Consumer Search and Firm Location: Theory and Evidence from the Garment Sector in Uganda

Anna Vitali
University College London

NBER Summer Institute
July 25, 2023
In low-income cities, economic activity is spatially concentrated, with firms producing similar goods locating next to one another.
In low-income cities, economic activity is spatially concentrated, with firms *producing* similar goods locating next to one another.
Motivation

► In low-income cities, economic activity is spatially concentrated, with firms producing similar goods locating next to one another.

Source: 2010 Ugandan Census of business establishments
Motivation: what drives firms’ spatial concentration?

- Quantitative spatial models: focus on production externalities (Ahlfeldt et al., 2015; Allen, et al., 2015; Monte et al., 2018; Owens et al., 2020)

- This paper: theory and data to quantify role of consumers’ information frictions
  - Agglomerated firms attract consumers by lowering cost of gathering information
  - More firms to compete with, but larger customer base (Stahl, 1982; Wolinski, 1983)
Motivation: what drives firms’ spatial concentration?

- Quantitative spatial models: focus on *production externalities* (Ahlfeldt et al., 2015; Allen, et al., 2015; Monte et al., 2018; Owens et al., 2020)

- This paper: theory and data to quantify role of *consumers’ information frictions*
  - Agglomerated firms attract consumers by lowering cost of gathering information
  - More firms to compete with, but larger customer base (Stahl, 1982; Wolinski, 1983)

- Salient in low-income settings: (i) in person search, (ii) bundling of production and sale
Motivation: what drives firms’ spatial concentration?

- Quantitative spatial models: focus on *production externalities* (Ahlfeldt et al., 2015; Allen, et al., 2015; Monte et al., 2018; Owens et al., 2020)

- This paper: theory and data to quantify role of *consumers’ information frictions*
  - Agglomerated firms attract consumers by lowering cost of gathering information
  - More firms to compete with, but larger customer base (Stahl, 1982; Wolinski, 1983)

- Salient in low-income settings: (i) in person search, (ii) bundling of production and sale

- Why does it matter?
  - Different welfare implications of urban / spatial policies
  - Consequences of demand-side constraints for misallocation

This paper

Information frictions → Consumer search → Firm location → Welfare

This paper

Information frictions → Consumer search → Firm location → Welfare

- **Data**: Collect data from garment firms and customers in Kampala
  1. *Transaction data*: estimate demand
  2. *Customer data*: evidence on search
  3. *Mystery shoppers data*: quality and prices
This paper

Information frictions $\rightarrow$ Consumer search $\rightarrow$ Firm location $\rightarrow$ Welfare

- **Data**: Collect data from garment firms and customers in Kampala
  1. *Transaction data*: estimate demand
  2. *Customer data*: evidence on search
  3. *Mystery shoppers data*: quality and prices

- **Model**: Quantitative equilibrium model of consumer search and firm location
  - *Agglomeration*: information frictions, proximity to suppliers
  - *Congestion*: within-location competition, transport costs, factor prices
This paper

Information frictions → Consumer search → Firm location → Welfare

▶ Data: Collect data from garment firms and customers in Kampala
  1. *Transaction data*: estimate demand
  2. *Customer data*: evidence on search
  3. *Mystery shoppers data*: quality and prices

▶ Model: Quantitative equilibrium model of consumer search and firm location
  • *Agglomeration*: information frictions, proximity to suppliers
  • *Congestion*: within-location competition, transport costs, factor prices

▶ Counterfactuals:
  • Equilibrium outcomes in the absence of information frictions
  • Assess welfare effects of policies on Ugandan policymakers’ agenda
    ▶ E-commerce
    ▶ Decongestion policies
Preview of findings

1. **Information frictions matter for the spatial distribution of firms**
   - Eliminating frictions leads to a 42% drop in sales in the core
1. **Information frictions matter for the spatial distribution of firms**
   - Eliminating frictions leads to a 42% drop in sales in the core

2. **Frictions limit the ability of high-quality firms to attract customers**
   - When removed: ↑ profits of high quality, ↓ profits of low-quality firms
   - 37% of low-quality firms make losses and are better off exiting the market
1. Information frictions matter for the spatial distribution of firms
   • Eliminating frictions leads to a 42% drop in sales in the core

2. Frictions limit the ability of high-quality firms to attract customers
   • When removed: \[\uparrow\] profits of high quality, \[\downarrow\] profits of low-quality firms
   • 37% of low-quality firms make losses and are better off exiting the market

3. Opposite effects of policies that target cause vs. symptoms of the inefficiency
   • E-commerce: 83% drop in sales in the core, primarily \textit{benefits} high-quality firms
   • Decongestion policies: disproportionately \textit{harm} high-quality firms
1. Quantitative spatial models of city structure
   - Role of production externalities (Ahlfeldt et al., 2015; Allen, et al., 2017; Monte et al., 2018; Davis et al., 2019; Agarwal et al., 2020; Owens et al., 2020; Miyauchi et al., 2021)
   - Contribution: additional role of information frictions

2. Consumer search
   - Impact on price elasticity and mark-ups (Hortaçsu and Syverson, 2004; Hong and Shum, 2006; De Los Santos et al., 2012; Murry and Zhou, 2020; Moraga-González et al., 2022)
   - Contribution: endogenize firm location

3. Information frictions and trade flows
   - Excessive price dispersion, survival of low-productivity firms (Arkolakis, 2010; Allen, 2014; Steinwender, 2018; Startz, 2021; Jensen, 2007; Aker, 2010; Goyal, 2010; Atkin et al., 2017; Jensen et al., 2018)
   - Contribution: demand-driven agglomeration to infer information frictions within a city
Outline

1. Setting and data
2. Motivating facts
3. Model
4. Estimation and counterfactuals
5. Conclusions
1. Setting and data
Kampala Garment sector

- Kampala: capital and economic hub of Uganda (60% of GDP)
- Garment sector: 42% of manufacturing firms, 15% of employment
- Median firm:
  - Descriptives
  - 1 worker, 3 machines, 3 m² surface, 93% informal
  - 5 years old, $100 revenues per month
- Hybrid between manufacturers and retailers
  - Production and sale done by the same person, in the same location
Data

1. Listing of 2,400+ establishments in Kampala

Panel A: Firms per square-km
Panel B: Selected Parishes
Data

1. Listing of 2,400+ establishments in Kampala
   2. Survey of 600 randomly selected firms (50% in Core, 50% in Periphery)
      • Firm and owner characteristics, location history, production process
      • Transaction data: type, quantity, price of product, final/retailer, origin of customer

3. Survey of 600 customers (50% final, 50% retailer)
   • Way in which consumers search for firms

4. Mystery shoppers exercise
   • Same garment commissioned to all firms
   • Quality assessment by expert tailor
Data

1. Listing of 2,400+ establishments in Kampala (Core/Periphery)
2. Survey of 600 randomly selected firms (50% in Core, 50% in Periphery)
   - Firm and owner characteristics, location history, production process
   - Transaction data: type, quantity, price of product, final/retailer, origin of customer
3. Survey of 600 customers (50% final, 50% retailer)
   - Way in which consumers search for firms
Data

1. Listing of 2,400+ establishments in Kampala (Core/Periphery)

2. Survey of 600 randomly selected firms (50% in Core, 50% in Periphery)
   - Firm and owner characteristics, location history, production process
   - Transaction data: type, quantity, price of product, final/retailer, origin of customer

3. Survey of 600 customers (50% final, 50% retailer)
   - Way in which consumers search for firms

4. Mystery shoppers exercise
   - Same garment commissioned to all firms
   - Quality assessment by expert tailor (Mystery Details)
2. Motivating Facts
Motivating Facts

▶ How do consumers search?

1. Pay large transport costs to travel to the core, but visit more firms prior to purchasing
   ▶ Average transport cost to core vs. periphery: $1.28 vs. $0.48, 34 vs. 17 minutes
   ▶ In core, visit 22% more firms before purchasing

Fact 1
Motivating Facts

- **How do consumers search?**
  1. Pay large transport costs to travel to the core, but visit more firms prior to purchasing
  2. **Customers visit the core to find more varieties and higher quality products**
    - Main reason for searching *Periphery*: proximity to home (64%)
    - Main reason for searching *Core*: number of tailors/varieties (55%), high-quality (58%)
Motivating Facts

▶ How do consumers search?

1. Pay large transport costs to travel to the core, but visit more firms prior to purchasing
2. Customers visit the core to find more varieties and higher quality products

▶ How do firms choose location?

3. **Firms in the core serve fewer, but larger customers, who grant them larger revenues**
   ▶ On average, firms in the core serve 18% fewer customers, but make 1.3 times the revenues of firms in periphery (Fact 3)
Motivating Facts

▶ How do consumers search?
1. Pay large transport costs to travel to the core, but visit more firms prior to purchasing
2. Customers visit the core to find more varieties and higher quality products

▶ How do firms choose location?
3. Firms in the core serve fewer, but larger customers, who grant them larger revenues
4. Are more likely to outsource intermediate tasks to nearby suppliers
   ▶ In core, 41% of workers involved in production are external (vs. 32% in periphery)
   ▶ Average distance to suppliers substantially lower in core (95% within 5 minutes walking)
Motivating Facts

▶ How do consumers search?

1. Pay large transport costs to travel to the core, but visit more firms prior to purchasing
2. Customers visit the core to find more varieties and higher quality products

▶ How do firms choose location?

3. Firms in the core serve fewer, but larger customers, who grant them larger revenues
4. Are more likely to outsource intermediate tasks to nearby suppliers

▶ Suggest: (i) consumers are affected by search frictions; (ii) firms internalize frictions in their choice of location

▶ Develop an equilibrium model to quantify the importance of this channel
Key features of the model

1. Demand:
   - **A1**: Consumers must pay *fixed* transport cost to the firm to observe preference over varieties
   - **A2**: Once in a location, observe preferences over all varieties sold in the location
     - All else equal, consumers prefer to search in locations with a high concentration of firms
     - Agglomeration stronger for large buyers due to *economies of scale in transport*
Key features of the model

1. Demand:
   - **A1**: Consumers must pay fixed transport cost to the firm to observe preference over varieties
   - **A2**: Once in a location, observe preferences over all varieties sold in the location
     → All else equal, consumers prefer to search in locations with a high concentration of firms
     → Agglomeration stronger for large buyers due to *economies of scale in transport*

2. Supply:
   - **A3**: In denser locations, cost of *outsourcing* is lower
     → Firms hire more external labor, face lower marginal cost
Key features of the model

1. **Demand:**
   - **A1:** Consumers must pay *fixed* transport cost to the firm to observe preference over varieties
   - **A2:** Once in a location, observe preferences over all varieties sold in the location
     → All else equal, consumers prefer to search in locations with a high concentration of firms
     → Agglomeration stronger for large buyers due to *economies of scale in transport*

2. **Supply:**
   - **A3:** In denser locations, cost of *outsourcing* is lower
     → Firms hire more external labor, face lower marginal cost

3. **Congestion:**
   - **A4:** Fiercer within-location *competition*, higher *commuting costs* and *factor prices* in high-density locations
3. Model
Set-up

Supply

- Finite number of firms $J$
  - Single-product, horizontally and vertically differentiated
  - Owned by individuals exogenously distributed across locations $l = \{1, 2, ..., N\}$
  - Idiosyncratic preferences over locations
Set-up

Supply

- Finite number of firms $J$
  - Single-product, horizontally and vertically differentiated
  - Owned by individuals exogenously distributed across locations $l = \{1, 2, ..., N\}$
  - Idiosyncratic preferences over locations

- They choose:
  1. Where to locate
  2. Once in a location, what price to charge
  3. What combination of land, internal and outsourced labor to employ
Set-up

Demand

- Finite number of consumers
  - Purchase one type of good, but heterogeneous in quantity $q$ demanded of a given good
  - Exogenously distributed across locations
  - Idiosyncratic preferences over products
Set-up

Demand

- Finite number of consumers I
  - Purchase one type of good, but heterogeneous in quantity $q$ demanded of a given good
  - Exogenously distributed across locations
  - Idiosyncratic preferences over products

- They choose:
  1. Which location to search
  2. Conditional on location, what product to buy
Set-up

Demand

▶ Finite number of consumers I
  • Purchase one type of good, but heterogeneous in quantity $q$ demanded of a given good
  • Exogenously distributed across locations
  • Idiosyncratic preferences over products

▶ They choose:
  1. Which location to search
  2. Conditional on location, what product to buy

Static model: formation of firm-customer matches and location choices that persist over time
Model overview

1. Demand:
   - Where to search
   - Which firm to buy from

2. Supply:
   - Production and outsourcing
   - Price choice

3. Firm location choice
Model overview

1. **Demand:**
   - Where to search
   - Which firm to buy from

2. **Supply:**
   - Production and outsourcing
   - Price choice

3. Firm location choice
Consumer utility

- Utility of consumer $i$ buying product $j$ in location $l$ is:

$$u_{ijl}^q = \left( \beta x_j + \xi_j + (1 - \sigma)\varepsilon_{ij} \right) q^\theta - \alpha p_{jl} q - C_{il}$$

- $x_j, \xi_j$: observable and unobservable product quality
  - *Vertical differentiation*: same ranking for all consumers

- $\varepsilon_{ij}$: idiosyncratic *taste* shock, iid $\sim$ standard $T1EV$
  - *Horizontal differentiation*: match specific (e.g. preference for style, color, fit, cut)

- $p_{jl}, q$: product $j$ price and quantity bought by individual $i$

- $C_{il}$: search cost
Search cost

\[ C_{il} = \tau_1 g(||z_i - z_l||) + \tau_2 \frac{N_l}{ar_l} + \omega_{il} \]

- \( ||z_i - z_l|| \): distance between consumer and firm location
  - Transport cost to location
- \( \frac{N_l}{ar_l} \): number of firms per square-km
  - Firm-specific search cost
- \( \omega_{il} \): individual-location specific search cost, iid \( \sim \) standard T1EV
  - E.g. idiosyncratic information
Timing of Consumer decision

1. Before searching, consumers do not observe $\varepsilon_{ij}$
2. Choose location based on available information
Timing of Consumer decision

1. Before searching, consumers do not observe $\varepsilon_{ij}$
2. Choose location based on available information
3. Upon paying the search cost, observe $\varepsilon_{ij}$ for all firms in the selected location
4. Choose firm that yields the highest utility
How does the number of firms affect demand?

1. **Market-size effect**: attracts consumers by increasing the number of available varieties
   - Let $s_{il}$ be the share of customers of type $i$ buying in location $l$
     \[
     s_{il}^q(L,p) = \frac{\left( \sum_{j=1}^{N_l} \exp \left( \frac{\delta_{jl}}{1-\sigma} \right) \right)^{q^\theta (1-\sigma)} \exp(-C_{il})}{\sum_{k=0}^{N} \left[ \left( \sum_{h=1}^{N_h} \exp \left( \frac{\delta_{hk}}{1-\sigma} \right) \right)^{q^\theta (1-\sigma)} \exp(-C_{ik}) \right]}
     \]
   - All else equal, location share is increasing in $N_l$
     - Observe more $\varepsilon$ draws (more varieties)
     - Higher probability of finding product with desired characteristics
How does the number of firms affect demand?

1. **Market-size effect**: attracts consumers by increasing the number of available varieties
   - Let $s_{il}$ be the share of customers of type $i$ buying in location $l$
   
   \[
   s_{il}^q(L,p) = \frac{\left( \sum_{j=1}^{N_l} \exp \left( \frac{\delta^q_{il}}{1-\sigma} \right) \right)^{q^q(1-\sigma)} \exp(-C_{il})}{\sum_{k=0}^{N} \left( \sum_{h=1}^{N_h} \exp \left( \frac{\delta^q_{hk}}{1-\sigma} \right) \right)^{q^q(1-\sigma)} \exp(-C_{ik})}
   \]
   
   ▶ All else equal, location share is increasing in $N_l$
   
   ▶ Effect is increasing in $q$
     - Large consumers benefit from a better match over all units of products bought
How does the number of firms affect demand?

1. Market-size effect: attracts consumers by increasing the number of available varieties

2. Market-share effect: increases competition within a location
   - Let $s_{j|i}$ be the share of customers of type $i$ buying from firm $j$ conditional location $l$

   $$s_{j|i}^q(p_l) = \frac{\exp \left( \frac{\delta_{jl}^q}{1-\sigma} \right)}{\sum_{h=1}^{N_l} \exp \left( \frac{\delta_{hl}^q}{1-\sigma} \right)}$$

   ▶ All else equal, conditional firm share is decreasing in $N_l$
Unconditional demand and sorting

Unconditional demand: \( s_{ijl}^q(L,p) = s_{il}^q(L,p) \times s_{jl\mid l}^q(p_l) \)

- Two opposite effects of number of firms \( N_l \):
  1. Market-size effect: \( \uparrow s_{il}^q \)
  2. Market-share effect: \( \downarrow s_{jl\mid l}^q \)

- Marginal effect depends on relative strength of these two forces
Unconditional demand and sorting

**Unconditional demand:** \( s_{ijl}^q(L,p) = s_{il}^q(L,p) \times s_{j|l}^q(p_l) \)

- Two opposite effects of number of firms \( N_l \):
  1. **Market-size effect:** \( \uparrow s_{il}^q \)
  2. **Market-share effect:** \( \downarrow s_{j|l}^q \)

- Marginal effect depends on relative strength of these two forces

**Sorting:** In absolute value, \( \frac{S_{ijl}}{N_l} \) larger for high-quality firms

- If positive, high-quality firms sort into larger locations
- Intuition: all firms equally benefit from market-size, but high-quality firms less affected by market-share effect
Model overview

1. Demand:
   - Where to search
   - Which firm to buy from

2. Supply:
   - Production and outsourcing
   - Price choice

3. Firm location choice
Firm location choice

- Static game of incomplete information
- Given spatial distribution of other firms \((L)\), firm \(j\)'s profits in \(l\) are:

\[
\Pi_{jl}(L, p) = \pi_{jl}(L, p) - FC_{jl}
\]

- Variable profits: \(\pi_{jl}(L, p) = (p_{jl} - c_{jl})Q_{jl}(L, p)\)
- Fixed cost: \(FC_{jl} = \tau_3 g(||z_j - z_l||) + e_{jl}\)
  - \(||z_j - z_l||\): distance between the owner's home and the firm location
  - \(e_{jl}\): idiosyncratic preference shock, \textit{unobserved} by other firms, iid \(\sim T1EV\)
Consumption and production externalities

▶ **Variable profits:** $\pi_{jl}(L, p) = (p_{jl} - c_{jl})Q_{jl}(L, p)$

\[
\frac{d\pi_{jl}(L, p)}{dN_l} = (p_{jl} - c_{jl}) \frac{\partial Q_{jl}(L, p)}{\partial N_l} - \frac{\partial c_{jl}}{\partial N_l} Q_{jl}(L, p) + \frac{\partial p_{jl}}{\partial N_l} Q_{jl}(L, p)
\]

- $\frac{\partial Q_{jl}(L, p)}{\partial N_l} > 0$: if *market-size* effect dominates *market-share* effect, demand increasing in $N_l$
- $\frac{\partial c_{jl}}{\partial N_l} < 0$: cost of outsourcing is lower in high $N_l$ locations (but land prices and wages higher)
- Impacts might be mitigated or enhanced by effect on equilibrium prices
4. Estimation and Counterfactuals
Bringing the model to the data

▶ CONSUMERS: Two types - **Final, retailers**

- Final consumers buy one unit of output
- Retailers buy ten units of output (median transaction size in data)
Bringing the model to the data

▶ CONSUMERS: Two types - **Final, retailers**
  - Final consumers buy one unit of output
  - Retailers buy ten units of output (median transaction size in data)

▶ FIRMS: Two types - **High, low quality**
  - Quality score from mystery shoppers exercise above/below median
  - Firms only know number of high and low-quality firms, but not owners’ origin
Bringing the model to the data

- **CONSUMERS**: Two types - Final, retailers
  - Final consumers buy one unit of output
  - Retailers buy ten units of output (median transaction size in data)

- **FIRMS**: Two types - High, low quality
  - Quality score from mystery shoppers exercise above/below median
  - Firms only know number of high and low-quality firms, but not owners’ origin

- **LOCATION**: Restrict firm choice to
  - Parish where owner lives
  - Core of the city
Roadmap for estimation procedure

1. **Demand**: $\{\alpha, \beta, \sigma, \theta, \tau_1, \tau_2\}$
   - Data: transaction and mystery shoppers (price and quality)
   - Method: maximum likelihood

2. **Supply**: $\{\delta, A_l, T(N_l)\}$
   - Data: firm survey (choice of land and labor, wages, rents)
   - Method: simulated method of moments

3. **Location**: $\{\tau_3\}$
   - Data: firm survey (firm location and owner’s origin)
   - Method: Nested Fixed Point Algorithm (Rust 1987)
Overview of counterfactuals

1. Shutting down information frictions
   - Consumers observe all product characteristics prior to purchasing
   - Transactions are in person: must pay transport cost to the firm

2. E-commerce
   - Consumers observe all product characteristics prior to purchasing
   - Can get products delivered at a flat fee

3. Decongestion policies
   - Cap the number of firms allowed to operate in the core
   - Motor-cycle taxis ban in the core
Overview of counterfactuals

1. Shutting down information frictions
   ▶ Consumers observe all product characteristics prior to purchasing
   ▶ Transactions are in person: must pay transport cost to the firm

2. E-commerce
   ▶ Consumers observe all product characteristics prior to purchasing
   ▶ Can get products delivered at a flat fee
Overview of counterfactuals

1. Shutting down information frictions
   - Consumers observe all product characteristics prior to purchasing
   - Transactions are in person: must pay transport cost to the firm

2. E-commerce
   - Consumers observe all product characteristics prior to purchasing
   - Can get products delivered at a flat fee

3. Decongestion policies
   - Cap the number of firms allowed to operate in the core
   - Motor-cycle taxis ban in the core
## Counterfactual 1: Shutting down information frictions

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>No information frictions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share of firms in core</td>
<td>0.365</td>
<td>0.335</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-8.2%</td>
</tr>
<tr>
<td>Share of sales in core</td>
<td>0.382</td>
<td>0.222</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-42%</td>
</tr>
<tr>
<td>Average price</td>
<td>20.44</td>
<td>17.52</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-14%</td>
</tr>
<tr>
<td>Average profits</td>
<td>476.0</td>
<td>391.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-18%</td>
</tr>
<tr>
<td>Average consumer welfare</td>
<td>19.22</td>
<td>21.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+11%</td>
</tr>
</tbody>
</table>
### Counterfactual 1: Shutting down information frictions

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>No information frictions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share of firms in core</td>
<td>0.365</td>
<td>0.335</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-8.2%</td>
</tr>
<tr>
<td>Share of sales in core</td>
<td>0.382</td>
<td>0.222</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-42%</td>
</tr>
<tr>
<td>Average price</td>
<td>20.44</td>
<td>17.52</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-14%</td>
</tr>
<tr>
<td>Average profits</td>
<td>476.0</td>
<td>391.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-18%</td>
</tr>
<tr>
<td>Average consumer welfare</td>
<td>19.22</td>
<td>21.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+11%</td>
</tr>
</tbody>
</table>
Counterfactuals 1: High vs. low-quality firms

Panel A: Share of firms in core

![Bar chart showing share of firms in core for high-quality and low-quality firms. The chart compares the baseline scenario with the scenario without info frictions.]
Counterfactuals 1: High vs. low-quality firms

Panel A: Share of firms in core

Panel B: Firm profits

Motivation

1. Setting and data

2. Motivating Facts

3. Model

4. Estimation and Counterfactuals

6. Conclusions
Counterfactuals 2 & 3: E-commerce and caps

▶ In Kampala, travel time estimated to be 13.5% of the city GDP (+4.2% with congestion) (Baertsch 2020)

▶ Simulate the impact of two policies on the agenda:

1. E-commerce platform: no information frictions, flat delivery fee
2. Cap number of owners that can operate in the core

KCCA removes mobile money kiosks off Kampala streets

KCCA officers remove a kiosk from Burton street In Kampala. URN photo
E-commerce vs. caps on high and low-quality firms

▶ Compare effect on **profits** of e-commerce and caps that induce same spatial dispersion
E-commerce vs. caps on high and low-quality firms

▶ Compare effect on **profits** of e-commerce and caps that induce same spatial dispersion
Conclusions

▶ Case study that highlights the importance of information frictions for firms’ location choice
▶ Framework extends to contexts with information frictions and costly search:
  • Low and high-income settings in which search is conducted in person
  • But might also contribute to concentration of sellers on online platforms
Case study that highlights the importance of information frictions for firms’ location choice

Framework extends to contexts with information frictions and costly search:
- Low and high-income settings in which search is conducted in person
- But might also contribute to concentration of sellers on online platforms

Broader implications:
1. Information frictions limit the ability of high-quality firms to attract customers
   → Role in explaining slow growth and firm size distribution
2. Firms rely on networks to achieve scale via outsourcing and machines rental (Bassi et al. 2022)
   → Within-firm contracting frictions preventing firms from merging/integrating
Source: 1919 US Census of Manufacturers
Manufacturing Firms, 2002 Census

Source: 2002 Census of Business Establishments
Firm Density in Selected Parishes

- Kisugu
- Mbuya I
- Kibuye II
- Kasubi
- Bwaise II
- Katwe I
- Kanwokya II
- Naguru I
- Wandegeya
- Kisenyi III
- Nakivubo-Sharuyako
- Nansana
- Kasubi
- Wandegeya
- Kisenyi II
- Nakasero IV

<table>
<thead>
<tr>
<th>Parishes</th>
<th>Firms per Square-km</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kisugu</td>
<td>1,642</td>
</tr>
<tr>
<td>Mbuya I</td>
<td>2,521</td>
</tr>
<tr>
<td>Kibuye II</td>
<td>3,277</td>
</tr>
<tr>
<td>Kasubi</td>
<td>21</td>
</tr>
<tr>
<td>Bwaise II</td>
<td>25</td>
</tr>
<tr>
<td>Katwe I</td>
<td>40</td>
</tr>
<tr>
<td>Kanwokya II</td>
<td>47</td>
</tr>
<tr>
<td>Naguru I</td>
<td>60</td>
</tr>
<tr>
<td>Wandegeya</td>
<td>71</td>
</tr>
<tr>
<td>Kisenyi III</td>
<td>137</td>
</tr>
<tr>
<td>Nakivubo-Sharuyako</td>
<td>196</td>
</tr>
<tr>
<td>Nansana</td>
<td>315</td>
</tr>
<tr>
<td>Kasubi</td>
<td>71</td>
</tr>
<tr>
<td>Wandegeya</td>
<td>71</td>
</tr>
<tr>
<td>Kisenyi II</td>
<td>1,642</td>
</tr>
<tr>
<td>Nakasero IV</td>
<td>2,521</td>
</tr>
</tbody>
</table>

Back
Appendix
## Firm Descriptives

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>Core</th>
<th>Periphery</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of workers</td>
<td>1.319</td>
<td>1.250</td>
<td>1.701</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>[1.000]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of machines</td>
<td>3.674</td>
<td>3.573</td>
<td>4.224</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td>[3.000]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size of premises (m²)</td>
<td>3.005</td>
<td>2.652</td>
<td>4.952</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>[2.000]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years of operation</td>
<td>8.001</td>
<td>7.974</td>
<td>8.151</td>
<td>.814</td>
</tr>
<tr>
<td></td>
<td>[5.000]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monthly revenues (USD)</td>
<td>167.039</td>
<td>179.402</td>
<td>100.611</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>[100.442]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rent per square-meter (USD)</td>
<td>19.459</td>
<td>20.847</td>
<td>11.717</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>[14.147]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monthly commuting cost (USD)</td>
<td>36.642</td>
<td>39.817</td>
<td>19.564</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>[40.743]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>601</td>
<td>302</td>
<td>299</td>
<td></td>
</tr>
</tbody>
</table>
## Mystery shoppers design and scoring

### Assessment Criteria

<table>
<thead>
<tr>
<th>BUSINESS ID:</th>
<th>SCORING GUIDE</th>
<th>MAX SCORE</th>
<th>SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 DARTS</td>
<td>Dart of 4 “long by 1”wide</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Correctly sewn</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Press to the right side</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Position of the Dart observed</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>2 COLLAR</td>
<td>Peter Pan/Baby Collar</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fixed correctly round the neckline</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>3 SLEEVES</td>
<td>Sleeved Well Gathered</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sleeve Length 8”</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Round sleeve 14”</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Correctly fixed on Bodice</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4 SKIRT</td>
<td>Skirt length 18”</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Skirt Equally Gathered</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Neatly fixed to Bodice</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Correct Seam Allowance</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Skirt bottom shaped round</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>5 ZIP</td>
<td>Zip attached to Centre back seam</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Right color of Zip</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Right length of Zip</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>6 SEAM</td>
<td>Right Seam Allowance “Y2-1”</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Correctly Pressed</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Neatly Finished Edges</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>7 HEM</td>
<td>Hemmed bottom of Dress</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hem lin-2ins</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hem Neatly sewn</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hem well pressed</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

### Appendix

Back
1. Firms and customers pay high transport costs to the Core

- Transport cost to Core: 22% of firms’ daily revenues, 14% of transaction value
  - More than twice the cost to travel to Periphery
- In the Core, customers visit 22% more firms prior to purchasing

PANEL A: Firms per square-km

PANEL B: Residents per square-km
2. Firms in the core sell higher quality products on average

- On average, firms in core are of higher quality ($p$-value $= 0.039$)
- Difference driven by tails, suggesting the best firms select into the core (Combes et al. 2012)
Customers’ reasons for searching in Core vs. Periphery

- Proximity to home
- Proximity to good transport
- Low prices
- Reputation of being reliable
- Large number of other-sector firms
- Reputation of good quality
- Large number of tailors/varieties
- Proximity to workplace
- Proximity to good transport
- Proximity to home

-6 -4 -2 0 .2 .4 .6

95% CI

Appendix
Correlation between enumerators’ and expert tailor’s scores
Product variety across locations

**Periphery**
- kids clothing
- dress
- jacket
- trousers/shorts
- skirt
- top/shirt
- gown/suit
- traditional clothing
- school uniform
- accessories
- repairs
- overall/apron
- bedsheets/curtains

**Core**
- kids clothing
- dress
- jacket
- trousers/shorts
- skirt
- top/shirt
- gown/suit
- traditional clothing
- school uniform
- accessories
- repairs
- overall/apron
- bedsheets/curtains
Typical dress
### Way in which customers would search for firms

<table>
<thead>
<tr>
<th></th>
<th>% of final customers</th>
<th>% of retail customers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walk into any firm</td>
<td>53.5</td>
<td>55.8</td>
</tr>
<tr>
<td>Ask friends/family members</td>
<td>43.9</td>
<td>42.4</td>
</tr>
<tr>
<td>Ask other tailoring firm</td>
<td>14.5</td>
<td>33.8</td>
</tr>
<tr>
<td>Ask firm in different sector</td>
<td>6.9</td>
<td>11.9</td>
</tr>
<tr>
<td>Look on the internet</td>
<td>7.9</td>
<td>4.0</td>
</tr>
</tbody>
</table>

**Note:** Data is from the baseline of customers.
3. Firms in Core serve fewer customers, but a larger share of retailers

<table>
<thead>
<tr>
<th></th>
<th>(1) Daily customers</th>
<th>(2) Daily revenues (USD)</th>
<th>(3) Share of retailers</th>
<th>(4) Transaction Value (USD)</th>
<th>(5) Quantity</th>
<th>(6) Unit price (USD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core</td>
<td>-0.163**</td>
<td>9.336***</td>
<td>0.446***</td>
<td>4.289***</td>
<td>14.35***</td>
<td>-0.215***</td>
</tr>
<tr>
<td></td>
<td>(0.0799)</td>
<td>(2.340)</td>
<td>(0.0294)</td>
<td>(0.923)</td>
<td>(1.493)</td>
<td>(0.243)</td>
</tr>
<tr>
<td>Quality score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.489***</td>
<td>-1.936**</td>
<td></td>
<td>0.836***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.571)</td>
<td>(0.781)</td>
<td></td>
<td>(0.183)</td>
</tr>
<tr>
<td>Mean</td>
<td>Periphery</td>
<td>0.980</td>
<td>7.423</td>
<td>0.102</td>
<td>6.763</td>
<td>3.628</td>
</tr>
<tr>
<td>N. Observations</td>
<td>546</td>
<td>546</td>
<td>512</td>
<td>2726</td>
<td>2726</td>
<td>2726</td>
</tr>
</tbody>
</table>

Note: Data is from transactions records and mystery shoppers. In Columns 1 to 3, the unit of observation is the firm. In Columns 4 to 6, it is the transaction. Standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$
3. Firms in Core serve fewer customers, but a larger share of retailers

<table>
<thead>
<tr>
<th></th>
<th>(1) Daily customers</th>
<th>(2) Daily revenues (USD)</th>
<th>(3) Share of retailers</th>
<th>(4) Transaction Value (USD)</th>
<th>(5) Quantity</th>
<th>(6) Unit price (USD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core</td>
<td>-0.163***</td>
<td>9.336***</td>
<td>0.446***</td>
<td>4.289***</td>
<td>14.35***</td>
<td>-0.215</td>
</tr>
<tr>
<td></td>
<td>(0.0799)</td>
<td>(2.340)</td>
<td>(0.0294)</td>
<td>(0.923)</td>
<td>(1.493)</td>
<td>(0.243)</td>
</tr>
<tr>
<td>Quality score</td>
<td></td>
<td></td>
<td></td>
<td>1.489***</td>
<td>-1.936**</td>
<td>0.836***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.571)</td>
<td>(0.781)</td>
<td>(0.183)</td>
</tr>
<tr>
<td>Mean Periphery</td>
<td>0.980</td>
<td>7.423</td>
<td>0.102</td>
<td>6.763</td>
<td>3.628</td>
<td>3.136</td>
</tr>
<tr>
<td>Product FEs</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
<td></td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>N. Observations</td>
<td>546 546 512</td>
<td></td>
<td></td>
<td>2726 2726 2726</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Data is from transactions records and mystery shoppers. In Columns 1 to 3, the unit of observation is the firm. In Columns 4 to 6, it is the transaction. Standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$
3. Firms in Core serve fewer customers, but a larger share of retailers

<table>
<thead>
<tr>
<th></th>
<th>(1) Daily customers</th>
<th>(2) Daily revenues (USD)</th>
<th>(3) Share of retailers</th>
<th>(4) Transaction Value (USD)</th>
<th>(5) Quantity</th>
<th>(6) Unit price (USD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core</td>
<td>-0.163***</td>
<td>9.336***</td>
<td>0.446***</td>
<td>4.289***</td>
<td>14.35***</td>
<td>-0.215***</td>
</tr>
<tr>
<td></td>
<td>(0.0799)</td>
<td>(2.340)</td>
<td>(0.0294)</td>
<td>(0.923)</td>
<td>(1.493)</td>
<td>(0.243)</td>
</tr>
<tr>
<td>Quality score</td>
<td></td>
<td></td>
<td></td>
<td>1.489***</td>
<td>-1.936**</td>
<td>0.836***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.571)</td>
<td>(0.781)</td>
<td>(0.183)</td>
</tr>
<tr>
<td>Mean</td>
<td>Periphery</td>
<td>0.980</td>
<td>7.423</td>
<td>0.102</td>
<td>6.763</td>
<td>3.628</td>
</tr>
<tr>
<td>Product FEs</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>N. Observations</td>
<td>546</td>
<td>546</td>
<td>512</td>
<td>2726</td>
<td>2726</td>
<td>2726</td>
</tr>
</tbody>
</table>

Note: Data is from transactions records and mystery shoppers. In Columns 1 to 3, the unit of observation is the firm. In Columns 4 to 6, it is the transaction. Standard errors in parentheses. * p < .10, ** p < .05, *** p < .01
Firms’ reasons for *locating* in Core vs. Periphery?

- Access to customers
- Proximity to potential suppliers
- Proximity to potential employees
- Proximity to other-sector firms
- Good transport/amenities
- Affordable rent
- Proximity to home

<table>
<thead>
<tr>
<th>Reason</th>
<th>Share Core</th>
<th>Share Periphery</th>
<th>Difference Core-Periphery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access to customers</td>
<td>0.5</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td>Proximity to potential suppliers</td>
<td>0.4</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>Proximity to potential employees</td>
<td>0.3</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>Proximity to other-sector firms</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Good transport/amenities</td>
<td>0.4</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>Affordable rent</td>
<td>0.5</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td>Proximity to home</td>
<td>0.5</td>
<td>0.3</td>
<td>0.2</td>
</tr>
</tbody>
</table>

95% CI
Firm relocation

<table>
<thead>
<tr>
<th>Category</th>
<th>% of firms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No Relocation</strong></td>
<td></td>
</tr>
<tr>
<td>never relocated</td>
<td>54.4</td>
</tr>
<tr>
<td><strong>Relocation to Core</strong></td>
<td></td>
</tr>
<tr>
<td>periphery to core</td>
<td>5.32</td>
</tr>
<tr>
<td>outside Kampala to core</td>
<td>6.16</td>
</tr>
<tr>
<td>relocated within core</td>
<td>11.3</td>
</tr>
<tr>
<td><strong>Relocation to Periphery</strong></td>
<td></td>
</tr>
<tr>
<td>core to periphery</td>
<td>2.83</td>
</tr>
<tr>
<td>outside Kampala to periphery</td>
<td>7.82</td>
</tr>
<tr>
<td>relocated within periphery</td>
<td>12.1</td>
</tr>
</tbody>
</table>
## Correlates of prices

<table>
<thead>
<tr>
<th></th>
<th>(1) Unit price</th>
<th>(2) Unit price</th>
<th>(3) Unit price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core</td>
<td>-0.733***</td>
<td>-0.215</td>
<td>0.620**</td>
</tr>
<tr>
<td></td>
<td>(0.227)</td>
<td>(0.248)</td>
<td>(0.311)</td>
</tr>
<tr>
<td>Quality score</td>
<td>1.032***</td>
<td>0.962***</td>
<td>0.836***</td>
</tr>
<tr>
<td></td>
<td>(0.195)</td>
<td>(0.191)</td>
<td>(0.183)</td>
</tr>
<tr>
<td>Number of items</td>
<td>-0.0361***</td>
<td>-0.0215***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00544)</td>
<td>(0.00674)</td>
<td></td>
</tr>
<tr>
<td>Business customer</td>
<td></td>
<td></td>
<td>-2.255***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.419)</td>
</tr>
<tr>
<td>Product FEs</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Mean</td>
<td>3.805</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Observations</td>
<td>2,458</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standard errors in parentheses

* $p < .10$, ** $p < .05$, *** $p < .01$
Profits decomposition

Core
- Revenues: 190
- Net of wages: 170
- Net of input costs: 150
- Net of rent: 130
- Net of commuting: 110

Periphery
- Revenues: 160
- Net of wages: 140
- Net of input costs: 120
- Net of rent: 100
- Net of commuting: 80
Large buyers travel further and pay lower unit transport costs

- Unit transport costs decreasing in quantity, despite buyers purchasing larger quantities travelling further
- Suggests transport costs are fixed $\rightarrow$ *Economies of scale in transport* (Grant and Startz 2022)
Firms in the Core are more likely to outsource intermediate tasks

<table>
<thead>
<tr>
<th></th>
<th>Core</th>
<th>Periphery</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PANEL A: Outsourcing</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total number of workers</td>
<td>2.240</td>
<td>1.927</td>
<td>[0.000]</td>
</tr>
<tr>
<td>Any external worker</td>
<td>0.726</td>
<td>0.583</td>
<td>[0.000]</td>
</tr>
<tr>
<td>Share of external workers</td>
<td>0.418</td>
<td>0.324</td>
<td>[0.000]</td>
</tr>
</tbody>
</table>

**PANEL B: Distance from External Workers**

<table>
<thead>
<tr>
<th></th>
<th>Core</th>
<th>Periphery</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within 5 minutes walking</td>
<td>0.954</td>
<td>0.557</td>
<td>[0.000]</td>
</tr>
<tr>
<td>Between 5 and 15 minutes walking</td>
<td>0.040</td>
<td>0.188</td>
<td>[0.000]</td>
</tr>
<tr>
<td>More than 15 minutes walking</td>
<td>0.005</td>
<td>0.257</td>
<td>[0.000]</td>
</tr>
</tbody>
</table>

**Note:** Data is from the baseline survey of garment firms. P-values in Panels A and B are from regressions that control for product type Fixed Effects.
Rent per square-meter

[Graph showing the distribution of rent per square meter (USD) across different percent ranges for Periphery and Core areas.]
Expected Utility from a Location

- Consumers choose location to maximize their expected utility:

\[
V_{il}^q = E_\varepsilon \left[ \max_{j \in l} u_{ijl}^q \right] = q^\theta (1 - \sigma) \ln \left( \sum_{j=1}^{N_l} \exp \left( \frac{\delta_{jl}^q}{1 - \sigma} \right) \right) - C_{il} + \gamma
\]

where \( \delta_{jl} = \beta x_{jl} - \alpha p_{jl} q^{1-\theta} + \xi_j \)

- All else equal, expected utility is increasing in \( N_l \)
  - Observe more \( \varepsilon \) draws (more varieties)
  - Higher probability of finding product with desired characteristics
High-quality firms are more likely to sort in agglomerated locations

- For illustration, assume there are two types of firms: *low* and *high-quality*
  - Let \( N^L_l \) and \( N^H_l \) be the number of low and high-quality firms in location \( l \)
  - Let \( s_{iLl} \) and \( s_{iHl} \) be the share of type-\( i \) customers buying from a low/high quality firm in \( l \)

- The marginal effects of the entry of a high-quality firm in \( l \) on \( s_{iLl} \) and \( s_{iHl} \) are:

\[
\frac{\partial s_{iLl}}{\partial N^H_l} = s_{iLl} \left( s^H_{Hl} (q^\theta (1 - \sigma)(1 - s_{iI}) - 1) - \frac{\tau_2}{\sigma r_l} (1 - s_{iI}) \right)
\]

\[
\frac{\partial s_{iHl}}{\partial N^H_l} = s_{iHl} \left( s^H_{Hl} (q^\theta (1 - \sigma)(1 - s_{iI}) - 1) - \frac{\tau_2}{\sigma r_l} (1 - s_{iI}) \right)
\]

- In absolute terms, marginal effect is larger for high-quality firms as \( s_{iHl} > s_{iLl} \)

→ If marginal effect is positive, high-quality firms more likely to sort into large locations
Unconditional shares

The share of customer $i$ buying products from firm $j$ in location $l$ is:

$$s_{ijl}(L,p) = s_{il}^q(L,p) \times s_{jl}^q(p_l) =$$

$$\exp\left(\frac{\delta_{jl}^q}{1-\sigma}\right) \left(\sum_{j'=1}^{N_l} \exp\left(\frac{\delta_{j'l}^q}{1-\sigma}\right)\right)^{q^\theta (1-\sigma)-1} \exp(-\tau_1 g(||z_i - z_l||) - \tau_2 \frac{N_l}{a_l})$$

$$= \frac{\exp(u_0^q) + \sum_{k=1}^{N} \left(\sum_{h=1}^{N} \exp\left(\frac{\delta_{hk}^q}{1-\sigma}\right)\right)^{q^\theta (1-\sigma)} \exp(-\tau_1 g(||z_i - z_k||) - \tau_2 \frac{N_l}{a_l})}{\sum_{j'=1}^{N_l} \exp\left(\frac{\delta_{j'l}^q}{1-\sigma}\right)^{q^\theta (1-\sigma)-1} \exp(-\tau_1 g(||z_i - z_l||) - \tau_2 \frac{N_l}{a_l})}$$
The demand for product $j$ sold in location $l$ is:

$$Q_{jl}(p, J) = \int q_i s_{ijl}(p, J) dF(q_i, z_i)$$

where,

- $p$ and $J$ are vectors of prices and number of firms across locations
- $dF(\cdot)$ the joint distribution of quantities and distances
Production function and outsourcing

- Production function: $A_l h^{1-\delta} \ell^{\delta}$
- Firms can employ **internal** or **external** labor (outsourcing)
- Face the following trade-off:
  - Procuring external task requires firms to pay transaction cost $T(N_l), T'(\cdot) < 0$ (Holmes 1995)
  - As firms produce more tasks internally, productivity decreases (Eckel and Neary 2010)
- At optimal outsourcing, cost of labor is $w_l T(N_l)$
- Marginal cost decreasing in $N_l$:
  $$c_{jl} = \frac{1}{A_l} \left( \frac{w_l T(N_l)}{\delta} \right)^{\delta} \left( \frac{r_l}{1 - \delta} \right)^{1-\delta}$$
Firms produce output using labor and land.

Cobb-Douglas CRS production function:

\[ f_j(h, \ell) = A_j \ell^\delta h^{1-\delta} \]

Labor is a composite input produced by combining a continuum of perfectly complementary tasks \( t \):

\[ \ell = \min\{x(t) | t \in [0, 1]\} \]
Outsourcing technology

- Tasks can be produced internally or outsourced:
  - **External Technology**: requires $x(t)$ units of labor
  - **Internal Technology**: requires $a(Z)x(t)$ units of labor, with $z$ the share of internally produced tasks

- Firms face the following trade-off:
  - Procuring external task requires firms to pay transport cost $T(N_l)$, $T'() < 0$ (Holmes 1995)
  - As firms produce more tasks internally, productivity decreases: $a'(Z) > 0$ (Eckel and Neary 2010)
Assume that cost of internal and external labor ($w$) is the same and constant across locations.

Firms will choose $Z^*$ to equalize the cost of producing tasks internally and externally:

$$a(Z^*) = T(N_i)$$

**Proposition:** As $N_i$ increases $Z^*$ decreases, namely firms outsource a larger share of tasks.
Price choice

- Nash-Bertrand pricing game
- Firms choose price to maximize:

\[ \pi_{jl}(p, L) = (p_{jl} - c_{jl})Q_{jl}(p, L) \]

- Net effect of agglomeration on prices is ambiguous:
  1. Production externalities reduce marginal costs and hence prices
  2. Market-share effect increases competition, lowering prices
  3. Market-size effect reduces competition, pushing prices upwards
Optimal price

- Optimal price is implicitly given by:

\[
p_{jl}^* = c_{jl} + \frac{(1 - \sigma) \int q_i s_{ijl} dF(\cdot)}{\alpha \int q_i s_{ijl} \left[1 - s_{ijkl} (1 - q_i^\theta (1 - \sigma) (1 - s_{il}))\right] dF(\cdot)}
\]

- Can write the system of best-response equations as (Berry 1994):

\[
p = c - \Lambda(p)^{-1} Q(p)
\]

where \( \Lambda(p) \) is the \( J \times J \) matrix of price derivatives

- A **Nash-Bertrand equilibrium** is a vector \( p^* \) that solves this fixed point mapping

- Equilibrium might not be unique
Equilibrium of location game

- Firms move simultaneously, forming expectations on behavior of other firms.
- Firm $j$’s probability of choosing location $l$ is:

$$\Psi_j(l|P) = \frac{\exp\left( \left( \sum_{l \neq j} \left[ \pi_{jl}(l, l) \prod_{h \neq j} P_h(l_h) \right] - \tau_3 g(||z_j - z_l||) \right) / \mu \right)}{1 + \sum_{k=1}^N \exp\left( \left( \sum_{k \neq j} \left[ \pi_{jk}(k, k) \prod_{h \neq j} P_h(k_h) \right] - \tau_3 g(||z_j - z_k||) \right) / \mu \right)}$$

- **Bayesian Nash Equilibrium** is the fixed point of the system of best response mappings.
  - Equilibrium exists, but might not be unique.
Demand Estimation

- Taking firm location and prices as given, maximize log-likelihood function:

\[ \ln L(\theta_1 | L, p) = \sum_{i,q,j,l} w_j \times I_{ijl}^q \times \ln s_{ijl}^q(L,p) \]

- \( s_{ijl}^q \): probability that type-\( q \) consumer from location \( i \) purchases products from firm \( j \) in \( l \)
- \( I_{ijl}^q \): indicator for whether consumer purchases a product from firm \( j \) in the data
- \( w_j \): sampling weights

- Assume \( \xi_j = 0 \), no price endogeneity

  - Use **price from mystery shoppers** exercise: same product, same bargaining
  - Strongly correlated with transaction prices and product quality
  - Uncorrelated with a number of variables that could be related to unobserved quality
Correlation between transaction and mystery shopper prices

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Transaction price</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mystery shoppers price</td>
<td>0.925***</td>
<td>0.808***</td>
<td>1.077***</td>
</tr>
<tr>
<td></td>
<td>(0.122)</td>
<td>(0.100)</td>
<td>(0.065)</td>
</tr>
<tr>
<td>Quality score</td>
<td></td>
<td></td>
<td>0.461***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.111)</td>
</tr>
<tr>
<td>Product FEs</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>2,571</td>
<td>2,571</td>
<td>2,541</td>
</tr>
</tbody>
</table>

Notes: Robust standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$
## Correlates of Mystery Shoppers prices

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price (USD)</td>
<td></td>
</tr>
<tr>
<td>Quality score (0-100 rating)</td>
<td>0.241***</td>
</tr>
<tr>
<td></td>
<td>(0.0761)</td>
</tr>
<tr>
<td>Customer care (0-10 rating)</td>
<td>0.0533</td>
</tr>
<tr>
<td></td>
<td>(0.0678)</td>
</tr>
<tr>
<td>Greeted upon entering the firm</td>
<td>-0.249</td>
</tr>
<tr>
<td></td>
<td>(0.460)</td>
</tr>
<tr>
<td>Given undivided attention</td>
<td>0.268</td>
</tr>
<tr>
<td></td>
<td>(0.354)</td>
</tr>
<tr>
<td>Pleasant closing comment</td>
<td>-0.451</td>
</tr>
<tr>
<td></td>
<td>(0.395)</td>
</tr>
<tr>
<td>Tidiness of premises (0-10 rating)</td>
<td>-0.0475</td>
</tr>
<tr>
<td></td>
<td>(0.0656)</td>
</tr>
<tr>
<td>Cleanliness of premises (0-10 rating)</td>
<td>0.164**</td>
</tr>
<tr>
<td></td>
<td>(0.0741)</td>
</tr>
<tr>
<td>Product ready by delivery date</td>
<td>-0.202</td>
</tr>
<tr>
<td></td>
<td>(0.199)</td>
</tr>
<tr>
<td>Offered something to come back</td>
<td>0.543</td>
</tr>
<tr>
<td></td>
<td>(0.515)</td>
</tr>
<tr>
<td>Told to advertise firm</td>
<td>0.147</td>
</tr>
<tr>
<td></td>
<td>(0.279)</td>
</tr>
</tbody>
</table>

Shopper FEs ✓
Parish FEs ✓
Number of Obs 529

Standard errors clustered at the parish level in parentheses. * p < .10, ** p < .05, *** p < .01
Demand Estimation Robustness

- Given estimates of \(\{\sigma, \theta, \tau_1, \tau_2\}\), can solve for \(\delta^q_{jl}\) from (Berry 1994):

\[
s^q_{jl}(p_l) = \frac{\exp\left(\frac{\delta^q_{jl}}{1-\sigma}\right)}{\sum_{h=1}^{N_l} \exp\left(\frac{\delta^q_{hl}}{1-\sigma}\right)}
\]

- Mean utilities defined as \(\delta^q_{jl} = \beta x_j + \alpha p_{jl} q^{1-\theta} + \xi_j\)

- Regress estimated \(\delta^q_{jl}\) on prices and observable characteristic to obtain \(\{\alpha, \beta\}\)

- Need IV for prices uncorrelated with unobserved shock \(\xi_j\)
  - **Cost shifter**: cost of material used in production of typical product
  - **BLP instrument**: share of high-quality firms in same location
### Demand Estimation Robustness

<table>
<thead>
<tr>
<th></th>
<th>(1) Transaction price</th>
<th>(2) Delta</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: First Stage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost of cloth (1 meter)</td>
<td>0.647***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.192)</td>
<td></td>
</tr>
<tr>
<td>Share high-quality firms</td>
<td>8.383***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.450)</td>
<td></td>
</tr>
<tr>
<td><strong>Panel B: Second Stage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transaction price</td>
<td></td>
<td>-0.092**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.038)</td>
</tr>
<tr>
<td>Product FEs</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>608</td>
<td>608</td>
</tr>
</tbody>
</table>

**Notes:** Robust standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$
Customers’ origin

- Missing consumer origin for 34% of transactions
- Uncorrelated with transaction and firm characteristics
- Imputed from origin of consumers of the same type, shopping in the same location via proportional random assignment

<table>
<thead>
<tr>
<th></th>
<th>(1) Missing Origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transaction value</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
</tr>
<tr>
<td>Number of daily customers</td>
<td>-0.040</td>
</tr>
<tr>
<td></td>
<td>(0.039)</td>
</tr>
<tr>
<td>Daily revenues</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
</tr>
<tr>
<td>Retail customer</td>
<td>-0.075*</td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
</tr>
<tr>
<td>Parish FEs</td>
<td>yes</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>2589</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
* $p < .10$, ** $p < .05$, *** $p < .01$
Supply Estimation

- Use optimal choice of labor \((\ell^*)\) and land \((h^*)\)
- Take input costs from data:
  - Wages assumed to be constant in periphery, but allowed to be different in Core
  - Rent exogenously varies across location, but not explicitly modelled
- Parametrization: \(T(N_l) = 1 + N_l^T\) (iceberg transport cost)
- Estimate using Simulated Method of Moments
  - **Parameters** to be estimated: \(\{\delta, T(N_l), A_l\}\)
  - **Targeted moments**: average ratio of land/labor, and variance of labor and land in each parish
Identification of Supply Parameters

\[ h_{jl}^* = \frac{Q_{jl}}{A_l} \left( \frac{(1 - \delta)wT(N_l)e_j}{\delta r_l} \right) \delta \]

\[ \ell_{jl}^* = \frac{Q_{jl}}{A_l} \left( \frac{\delta r_l}{(1 - \delta)wT(N_l)e_j} \right)^{1-\delta} \]

- Ratio across locations \( l, h \):
  \[ \frac{\ell_{jl}^*/h_{jl}^*}{\ell_{hk}^*/h_{hk}^*} = \frac{r_l w_k T(N_k)}{r_k w_l T(N_l)} \]

- Expected within location \( l \):
  \[ \frac{h_{jl}^*}{\ell_{jl}^*} = \frac{(1-\delta)wT_1(J_l)}{\delta r_l} \]

→ Ratios identify \( T_1(J_l) \) and \( \delta \)

- Given \( Q_{jl} \), recover location productivity \( A_l \) from levels \( h_{jl}^* \) and \( \ell_{jl}^* \)
Location Estimation

- A fixed point of Rust’s NFXP algorithm is a pair \( \{ \tau^*_3, P^* \} \) that satisfies:
  
  (i) \( \tau^*_3 = \arg \max_{\tau_3} \sum_j \sum_l \ln \Psi_j(l|P^*, \tau_3) l_j \)
  
  (ii) \( P^* = \Psi(P^*, \tau^*_3) \)

- Where the best response mapping is:

  \[
  \Psi_j(l|P) = \frac{\exp \left( \left( \sum_{l \neq j} \pi_{jl}(l, l_{-j}) \prod_{h \neq j} P_h(l_h) \right) - \tau_3 g(||z_j - z_l||) / \mu \right)}{1 + \sum_{k=1}^N \exp \left( \left( \sum_{k \neq j} \pi_{jk}(k, k_{-j}) \prod_{h \neq j} P_h(k_h) \right) - \tau_3 g(||z_j - z_k||) / \mu \right)}
  \]

- Must compute Nash-Bertrand equilibrium and variable profits \( \pi_{jl} \) for all possible configuration of firms in space \( l_{-j} \)

- \( \mu \) not separately identified: calibrated to 0.75 of a standard deviation of profits
Location Estimation: Reducing the state space

1. **Choice set**: stay in parish where the owner resides or move to Core
2. **Firm heterogeneity**: high or low-quality
   - But type also a function of owner’s parish: order of magnitude of state space is $10^{243}$
3. **Information**: know number of firms of each type $\{N_{low}, N_{high}\}$, but not parish of origin
   - Assume owners are uniformly distributed among periphery parishes
   - Limited knowledge: *representative parish* in the periphery
Location Estimation: Simplified best response mapping

- Let $n_{\text{low}}$ and $n_{\text{high}}$ be the number of low and high-quality firms (other than $j$) in the core.

- Let $\mathbf{P} = \{P_{\text{low}}, P_{\text{high}}\}$ be the vector of CCPs of any low and high-quality firm entering the core.

\[
\Psi_j(l|\mathbf{P}) = \frac{\exp \left( (\sum_{n_{\text{low}}, n_{\text{high}}}(\pi_{jl}(l, n_{\text{low}}, n_{\text{high}})Pr(n_{\text{low}}, n_{\text{high}})) - \tau_3 g(||z_j - z_l||))/\mu \right)}{1 + \sum_{k=1}^{N} \exp \left( (\sum_{k, k \neq j}(\pi_{jk}(k, n_{\text{low}}, n_{\text{high}})Pr(n_{\text{low}}, n_{\text{high}})) - \tau_3 g(||z_j - z_k||))/\mu \right)}
\]

- The probability that $n_{\text{low}}$ and $n_{\text{high}}$ low and high-quality firms enter the core is:

\[
Pr(n_{\text{low}}, n_{\text{high}}) = \binom{N_{\text{low}}}{n_{\text{low}}} (P_{\text{low}})^{n_{\text{low}}} (1 - P_{\text{low}})^{N_{\text{low}} - n_{\text{low}}} \times \binom{N_{\text{high}}}{n_{\text{high}}} (P_{\text{high}})^{n_{\text{high}}} (1 - P_{\text{high}})^{N_{\text{high}} - n_{\text{high}}}
\]
### Estimated Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std Error</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PANEL A: Demand</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price (USD)</td>
<td>$\alpha$</td>
<td>-0.064</td>
</tr>
<tr>
<td>Quality final customers</td>
<td>$\beta_f$</td>
<td>0.205</td>
</tr>
<tr>
<td>Quality retail customers</td>
<td>$\beta_r$</td>
<td>0.724</td>
</tr>
<tr>
<td>Taste shocks correlation</td>
<td>$\sigma$</td>
<td>0.329</td>
</tr>
<tr>
<td>Quantity multiplier</td>
<td>$\theta$</td>
<td>0.316</td>
</tr>
<tr>
<td>Travel cost</td>
<td>$\tau_1$</td>
<td>-0.139</td>
</tr>
<tr>
<td>Within location search cost</td>
<td>$\tau_2$</td>
<td>-0.0004</td>
</tr>
<tr>
<td><strong>PANEL B: Supply</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labor share</td>
<td>$\delta$</td>
<td>0.651</td>
</tr>
<tr>
<td>Outsourcing cost</td>
<td>$T$</td>
<td>-0.521</td>
</tr>
<tr>
<td>Productivity Core</td>
<td>$A_{core}$</td>
<td>17.950</td>
</tr>
<tr>
<td>Productivity Periphery (mean)</td>
<td>$A_{per}$</td>
<td>10.886</td>
</tr>
<tr>
<td><strong>PANEL C: Location</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commuting cost</td>
<td>$\tau_3$</td>
<td>-5.739</td>
</tr>
</tbody>
</table>

**Note:** Standard errors bootstrapped using 100 bootstrapped samples.
Model Fit: Parish Demand Shares

![Graph showing data and estimate for share of customers across different parishes.](image)

- **Kibuye II**
- **Katwe I**
- **Bwaise II**
- **Kisugu**
- **Nakivubo-Shauriyako**
- **Kamwokya II**
- **Wandegeya**
- **Kisenyi III**
- **Naguru I**
- **Mbuya I**
- **Kasubi**
- **Nakasero IV**

**X-axis:** Parish

**Y-axis:** Share of customers

**Legend:**
- Blue line: Data
- Red line: Estimate
Model Fit: Within Location Firm Shares

![Graph showing model fit with data and estimate lines]

- Share of customers | parish data estimate
- X-axis: Share of customers
- Y-axis: Parish

Appendix
## Model Fit: Choice of Land and Labor

<table>
<thead>
<tr>
<th>Parish</th>
<th>Data</th>
<th>Sim</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bwaise II</td>
<td>6.050</td>
<td>5.890</td>
</tr>
<tr>
<td>Kamwokya II</td>
<td>5.450</td>
<td>5.466</td>
</tr>
<tr>
<td>Kasubi</td>
<td>5.003</td>
<td>4.736</td>
</tr>
<tr>
<td>Katwe I</td>
<td>1.750</td>
<td>1.989</td>
</tr>
<tr>
<td>Kibuye II</td>
<td>2.857</td>
<td>2.671</td>
</tr>
<tr>
<td>Kisenyi III</td>
<td>3.450</td>
<td>3.823</td>
</tr>
<tr>
<td>Kisugu</td>
<td>7.750</td>
<td>7.347</td>
</tr>
<tr>
<td>Mbuya I</td>
<td>9.394</td>
<td>9.194</td>
</tr>
<tr>
<td>Naguru I</td>
<td>3.862</td>
<td>4.192</td>
</tr>
<tr>
<td>Core</td>
<td>2.671</td>
<td>2.808</td>
</tr>
<tr>
<td>Nakivubo-Shauriyako</td>
<td>4.533</td>
<td>3.282</td>
</tr>
<tr>
<td>Wandegeya</td>
<td>2.478</td>
<td>2.217</td>
</tr>
</tbody>
</table>

**Notes:** Robust standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$
### Counterfactual 2: E-commerce platform

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>E-Commerce</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share of firms in core</td>
<td>0.365</td>
<td>0.222</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-39%</td>
</tr>
<tr>
<td>Share of sales in core</td>
<td>0.382</td>
<td>0.065</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-83%</td>
</tr>
<tr>
<td>Average price</td>
<td>20.44</td>
<td>17.38</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-15%</td>
</tr>
<tr>
<td>Average profits</td>
<td>476.0</td>
<td>411.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-14%</td>
</tr>
<tr>
<td>Average consumer welfare</td>
<td>19.22</td>
<td>31.80</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+71%</td>
</tr>
</tbody>
</table>
Counterfactual 3: Capping the number of firms in the core

Panel A: Firm Profits

Panel B: Consumer Welfare
Banning boda-bodas from the core

- Travel time separately for cars and motorcycles
  - No ban: min\{car, motorcycle\}; Ban: car
  - Applies to part of the journey within central district
- Increase in travel for customers (5.2%) and firm owners (6.8%)
- 9.8% of firms relocate in the periphery
- Small increase in average profits, but:
  - −3.6% in core, +3.3% in periphery