Agglomeration Economies

Urban Economics

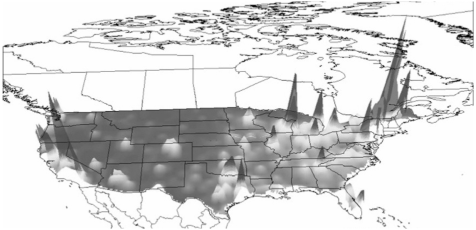
Ignacio Sarmiento-Barbieri

Universidad de los Andes

October 17, 2023

Spatial Distribution

- Why do we see such a remarkable clustering of human activity in a small number of urban areas?
- Cities exist because they are areas with high levels of productivity, which might occur because people come to places that are innately more productive or because density itself enhances productivity because of agglomeration economies



Spatial Distribution: TFP

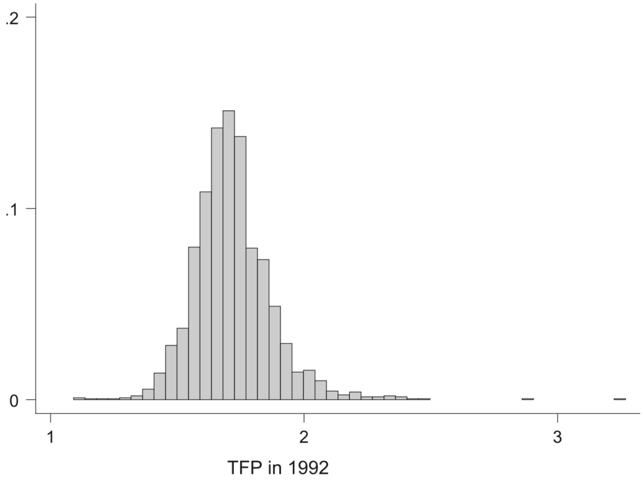


Figure 5 ***Distribution of total factor productivity in manufacturing establishments, by county.***

Introduction to a basic quantitative spatial model

- We begin with a twist to Rosen-Roback
  - We'll work through $n = 2$ case to develop intuition, but it can be easily extended to n locations
  - Can be used for other applications (trade, commuting, etc.)

Set up

- Assume wages, rents, amenities are exogenous
- Two cities A and B
- Person i's indirect utility of being in A:

$$V_A^i = w_A - r_A + A_A + \epsilon_A^i \tag{1}$$

- Person i's indirect utility of being in B:

$$V_B^i = w_B - r_B + A_B + \epsilon_B^i \tag{2}$$

$$\epsilon_A^i - \epsilon_B^i \sim U[-s, s] \tag{3}$$

Spatial Equilibrium with Agglomeration Economies

- Assume

$$X_c = x_c + \gamma N_c \tag{4}$$

- The MPL

$$w_c = x_c + (\gamma - (1-h))\, N_c + (1-h)\, K_c + ln(h) \tag{5}$$

Spatial Equilibrium with Agglomeration Economies

- Two periods
  - Period 1 both cities are identical
  - Period 2 productivity increases in B: $X_{B2} = X_{B1} + \Delta$ where $\Delta > 0$

Spatial Equilibrium with Agglomeration Economies

Change in nominal wages?

$$w_{B2} - w_{B1} = \frac{h\left(N\left(k_B + k_A\right) + 2s\right) - \gamma N}{h\left(N\left(k_B + k_A\right) + 2s\right) - 2\gamma N}\Delta \geq \Delta \geq 0 \tag{6}$$

$$\frac{\partial\left(w_{B2} - w_{B1}\right)}{\partial \gamma} = \frac{Nh\left(N\left(k_B + k_A\right) + 2s\right)}{\left(h\left(N\left(k_B + k_A\right) + 2s\right) - 2\gamma N\right)^2}\Delta \geq 0 \tag{7}$$

Spatial Equilibrium with Agglomeration Economies

Change in population?

$$(N_{B2} - N_{B1}) = \frac{Nh}{h\left(N\left(k_B + k_A\right) + 2s\right) - 2\gamma N}\Delta \geq 0 \tag{8}$$

Spatial Equilibrium with Agglomeration Economies

Change in housing markets?

$$r_{B2} - r_{B1} = \frac{hNk_B}{h\left(N\left(k_B + k_A\right) + 2s\right) - 2\gamma N}\Delta \geq 0 \tag{9}$$

Spatial Equilibrium with Agglomeration Economies

Change in Real wages?

$$(w_{B2} - w_{B1}) - (r_{B2} - r_{B1}) = \frac{(k_A N + 2s)\, h - \gamma N}{h\left(N\left(k_B + k_A\right) + 2s\right) - 2\gamma N}\Delta \tag{10}$$

$$\frac{\partial\left(w_{B2} - w_{B1}\right) - \left(r_{B2} - r_{B1}\right)}{\partial\gamma} = \frac{\left(Nh\left(k_A - k_B\right) + 2s\right)}{h\left(N\left(k_B + k_A\right) + 2s\right) - 2\gamma N}\Delta \tag{11}$$

Agglomeration and Empirics

- We have seen how agglomeration forces, can give rise to concentration of economic activity in certain locations (rosen roback) and in city centers(monocentric model).

Agglomeration and Empirics

- We have seen how agglomeration forces, can give rise to concentration of economic activity in certain locations (rosen roback) and in city centers(monocentric model).
- However, we need to see if these exist and how important they are.

Agglomeration and Empirics

- We have seen how agglomeration forces, can give rise to concentration of economic activity in certain locations (rosen roback) and in city centers(monocentric model).
- However, we need to see if these exist and how important they are.
- Let's review some empirical papers that estimate the extent of agglomeration forces.

Agglomeration and Empirics

- We would like to estimate how strong are these increasing returns arising from agglomeration.

Agglomeration and Empirics

- We would like to estimate how strong are these increasing returns arising from agglomeration.
- A good starting point to tackle these empirical issues is to think of the ideal experiment that would give us the estimate we want.

Agglomeration and Empirics

Ciccone and Hall (1996)

- Ciccone and Hall (1996) estimate the effects of density on productivity using data for U.S. states.
- Their starting point is a production function with increasing returns to scale at the county level:

$$q_c = a_c \left(\frac{n_c}{a_c}\right)^{\alpha} \left(\frac{q_c}{a_c}\right)^{\frac{\lambda-1}{\lambda}} \tag{12}$$

Agglomeration and Empirics

Ciccone and Hall (1996)

- Ciccone and Hall (1996) estimate the effects of density on productivity using data for U.S. states.
- Their starting point is a production function with increasing returns to scale at the county level:

$$q_c = a_c \left(\frac{n_c}{a_c}\right)^{\alpha} \left(\frac{q_c}{a_c}\right)^{\frac{\lambda-1}{\lambda}} \tag{12}$$

- Solving for q_c

$$q_c = \left(\frac{n_c}{a_c}\right)^{\gamma} \tag{13}$$

Agglomeration and Empirics

Ciccone and Hall (1996)

- Adding up over counties to the state level

$$\frac{Q_s}{N_s} = \frac{\sum_{c \in C_s} n_c^{\gamma} a_c^{-(\gamma-1)}}{N_s} \tag{14}$$

Agglomeration and Empirics

Ciccone and Hall (1996)

- Ciccone and Hall extend this basic model to account for physical and human capital, but the basic intuition remains the same. The estimating equation is the following:

$$log\left(\frac{Q_s}{N_s}\right) = log\phi + logD_s(\theta,\eta) + u_s \tag{15}$$

- with

$$D_s(\theta,\eta) = \frac{\sum_{c\in C_s}(n_c h_c^{\eta})^{\theta} a_c^{1-\theta}}{N_s} \tag{16}$$

Agglomeration and Empirics

Ciccone and Hall (1996)

- **Estimation Issues**
  - Non-linearity
  - Endogenity

Agglomeration and Empirics

Aside: IV

- Consider we want to estimate

$$y = X\beta + u \tag{17}$$

Agglomeration and Empirics

Aside: IV

- Two key assumptions for IV
  - Relevance
  - Validity o Exclusion

Agglomeration and Empirics

Aside: Non Linear Estimation

- Write the outcome as a non-linear function of the parameters:

$$y = f(\beta) + u \tag{18}$$

- With an instrumental variable in hand, the orthogonality assumption continues to be that Z and u are not correlated.
- Consider minimizing the following function in β:

$$(y - f(\beta))' Z\Omega Z' (y - f(\beta)) \tag{19}$$

Agglomeration and Empirics

Estimation and Results

- The endogenous variable here is D_s,

$$D_s(\theta,\eta) = \frac{\sum_{c\in C_s}(n_c h_c^{\eta})^{\theta} a_c^{1-\theta}}{N_s} \tag{20}$$

- for which Ciconne and Hall use 4 instruments:
  1. Presence of absence of a railroad in the state in 1860
  2. Population in 1950
  3. Population density in 1880
  4. Distance from the sea

Agglomeration and Empirics

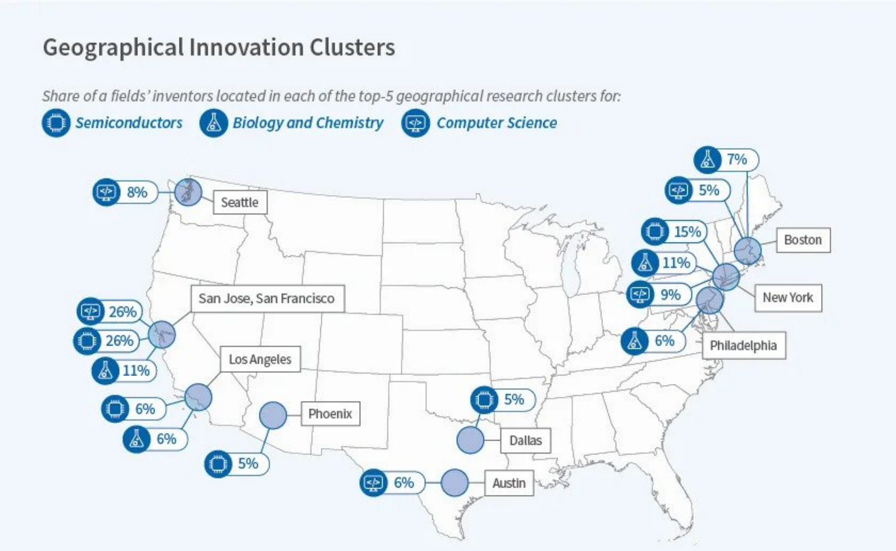
Ciccone and Hall (1996)

TABLE 1—ESTIMATION RESULTS

Instrument	Density elasticity, θ (standard error)	Education elasticity, η (standard error)	R^2
None (NLLS)	1.052	0.410	0.551
	(0.008)	(0.396)	
Eastern seaboard	1.055	0.460	0.548
	(0.017)	(0.51)	
Railroad in 1860	1.061	0.330	0.537
	(0.011)	(0.450)	
Population in 1850	1.060	0.350	0.539
	(0.015)	(0.510)	
Population density in 1880	1.051	0.530	0.549
	(0.019)	(0.550)	
All	1.06	0.060	0.536
	(0.01)	(0.82)	

Notes: The equation estimated is (24). The data are value added for 46 states and Washington DC. For the 46 states we have used data on employment and average years of education at the county level.

Agglomeration and Empirics
Moretti (2021)



Agglomeration and Empirics

Moretti (2021)

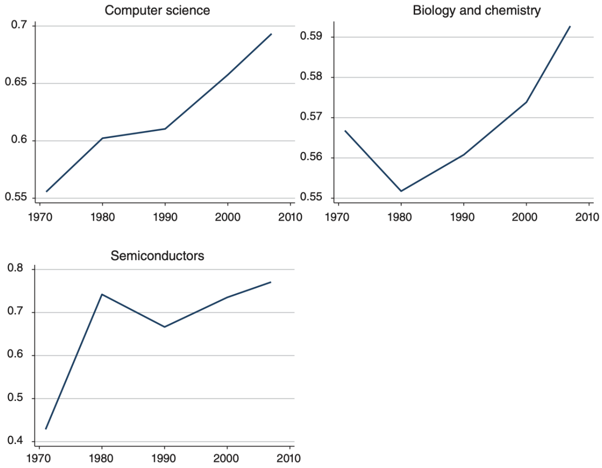
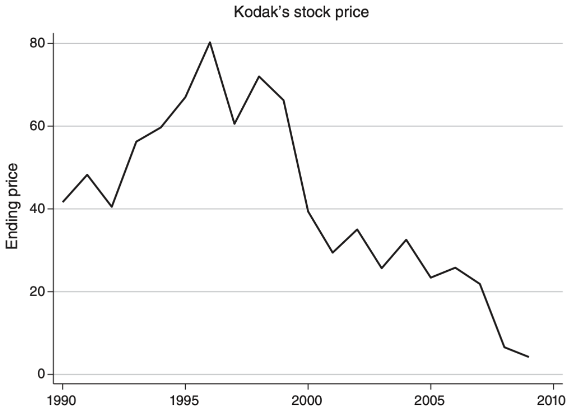


FIGURE 1. SHARE OF TOP TEN CITIES OVER TIME

Agglomeration and Empirics

Moretti (2021)



Agglomeration and Empirics

Moretti (2021)

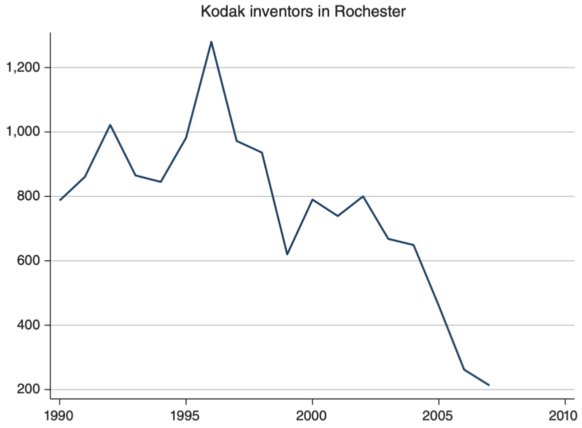


FIGURE 2. KODAK'S DECLINE

Agglomeration and Empirics

Moretti (2021)

FIGURE 3. AVERAGE INVENTOR PRODUCTIVITY IN ROCHESTER OUTSIDE KODAK

Note: Controls include research field dummies.

Agglomeration and Empirics
Moretti (2021)

TABLE 2—DIFFERENCE-IN-DIFFERENCE ESTIMATES: 1996–2007 PRODUCTIVITY CHANGE OF NON-KODAK INVENTORS IN ROCHESTER COMPARED TO OTHER CITIES

	(1)	(2)	(3)	(4)	Weighted (5)
Panel A					
Rochester × 2007	−0.0641	−0.0673	−0.0805	−0.0916	−0.0947
	(0.00757)	(0.00674)	(0.00631)	(0.00665)	(0.00860)
Rochester	−0.0148	−0.0364	−0.0317		
	(0.0105)	(0.0101)	(0.00987)		
2007	−0.190	−0.189			
	(0.00757)	(0.00713)			
Observations	194,120	194,120	194,120	194,120	193,331
Field		Yes	Yes	Yes	Yes
Field × year			Yes	Yes	Yes
Field × city				Yes	Yes

Agglomeration and Empirics

Moretti (2021)

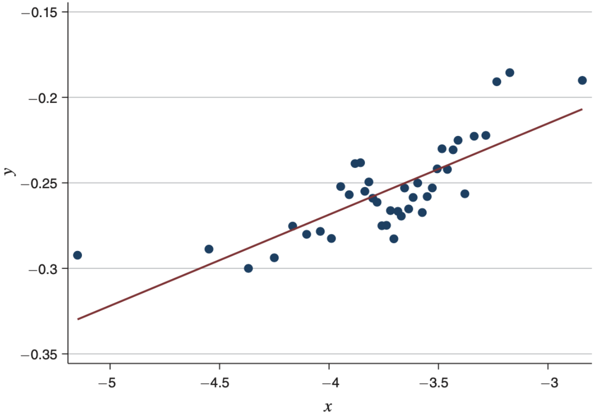


FIGURE 4. AVERAGE LOG NUMBER OF PATENTS PER INVENTOR PER YEAR AND LOG CLUSTER SIZE: ALL YEARS AND FIELDS

Agglomeration and Empirics

Moretti (2021)

TABLE 3—EFFECT OF CLUSTER SIZE ON INVENTOR PRODUCTIVITY: BASELINE MODELS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log size	0.0518							
	(0.00815)							
Observations	932,059							
Year	Yes							
City	Yes							
Field	Yes							
Class	Yes							
City × field								
City × class								
Field × year								
Class × year								
Inventor								
City × year								
Firm								

Notes: Each column is a separate regression. The level of observation in the regressions is inventor-year. The dependent variable is log of number of patents filed in a year. The model estimated is equation (1). Standard errors are clustered by city × research field.

Agglomeration and Empirics

Moretti (2021)

TABLE 3—EFFECT OF CLUSTER SIZE ON INVENTOR PRODUCTIVITY: BASELINE MODELS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log size	0.0518	0.0762	0.0881	0.0907				
	(0.00815)	(0.0167)	(0.0187)	(0.00926)				
Observations	932,059	932,059	932,059	932,059				
Year	Yes	Yes	Yes	Yes				
City	Yes	Yes	Yes	Yes				
Field	Yes	Yes	Yes	Yes				
Class	Yes	Yes	Yes	Yes				
City × field		Yes	Yes	Yes				
City × class			Yes	Yes				
Field × year				Yes				
Class × year								
Inventor								
City × year								
Firm								

Notes: Each column is a separate regression. The level of observation in the regressions is inventor-year. The dependent variable is log of number of patents filed in a year. The model estimated is equation (1). Standard errors are clustered by city × research field.

Agglomeration and Empirics

Moretti (2021)

Table 3—Effect of Cluster Size on Inventor Productivity: Baseline Models

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log size	0.0518	0.0762	0.0881	0.0907	0.0677			
	(0.00815)	(0.0167)	(0.0187)	(0.00926)	(0.00862)			
Observations	932,059	932,059	932,059	932,059	932,059			
Year	Yes	Yes	Yes	Yes	Yes			
City	Yes	Yes	Yes	Yes	Yes			
Field	Yes	Yes	Yes	Yes	Yes			
Class	Yes	Yes	Yes	Yes	Yes			
City × field		Yes	Yes	Yes	Yes			
City × class			Yes	Yes	Yes			
Field × year				Yes	Yes			
Class × year					Yes			
Inventor								
City × year								
Firm								

Notes: Each column is a separate regression. The level of observation in the regressions is inventor-year. The dependent variable is log of number of patents filed in a year. The model estimated is equation (1). Standard errors are clustered by city × research field.

Agglomeration and Empirics

Moretti (2021)

TABLE 3—EFFECT OF CLUSTER SIZE ON INVENTOR PRODUCTIVITY: BASELINE MODELS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log size	0.0518	0.0762	0.0881	0.0907	0.0677	0.0923		
	(0.00815)	(0.0167)	(0.0187)	(0.00926)	(0.00862)	(0.00990)		
Observations	932,059	932,059	932,059	932,059	932,059	932,059		
Year	Yes	Yes	Yes	Yes	Yes	Yes		
City	Yes	Yes	Yes	Yes	Yes	Yes		
Field	Yes	Yes	Yes	Yes	Yes	Yes		
Class	Yes	Yes	Yes	Yes	Yes	Yes		
City × field		Yes	Yes	Yes	Yes	Yes		
City × class			Yes	Yes	Yes	Yes		
Field × year				Yes	Yes	Yes		
Class × year					Yes	Yes		
Inventor						Yes		
City × year								
Firm								

Notes: Each column is a separate regression. The level of observation in the regressions is inventor-year. The dependent variable is log of number of patents filed in a year. The model estimated is equation (1). Standard errors are clustered by city × research field.

Agglomeration and Empirics

Moretti (2021)

TABLE 3—EFFECT OF CLUSTER SIZE ON INVENTOR PRODUCTIVITY: BASELINE MODELS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log size	0.0518	0.0762	0.0881	0.0907	0.0677	0.0923	0.0545	
	(0.00815)	(0.0167)	(0.0187)	(0.00926)	(0.00862)	(0.00990)	(0.0116)	
Observations	932,059	932,059	932,059	932,059	932,059	932,059	932,059	
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
City	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Field	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Class	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
City × field		Yes	Yes	Yes	Yes	Yes	Yes	
City × class			Yes	Yes	Yes	Yes	Yes	
Field × year				Yes	Yes	Yes	Yes	
Class × year					Yes	Yes	Yes	
Inventor						Yes	Yes	
City × year							Yes	
Firm								

Notes: Each column is a separate regression. The level of observation in the regressions is inventor-year. The dependent variable is log of number of patents filed in a year. The model estimated is equation (1). Standard errors are clustered by city × research field.

Agglomeration and Empirics

Moretti (2021)

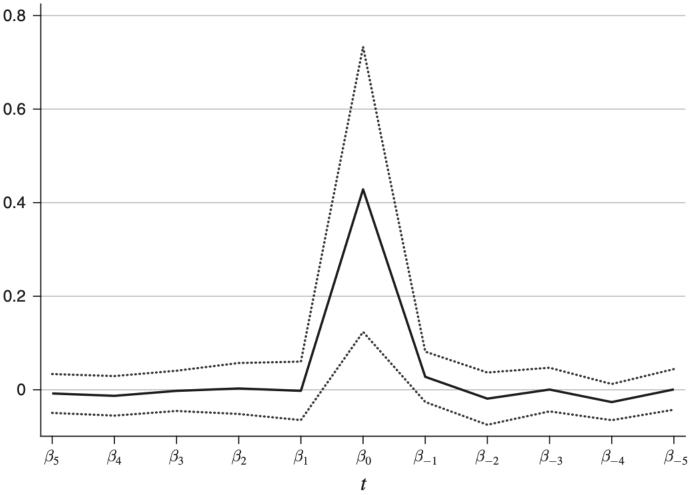
TABLE 3—EFFECT OF CLUSTER SIZE ON INVENTOR PRODUCTIVITY: BASELINE MODELS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log size	0.0518	0.0762	0.0881	0.0907	0.0677	0.0923	0.0545	0.0676
	(0.00815)	(0.0167)	(0.0187)	(0.00926)	(0.00862)	(0.00990)	(0.0116)	(0.0139)
Observations	932,059	932,059	932,059	932,059	932,059	932,059	932,059	823,375
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Field	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Class	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City × field		Yes	Yes	Yes	Yes	Yes	Yes	Yes
City × class			Yes	Yes	Yes	Yes	Yes	Yes
Field × year				Yes	Yes	Yes	Yes	Yes
Class × year					Yes	Yes	Yes	Yes
Inventor						Yes	Yes	Yes
City × year							Yes	Yes
Firm								Yes

Notes: Each column is a separate regression. The level of observation in the regressions is inventor-year. The dependent variable is log of number of patents filed in a year. The model estimated is equation (1). Standard errors are clustered by city × research field.

Agglomeration and Empirics

Moretti (2021)



Agglomeration and Empirics

Moretti (2021)

Table 5—Models in Differences: Effect of Changes in Cluster Size on Changes in Inventor Productivity: OLS and IV Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. OLS						
Δlog size	0.0141	0.0145	0.0153	0.0164	0.0162	0.0159
	(0.00394)	(0.00392)	(0.00376)	(0.00397)	(0.00392)	(0.00385)
Panel B. 2SLS						
Δlog size	0.0422	0.0630	0.0502	0.0496	0.0502	0.0491
	(0.0186)	(0.0211)	(0.0189)	(0.0131)	(0.0137)	(0.0144)
First stage	1.109	1.076	1.096	1.431	1.475	1.488
	(0.151)	(0.170)	(0.167)	(0.214)	(0.189)	(0.185)
F-statistic	53.8	40.2	43.0	44.5	60.8	64.2
Observations	419,596	419,596	419,565	405,111	405,111	403,955
Year	Yes	Yes	Yes	Yes	Yes	Yes
Field		Yes	Yes	Yes	Yes	Yes
Class			Yes	Yes	Yes	Yes
Firm				Yes	Yes	Yes
Field × year					Yes	Yes
Class × year						Yes