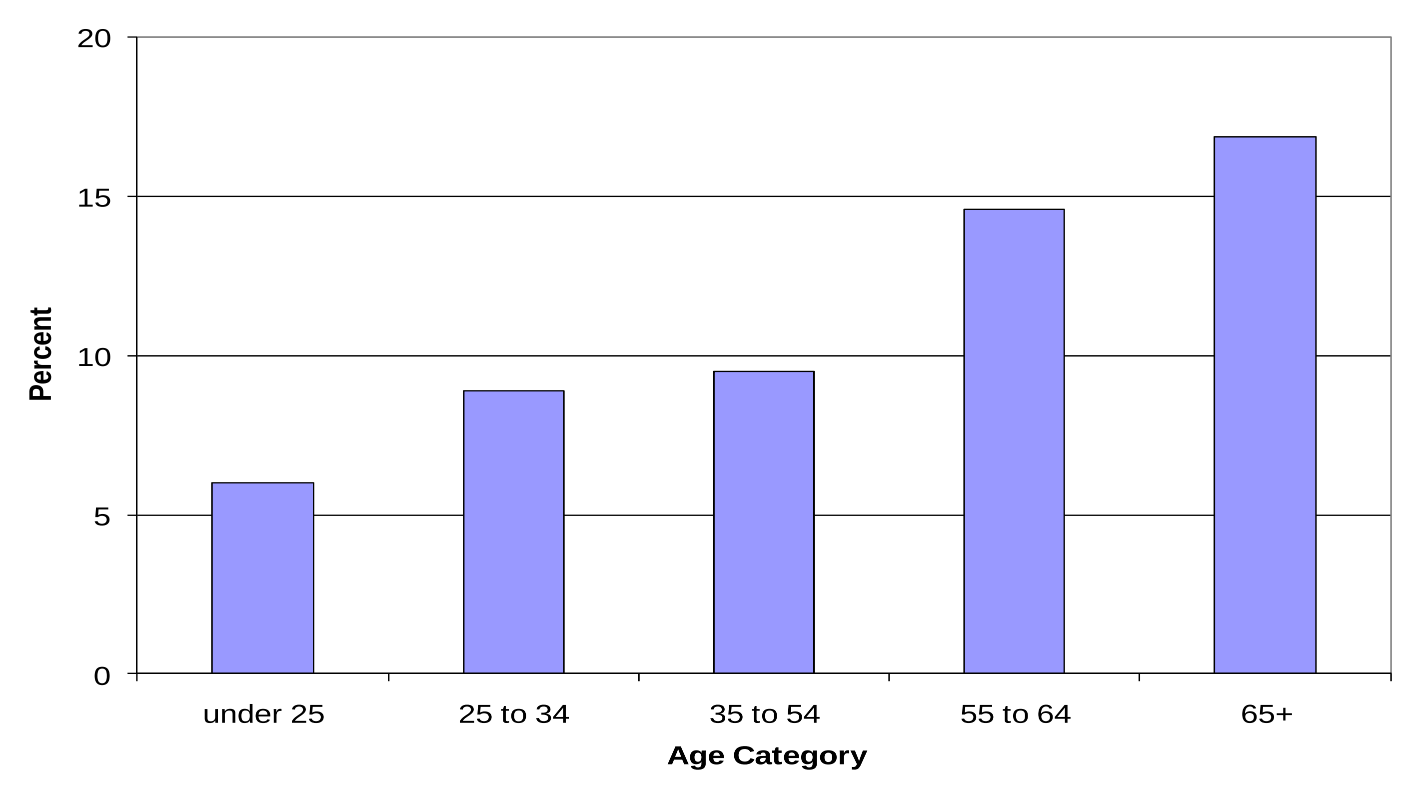
Source: Davis et al. (2008)

LEHD: Self Employment Dynamics Database



Fraction of workforce that try self-employment…

LEHD: Self Employment Dynamics Database



Of those how many move to full self-employment…

Nuts and Bolts

- How do we model and analyze the extent to which an economy exhibits patterns consistent with static and dynamic allocative efficiency?
- How do we explore empirically?
- How do we measure outputs, inputs, prices and productivity?
- What data are available for U.S. and other countries?
- In what follows, we provide some examples of all of these issues…

Standard Heterogeneous-Producer Industry Models

The Workhorse:

- Producers i differ in a profitability component ω_i, usually taken to represent costs/productivity
- Profits depend on ω_i and industry state S: $\pi_i = \pi_i(\omega_i, S) \quad \omega_i \sim G(\omega)$
- There is some critical ω^* such that producers with $\omega_i < \omega^*$ have NPVs below outside option and therefore exit the industry
- Industry state S typically depends on endogenously determined distribution of ω_i among producers (add'l free entry assumption)
- Examples: Jovanovic (1982), Hopenhayn (1992), Melitz (2003), Asplund and Nocke (2007)

Closely Related Issue – Size Distribution of Activity

- $\pi_i = \pi_i(\omega_i, S)$ has curvature either from decreasing returns (e.g., Lucas (1978)) or product differentiation (e.g., Melitz (2003))
- Curvature pins down the size distribution of activity and permits studying the evolution of the size distribution of activity
- In healthy market economies, most productive plants are the largest – allocative efficiency
- Active literature attempting to explain cross country differences in productivity (e.g., Hsieh and Klenow (2009)) using distortions on this margin

Model: Melitz/Ottaviano (2005) and FHS (2008)

Industry is comprised of a continuum of producers of measure N.

Each produces a single variety (indexed by i) of industry product.

Representative consumer's utility function

$$U = y + \int_{i\in I}(\alpha+\delta_i)q_i di - \frac{1}{2}\eta\left(\int_{i\in I} q_i di\right)^2 - \frac{1}{2}\gamma\int_{i\in I} q_i^2 di$$

$$= y + \alpha\int_{i\in I} q_i di - \frac{1}{2}\left(\eta+\frac{\gamma}{N}\right)\left(\int_{i\in I} q_i di\right)^2 + \int_{i\in I}\delta_i q_i di - \frac{1}{2}\gamma\int_{i\in I}(q_i-\overline{q})^2 di$$

$\alpha > 0$, $\eta > 0$, and $\gamma \geq 0$.

y = numeraire good

δ_i = variety-specific, mean-zero taste shifter

q_i = quantity of good i consumed

$$\overline{q} = \frac{1}{N}\int_{i\in I} q_i di$$

The implied demand curve:

$$q_i = \frac{1}{\eta N+\gamma}\alpha - \frac{1}{\gamma}\frac{\eta N}{\eta N+\gamma}\overline{\delta} + \frac{1}{\gamma}\frac{\eta N}{\eta N+\gamma}\overline{p} + \frac{1}{\gamma}\delta_i - \frac{1}{\gamma}p_i$$

Model: Supply

Production Function: $q_i = \omega_i x_i$

Producers face (potentially idiosyncratic) factor price w_i

$\Rightarrow$ marginal cost = w_i/ω_i

Profits:

$$\pi_i = \left(\frac{1}{\eta N + \gamma}\alpha - \frac{1}{\gamma}\frac{\eta N}{\eta N + \gamma}\bar{\delta} + \frac{1}{\gamma}\frac{\eta N}{\eta N + \gamma}\bar{p} + \frac{1}{\gamma}\delta_i - \frac{1}{\gamma}p_i \right)\left(p_i - \frac{w_i}{\omega_i} \right)$$

Profit-maximizing price (constant marginal cost c_i):

$$p_i = \frac{1}{2}\frac{\gamma}{\eta N + \gamma}\alpha - \frac{1}{2}\frac{\eta N}{\eta N + \gamma}\bar{\delta} + \frac{1}{2}\frac{\eta N}{\eta N + \gamma}\bar{p} + \frac{1}{2}\delta_i + \frac{1}{2}\frac{w_i}{\omega_i}$$

Deviation from industry-average price:

$$p_i - \bar{p} = \frac{1}{2}\left(\delta_i - \bar{\delta}\right) + \frac{1}{2}\left(\frac{w_i}{\omega_i} - \left(\frac{w}{\omega} \right) \right)$$

Maximized profits:

$$\pi_i = \frac{1}{4\gamma}\left(\frac{\gamma}{\eta N + \gamma}\alpha - \frac{\eta N}{\eta N + \gamma}\bar{\delta} + \frac{\eta N}{\eta N + \gamma}\bar{p} + \delta_i - \frac{w_i}{\omega_i} \right)^2$$

Model: Equilibrium

Equilibrium Condition 1: The marginal producer in the industry makes zero profits

Define "profitability index" $\phi_i \equiv \delta_i - \frac{w_i}{\omega_i}$. Then marginal producer has index equal to:

$$\phi^* = -\frac{\gamma}{\eta N + \gamma}\alpha + \frac{\eta N}{\eta N + \gamma}\overline{\delta} - \frac{\eta N}{\eta N + \gamma}\overline{p}$$

Profits can be rewritten in terms of this marginal profitability level

$$\pi_i = \frac{1}{4\gamma}(\phi_i - \phi^*)^2$$

Profits increase in demand (δ_i) and efficiency (ω_i), decrease in factor price (w_i)

Equilibrium Condition 2: Potential entrants decide whether to pay sunk entry cost s to learn δ_i, ω_i, w_i. Expected value of entry is 0.

$$V^e = \int_0^{w_u}\int_{\omega_l}^{\omega_u}\int_{\phi^*+\frac{w}{\omega}}^{\delta_e} \frac{1}{4\gamma}(\phi_i - \phi^*)^2 f(\delta,\omega,w)\,d\delta\, d\omega\, dw - s = 0$$

Selection effect:

- Only high-profitability producers operate in equilibrium
- Low types exit

Sunk costs, market power and dispersion:

- Sunk costs make entry costly
- Curvature yields equilibrium size distribution

Many models of selection also include fixed costs of operating each period

Model: Empirical Implications

Output-based productivity:

$$TFPQ_i = \frac{q_i}{x_i} = \frac{\omega_i x_i}{x_i} = \omega_i$$

Revenue-based productivity (literature standard):

$$TFPR_i = \frac{p_i q_i}{x_i} = p_i \omega_i = \frac{1}{2}\frac{\gamma\alpha}{\eta N + \gamma}\omega_i + \frac{1}{2}\frac{\eta N}{\eta N + \gamma}\left(\overline{p} - \overline{\delta}\right)\omega_i + \frac{1}{2}\delta_i \omega_i + \frac{1}{2} w_i$$

Plant price deviation from industry deflator depends on both demand (enters positively into profits) and costs (enter negatively):

$$p_i - \overline{p} = \frac{1}{2}\left(\delta_i - \overline{\delta}\right) + \frac{1}{2}\left(\frac{w_i}{\omega_i} - \overline{\left(\frac{w}{\omega}\right)}\right)$$

Comparative Statics:

- $\frac{d\phi^*}{d\gamma} < 0$: Lower substitutability (higher γ) lowers ϕ^*
- $\frac{d\phi^*}{ds} < 0$: Higher sunk entry cost lowers ϕ^*

Measurement of Plant-level Productivity

$$tfp_i = y_i - \alpha_l l_i - \alpha_k k_i - \alpha_m m_i - \alpha_e e_t$$

All variables in logs, difficult measurement Issues on outputs and inputs and factor elasticities

Typical to assume Cobb-Douglass or to have Divisia index approach approximation

Measurement issues

- Factor inputs:
  - Labor quality
  - Capital stock (book value vs. perpetual inventory)
- Factor elasticities:
  - Cost shares, estimated elasticities using OLS, IV, proxy methods
  - All typically estimate factor elasticities at the industry level
    - Time invariant with estimated approach typically given Cobb-Douglass assumptions
  - Estimates vary in literature but measures of TFP highly correlated across these methods. Other issues (below) appear to matter more.
- Plant-level heterogeneity in output and input prices
- Plant-level heterogeneity in factor elasticities

Example of proxy method

$$y_{jt} = \beta_0 + \beta_k k_{jt} + \beta_a a_{jt} + \beta_l l_{jt} + \omega_{jt} + \eta_{jt} \quad (24)$$

$$i_{jt} = i(k_{jt}, a_{jt}, \omega_{jt}, \Delta_t) = i_t(k_{jt}, a_{jt}, \omega_{jt}). \quad (27)$$

$$\omega_{jt} = h_t(k_{jt}, a_{jt}, i_{jt}). \quad (28)$$

$$y_{jt} = \beta_0 + \beta_k k_{jt} + \beta_a a_{jt} + \beta_l l_{jt} + h_t(k_{jt}, a_{jt}, i_{jt}) + \eta_{jt}. \quad (29)$$

$$y_{jt} - \beta_l l_{jt} = \beta_0 + \beta_k k_{jt} + \beta_a a_{jt} + \omega_{jt} + \eta_{jt}. \quad (33)$$

$$\begin{aligned} y_{jt} - \beta_l l_{jt} & \\ &= \beta_0 + \beta_k k_{jt} + \beta_a a_{jt} + g(\omega_{jt-1}) + \xi_{jt} + \eta_{jt} && (34a) \\ &= \beta_0 + \beta_k k_{jt} + \beta_a a_{jt} + g(\phi_{jt-1} - \beta_0 - \beta_k k_{jt-1} - \beta_a a_{jt-1}) + \xi_{jt} + \eta_{jt} \\ &= \beta_k k_{jt} + \beta_a a_{jt} + \tilde{g}(\phi_{jt-1} - \beta_k k_{jt-1} - \beta_a a_{jt-1}) + \xi_{jt} + \eta_{jt}, && (34b) \end{aligned}$$

Depends critically on the invertibility amongst other assumptions

Start with Foster, Haltiwanger and Syverson (2008)

- Source data: Census of Manufactures
  - High quality coverage
  - Limited number of products with physical quantity data

Correlations								
Variables	Trad'l. Output	Revenue Output	Physical Output	Price	Trad'l. TFP	Revenue TFP	Physical TFP	Capital
Traditional Output	1.00							
Revenue Output	0.99	1.00						
Physical Output	0.98	0.99	1.00					
Price	-0.03	-0.03	-0.19	1.00				
Traditional TFP	0.19	0.18	0.15	0.13	1.00			
Revenue TFP	0.17	0.21	0.18	0.16	0.86	1.00		
Physical TFP	0.17	0.20	0.28	-0.54	0.64	0.75	1.00	
Capital	0.86	0.85	0.84	-0.04	0.00	-0.00	0.03	1.00
Standard Deviations								
Standard Deviations	1.03	1.03	1.05	0.18	0.21	0.22	0.26	1.14

Measuring Plant-Level Demand

Estimate product demand curves; plant-specific residual is idio. demand

$$\ln q_{it} = \alpha_o + \alpha_1 \ln p_{it} + \alpha_2 \ln(INCOME_{mt}) + \sum_t \alpha_t YEAR_t + \eta_{it}$$

q_{it}—physical output of plant i in year t
p_{it}—plant unit price
$INCOME_{mt}$—average income in the plant's local market m
$YEAR_t$—year dummy
η_{it}—plant-year disturbance term

Plant demand:

$$\hat{\delta}_{it} = \hat{\eta}_{it} + \hat{\alpha}_2 \ln(INCOME_{mt}) = \ln q_{it} - \hat{\alpha}_o - \hat{\alpha}_1 \ln p_{it} - \sum_t \hat{\alpha}_t YEAR_t$$

I.e., residual is plant quantity sold that can't be accounted for by unit price or local income differences

- Use $TFPQ_{it}$ to instrument for prices (captures production costs)

	IV Estimation			OLS Estimation	
Product	Price Coefficient (α_1)		Income Coefficient (α_2)	Price Coefficient (α_1)	Income Coefficient (α_2)
Boxes	-3.02		-0.03	-2.19	-0.03
	0.17	[0.61]	*0.02*	*0.12*	*0.02*
Bread	-3.09		0.12	-0.89	0.07
	0.42	[0.33]	*0.05*	*0.15*	*0.04*
Carbon Black	-0.52		-0.21	-0.57	-0.21
	0.38	[0.50]	*0.11*	*0.21*	*0.11*
Coffee	-3.63		0.22	-1.03	0.20
	0.98	[0.41]	*0.14*	*0.32*	*0.13*
Concrete	-5.93		0.13	-0.83	0.15
	0.36	[0.10]	*0.01*	*0.09*	*0.01*
Hardwood Flooring	-1.67		-0.20	-0.87	-0.24
	0.48	[0.61]	*0.18*	*0.47*	*0.18*
Gasoline	-1.42		0.23	-0.16	0.23
	2.72	[0.20]	*0.07*	*0.80*	*0.07*
Block Ice	-2.05		0.00	-0.63	0.16
	0.46	[0.32]	*0.11*	*0.20*	*0.07*
Processed Ice	-1.48		0.18	-0.70	0.16
	0.27	[0.37]	*0.03*	*0.13*	*0.03*
Plywood	-1.21		-0.23	-1.19	-0.23
	0.14	[0.89]	*0.10*	*0.13*	*0.10*
Sugar	-2.52		0.76	-1.04	0.72
	1.01	[0.15]	*0.13*	*0.55*	*0.12*

Dependent Variable	Five-Year Horizon		Implied One-Year Persistence Rates	
	Unweighted Regression	Weighted Regression	Unweighted Regression	Weighted Regression
Traditional TFP	0.249 *0.017*	0.316 *0.042*	0.757	0.794
Revenue TFP	0.277 *0.021*	0.316 *0.042*	0.774	0.794
Physical TFP	0.312 *0.019*	0.358 *0.049*	0.792	0.814
Price	0.365 *0.025*	0.384 *0.066*	0.817	0.826
Demand Shock	0.619 *0.013*	0.843 *0.021*	0.909	0.966

Variable	Plant Age Dummies Exit	 Entry	 Young	 Medium
Unweighted Regressions				
Traditional TFP	-0.0211	0.0044	0.0074	0.0061
	0.0042	*0.0044*	*0.0048*	*0.0048*
Revenue TFP	-0.0220	0.0133	0.0075	0.0028
	0.0044	*0.0047*	*0.0051*	*0.0053*
Physical TFP	-0.0186	0.0128	0.0046	-0.0039
	0.0050	*0.0053*	*0.0058*	*0.0062*
Price	-0.0034	0.0005	0.0029	0.0067
	0.0031	*0.0034*	*0.0038*	*0.0042*
Demand Shock	-0.3466	-0.5557	-0.3985	-0.3183
	0.0227	*0.0264*	*0.0263*	*0.0267*

Determinants of Market Selection

Specification:	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Traditional TFP	-0.073						
	0.015						
Revenue TFP		-0.063					
		0.014					
Physical TFP			-0.040			-0.062	-0.034
			0.012			*0.014*	*0.012*
Prices				-0.021		-0.069	
				0.018		*0.021*	
Demand Shock					-0.047		-0.047
					0.003		*0.003*

Note: Much greater dispersion in demand shocks than physical TFP

Establishment-level Productivity Empirical Patterns

- Dispersion (large), persistence (high) evolution (consistent with learning and selection)
- Selection
  - Lower productivity plants exit
  - Other determinants of productivity matter
  - Open questions: Impact of distortions on selection?
    - Models like Melitz (2003) and Restuccia and Rogerson (2007) imply reduced distortions will improve selection
    - Eslava et. al. (2009) find evidence that trade liberalization improves market selection
- These patterns both support basic models and can be used to test and estimate models
- One other approach has to been to explore the covariance between size and productivity within industries.
  - Basic prediction of virtually all models is positive correlation between size and profitability/productivity

Size/productivity relationship within industries

$$\Omega_t = \sum_i s_{it}\omega_{it}$$
$$= (1/N_t)\sum_i \omega_{it} + \sum_i (s_{it} - (1/N_t)\sum_i s_{it})(\omega_{it} - (1/N_t)\sum_i \omega_{it}))$$

Olley and Pakes (1996) decomposition

$$\Delta\,\Omega_t = \sum_i s_{it}\omega_{it} - \sum_i s_{it-1}\omega_{it-1}$$
$$= \sum_{i\in C}\overline{s_{it}}\Delta\omega_{it} + \sum_{i\in C}\Delta s_{it}\overline{(\omega_{it}} - \overline{\Omega_t}) + \sum_{i\in N} s_{it}\left(\omega_{it} - \overline{\Omega_t}\right) - \sum_{i\in X} s_{it-1}\left(\omega_{it-1} - \overline{\Omega_t}\right)$$
$$= within \quad + \quad reallocation \quad + \quad entry \quad - \quad exit$$

Modified Baily, Hulten and Campbell (1992) and Griliches and Regev (1995) decomposition

Comments on Decomposition in Literature

- Some questions about how to interpret industry-level index defined in this manner
  - Typical check (e.g., BHC and FHK) to see how this index performs relative to standard aggregate ***industry*** measures
    - Common result – magnitudes very similar and correlations high in most studies
    - Cautions:
      - These measures very sensitive to measurement error since depend on measuring within industry productivity (log) level dispersion accurately
      - Not appropriate for decompositions that exploit between industry variation (measurement and index problems)
  - Standard decomposition summarizes changes in activity weighted micro distribution
  - Decompositions more closely tied to aggregate welfare and productivity have been developed (Petrin and Levinsohn (2008), Basu and Fernald (2002)
  - Alternatively, these decompositions can be used as moments to match in a calibration or indirect inference approach (see, e.g., Bartelsman, Haltiwanger and Scarpetta (2009)

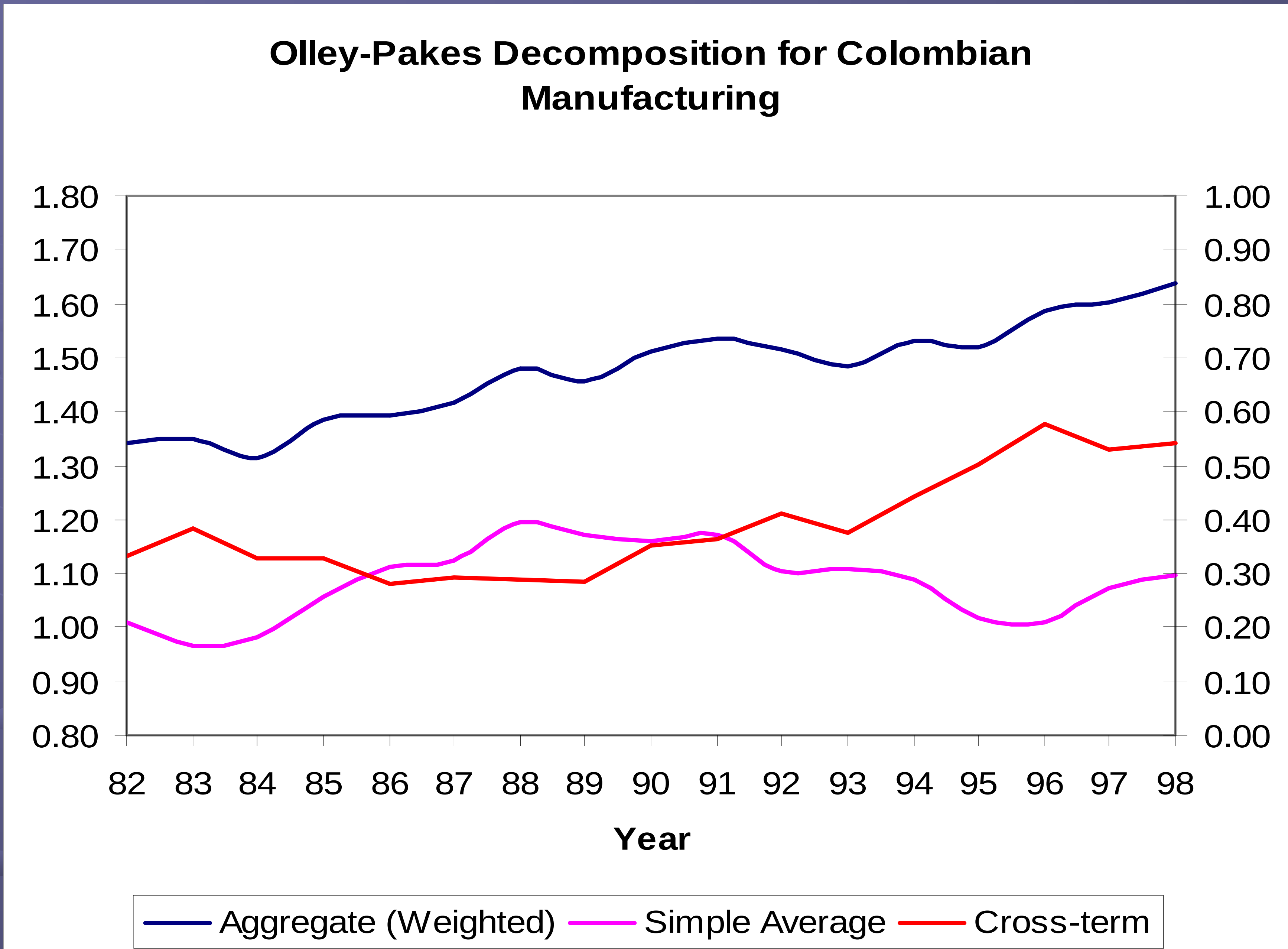
Olley and Pakes (1996) results for Telecommunications equipment

TABLE XI

DECOMPOSITION OF PRODUCTIVITY[a]
(EQUATION (16))

Year	p_t	$\bar{p}_t$	$\Sigma_i \Delta s_{it} \Delta p_{it}$	$\rho(p_t, k_t)$
1974	1.00	0.90	0.01	−0.07
1975	0.72	0.66	0.06	−0.11
1976	0.77	0.69	0.07	−0.12
1977	0.75	0.72	0.03	−0.09
1978	0.92	0.80	0.12	−0.05
1979	0.95	0.84	0.12	−0.05
1980	1.12	0.84	0.28	−0.02
1981	1.11	0.76	0.35	0.02
1982	1.08	0.77	0.31	−0.01
1983	0.84	0.76	0.08	−0.07
1984	0.90	0.83	0.07	−0.09
1985	0.99	0.72	0.26	0.02
1986	0.92	0.72	0.20	0.03
1987	0.97	0.66	0.32	0.10

[a] See text for details.



Source: Eslava et al. (2005)

Within Industry Dynamic Decomposition Applied to FHS (2008) data

		Components of Decomposition (GR)				
		Within	Between	Entry	Exit	Net Entry
Traditional	2.30	1.40	0.18	0.44	0.27	0.72
Revenue	5.13	4.03	0.16	0.55	0.39	0.94
Physical	5.13	3.82	-0.05	1.04	0.32	1.36

More Basic Measures of Productivity Are Often Used

- Labor productivity Measures at the Establishment (or Firm level)
  - Real Value Added Per Worker

$$RLP_{et} = (VA_{et} / TE_{et}) = (Y_{et} - M_{et}) / TE_{et}$$

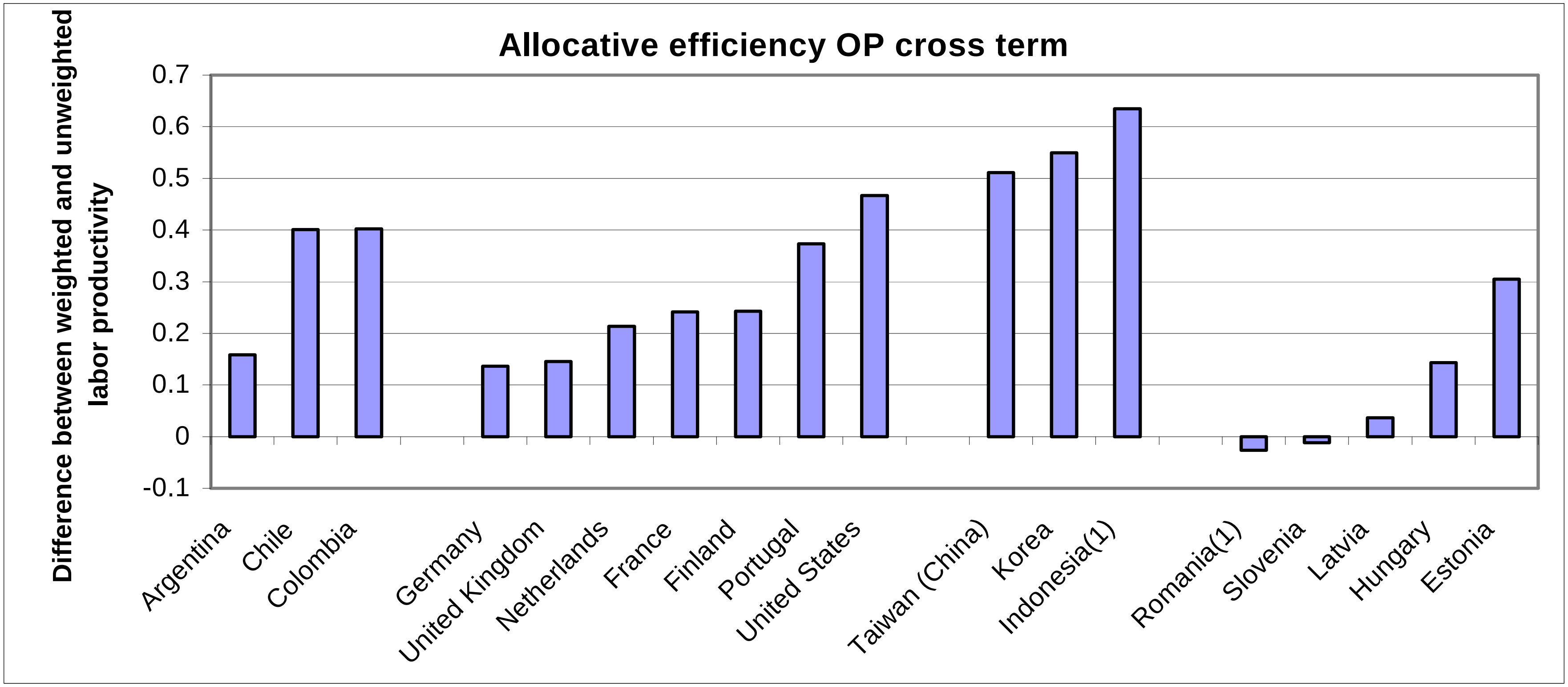
Where Y_{et} = Real Gross Output
M_{et}= Real Materials (including energy)
Te_{et}= Total Employment
Use detailed industry output and material price deflators

Often best available measure is real gross output per worker – comparable within industries

Data sources for Firm Level Dynamics Project (OECD and World Bank)

- **Business registers for firm demographics**
  - Firm level, at least one employee, 2/3-digit industry
- **Production Stats, enterprise surveys for productivity analysis**
- **Countries:**
  - 10 OECD
  - 5 Central and Eastern Europe
  - 6 Latin America
  - 3 East Asia
- **Data are disaggregated by:**
  - industry (2-3 digit);
  - size classes 1-9; 10-19; 20-49; 50-99; 100-249; 250-499; 500+ (for OECD sample the groups between 1 and 20 and the groups between 100 and 500 are combined)
  - Time (late 1980s – late 1990s)

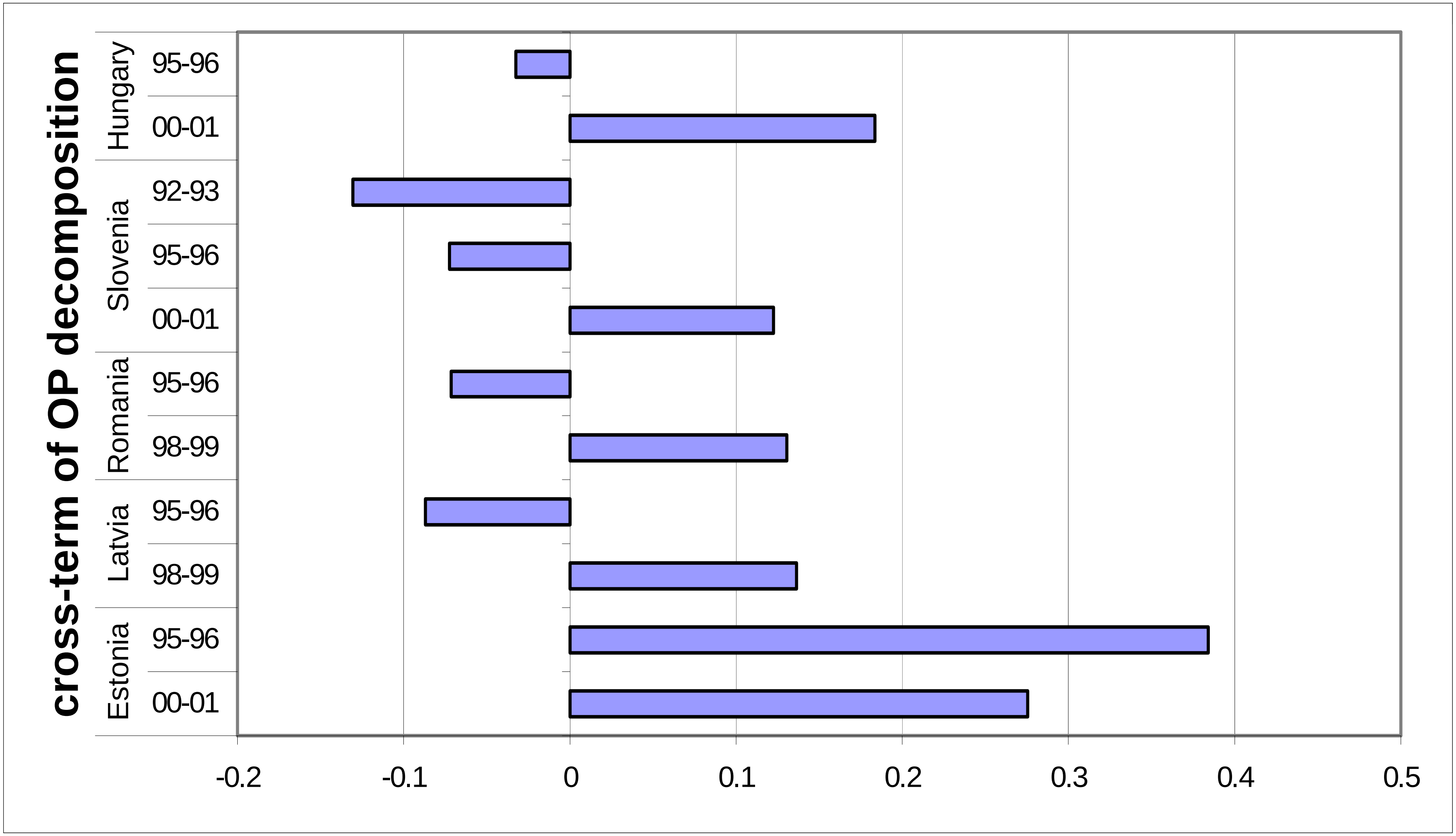
Allocative efficiency (Olley Pakes decomposition -- cross term)
(weighted averages of industry level cross terms from OP decomposition)



1. Based on the three-year differences

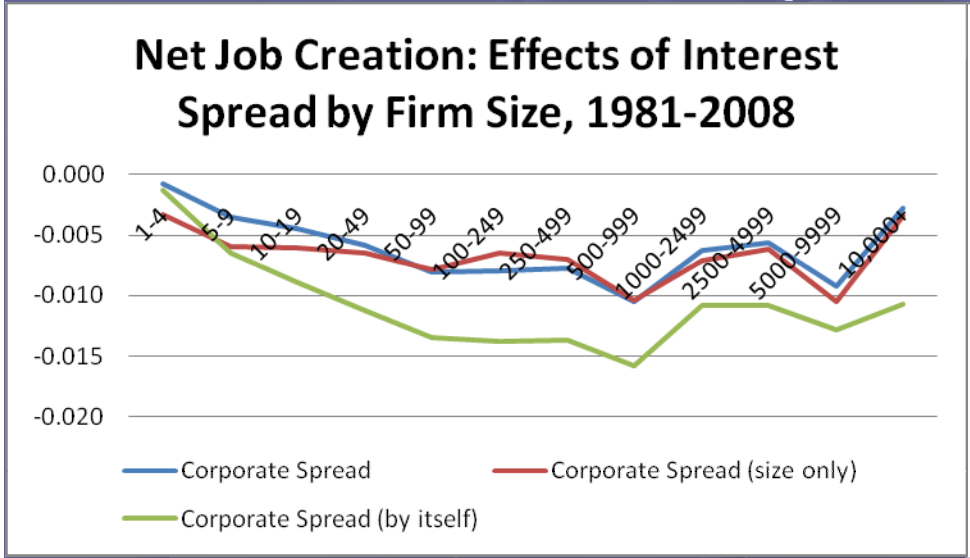
Source: Bartelsman, Haltiwanger and Scarpetta (2006)

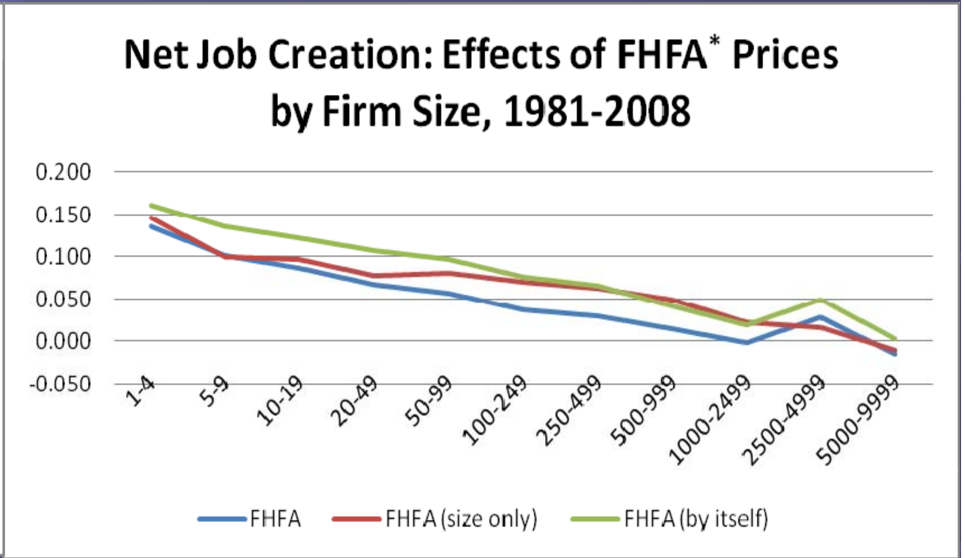
Evolution of allocative efficiency during the transition -- Eastern Europe, manufacturing
(weighted averages of industry level cross terms from OP decomposition)



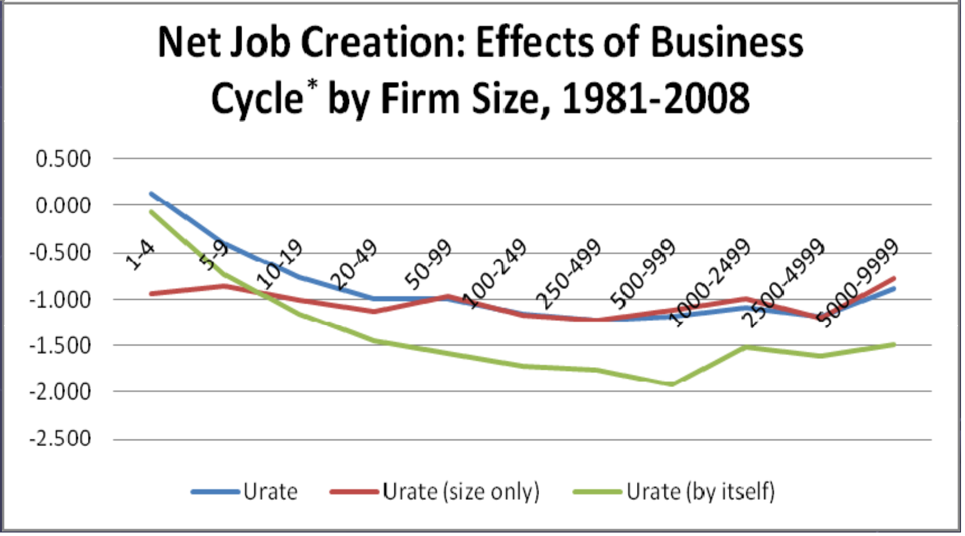
Source: Bartelsman, Haltiwanger and Scarpetta (2006)

LBD: The effect of business cycle dynamics and credit conditions on firms and job creation





Firm Size Effects:
Large firms are more sensitive to cycle…
Forms of financing differ for small and large firms…



Data

- Tracking U.S. Business Dynamics
  - The Longitudinal Business Database
    - 1975-2005 (08) – long time series permits analysis by firm age
    - Private Non Farm Economy
    - Establishment level with Firm identifiers
    - High quality establishment links to identify entry/exit
      - Need both firm and establishment level data to get dynamics right
    - Firm Size: constructed by aggregating employment up to firm
    - Firm Age: constructed from age of oldest establishment at time of firm birth
    - Other: Payroll, Industry, Location (Lat/Lon possible)
    - Can be integrated with data from Economic Censuses and Annual Surveys as well as external data (COMPUSTAT, Venture Capital, Private Equity)

Micro Productivity Data in U.S.

- Manufacturing:
  - Annual Survey of Manufactures and Census of Manufactures
    - Nominal revenue and expenditures
    - Can construct measures of real outputs and inputs
    - Five year panel rotation so longitudinal analysis possible (but requires careful treatment of data)
    - Selected products have physical quantities
- Retail Trade
  - Census of Retail Trade
    - Nominal revenue so a gross output per store measure feasible

New data on micro businesses

- ILBD:
  - Tracks all nonemployer and employer businesses including transitions
- LEHD:
  - Tracks all employer-employee matches in U.S.
  - Can be integrated with ILBD
  - Enables tracking of transitions between W&S, an owner of nonemployer business and owner of employer business

Availability of data

- Public domain tabulations available at:

http://www.ces.census.gov/index.php/bds/bds_home

- Census NSF/RDC access at:

http://www.ces.census.gov/index.php/ces/researchguidelines

- Sensitive data:
  - Must work in enclave (NBER, NYCRDC, Washington, D.C., Chicago Fed, Duke, UCLA, UC-Berkeley, Univ. of Michigan, Cornell, Stanford , Univ. of Minn., Atlanta, …)
  - Predominant purpose must benefit U.S. Census

Extra Slides

Growth Identities: Establishment

$$g_{it} = (E_{it} - E_{it-1}) / X_{it}$$

where

$$X_{it} = .5*(E_{it} + E_{it-1})$$

Then

$$JC_{it} = \max(g_{it}, 0)$$

$$JD_{it} = \max(-g_{it}, 0)$$

From Entry/Exit

$$JC_{it} = \max(g_{it}, 0) * I\{g_{it} = 2\}$$

$$JD_{it} = \max(-g_{it}, 0) * I\{-g_{it} = 2\}$$

Aggregate Measures (any level)

$$JC_t = \sum_i (X_{it} / X_t) \max\{g_{it}, 0\} \qquad JD_t = \sum_i (X_{it} / X_t) \max\{-g_{it}, 0\}$$

$$JC\_Entry_t = \sum_i (X_{it} / X_t) I\{g_{it} = 2\} \max(g_{it}, 0)$$

$$JD\_Exit_t = \sum_i (X_{it} / X_t) I\{g_{it} = -2\} \max(-g_{it}, 0)$$

$$g_t = JC_t - JD_t$$

$$JC_t = JC\_Cont_t + JC\_Entry_t$$

$$JD_t = JD\_Cont_t + JD\_Exit_t$$