Customers and Retail Growth

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

September 8, 2020

Wharton Macro seminar

What we do

- Quantify the role of customer growth in the growth of retail merchants and stores using Visa data in the U.S. from 2016–2019
- Trace *aggregate* retail sales changes to merchants with rapidly-growing vs. rapidly-shrinking numbers of customers
- Model firm growth through innovation and customer acquisition to see how the customer margin affects aggregate growth

Some related papers

- Models
  - Arkolakis (2010, 2016)
  - Perla (2019)
- Evidence
  - Foster, Haltiwanger and Syverson (2008, 2016)
  - Hottman, Redding and Weinstein (2016)
  - Fitzgerald and Priolo (2018), Fitzgerald, Haller and Yedid-Levi (2019)
  - Baker, Baught and Sammon (2020)
- Both
  - Gourio and Rudanko (2014)
  - Eslava, Tybout, Jinkins, Krizan and Eaton (2015)
  - Bornstein (2018)
  - Bernard, Dhyne, Magerman, Manova and Moxnes (2019)

Visa data

- Transaction amount and day
- Unique card identifiers (credit and debit)
- Unique merchant identifiers (firms with one or more stores)
- Store NAICS
- Store address
- January 2016 through December 2019
- No detail on items bought or prices paid
- Cannot tie multiple card numbers to single households

Visa data confidentiality

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 from very recent months.

Visa summary statistics

U.S. annual averages from 2007 through 2019

- 430 million accounts
- 32 billion transactions
- 24% of all consumption
- 60% credit, 40% debit

Visa sub-samples

- All NAICS
- Retail (+ restaurant) NAICS
- **Offline retail** (card present transactions)
- “Named” merchants back to 2007

Decomposing sales

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

For merchants, we can further decompose accounts:

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

Sales regressions

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

Coefficients decompose sales:

- Across merchants in 2019 (NAICS fixed effects)
- Across stores within merchants in 2019 (merchant fixed effects)
- Over time within stores/merchants 2016–2019 (store/merchant and year fixed effects)

Decomposing sales across merchants in 2019

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

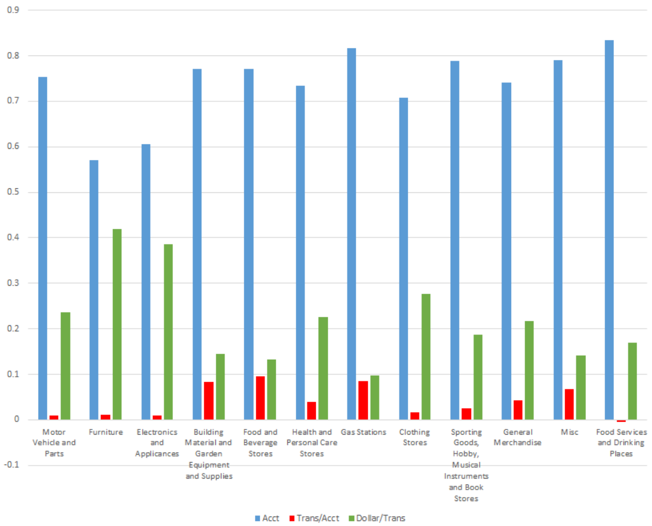
Dep. var. →	Accounts	Trans/Account	Sales/Trans	# obs.
All NAICS	0.743	0.037	0.221	2.18m
Online	0.673	0.073	0.254	606k
Offline	0.813	0.031	0.159	1.79m
Offline retail	0.812	0.035	0.166	954k

The coefficients add up to 1 by construction.

Decomposing sales in offline retail

Dependent variable →	Accounts	Trans/Account	Sales/Trans	# obs.
Across merchants in 2019	0.812	0.035	0.166	954k
Within merchants 2016–2019	0.844	0.101	0.055	3.91m
Across stores within merchants 2019	0.843	0.079	0.078	2.02m
Within stores 2016–2019	0.818	0.137	0.045	8.19m

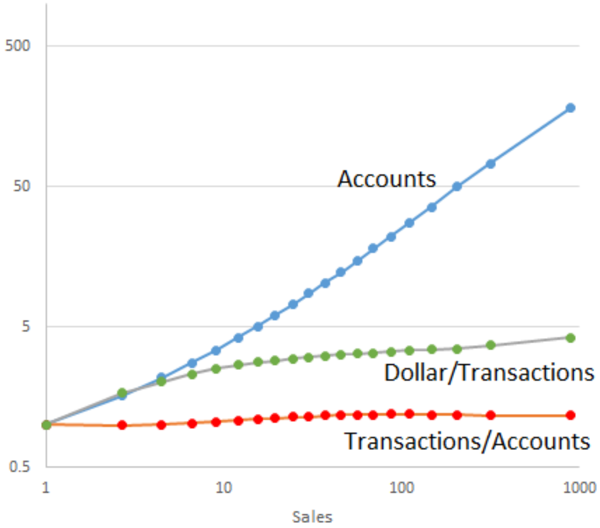
Decomposing 2019 merchant sales growth, by NAICS



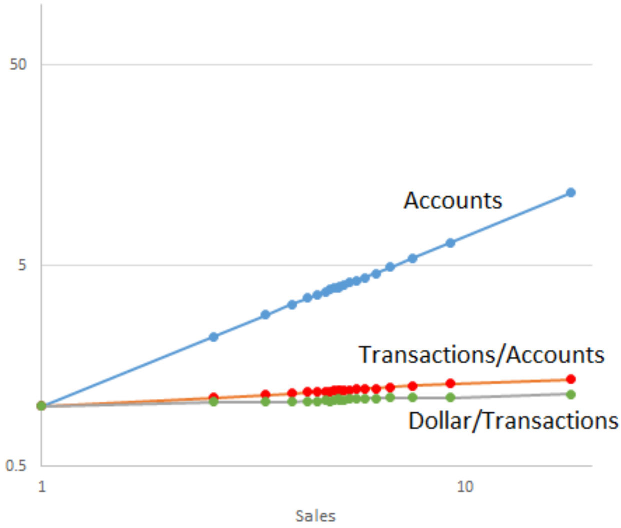
Ventile Figures

- 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



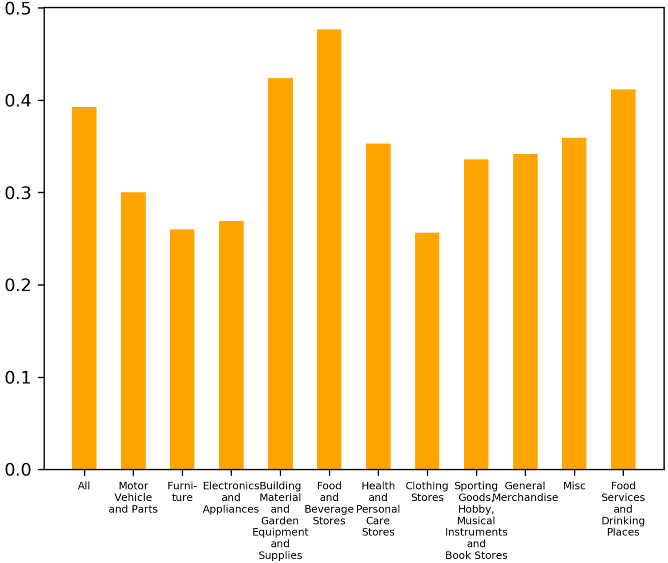
Within merchants over time, 2016–2019



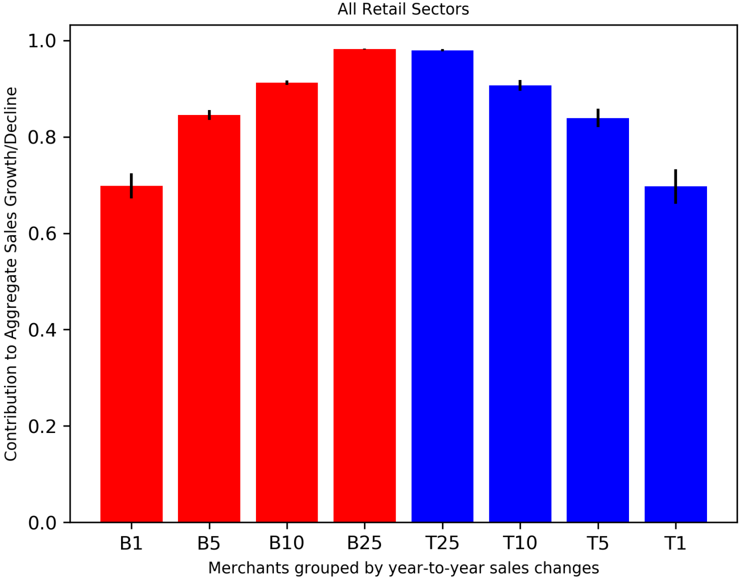
Stores vs. accounts per store across merchants in 2019

Stores vs. accounts per store over time within merchants, 2016–2019

Growth in spending per *returning* customer, merchants 2016–2019



Merchant contributions to aggregate sales changes, 2016–2019



The importance of customers in the tails

A merchant's sales changes can be decomposed as:

$$\Delta S_{it} \equiv \Delta N_{it} \cdot \overline{S/N}_{it} + \Delta (S/N)_{it} \cdot \overline{N}_{it}$$

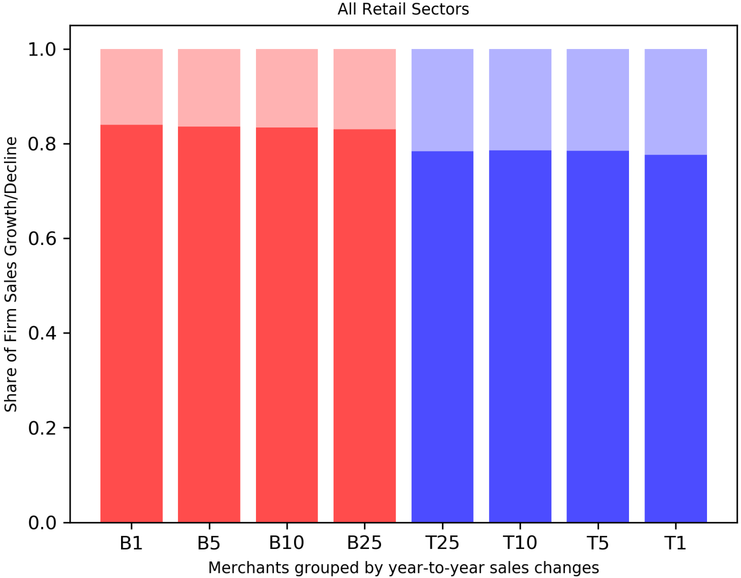
where

$$\overline{N}_{it} \equiv \frac{N_{it} + N_{i,t-1}}{2}$$

and

$$\overline{S/N}_{it} \equiv \frac{S_{it}/N_{it} + S_{i,t-1}/N_{i,t-1}}{2}$$

Customers vs. sales/customer and firm sales changes, 2016–2019



Visa patterns and a model

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

Motivates a model with these features:

1. Heterogeneous firm innovation
2. The number of customers respond strongly to innovation

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 varieties:

$$C_t = \left(\int_0^1 n_{jt} \left(q_{jt} c_{jt} \right)^{\frac{\theta-1}{\theta}} dj \right)^{\frac{\theta}{\theta-1}} .$$

Note the fixed unit mass of varieties $j \in [0, 1]$, with variety j having quality q_{jt}

$n_{jt} \in [0, 1]$ is the probability that a consumer has the option to purchase variety j

c_{jt} is the quantity of variety j purchased by consumers who have access to j

Customer demand and the aggregate price index

Per customer demand conditional on access to variety j:

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

Implies the ideal price index

$$P_t \equiv \left(\int_0^1 n_{jt} \left(\frac{p_{jt}}{q_{jt}}\right)^{1-\theta} dj\right)^{\frac{1}{1-\theta}} .$$

Total demand facing firm j, summed across its customers:

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

Production, customer acquisition and flow profits

Each firm uses production labor l_{jt} to produce its single variety:

$$y_{jt} = l_{jt}$$

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

$$n_{jt} = \left(\frac{\gamma}{\phi} \cdot \frac{m_{jt}}{M_t}\right)^{1/\gamma}, \quad \text{where} \quad \gamma > 1, \quad M_t \equiv \int_0^1 m_{it} \mathrm{d}i.$$

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

$$\max_{p_{jt}, m_{jt}} (p_{jt} - 1)\, y_{jt} - m_{jt}$$

Firm pricing, customers, and variable profits

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

$$n_{it} = \min \left\{ \left(\frac{q_{it} P_t}{\mu} \right)^{\theta-1} \frac{P_t C_t}{\theta \phi M_t}, 1 \right\}^{\frac{1}{\gamma-1}}$$

$$\Pi_{it} = \left[\left(\frac{q_{it} P_t}{\mu} \right)^{\theta-1} \cdot \frac{P_t C_t}{\theta \phi M_t} \right]^{\Gamma} \cdot \frac{\phi M_t}{\Gamma}, \quad \text{where} \quad \Gamma \equiv \frac{\gamma}{\gamma-1}$$

Innovation

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

$$q_{jt+1} = \begin{cases} q_{jt}e^{\Delta} & \text{w/ prob. } x_{jt} \\ q_{jt} & \text{w/ prob. } 1 - x_{jt} \end{cases} \quad \text{and} \quad s_{jt} = b_0 \cdot e^{b_1 x_{jt}} \cdot z_{jt}^{b_2}$$

Δ, b_0, b_1 and b_2 are all strictly positive

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

Imitation

A continuing firm's value function is given by:

$$v_t(z) = \Pi_t(z) + \max_{x \in [0,1]} \left\{ R_t^{-1} \left[x V_{t+1}\left(z e^{\Delta - g_t}\right) + (1-x) V_{t+1}\left(z e^{-g_t}\right) \right] - s_t(z,x) \right\}$$

Taking into account the imitation option, the value function is:

$$V_{t+1}(z) = \max \left\{ v_{t+1}(z), \int_{\underline{z}}^{\infty} v_{t+1}(z) \mathrm{d}F_t(z) - \epsilon \right\}$$

The lower bound $\underline{z}$ is determined by the "free re-entry condition":

$$v_{t+1}(\underline{z}) = \int_{\underline{z}}^{\infty} v_{t+1}(z) \mathrm{d}F_t(z) - \epsilon$$

Recap on firm decisions

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

1. Hire marketing labor to access customers (as a function of z)
2. Hire production labor to sell to customers (as a function of z)
3. Hire research labor to achieve a probability of research success $x \in [0, 1]$
4. Based on research outcome, continue or pay ϵ to re-enter with quality draw $F(z)$

Labor market clearing

$$L_t = \int_{\underline{z}}^{\infty} l(z)\, \mathrm{d}F_t(z)$$

$$M_t = \int_{\underline{z}}^{\infty} m(z)\, \mathrm{d}F_t(z)$$

$$S_t = \int_{\underline{z}}^{\infty} s(z)\, \mathrm{d}F_t(z)$$

$$E_t = \int_{\underline{z}}^{\infty} \delta(z)\, \mathrm{d}F_t(z) = \delta_t \epsilon$$

$$L_t + M_t + S_t + E_t = 1$$

Some aggregates

$$P_tC_t = \int_0^1 p_{it}n_{it}c_{it}\mathrm{d}i = L_t + M_t + \int_0^1 \Pi_{it}\mathrm{d}i$$

$$P_tC_t = \frac{\theta\Gamma\left(L_t + M_t\right)}{\theta\Gamma - 1}$$

$$M_t = L_t / \left[\gamma\left(\theta - 1\right)\right]$$

$$L_t = P_tC_t\left(\theta - 1\right)/\theta$$

Aggregate quality growth

$$1 + g_t = \left(\int_{\underline{z}}^{\infty} \left[\left(x_t(z) e^{\Delta} + (1 - x_t(z)) \cdot \left(1 - \delta_t(z) + \delta_t(z) \frac{\bar{z}}{z} \right) \right) \cdot z \right]^{\Gamma(\theta - 1)} \mathrm{d}F_t(z) \right)^{\frac{1}{\Gamma(\theta-1)}}$$

This produces the gross real interest rate via the usual Euler equation:

$$R_t = \frac{(1 + g_t)^{1/\theta}}{\beta}$$

Parameter values

Symbol	Parameter	Value
σ	Intertemporal elasticity of substitution	0.5
θ	Elasticity of substitution between varieties	3
β	Discount factor	0.991
ϕ	Scale of marketing costs	$2.75 \cdot 10^4$
γ	Elasticity of marketing costs wrt customers	1.25
Δ	Quality step size	0.04
b_0	Linear research cost	$3.44 \cdot 10^{-6}$
b_1	Convex research cost	14
b_2	Research spillover parameter	10
ϵ	Re-entry cost	1.18

Intensive and extensive margin elasticities

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:

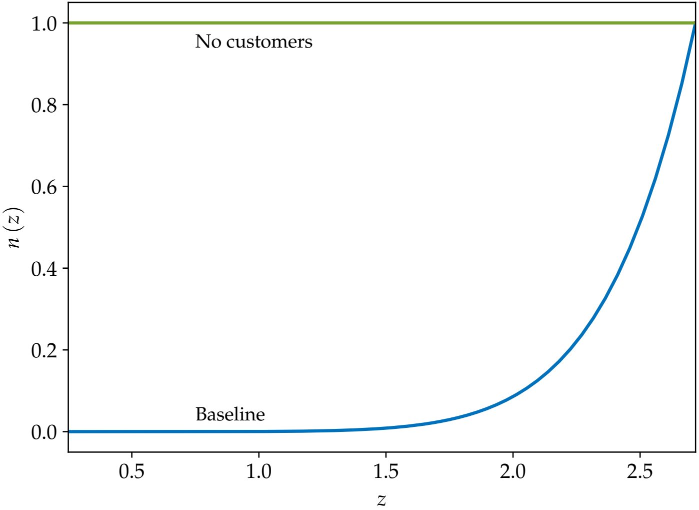
$$\xi_{y,q} = \xi_{n,q} + \xi_{c,q} = \frac{\theta - 1}{\gamma - 1} + \theta - 1$$

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

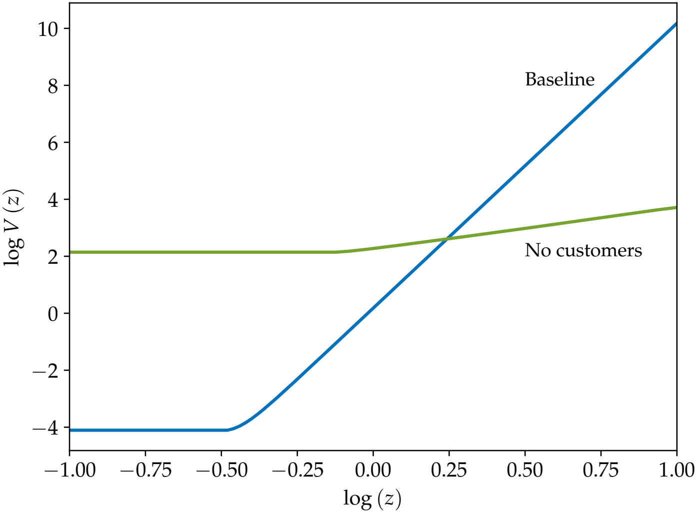
Steady state values of endogenous variables

Symbol	Parameter	Baseline	No customers
g	Growth rate	2.00%	2.04%
r	Interest rate	5.00%	5.07%
L	Production labor	68.3%	93.7%
M	Marketing labor	27.3%	0.0%
S	Research labor	3.13%	3.10%
E	Adoption labor	1.25%	3.24%
δ	Re-entry rate	1.11%	2.93%

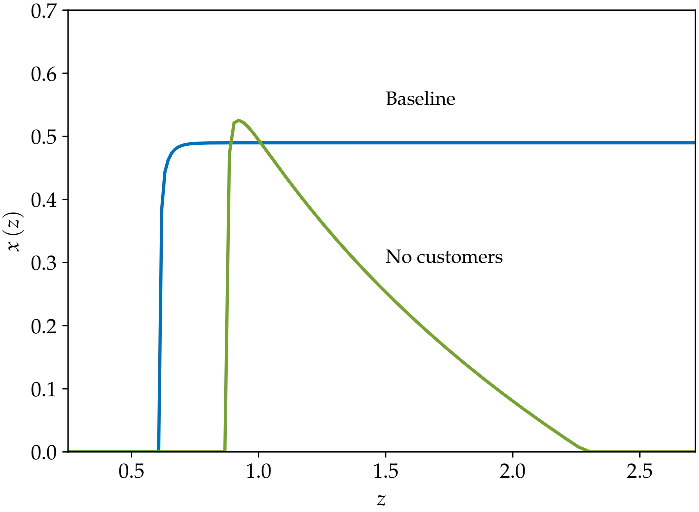
Customers and firm quality



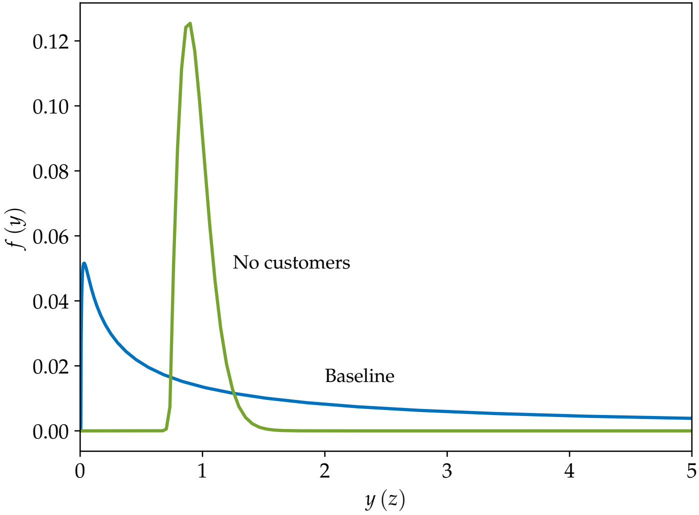
Customers and firm value



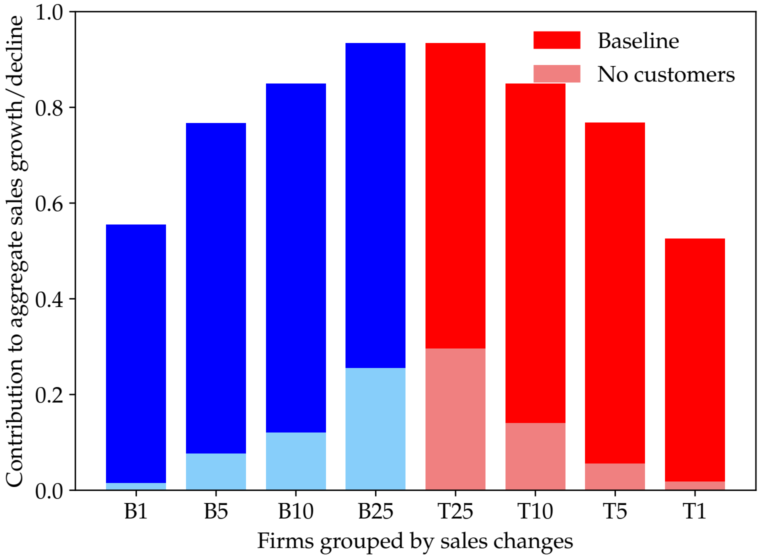
Customers and firm innovation



The distribution of sales



Firm contributions to aggregate sales changes



Conclusion

- We looked at Visa debit and credit card data on transactions at all offline retail merchants from 2016–2019

- 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 firms but not overall innovation and growth