Geospatial Heterogeneity in Inflation:

A Market Concentration Story

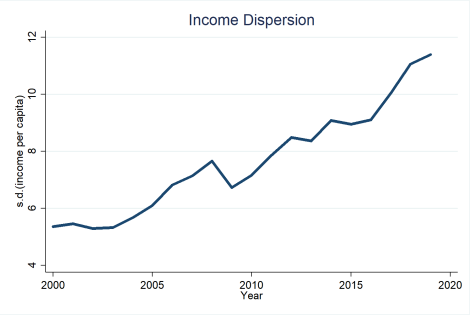
Seula Kim[1] Michael A. Navarrete[2]

[1] Princeton University

[2] University of Maryland

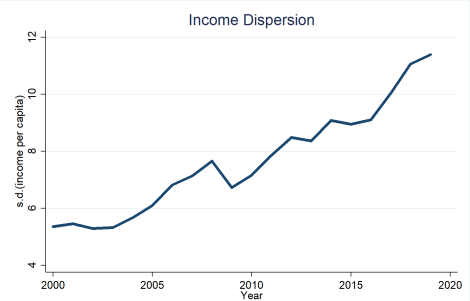
February 13, 2024

Motivation



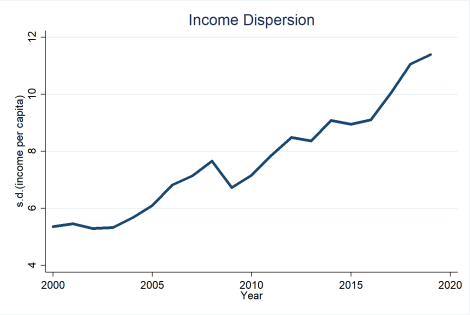
- Spatial income inequality has been on the rise in the United States
  - The dispersion of MSA-level (log) income has increased from 5.36 to 11.40 from 2000 to 2019

Motivation



- Spatial income inequality has been on the rise in the United States
  - The dispersion of MSA-level (log) income has increased from 5.36 to 11.40 from 2000 to 2019
- If inflation varies b/w rich and poor regions → "real" income inequality might have a different story

Motivation



- Spatial income inequality has been on the rise in the United States
  - The dispersion of MSA-level (log) income has increased from 5.36 to 11.40 from 2000 to 2019
- If inflation varies b/w rich and poor regions → "real" income inequality might have a different story
- Inflation is typically measured at the national level and presumed to be uniform across regions

Research Questions

1. Do inflation rates systematically vary across MSAs having different income level?
2. How is it related to local retailer market structure and power?

This Paper

- Uses Nielsen Retail Scanner and Business Dynamic Statistics

This Paper

- Uses Nielsen Retail Scanner and Business Dynamic Statistics
- Finds the following evidence:
  1. Food inflation rates vary across regions with different income level
  2. The poorest decile experiences about 10 p.p. ↑ inflation than the richest decile over 2006-2016
  3. The pattern holds for both aggregate and disaggregated food categories
  4. Larger (smaller) share of large (small) retailers in poorer areas
  5. The degree of market concentration is higher in poorer areas

This Paper

- Uses Nielsen Retail Scanner and Business Dynamic Statistics
- Finds the following evidence:
  1. Food inflation rates vary across regions with different income level
  2. The poorest decile experiences about 10 p.p. ↑ inflation than the richest decile over 2006-2016
  3. The pattern holds for both aggregate and disaggregated food categories
  4. Larger (smaller) share of large (small) retailers in poorer areas
  5. The degree of market concentration is higher in poorer areas
- Documents suggestive evidence about heterogeneous market power acting as a potential source behind the differential inflation rates

Data and Measures

Nielsen Retail Scanner (RMS)

- Nielsen contains detailed information for retail chains across U.S. markets
  - Covers 100 chains and over 40,000 individual stores
  - Weekly pricing, volume, store location, and merchandising conditions, etc.
  - Total sales worth over $200 billion/yr; 50% of total sales in grocery stores; 55% in drug stores; 32% in mass merchandisers; and 2% in convenience stores
  - Over 2.6 million 12-digit universal product codes (UPCs), aggregated to product modules and groups

Nielsen Retail Scanner (RMS)

- Nielsen contains detailed information for retail chains across U.S. markets
  - Covers 100 chains and over 40,000 individual stores
  - Weekly pricing, volume, store location, and merchandising conditions, etc.
  - Total sales worth over $200 billion/yr; 50% of total sales in grocery stores; 55% in drug stores; 32% in mass merchandisers; and 2% in convenience stores
  - Over 2.6 million 12-digit universal product codes (UPCs), aggregated to product modules and groups
- Build on the BLS concordance and construct a mapping b/w Nielsen and PCE food categories
- Map MSAs into into deciles based on income per capita

Business Dynamic Statistics (BDS)

- BDS provides annual measures of business dynamics
- Public version of the Census firm data (Longitudinal Business Dynamics)

Business Dynamic Statistics (BDS)

- BDS provides annual measures of business dynamics
- Public version of the Census firm data (Longitudinal Business Dynamics)
- Contains detailed information about establishment and firm characteristics
  - Total number of firms, employment, job creation and destruction
  - By different firm age/size bins, industry, and regions (MSAs)

Business Dynamic Statistics (BDS)

- BDS provides annual measures of business dynamics
- Public version of the Census firm data (Longitudinal Business Dynamics)
- Contains detailed information about establishment and firm characteristics
  - Total number of firms, employment, job creation and destruction
  - By different firm age/size bins, industry, and regions (MSAs)
- We use it for retail trade sector (NAICS 44-45)

Business Dynamic Statistics (BDS)

- BDS provides annual measures of business dynamics
- Public version of the Census firm data (Longitudinal Business Dynamics)
- Contains detailed information about establishment and firm characteristics
  - Total number of firms, employment, job creation and destruction
  - By different firm age/size bins, industry, and regions (MSAs)
- We use it for retail trade sector (NAICS 44-45)
- Use employment size and define large (500+) and small (20-) firms

Price Indices

Demand-based

$$\ln \Psi_t^G = \sum_{k \in \mathbb{C}_{t-1,t}} w_{kt} \ln \frac{p_{kt}}{p_{kt-1}},$$

- w_{kt} is a weight assigned to product k (typically based on the product's market share) in quarter t
- The set $\mathbb{C}_{t-1,t}$ is the set of all "continuing" goods that are sold both in period t and in period $t-1$

Price Indices

Demand-based

$$\ln \Psi_t^G = \sum_{k \in \mathbb{C}_{t-1,t}} w_{kt} \ln \frac{p_{kt}}{p_{kt-1}},$$

- w_{kt} is a weight assigned to product k (typically based on the product's market share) in quarter t
- The set $\mathbb{C}_{t-1,t}$ is the set of all "continuing" goods that are sold both in period t and in period $t-1$
- Our main focus is the Laspeyres index
  - Laspeyres index uses lagged expenditure shares as weights ($w_{kt} = s_{kt-1}$)
  - Paasche index uses current expenditure shares ($w_{kt} = s_{kt}$)

Price Indices

Demand-based

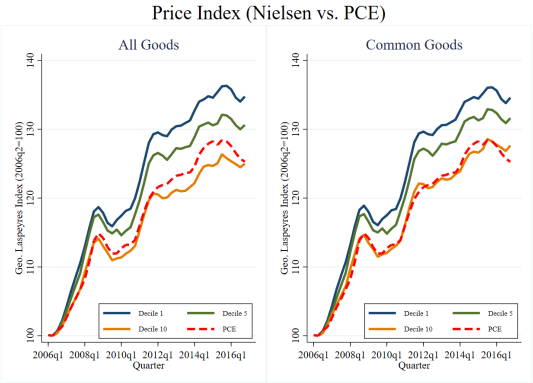
$$\ln \Psi_t^G = \sum_{k \in \mathbb{C}_{t-1,t}} w_{kt} \ln \frac{p_{kt}}{p_{kt-1}},$$

- w_{kt} is a weight assigned to product k (typically based on the product's market share) in quarter t
- The set $\mathbb{C}_{t-1,t}$ is the set of all "continuing" goods that are sold both in period t and in period $t-1$
- Our main focus is the Laspeyres index
  - Laspeyres index uses lagged expenditure shares as weights ($w_{kt} = s_{kt-1}$)
  - Paasche index uses current expenditure shares ($w_{kt} = s_{kt}$)
- As robustness check, we have used demand-based indices

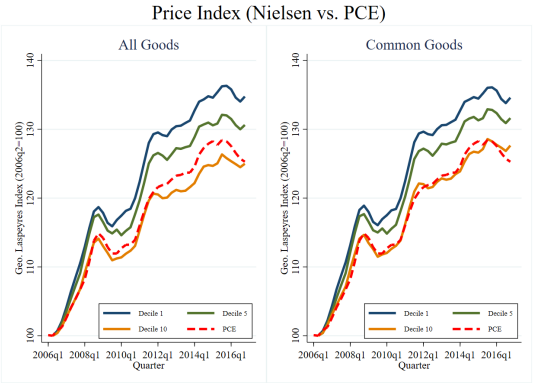
  e.g. Sato-vartia ($w_{kt} = \frac{\frac{(s_{k,t} - s_{k,t-1})}{(\ln s_{k,t} - \ln s_{k,t-1})}}{\sum_{k \in \mathbb{C}_{t-1,t}} \frac{(s_{k,t} - s_{k,t-1})}{(\ln s_{k,t} - \ln s_{k,t-1})}}$)

Main Findings

Spatial Heterogeneity in Inflation: Aggregate Food



Spatial Heterogeneity in Inflation: Aggregate Food



- Food price has been growing faster in poorer areas

Spatial Heterogeneity in Retailer Dynamics

Regression

Figure: Share of large firms (emp≥ 500)

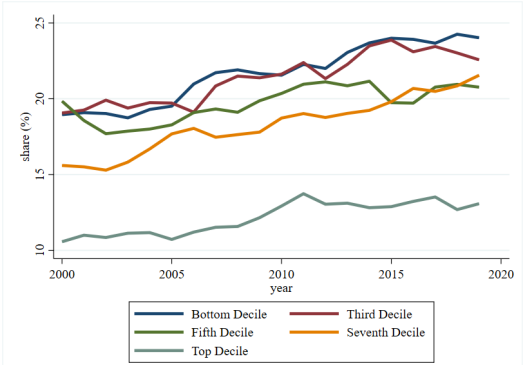
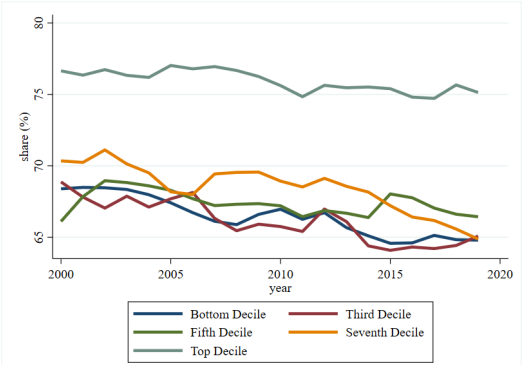


Figure: Share of small firms (emp< 20)



Spatial Heterogeneity in Retailer Dynamics

Regression

Figure: Share of large firms (emp≥ 500)

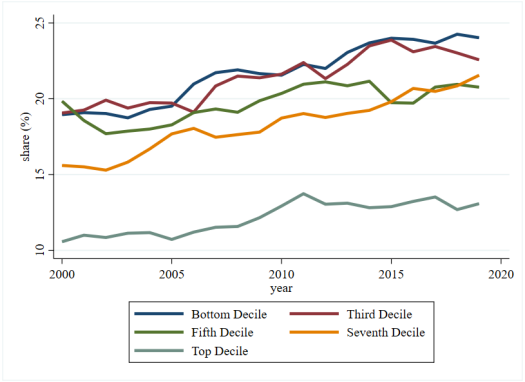
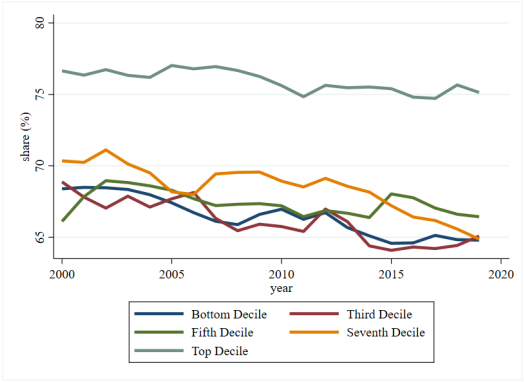


Figure: Share of small firms (emp< 20)



- More (less) large firms located in the bottom (top) decile
- Less (more) small firms located in the bottom (top) decile

Market Concentration across Income Deciles

$$HHI_{idt} = \beta_0 + \beta_1 Decile_{dt} + \delta_i + \delta_t + \varepsilon_{idt}$$

Market Concentration across Income Deciles

$$HHI_{idt} = \beta_0 + \beta_1 Decile_{dt} + \delta_i + \delta_t + \varepsilon_{idt}$$

- HHI_{idt} is the Herfindahl–Hirschman index of retailer sales for PCE food category i, MSAs in income decile d in quarter t
- $Decile_{dt}$ is an indicator for income decile
- δ_i, δ_t: PCE food category, year fixed effects

Market Concentration across Income Deciles

	HHI
Decile	-0.004***
	[0.000]
Constant	0.145***
	[0.001]
Observations	10,920

*** p<0.01, ** p<0.05, * p<0.1

- Higher retailer concentration is observed in lower income decile

Potential Mechanism: Retailers' Market Power

The 2015 Bird Flu

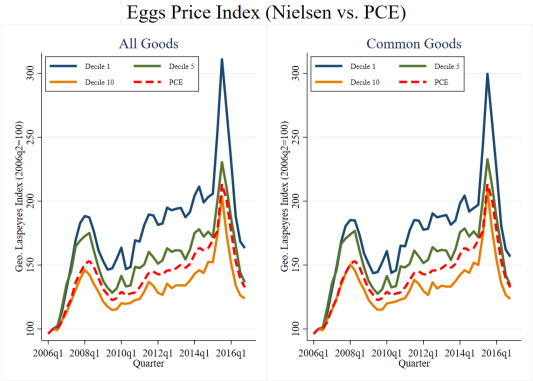
- In the USDA report: **the outbreak starts in December 2014 and starts to taper in June 2015**[1]
- By the end of June 2015, USDA estimated 36 million layers (birds that lay eggs) were lost due to the bird flu
- The USDA report along with the GAO report: geospatial heterogeneity

  → predominantly affected the central and western part of the US[2]

  → We exploit a USDA report detailing the farms that received subsidies for culling their layers
- PCE also captures this surge in inflation for eggs during the 2015 bird flu
  - Use a triple difference estimator to see the impact on inflation in eggs between higher HHI MSAs who received a government subsidy during the bird flu relative low HHI MSAs who received a government subsidy

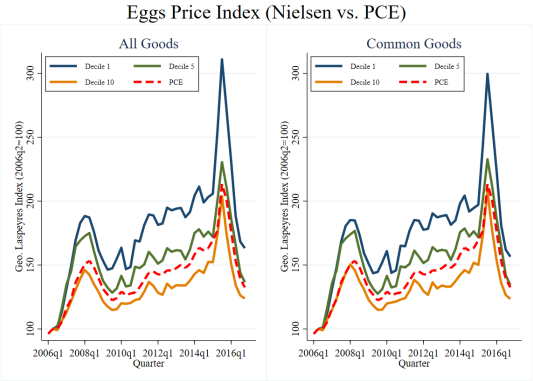
Spatial Heterogeneity in Inflation: Eggs

Demand-based



Spatial Heterogeneity in Inflation: Eggs

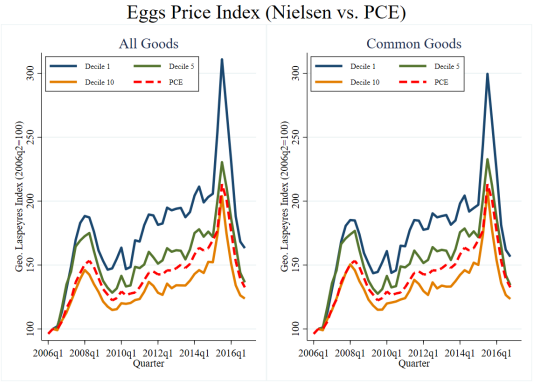
Demand-based



- The pattern stays robust

Spatial Heterogeneity in Inflation: Eggs

Demand-based



- The pattern stays robust
- There was a bird flu episode in 2015 causing the price spike

Simple OLS Estimator

$$P_{st} = \beta_0 + \beta_1 HHI_{st} + \delta_s + \delta_t^{yr} + \delta_t^{qtr} + \varepsilon_{st}$$

- P_{st} is the (geometric) Laspeyres index of eggs in MSA s, quarter t
- HHI_{st} is the HHI of retailer sales in MSA s, quarter t
- δ_s, δ_t^{yr}, δ_t^{qtr}: MSA, year, quarter fixed effects

OLS Estimation Results

	Price
HHI	0.011*
	[0.006]
Constant	1.009***
	[0.003]
Observations	9,484

*** p<0.01, ** p<0.05, * p<0.1

OLS Estimation Results

	Price
HHI	0.011*
	[0.006]
Constant	1.009***
	[0.003]
Observations	9,484

*** p<0.01, ** p<0.05, * p<0.1

- HHI increases price level
- Potential endogeneity issue exists

Triple Difference Estimator

$$
\begin{aligned}
P_{st} = {} & \beta_0 + \beta_2 HHI_{st} + \beta_4 (Treated_s \times HHI_{st}) \\
& + \beta_5 (Treated_s \times HHI_{st}) + \beta_6 (Post_t \times HHI_{st}) \\
& + \beta_7 (Treated_s \times HHI_{st} \times Post_t) + \delta_s + \delta_t^{yr} + \delta_t^{qtr} + \varepsilon_{st}
\end{aligned}
$$

Triple Difference Estimator

$$
\begin{aligned}
P_{st} = \beta_0 + \beta_2 HHI_{st} + \beta_4 (Treated_s \times HHI_{st}) \\
+ \beta_5 (Treated_s \times HHI_{st}) + \beta_6 (Post_t \times HHI_{st}) \\
+ \beta_7 (Treated_s \times HHI_{st} \times Post_t) + \delta_s + \delta_t^{yr} + \delta_t^{qtr} + \varepsilon_{st}
\end{aligned}
$$

- $Treated_s$ is a binary variable indicating whether MSA s is near to where egg layers were culled during the 2015 Bird Flu according to the USDA report.
- $Post_t$ is a binary variable equal to 1 if quarter t is after 2015q1
- P_{st}, HHI_{st}, and fixed effects are the same as before

Triple Difference Estimation Results

	Price	Price	Price
Bird Flu × HHI × Post		0.033***	0.018**
		[0.011]	[0.008]
Bird Flu × Post	-0.006***	-0.023***	-0.017***
	[0.002]	[0.007]	[0.005]
HHI × Post		-0.014**	-0.008*
		[0.006]	[0.005]
Bird Flu × HHI		-0.003	-0.030**
		[0.009]	[0.015]
HHI		0.013***	0.014***
		[0.002]	[0.005]
Fixed Effects	Yes	No	Yes
Observations	9,484	9,484	9,484

*** p<0.01, ** p<0.05, * p<0.1

Conclusion

Concluding Remarks

- Systematic diffs. in inflation rates and retailer market structure observed b/w poor and rich MSAs
- The poorest decile of MSAs faces (than the richest)
  1. Higher inflation rates for both aggregate and disaggregated food items
  2. Higher (Smaller) fraction of large (small) retailers
  3. Higher concentration rate of retailers
- Exploiting the 2015 bird flu episode, we find that more concentrated retailers charge higher prices
- Future work:
  1. Structural estimation of market power and its contribution to price growth
  2. Identify and quantify the impact on spatial inequality

THANK YOU! ☺

sk6285@princeton.edu / mnav@umd.edu

Appendix

Price Indices (Demand-based)

Main

- Potential issues with Laspeyres or Paasche: no consideration on substitution effects
- Demand-based indices with CES assumption
  - Sato-Vartia: based on common goods (b/w t and $t-1$)

$$\ln \Psi_t^{SV} = \sum_{k \in \mathbb{C}_{t-1,t}} w_{kt} \ln \frac{p_{kt}}{p_{kt-1}}, \quad \text{where } w_{kt} = \frac{\frac{(s_{k,t} - s_{k,t-1})}{(\ln s_{k,t} - \ln s_{k,t-1})}}{\sum_{k \in \mathbb{C}_{t-1,t}} \frac{(s_{k,t} - s_{k,t-1})}{(\ln s_{k,t} - \ln s_{k,t-1})}}$$

  - Feenstra-adjusted Sato-Vartia: further take into account product turnover

$$\ln \Psi_t^{Feenstra-SV} = \ln \Psi_t^{SV} + \frac{1}{\sigma - 1} \ln \frac{\lambda_{t,t-1}}{\lambda_{t-1,t}}, \quad \text{where } \lambda_{t,t-1} = \frac{\sum_{k \in \mathbb{C}_{t-1,t}} p_{k,t} q_{k,t}}{\sum_{k \in \Omega_t} p_{k,t} q_{k,t}}, \lambda_{t-1,t} = \frac{\sum_{k \in \mathbb{C}_{t-1,t}} p_{k,t-1} q_{k,t-1}}{\sum_{k \in \Omega_{t-1}} p_{k,t-1} q_{k,t-1}}$$

Spatial Heterogeneity in Inflation: Eggs (Demand-based)

Main

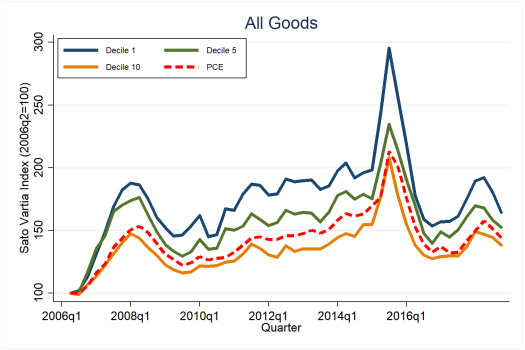


Figure: Sato Vartia

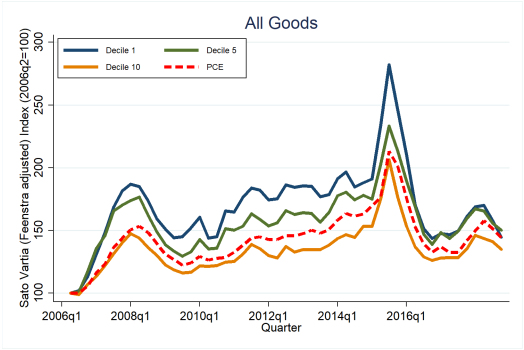


Figure: Feenstra-adjusted Sato Vartia

- The patterns stays robust (even after considering product turnover)
- Entering goods have larger sales value than exiting goods across all deciles (more so in decile 1)

Spatial Heterogeneity in Large Firm Activity

Main

$$LargeFirm_{st} = \beta_0 + \beta_1 Income_{st} + \delta_s + \delta_t + \varepsilon_{st}$$

- $LargeFirm_{st}$ is the (employment) share of large firms in MSA s, year t
  - Large firms: firms with 500+ employees
- $Income_{st}$ is income per capita in MSA s
- δ_s, δ_t: MSA, year fixed effects

Spatial Heterogeneity in Large Firm Activity

Main

	Large firm share	Large firm emp. share
Income	-0.040***	-0.040***
	[0.006]	[0.009]
Constant	19.896***	61.713***
	[0.214]	[0.345]
Observations	7,620	7,620

*** p<0.01, ** p<0.05, * p<0.1

- Larger firms are more active in lower income decile