#### Customers and Retail Growth

Liran Einav, Stanford and NBER Pete Klenow, Stanford and NBER Jonathan Levin, Stanford and NBER Raviv Murciano-Goroff, Boston University

> Payment System Studies Seminar Federal Reserve Board

> > September 2022

- How do firms grow?
  - Customers?
  - Sales per customer?
- Does customer acquisition affect productivity growth?
  - Does it stimulate innovation as a complement?
  - Or just marketing effort?

 Quantify the role of customers in the sales growth of retail merchants and stores using Visa, Inc. data in the U.S. from 2017 to 2021

Trace *aggregate* retail sales changes to merchants with big increases and big decreases in their numbers of customers

Model retailer growth through innovation and customer acquisition to see how the customer margin might influence aggregate productivity growth

## **Related** papers

#### Models

- Fishman and Rob (2003)
- Luttmer (2006)
- Arkolakis (2010, 2016)
- Perla (2019)
- Evidence
  - ► Foster, Haltiwanger and Syverson (2008, 2016)
  - ▶ Hottman, Redding and Weinstein (2016)
  - Baker, Baugh and Sammon (2021)
  - ► Fitzgerald, Haller and Yedid-Levi (2022)

- Both models and evidence
  - Gourio and Rudanko (2014)
  - Eaton, Eslava, Jinkins, Krizan, and Tybout (2021)
  - Bornstein (2021)
  - Afrouzi, Drenik and Kim (2021)
  - Argente, Fitzgerald, Priolo and Moreira (2021)
  - Bernard, Dhyne, Magerman, Manova and Moxnes (2022)

Our distinct focus: the customers of retailers

#### Visa data

- Transaction amount and day
- Unique card identifiers (credit and debit)
- Unique merchant identifiers (retailers with one or more stores)
- Store industry (3-digit NAICS)
- Store location (address)
- May 2017 through December 2021
- No detail on items bought or prices paid
- Cannot tie multiple cards to a single household, except for a subset of credit cards

All results have been reviewed to ensure that no confidential information about Visa merchants or cardholders is disclosed.

Cards are anonymized, and we report no data on individual cards. Cardholder information is based solely on the card's transactions.

We report no data on specific merchants or recent months (nothing from 2020 onward).

- U.S. annual averages from 2017 through 2021
  - 379 million cards
  - 29 billion transactions
  - 24% of all consumption
  - 60% credit, 40% debit

#### • All NAICS

• Retail (+ restaurant) NAICS

• Offline retail (card-present transactions) – our baseline sample

• A much smaller set of large merchants back to 2010 (instead of just 2017–2021)

$$Sales \equiv Cards \cdot \frac{Transactions}{Cards} \cdot \frac{Sales}{Transactions}$$

For merchants, we can further decompose cards:

$$Cards \equiv Stores \cdot \frac{Cards}{Stores}$$

We take logs and regress each RHS variable on the LHS (log of Sales).

Coefficients decompose sales:

- Across merchants in 2019 (with NAICS fixed effects)
- Across stores within merchants in 2019 (with merchant fixed effects)
- Over time within stores/merchants 2017–2021 (store/merchant and year fixed effects)

#### Decomposing sales across merchants in 2019

Each entry is from a single univariate log-log regression on Sales

| Dep. var. $\rightarrow$ | Cards | Trans/Cards | Sales/Trans | # obs. |
|-------------------------|-------|-------------|-------------|--------|
| All NAICS               | 0.727 | 0.037       | 0.236       | 2.24m  |
| Online                  | 0.646 | 0.049       | 0.305       | 501k   |
| Offline                 | 0.749 | 0.031       | 0.220       | 1.82m  |
| <b>Offline retail</b>   | 0.786 | 0.038       | 0.176       | 928k   |

The coefficients in each row add up to 1 by construction; standard errors are tiny.

| Dependent variable $\rightarrow$    | Cards | Trans/Cards | Sales/Trans | # obs. |
|-------------------------------------|-------|-------------|-------------|--------|
| Across merchants in 2019            | 0.786 | 0.038       | 0.176       | 928k   |
| Within merchants 2017–2021          | 0.845 | 0.103       | 0.053       | 4.51m  |
| Across stores within merchants 2019 | 0.838 | 0.091       | 0.071       | 1.96m  |
| Within stores 2017–2021             | 0.816 | 0.138       | 0.046       | 9.58m  |

• Experian credit bureau data merged in by Visa

• Can link Visa credit cards across these households

• We run the same regressions using households rather than cards

• Household margin is  $\sim 1$  percentage point lower than the card margin

#### Decomposing 2019 merchant sales growth, by NAICS



Cards Trans/Cards Dollar/Trans

- For each variable, subtract any fixed effects in logs
- Sort observations (merchant-years or store-years) into 20 groups based on their sales
- For each variable, compute its average within each ventile
- Exponentiate the average in each ventile, and normalize the lowest ventile to 1

#### Across merchants in 2019



Sales

#### Within merchants over time, 2017–2021



#### Stores vs. cards per store across merchants in 2019



#### Stores vs. cards per store over time within merchants, 2017–2021



## Across stores within merchants in 2019



#### Within stores over time, 2017–2021



Merchant contributions to aggregate sales changes, 2017–2021

# All Retail Sectors



## Merchant contributions to sales changes over 2017–2021, by NAICS



## Customers vs. sales/customer and merchant sales changes, 2018–2021

Offline Retail, 2018-2021



### % of sales from returning customers, 2018–2021



## A growth model fit to the Visa customer margin

- The number of customers drive Visa merchant growth
- Rapid-growers and shrinkers dominate aggregate Visa sales changes

Motivates a model with these features:

- **()** Heterogeneous firm innovation
- In the number of customers responds strongly to innovation

Customer dynamics:

- 40% annual turnover of cards for individual merchants (weighted by card spending)
- More turnover in durables NAICS, less in nondurables and services NAICS

Productivity growth in retail:

- 3.0% annual growth in labor productivity from 1988-2020 (BLS)
- Foster, Haltiwanger and Krizan (2006) identify it as a key driver of 1990s growth

#### **Consumer preferences**

A unit mass of customers have the following preferences over consumption:

$$U = \sum_{t=0}^{\infty} \beta^t \, \frac{C_t^{1-1/\sigma}}{1-1/\sigma}$$

Consumption is an aggregate of spending on retailers:

$$C_t = \left(\int_0^1 n_{it} \left(q_{it}c_{it}\right)^{\frac{\theta-1}{\theta}} \mathrm{d}i\right)^{\frac{\theta}{\theta-1}}$$

Note the fixed unit mass of retailers  $i \in [0, 1]$ , with retailer *i* having quality  $q_{it}$ .  $n_{it} \in [0, 1]$  is the probability that a consumer has the option to purchase from retailer *i*.  $c_{it}$  is the quantity of retailer *i* purchased by each consumer with access to *i*.

#### Customer demand and the aggregate price index

Per customer spending for those customers with access to retailer *i*:

$$c_{it} = \left(\frac{p_{it}}{P_t}\right)^{-\theta} q_{it}^{\theta-1} C_t$$

This yields the ideal price index:

$$P_t \equiv \left(\int_0^1 n_{it} \left(\frac{p_{it}}{q_{it}}\right)^{1-\theta} \mathrm{d}i\right)^{\frac{1}{1-\theta}}$$

Total demand facing retailer i, summed across its customers:

$$y_{it} = n_{it} \cdot c_{it}$$

## Production, customer acquisition, and static profit maximization

Each retailer uses production labor  $l_{it}$  to produce its output:

$$y_{it} = l_{it}$$

It uses marketing labor  $m_{it}$  to reach fraction  $n_{it}$  of customers:

$$n_{it} = \left(\frac{\gamma}{\phi} \cdot m_{it}\right)^{1/\gamma}$$
 where  $\gamma > 1$ 

Normalizing the nominal wage to 1, the retailer's static profit maximization problem is:

$$\max_{p_{it},m_{it}} \left( p_{it} - 1 \right) y_{it} - m_{it}$$

## Retailer pricing and aggregate quality

Through monopolistic competition, retailers choose pricing:

$$p_{it} = \frac{\theta}{\theta - 1} \equiv \mu$$

Aggregate quality:

$$Q_t \equiv \left(\int_0^1 q_{it}^{\Gamma(\theta-1)} \mathrm{d}i\right)^{\frac{1}{\Gamma(\theta-1)}} \quad \text{where} \quad \Gamma \equiv \frac{\gamma}{\gamma-1}$$

## A retailer's customers and variable profits

$$n_{it} = \left(\frac{\gamma z_{it}^{\Gamma(\theta-1)}}{\phi}\right)^{\frac{1}{\gamma}}$$

and

$$\pi_{it} = \frac{z_{it}^{\Gamma(\theta-1)}}{\Gamma(\theta-1)} \cdot L_t$$

where  $z_{it} \equiv \frac{q_{it}}{Q_t}$ 

A retailer with absolute quality  $q_{it}$  and relative quality  $z_{it}$  that hires research labor  $s_{it}$  sees its quality follow a controlled binomial process with probability  $x_{it} \in [0, 1]$ :

$$q_{it+1} = \begin{cases} q_{it} \ e^{\Delta} & \text{w/ prob. } x_{it} \\ & \text{and} & s_{it} = \lambda \cdot \log\left(\frac{1}{1 - x_{it}}\right) \cdot z_{it}^{\zeta} \\ q_{it} & \text{w/ prob. } 1 - x_{it} \end{cases}$$

 $\Delta$ ,  $\lambda$  and  $\zeta$  are all strictly positive parameters

A retailer's value function is given by:

$$V_{t}(z) = \pi_{t}(z) + \max_{x \in [0,1]} \left\{ R_{t}^{-1} \left[ x V_{t+1} \left( z e^{\Delta - g_{t}} \right) + (1-x) V_{t+1} \left( z e^{-g_{t}} \right) \right] - s(z,x) \right\}$$

The growth rate of this economy is:

$$1 + g_t = \left(1 + \int x(z) z^{\Gamma(\theta-1)} \left(e^{\Delta \Gamma(\theta-1)} - 1\right) \mathrm{d}F_t(z)\right)^{\frac{1}{\Gamma(\theta-1)}}$$

In each period, a retailer makes the following ordered decisions:

- 1. Hire marketing labor m(z) to access customers n(z)
- 2. Hire production labor l(z) to sell to their customers
- 3. Hire research labor to achieve a probability of research success  $x(z) \in [0,1]$

# Labor market clearing

$$L_{t} = \int l_{t}(z) dF_{t}(z)$$
$$M_{t} = \int m_{t}(z) dF_{t}(z)$$
$$S_{t} = \int s_{t}(z) dF_{t}(z)$$

$$L_t + M_t + S_t = 1$$

# Aggregates in equilibrium

$$L_t = \frac{\gamma \left(\theta - 1\right) \left(1 - S_t\right)}{\gamma \left(\theta - 1\right) + 1}$$

$$M_t = \frac{1 - S_t}{\gamma \left(\theta - 1\right) + 1}$$

$$C_t = (\gamma/\phi)^{\frac{1}{\gamma(\theta-1)}} \cdot L_t \cdot Q_t$$

\_

| Symbol    | Parameter                 | Value                | Target                           |
|-----------|---------------------------|----------------------|----------------------------------|
| σ         | IES                       | 0.5                  | Hall (2009)                      |
| heta      | CES between retailers     | 3                    | Aghion et al. (2019)             |
| eta       | Discount factor           | 0.992                | Farhi and Gourio (2018)          |
| $\phi$    | Scale of marketing costs  | $1.83 \cdot 10^{35}$ | Largest retailers have $n = 0.5$ |
| $\gamma$  | Marketing cost elasticity | 1.25                 | Visa sales decomposition         |
| $\Delta$  | Quality step size         | 0.063                | Average $x$ of 0.5               |
| $\lambda$ | Linear research cost      | 0.094                | 3.05% BLS growth rate            |
| $\zeta$   | Convex research cost      | 10.04                | Top 1% contribution in Visa      |

The elasticity of sales with respect to quality is the sum of the elasticity of customers and the elasticity of spending per customer with respect to quality:

$$\xi_{y,q} = \xi_{n,q} + \xi_{c,q} = \frac{\theta - 1}{\gamma - 1} + \theta - 1$$

With our calibration ( $\gamma = 1.25$  and  $\theta = 3$ ), the customer share of the sales elasticity is fixed at 80%, which matches our finding in the Visa data.

## A comparison economy with a marketing tax

Instead of paying  $w \cdot m$ , retailers pay a progressive marketing tax so that their gross-of-tax marketing costs are:

$$w \cdot \tau \cdot m^{1+\omega}$$

Here  $\tau > 0$  controls the average tax and  $\omega > 0$  the progressivity of the tax.

We choose  $\omega$  and  $\tau$  so that the customer margin is 20% rather than 80% and yet total marketing labor in the economy is the same as in the baseline economy.

#### Customers and retailer quality



#### Customers and retailer value



#### Customers and retailer innovation



#### Customers and cumulative innovation effort



## The distribution of retailer quality



## The distribution of retailer sales



## Firm contributions to aggregate sales changes (in the model)



| Symbol | Variable         | Baseline | Marketing tax |
|--------|------------------|----------|---------------|
| g      | Growth rate      | 3.05%    | 2.79%         |
| r      | Interest rate    | 7.01%    | 6.47%         |
| L      | Production labor | 67.9%    | 67.9%         |
| M      | Marketing labor  | 27.2%    | 27.2%         |
| S      | Research labor   | 4.92%    | 4.97%         |

|                          | Baseline | Marketing tax |
|--------------------------|----------|---------------|
| True growth rate         | 3.05%    | 2.79%         |
| Approximated growth rate | 3.10%    | 2.84%         |
| 1st order term           | 2.60%    | 2.72%         |
| 2nd order term           | 0.50%    | 0.12%         |

• We looked at Visa debit and credit card transactions data at all offline retail merchants from 2017–2021

• We documented a dominant role for the customer extensive margin in the dispersion of sales and sales growth across merchants

• In a simple growth model, the customer margin stimulates innovation and marketing by large retailers, and overall growth in the process

## Potential policy implications?

• Knowledge spillovers from research (a force for too little research)

• Business stealing (a force for too much research and marketing)

• No negative marketing externality (otherwise a force for too much marketing)

• Would need to analyze transition dynamics from R&D subsidy and/or marketing taxes

## Why is the distribution of (relative) retailer quality stationary?

• Retailers never enter or exit (no fixed costs)

• Research gets harder and harder for high quality retailers

• Research is comparatively easy for low quality retailers

• Thus mean reversion in quality offsets random realizations of quality steps