Agglomeration, Misallocation, and (the Lack of) Competition Online Appendix

Wyatt J. Brooks, Joseph P. Kaboski and Yao Amber Li

A Online Appendix

A.1 Robustness Results

Here we present the following empirical results:

- Table A1: Placebo Test Using Affiliate Sample
- Table A2: Placebo Test Using SOE Sample
- Table A3: Rauch Product Classification Results
- Table A4: Other Dependent Variables
- Table A5: Instrumental Variable Estimation Results Using Low CV Deciles
- Table A6: Baseline Results Using Overall Sample without Firm Fixed Effects
- Table A7: Baseline Results Using Sales-to-cost-ratio as Alternative Measure of Markups

Table A.1: Appendix Table–Placebo Test Using Affiliate Sample

	Dependent Variable: $\frac{1}{\mu_{nit}}$									
	(1)	(2) Province	(3) City	(4) County	(5) Province	(6) City	(7) County	(8) Province	(9) City	(10) County
Firm's share	-0.009				-0.029	-0.016	-0.039	-0.067	-0.048	-0.065
	(0.060)				(0.063)	(0.067)	(0.080)	(0.065)	(0.070)	(0.082)
Region's share		0.019	0.004	0.013	0.021	0.007	0.030	0.028	0.009	0.026
		(0.021)	(0.030)	(0.040)	(0.022)	(0.033)	(0.054)	(0.023)	(0.035)	(0.056)
SEZ*Firm's share								0.094	0.100	0.100
								(0.137)	(0.137)	(0.137)
SEZ*Region's share								0.019	0.023	0.023
								(0.039)	(0.039)	(0.039)
SEZ Dummy								0.004	0.004	0.004
								(0.005)	(0.005)	(0.005)
Year FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Firm FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	24771	24771	24771	24771	24771	24771	24771	20679	20679	20679
Overall R^2	0.010	0.012	0.010	0.010	0.012	0.010	0.010	0.010	0.009	0.009

Notes: Robust standard errors clustered at firm level in parentheses. Significance: ***: 1%, **: 5%, *: 10%. Regions are defined at various aggregation levels, including province (in specifications 2, 5, and 8), city (in specifications 3, 6, and 9), and county (in specifications 4, 7, and 10). All specifications are regressions weighted by the number of observations for each two-digit CIC sector production function estimation reported (following De Loecker et al. 2014). All regressions include a constant term.

Table A.2: Appendix Table–Placebo Test Using SOE Sample

	Dependent Variable: $\frac{1}{\mu_{nit}}$									
	(1)	(2) Province	(3) City	(4) County	(5) Province	(6) City	(7) County	(8) Province	(9) City	(10) County
Firm's share	-0.047				-0.029	-0.066	-0.061	-0.078	-0.101*	-0.147*
	(0.055)				(0.056)	(0.058)	(0.070)	(0.054)	(0.060)	(0.076)
Region's share		-0.024*	0.005	-0.020	-0.021	0.020	0.013	-0.010	0.016	0.060
		(0.013)	(0.023)	(0.037)	(0.014)	(0.025)	(0.048)	(0.016)	(0.028)	(0.055)
SEZ*Firm's share								0.052	0.050	0.050
								(0.120)	(0.120)	(0.119)
SEZ*Region's share								-0.014	-0.016	-0.016
								(0.039)	(0.039)	(0.039)
SEZ dummy								-0.000	-0.000	-0.000
								(0.004)	(0.004)	(0.004)
Year FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Firm FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	106434	106434	106434	106434	106434	106434	106434	69608	69608	69608
Overall R^2	0.034	0.030	0.032	0.033	0.031	0.033	0.034	0.034	0.035	0.035

Notes: Robust standard errors clustered at firm level in parentheses. Significance: ***: 1%, **: 5%, *: 10%. Regions are defined at various aggregation levels, including province (in specifications 2, 5, and 8), city (in specifications 3, 6, and 9), and county (in specifications 4, 7, and 10). All specifications are regressions weighted by the number of observations for each two-digit CIC sector production function estimation reported (following De Loecker et al. 2014). All regressions include a constant term.

Table A.3: Appendix Table–Rauch Product Classification Results

	Dependent Variable: $\frac{1}{\mu_{nit}}$					
	(1) homo/ref	(2) diff.	(3) overall	(4) homo/ref	(5) diff.	(6) overall
Firm's Share	-0.167**	-0.074***	-0.127**	-0.141	-0.046*	-0.179***
	(0.080)	(0.023)	(0.051)	(0.202)	(0.024)	(0.040)
Region's Share	-0.076***	-0.026***	-0.067***	-0.296***	-0.016*	-0.070***
	(0.011)	(0.008)	(0.010)	(0.089)	(0.009)	(0.009)
Differentiated X Firm's Share			0.042			0.126***
			(0.056)			(0.044)
Differentiated X Region's Share			0.045***			0.056***
			(0.013)			(0.012)
Differentiated Dummy			-0.002			-0.001
			(0.001)			(0.001)
Year FEs	YES	YES	YES	YES	YES	YES
Firm FEs	YES	YES	YES	YES	YES	YES
Observations	273327	935702	1279149	75692	642132	1279149
Overall R^2	0.036	0.024	0.027	0.015	0.019	0.027

Notes: Robust standard errors clustered at firm level in parentheses. Significance: ***: 1%, **: 5%, *: 10%. Specifications 1-3 refer to product classification using "most frequent" principle; specifications 4-6 refer to product classification using "pure" principle. All specifications are regressions weighted by the number of observations for each two-digit CIC sector production function estimation reported (following De Loecker et al. 2014). All regressions include a constant term.

Table A.4: Appendix Table–Robustness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Dependent Variable = $1/\mu_{nit}$								
Firm's Share	-0.143***		-0.099***	-0.163***	-0.083***		-0.046**	-0.089***
	(0.021)		(0.022)	(0.027)	(0.017)		(0.018)	(0.023)
Region's Share		-0.053***	-0.040***	-0.031***		-0.044***	-0.036***	-0.024***
		(0.006)	(0.006)	(0.007)		(0.006)	(0.006)	(0.007)
SEZ*Firm's Share				0.095**				0.073**
				(0.039)				(0.035)
SEZ*Region's Share				-0.018*				-0.021**
				(0.010)				(0.010)
Observations	1346860	1346860	1346860	1105162	1346860	1346860	1346860	1105162
Overall R^2	0.028	0.027	0.028	0.026	0.028	0.027	0.027	0.026
Panel B: Dependent Variable = $\mu_{nit}/(\mu_{nit}-1)$ *(full sample)*								
Firm's Share	225.521		195.421	388.275	469.783		439.703	800.015
	(200.154)		(204.416)	(318.785)	(474.848)		(480.277)	(772.947)
Region's Share		53.610	27.789	18.889		109.975	28.807	15.500
		(33.033)	(25.908)	(31.765)		(88.601)	(33.949)	(41.299)
SEZ*Firm's Share				-369.619				-694.560
				(251.815)				(525.775)
SEZ*Region's Share				28.451				38.929
				(26.538)				(35.645)
Observations	1346860	1346860	1346860	1105162	1346860	1346860	1346860	1105162
Overall R^2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Panel C: Dependent Variable = $\mu_{nit}/(\mu_{nit}-1)$ *(drop* $\mu_{nit} < 1.06$*)*								
Firm's Share	-3.229***		-1.985***	-3.783***	-2.471***		-1.311***	-2.488***
	(0.561)		(0.594)	(0.745)	(0.464)		(0.500)	(0.692)
Region's Share		-1.404***	-1.140***	-1.051***		-1.339***	-1.103***	-0.932***
		(0.170)	(0.183)	(0.220)		(0.161)	(0.175)	(0.208)
SEZ*Firm's Share				3.248***				2.545***
				(0.964)				(0.905)
SEZ*Region's Share				-0.272				-0.372
				(0.259)				(0.251)
Observations	1228255	1228255	1228255	1006748	1228255	1228255	1228255	1006748
Overall R^2	0.010	0.009	0.009	0.009	0.010	0.009	0.009	0.009
Panel D: Dependent Variable = $log(mu)_{nit}$								
Firm's Share	0.174***		0.118***	0.193***	0.099***		0.051**	0.102***
	(0.027)		(0.028)	(0.035)	(0.022)		(0.024)	(0.030)
Region's Share		0.067***	0.052***	0.035***		0.055***	0.046***	0.028***
		(0.008)	(0.008)	(0.009)		(0.007)	(0.008)	(0.009)
SEZ*Firm's Share				-0.103*				-0.085*
				(0.053)				(0.047)
SEZ*Region's Share				0.034**				0.035***
				(0.014)				(0.013)
Observations	1346860	1346860	1346860	1105162	1346860	1346860	1346860	1105162
Overall R^2	0.028	0.027	0.028	0.027	0.028	0.027	0.027	0.026
All Panels								
Year FEs	YES	YES	YES	YES	YES	YES	YES	YES

Table A.5: Appendix Table–Instrumental Variable Estimation Results Using Low CV Deciles

	Dependent Variable: $\frac{1}{\mu_{nit}}$								
	(1) Province	(2) City	(3) County	(4) Province	(5) City	(6) County	(7) Province	(8) City	(9) County
Firm's share	0.007	-0.383***	-0.224***				0.100	-0.229*	-0.113***
	(0.130)	(0.114)	(0.052)				(0.178)	(0.121)	(0.037)
Region's share				0.010	-0.042***	-0.057***	0.018	-0.030***	-0.048***
				(0.016)	(0.011)	(0.016)	(0.022)	(0.012)	(0.017)
Year FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES
Firm FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	258745	154012	187050	258745	154012	187050	258745	154012	187050
Overall R^2	0.031	0.014	0.023	0.031	0.015	0.023	0.031	0.015	0.023
First-Stage Instruments:	(Sum of other firms' productivity; Sum of outside-cluster firms' productivity)								
Weak Instrument (Prob > F)	0.0000								

Notes: Robust standard errors clustered at firm level in parentheses. Significance: ***: 1%, **: 5%, *: 10%. Regions are defined at various aggregation levels, including province (in specifications 1, 4, and 7), city (in specifications 2, 5, and 8), and county (in specifications 3, 6 and 9). All specifications are regressions weighted by the number of observations for each two-digit CIC sector production function estimation reported (following De Loecker et al. 2014). All regressions include a constant term.

Table A.6: Baseline Results Using Overall Sample without Firm Fixed Effects

	Dependent Variable: $\frac{1}{\mu_{nit}}$									
	(1)	(2) Province	(3) City	(4) County	(5) Province	(6) City	(7) County	(8) Province	(9) City	(10) County
Firm's share in industry	-0.467***				-0.455***	-0.436***	-0.412***	-0.506***	-0.489***	-0.465***
	(0.033)				(0.033)	(0.033)	(0.033)	(0.036)	(0.036)	(0.038)
Region's share in industry		-0.018***	-0.050***	-0.100***	-0.012***	-0.030***	-0.053***	-0.009***	-0.025***	-0.048***
		(0.003)	(0.006)	(0.011)	(0.002)	(0.006)	(0.009)	(0.003)	(0.006)	(0.010)
SEZ*Firm's share in industry								0.171**	0.180***	0.181***
								(0.067)	(0.064)	(0.067)
SEZ*Region's share in industry								-0.001	-0.007	-0.009
								(0.005)	(0.010)	(0.015)
SEZ Dummy								-0.003***	-0.003***	-0.003***
								(0.001)	(0.001)	(0.001)
Year FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Industry FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Location FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	1346860	1346860	1346860	1346860	1346860	1346860	1346860	1105162	1105162	1105162
R^2	0.263	0.262	0.262	0.262	0.263	0.263	0.263	0.248	0.248	0.248

Notes: Robust standard errors clustered at firm level in parentheses. Significance: ***: 1%, **: 5%, *: 10%. Regions are defined at various aggregation levels, including province (in specifications 2, 5, and 8), city (in specifications 3, 6, and 9), and county (in specifications 4, 7, and 10). The last two rows report the average number of firms per region-industry and the average number of industries per region. All specifications are regressions weighted by the number of observations for each two-digit CIC sector production function estimation reported (following De Loecker et al. 2014). All regressions include a constant term, 4-digit industry fixed effects, and the county-level location fixed effects.

Table A.7: Baseline Results Using Sales-to-cost-ratio as Alternative Measure of Markups

	Dependent Variable: $\frac{1}{\mu_{nit}}$						
	(1)	(2) Province	(3) City	(4) County	(5) Province	(6) City	(7) County
Firm's share in industry	-0.443***				-0.427***	-0.410***	-0.348***
	(0.110)				(0.111)	(0.113)	(0.115)
Region's share in industry		-0.020*	-0.051**	-0.122***	-0.014	-0.027	-0.082***
		(0.011)	(0.021)	(0.030)	(0.011)	(0.022)	(0.032)
Year FEs	YES	YES	YES	YES	YES	YES	YES
Firm FEs	YES	YES	YES	YES	YES	YES	YES
Observations	675734	675734	675734	675734	675734	675734	675734
R^2	0.022	0.022	0.022	0.022	0.022	0.022	0.022

Notes: Robust standard errors clustered at firm level in parentheses. Significance: ***: 1%, **: 5%, *: 10%. All specifications are regressions weighted by the number of observations for each two-digit CIC sector production function estimation reported (following De Loecker et al. 2014). All regressions include a constant term.

A.2 Derivation of the Demand Function

Suppose the household solves the following problem:

$$\max_{\{Y_i\}} \left(\sum_i D_i^{1/\gamma} Y_i^{\frac{\gamma-1}{\gamma}} \right)^{\frac{\gamma}{\gamma-1}} \tag{1}$$

subject to:

$$\sum_i P_i Y_i \leq P$$

We take the budget of the household P to be exogenous. Cost minimization on the part of the representative household implies the demand function:

$$Y_i = D_i \left(\frac{P_i}{P} \right)^{-\gamma} \tag{2}$$

The final product in each industry is assembled by competitive firms in each industry that solves:

$$P_i Y_i = \min_{\{y_{ni}\}} \sum_{n \in \Omega_i} p_{ni} y_{ni} \tag{3}$$

subject to:

$$Y_i = \left(\sum_i y_{ni}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$$

Cost minimization on the part of these competitive firms implies:

$$y_{ni} = Y_i \left(\frac{p_{ni}}{P_i} \right)^{-\sigma} \tag{4}$$

Combining equations (2) and (4) implies:

$$y_{ni} = D_i \left(\frac{P_i}{P} \right)^{-\gamma} \left(\frac{p_{ni}}{P_i} \right)^{-\sigma} \tag{5}$$

A.3 Proof of Proposition 1

Suppose marginal costs of all firms are bounded and non-decreasing. Proposition 1 has the following five parts:

1. If operating independently, firm markups are increasing in a firm's own market share,
2. If operating as a cartel, cartel markups are increasing in total cartel market share with each firm's own market share playing no additional role,
3. Firm markups are higher under cartel decisions than when operating independently,
4. Firm markups are more similar when operating as a cartel than when operating independently,
5. Firm market shares are more similar when operating independently than when operating as a cartel

Proof Suppose any firm n in industry i weights the profits of the set of firms $S \subset \Omega_i$ with constant $\kappa \in [0,1]$. Then their objective is:

$$\max_{y_{ni}} p(y_{ni})y_{ni} - C(y_{ni}; X_{ni}) + \kappa \sum_{m \in S} \left[p(y_{mi})y_{mi} - C(y_{mi}; X_{mi}) \right] \tag{6}$$

Then for μ_{ni} defined as price divided by marginal cost and share defined as the firm's revenue divided by the sum of firm revenues in the industry, the firm's first order condition can be rewritten as:

$$\frac{1}{\mu_{ni}} = 1 + (1-\kappa)\frac{\partial \log(p_{ni})}{\partial \log(y_{ni})} + \kappa \sum_{m \in S} \frac{s_{mi}}{s_{ni}} \frac{\partial \log(p_{mi})}{\partial \log(y_{ni})} \tag{7}$$

If inverse demand is given by:

$$p_{ni} = D_i y_{ni}^{-1/\sigma} \left(\sum_{m \in \Omega_i} y_{mi}^{1-1/\sigma} \right)^{\frac{\sigma}{\gamma}\frac{\gamma-1}{\sigma-1}-1} \tag{8}$$

Then the cross-price elasticities are:

$$\frac{\partial \log(p_{mi})}{\partial \log(y_{ni})} = \left(\frac{1}{\sigma} - \frac{1}{\gamma} \right) s_{ni} \tag{9}$$

The own-price elasticity is:

$$\frac{\partial \log(p_{ni})}{\partial \log(y_{ni})} = -\frac{1}{\sigma} + \left(\frac{1}{\sigma} - \frac{1}{\gamma}\right) s_{ni} \tag{10}$$

Together these imply that:

$$\frac{1}{\mu_{ni}} = 1 - \frac{1}{\sigma} + \left(\frac{1}{\sigma} - \frac{1}{\gamma}\right)\left((1-\kappa)s_{ni} + \kappa \sum_{m\in S} s_{mi}\right) \tag{11}$$

Firms operating independently is the case where $\kappa = 0$, so then:

$$\frac{1}{\mu_{ni}} = 1 - \frac{1}{\sigma} + \left(\frac{1}{\sigma} - \frac{1}{\gamma}\right) s_{ni} \tag{12}$$

This implies result 1, when $\sigma > \gamma$. Likewise, if firms are operating as a perfect cartel, then $\kappa = 1$:

$$\frac{1}{\mu_{ni}} = 1 - \frac{1}{\sigma} + \left(\frac{1}{\sigma} - \frac{1}{\gamma}\right) \sum_{m\in S} s_{mi} \tag{13}$$

This immediately implies the second result. Moreover, equations (12) and (13) together imply the fourth result, as cartels have no variation in markups (even if they have variation in market shares) while independent firms have markups that vary with their shares.

To compare firms in a cartel to those operating independently, we construct an artificial single firm that is equivalent to the cartel. That is, suppose $\kappa = 1$ so that the cartel solves:

$$\max_{\{y_{mi}\}} \sum_{m\in S} \left(p_{mi} y_{mi} - C(y_{mi}; X_{mi})\right) \tag{14}$$

where p_{mi} is given by (8). Now define a cartel aggregate of production:

$$Y = \left(\sum_{m\in S} y_{mi}^{1-1/\sigma}\right)^{\frac{\sigma}{\sigma-1}} \tag{15}$$

Let $\tilde{C}(Y)$ be the cost function of the cartel defined as:

$$\tilde{C}(Y) = \min_{\{y_{mi}\}} \sum_{m\in S} C(y_{mi}; X_{mi}) \tag{16}$$

$$\text{subject to:} \quad Y = \left(\sum_{m \in S} y_{mi}^{1-1/\sigma} \right)^{\frac{\sigma}{\sigma-1}}$$

Then the following problem is equivalent to (14):

$$\max_{Y} D_i Y^{1-1/\sigma} \left(Y^{1-1/\sigma} + \sum_{n \notin S} y_{ni}^{1-1/\sigma} \right)^{\frac{\sigma}{\gamma}\frac{\gamma-1}{\sigma-1}-1} - \tilde{C}(Y) \tag{17}$$

First notice that the Envelope Theorem applied to the problem in (16):

$$\forall m \in S, \qquad \tilde{C}'(Y) = \lambda = \frac{C'(y_{mi}; X_{mi})}{y_{mi}^{-1/\sigma} Y^{1/\sigma}} \tag{18}$$

Then we can relate the size of the cartel to the cost of the cartel's production.

Lemma 1 *Consider a cartel made up of in $T \subset S$. Then for every level of production Y, the marginal cost in the cartel composed of T is strictly higher than in the cartel composed of S.*

To prove this lemma, suppose y_{mi}^T is how much firm m produces when part of the cartel composed of T and y_{mi}^S is how much the same firm produces when part of the cartel composed of S. Then for any given Y it must be the case that:

$$y_{mi}^S < y_{mi}^T \implies \frac{C'(y_{mi}^S; X_{mi})}{{y_{mi}^S}^{-1/\sigma} Y^{1/\sigma}} < \frac{C'(y_{mi}^T; X_{mi})}{{y_{mi}^T}^{-1/\sigma} Y^{1/\sigma}} \implies \tilde{C^S}'(Y) < \tilde{C^T}'(Y)$$

where the second implication follows from the fact that all firms have non-decreasing marginal costs. The first inequality follows from bounded marginal costs and Inada conditions in the aggregation of individual firm production to cartel-level production. Therefore, if more firms are added to a cartel, marginal costs for the cartel are reduced for every level of output.

Given this lemma, notice that as a cartel grows, the markup that the cartel charges strictly increases. This follows immediately from that fact that, given the lemma, marginal costs decline so cartel production increases, and as another firm from within the same industry is brought into the cartel, that firm's production is no longer counted in the denominator when computing the cartel's market share. Therefore, the cartel's market share strictly increases as more firms are added. Hence, by (13), the markup charged by the cartel increases.

A special case of this result is part 3 of Proposition 1. If a firm is operating outside of an existing cartel then is brought into it, the new cartel would have strictly higher markups than either the original cartel or the formerly independent firm.

To demonstrate the last result, consider any two firms n and m within the same cartel. Manipulating (18) gives:

$$\frac{C'(y_{mi};X_{mi})}{C'(y_{ni};X_{ni})} = \left(\frac{y_{mi}}{y_{ni}}\right)^{-\frac{1}{\sigma}} = \left(\frac{s_{mi}}{s_{ni}}\right)^{\frac{1}{1-\sigma}} \tag{19}$$

Then consider two other firms v and w that are operating independently. Then the relationship between marginal cost and market share is:

$$\frac{C'(y_{vi};X_{vi})}{C'(y_{wi};X_{wi})} = \left(\frac{s_{vi}}{s_{wi}}\right)^{\frac{1}{1-\sigma}} \frac{1-1/\sigma+(1/\sigma-1/\gamma)s_{vi}}{1-1/\sigma+(1/\sigma-1/\gamma)s_{wi}} \tag{20}$$

Suppose these two pairs of firms have the same relative marginal costs. Then:

$$\frac{C'(y_{mi};X_{mi})}{C'(y_{ni};X_{ni})} = \frac{C'(y_{vi};X_{vi})}{C'(y_{wi};X_{wi})} \implies \tag{21}$$

$$\left(\frac{s_{mi}}{s_{ni}}\right)^{\frac{1}{1-\sigma}} = \left(\frac{s_{vi}}{s_{wi}}\right)^{\frac{1}{1-\sigma}} \frac{1-1/\sigma+(1/\sigma-1/\gamma)s_{vi}}{1-1/\sigma+(1/\sigma-1/\gamma)s_{wi}}$$

Without loss, if firms v and m have relatively high costs, then:

$$\frac{C'(y_{mi};X_{mi})}{C'(y_{ni};X_{ni})} = \frac{C'(y_{vi};X_{vi})}{C'(y_{wi};X_{wi})} > 1 \implies \tag{22}$$

$$\frac{1-1/\sigma+(1/\sigma-1/\gamma)s_{vi}}{1-1/\sigma+(1/\sigma-1/\gamma)s_{wi}} > 1 \implies \frac{s_{ni}}{s_{mi}} > \frac{s_{wi}}{s_{vi}}$$

Therefore, independently operating firms have wider variation in market shares conditional on marginal cost than do firms operating as a cartel. This completes the proof.

A.4 Simulation of Model with Shocks to Demand and Costs

We now consider a version of the model where some uncertainty in costs or demand is realized after production choices are made. Firm i in industry j located in region k in year t solves the following problem:

$$\max_{l_{ijkt}} \int_{S_\varepsilon} \int_{S_\rho} \left[(1-\kappa)\pi_{ijkt}(l,\varepsilon,\rho) + \kappa \sum_{m \in \omega_{jkt}} \pi_{mjkt}(l,\varepsilon,\rho) \right] dF(\varepsilon) dG(\rho)$$

where:

$$\pi_{ijkt}(l,\varepsilon,\rho) = D_j (\varepsilon_{ijkt} l_{ijkt}^{1/\eta})^{1-1/\sigma} \left(\sum_{m \in \Omega_{jt}} (\varepsilon_{mjkt} l_{mjkt}^{1/\eta})^{1-1/\sigma} \right)^{\frac{\sigma}{\gamma}\frac{\gamma-1}{\sigma-1}-1} - \rho_{ijkt} \frac{l_{ijkt}}{z_{ijkt}}$$

Here ε is the vector of demand shocks, ρ is the vector of cost shocks, and l is the vector of production choices. The set of firms operating in industry j at time t is Ω_{jt}, and its subset of firms operating within region k is ω_{jkt}. For any given firm, z_{ijkt} is the component of their costs that is known before production decisions are made. Without heterogeneity in this, there would be no heterogeneity in l_{ijkt}. The parameter η allows for curvature in the cost function.

Notice that F and G are probability distributions over vectors, and we will consider covariance at the cluster, industry and year levels.

The first order condition implies:

$$\int_{S_\rho} \frac{\eta \rho_{ijkt} l_{ijkt}^{1-1/\eta}}{z_{ijkt}} dG(\rho) =$$

$$= \int_{S_\varepsilon} p_{ijkt}(l,\varepsilon) \left[\frac{\sigma-1}{\sigma} + \left(\frac{1}{\sigma} - \frac{1}{\gamma}\right) \frac{\kappa(\varepsilon_{ijkt} l_{ijkt}^{1/\eta})^{1-1/\sigma} + (1-\kappa) \sum_{n \in \omega_{jkt}} (\varepsilon_{njkt} l_{njkt}^{1/\eta})^{1-1/\sigma}}{\sum_{m \in \Omega_{jt}} (\varepsilon_{mjkt} l_{mjkt}^{1/\eta})^{1-1/\sigma}} \right] dF(\varepsilon)$$

where:

$$p_{ijkt}(l,\varepsilon) = D_j \varepsilon_{ijkt}^{1-1/\sigma} l_{ijkt}^{-1/\eta\sigma} \left(\sum_{m \in \Omega_{jt}} (\varepsilon_{mjkt} l_{mjkt})^{1/\eta(1-1/\sigma)} \right)^{\frac{\sigma}{\gamma}\frac{\gamma-1}{\sigma-1}-1}$$

Firms face a variety of shocks at different levels:

$$\varepsilon_{ijkt} = \nu_1 \varepsilon_t^1 + \nu_2 \varepsilon_{jt}^2 + \nu_3 \varepsilon_{ijkt}^3 + \nu_4 \varepsilon_{jkt}^4 + \nu_5 \varepsilon_{kt}^5$$

$$\rho_{ijkt} = \mu_1 \rho_t^1 + \mu_2 \rho_{jt}^2 + \mu_3 \rho_{ijkt}^3 + \mu_4 \rho_{jkt}^4 + \mu_5 \rho_{kt}^5$$

Therefore, we can separately analyze shocks at different levels.

A.4.1 Computational Implementation

The simulated dataset has T years, J industries and K regions. Every industry-region-year has I firms within it. The vectors ε and ρ are therefore of length $I \times J \times K \times T$. First, both ε and ρ are simulated M times. Then a vector L is drawn. Then L is input as the vector of production choices of firms. Using the first order condition, we then solve for the vector Z of anticipated costs that rationalizes the vector L. Together, Z, L, and the realization of shocks implies markups (using the method of De Loecker and Warzynski) and market shares for each firm. Then, for each realization, the regression described in the paper is run on the simulated data. This is done M times.

For these results we choose $\sigma = 5$, $\gamma = 3$, and $\kappa = 0.3$. We set $T = 11$, $J = 5$, $K = 8$, $I = 10$ and $M = 1000$. We assume that the log of each shock is a standard normal random variable.

A.4.2 Effects of Shocks: Comparative Statics

First we look at the effects of all twelve types of shocks individually. The table below presents the results of setting $\mu_1 = ... = \mu_5 = \nu_1 = ... = \nu_5 = 0$, then individually setting each to 1.

In each iteration of the simulation we run the following regression:

$$\frac{1}{\text{markup}_{ijkt}} = \alpha + \beta_1 s_{ijkt} + \beta_2 c_{jkt} + \delta_{ijkt}$$

where:

$$s_{ijkt} = \frac{\left(\varepsilon_{ijkt} y_{ijkt}\right)^{1-1/\sigma}}{\sum\limits_{m \in \Omega_{jt}} \left(\varepsilon_{mjkt} y_{mjkt}\right)^{1-1/\sigma}}$$

Table A.8: Simulation Results: Ex Post Shocks

	No Fixed Effects			Region-Year and Firm FEs		
Cost Shocks:	Avg. $\hat{\kappa}$	St. Dev. $\hat{\kappa}$	Adj. R^2	Avg. $\hat{\kappa}$	St. Dev. $\hat{\kappa}$	Adj. R^2
Year	0.2997	0.0059	-0.0000	0.3000	0.0094	0.5525
Industry-Year	0.2997	0.0038	-0.0003	0.3000	0.0116	-0.0003
Firm-Year	0.2009	2.1071	-0.0000	0.1889	17.9759	-0.0000
Cluster-Year	0.5954	43.1666	0.0016	0.6822	25.4955	0.0024
Region-Year	0.5571	8.9698	0.0015	0.3029	0.0163	0.4564
Demand Shocks:	Avg. $\hat{\kappa}$	St. Dev. $\hat{\kappa}$	Adj. R^2	Avg. $\hat{\kappa}$	St. Dev. $\hat{\kappa}$	Adj. R^2
Year	0.2998	0.0037	0.0006	0.3003	0.0083	0.7246
Industry-Year	0.2998	0.0024	0.0001	0.2998	0.0064	0.0001
Firm-Year	0.0927	0.1097	0.0024	0.0950	0.1252	0.0022
Cluster-Year	0.8398	0.2227	0.0205	0.8362	0.0780	0.0212
Region-Year	0.5769	12.6588	0.0018	0.3006	0.0117	0.4516

	Firm FEs			Region-Year FEs		
Cost Shocks:	Avg. $\hat{\kappa}$	St. Dev. $\hat{\kappa}$	Adj. R^2	Avg. $\hat{\kappa}$	St. Dev. $\hat{\kappa}$	Adj. R^2
Year	0.3004	0.0057	-0.0001	0.3001	0.0091	0.5538
Industry-Year	0.2992	0.0088	-0.0003	0.3001	0.0087	-0.0003
Firm-Year	0.1676	3.0608	-0.0000	0.1489	26.1445	-0.0001
Cluster-Year	0.6232	6.6294	0.0019	0.6574	13.2127	0.0025
Region-Year	0.6075	10.8352	0.0020	0.3010	0.0081	0.4529
Demand Shocks:	Avg. $\hat{\kappa}$	St. Dev. $\hat{\kappa}$	Adj. R^2	Avg. $\hat{\kappa}$	St. Dev. $\hat{\kappa}$	Adj. R^2
Year	0.2993	0.0052	0.0004	0.3004	0.0049	0.7179
Industry-Year	0.3001	0.0034	0.0002	0.3003	0.0051	0.0002
Firm-Year	0.0754	0.1022	0.0027	0.0937	0.1140	0.0228
Cluster-Year	0.8169	0.8512	0.0171	0.8479	0.0677	0.0228
Region-Year	0.6283	42.3770	0.0022	0.3003	0.0107	0.4519

$$c_{jkt} = \sum_{l \in \omega_{jkt}} s_{ljkt}$$

Here we present the simulated moments of $\hat{\kappa}$ defined by:

$$\hat{\kappa} \equiv \frac{\beta_2}{\beta_1 + \beta_2}$$

The results from these experiments are given in Table A.8. We provide four sets of results based on the set of fixed effects considered, and for each case we provide the average and standard deviation of κ across the 1000 simulations. We also provide the adjusted R^2 averaged across the 1000 simulations.

These results demonstrate two important things to help understand how our estimates of κ could be biased. Firm-year shocks bias estimates of κ downward, and cluster-year and region-year shocks bias estimates upward. The region-year shocks can be mitigated with region-year fixed effects: the bias is almost eliminated for cost shocks and is less severe for demand shocks. In the other cases, the adjusted R^2 of the model can fall considerably, but we see little evidence of bias in estimates of κ.

A.4.3 Calibrated Example

The previous subsection demonstrates that the most serious bias arises when ex post shocks are at the firm-year and cluster-year level. We now repeat the numerical exercise from the previous section but now we parameterize the model to replicate the results of our baseline results in column 7 of Table 4 in the paper. As in that regression, we include firm and year fixed effects and cluster standard errors at the firm level. We consider ex post shocks to productivity at the firm-year level and the cluster-year level, and we include measurement error at the firm-level. We also have idiosyncratic firm-year ex ante shocks. Each shock is assumed to be log-normal.

We calibrate six parameters: the variance of the three shocks, γ, σ, and κ. We match six moments: the coefficient estimate on the firm's own share and on the cluster's share, the standard errors on the firm's own share and on the cluster share, the average markup, and the regression's within-R^2.

The calibrated value of κ is 0.29, while the value in the model, as in the data, is 0.28. This demonstrates that, in this case, we actually underestimate the degree of collusion with our procedure relative to its true value. The calibrated standard deviation of the firm-year productivity shock is 0.011 while that of the cluster-year shock is 0.009. Our estimate of σ is 4.57 and γ is 2.74. The standard deviation of the measurement error is 0.044.

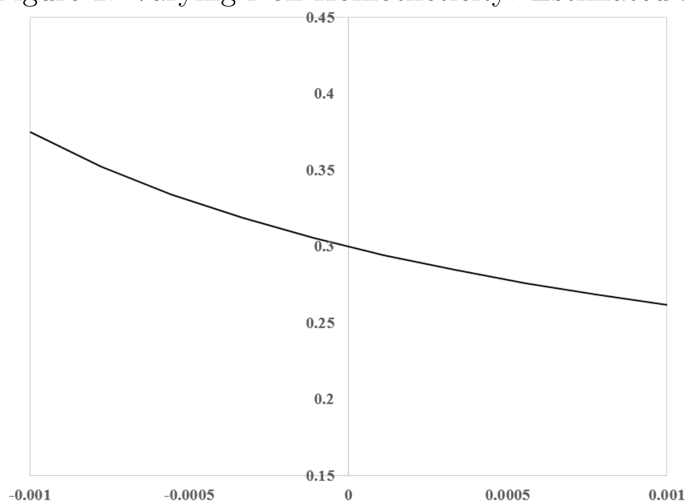
A.4.4 Departure from CES Demand

Next we consider the case where the demand system is instead given by:

$$q_{ijkt} = \left(\frac{p_{ijkt} + \bar{p}}{P_{jt}}\right)^{-\sigma} \left(\frac{P_{jt}}{P_t}\right)^{-\gamma} \tag{23}$$

Proceeding with the same simulation technique as above, we consider the case where there are no ex post shocks and vary the magnitude of $\bar{p}$.

Figure 1: Varying Non-Homotheticity: Estimated κ



The results are summarized below in Figures 1 and 2. As the value of $\bar{p}$ varies, as shown on the horizontal axis in both figures, on average our measure of κ will be affected monotonically as shown in Figure 1. As before, the true value of κ in this simulation is equal to 0.3. Figure 2 shows that this bias is entirely due to bias in the coefficient on firms' own shares. In fact, the coefficient on cluster shares is unbiased by $\bar{p}$.

This supports our conclusion that a non-CES demand system of this type affects our estimate of the magnitude of collusion. However, if we interpret the t-test of whether or not the coefficient on the cluster share is positive to be a test of collusion, that test is unaffected by non-CES demand systems of this form.

A.4.5 Measurement Error

Next, we consider the case where revenues are measured with error. We proceed as before, but now instead of unanticipated shocks, we study the effect of increases in the variances of the measurement error.

Following the parameterization in the first simulation exercise, Table A.9 shows the effects of measurement error. In the "Idiosyncratic" columns, we assume that measurement error has no correlation across firms. In the

Figure 2: Varying Non-Homotheticity: Coefficient Estimates

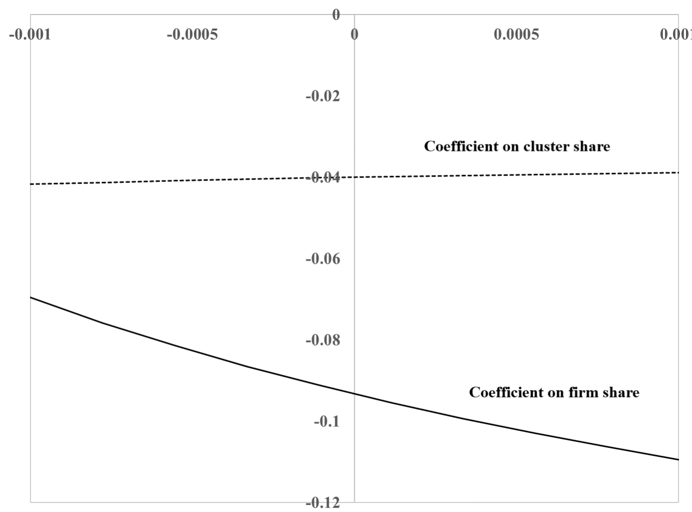


Table A.9: Effects of Measurement Error

	Measurement Error, Idiosyncratic			
Var. of Error	Avg. $\hat{\kappa}$	St. Dev. $\hat{\kappa}$	Avg. $\hat{\beta}_1$	$Avg.\hat{\beta}_2$
0.1	0.2935	0.1401	-0.1096	-0.0461
0.2	0.2885	12.0245	-0.1173	-0.0480
0.3	0.2821	4.6233	-0.1251	-0.0516
0.4	0.2825	37.1823	-0.1307	-0.0533
0.5	0.2582	16.6251	-0.1348	-0.0542
	Measurement Error, Cluster			
Var. of Error	Avg. $\hat{\kappa}$	St. Dev. $\hat{\kappa}$	Avg. $\hat{\beta}_1$	$Avg.\hat{\beta}_2$
0.1	0.2978	0.3590	-0.1088	-0.0461
0.2	0.3121	16.0014	-0.1139	-0.0505
0.3	0.3247	11.4702	-0.1202	-0.0532
0.4	0.3371	7.8773	-0.1261	-0.0565
0.5	0.3886	44.0116	-0.1327	-0.0569

"Cluster" columns, we consider the extreme case of correlation within clusters where measurement errors are equal in all firms of the same cluster.

A.4.6 Bias in Production Function Estimation

To estimate markups, we apply the methods of De Loecker and Warzynski (2012). This requires us to estimate the elasticity of output with respect to inputs, which we do by estimating production functions. We apply the methods of Ackerberg, Caves and Frazer (2015). However, we do not have data on physical quantities of output (we observe revenues) and our model explicitly includes imperfect competition instead of exogenous output prices. We find that this leads to downward bias in our estimates of production elasticities. However, we find no evidence that this biases our estimates of κ, the measure of the extent of cooperation that we derive.

To do this, first we replicate the exercise in Ackerberg, Caves and Frazer (2015) using the code made available by Kim, Luo and Su (2019). We then modify this code to reflect our demand system. We then conduct production function estimation in 5 cases: perfect competition ($\sigma = \gamma \to \infty$), monopolistic competition ($\sigma = \gamma = 5$), and three cases with imperfect competition and variable markups ($\sigma = 5$, $\gamma = 3.5$, $\kappa \in \{0, 0.3, 1\}$).

Results from the production function estimation are given in Table A.10. We find that the perfect competition case recovers the true production parameters, as expected. However, with monopolistic competition the parameters are biased downward. Yet we do not see any change in this downward bias in the cases with variable markups as κ changes. This demonstrates that imperfect competition is potentially a problem for recovering unbiased estimates of production function parameters. However, this problem is not affected by the presence of cooperation among firms.

Finally, we evaluate the possibility that bias in production function estimation influences our estimates of κ, the extent of cooperation among firms. To do this, we take the estimated production function parameters from the simulations described before, use those to measure markups with the De Loecker and Warzynski (2012) formula and run the regressions that constitute the main results in our paper to estimate κ using the simulation methods described in the previous subsection. The results of this are reported in the last two columns of Table A.10. We find that the estimated values of κ are close to their true values. In the cases with perfect competition or monopolistic competition there is no scope for cooperation and the estimated values are near zero, as expected.

Table A.10: Production Function Estimates, Varying Demand Structures

Demand System	Capital (β_k) Mean	Capital (β_k) St. Dev.	Labor (β_l) Mean	Labor (β_l) St. Dev.	Cooperation (κ) Mean	Cooperation (κ) St. Dev
Perfect Competition	0.3996	0.0153	0.6000	0.0091	0.0607	13.5859
Constant Markups	0.3200	0.0165	0.4799	0.0029	0.0527	11.5285
Variable Markups, $\kappa = 0$	0.2795	0.0125	0.4637	0.0256	0.0004	0.0164
Variable Markups, $\kappa = 0.3$	0.2736	0.0106	0.4627	0.0032	0.2996	0.0177
Variable Markups, $\kappa = 1$	0.2965	0.2459	0.4620	0.0060	0.9976	0.0355

Table A.11: Varying the Fraction of Firms Collaborating

Fraction Collaborating (ρ)	Avg. $\hat{\kappa}$	St. Dev. $\hat{\kappa}$	Avg. $\hat{\beta_1}$	$Avg.\hat{\beta_2}$	$\frac{\text{Avg. } \hat{\kappa}}{\rho}$
0.05	0.0151	0.0157	-0.1440	-0.0022	0.3012
0.15	0.0432	0.0167	-0.1401	-0.0063	0.2879
0.25	0.0733	0.0164	-0.1354	-0.0108	0.2932
0.35	0.1027	0.0160	-0.1312	-0.0150	0.2935
0.45	0.1304	0.0158	-0.1274	-0.0191	0.2898
0.55	0.1591	0.0160	-0.1229	-0.0232	0.2892
0.65	0.1881	0.0164	-0.1188	-0.0275	0.2894
0.75	0.2183	0.0170	-0.1142	-0.0318	0.2910
0.85	0.2472	0.0156	-0.1099	-0.0361	0.2908
0.95	0.2766	0.0173	-0.1057	-0.0403	0.2912

A.4.7 Unknown Sets of Cooperators

Our main exercises assume that we know the set of firms that are cooperating. Here we dispense with this assumption and see how our estimates of κ are affected. In particular, we assume that a fraction ρ of firms cooperate with one another, but that all firms are pooled together when estimating our main regression specifications. Here we assume that the true value of κ is 0.29, and we vary the fraction of firms that collaborate ρ from 5% to 95%. The resulting estimates of κ are presented in Table A.11. In the right-most column, we can see that dividing the estimated value of $\hat{\kappa}$ by the fraction of collaborating firms ρ generates values close to the true value of κ for each ρ. Hence, our estimated κ could be interpreted either as a measure of the fraction of firms that cooperate, or the intensity with which they cooperate.

A.5 Profitability of Cooperation by Productivity Level

We now show how profits change with κ by productivity level. We follow the simulation described in the previous subsection of this appendix. We then put productivity levels into bins, and compute average profit by bin for three

Table A.12: Profit by Productivity Level

Productivity Range		$\kappa = 0$	$\kappa = 0.3$	$\kappa = 1$
-1	0	2.24×10^{-5}	2.26×10^{-5}	2.33×10^{-5}
0	1	4.58×10^{-4}	4.39×10^{-4}	4.39×10^{-4}
1	2	0.0187	0.0189	0.0195
2	3	0.3609	0.2931	0.2893
3	4	10.272	10.384	10.706
4	5	110.84	112.83	113.85
5	6	1255.6	1266.5	1268.8
6	7	19243	19488	20248

values of κ: 0, 0.3 and 1. The results appear in Table A.12. We see that profits increase for each productivity bin as κ gets larger.

References

Ackerberg, Daniel, Kevin Caves, and Garth Frazer. 2015. "Identification Properties of Recent Production Function Estimators." *Econometrica*, 83(6): 2411–2451.

De Loecker, Jan, and Frederic Warzynski. 2012. "Markups and Firm-Level Export Status." *American Economic Review*, 102(6): 2437–71.

Kim, Kyoo il, Yao Luo, and Yingjun Su. 2019. "A Robust Approach to Estimating Production Functions: Replication of the ACF procedure." *Journal of Applied Econometrics*, 34(4): 612–619.