NBER WORKING PAPER SERIES

ECONOMIES OF DENSITY IN E-COMMERCE: A STUDY OF AMAZON'S FULFILLMENT CENTER NETWORK

Jean-François Houde
Peter Newberry
Katja Seim

Working Paper 23361
http://www.nber.org/papers/w23361

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2017, Revised May 2021

A previous version of this paper was circulated under the title "Sales Tax, E-Commerce, and Amazon's Fulfillment Center Network." We thank Josalyn Geiman, Mallick Hossain, Olleskii Khvastunov, Nitin Krishnan, and Xinrong Zhu for research assistance. We have benefited from conversations with participants at various conferences and seminars, as well as discussions with John Asker, Alan Collard-Wexler, Paul Grieco, Jin-Hyuk Kim, Brad Larsen, Robin Lee, Francesca Molinari, Charles Murry, Amedeo Piolatto, Imke Reimers, Fanyin Zheng, and others. The views expressed herein are those of the authors and do not necessarily reflect the views of the Federal Communications Commission, the U.S. Government, or the National Bureau of Economic Research.

Economies of Density in E-Commerce: A Study of Amazon’s Fulfillment Center Network
Jean-François Houde, Peter Newberry, and Katja Seim
NBER Working Paper No. 23361
April 2017, Revised May 2021
JEL No. H71,L23,L81

ABSTRACT

We study the importance of economies of scale and density in e-commerce, using the expansion of Amazon’s distribution network between 1999 and 2018 as a case study. We highlight the role of two features: densification of the network of distribution facilities and vertical integration into package sortation. The resulting reduction in the cost of shipping orders comes at the expense of additional sales tax liabilities due to nexus tax laws, higher facility operating costs in more expensive areas, and lower scale economies of processing shipments. We combine data on household spending across online and offline retailers with detailed data on Amazon’s expansion in order to estimate this trade-off through a static model of demand and a dynamic model of investment. Our results suggest that Amazon’s expansion led to significant shipping cost savings, facilitated the realization of aggregate economies of scale, and lowered the external costs of e-commerce. We use the estimated model to quantify the distortionary effects of nexus tax laws on the firm’s distribution network and shipping cost, relative to non-discriminatory taxation.

Jean-François Houde
Department of Economics
University of Wisconsin-Madison
1180 Observatory Dr
Madison, WI 53706
and NBER
houde@wisc.edu

Katja Seim
School of Management
Department of Economics
Yale University
165 Whitney Ave
New Haven, CT 06511
and NBER
katja.seim@yale.edu

Peter Newberry
Department of Economics, Penn State University
510 Kern Graduate Building
University Park, PA 16802
pwnewberry@gmail.com

1 Introduction

Online retail has grown substantially over the two last decades, accounting for 12% of all retail spending in Q1 of 2020.[1] The largest online retail platform, Amazon.com, henceforth Amazon, is central to this growth. Between 1999 and 2018, Amazon's share of online spending has grown from 10% to 45%, contributing to the rise in the concentration in retail markets (Autor et al., 2020). This increasing dominance suggests that e-commerce is associated with important economies of scale and scope, leading to a winner-take-all trajectory.

The importance of demand-side increasing returns to scale on platform competition is well understood (Greenstein, 1993; Chu and Manchanda, 2016; Cao et al., 2018, see). A less studied source of supply-side economies of scale is investment in distribution networks. Rather than relying on existing hub-and-spoke networks operated by independent logistic companies, Amazon is now the third largest delivery company globally, and industry experts forecast that it will soon operate a fully integrated logistic supply chain.

While this investment strategy mimics that of vertically integrated brick-and-mortar retail chains, such as Walmart, Kmart and Target (Barwick, 2008; Holmes, 2011; Zheng, 2016), a fundamental difference between e-commerce platforms and traditional retailers is that consumers make purchases online without being physically close to product inventories.[2] This allows online retailers to optimize the configuration of their logistic networks to minimize costs, accounting for economies of density and scale. However, differential state tax policy has the potential to affect the network configuration. We focus on nexus tax laws that, during our sample period, required online retailers to collect sales tax only in states where they maintain a physical presence. Such laws favor retailers with small geographic footprints and introduce demand-side considerations to the network configuration problem, which distorts investments in cost-saving technologies.

The goal of this paper is to measure these distortions by evaluating the effect of discriminatory tax policies such as nexus tax laws on the distribution network of online retailers. To accomplish this objective, we study Amazon's growth between 1999 and 2018 through the lens of a dynamic model of investment in a distribution network. We first quantify the revenue trade-off associated with nexus tax laws and use the model to measure the importance of economies of density at Amazon. We then use these cost and demand estimates to assess the distortions caused by nexus laws on the average cost of fulfilling orders and on the distance to final consumers.

In the model, Amazon chooses the locations of new logistic facilities anticipating the impact of its network configuration on current and future revenue and on the cost of fulfilling orders. We focus on two types of facilities: (i) fulfillment centers and (ii) sortation centers. Fulfillment centers are facilities where goods are stored, orders are packed, and packages are transferred to downstream facilities for sorting and final delivery. In contrast, sortation centers are facilities where packages are sorted by destination in preparation for final delivery. By locating fulfillment and sortation

[1] See `https://www.census.gov/retail/mrts/www/data/pdf/ec_current.pdf`.

[2] In Houde et al. (2021), we show that consumer spending at Amazon is independent of distance to distribution centers.

centers close to each other, Amazon can integrate the order fulfillment process vertically and fulfill an order entirely in-house, up to the last mile.

We model the cost of fulfilling orders as a combination of three components. Fixed costs include the cost of warehouse space, which depends on the local rental rate and a congestion penalty of operating in urban areas. Variable costs include labor costs of processing the orders at a facility and shipping costs, which depend on total shipping distance and whether orders are handled by an integrated sortation center or by a independent shipping company.

The model predicts that the optimal network configuration is the outcome of two trade-offs. First, Amazon faces a trade-off between shipping cost savings (density) and order processing and fixed costs (scale), given current and future demand. Second, nexus tax laws induce a trade-off between economies of density and revenue. Upon building a distribution facility in a new state, online retailers have to collect sales taxes on all purchases from that state. The magnitude of this opportunity cost is modulated by the elasticity of demand, as well as tax collection delays negotiated between Amazon and state governments.

We measure the relative importance of each of these forces in determining the location of new facilities in three steps. In the first step, we estimate a model of demand for online and offline goods. We combine data on household-level online spending at various websites from the comScore Web Behavior database, data on retail spending in total from the Consumer Expenditure Survey, and data on the propensity to shop online from Forrester Research. We estimate a CES specification that allows us to predict household spending and Amazon's revenue in each county and measure the elasticity of Amazon's revenue to local tax changes, a key ingredient to quantifying the revenue side of the network expansion trade-offs. We identify the demand elasticity using quasi-random variation in Amazon's tax-inclusive price caused by past expansions of the distribution network into different states. We control for nationwide improvements in quality, convenience, and product variety with year-platform fixed effects. The results confirm that consumers are responsive to price, with a demand elasticity of approximately -1.5, similar to the estimates from Einav et al. (2014) at -1.8 and Baugh et al. (2018) at -1.5. The estimate implies that a customer in a non-taxed location spends 9.8% more at Amazon than the same customer in a location with the average sales tax rate of 6.5%.

In the second step, we use the volume of orders originating from each county in each year, as predicted by the demand model, to estimate the production technology of processing orders. Since we do not observe how Amazon allocates orders to fulfillment centers, we use a simple probabilistic model of product availability to predict order flows. We rely on data on the distributions of capacity and number of employees across facilities to identify the likelihood of product availability and the order processing technology. We estimate that a given fulfillment center is able to satisfy an order with a probability of approximately 50%. As orders are therefore frequently allocated to more distant fulfillment centers with higher shipping costs, this creates an incentive to use a denser network of facilities. At the same time, the results imply that there are sizable increasing returns to scale to order processing, which leads to an incentive to build fewer, but larger facilities with higher capacity utilization rates.

In the final step, we estimate shipping and fixed costs using a moment inequality estimator. Following Holmes (2011), we specify a set of moment inequalities that rationalize Amazon's observed network expansion strategy. We compare the firm's discounted profit stream under the observed distribution center locations to its profit under alternative locations Amazon could have chosen, but elected not to. For a given fulfillment center, we define alternative locations as the locations of other facilities opened at later dates. We find values of cost parameters that render the observed network roll-out more profitable than these alternatives.

We exploit variation in taxes and input prices across counterfactual networks to identify the effects of distance and vertical integration into sortation on shipping costs and the effect of local congestion on fixed costs. For example, moving up the opening date of a facility in a low-population, low-tax state and in turn delaying the opening of a facility in a more populous, high-tax state implies both higher revenue streams due to lower exposure to sales tax early on and longer shipping distances and thus higher cost. By independently varying the relative magnitudes of the predicted tax implications and input prices, we estimate bounds on each of the cost parameters.

Our estimates imply that Amazon's average cost of shipping an order decreased from \$2 to \$0.30 over the sample period. Highlighting the potential supply chain efficiencies associated with vertical integration into sortation, we find that a significant contributor to these lower shipping costs is in-house sortation, reducing shipping costs by 40% in 2018. Overall, the results suggest that the economies of density in shipping costs exceed the scale economies to order processing. As a result, Amazon's implied long-run average total cost of order fulfillment exhibits substantial economies of scale. By expanding its network, Amazon reduced its total average fulfillment cost by about 46%, despite facing larger labor and fixed costs. Finally, we demonstrate that Amazon's expansion reduced an order's average shipping distance from 450 miles to 141 miles. Using estimates of pollution and congestion costs of road transport from the environmental economics literature, we show that the increased proximity to the consumer has slowed the growth of external costs of long-haul trucking due to e-commerce.

We next use the estimated demand and fulfillment cost functions to illustrate the effect of nexus tax laws on the configuration of the distribution network. We simulate counterfactual profit-maximizing network configurations under two tax regimes. We contrast the current nexus tax laws with a non-discriminatory tax policy, where Amazon collects sales tax in all states. Since the computational burden of solving the dynamic investment problem that Amazon faces is prohibitive, we approximate its solution with a series of static profit maximization problems evaluated at different levels of demand. The solutions to these static problems show that the estimated demand and cost functions are able to predict the network expansion observed in our sample period well.

We find that under the discriminatory tax system entailed by nexus tax laws, the simulated networks are both smaller and more centralized than under a non-discriminatory tax collection policy. Evaluated at the 2018 demand, we find that the average shipping cost and distance are 28% and 15% higher, respectively, under the nexus policy. The higher shipping costs are partially offset by savings in labor and fixed costs, as total average costs are only 6% higher than under a non-discriminatory tax regime. The nexus policy also leads to an increase in external costs from long-haul trucking of 22%.

Our work relates to several strands of literature. An increasing body of work focuses on the estimation of demand in online retail markets (e.g., Dolfen et al., 2019), the most relevant being studies on the responsiveness to sales tax and the gains from variety. Einav et al. (2014) estimate the demand response to sales tax using eBay data, exploiting the fact that a buyer has the option to buy from an out-of-state seller who does not charge sales tax. Baugh et al. (2018) use a differences-in-differences approach to estimate the effect of the 'Amazon tax', or changes in sales-tax collection on Amazon between 2013 and 2015. In estimating the tax sensitivity, neither set of authors is able to consider substitution to other taxed or non-taxed online outlets. We thus contribute to this literature by expanding the estimation of demand for retail goods beyond a single online firm to a large number of online and offline retailers.

Our analysis is related to recent operations research and economics literature on the management of online firms' distribution networks (see Agatz et al., 2008, for an overview) and on the relationship between a brick-and-mortar retailer's store locations and its distribution network. Building on Barwick (2008) who estimates the aggregate scale economies, irrespective of source, from operating multiple stores in close geographic proximity, Holmes (2011) estimates the savings in distribution costs associated with clustering stores near a fulfillment center. Zheng (2016) relates the proximity of a rival's fulfillment center to the chain's expected future entry. These studies take the configuration of the distribution network as given and rely on variation in the distances from a fulfillment center to potential store locations to identify scale economies. Instead, we study the development of the network of fulfillment centers as a strategic choice for the firm. Little work to date has studied such classic industrial organization questions as the role of cost differences in affecting firms' competitive positions in the context of distribution in online markets.

The remainder of the paper is organized as follows. The next section describes the consumer spending data and summarizes expansion of the network. Section 3 specifies Amazon's optimization problem and the components of the profit function, while Section 4 presents the estimates of the model parameters. Section 5 uses these estimates to analyze the impact of the distribution network expansion and the role of sales tax laws therein. Section 6 concludes.

2 Data and Stylized Facts

To analyze the expansion of Amazon's distribution network, we rely on several data sources. The available network data covers the years 1999 to 2018, which defines our sample period. The other data do not cover the entire sample period, so we use a combination of our estimated models, interpolation, and extrapolation to construct the missing observations. Section 4 and Appendices A and B.1 describe these processes.

We begin by briefly discussing each source and the process by which we construct our primary variables, although we leave many of the details for the appendices. We then present important facts and trends in the data that we rely on to identify the determinants of Amazon's profit function.

2.1 Retail Spending and Orders: Data

The main input into our demand model is the spending on retail goods by a county's representative household. We construct a panel of annual spending, differentiating between different types of shopping outlets, or shopping 'modes'. We consider three online shopping modes that vary in their exposure to sales tax and a single offline shopping mode.

We rely on the comScore Web Behavior Database from 2006 to 2016 to construct the spending series for the three online modes. comScore records the price, the domain, and the date of every transaction made by a random sample of online shoppers.[3] For each respondent, comScore also records the five-digit ZIP code and a number of demographics. The sample includes 40 (12) thousand households and covers 85% (54%) of US counties in its largest (smallest) sample year. The counties covered are the most populous, as 99% (92%) of US households live in the represented counties in the largest (smallest) sample year (see Table A-1).

We manually classify each transaction's seller into a retailer type based on the seller's physical footprint across the US. Amazon, which we denote as shopping mode 1, has sales tax liabilities in select states due to its growing distribution network, but, during most of our sample, not in states where it does not have a physical presence. We note that Amazon's sales are comprised of both direct sales and sales through third-party sellers (known as Amazon Marketplace). Third-party sales on Amazon have grown substantially over time from a 3% share of Amazon's worldwide gross sales in 1999 to 58% in 2018.[4] While Amazon's tax obligations and margin differ between direct and third-party sales, we cannot distinguish expenditures across them in our data and, thus, cannot model them as separate modes. Instead we use data on the aggregate share of sales by third-party sellers, observed in Amazon's 2018 financial report, to account for these differences. See the discussion in Section 3.

Shopping mode 2 consists of the online arms of nationwide retailers with a broad physical store network, such as `walmart.com`, which we denote as taxed online retailers. Mode 3, which we denote as non-taxed online retailers, covers firms that rely on the online sales channel only but do not operate extensive distribution networks; they thus lack a physical presence across states (e.g., `overstock.com`). Therefore, the variation in retailer tax collection obligations across shopping modes, combined with within- and across-state variation in sales tax rates, changes the relative price of a given purchase across locations and modes. Using this classification, we calculate the mode-level annual spending for each household in the comScore sample.

We make two important adjustments to the comScore spending variables. First, we account for

---

[3] comScore also records browsing behavior for households who do not shop online. We do not rely on these data here.

[4] See Amazon's Annual Letter to Shareholders, 2018, available at `https://www.sec.gov/Archives/edgar/data/1018724/000119312519103013/d727605dex991.htm`, accessed 6/30/2020.

households who do not shop online by using survey data from Forrester on the prevalence of online shopping as a function of demographics. Second, we correct for the fact that comScore tracks only the subset of user transactions made on a single registered device by scaling up household expenditures to match average spending per household from Amazon's financial statements. We similarly scale other online spending using reports from the US Department of Commerce. We calculate spending on the offline mode, mode 0, by combining data from the annual Consumer Expenditure Survey and Census tables. We then calculate the county-level weighted average annual spending on each mode using population weights from the US Census, giving us annual mode-level spending for the representative household in each county. Appendix A provides details.

In addition, the demand model includes a number of consumer demographics, which we collect from the American Community Survey and the Decennial Census, and measures of local concentration of offline retailers, which we collect from the Census' County Business Patterns. See Appendix A.

2.2 Distribution Network: Data

We obtain information on Amazon's distribution network from the supply-chain consulting company MWPVL, International (`http://www.mwpvl.com/`). For each distribution facility, MWPVL provides information on the location, size in square feet of floor space, employment, facility type, opening date and closing date, where applicable.

We rely on the facility type to identify two primary types of distribution centers. First, we group all distribution centers that store non-grocery items into a single category of fulfillment center.[5] We drop specialized distribution centers, including 'PrimeNow Hubs', Amazon Fresh grocery delivery centers, return centers, and distribution centers for select high-value items such as jewelry.

The second type of facility, the sortation center, is a downstream facility in the delivery process. For most of our sample, independent shippers such as UPS and FedEx handled much of the order fulfillment process, including the shipping process, routing the packages through their own network of sortation and delivery facilities from the fulfillment center to the final destination. However, starting in 2014, Amazon began to build its own sortation network, which brings the majority of the fulfillment process in-house: co-located fulfillment centers continuously send shipments to the sortation centers, where these are grouped into three-digit destination zip code areas and transferred for 'last-mile' delivery by either the postal service or another local courier.[6] Figure 1 summarizes the process for both 'outsourced' and 'integrated' shipments.

According to MWPVL, packages that are routed through Amazon's own sortation network satisfy two conditions. First, the final destination must be within the sortation facility's coverage region. MWPVL suggests that a sortation center's catchment area includes destination zip codes

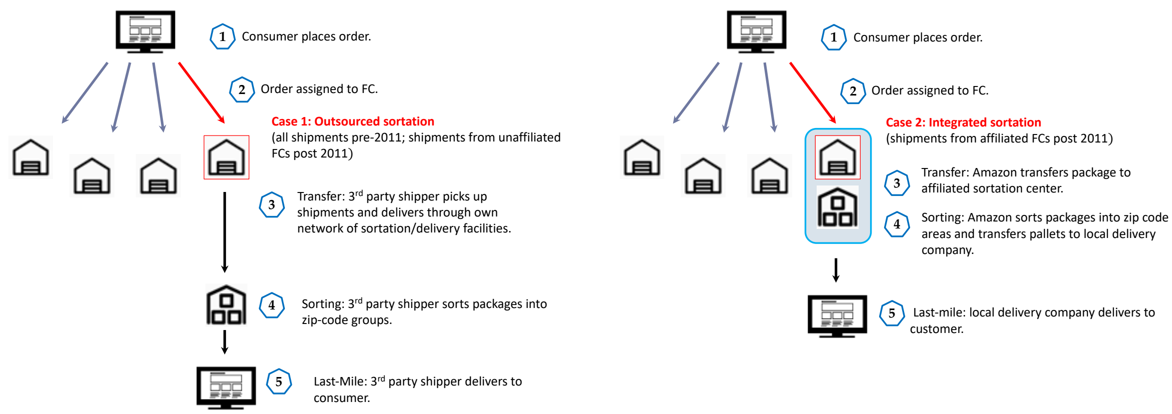
---

[5] This includes non-sortable centers, handling large items that the firm cannot ship in combination with any other products, and small and large sortable centers, handling items that the firm is able to combine into a single package.

[6] Since the end of our sample period, Amazon has also begun to invest in last-mile delivery, in particular in urban areas.

Figure 1: Logistic of Online Transactions

(a) Independent Sortation (b) Integrated Sortation



within 150 miles from the facility. Second, the sortation center must be near one or more fulfillment centers, at a distance of at most 25 miles.

2.3 Cost components: Data

We assume that the costs Amazon faces when making network decisions are split into fixed costs and variable costs per order. Fixed costs consist of the cost of warehouse space. We recognize, per square foot of warehouse space, the observed rental rate and a congestion penalty to urban locations. We approximate the rental rate with local commercial rents for warehousing space. The primary source for these data is Moody's Analytics REIS database which covers most MSAs between 2006 and 2018. Appendix A describes how we construct county-level rents per square foot using these data. The congestion penalty allows for the possibility that fixed costs are higher in urban counties, where, for example, integrating facilities into the highway network is more difficult. We rely on the population density of the facility's county, obtained from the Census, to be a proxy for such costs. We estimate the total fixed cost of a facility as the square footage reported by MWPVL times the sum of rent and congestion payment.

The variable costs of an order include the wholesale cost of the product, the labor cost of processing the order, and the shipping cost. We follow Holmes (2011) to infer wholesale costs, or the cost-of-goods sold, from the gross margins reported in Amazon's financial statements. See Online Appendix OA.1.

We scale the number of employees at a facility by an average annual wage to obtain annual labor costs of processing shipments at that facility. We use the average annual county-level wage of a retail employee from the Bureau of Labor Statistics as our measure of wages for fulfillment and sortation center employees. We obtain employment data for 131 of the 165 facilities in 2017 from a combination of industry reports and Amazon's financial statements (see Appendix A). We also observe system-wide employment in 2017 from an Amazon press release.[7]

Finally, we assume that shipping cost increases with shipping distance and calculate the distance a shipment travels from a given fulfillment center to the consumer in a given county. We use the Haversine formula to calculate the straight-line distances from the fulfillment center to each county's population-weighted centroid.

2.4 Sales Taxes: Data

The sales tax data come from two primary sources. First, we obtain state, county, and local sales tax rates from Thomson Reuters' Tax Data Systems for the years 2006-2018. For each year and county, we calculate the average tax rate, as tax rates can vary within a county and may change mid year. We assume that this sales tax rate applies to all consumer transactions at taxed online and brick-and-mortar retailers, as well as to taxable transactions on Amazon.[8] The average sales tax rate is 6.5% across counties and years, and there is a significant of amount of time-series and cross sectional variation in rates (see Table A-5).

Second, we observe the date on which Amazon began to collect sales tax in every US state. We rely on data from Baugh et al. (2018) for states that realized the change before the end of 2015. For the remaining states, we obtain the date of the change using various news sources. In 2017, Amazon voluntarily began collecting sales tax on all of its transactions, regardless of consumer location. This change was largely inconsequential, however; at the time, the company was already collecting sales tax on orders from over 90% of US households. As our demand model is at the annual level, we assume that the sales tax collection obligation applies to a given year and all subsequent years if it is effective before August of the year; otherwise we assign it to the following year.

Changes in the sales tax status are triggered by Amazon's expanding distribution network due to nexus tax laws. As we discuss above, these laws require retailers to collect sales tax from consumers if they have a physical presence in the consumer's state of residence. Otherwise, the consumer is responsible for remitting a use tax, but few consumers do so.[9] Not surprisingly, with the rapid growth of e-commerce, brick-and-mortar retailers and policymakers, fearing significant tax revenue losses, began supporting legislation to revise the definition of nexus as early as 2008 (see Bruce et al., 2009). This culminated in a 2018 Supreme Court decision, where the court ruled that states could put tax collection responsibility on online retailers that exceed a minimum transaction threshold.

The onset of Amazon's tax collection does not always coincide with the opening of the first

[7] Source: https://press.aboutamazon.com/news-releases/news-release-details/amazon-now-hiring-over-120000-jobs-us-holiday-season.

[8] Note that in five states, clothing purchases are generally sales tax exempt. We conduct two robustness checks where we remove either all households from the affected states or all purchases categorized as "Apparel" from the sample. The exemptions do not appear to drive the estimated tax responsiveness. See Houde et al. (2021).

[9] See, e.g., http://www.npr.org/sections/money/2013/04/16/177384487/most-people-are-supposed-to-pay-this-tax, accessed on 1/12/2017.

facility in the state. For example, Amazon opened its first facility in Tennessee in 2011 and did not start collecting sales tax in the state until 2014. Such gaps reflect negotiations about tax collection responsibilities between state governments and Amazon. Amazon's ability to extract tax delay concessions from state governments varies with the growth of its network and its demand, as we discuss in Appendix A.3. Hereafter, we call such delays 'tax abatement' agreements, although we are being somewhat loose with this terminology.[10] We account for the delays in our model by assuming a deterministic schedule that maps the year of the first entry into a state into the year when Amazon begins collecting sales tax. See Section 4.3.

2.5 Trends: Retail Spending and Orders

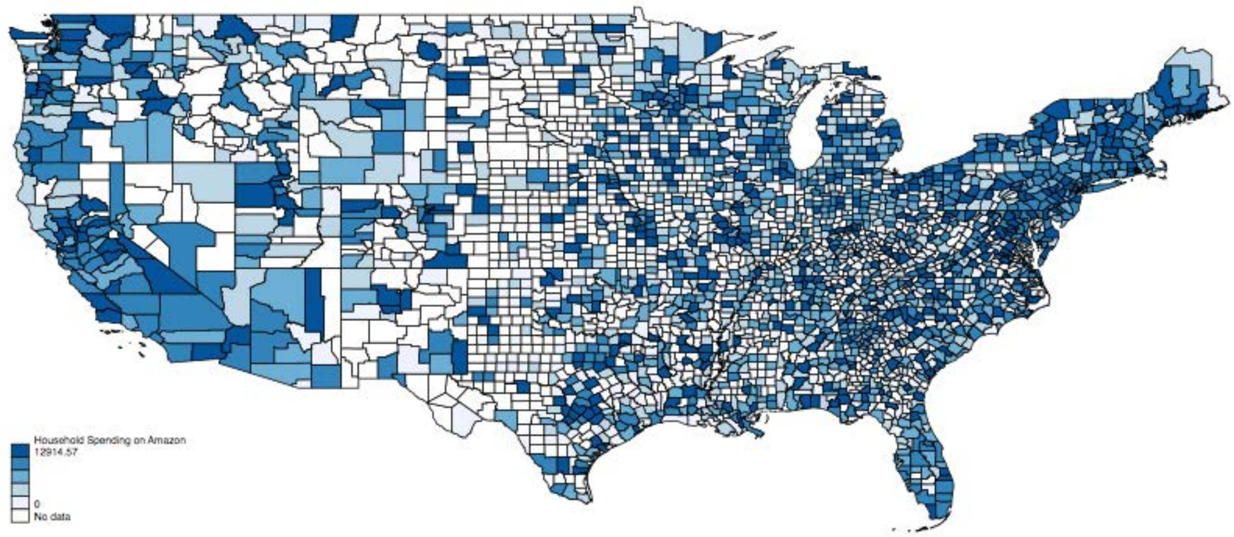
The basic descriptive patterns in spending demonstrate, not surprisingly, significant growth in online shopping during our sample, with an average annual growth rate in household spending of about 12%. At the same time, offline spending experiences an average annual decline of 5%. The online growth reflects, in large part, Amazon's expansion. The company's sales grow on average 33% per year during our sample, while non-taxed and taxed online retail grow 10% and 8%, respectively, per year. Amazon's share of online retail thus increases from about 6% in 2006 to 31% in 2016. Table A-1 and Figure A-1 in Appendix A.1 demonstrate these patters.

While the growth in online spending in many locations in the US mirrors the aggregate trends, there is significant cross-sectional variation in the level of spending on Amazon. We explore this geographic variation in Figure 2, where we categorize counties based on the quintiles of the distribution of spending on Amazon by the county's representative household in 2016. The map indicates that counties in larger markets, such as counties around San Francisco, Chicago, and New York, fall into the top or second spending quintiles of all counties. At the same time, a number of counties in less densely populated areas, such as counties east of the Mississippi River, also exhibit high levels of spending, placing them in the same quintiles. This spatial variation in spending provides incentives for Amazon to decentralize its distribution network and to reduce shipping cost to not just major metropolitan areas, but also some of these less urban areas.

Below the map, we break down average household spending on Amazon in 2016 across urban and rural counties, wealthy and non-wealthy counties, and by Census regions. The data suggest that households in urban or high-income counties spend more on Amazon than rural counties and low-income counties. There is also regional variation in spending with the Northeast and the West having higher spending levels. See Table A-4 for these spending averages for each of the sample years.

[10] Amazon benefits from the fact that they are not required to collect sales tax because consumers rarely remit the alternative use taxes. In practice, a sales tax collection delay is thus similar to an abatement agreement where firms do not have to pay property taxes for a certain length of time in exchange for economic development.

Figure 2: Geographic Distribution of Spending on Amazon (2016)



Average Annual Household Spending by Demographic Group ($, 2016)

Density		Income		Region			
Urban	Rural	High	Low	Northeast	Midwest	South	West
1,125	1,020	1,219	989	1,159	1,019	1,007	1,181

Notes: We classify counties with a population density of at least (less than) 500 people per square mile as urban (rural) and counties with median household income of at least (less than) $80,000 as high (low) income.

2.6 Trends: Distribution Network

Table 1 summarizes the roll-out of Amazon's fulfillment and sortation centers over the period 1999 to 2018, in three-year increments. The company expanded its fulfillment center network from five individual facilities in 1999 to 128 by 2018.

Often, this expansion takes the form of locating new facilities within close proximity of existing facilities, which we treat as co-location going forward. We use a clustering algorithm to define groups of co-located fulfillment centers, or 'clusters', as of the end of the sample period. Roughly, this amounts to grouping facilities that are within 20 miles of each other. We assign the centroid of the locations of all clustered facilities at the end of our sample as the cluster's location, recognizing that shipping costs and distances to the various facilities within the cluster are largely the same. Therefore, in the logistic model we describe below, we calculate the shipping cost at the level of the cluster.

As an example, in 2018, Amazon operates six fulfillment centers near Harrisburg, PA, which we group into a cluster located at the centroid of the six facilities. This cluster first came into existence in 2010 when Amazon opened two facilities in Harrisburg. Over the next eight years, it opened four additional nearby facilities, expanding the cluster in both number and size. The number of clusters, which we list in parentheses next to the number of fulfillment center locations in Table 1, grew from five in 1999 to 75 in 2018. Hereafter, we use the terms 'cluster' and 'location' interchangeably, as they both to indicate a potential site for expansion, either in the form of de-novo entry, if no fulfillment center is active in the cluster at a point in time, or in the form of incremental expansion through the addition of new facilities. We refer to clusters with at least one operating fulfillment center as 'active clusters'.

Table 1: Expansion of the Distribution Center Network

	Facilities			Size (100k ft^2)			Households	
Year	FC	SC	States	FC Cluster	SC	Employ-ees	Distance	With SC
1999	5 (5)	0	4	5.9	-	305.8	307.6	-
2002	6 (6)	0	4	6.0	-	329.3	301.2	-
2005	9 (9)	0	6	5.6	-	337.3	268.8	-
2008	15 (12)	0	8	7.8	-	455.3	253.3	-
2011	25 (14)	1	9	13.2	3.2	1009.8	240.1	1.8
2014	57 (33)	12	17	14.0	3.5	965.1	127.2	48.7
2017	98 (52)	30	28	15.2	3.2	1984.1	73.8	81.7
2018	128 (70)	35	30	15.1	3.3	1906.7	67.4	82.8

Notes: Under facilities, we depict the numbers of fulfillment centers, with the number of active clusters in parenthesis, sortation centers, and states with a facility. Size is the average square footage of fulfillment and sortation centers and employees the number of employees of a cluster. Distance denotes the population-weighted average distance in miles from a county's centroid to the closest fulfillment center location. Households with SC is the percent of US households with a sortation center within 150 miles.

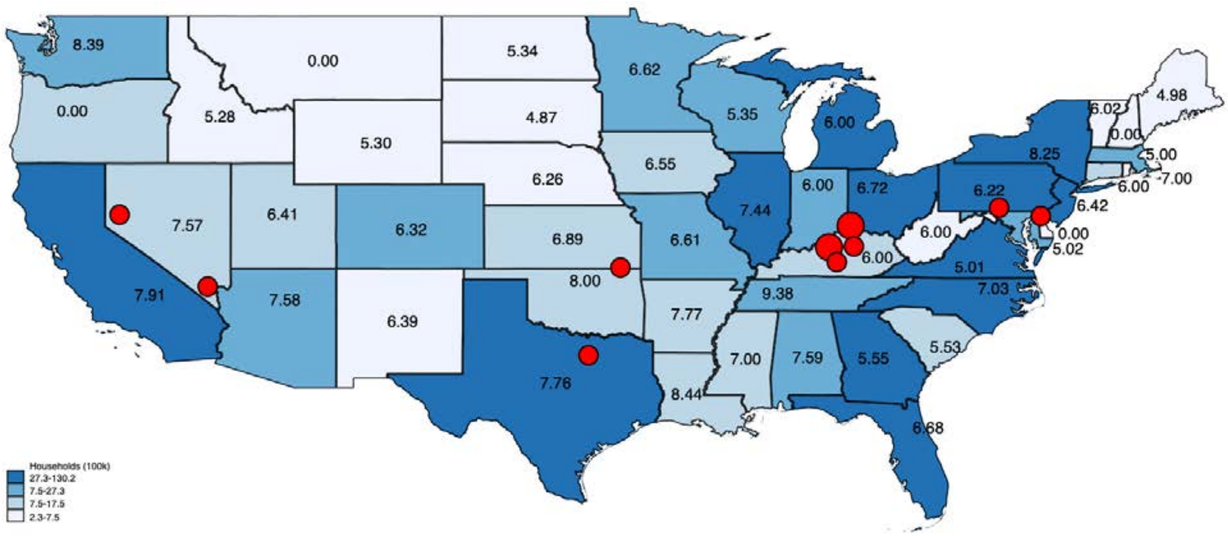
The growth in fulfillment centers has been accompanied by a densification of the network. One indication of this densification is the increase in the number of states with at least one active cluster, which rose from four in 1999 to 30 in 2018. Table 1 also depicts network density in miles, measuring the population-weighted average great-circle distance between each consumer location, which we take to be a county's population-weighted centroid, and its closest cluster. The average distance to the consumer fell from 308 miles in 1999 to only 67 miles in 2018. Most of this decline is due to the expansion of the distribution network into the most densely populated states along the coasts in the mid-2010s. We illustrate these patterns in Figure 3, which maps locations of clusters (red) in 2006 and 2018, shading states by the number of households and presenting the state average sales tax rate. The size of each bubble indicates the number of fulfillment centers in the active cluster.

Panel (b) of Figure 3 adds the location of the 35 sortation centers (yellow) that Amazon built by 2018. As the map and Table 1 illustrate, sortation centers primarily serve large urban areas: by 2018, the relatively small number of sortation centers is able to serve a set of counties that together account for 83% of US households.

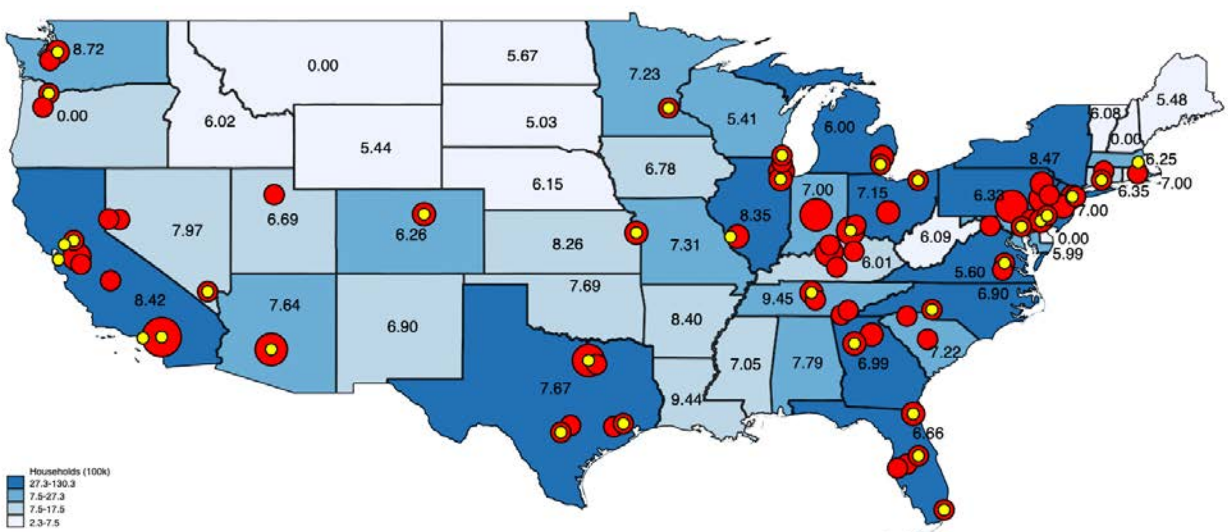
Table 1 shows that both the size and the number employees per cluster have increased substantially over time. In addition, Amazon has expanded into more expensive areas: Table 2 shows that the average rent per square foot in Amazon's facility locations is below the national average

Figure 3: Distribution Center Network, 2006 and 2018

(a) Amazon fulfillment center network, 2006



(b) Amazon fulfillment center (red) and sortation center (yellow) network, 2018



Notes: Counties categorized into and shaded by quintiles of the distribution of number of households. Size of each bubble scales with number of fulfillment centers.

(in parentheses) in 1999, but about 10% higher than the national average in 2018. Similarly, the average population density of counties with a facility, our measure of congestion, has increased substantially over time, and the wages in these counties are about $1,000 above the national average by the end of the sample. This suggests that, while the distribution network's expansion has reduced the distance to the average consumer and thus shipping cost, it has also disproportionately increased fixed and labor costs.

Table 2: Local Determinants of Profitability

Year	Rent ($/sq ft)	Population Density (pop/sq mile)	Wage ($k)	Sales Tax (%)	Taxed Households (%)
1999	3.5 (3.6)	302.7 (245.8)	14.4 (13.8)	5.1 (6.0)	0
2002	3.7 (3.8)	414.7 (248.9)	15.7 (15.3)	5.1 (6.0)	0
2005	4.1 (4.0)	562.2 (252.1)	17.2 (16.7)	5.7 (6.0)	2.7
2008	4.2 (4.2)	512.1 (257.0)	18.8 (18.5)	6.0 (6.1)	11.3
2011	4.0 (4.0)	595.6 (259.2)	20.0 (19.3)	6.6 (6.2)	11.3
2014	4.2 (4.2)	685.0 (266.0)	21.1 (20.4)	6.8 (6.3)	69.2
2017	4.7 (4.4)	1059.5 (270.8)	22.8 (22.0)	6.6 (6.4)	97.6
2018	5.0 (4.6)	1156.7 (272.0)	23.6 (22.7)	6.6 (6.4)	97.6

Notes: We display average commercial rents, population per square mile, and annual wages at the county level and average state-level sales tax rates for counties with an active fulfillment center cluster, compared to the average across all counties or states in parentheses. Taxed households represent the share of US households who live in states with positive sales tax and for whom Amazon remits sales taxes.

Some additional aspects of the distribution network expansion are noteworthy. Until 2010, Amazon placed fulfillment centers in relatively low-population and low-tax states near highly populated or high-tax areas. For example, Amazon opened two facilities in Nevada on the California border, close to that state's major cities. The company also opened one fulfillment center each in New Hampshire and Delaware, both of which are close to major East Coast cities and have zero sales tax. Indeed, the second to last column of Table 2 highlights that prior to 2011, the average sales tax rate in states with a fulfillment center is lower than the nationwide average. At the same time, the percentage of taxed consumers only increased by about 11% (last column of Table 2) between 1999 and 2011, even though, as Table 1 indicates, the average distance between consumers and their closest fulfillment center fell by 22% over that period. These patterns suggest that the company actively chose locations that mitigate consumer exposure to sales taxes, in terms of either the tax level or the size of the customer base exposed to taxes.[11]

As Amazon grew in scale, however, the network of fulfillment centers expanded beyond such locations, presumably to be closer to population hubs despite sales tax implications and higher fixed costs of warehousing in densely populated areas. For example, by 2014, we see entry into

[11] Work documenting consumers' sensitivity to sales taxes, in terms of the extensive margin of shopping online, includes Goolsbee (2000a), Goolsbee (2000b), Alm and Melnik (2005), Ballard and Lee (2007), and Scanlan (2007). Ellison and Ellison (2009), Smith and Brynjolfsson (2001), Anderson et al. (2010), and Goolsbee et al. (2010) study the intensive margin response to sales tax. Numerous papers, including Asplund et al. (2007), Agarwal et al. (2017), and Chetty et al. (2009), study the response to sales taxes in offline markets.

highly populated states, such as California and Virginia, and high tax states, such as Tennessee (9.5% tax rate in 2013). By 2018, Amazon added fulfillment centers in Georgia, Illinois, North Carolina, and Ohio. The average sales tax rate in states with a fulfillment center surpassed the overall average in 2011. Further, the growth in the share of households subject to sales tax rises to over 90% by 2017, when Amazon began to collect tax from all consumers.

3 Model

We are interested in modeling the investment decisions leading to the expansion of Amazon's distribution network. Here, we detail components of the online retailer's underlying per-period payoff function, paying particular attention to the firm's logistic problem, which drives the trade-offs it faces in choosing the network of distribution facilities dynamically, as demand and order flows evolve. We conclude this section with a description of this dynamic optimization problem.

3.1 Distribution network

We use the following notation to describe the distribution network. Facilities are indexed by $j = 1, \ldots, n$, and are fully described by their entry year (a_j), capacity (k_j), location (l_j), and type ($m_j = \{FC, SC\}$).[12] The index j represents a unique label attached to each facility. We assume that locations are chosen from a finite set of possible locations, $l = 1, \ldots, L$. In line with the co-location of facilities we observe in the data, each location l can accommodate multiple fulfillment centers and up to one sortation center. We thus use l to describe both the cluster and its location. We use $n_t = \sum_{j=1}^{n} 1(a_j \leq t)$ to denote the number of active facilities in period t. Similarly, $K_{lt} = \sum_{j=1}^{n} 1(a_j \leq t) \times 1(l_j = l \text{ and } m_j = FC) \cdot k_j$ measures the fulfillment capacity of cluster l in period t.

We assume that the square-footage and type of each facility are fixed, so that the distribution network evolves over time based on the opening of new facilities. We summarize the network in period t, given the chosen sequence of opening decisions $\mathbf{a} = \{a_1, \ldots, a_n\}$, by an $n_t \times 3$ matrix:

$$N_t(\mathbf{a}) = \left\{(k_j, l_j, m_j) \middle| \forall j \text{ s.t. } a_j \leq t\right\}.$$

We use $\mathbf{a}^0$, n_t^0, and N_t^0 to denote the observed sequence of entry decisions and observed number of active facilities and characteristics of these facilities in period t.

3.2 Demand and revenue function

We use a representative agent framework to predict orders and revenues as a function of the network configuration and sales tax laws. Consumers are located at the population-weighted center of their county, indexed by $i = 1, \ldots, I$. They allocate their retail budget B_i between brick-and-mortar retailers ($k = 0$) and the three online modes ($k = 1, 2, 3$).

[12] We abstract from facility exits as only one fulfillment center closes during the sample.

Preferences for each mode reflect differences in tax exposure. For Amazon, or mode 1, tax nexus laws imply that the distribution of sales tax across counties at a point in time reflects its distribution network.[13] For the remaining modes, exposure to sales tax does not vary across counties within a year.

Beyond differences in tax exposure, we allow for two dimensions of differentiation between online modes: quality and variety. We assume that consumers have CES preferences over modes where each mode offers a mass of varieties ω with distribution $F_{kt}(\omega)$. We treat ω as a quality index by assuming that the marginal utility of consumption of variety ω depends on a separable function of ω and consumer i's mode-specific taste for variety at time t, α_{ikt}. We assume that the offline mode offers a single variety.

We show in Online Appendix OA.1.1 that consumers' expenditure function is given by:

$$e_{ikt} = \int \alpha_{ikt}\omega\tilde{p}_{ikt}(\omega)^{1-\sigma}P_{it}^{\sigma-1}B_{it}dF_{kt}(\omega). \tag{1}$$

$\tilde{p}_{ikt}(\omega)$, the tax-inclusive price of variety ω on mode k, equals $p_{ikt}(\omega)(1+\tau_{ikt})$, where p_{ikt} is the pre-tax price and τ_{ikt} is the sales tax a resident in county i pays on purchases from mode k (see below). P_{it} is the tax-inclusive Dixit-Stiglitz price index (Dixit and Stiglitz, 1977).

We use consumer expenditure to derive aggregate revenue for Amazon, or mode 1, net of tax remitted to state and local taxing authorities, by summing across locations:

$$R_t(N_t) = \sum_i M_{it}\frac{e_{i1t}}{1+\tau_{i1t}(N_t)} \tag{2}$$

where M_{it} is the number of households. Similarly the number of orders originating from county i is given by:

$$Q_{it}(N_t) = M_{it}\frac{e_{i1t}}{(1+\tau_{i1t}(N_t))\bar{p}_t} \tag{3}$$

where $\bar{p}_t$ is the average pre-tax price of Amazon's varieties.

We impose the following assumptions on pricing and the distribution of orders across third-party and Amazon's direct sales.

Pricing Assumptions.

a. *The price of variety ω does not vary across counties.*

b. *The pre-tax price of variety ω is set competitively by Amazon and a continuum of third-party sellers according to a log-linear hedonic function:*

$$\ln p_{kt}(\omega) = \rho_{kt} + \ln\omega.$$

[13] In theory, the network configuration can also affect consumers' willingness-to-pay because of proximity to facilities. Houde et al. (2021) do not find evidence of this type of "local convenience", so we focus solely on the tax and fulfillment cost trade-off.

c. *The variety of goods is measured by the coefficient of variation of* $\ln \omega$*, normalized to one in 2018.*[14]

$$\mathit{Variety}_t = CV(\ln p_{1t})/CV(\ln p_{1,2018}).$$

d. *Amazon earns variable profit according to common revenue sharing agreements with manufacturers and third-party sellers.*

e. *Third-party sales are non-taxable and have a constant market share across consumer locations.*

Assumption a reflects that Amazon does not employ spatial price discrimination on its platform and such discrimination is not widespread on other platforms; thus the tax-inclusive price varies across locations solely due to differences in sales taxes.[15]

Assumption b allows us to treat prices as fixed, capturing that a large number of sellers are active on the platform, selling products that are available from a variety of other online and offline retailers. Together with Assumption c, it also allows us to measure the distribution of product variety using the distribution of item-level prices. The hedonic functional form of Assumption b is consistent with a model of product assortment that leads to an approximately log-normal distribution of variety. We use the coefficient of variation of the price of items on Amazon as a measure of the dispersion in available qualities. Finally, we normalize the variety index to one in 2018 to facilitate the interpretation of the model parameters.

We make Assumption d mostly for convenience since we observe limited information about the characteristics of products and contractual terms between Amazon and its suppliers. Assuming a common revenue-sharing formulation, together with Assumption e, allows us to write the revenue net of cost of goods sold as a linear function of gross revenue:

$$\text{Revenue Net Cost}_t = R_t(N_t) \times \bar{\mu}_t.$$

Here, $\bar{\mu}_t$ denotes the average markup Amazon earns on direct sales, μ_t^{own}, and on third-party sales, μ_t^{3PS}, weighted by the contribution of each channel to orders:

$$\bar{\mu}_t = s_t^{3PS}\mu_t^{3PS} + (1 - s_t^{3PS})\mu_t^{own},$$

where s_t^{3PS} denotes the share of third-party orders in year t.

Assumption e implies that consumers can always find an untaxed third-party seller, which allows us to recognize the effect of third-party sellers on local sales tax liabilities.[16] As a result, the sales

[14] We measure the coefficient of variation as the ratio of the predicted interquartile range to the predicted average price in each year to limit the importance of outlier prices on our variety measure.

[15] Evidence in Cavallo (2017) shows that over the period 2014-2016, online prices in the US do not change with the location of the consumer. The retailers covered in the analysis include Amazon's biggest competitors, such as target.com and walmart.com, who reportedly priced nationally during the full sample. Cavallo's work suggests furthermore that reported efforts at spatial price discrimination by specialized online retailers in the early 2010s, most notably Staples.com, were short-lived and temporary, presumably due to consumer backlash.

[16] This assumption entails an upper bound on the sales tax sensitivity. We have also conducted our analysis assuming

tax rate on each mode is given by:

$$\tau_{ikt} = \begin{cases} (1 - s_t^{3PS})1_{i1t}^{taxable}\tilde{\tau}_{it} & \text{if } k = 1 \text{ (Amazon)} \\ \tilde{\tau}_{it} & \text{if } k = 0, 2 \text{ (taxed offline and online retailers)} \\ 0 & \text{if } k = 3 \text{ (non-taxed online retailers)} \end{cases}$$

where $\tilde{\tau}_{it}$ is the sales tax rate in county i and year t and $1_{i1t}^{taxable}$ is an indicator variable equal to one if Amazon collects sales tax in county i in year t. In estimating the demand model, we use Amazon's observed physical presence and each state's observed sales tax abatement schedule to determine the tax rate residents of each county pay.

3.3 Order flow

The demand model allows us to predict the number of orders originating from county i. We use a simple logistic model to assign a given order to a fulfillment cluster, depending on the availability of goods across clusters and the distance between the clusters and the order's county of origin. The logistic model returns the flow of orders in the network, which we use to determine the variable shipping and order processing costs.

First, we assume that product availability at a cluster is determined by an IID binary random variable with probability that a product is available given by

$$\phi_t(K_{lt}) = 1 - \exp\left(-\psi \frac{K_{lt}}{\text{Variety}_t}\right). \tag{4}$$

The availability probability increases in the total fulfillment capacity of cluster l, K_{lt}, but decreases in the variety of goods available on Amazon in period t. Equation (4) thus implies that an increase in variety leads to a higher stock-out probability, giving Amazon an incentive to add fulfillment centers and thus capacity to a given location as variety increases.

We rule out the possibility of a network-wide stock-out and assume that orders are processed by the closest active cluster with available inventory. This leads to an $L \times I$ origin-destination fulfillment matrix $\Omega(N_t)$, where, in light of the large number of orders, element (l, i), $\Omega_{l,i}$, measures the fraction of orders from county i that are fulfilled by location l. As $\psi \to \infty$, the probability that the nearest active cluster fulfills a given order approaches one.

Conditional on orders being assigned to cluster l, we assume that they are distributed across the individual fulfillment centers that make up cluster l in proportion to facility size k_j.

Underlying our order assignment model is the assumption that conditional on fulfillment capacity, all fulfillment clusters and centers carry every variety with the same probability. In practice, while most fulfillment centers store "general merchandise" goods, some facilities are specialized in certain categories, such as large electronics. Unfortunately our demand and network data are not

that all third-party transactions are taxed, resulting in a lower-bound, but significant, sensitivity to sales-tax induced price variation. The true price sensitivity likely lies in-between these extreme cases, but Assumption e's is closer to the literature.

detailed enough to distinguish between types of goods. As a result, in the model, the fulfillment capacities of a cluster and individual facilities affect only the probability of product availability, but not the range of varieties available at the location and the facility. Thus, we allow for exogenous changes in product variety to influence Amazon's network expansion.

Last, we rely on the industry evidence on sortation center catchment areas discussed in Section 2.2, together with the assignment of orders to clusters, to determine orders that are fulfilled fully by Amazon. We summarize the share of orders from county i that is routed through a sortation center at cluster l by an $L \times I$ matrix $\Omega^{sc}(N_t)$. Element (l,i) of Ω^{sc} equals to the sum of orders from a household in county i, in percent, that are fulfilled by any cluster l', $\Omega_{l',i}$, that is within 25 miles of a sortation center l, provided the county falls in the 150 mile delivery catchment area of the sortation center at l.

We use these assumptions to predict the quantity of orders handled by facility j in period t as a function of the availability parameter ψ, the aggregate variety carried by Amazon, and the network:

$$q_{jt}(\psi) = \begin{cases} \sum_i Q_{it}(N_t)\Omega_{l,i}(N_t)\frac{k_j}{K_{lt}}1(a_j \leq t) & \text{if } m_j = FC \text{ and } l_j = l, \\ \sum_i Q_{it}(N_t)\Omega^{sc}_{l,i}(N_t)1(a_j \leq t) & \text{if } m_j = SC \text{ and } l_j = l. \end{cases} \tag{5}$$

Online Appendix OA.1.2 provides detail on the derivation of this function.

3.4 Cost function

The cost of fulfilling orders has three components: (i) shipping costs, (ii) labor costs in processing orders, and (iii) fixed costs of facilities consisting of rents and congestion costs.

We begin with the shipping cost, which has two components. First, is a 'distance cost' that is assumed to be a linear function of the delivery distance. Second, orders that Amazon routes through its own sortation centers, which we refer to as "vertically integrated", entail a constant, per-order reduction in shipping cost relative to non-integrated orders.

This leads to the following system-wide shipping cost function:

$$\begin{aligned} C_t^{\text{Shipping}}(N_t) &= \theta_d \cdot \left(\sum_l \sum_i \Omega_{l,i}(N_t) Q_{it}(N_t) d_{il} \right) + \theta_{vi} \cdot \left(\sum_j q_{jt} 1(m_j = SC) \right) \\ &= \theta_d \cdot D_t(N_t) + \theta_{vi} \cdot Q_t^{vi}(N_t) \end{aligned} \tag{6}$$

where d_{il} and $D_t(N_t)$ are the shipping distance in miles from county i to cluster l and in total, respectively, and $Q_t^{vi}(N_t)$ is the volume of vertically integrated orders. If $\theta_d > 0$, Amazon can reduce the cost of shipping to a particular county i either by opening up fulfillment centers in new locations, reducing the distance to the nearest cluster, or by expanding the capacity of the nearest cluster, thereby increasing the likelihood of product availability at the nearest cluster. If $\theta_{vi} < 0$, shipping costs can be further reduced by locating a sortation center near existing fulfillment capacity. This creates a complementarity in the choices of sortation center and fulfillment facility locations. Network density and vertical integration capture the idea that Amazon can lower shipping cost by reducing its reliance on independent shipping companies either through a lower distance that the shippers must cover, which in turn allows Amazon to put downward pressure on negotiated rates, or through in-house sorting of packages.

While not specifically modeled, the benefits from in-house sorting, rather than contracting with UPS or FedEx, stem from two sources. First, there are the cost efficiencies of bypassing the downstream shippers, who hold a significant amount of market power. Second, Amazon and the downstream shippers face a hold-up problem. UPS or FedEx would have to invest a significant amount in their network in order to handle Amazon's volume and to guarantee shipping times without congestion delays. The reputational concerns that shipping delays raise for Amazon provide the company with incentives to invest in its own network.[17]

We model the labor cost of processing using a Cobb-Douglas function, resulting in the following labor demand functions for each facility type:

$$L_{jt}(N_t) = A_{m_j} q_{jt}(\psi)^{\gamma} \tag{7}$$

where $q_{jt}(\psi)$ is given by Equation (5) and A_{m_j} is the productivity of a facility of type m_j. The total labor cost is obtained by aggregating across facilities:

$$C_t^{\text{Labor}}(N_t) = \sum_j \sum_l w_{lt} L_{jt}(N_t) 1(l_j = l), \tag{8}$$

where w_{lt} is the annual wage in cluster l. Importantly, if $\gamma < 1$ the production technology exhibits economies of scale. Therefore, in addition to locating facilities in low-wage areas, Amazon can lower labor cost by concentrating fulfillment capacity in a small number of fulfillment centers.

Last, the fixed cost of operating network N_t is given by:

$$\begin{aligned} F_t(N_t) &= \sum_j \sum_l k_j \cdot (r_{lt} + \kappa \text{Pop Density}_{lt})\, 1(l_j = l) \\ &= C_t^{\text{Rent}}(N_t) + \sum_j \sum_l k_j \cdot (\kappa \text{Pop Density}_{lt})\, 1(l_j = l). \end{aligned} \tag{9}$$

C_t^{Rent} is the fixed cost of space, which scales with a rental rate of r_{lt}. The parameter κ measures how the fixed cost per square foot increases as the population density of location l increases. Such additional penalties reflect either congestion or measurement error in rental rates, both of which are likely more pronounced for large facilities.

[17] The logistics company MWPVL provides a summary of the primary benefits of Amazon's vertical integration. See the article at http://www.mwpvl.com/html/amazon_building_new_sortation_network.html, accessed on 11/10/2018, for details.

3.5 Optimization problem

Putting together revenue and the various cost components yields the firm's flow profits associated with network N_t:

$$\pi_t\left(N_t\right) = \bar{\mu}_t R_t(N_t) - C_t^{\text{Shipping}}(N_t) - C_t^{\text{Labor}}(N_t) - F_t(N_t). \tag{10}$$

We are interested in understanding the trade-offs associated with the location choice of each new active facility, conditional on the number and characteristics of facilities built every year. Like Holmes (2011), we characterize the expansion of the network as the outcome of a constrained dynamic optimization problem with perfect foresight:

$$\begin{aligned} \Pi\left(\mathbf{a}^0\right) \quad &= \quad \max_{\mathbf{a}} \sum_{t=0}^{\infty} \beta^t \pi_t\left(N_t\right) \\ \text{s.t.} \quad & N_t = N_t^0(\mathbf{a}) \\ & \sum_j 1(a_j = t) = n_t^0 - n_{t-1}^0 \end{aligned} \tag{11}$$

where $\beta = 0.95$ is Amazon's discount factor and $n_t^0 - n_{t-1}^0$ is the observed number of facilities opened in period t.[18] The solution to this problem, which describes the optimal roll-out of the distribution network, is, under revealed preference, given by $\mathbf{a}^0$.

4 Empirical Analysis

In this section we discuss our econometric approach to estimating the above demand and cost functions and the empirical results. In Section 4.1, we discuss the estimation of the demand model, which we use to derive Amazon's total revenue, $R_t(N_t)$, and the geographic distribution of orders, $Q_{it}(N_t)$, as a function of the network. Section 4.2 focuses on the estimation of the order flow matrices, $\Omega(N_t)$ and $\Omega^{sc}(N_t)$, and the labor demand function, $L_{jt}(N_t)$. Finally, in Section 4.3, we lay out the estimation of the fixed costs and the shipping costs associated with fulfilling an order.

4.1 Demand

The CES model predicts mode-level spending by a county's representative household. From Equation (1) in Section 3, we construct a log-linear function describing the spending in county i on mode

[18] The perfect foresight assumption is common to the literature and is justified by the fact that Amazon is a large forward-looking company that utilizes sophisticated analyses to make predictions about the market. Therefore, it is reasonable to believe that Amazon makes predictions about future demand and supply conditions with limited error.

k relative to the offline option, mode 0:

$$\begin{aligned}
\tilde{e}_{ikt} &= \ln e_{ikt} - \ln e_{i0t} \\
&= \ln(\alpha_{ikt}) - ln(\alpha_{i0t}) + (1-\sigma)(\ln(\rho_{kt}) - \ln(p_{i0t})) \\
&\quad +(1-\sigma)\underbrace{(\ln(1+\tau_{ikt}) - \ln(1+\tau_{i0t}))}_{\Delta\tau_{ikt}} + \ln\left(\int \omega^{2-\sigma} dF_{kt}(\omega)\right) \\
&= \xi_{kt} + \lambda_k Z_{it} + \gamma_k C_{it} + \bar{\xi}_i + \Delta\xi_{ct} + (1-\sigma)\Delta\tau_{ikt} + \epsilon_{ikt} \qquad (12)
\end{aligned}$$

where ξ_{kt}, as a mode-year fixed effect, captures mode-level determinants of expenditure, including price, aggregate quality, and variety – the last term in the second line of Equation (12). The vector Z_{it} contains demographics of the representative household, and C_{it} contains variables that measure the level of local offline competition. The next two terms represent unobserved relative preferences for online shopping. We assume that these preferences consist of a county-level, time-invariant, component, $\bar{\xi}_i$, and a time trend that captures changes in preferences at the level of the county's Census Division, $\Delta\xi_{ct}$. Since we do not observe prices for the offline option, the time-varying controls and fixed effects account for county-level and time-series variation in prices at brick-and-mortar retailers in county i. $\Delta\tau_{ikt}$ is the difference in log sales tax rates between mode k and the offline option. Last, the model residual ϵ_{ikt} includes time-varying shocks to the willingness to pay of consumers for the three retail modes relative to the outside option, as well as measurement error in spending shares. See Online Appendix OA.1.1 for a step-by-step derivation of Equation (12). We estimate the model using weighted least-squares, based on the number of observations in the comScore sample for county i that we use to calculate $\tilde{e}_{ikt}$.

The demographic variables in Z_{it} are the income, age, and race of the representative household. We measure race as the share of people in county i who are of a given race. To allow for non-linearities in the effects of age and income, we also include the share of people in the county with incomes above $100,000 and the share of people in the county who are under the age of 35. The competition variables in C_{it} include the log of county i's population density as a proxy for travel costs and the number of all and of small (under 50 employees) retail establishments per 1,000 people in county i. We include an interaction between both Z_{it} and C_{it} and an Amazon mode indicator to capture varying preferences for Amazon by demographic group. These preferences could represent mode-specific targeting, heterogeneity in preferences for quality, variety, and convenience, or variation in price sensitivity.

The identification of the elasticity of substitution, σ, comes from variation in relative taxes across counties and time ($\Delta\tau_{ikt}$). The within-county variation exploits the timing of tax changes triggered by Amazon's expansion decisions and tax abatements, similar to a difference-in-difference regression. The intensity of the "treatment" via tax rates varies across counties within the same state, contributing to the identification of σ. Our main identifying assumption is that the timing and magnitude of tax changes are independent of changes in the demand residual, conditional on aggregate regional and mode-level trends.

Our expenditure model does not depend on the network through, e.g., the inclusion of the distance to the closest fulfillment cluster to county i. This would bias our estimate of σ if (a) changes in shipping distance are correlated with changes in relative tax changes, and (b) consumers value short shipping distances to the extent that they result in shorter shipping times. Note, however, that the tax changes do not necessarily coincide with a decrease in distance, as the closest fulfillment cluster may be located in a nearby state and as tax abatements drive a wedge between the timing of tax and shipping distance changes. In the short run, the entry of new facilities thus varies taxes independently of shipping distance.

In Houde et al. (2021), we exploit this feature of the data to measure the effect of local improvements in the network on Amazon's demand. We use multiple proxies of shipping speed from the closest fulfillment cluster to the consumer and cannot reject the hypothesis that consumer spending is independent of the proximity to an Amazon facility. We thus do not find any evidence that there are location-specific trends in the valuation of the platform that would be consistent with a local convenience effect. This is consistent with the limited variation in shipping times on Amazon in practice: since 2005, Amazon has offered the same shipping terms, free two-day shipping on eligible purchases, to all Prime members irrespective of location.[19] Instead, Equation (12) allows for uniform, mode-level improvements in shipping speed through the mode-level fixed effect, or a global effect of the increasing convenience of the platform to consumers.

Table 3: Demand Estimates

	Specification I: Homogeneous σ				Specification II: Heterogeneous σ			
		Est	SE			Est	SE	
Elasticity (σ):								
Constant		-1.516	0.399			-2.488	0.507	
% income 100k+						4.179	1.343	
	Incremental Amazon		Base		Incremental Amazon		Base	
	Est	SE	Est	SE	Est	SE	Est	SE
Demographics (λ_k):								
Age	0.013	0.008	-0.020	0.010	0.013	0.008	-0.020	0.010
% under 35	1.701	0.397	-0.043	0.402	1.677	0.397	-0.037	0.402
log(Income)	0.222	0.22	0.013	0.008	0.226	0.22	0.013	0.008
% income 100k+	0.257	0.395	-1.857	0.533	0.311	0.396	-1.741	0.534
% black	-0.361	0.077	-0.647	0.637	-0.365	0.077	-0.648	0.637
% asian	0.873	0.209	2.252	1.215	0.881	0.209	2.162	1.216
Offline Competition (γ_k):								
log(Pop density)	0.014	0.008	0.626	0.228	0.014	0.008	0.617	0.228
Retailers/pop	-0.380	0.307	-0.828	0.415	-0.376	0.307	-0.821	0.415
Small retailers/pop	0.387	0.317	0.894	0.426	0.382	0.317	0.89	0.425

We present the estimated parameters and standard errors of two specifications in Table 3.

[19] See https://money.cnn.com/2018/04/28/technology/amazon-prime-timeline/index.html, accessed in Oct 2019.

Specification I restricts σ to be the same for all consumers, while Specification II allows σ to vary with the income of the representative consumer. We interact the relative tax rate with the share of households in county i who have an income above $100,000.

For each specification, we estimate a base effect of demographics and offline competition on the demand for all modes relative to the offline option. We also allow for an incremental effect of the covariates on the demand for Amazon. The estimated coefficients are consistent across specification, so we focus our discussion on the heterogeneous-σ model.

The estimates suggest that the propensity to shop online falls with age and that households with a head under the age of 35, in particular, prefer to shop on Amazon. Log-income does not have any significant linear impacts on spending, but high-income households have lower preferences for online shopping. Asian households prefer online shopping, and in particular shopping on Amazon, while black households shop disproportionately offline. Online shopping does not vary with population density, but the total number of offline competitors decreases spending across all three modes. However, if those retailers are small, then the effects of competition disappear.

The estimated elasticity of substitution in the homogeneous model, which is approximately the own price elasticity in a CES model for choices with small shares, is a precisely estimated -1.52.[20] This result is in line with the findings of Einav et al. (2014) and Baugh et al. (2018). The magnitude of σ suggests that a move from no taxes to the average tax rate of 6.5% results in a decrease in demand of about 10% (6.5 times 1.52). The results of the heterogeneous specification indicate that the price sensitivity is lower for high-income households. Specifically, a county at the 25$^{\text{th}}$ percentile, with 8% of households having incomes above $100,000, has an elasticity of -2.21, while a county with 15% of high-income households, the 75$^{\text{th}}$ percentile, has an elasticity of -1.89. We investigate the robustness of our estimated elasticity to alternative spending measures and demand specifications in Online Appendix OA.3.

With the estimates of the heterogeneous demand model in Table 3, we can calculate Amazon's total revenue (Equation 2) and orders (Equation 3) for each year from 1999-2018 as a function of the network configuration. To do so, we first generate the expenditure function for each county and year outside of the comScore sample using a combination of extrapolation, imputation, and data from Amazon's financial statements. Appendix B.1 describes this process. We then use the expenditure function to predict the number of orders for each county and year by dividing predicted expenditure by a yearly price index for goods sold on Amazon. See Appendix A.6 for details on how we construct the index. To validate this approach, we compare the predicted orders for the representative household in each county to orders calculated directly from the comScore data for the subset of periods and counties of overlap; the correlation is 0.65.

Finally, to illustrate the importance of the estimated elasticity, we use the estimated Equation (2) to calculate Amazon's counterfactual revenue assuming a zero tax rate on Amazon transactions throughout the sample and compare it to the firm's predicted revenue under its actual tax obligations. Few states collected taxes before 2006, and so the effect of taxes is growing over time,

[20] The maximum spending share of any of the online modes is under 5%.

especially after 2012. The loss in revenue under the actual tax path, relative to a world with zero taxes, amounts to approximately \$9 billion by the end of the sample in 2018, or 4% of sales, suggesting that changes in Amazon's sales tax liability significantly impact revenue.

4.2 Order flow and labor demand

In this section, we discuss the estimation of the first set of cost parameters that enter (i) the availability probability (ψ) and (ii) the labor demand function (A_{fc}, A_{sc}, γ). Since we do not observe the flow of orders across the network directly, we identify the parameters by combining the predicted distribution of orders from the demand model with data on employment across facilities.

We derive our main estimating equations by substituting the predicted flow of orders, $\hat{Q}_t(N_t)$, into Equation (5), resulting in the predicted number of orders for each facility j in year t, $\hat{q}_{jt}$. Plugging $\hat{q}_{jt}$ into Equation (7) yields the predicted employment of facility j in year t, conditional on its size k_j and location l:

$$L_{jt}(N_t|\theta^1) = A_{m_j}\hat{q}_{jt}(\psi)^{\gamma} \tag{13}$$

The origin-destination probability matrices, which enter $\hat{q}_{jt}$, depend on the product availability parameter ψ. Let $\theta^1 = (A_{fc}, A_{sc}, \gamma, \psi)$ denote the vector of parameters determining labor demand.

As noted in Section 2, we observe the number of employees for most fulfillment and sortation facilities in 2017 (denoted by $\hat{L}_{j,2017}$ for facility j) and system-wide employment in 2017 (denoted by $\hat{L}_{.,2017}$). We use both sources of information to construct a minimum-distance estimator:

$$\min_{\theta^1}\sum_j \Big(L_{j,2017}(N_{2017}|\theta) - \hat{L}_j\Big)^2 + \Big(L_{.,2017}(N_{2017}|\theta) - \hat{L}_{.,2017}\Big)^2. \tag{14}$$

Heuristically, the parameters are identified as follows. The two productivity parameters (A_{fc}, A_{sc}) enter linearly in the labor demand function and are therefore identified from the covariation in observed employment across facilities, by type, and the predicted number of orders. The returns to scale parameter (γ) is identified from covariation across facilities in capacity and predicted orders, which translates into differences in predicted orders per employee. Finally the order flow parameter (ψ) is identified from the covariation in the spatial distribution of employees and orders predicted by the model.

We report the estimates of θ^1 for two specifications in Table 4a and summary statistics for three years of our sample calculated using our main model in 4b. The 'Stockout' specification is our main model and assumes that orders originate from the closest fulfillment center with availability. We also report estimates for a specification where the order originates from the closest facility, abstracting from availability. The estimate of γ is significantly less than one in both specifications, implying that the production function exhibits increasing returns to scale. Increasing returns to scale in fulfillment is expected since the process of packaging orders is largely automatized and relies on robotic technologies.

Table 4: Order Flow Model Estimates

(a) Estimates summary

	Specification I: Stockout		Specification II: Nearest		Goodness of Fit			
						Data	Specification	
	Est	SE	Est	SE		Est	I	II
Availability (ψ)	0.49	0.21	–		OLS: $Y =$	$\ln L$	$\ln \hat{L}$	$\ln \hat{L}$
Orders (γ)	0.47	0.10	0.31	0.08	Intercept	5.39*	6.22*	6.40*
A_{fc}	1.81	0.75	3.52	1.00	FC x Dens.	0.25*	0.11*	0.08*
A_{sc}	0.41	0.17	0.68	0.25	FC x k_j	0.64*	0.71*	0.38*
					SC	0.21	-0.62*	-0.80*
SSR	37.62		40.08		2017 $\ln L$	4.83	4.92	4.93

(b) Summary statistics of the flow of orders

Years	Ave Shipping Distance (miles per order)	VI Orders (%)	Fullfillment prob. Closest	2nd Closest	3rd Closest
2006	450.43	0.00	0.48	0.26	0.17
2012	303.82	0.02	0.65	0.21	0.08
2018	141.80	0.37	0.51	0.23	0.12

Notes: The goodness-of-fit regression results are from OLS regressions of log employment at a facility on local shifters of demand and facility characteristics. The * indicates the coefficient is significant at the 5% level. The last row of Table 4a shows the distribution of log employment in the data and that predicted by the model. The summary statistics in the bottom panel are calculated using parameters from Specification I.

The availability parameter, which drives the allocation of orders to facilities, is equal to 0.49. Rows 3-5 of Table 4b show that across years, this value implies that almost 90% of orders come from one of the three closest facilities, with roughly 50% coming from the closest in 2018. Due to expansion in variety and a resulting higher probability of stockouts, delivery from the closest facility is slightly declining over time, despite expansion in capacity. Comparing Table 4b and Table 1 shows that this manifests itself in an average shipping distance per order of 142 miles in 2018, while the average distance to the closest fulfillment center is only about 70 miles in 2018. Nevertheless, the expansion of the network of fulfillment centers led to a substantial decrease in the average shipping distance over the period 2006 to 2018, from 450 to 142 miles. The expansion of the sortation center network also led to an increase in the share of orders that Amazon handles fully in-house; the model predicts that in-house sortation applies to 37% of orders in 2018. This matches estimates from outside sources, which suggest that in 2014, at most 40% of Amazon's orders were sorted in-house.[21]

The last three columns of Table 4a evaluate the goodness-of-fit of the models. We report results of three OLS projections of observed (labeled 'Data') and predicted 2017 facility employment on the type of the facility and, for fulfillment centers, the capacity of the facility and the population

[21] See `https://nypost.com/2017/12/29/trump-says-amazon-is-making-us-postal-service-dumber-and-poorer/`, which states that USPS handled 40% of Amazon's shipments in 2014. Because USPS' agreement with Amazon covers primarily last-mile delivery, this estimate should be the maximum share of shipments that go through Amazon's sortation centers.

density of the surrounding area. The asterisk indicates coefficients that are significant the 5% level. The last row displays actual and predicted average log employment calculated using the regression coefficients.

Overall, both models fit the data well, in particular in relating fulfillment center capacity to labor. The only coefficient that differs from the data significantly is the one on the SC indicator, which is negative and significant when using the predicted employment but positive and insignificant in the data. The difference is likely due to the fact that the data is limited in both the number of sortation centers for which we observe employment and the variation in observed employment across sortation centers. In contrast, the model predicts rich variation in employment across the entire set of sortation centers. Comparing the two models suggests that the 'Stockout' model is a better fit. As an additional validity check, we compare the year-over-year growth of system-wide employment predicted from the model to a measure of growth calculated from Amazon's financial statements, and we are able to match the observed growth well.[22]

4.3 Shipping and fixed costs

The parameters of the cost function that remain to be estimated are the ones that enter the shipping cost and fixed cost functions: $\theta^2 = \{\theta_d, \theta_{vi}, \kappa\}$, following the functional forms in Equations (6) and (9). We take a revealed preference approach by finding parameter values that render the observed network more profitable than alternative, perturbed networks. This leads to a moment inequality estimator.

Approach. We focus on alternative network roll-outs where we swap the opening dates of two facilities to construct revealed-preference inequalities. Under our perfect foresight assumption, changing the opening date of facility j to that of facility j', and vice-versa, holding the facility's other characteristics of location and size fixed, must result in a lower discounted net-present value of profits than under the observed network $\mathbf{a}^0$:

$$\Pi(\mathbf{a}^0; \theta^2) - \Pi(\mathbf{a}^{j,j'}; \theta^2) \geq 0$$

where $\Pi(\mathbf{a})$ is defined in Equation (11). We use three criteria to select potential entry date swaps: (i) facility j opened more than one year before facility j', (ii) facilities j and j' are of the same type (FC or SC), and (iii) the difference between the sizes of j and j' is less than 550,000 sq ft, the inter-quartile range of capacity differences. This leads to $M = 5,577$ potential permutations.

Importantly, the condition that strategy $\mathbf{a}^0$ yields higher profit than $\mathbf{a}^{j,j'}$ only depends on the profit flow differences between the entry dates of facilities j and j', which we can calculate without having to solve the infinite horizon dynamic programming problem. In contrast, counterfactual network roll-outs that involve locations that Amazon does not choose in our sample lead to a

[22] See Appendix A.2 for how we construct the measures of system-wide employment. Based on a regression of system-wide employment on a time trend, the year-over-year growth in the data is 0.24, and the estimate using our predictions is 0.21, though the latter is not significant.

network configuration that differs from Amazon's chosen one by the end of the sample. In this case we would need to make additional assumptions about Amazon's expectation regarding future market conditions post-sample including, for instance, strategic deterrence considerations regarding the entry of potential e-commerce rivals. By focusing only on deviations involving observed facility locations, our estimation results are robust to the presence of these dynamic considerations, while exploiting the significant cross-sectional variation in the attributes of chosen locations.

We decompose the value function differences, $\Pi(\mathbf{a}^0;\theta^2) - \Pi(\mathbf{a}^{j,j'};\theta^2)$, into a return function $\Delta\Pi(\mathbf{a}^0,\mathbf{a}^{j,j'};\theta^2)$ of predicted differences and an unobserved error associated with swap (j,j'), $\epsilon^{j,j'}$, capturing deviations of the true value function difference from our prediction. This leads to the following inequality condition:

$$\Pi(\mathbf{a}^0;\theta^2) - \Pi(\mathbf{a}^{j,j'};\theta^2) = \Delta\Pi(\mathbf{a}^0,\mathbf{a}^{j,j'};\theta^2) + \epsilon^{j,j'} \geq 0. \tag{15}$$

Separating the contribution of changes in fulfillment cost from the remaining components of the return function $\Delta\Pi$ highlights that the return function is linear in θ^2, reflecting the linearity of the fulfillment cost function:

$$\Delta\Pi(\mathbf{a}^0,\mathbf{a}^{j,j'};\theta^2) = Y^{j,j'} - (\theta_d X_d^{j,j'} + \theta_{vi} X_v^{j,j'} + \kappa X_p^{j,j'}) \tag{16}$$

Here, $Y^{j,j'}$ is the difference in discounted gross profit net of wages and rents between the chosen and counterfactual networks:

$$\begin{aligned} Y^{j,j'} \quad = \quad & \sum_{t=t(j)}^{t(j')} \beta^t \left(\hat{R}_t(N_t|\mathbf{a}^0) - \hat{C}_t^{\text{Labor}}(N_t|\mathbf{a}^0) - C_t^{\text{Rent}}(N_t|\mathbf{a}^0)\right). \\ & - \sum_{t=t(j)}^{t(j')} \beta^t \left(\hat{R}_t(N_t|\mathbf{a}^{j,j'}) - \hat{C}_t^{\text{Labor}}(N_t|\mathbf{a}^{j,j'}) - C_t^{\text{Rent}}(N_t|\mathbf{a}^{j,j'})\right) \end{aligned} \tag{17}$$

where a hat indicates that the function was estimated in a previous step. $Y^{j,j'}$ thus measures the net effect of the network configuration on gross revenue through tax changes and on wages and rents through adjustments in the timing of location choices.

Similarly, the term $X_d^{j,j'}$ is the discounted difference in total shipping distance, calculated as:

$$X_d^{j,j'} \quad = \quad \sum_{t=t(j)}^{t(j')} \beta^t \left(\hat{D}_t(N_t|\mathbf{a}^0) - \hat{D}_t(N_t|\mathbf{a}^{j,j'})\right). \tag{18}$$

We define the differences in the discounted sum of vertically integrated orders, $X_{vi}^{j,j'}$, and in the discounted sum of our congestion proxy - population density scaled by facility square-footage - $X_p^{j,j'}$, analogously (See Online Appendix OA.2.1).[23]

[23] In considering counterfactual networks, we abstract from changes in the cost of tying manufacturers into the network as manufacturers already serve the many locations of retailers like Wal-Mart and Target. We also assume that the

In calculating the components of the return function, we rely on the previously estimated demand and order flow models to predict revenue, wages, and rents under the two networks in the years between the swapped facilities' opening dates. This entails predicting Amazon's revenue and total number of orders in each county and the assignment of these orders to fulfillment centers for the observed and counterfactual networks during the relevant time period.

To predict the consumer's tax exposure under the two alternative networks, we assume that the first entry into a state triggers the nexus laws' physical presence rule, but that the tax status of consumers in all counties adjusts only after a period of tax abatement. We assume a deterministic schedule for the abatement period depending on the year of first entry of (i) five years if $t < 2008$, (ii) two years if $2008 \leq t \leq 2010$, and (iii) immediate if $t > 2010$.[24]

Estimator Set-up. The residual value function difference, ϵ, arises from various potential sources: measurement error in the demand model, unobserved fixed cost components, including unobserved subsidy payments by state and local governments, or mis-specification of the firm's beliefs regarding sales tax changes. We focus on the interpretation of ϵ as measurement error. One challenge is that ϵ is potentially correlated with both gross profit differences $Y^{j,j'}$ and, through variable cost channels, differences in distance $X_d^{j,j'}$ and vertically integrated orders $X_{vi}^{j,j'}$.

We address this simultaneity problem by constructing a vector of H non-negative instrumental variables $Z^{j,j'}$ that are correlated with changes in the profit components, but uncorrelated with $\epsilon^{j,j'}$. This allows us to consistently estimate θ^2 using the following moment inequalities conditions:

$$E\left[Z_{j,j'} \cdot \left(Y^{j,j'} - (\theta_d X_d^{j,j'} + \theta_{vi} X_{vi}^{j,j'} + \kappa X_p^{j,j'})\right)\right] + \underbrace{E\left[Z^{j,j'} \cdot \epsilon^{j,j'}\right]}_{=0} \geq 0. \tag{19}$$

Following Pakes et al. (2015), we use this condition to construct sample moment inequalities:

$$\frac{1}{M}\sum_{j,j'} Z_h^{j,j'} \cdot \underbrace{\left(Y^{j,j'} - \left(\theta_d X_d^{j,j'} + \theta_{vi} X_{vi}^{j,j'} + \kappa X_p^{j,j'}\right)\right)}_{\Delta\Pi(\mathbf{a}^0, \mathbf{a}^{j,j'}; \theta^2)} = \tilde{m}_h(\theta^2) \geq 0, \forall h = 1, \ldots, H. \tag{20}$$

To construct valid moment conditions, we assume that the econometric error $\epsilon^{j,j'}$ is mean zero and independent of the sequence of predetermined demand and cost shifters that enter profits, including county demographic characteristics on the demand side and county wages, rents, and sales tax rates on the cost side.

We use this assumption to construct proxies for the gross profit and cost components entering the return function $\Delta\Pi(\mathbf{a}^0, \mathbf{a}^{j,j'}; \theta^2)$ that are orthogonal to $\epsilon^{j,j'}$. We calculate population-weighted (rather than demand- or revenue-weighted) changes in distance and vertical integration associated

demand-side aggregate platform quality ξ_{kt} does not respond to counterfactual changes in network density. Most swaps entail small changes in average delivery distances and thus shipping speeds, and our short time-series does not allow us to identify the effect of network density on platform quality.

[24] We recognize that in New York, Amazon remitted sales tax prior to opening any fulfillment center, and that in North Dakota and Washington State, Amazon operates non-logistic facilities that triggered sales tax liabilities.

with each swap (j, j'). Let $\hat{X}_d^{j,j'}$ and $\hat{X}_{vi}^{j,j'}$ denote these pre-determined proxies for the order-weighted distance and vertical integration variables.

Similarly, we use measures of population-weighted average tax and cost differences as shifters for gross profit differences $Y^{j,j'}$. We measure cost differences using changes in average input prices and population density across active locations induced by swap (j, j'), $\Delta\text{Input prices}^{j,j'}$ and $\Delta\text{Density}^{j,j'}$, respectively. We calculate tax differences, $\Delta\text{Tax}^{j,j'}$, using population weighted average tax rates under the observed and counterfactual roll-out strategies. See Appendix B.2 for details on these variables.

Identification based on Swap Groupings. We use these proxies to construct H categorical instruments that indicate whether a particular swap (j, j') informs the different economic trade-offs Amazon faces.

To understand how such groupings of swaps facilitates parameter identification, consider first a case where fulfillment cost depends only on shipping distance and $\theta_{vi} = \kappa = 0$. We observe two types of decisions that affect distance to the customer: enter early in a densely populated location or enter late in the same type of location. To construct moments that explain these decisions, the instruments must capture the trade-off between changes in the proximity to final consumers ($\hat{X}_d^{j,j'}$) and changes in gross profit ($Y^{j,j'}$).

Consider first instances where the firm chose to open a fulfillment center in a densely populated area early and open a comparable fulfillment center in a less densely populated area late. The firm's profit under this chosen network roll-out must exceed its profit under the alternative roll-out where we swap the opening dates of these two facilities. We pair fulfillment center j with all fulfillment centers j' such that swapping each resulting pair's opening dates yields the following changes in profit shifters and distance. First, the difference in population-weighted distance ($\hat{X}_d^{j,j'}$) is negative; the perturbed network delays the expansion into a densely populated area, and therefore, increases the aggregate shipping distance. Second, the difference in population-weighted tax rates ($\Delta\text{Tax}^{j,j'}$) is positive; the firm delays moving into areas with higher tax rates and with a larger population that has to pay sales tax. We categorize swaps that satisfy both conditions using an indicator variable $Z_h^{j,j'} = \mathbf{1}\left(\hat{X}_d^{j,j'} < 0 \text{ and } \Delta\text{Tax}^{j,j'} > 0\right)$.

To the extent that the tax changes associated with these swaps are associated with negative gross profit differences on average, the following moment restriction identifies a lower bound for θ_d:

$$E\left[Y^{j,j'} - \theta_d X_d^{j,j'} | Z_h^{j,j'}\right] + E\left[\epsilon^{j,j'} | Z_h^{j,j'}\right] \geq 0 \longrightarrow \theta_d \geq \frac{E\left[Y^{j,j'} | Z_h^{j,j'}\right]}{E\left[X_d^{j,j'} | Z_h^{j,j'}\right]}. \tag{21}$$

where $E\left[X_d^{j,j'} | Z_h^{j,j'}\right] = E\left[X_d^{j,j'} | Z_h^{j,j'} = 1\right]$ is the average total difference in shipping distance conditional on belonging to the group $Z_h^{j,j'} = 1$. The numerator $E\left[Y^{j,j'} | Z_h^{j,j'}\right]$ is defined analogously. Intuitively, these type of swaps determine the lowest level of θ_d such that the shipping costs savings from opening in a populated area outweigh the lost revenue due to taxes. An upper bound can be constructed using the opposite trade-off, where we select network swaps such that the firm's actual entry decision into a densely populated, high-tax area occurs late and we compare this network to counterfactual networks where the firm enters in densely populated, high-tax areas early.

Since the gross profit differences $Y^{j,j'}$ net out wages and rents, we can construct similar moment conditions by exploiting the trade-off between distance and input prices across locations. For instance, if we observe the firm entering early in areas with higher wages or rents, the change in shipping cost savings from opening in these areas must outweigh the net profit declines due to higher wage or rental bills.

Table B-1 in Appendix B.2 defines the set of moment conditions we use in estimation. Our swap groupings $Z^{j,j'}$ capture seven "trade-offs", leading to fourteen lower and upper-bound moments. We use six instruments to capture the trade-offs in the timing of the network roll-out induced by nexus tax laws. The first two instruments group swaps that trade off tax and distance. We similarly construct lower- and upper-bound instruments by grouping swaps that (a) the capture tax and vertical integration trade-off (e.g., moving up the opening of a sortation center reduces shipping cost, but increases tax exposure), and (b) the tax trade-off alone, unconditional of changes in shipping cost associated with changes in the timing of a facility opening. Similarly, we construct six instruments to capture the above trade-offs between higher input cost bills, distance, and vertical integration. Finally, we use two instruments that capture the trade-off in the network roll-out between fixed cost savings from lower congestion, proxied by population density, and the distance to populated areas.

Table 5: Moment Conditions and Distance Trade-offs

	Lower bound: θ_d $Z^{j,j'} = 1(\Delta\text{Shifter}^{j,j'} > 0 \ \& \ \hat{X}_d^{j,j'} < 0)$		Upper bound: θ_d $Z^{j,j'} = 1(\Delta\text{Shifter}^{j,j'} < 0 \ \& \ \hat{X}_d^{j,j'} > 0)$	
	$E(Y\|Z)$	$E(X_d\|Z)$	$E(Y\|Z)$	$E(X_d\|Z)$
ΔShifter (Gross Profit)				
(a) Tax	-13.30	-93.60	37.10	140.19
(b) Input prices	-5.94	-82.68	7.60	128.21
Bounds: $\frac{E(Y\|Z)}{E(X_d\|Z)}$				
(a) Tax	0.14		0.26	
(b) Input prices	0.07		0.06	

Notes: In selecting swaps for inclusion in each instrument category, we condition on population-weighted tax, input price, and distance changes. The statistics in the body of the table, however, represent order-weighted aggregates. The variable ΔShifter refers to the change in one of two population-weighted profit shifters: taxes and average input prices.

Interim Estimates. Above, we motivate the value of grouping swaps in identifying the parameters of interest based on the example of the tax and shipping distance trade-off associated with adjusting the opening date of a facility. Before discussing the estimates of the full model, we use this example to derive initial estimates of the lower and upper bound for the distance coefficient, θ_d, using swaps that isolate the tax and distance trade-off. We also provide estimates of these bounds using swaps that isolate the trade-off between higher input costs and shorter shipping distances. The results of this exercise are displayed Table 5. In the first row of the table, we report the change in discounted gross profit (Y) and discounted aggregate shipping distance (X), averaged across swaps that capture the economic trade-off between distance and taxes. To isolate this trade-off, we only include swaps in this subset where the other profit components (e.g., population, input prices, etc.) are 'fixed'. In practice, we condition on swaps that exhibit small differences in these other components relative to the focal fulfillment center j.[25]

The first entry in the first row of the table shows that the change in discounted gross profits, averaged over swaps that entail a decrease in the population-weighted average tax rate and an increase in population-weighted distance relative to the observed network, is $-\$13.30$ million. The average change in the discounted aggregated shipping distance for the same subset of swaps is -93.60 hundred million miles. Per the discussion above, these swaps identify the lower bound of the cost parameter. Using the intuition of Equation (21), we calculate this bound as the ratio of the average change in gross profits to the average change in distance, holding all remaining exogenous cost contributions approximately fixed. The resulting estimate, presented in the lower portion of the table, suggests that the lower bound of θ_d is \$0.14 per 100 miles. Moving to the right in the first row, we perform the same exercise, but this time we condition on swaps that increase the population-weighted average tax rate and decrease the population-weighted distance relative to the observed network. These swaps identify the upper bound of the distance parameter, which we calculate to be \$0.26 per 100 miles.

Similarly, the second row of Table 5 presents the average statistics for swaps that feature changes in population-weighted input prices, rather than taxes, and distance. The bottom of the table displays the bounds of θ_d calculated based on theses swaps, with \$0.07 being the lower bound and \$0.06 being the upper bound. Therefore, at the midpoint of the smallest lower bound and highest upper bound, a 100-mile increase in distance raises the average shipping cost by \$0.17.[26]

Besides providing initial estimates of the bounds, this exercise also demonstrates that the instruments are inducing the expected trade-off between distance and gross profits. For example, swaps that feature a decrease in taxes and an increase in population-weighted distance relative to the observed network (first row, left side), result in a decrease in gross profit and an increase in aggregate order-weighted shipping distance. The population-weighted proxy variables that define the instruments (i.e., $\Delta\text{Tax}^{j,j'}$ and $\hat{X}_d^{j,j'}$) thus correctly predict a positive correlation between changes in gross profit and shipping distance. This is true for the other instruments in Table 5 as well.

We note that the distance and gross profit trade-offs in Table 5 generate only four of the fourteen moment conditions discussed above. We exploit five additional trade-offs in constructing

[25] We select swaps such that the value of each of the other variables falls within the variable's interquartile range from the focal fulfillment center's realization.

[26] Note that the estimated lower bound is not necessarily below the estimated upper bound due to sampling error in the moments and our inability to perfectly hold fixed the remaining exogenous cost components, which are correlated with distance.

the remaining moments that identify θ_{vi} and κ and aid in identifying θ_d. We repeat the exercise behind Table 5 for the remaining trade-offs and present preliminary estimates of θ_{vi} and κ in Table OA-1. In the following section, we present the results of the full model, which differs from this exercise in that we jointly estimate the cost function parameters using information contained in all the groupings of swaps that define the fourteen moments.

Full Model Results. We present the estimates of the full model in Table 6. The column labelled "Est." corresponds to the parameter estimate that minimizes the objective function. In all cases this is a single point because the moment conditions are not jointly satisfied in our sample. To conduct inference we need to account for the fact that the model parameters are partially identified, and the literature suggests several approaches for doing so. Since our main specification includes multiple parameters, we construct confidence intervals based on the "profiled test statistic" approach proposed by Bugni et al. (2017).[27] We construct the confidence interval of each individual parameter by testing repeated null hypotheses that the parameter is equal to a range of candidate values. We define the confidence interval of each parameter as the set of values such that the null hypothesis cannot be rejected at the 5% confidence level. The resulting confidence interval is therefore constructed using the marginal distribution of each parameter, effectively "profiling-out" the other two parameters.

Table 6: Cost Function Estimates

	Specification 1			Specification 2			Specification 3		
	Est.	CI		Est.	CI		Est.	CI	
θ_d: Dist. (x100 miles)	0.16	0.15	0.17	0.59	0.41	0.88	0.34	0.26	0.49
θ_{vi}: VI orders							-0.52	-0.91	0.01
κ: Density (x100)				2.06	1.48	3.17	0.98	0.69	1.56
Moments	4			8			14		

Notes: "Est." denotes the parameter value at which the objective function is minimized and CI is the 95% confidence interval calculated as described in the text. All specifications utilize 5,577 total swaps.

We present results for three specifications of the cost function here and relegate estimates under alternative assumptions on the demand and order-flow models to Online Appendix OA.3. The first specification corresponds to a model where shipping cost depends only on distance and where the logistic network's fixed costs consist only of observed rents. This leads to an estimated shipping cost per 100 miles of \$0.16, which is similar to the midpoint of the preliminary bounds presented above. Controlling for the effect of density on fixed costs in Specification (2) and then also the cost savings to vertical integration in Specification (3) substantially increases the estimated shipping cost, however, suggesting omitted variables biases in Specifications (1) and (2). The change in θ_d

[27] To account for correlation between swaps, we estimate the empirical correlation between moments sharing the same facility choice, evaluated at first-stage parameter estimates that assume zero correlation. We use this correlation (0.3) when sampling shocks in the parametric bootstrap procedure described by Bugni et al. (2017). See also Kaido et al. (2019) for a related approach to constructing profiled test statistics.

in going from Specification (1) to (2) reflects that opening a facility in high-density areas not only reduces distance to the consumers, but also increases fixed fulfillment costs. The change in θ_d from Specification (2) to (3) reflects that opening a fulfillment center in an urban area also serves the purpose of vertical integration. Our preferred Specification (3) suggests average estimated shipping cost of $0.34 per 100 miles.

Under Specification (3), we also find significant cost to density; our estimate of κ implies that rents account for only approximately one half or less of the fixed costs to locating in an urban area with a density of 1,000 people per square mile. We interpret this as evidence that traffic congestion in urban areas increases the fixed cost of managing large fulfillment centers.

Finally, we estimate cost savings of $0.52 per order from vertical integration into sortation. To put this estimate into perspective, consider that the order-flow model predicts an average shipping distance of 303 miles in 2012. The variable component of the shipping cost for this average order without an integrated sortation facility is $1.02, compared to $0.50 with vertical integration.

Predicted Order Fulfillment Costs. To illustrate our model results, we analyze trends in the implied average order fulfillment costs. Table 7 summarizes the evolution of average cost over time separately by each cost component. We measure average cost by dividing each cost component by the predicted quantity of orders. In this exercise and all of the following analyses, we utilize the point estimates from Specification (3) in Table 6.

Table 7: Average Cost Decomposition

			Average cost components				
	FC	SC	Shipping	Labor	Rent	Density	Total
2000	5	0	1.99	0.55	0.12	0.05	2.71
2003	7	0	1.61	0.60	0.11	0.08	2.41
2006	12	0	1.55	0.62	0.11	0.13	2.40
2009	16	0	1.39	0.53	0.08	0.10	2.11
2012	31	1	1.03	0.43	0.07	0.07	1.60
2015	67	21	0.50	0.51	0.09	0.19	1.28
2018	128	35	0.29	0.51	0.08	0.22	1.11
2018*	128	0	0.49	0.48	0.07	0.20	1.23

Notes: 2018* corresponds to a counterfactual network with no sortation centers. Average cost components are calculated as the total network cost divided by the total orders predicted by the model.

Investments in the logistic network led to a large decrease in shipping cost from nearly $2 per unit in 2000 to $0.29 in 2018. The drop was most pronounced between 2009 and 2015, a period during which Amazon quadrupled the number of fulfillment centers. By 2015, shipping no longer represents the largest component of average order fulfillment costs.

Much of the drop in shipping cost from 2015 to 2018 is due to the build-up of the sortation network. In the last row, we calculate the average cost the firm would have incurred in 2018 in the absence of sortation facilities. Eliminating sortation centers, which are located primarily in relatively urban locations, would have increased average shipping cost by 69% (from 0.29 to 0.49), but decreased rents plus wages by 7% (from 0.81 to 0.75), resulting in an overall average cost increases of 11% (or $0.12).

The labor cost per order remains largely constant over the period, reflecting that the expansion in the volume of transactions increases economies of scale in order processing even as facilities are added. These scale economies offset the increase in employment and higher wages to workers, as fulfillment center openings during this period took place in higher wage counties. Turning to fixed costs, the combined rent and density costs per order increased from 2009 to 2018 due to expansion into more urban areas.

Overall, Amazon was able to decrease total average cost by 55%, as seen in the final column. We note that our order fulfillment cost estimates are lower than ballpark figures collected from outside sources.[28] While such measures are not directly comparable to ours for a number of reasons, they point to the fact that our estimates are likely a lower bound on the fulfillment cost. One reason for this is that our moment inequalities estimator is only able to capture costs that vary across the locations in the network. We are therefore not able to, e.g., identify a constant base cost of shipping a package (i.e., a cost function intercept) or the cost contribution of system-wide investments in robotics. Another reason is that our model under-estimates the total revenue Amazon earns from third-party sellers as we assume a constant mark-up; our cost estimate is proportional to revenue.

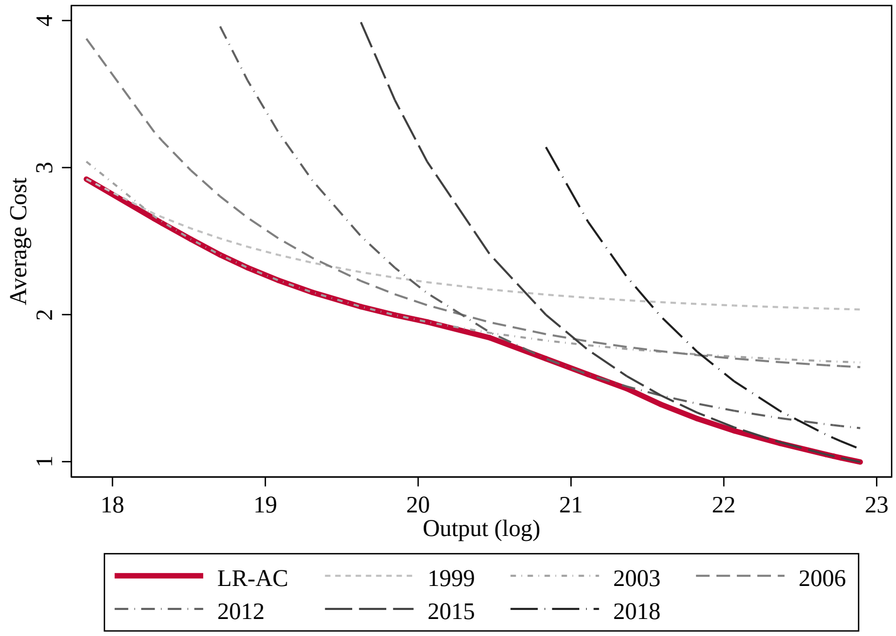
We note, however, that it is not the level of costs that is key to our analysis, it is the trends therein and the relative magnitude of fulfillment costs to the other profit components that ultimately drive Amazon's expansion. The fact that, as we illustrate in Section 5, our model predicts the evolution of the distribution network suggests that we are able to capture these two features of the firm's cost.

Figure 4 illustrates the importance of economies of scale by plotting the implied long-run average cost function across observed network configurations from 1999 to 2018. Each grey curve represents the short-run average cost function calculated with a given network configuration; more recent networks are represented with darker colors. The x-axis represents a grid of (log) aggregate output covering the range of orders that we observe during out sample, extended by 25% beyond the 2018 level. The lower envelope (red) represents the long-run average cost function, or the most efficient technology associated with each output level. The elasticity of the long-run total cost function with respect to output is approximately equal to -0.8.

Amazon's long-run cost minimization problem is characterized by a standard trade-off between

[28] For example, as shown on `https://services.amazon.com/fulfillment-by-amazon/pricing.htm?ld=NSGoogleAS`, the fee Amazon charges third-party sellers for order fulfillment ranges from $2.50 to $3.50 in Q4 of 2020. However, these fees include costs of "picking and packing [the] orders", "customer service", inventory management, and a mark-up, none of which are included in our estimate. The company Lojistics (`https://www.lojistic.com/white-papers`) reports that in 2017/2018, the cost of residential shipping via FedEx or UPS was around $3.40 per package, up from $2.40 in 2011. These appear to be base rates for a generic shipper and, therefore, do not include any high-volume discounts that Amazon negotiates. Finally, news reports during a 2020 feud between Amazon and the federal government reported that Amazon paid the USPS about $2 per package, $1.46 below the cost to USPS of handling such deliveries.See `https://nypost.com/2017/12/29/trump-says-amazon-is-making-us-postal-service-dumber-and-poorer/` and `https://www.businessinsider.com/trump-amazon-post-office-deal-2018-4`

Figure 4: Long-run Average Cost



Notes: Each curve represents a different network configuration. The average cost is the total average cost of fulfilling an order.

fixed and variable costs. As Figure 4 illustrates, denser networks (recent vintages) are associated with lower variable costs (limit of AC as $Q \to \infty$) and higher fixed costs. As the volume of orders expands beyond 500 million units ($\ln Q = 20$), it becomes more efficient to operate a decentralized network with 30+ facilities, instead of the 2006 network with 12 facilities. A volume of orders beyond 3,000 million units ($\ln Q = 22$) justifies investing in a network of over 90 facilities, composed of fulfillment and sortation centers (i.e., the 2015 network configuration).

External Cost Implications. The previous discussion focused on the private benefits of Amazon's expansion. Long-haul shipping also generates significant external costs, however, which is a concern that has been raised about the rising reliance on e-commerce.[29] Amazon's distribution expansion, by significantly reducing the average shipping distance of each order, has the potential to curb aggregate increases in external costs as the volume of online orders grows. In Table 8, we rely on estimates of the external costs of freight shipping per ton-mile from the literature to examine how the reduction in distance and the increases in order volume combine into estimated aggregate external costs.[30]

We calculate the counterfactual flow of orders under the observed 2006 and 2018 network for three levels of demand, 2006, 2012, and 2018, which provides us with an estimate of the total

[29] See, for example, https://www.theglobalcurrent.com/home/e-commerce-has-a-climate-change-problem.

[30] As we did in the model, we assume that the mileage incurred by the upstream suppliers is unchanged.

shipping distance for the hypothetical network and demand combination. We transform the shipping distance into ton-miles, assuming an average weight of 2.31 pounds per shipment[31], and apply three different external cost estimates from the literature ranging from 2.63 cents/ton-mile to 13.96 cents/ton-mile.[32] Focusing on 2018 demand in the bottom row of Table 8, the intermediate estimate (Van Essen et al., 2019) implies that the expansion of the network from 2006 to 2018 led to a three-fold reduction in external cost from $255 to $83 million. To put this number in perspective, note that the growth in e-commerce alone increased external costs by a factor of nearly 30, as the costs under the 2006 network increased from $9 to $255 million over 2006-2018. The network expansion limited this growth in external costs to a factor of 9, from $9 to $83 million.

Table 8: Network Expansion and External Cost

	External costs from truck mileage ($/millions)					
	Austin		Delucchi-McCubbin		Van Essen et al.	
	N_{2006}	N_{2018}	N_{2006}	N_{2018}	N_{2006}	N_{2018}
2006	5.35	1.56	17.63	5.14	9.14	2.67
2012	32.15	9.55	105.86	31.43	54.90	16.30
2018	149.53	48.62	492.32	160.06	255.33	83.01

We rely on the following external cost estimates (cents/ton-mile): Austin (2015), 2.63; Delucchi and McCubbin (2011), 13.96; Van Essen et al. (2019), 7.24.

4.4 Illustration: California entry

Beyond quantifying the role of economies of density in network expansion, we aim to assess the role of tax laws in Amazon's investment decisions during our sample period. To motivate how sales tax liabilities through the revenue channel affect firm profitability and thus the return to investing in fulfillment facilities near population centers, we conclude this section with an illustration of the tax-distance trade-off that our estimated model implies.

As an example, we consider a cluster of fulfillment centers in San Bernardino, CA. Amazon opened the first facility at this location in 2012, which also marks Amazon's first entry into California. The firm added two facilities to the cluster in subsequent years. In 2011, the year prior to the opening of the cluster, our model predicts that a majority of southern California orders were fulfilled by the then-closest cluster in Arizona, with the remaining orders coming from clusters in Nevada, Washington state, and Texas.

[31] We use the average weight in 2018 of a USPS "Parcel Select" shipment, the type of shipment Amazon uses for USPS shipments. See `https://about.usps.com/who-we-are/financials/revenue-pieces-weight-reports/fy2018-q2.pdf`.

[32] We take external cost estimates from Austin (2015), Delucchi and McCubbin (2011), and Van Essen et al. (2019). As these authors report average estimates per ton-mile, we do not quantify the external cost implications of the firm expanding into urban areas where it may generate above-average congestion externalities.

Table 9: Effect of Entering California (2011)

		Nexus (%Change)				Uniform Tax (%Change)		
Distance (m)	State	Orders	Profit	Shipping	L+FC	Orders	Shipping	L+FC
302	AZ	-0.88	-0.68	-0.20	0.99	-0.83	-0.21	0.95
399	NV	-0.17	-0.18	-0.03	0.12	-0.13	-0.04	0.10
965	WA	-0.03	-0.08	-0.24	0.08	-0.03	-0.23	0.07
1169	TX	-0.01	-0.01	-0.04	0.01	-0.01	-0.04	0.01
Total		-0.01	-0.00	-0.08	0.05	0	-0.08	0.05

Notes: Distance denotes the distance in miles from each facility to the San Bernardino cluster. "Orders", "Profit", "Shipping", and "L+FC" denote the percentage change in facility-level output, profit, shipping cost, and labor and fixed costs due to the hypothetical earlier opening of three fulfillment centers in San Bernardino, CA. The last row labeled "Total" measures the system-wide change in the outcomes of interest.

We compare outcomes under this actual opening sequence to outcomes under a network where Amazon moved up the opening of all three San Bernardino facilities to 2011. The left panel of Table 9 displays the results of this experiment under nexus tax laws. We first focus on the bottom row, which summarizes the overall percentage change in orders, profit, shipping costs, and labor and fixed costs, relative to the actual opening sequence. Under the counterfactual opening dates, the total number of orders decreases by 1% due to the earlier onset of tax liabilities in California. However, Amazon is able to realize savings in shipping costs of about 8%, as orders from southern California are now mostly fulfilled by San Bernadino, instead of further-away facilities. Labor and fixed costs increase by about 5% due to higher wages, rents, and congestion costs, as well as reduced economies of scale across facilities, as each facility handles fewer shipments with the California entry.

We further break down the aggregate effects in the first four rows of the table, where we present outcomes for the most affected clusters. The entry in California results in a redistribution of orders across facilities, as we show in the 'Orders' column. The Arizona facility experiences the biggest drop in fulfilled orders of 88%, suggesting that, once there is entry in southern California, the facility is largely redundant. However, because the remaining shipments are now mostly local, the average cost of shipping orders from this facility decreases by 20%. The large drop in orders also results in a near doubling of labor and fixed costs due to loss of economies of scale. Taken together, the profit contribution of the Arizona facility drops by 68%.

The remaining, further-away, facilities experience smaller declines in volume. These small changes in volume can, nevertheless, translate into sizable changes in shipping costs, as the Washington state facility illustrates. It sees a drop in volume of only 3%, which due to its initially low order volume relative to fixed costs results in an increase in combined labor and fixed costs of 8%. At the same time, the reduction in shipping costs amounts to a sizable 24%, larger than at the Arizona facility. Due to distance, shipments from Washington state to southern California are very costly, so even though only a small number of the Washington state shipments are reallocated to the San Bernardino cluster, these shipments have a large impact on shipping costs for the Washington state cluster. The distance between a new facility and existing facilities thus does not have a monotonically declining impact on the shipping costs at the existing facilities. While the profit impact at existing facilities declines monotonically with distance to the new facility in this particular example, the non-monotonicity in shipping costs could in principal translate into similar patterns in facility-level profit. This is similar to Holmes' (2011) approach who allows for a trade-off between cannibalization and density. At the same time, the non-monotone impact of distance between facilities on profit precludes us from using Barwick's (2008) methods in solving counterfactual network optimization problems under alternative tax structures in the following section.

In order to illustrate the impact of tax nexus laws on the firm's profitability, the right panel of Table 9 summarizes profits and costs under an alternative tax treatment where Amazon is responsible for remitting sales tax on all transactions, irrespective of presence in the state, which we term a 'non-discriminatory tax' law. Here, entering a new state no longer triggers a new tax collection and, thus, total demand remains unchanged (see last row) when we move up the San Bernardino opening date. Since demand in California is unaffected, we observe a less significant redistribution of orders as under the nexus laws, when the San Bernardino cluster is able to largely handle all of the southern California orders.

On net, the aggregate reduction in shipping costs from entering San Bernardino earlier would be slightly larger and the increase in other costs slightly lower under non-discriminatory taxes (even though the magnitudes appear to be the same in the table due to rounding). Not shown in Table 9 is the fact that overall profit for Amazon increases by 1% due to the earlier entry in California under non-discriminatory tax laws. Comparing this to the small but negative impact of entry on profit under the nexus laws demonstrates how Amazon's entry incentives change under different tax regimes.

5 Taxes and Investment

We now use the estimated model to quantify the combined effect of demand expansion and tax policy on investment in the distribution network. We first use the model to illustrate the optimal growth of the network as online demand expands. We then quantify the effect of nexus tax laws on the growth and configuration of the network to illustrate how the laws distort Amazon's investments.

Finding Counterfactual Networks. Solving the dynamic optimization problem is beyond the scope of this paper. We instead approximate the solution to this problem with a series of static profit maximization problems at different stages of e-commerce demand. We thus find the optimal network that maximizes the static flow profit in Equation (10), which we replicate here:

$$N_t^*(\Theta) = \arg\max_{N_t} \quad \bar{\mu}_t \hat{R}_t(N_t) - \hat{C}_t^{\text{Shipping}}(N_t) - \hat{C}_t^{\text{Labor}}(N_t) - \hat{F}_t(N_t), \tag{22}$$

We consider four years during our sample period – 1999, 2006, 2012 and 2018 – that are exemplary of the demand expansion that Amazon has experienced and use Θ to denote the dependence of the optimal network on tax policy. In the constrained dynamic problem we relied on in estimation, we hold the locations of facilities fixed and exploit the optimality of the opening sequence only. Here, we now allow the platform to choose both the location and number of facilities of each type, which are summarized by $N_t^*(\Theta)$, thereby studying the effect of tax policy and demand on aggregate levels of investment.

To implement this procedure, we first define a set of potential locations. We start with the locations of the roughly 150 Amazon facilities in 2018, including facilities that we excluded from our prior analysis (e.g., Amazon Fresh grocery delivery centers). We further consider the approximately 770 locations of distribution centers operated by Target and Walmart, from MWPVL, and UPS from Reference USA, in 2018. We keep only the subset of non-Amazon locations that have similar levels of income and population to those chosen by Amazon. Taken together, this results in about 330 unique potential facility locations. Using a distance radius of 20 miles, we use a hierarchical clustering algorithm to group nearby facilities, resulting in $N = 253$ unique potential locations spanning 39 of the 48 contiguous states. See Figure OA-1 in Online Appendix OA.2.3 for a map of the final set of potential locations.

We make the following simplifications to facilitate the comparison of networks across years. First, we assume that all input prices are fixed at their average levels for all time periods. Therefore, the only time varying component is the growth in Amazon's demand relative to other retail modes (i.e. α_{ikt} in our demand model). Second, we abstract away from differential tax treatment of sales by third-party sellers and assume that all Amazon transactions are subject to the sales tax policy. Third, we abstract from capacity choice by setting the size of each fulfillment and sortation center equal to $1,000,000$ and $300,000$ sq. ft., respectively, the approximate averages in the data.

Despite these simplifications, with approximately 250 locations and two facility types, the sheer number of potential networks renders it infeasible to solve the optimal network problem exactly (except in $t = 1999$). We therefore approximate the optimal network using simulation techniques borrowed from the operations literature. We use the simulation-based Population-Based Incremental Learning algorithm developed by Baluja (1994) that combines ingredients of genetic and standard hill-climbing optimization algorithms. In theory, as the number of simulations grows large, the algorithm converges to a global maximum. In practice, given the scale of our problem, the procedure may not identify the global maximum, in particular in the later years when the optimal number of facilities is large. Therefore, we refer to the solution as an approximation to the optimal network. We find through full enumeration, however, that the procedure predicts the globally optimal network in 1999. Online Appendix OA.2.3 describes the algorithm in detail.

Predicted Optimal Networks under Nexus Tax Laws. Figure 5a maps the predicted network evolution across four demand states, 1999, 2006, 2012, and 2018, under the nexus tax laws. The color of each dot indicates the first year in which we predict a facility to operate in a given location. This location may or may not have a facility operating again in later years.

To analyze model fit, we present a map-based comparison of the observed and predicted network roll-out in Appendix B.3. Overall, across years, the static model matches the regional distribution of actual fulfillment centers well, even though it is not always able to predict the exact set of states with physical presence. In 2018, when the network is larger, we predict the set of states with fulfillment capacity well, though the model slightly under-predicts the total of number locations.[33] Our estimates are thus largely able to capture the relative levels of revenue and cost that drive Amazon's network decisions.

We also use Figure 5a to highlight the three primary trade-offs that Amazon faces. First, there is a trade-off between economies of scale (fewer facilities) and density (more facilities): we observe large networks only in 2012 and 2018, once demand has grown sufficiently. Second, sales tax liabilities favor placement of facilities in low-population states, but such placement does not allow the company to benefit from economies of density. This trade-off drives the predicted (and actual) opening of facilities in high population states like California, Texas, and Florida only in 2012 and in 2018. In these years, local demand is sufficiently high for the benefits of lower shipping costs to outweigh the tax implications. Finally, there is a trade-off between high labor, rent and congestion costs that favor placement of facilities in low-cost areas, which also typically are low demand and thus do not generate significant economies of density. The patterns of fulfillment center placement in Oregon illustrate this trade-off. In 2006, we predict the optimal fulfillment center location to be remote and thus, low-cost. In later years, when demand is higher, we predict that locations closer to Portland are optimal.

Predicted Optimal Networks under Non-discriminatory Tax Laws. Figure 5b presents the roll-out under a non-discriminatory tax law. A comparison of the two maps suggests that nexus tax laws, not surprisingly, impact the network configuration primarily in high-population or high-tax areas where the revenue implications of raising tax-inclusive prices is largest. Texas provides a good example as our simulations show that Amazon would have entered in 2006 under non-discriminatory tax laws. Under nexus tax laws, however, the model predicts early entries in nearby Oklahoma and Louisiana instead, and entry occurs in Texas itself only in 2018, when cost savings from density finally outweigh the effects of sales tax.

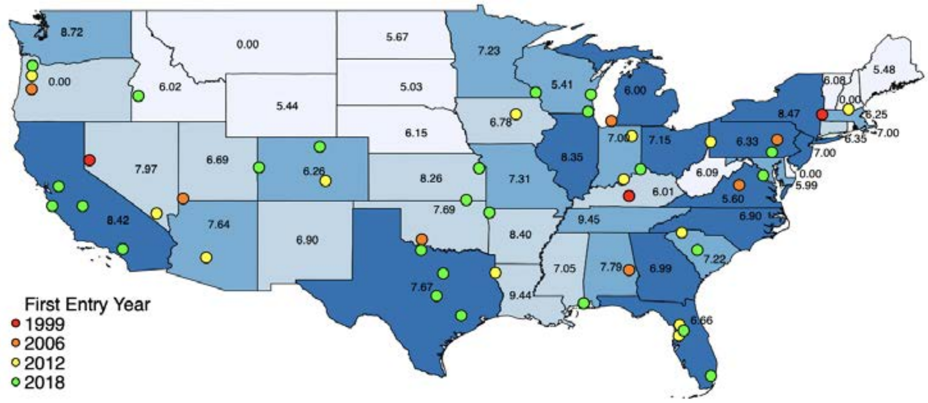
California provides another clear example of the effect of the nexus laws. The model predicts that under non-discriminatory taxes, Amazon would have entered in relatively remote areas of the state in 1999 and expanded closer to high density areas (San Francisco, Los Angeles, and San Diego) as demand grows, consistent with the presence of large fixed costs in those areas. In contrast, with nexus laws, the model predicts that Amazon would enter California only in 2018, with California orders being fulfilled from Nevada, Oregon, and Arizona beforehand.

Table 10 summarizes how demand growth and tax distortions affect Amazon's profit and average cost in aggregate, with each row corresponding to a different state of demand. We present Amazon's profit (column 1), average fulfillment cost in total and broken down into its components (columns 2-5), the sales-weighted average shipping distance (column 6), the average sales tax rate paid across

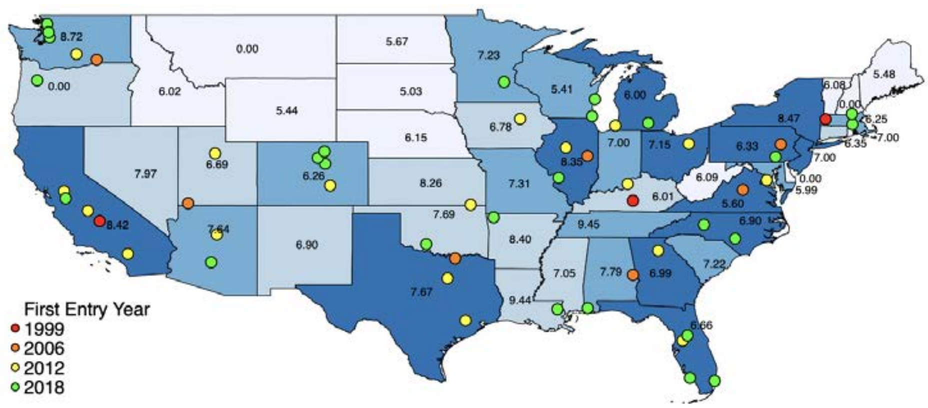
[33] The fact that the actual network includes more facilities is due to the fact that Amazon's investment decision is influenced by dynamic considerations such as entry deterrence and anticipated future growth.

Figure 5: Fulfillment Center Network and Nexus Tax Laws

(a) Nexus Tax



(b) Non-discriminatory Tax



Notes: Colors of dots denote the first year in which we predict entry into that location. We shade states by quintiles of the number of households, with the top quintile taking on the darkest color, and display the average sales tax rate in each state.

orders (column 7), and the fraction of orders subject to positive sales tax. Finally, we report the number of facilities of each type in the optimal network.

Relative to a non-discriminatory tax policy, Amazon's overall profit is higher under the nexus laws, despite the fact that average cost are generally higher, reflecting the direct effect of lower sales taxes on revenue.

The number of facilities under nexus laws is also smaller than with non-discriminatory taxes due to the fact that Amazon has a lower incentive to invest. This difference is especially pronounced when demand is high (2018): the model predicts both a smaller sortation network (9 versus 11 facilities) and fewer fulfillment centers (65 versus 74) under nexus laws, in part due to the complementarity between the two types of facilities. The nexus tax policy has a large effect on the configuration of the sortation network because the cost savings from vertical integration can only be realized by locating sortation centers close to high population areas; this makes tax arbitrage

Table 10: Summary statistics Across Tax Regimes

(a) Tax Nexus

		Average Cost ($/order)					Amazon Tax		Facilities	
	Profit	Shipping	Labor	FC	Total	Dist.	Avg.	% > 0	FC	SC
1999	88	1.29	0.86	0.32	2.47	377	0.26	0.05	3	0
2006	537	1.14	0.66	0.22	2.02	347	0.95	0.15	8	1
2012	4,784	0.71	0.48	0.14	1.32	233	1.71	0.24	25	4
2018	27,626	0.50	0.35	0.08	0.93	173	3.08	0.42	65	9

(b) Non-Discriminatory Taxes

		Average Cost ($/order)					Amazon Tax		Facilities	
	Profit	Shipping	Labor	FC	Total	Dist.	Avg.	% > 0	FC	SC
1999	79	1.26	0.88	0.37	2.51	367	6.74	0.98	3	0
2006	488	1.10	0.69	0.24	2.03	336	6.75	0.98	8	1
2012	4,499	0.59	0.52	0.15	1.27	201	6.96	0.98	27	5
2018	26,406	0.39	0.39	0.10	0.88	150	7.13	0.98	74	11

Notes: We measure profit in $ million and distance in miles. The columns labeled 'Amazon Tax' are the population-weighted average tax rate paid and the share of the population that pays a positive tax rate on Amazon transactions.

strategies based on state boundaries more difficult.

The distortions induced by nexus laws lead to a small increase in the average total fulfillment cost (from $0.88 to $0.93) in 2018. The effects are more pronounced for shipping costs: across years, the distorted network exhibits larger shipping costs, amounting to 28% in 2018, due to higher average shipping distances (of 15% in 2018), suggesting that the nexus laws reduce the incentive to realize economies of density. These cost increases are balanced by reductions in labor and fixed costs. With economies of scale in labor and large fixed costs of investment, a more concentrated network leads to cost savings due to a higher capacity utilization rate.

Finally, Table 11 quantifies the effect of the increase in shipping distance under nexus laws on the external cost of long-haul trucking. Recall from Table 8 that for the intermediate estimate of external costs (Van Essen et al., 2019), the model predicts a growth in external cost from $9 million to $83 million between 2006 and 2018 under the actual network. Table 11 shows that, using these same external cost estimates, the counterfactual network predicted by the model under nexus laws implies a slightly larger increase in external costs from $7 to $102 million, as the static optimization model under-predicts the number of facilities in 2018 slightly. Comparing the non-discriminatory and nexus columns, we see that the implementation of nexus laws alone increased external costs by $20 million. We thus estimate that the distortions in investment induced by nexus tax laws resulted in a 22% increase in the external cost of trucking in 2018, with similar proportions predicted in earlier years.

Table 11: External costs from truck mileage ($m) under Alternative Tax Laws

	Austin		Delucchi-McCubbin		Van Essen et al.	
	Nexus	Non-dis-criminatory	Nexus	Non-dis-criminatory	Nexus	Non-dis-criminatory
2006	4.08	3.59	13.43	11.81	6.96	6.13
2012	16.65	13.25	54.80	43.64	28.42	22.63
2018	59.81	48.86	196.92	160.87	102.13	83.43

Notes: External cost estimates (cents/ton-mile): Austin (2015), 2.63; Delucchi and McCubbin (2011), 13.96; Van Essen et al. (2019), 7.24.

6 Conclusion

We aim to make three primary contributions. First, we estimate consumers' sensitivity to sales tax across a wide range of online and brick-and-mortar retailers, thereby demonstrating the competitive advantage of online firms over their brick-and-mortar counterparts. We find that consumers are sensitive to sales tax, suggesting that online retailers, such as Amazon, consider the tax implications of their distribution network on demand. Second, we endogenize the distribution network for a large retailer, allowing us to quantify the benefits of density and vertical integration. This departs from previous literature, which generally takes the network of distribution facilities as given and relies on it to identify models of retail store location choice. Our estimates suggest that Amazon realizes significant cost savings and economies of scale from network expansion. Third, we quantify the distortionary effects of discriminatory nexus tax laws, the topic of intense political debate. We find that the laws reduced Amazon's incentive to invest in the network, increasing both the firm's overall shipping costs and the external costs from long-haul trucking.

The paper represents a first step towards understanding the sources of Amazon's dominance in the market and opens up several avenues for future research. One of these is the relationship between tax incentives and investment. As demonstrated, there is often a discrepancy between the timing of Amazon's entry and tax collection. This is likely an outcome of negotiations between state officials and Amazon. We do not incorporate this bargaining into our model, only allowing for a deterministic delay schedule based on the timing of entry. With a more expansive data set, one could model bargaining between firms and governments regarding subsidies and apply it to our dynamic framework. Slattery (2020) studies this type of environment, but does not consider the dynamic implications of entry on bargaining power.

Second, we take the growth of third-party sellers as given in our model, but this growth implies potential benefits to the platform beyond what we measure. As Amazon's distribution network has grown, the firm has become a vital distributor for these sellers, leading to potentially anti-competitive rent-seeking behavior. Such behavior is commonly raised by proponents of antitrust enforcement efforts against the firm.[34] With more detailed data that separates third-party sales from Amazon's direct sales, one could use our framework to quantify the market power Amazon commands from its status as a key logistics supplier to third-party retailers.

[34] See https://www.nytimes.com/2020/11/10/business/amazon-eu-antitrust.html, accessed on 3-9-2021.

References

Agarwal, S., N. Marwell, and L. McGranahan (2017): "Consumption responses to temporary tax incentives: Evidence from state sales tax holidays," *American Economic Journal: Economic Policy*, 9, 1–27.

Agatz, N. A., M. Fleischmann, and J. A. Van Nunen (2008): "E-fulfillment and multi-channel distribution–A review," *European Journal of Operational Research*, 187, 339–356.

Alm, J. and M. I. Melnik (2005): "Sales taxes and the decision to purchase online," *Public Finance Review*, 33, 184–212.

Anderson, E. T., N. M. Fong, D. I. Simester, and C. E. Tucker (2010): "How sales taxes affect customer and firm behavior: The role of search on the Internet," *Journal of Marketing Research*, 47, 229–239.

Asplund, M., R. Friberg, and F. Wilander (2007): "Demand and distance: evidence on cross-border shopping," *Journal of Public Economics*, 91, 141–157.

Austin, D. (2015): "Pricing freight transport to account for external costs," *Congressional Budget Office Working Paper 2015-03*.

Autor, D., D. Dorn, L. F. Katz, C. Patterson, and J. Van Reenen (2020): "The fall of the labor share and the rise of superstar firms," *The Quarterly Journal of Economics*, 135, 645–709.

Ballard, C. L. and J. Lee (2007): "Internet Purchases, Cross-Border Shopping, and Sales Taxes," *National Tax Journal*, 60, 711–725.

Baluja, S. (1994): "Population-based incremental learning. a method for integrating genetic search based function optimization and competitive learning," CMU-CS-94-163.

Barwick, P. J. (2008): "What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry," *Econometrica*, 76, 1263–1316.

Baugh, B., I. Ben-David, and H. Park (2018): "Can taxes shape an industry? Evidence from the implementation of the Amazon tax," *The Journal of Finance*, 73, 1819–1855.

Blundell, R., L. Pistaferri, and I. Preston (2008): "Consumption inequality and partial insurance," *American Economic Review*, 98, 1887–1921.

Bruce, D., W. F. Fox, and L. Luna (2009): "State and local government sales tax revenue losses from electronic commerce," *State Tax Notes*, 52, 537–558.

Bugni, F. A., I. A. Canay, and X. Shi (2017): "Inference for subvectors and other functions of partially identified parameters in moment inequality models," *Quantitative Economics*, 8, 1–38.

Cao, G., G. Z. Jin, X. Weng, and L.-A. Zhou (2018): "Market expanding or market stealing? Competition with network effects in bikesharing," National Bureau of Economic Research Working Paper 24938.

Cavallo, A. (2017): "Are Online and Offline Prices Similar? Evidence from Large Multi-channel Retailers," *American Economic Review*, 107, 283–303.

Chetty, R., A. Looney, and K. Kroft (2009): "Salience and Taxation: Theory and Evidence," *The American Economic Review*, 99, 1145–1177.

Chu, J. and P. Manchanda (2016): "Quantifying cross and direct network effects in online consumer-to-consumer platforms," *Marketing Science*, 35, 870–893.

De los Santos, B., A. Hortaçsu, and M. R. Wildenbeest (2012): "Testing models of consumer search using data on web browsing and purchasing behavior," *The American Economic Review*, 102, 2955–2980.

Delucchi, M. and D. McCubbin (2011): "External costs of transport in the United States," in *A handbook of transport economics*, Edward Elgar Publishing.

Dixit, A. K. and J. E. Stiglitz (1977): "Monopolistic competition and optimum product diversity," *The American economic review*, 67, 297–308.

Dolfen, P., L. Einav, P. J. Klenow, B. Klopack, J. D. Levin, L. Levin, and W. Best (2019): "Assessing the gains from e-commerce," National Bureau of Economic Research Working Paper 25610.

Einav, L., J. Levin, and N. Sundaresan (2014): "Sales Taxes and Internet Commerce," *The American Economic Review*, 104, 1–26.

Ellison, G. and S. F. Ellison (2009): "Tax Sensitivity and Home State Preferences in Internet Purchasing," *American Economic Journal: Economic Policy*, 1, 53–71.

Goolsbee, A. (2000a): "In a World Without Borders: The Impact of Taxes on Internet Commerce," *The Quarterly Journal of Economics*, 115, 561–576.

——— (2000b): "Internet Commerce, Tax Sensitivity, and the Generation Gap," *Tax Policy and the Economy*, 45–65.

Goolsbee, A., M. F. Lovenheim, and J. Slemrod (2010): "Playing with fire: Cigarettes, taxes, and competition from the internet," *American Economic Journal: Economic Policy*, 2, 131–154.

Greenstein, S. M. (1993): "Did installed base give an incumbent any (measureable) advantages in federal computer procurement?" *The RAND Journal of Economics*, 19–39.

Holmes, T. J. (2011): "The diffusion of Wal-Mart and economies of density," *Econometrica*, 79, 253–302.

Houde, J.-F., P. Newberry, and K. Seim (2021): "The Impact of Distance and Taxes on Amazon," Working Paper.

Kaido, H., F. Molinari, and J. Stoye (2019): "Confidence intervals for projections of partially identified parameters," *Econometrica*, 87, 1397–1432.

Pakes, A., J. Porter, K. Ho, and J. Ishii (2015): "Moment Inequalities and Their Application," *Econometrica*, 83, 315–334.

Scanlan, M. A. (2007): "Tax Sensitivity in Electronic Commerce," *Fiscal Studies*, 28, 417–436.

Slattery, C. (2020): "Bidding for firms: Subsidy competition in the US," Working Paper.

Smith, M. D. and E. Brynjolfsson (2001): "Consumer Decision-Making at an Internet Shopbot: Brand Still Matters," *Journal of Industrial Economics*, 49, 541–558.

Van Essen, H., L. van Wijngaarden, A. Schroten, D. Sutter, C. Bieler, S. Maffii, M. Brambilla, D. Fiorello, F. Fermi, R. Parolin, et al. (2019): *Handbook on the external costs of transport, version 2019*, 18.4 K83. 131, CE Delft.

Zheng, F. (2016): "Spatial competition and preemptive entry in the discount retail industry," Columbia Business School Research Paper.

Appendix

A Appendix: Data

A.1 Expenditures for the Representative Household

In this Appendix, we provide details on the spending data and construction of our measures of online spending by shopping mode and of offline expenditure for the representative household in county i and year t. The primary source for our online spending data is the comScore Web Behavior Database, which tracks the online purchasing and browsing activity of a random sample of internet users. With their permission, comScore records any activity on the users' registered computer, including that of other household members, and also collects the zip-code and demographic characteristics of participating households. For a given household, comScore records every online order during its time in the sample. For each order, we observe the seller's domain, the date and time of the order, the product category of item(s) purchased, the list price for each individual item, and a 'basket total', representing the order's total cost, including shipping and taxes. Coverage of the sample is reported in the first three columns of Table A-1.

Table A-1: comScore Sample Coverage and Spending by Retail Channel

Year	Households (k)	Coverage (%) Counties	Coverage (%) Households	Average Expenditure ($) [Orders] Offline	Amazon	Taxed	Non-taxed
2006	38.3	84	99	5,341	62 [2]	510 [13]	486 [11]
2007	39.5	85	99	5,474	86 [3]	588 [14]	502 [11]
2008	20.2	71	96	5,520	106 [4]	569 [14]	450 [10]
2009	15.9	64	95	4,912	129 [4]	619 [15]	514 [11]
2010	15.4	66	95	4,662	195 [7]	687 [16]	569 [12]
2011	17.3	68	96	4,792	286 [9]	865 [20]	542 [12]
2012	18.2	67	96	4,494	378 [12]	1,015 [23]	553 [12]
2013	16.5	62	95	3,816	491 [16]	1,062 [24]	693 [15]
2014	12.0	54	92	3,846	604 [19]	1,107 [25]	742 [16]
2015	16.3	64	95	3,219	798 [25]	1,250 [28]	853 [18]
2016	24.8	72	97	3,272	1,040 [32]	1,272 [28]	1,031 [22]

Notes: County and household coverage are the percentage of US counties and households residing in the comScore data. Expenditures and orders are the average across households in the given year.

We begin the process of constructing our spending variables by removing transactions for product categories (defined by comScore) that Amazon does not compete in, for example travel and dating services. Next, we classify each of the remaining domains into one of the three online modes, where the classification depends on the retailers offline footprint and, therefore, their tax liabilities. Appendix Table A-2 lists the top ten domains each that fall into the non-Amazon categories of sellers. Then, we aggregate the 'basket total' across all transactions through mode k for household h in year t. This results in annual 'household expenditures' in each mode.

The next step is to account for the fact that we only observe 'household expenditures' for households that made at least one online purchase in year t. To do this, we supplement the comScore data with survey data from Forrester Research, Inc.'s "North American Technographics Online Benchmark Survey". Among other information, the survey records for the period 2006 to

Table A-2: Taxed and Non-Taxed Competitors

Sales Rank	Taxed	Non-Taxed
1	walmart.com	ebay.com
2	jcpenney.com	qvc.com
3	bestbuy.com	dell.com
4	macys.com	yahoo.net
5	apple.com	hsn.com
6	homedepot.com	overstock.com
7	victoriassecret.com	fingerhut.com
8	target.com	amway.com
9	staples.com	newegg.com
10	gap.com	orientaltrading.com
Total	102	292

Notes: Table displays top 10 domains in terms of 2013 expenditures in the comScore data that we define as taxed and non-taxed.

2007 and 2010 to 2014 whether a responding household indicates having made an online purchase over the three months prior to the data collection, together with the age, income, race, and zip code of the respondent. Patterns in the Forrester data suggest an increasing take-up of e-commerce (see column (1) in Table A-3).

We employ a linear probability model to project each Forrester respondent's propensity to purchase online on household demographic categories (race, age and census region of head of household), household income, and time trends. We then use the estimates of the model to predict the probability that a household in the comScore data made an online purchase in a given year, based on their demographics. For the years 2008 and 2009, we use linear interpolations for a given demographic group based on the predictions in 2007 and 2010. Using the predicted online purchase propensities from this model, we calculate 'expected annual expenditures' on each shopping mode for the households in the comScore sample, assuming that the propensity to purchase online applies to all modes equally.

Next, we aggregate across households to derive the expenditure of a representative household in county i in year t on the three online shopping modes. To do this, we first calculate the average expenditures for demographic group z in county i and year t for mode k. We then apply demographic sampling weights that measure the relative prevalence of demographic group z in the comScore data and data from the Census to derive expenditures for mode k for a representative household in county i and year t.[35],[36]

Finally, we address the intensive-margin bias in comScore spending introduced by the firm not recording the full universe of household online activity across devices. We assume that under-

[35] De los Santos et al. (2012) compare the sample of comScore users in 2002 and 2004 to the Computer Use Supplement of the Current Population Survey and find that the sample generally compares well with the population of online shoppers.

[36] The sampling weights are constructed based on the relative number of households that fall into different demographics bins in the comScore sample and in the population. The population data comes from the American Community Survey ('ACS-5 year') from 2009-2016. To extrapolate data back to 2006, we assume a county level constant growth rate in population belonging to each bin between 2000 and 2009, where the 2000 data come from the decennial Census.

Table A-3: Unscaled Household Online Purchasing

	Forrester	Average Expenditures w/out Scaling			
	Offline Only (%)	Amazon ($/year)	Taxable ($/year)	Non-Taxable ($/year)	All ($)
2006	55.6	18.2	99.8	95.1	213.0
2007	60.8	19.8	102.8	87.7	210.3
2008	-	21.5	99.8	78.9	200.3
2009	-	25.3	81.5	67.8	174.6
2010	32.1	34.1	76.7	63.6	174.3
2011	23.0	52.0	88.3	55.4	195.6
2012	23.9	84.1	85.0	46.3	215.5
2013	27.2	78.9	81.0	52.8	212.7
2014	24.6	77.3	85.2	57.1	219.6
2015	23.0	98.0	107.0	73.1	278.1
2016	22.6	116.2	98.5	79.8	294.5

Notes: Expenditures are the average across households. Offline Only denotes the share of respondents who answered no to the question whether they had shopped online in the previous three months in the Forrester Technographics Survey.

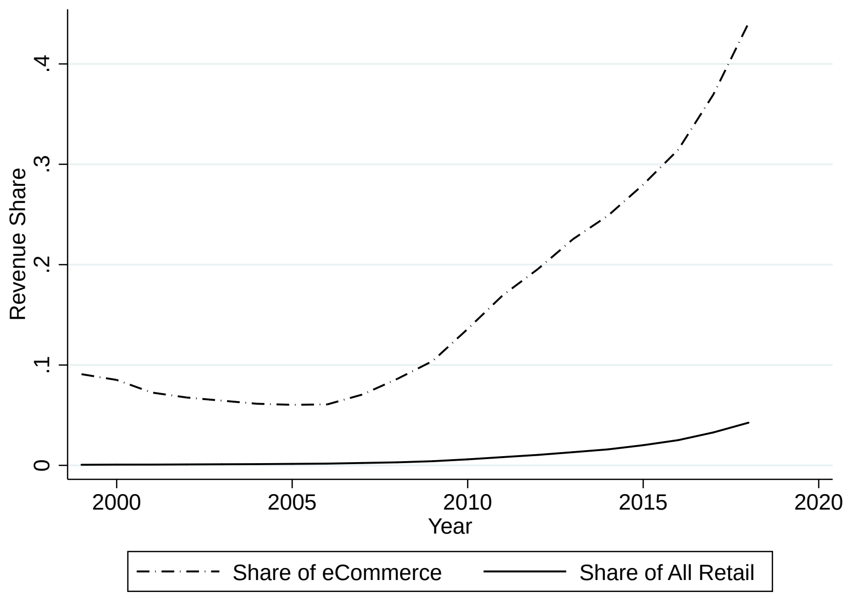
reporting in comScore online spending is uniform across counties and scale up each representative household's spending by a year and shopping-mode varying scale factor. We determine the scale factor by matching the average household spending calculated using our sample to the average spending per household on Amazon calculated using Amazon's annual reports and spending on the other modes calculated using the U.S. Census Bureau's quarterly e-commerce retail sales reports. Specifically, we multiply the expected expenditures for each representative household by the scaling factor and then calculate the average household spending across the United States on each mode. We search for the scaling factors for each mode and year where the resulting household averages match the averages from the supplemental sources.

In order to construct a spending figure from Amazon's annual financial statements that is comparable to the comScore data, we only include the reported sales from the "Media" and "Electronics and Other General Merchandise" categories in North America. An additional issue in matching the spending data is that the financial statement excludes sales tax paid and includes only the royalties and fees earned off of third-party sales. The comScore data, on the other hand, includes sales taxes and the full revenue from third-party sales. To account for this when determining the scaling factors, we adjust the comScore data to exclude taxes and to include only the portion of third-party sales that Amazon retains. We calculate the latter using Amazon's royalty rate on third-party transactions (see description in Section A.7 below) and data on the evolution of the share of total sales accounted for by third-party sellers over the sample period reported in the 2018 annual report cover letter.

The right panel of Table A-1 summarizes the average household spending by shopping mode. Spending on all online modes have increased substantially over time as people substitute away from offline shopping, with Amazon displaying the most pronounced growth. By 2016, the average household spending is about $1,000 on Amazon, while it spends about the same amount on non-taxed competitors. For taxed competitors, the spending is higher at $1,300. Finally, households spend about $3,300 at offline retailers. In brackets, we display the number of orders for the online modes, which are calculated by dividing the household spending by the price index calculated in Appendix A.6. The orders have roughly the same growth patterns as spending.

Figure A-1 demonstrates the growth in Amazon's share of online and all retail spending over time. These numbers are calculated using data from Amazon's financial statements and reports by the US Census bureau, which is why we have data points outside of our comScore sample. Amazon's online share remains relatively stable, and even decreases, until about 2006. Thereafter, we see a rapid increase, culminating at over over 40% by 2018. Amazon's share of total retail has similar patterns, reaching about 5% in 2018.

Figure A-1: Amazon's Growth



Notes: Calculations are made using reported sales in Amazons financial statements and reported online and total retail sales reported in the U.S. Census Bureau's quarterly e-commerce retail sales reports.

We explore heterogeneity in spending on Amazon across demographics and geography in Table A-4, where break down average annual household spending on Amazon across urban and rural counties, wealthy and non-wealthy counties, and the counties belonging to the four different Census regions. Overall, the growth rates are similar across demographic groups and regions, but there exists significant cross-sectional differences in spending based on region, income, and whether or not the consumer lives in a city.

Finally, we report the unscaled average household spending data in table A-3 and compare it to the scaled data in Table A-1. For Amazon, the pattern of increasing household spending matches the scaled data, but the magnitude of the growth is smaller. This is inline with the intuition that the share of 'missing' transactions increases over time due to the take up of mobile purchasing. Modes 2 and 3 generally show more of a U-shaped or flat pattern, which is largely inconsistent with what we see in the data from the US Department of Commerce. Again, this is likely because of the data we are missing from mobile or other-computer transactions for these modes.

Note that the scaling is mode-year specific, so that we preserve the rich variation in spending across geographies and demographics from the comScore and Forrester data. This variation identifies the consumer response to sales-tax, while the scaling acts to correct the *level* of spending and revenue by mode, which is important in identifying the supply side of the model. Therefore, the scaling exercise mostly serves to capture the degree of Amazon's growth relative to the other modes in line with the data observed in the reports.

Table A-4: Average Spending on Amazon, by Type of County ($/year)

			Income					
Year	Urban	Rural	High	Low	Northeast	Midwest	South	West
2006	70	55	87	57	80	54	52	74
2007	98	76	121	79	105	79	76	100
2008	120	99	131	102	131	93	99	129
2009	152	118	175	124	155	126	117	161
2010	212	197	255	189	234	187	180	238
2011	322	273	372	267	360	275	263	327
2012	423	367	464	367	436	384	353	447
2013	534	505	597	487	558	498	468	595
2014	656	665	752	615	721	603	589	774
2015	883	793	981	765	949	755	798	910
2016	1,125	1,020	1,219	989	1,159	1,019	1,007	1,181

Notes: We define rural counties as counties with a population density of less than 500 residents per square mile and low-income counties as counties with average household income below $80,000. Region is as defined by the US Census Bureau.

A.2 Employment

The employment data comes from industry sources and Amazon's financial statements and press releases. We construct a cross section of the number employees at a subset of fulfillment and sortation centers in 2017. Our main source of information is MWPVL and the establishment survey YTS (Your-economy Time Series).[37] MWPVL reports the target number of employees for a subset of facilities. We match the list of facilities and the panel of Amazon establishment from YTS. This survey provides employment data for most facilities, but the overlap is not perfect. In addition, although the YTS data is annual, the data exhibit very little adjustments over time, and so we use the more recent cross-section (2017). When two facilities are covered by YTS and MWPVL, we use an average of the two estimates. Finally, for a small number of facilities that are unmatched, we use data from Reference USA which includes information on employees for a limited set of Amazon facilities.[38] Overall, we use employment information for 131 out of the 163 facilities active in 2018 (including 26 sortation centers).

We also observe the total number of fulfillment employees in 2017, 125 thousand, from an Amazon press release.[39] We combine the latter and information from Amazon's financial statements to get an estimate of the total number of fulfillment employees for the other years in our sample. We observe the logistic cost share of total cost and the total number of employees in Amazon's financial statements for our entire sample. With these data and the number of logistic employees from the press release, we construct the ratio of the cost share of logistics to the employment share of logistics in 2017. Assuming this ratio is constant over time allows us to back out the number of logistics employees for all of the years in our sample.

Figure A-2 provides a summary of these data. Figure A-2a demonstrates that the increase in fulfillment employment coincides with the accelerating expansion of the network starting around 2010. The histograms in Figures A-2b and A-2c show that both types of facilities have a significant

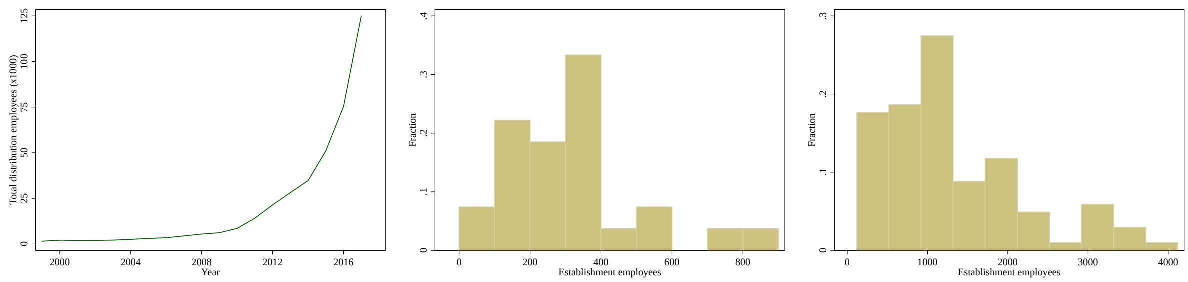
[37] See https://wisconsinbdrc.org/data/

[38] See http://www.referenceusa.com/Home/Home

[39] See https://press.aboutamazon.com/news-releases/news-release-details/amazon-now-hiring-over-120000-jobs-us-holiday-season

Figure A-2: Employment by facility types and year

(a) Total Employment (b) Sortation Center(2017) (c) Fulfillment Center(2017)



Notes: The left graph is the time series of overall fulfillment employment. The center and right graphs are the distributions of employment across sortation centers and fulfillment centers in 2017, respectively.

amount of variation in employment in 2017, with the level of employment being much higher at fulfillment centers.

A.3 Sales Tax

Our tax rates come from Thomson Reuters' Tax Data Systems, which provide state, county, and local tax rates for the years 2006-2018. We add these rates together and compute the average tax rate across municipalities in each county for each year of the sample. Because we do not observe rates before 2006, we assume that the tax rates from 1999-2005 are equal to the tax rate in 2006. Table A-5 summarizes the tax rate data. The household weighted average sales tax rate varies between 6.75% and 7.13%, with a standard deviation of between 1.5% and 1.6% across counties in every year. In addition, tax rates vary across time, as between 42 and 68% of households live in counties that experience tax rate changes from year to year and all counties experience at least one tax rate change over our sample period. Finally, higher population counties tend to have higher tax rates, as demonstrated by the final column.

Table A-5: Tax Rates

Year	Ave Tax	StDev Tax	% HHs w/ Change	Corr(#HHs, Tax)
2006	6.75	1.54	55.33	0.09
2008	6.82	1.51	67.47	0.10
2010	7.03	1.66	67.97	0.13
2012	6.96	1.56	64.60	0.11
2014	7.02	1.56	42.13	0.11
2016	7.13	1.59	42.69	0.10

Notes: The average and standard deviation of tax are moments from the distribution of the county level average tax rates. HHs with change is the percentage of households that lived in a county where the average sales tax rate changed that year. Corr is the county level correlation between households and the average sales tax rate.

This tax rate is assumed to apply to all transactions at brick-and-mortar and taxed online retailers. Amazon transactions are taxed, and the local sales tax rate applies, if the consumer lives in a state where Amazon collects taxes. As mentioned in the text, nexus tax laws would suggest that this occurs when Amazon operates a facility in the consumer's state. However, it is not always the case that these two events coincide perfectly, and here, we provide anecdotal and empirical evidence that the gaps in timing are due to negotiation between state governments and Amazon

We begin by providing a few pieces of anecdotal evidence of this connection. First, in the 1999 through 2011 financial statements, Amazon states that "a successful assertion by one or more states or foreign countries that we should collect sales or other taxes on the sale of merchandise or services could result in substantial tax liabilities for past sales, decrease our ability to compete with traditional retailers, and otherwise harm our business." (p. 15 in 1999) There are similar quotes in later statements about the repercussions of states "requiring [Amazon] to collect of taxes where [they] do not" (e.g., 2017 page 12). This suggests that Amazon considered sales tax to be a first-order issue impacting their bottom line.

A second piece of anecdotal evidence comes from documented negotiations between specific states, such as Nevada and Texas, where the debate between Amazon and the state governments over locating in the state centered around sales tax and Amazon's physical presence in the state. Amazon also chose to shut down its affiliate program in Illinois in order to avoid sales tax when the state changed the nexus laws to include affiliates. These examples provide suggestive evidence of the importance of the relationship between sales tax and Amazon's strategic decisions.

There are occasions, mostly towards the end of our sample, where Amazon began to charge sales tax before opening a facility. However, to the best of our knowledge, in all but one of these cases (NY), Amazon had plans to build a facility in that state soon after the onset of sales tax collection. Of course, it could be that Amazon decided to build a facility only after conceding on the sales tax issue, but this seems unlikely given the nature and history of nexus laws. States have unsuccessfully fought for years to change the nexus laws to include companies without a physical presence in their jurisdiction, like Amazon. However, Amazon did start to charge sales tax in all US states in 2017, irrespective of physical presence. This appears to be mostly a move for good publicity and/or to put pressure on federal regulators to change the nexus laws, as Amazon had already been charging sales tax to over 90% of the US population by this time.

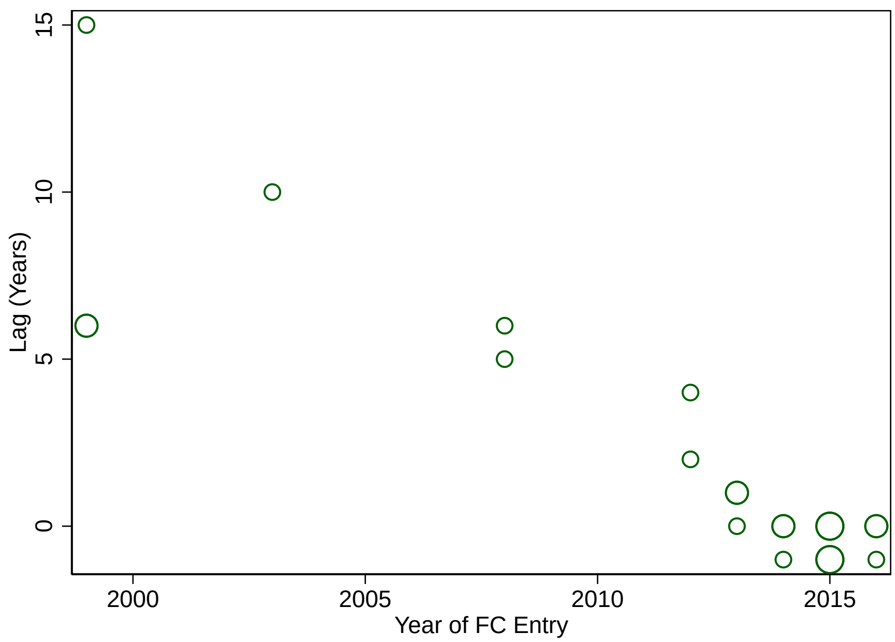
The final pieces of anecdotal evidence are the observed location decisions of Amazon. Specifically, placing facilities on the western border of Nevada, the southern border of New Hampshire, the southern border of Wisconsin, and in Delaware suggest that sales tax played an important strategic role in determining their locations.

Given these anecdotes, we argue that the change in sales tax obligation is triggered by entry, but that this change may only take effect after negotiations with state officials. The way that we think about these negotiations is based the observed changes in the time between entry and sales tax collection over time.

We observe that in the early period of our sample, there was often a significant lag between the date of entry and the change in sales tax laws. For example, Amazon did not start collecting sales tax in Pennsylvania until nine years after it opened its first fulfillment center in the state. This was a time when Amazon had significant bargaining power with the states. One reason is the lower demand in early time periods, so it was not as important for Amazon to be close to its customers. Additionally, in early periods they were not located in very many states, meaning there may be competing options of other, possibly lower sales-tax, states. For both of these reasons, Amazon's threat point (the value of opening in a different state) was higher in the early periods. An example of this comes from Nevada in 2011, where a law that would have forced Amazon (and other online retailers) to collect sales tax failed due (partially) to concerns that Amazon would close its fulfillment center in the state.[40] Later, Amazon negotiated a deal with the state to start collecting sales tax in 2014.[41] In another example in the same year, Amazon closed a fulfillment center and abandoned plans for expansion in Texas due to a dispute about paying uncollected taxes. These two examples show Amazon using its leverage in order gain preferential tax treatment in the early years of our sample.

However, as Amazon continued to expand their network, their bargaining power with states lessened. State governments knew that, given the level of demand, it was important for Amazon to be close to customers and that there were not many tax-friendly fallback states remaining. Because of this, we see Amazon agreeing to charge sales tax quickly after the opening of the first facility in a state and sometime even before. We note that in all but one of the states in which they started to collect sales tax 'early', they had agreements in place to build a facility soon after.[42] To the best of our knowledge, there is only one example, New York, where Amazon started to collect sales tax without an announced plan to build an fulfillment center

Figure A-3: Tax Date Lags



Notes: Each point represents a year of first entry into the state and the corresponding lag or lead in the implementation of the collection of sales tax. The size of each point indicates the number of facilities with that entry and lag.

To formalize this argument, we use data on the time between entry and sales tax. First, we present Figure A-3, which shows, for each first entry into a given state, the time between entry and tax collection. This demonstrates a negative relationship between date of first entry and the

[40] See https://www.reviewjournal.com/uncategorized/taxation-committee-drops-internet-sales-tax-amendment/?ref=894

[41] See https://lasvegassun.com/news/2012/apr/23/nevada-reaches-agreement-amazon-collection-sales-t/

[42] See, for example, the situation in New Jersey: https://www.nj.com/news/2012/05/amazoncom_to_begin_collecting.html

lag in tax collection dates. Second, we run a regression where the dependant variable is the time between first entry into a state and the beginning of tax collection. The regressors are a linear time trend, a dummy variable indicating that the first entry in that state occurred post 2012, and the minimum tax rate in a nearby state (i.e., same census region) that has yet to have an fulfillment center. The time trend is intended to capture the reduction in bargaining power due to the increase in demand, while the post-2012 dummy variable captures a discrete shift in bargaining strategy at this time period. The tax variable is included to proxy for the value of the 'outside option'. Table A-6 displays the results, which show that the lag gets shorter post-2012 and for each additional year. Additionally, the more attractive the outside option (i.e., lower tax rates), the higher the lag in collection.

Table A-6: Tax Lag Regression

VARIABLES	(1) Tax year - Entry year
Minimum tax rate in unoccupied states from the same region	-41.80**
	(17.05)
State entry year	-0.418***
	(0.107)
1(Entry year>2012)	-2.295*
	(1.251)
Constant	847.1***
	(214.9)
Observations	23
R-squared	0.861
Mean dep. variable	2.217
SD dep. variable	4.112
Mean tax variable	0.0393
SD tax variable	0.0225

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses. The dependant variable is the first year of tax collection minus the year of the first entry into a state. Minimum tax rate is the smallest tax rate of states in the same region that do not yet have a facility. Regions are defined by the US Census Bureau.

We take this into account in the supply side model and in the counterfactuals by assuming an deterministic tax abatement schedule that depends on the entry year. We believe that this schedule is able to capture the most important aspect of the negotiations (the time series) and that a more comprehensive bargaining model between the states and Amazon, although interesting, is out of the scope of this paper.

A.4 County characteristics

We explain our procedure for collecting and imputing the county level data in our analysis. We separate the sections by the different categories of data we collect.

Demographics

The estimation of the model requires observing the demographic characteristics of the representative household and the number of households in each county in the US for the years 1999 to 2018. The characteristics we include are income, age, and race. To measure the income of the representative household, we include the average household income in a county as well as the share of households that have household income above 100 thousand. To measure age, we include the share of households with a head of the household above 35 years old and to measure race, we include the share of households with a head of the household who is black, Asian, or another race.

These variables, as well as the total number of households, are available during census years (2000 and 2010) for every county, but between census years, we rely on the Census's American Community Survey (ACS). The drawback of the ACS is that the data are only available for high population counties for the years 2006-2008 via the "ACS-1 Year" (depending on the variable, missing percentage ranges from 41% to 75%). The "ACS-5 Year", on the other hand, is available between 2009-2016 for all counties. At the time when the data were collected, the 2017 and 2018 "ACS-5 year" were not yet available. To summarize, we observe the necessary variables (1) for the complete set of counties in 2000 and from 2009-2016 (2) for only large counties from 2006-2008 and (3) for no counties in 1999, 2001-2005, and 2017-2018.

Therefore, we impute demographics and the number of households for the missing counties between 2006-2008 and all counties in 1999 and between 2001-2005 using. We regress each county level demographic variable or number of households on state×year fixed-effects and county fixed-effects:

$$y_{it} = \mu_i + \tau_{state(i),t} + e_{it}.$$

This regression is estimated using observations for all counties during the census years, as well as non-missing ACS counties for other years. We use the predicted value of this regression to impute the data for missing counties from 2006-2008.

In order to use the model to predict the data for all counties in 1999 and 2001-2005, it is necessary to have values of the state/year fixed effects for these years. We assume a constant annual growth rate of these effects between 1999 and 2006, which we are able to determine using the estimates of the fixed effects for 2000 and 2006.

Finally, in order to predict the variables for 2017 and 2018, we assume a constant annual growth rate of each demographic variable and the number of households from 2015-2018. We calculate the growth rate using the data from 2015-2016.

Retail Establishments

Our measure of offline competition is the count of local retail establishments. We separate these into large and small retailers by the number of employees (greater or less than 50). For these counts, we use data from the County-Business Pattern (CBP) between 1999-2016 for all counties. Similar to the demographic data, we assume a constant growth from 2015-2018 in order to predict the offline competition variables.

Retail wage

To estimate the county-level average retail wage, we use the annual wage in the retail sector observed in the BLS Quarterly Census of Employment and Wages from 1999-2016. In each year, there are around 37% counties with no wage data and between 0.01%-0.03% of counties are outliers (wage out of range \$10,000-\$50,000). To impute the wage for these counties, we calculate the weighted average wage of the twenty nearest neighboring counties that have data, using an inverse distance weight. The data is then extrapolated to 2017 and 2018 using the same procedure as the one used with the demographics, households, and retail establishments.

Rent

We observe rents for distribution and industrial establishments at the MSA level for 2010-2018 for 47 MSAs from the Regional Economic Information System (REIS), and for industrial only establishments between 2006-2010. Assuming the MSA-level ratio of industrial to distribution rent stayed constant over the period, we use both data sets to predict MSA-level distribution rent from 2006-2018.

We then proceed in two steps to form a county level measure of rent for the years 1999-2018. In the first two steps, we use the REIS and other data to predict rents for all MSAs from 1999-2018 and in the third step, we use these predictions to estimate a county level measure. The process proceeds as follows:

1. Using the 2006-2018 data for the observed MSAs, we regress the (log) average rent on the MSA-level CPI, which is available for all years between 1999 and 2018 from the BLS, while controlling for MSA fixed-effects. We use the predicted values of this regression to predict rent prior to 2006 for the observed MSAs.

2. To predict rent for the remaining MSAs, we regress (log) average rent on the following MSA characteristics observed in the BLS and Census demographic data and aggregated to the county level (in log): population, pop. density, employment rate, industrial employment rate, office employment rate, nb of of employees, nb of establishments, median house value, mean income, labor force participation rate, and year and census division fixed effects. We use the predicted value from this regression to predict rents for the missing MSAs for all years.

3. To predict rent at the county level, we first assign the MSA rent measure to all counties located with 20 miles of the MSA centroid. To predict outside of this radius, we regress MSA rents on the characteristics of counties located within the radius, where we include the following county characteristics: median house value, land size, population density, wage and median income. The counties outside of the radius are then assigned the predicted value of this regression. Essentially, we extrapolate outside the MSA core using the observed relationship between housing cost and density and average distribution rent.

A.5 Total Retail spending

We use micro data from the Consumer Expenditure Survey (CEX) between 1999 and 2016 to construct a measure of average retail at the county-year level. The micro-data sample includes repeated cross-sections of roughly 60,000 households reporting annual spending on categories covered by Amazon and other general merchandise retailers. We choose categories based on their Universal Classification Code (UCC). The categories cover items that can roughly be put in to the following larger categories: beauty supplies, household items, electronics, apparel and accessories, office/school supplies, books, pet supplies, and sporting goods. A detailed description of which UCC codes are included is available upon request.

We use these data and an imputation approach similar to Blundell et al. (2008) to construct our retail spending for the representative consumer. The starting point is a linear regression relating spending with household characteristics and time/region fixed-effects:

$$y_{it} = X_{it}\beta_t + \mu_{r(i),t} + \epsilon_{it} \tag{A-1}$$

where X_i includes indicators for age and income groups, education, employment status, rate and family composition, and μ is a region/year fixed effect. Note that the regression is estimated separately for each year.

The second step measures the mean of X_{it} for each county/year. We use the aggregate census tables described in the previous section for this task. Let $\bar{X}_{j,t}$ denotes this average for county j in year t.

We then use the estimates $(\hat{\beta}_t, \hat{\mu}_t)$ to calculate the conditional expectation of (log) annual spending in each county:

$$\bar{y}_{jt} = \bar{X}_{jt}\hat{\beta}_t + \hat{\mu}_{r(j),t}. \tag{A-2}$$

Finally, we transform this conditional expectation into levels assuming that spending is distributed according to a log normal distribution:

$$\text{Spending}_{jt} = \exp\left((\bar{y}_{jt} + \hat{\sigma}_\epsilon/2)\right) \tag{A-3}$$

Since the log-normal distribution is sensitive to outliers, we windsorize the distribution of $\bar{y}_{jt}$ by truncating the values to the 99.9% percentile. To form data for 2017 and 2018, we extrapolate assuming a constant annual growth rate as in the previous sections.

A.6 Prices and Variety

In this section we describe how we construct the price indices and the inter-quartile range of prices (IQR) on Amazon, which is the measure of product variety. We first discuss the price indices. We calculate the average price of goods purchased over the course of a year for the representative household in each county using the comScore data. Specifically, we calculate the average price of goods purchased on Amazon for each household and then use population weights to aggregate to the county level. Therefore, the construction of this is the same as the construction of the weighted average spending discussed in Section A.1. To smooth any remaining noise in the pricing data and to make out-of-sample predictions, we regress the county level prices on a linear time trend and the share of households in the county who have a head of the household that makes over $100 thousand in income. Including the latter accounts for the fact that higher income households may be buying different things on Amazon.

The results of these regressions are in the first two columns of Table A-7. In the first column, the dependent variable is in levels, while in the second column, it is in logs. Both specifications indicate that there is a positive trend in the prices on Amazon over time and the log specification shows that higher income households buy more expensive items. The time trend in the log-linear model suggests that prices increase by just under 1% each year. We use the results of these models to predict the transacted prices for each county and each year. These predictions are in the first two columns of Table A-8. Both specifications predict a 1999 price of about $25. The prices then increase until 2018, with the increase being slightly more for the linear model.

We now turn to the IQR of prices. Because we use this as our measure of variety across Amazon, we do not utilize the variation in prices at the county level, which represents only the variation in products purchased by a single household (or county) and not necessarily the products available on the platform. Therefore, we utilize the variation in prices across the entire US. Specifically, we calculate the IQR for each year using all the transacted prices observed in the comScore data.

Similar to the price index, we use a regression to smooth the data and to make out of sample predictions. However, here we only include the time trend as the dependent variable because the dependant variable is at the year level rather than the county/year level. The results are displayed in the third and fourth columns of A-7. Despite the fact that we only have 11 observations, the model fits the data well. Additionally, the estimates show that the variation in prices is increasing over time, with the log-linear model indicating that the IQR increases by about 8% each year. We

Table A-7: Price Regressions

Dep Var	P	Log(P)	IQR	Log(IQR)
Linear Time Trend	0.436***	0.009***	1.970***	0.084***
	(0.044)	(0.001)	(0.263)	(0.008)
Share of Pop w/ Inc >100k	0.069	0.581***		
	(1.566)	(0.038)		
Obs	17,338	17,338	11	11
R-Sq	0.006	0.021	0.862	0.922

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses. The first two columns are price regressions using county level data and the last two columns are variety regression using annual data.

again use the results to predict the IQR across all years from 1999-2018. The last two columns of Table A-8 indicate that the IQR increases substantially over time. While the linear model predicts a negative value in the early years, the middle years (during our sample) and the overall increase are comparable across specifications. In our remaining analysis, we use the predictions of the log-linear model for prices and IQR.

Table A-8: Predicted Price Moments

	Predicted Price Index		Predicted IQR	
Year	Linear	Log	Linear	Log
1999	25.01	25.16	-0.73	8.05
2002	26.32	25.87	5.18	10.36
2005	27.63	26.71	11.09	13.34
2008	28.94	27.81	17.00	17.17
2011	30.24	28.66	22.91	22.10
2014	31.55	29.67	28.82	28.45
2017	32.86	30.51	34.73	36.63
2018	33.30	30.77	36.70	39.85

Notes: The predictions of the log models account for the variance of the residual.

A.7 Marketplace Spending and Approximating the Gross Margin

Recall that $\bar{\mu}$ is the weighed average margin between purchases directly through Amazon and purchases through third party sellers:

$$\bar{\mu}_t = s_t^{3PS}\mu^{3PS} + (1 - s_t^{3PS})\mu^{own}$$

where s_t^{3PS} is the share of marketplace sales observed in the 2018 annual report (cover letter). We use the annual reports of Amazon from 1999-2018 along with outside sources to determine the values of μ^{3PS} and μ^{own}, which we assume are fixed over time.

We begin with our approximation of μ^{own}. Amazon's annual report displays their gross margin each year (e.g., 2013 p. 26), but there are three problems with using this number directly. First, this includes the cost and revenue of third party sales. Second, this margin includes some of the costs that we will estimate such as labor, land, density, and shipping costs. Third, there are additional costs that we want to include that are not included in "Cost of Goods Sold", for example robotics, electricity, etc.

Instead, we calculate the margin as:

$$\mu^{own} = \frac{\text{Amazon Sales} - (\text{Cost of Goods Sold} + \text{Other Costs} - \text{Shipping Costs})}{\text{Amazon Sales}}$$

We observe "Amazon Sales" in their annual report (e.g., 2016 p.26). Specifically, Amazon reports their total revenue in "Media" and "Electronics and Other General Merchandise" categories in North America, which is roughly equivalent to the revenue we predict from our model. Revenue from Canada and Mexico is not included in the comScore data, but this comprises a small share of total revenue.[43] This figure includes both direct product sales as well as 'service sales', which include revenue from third-party sales. Therefore, we net out the third-party revenue using the share of third-party transactions and the margin earned on third-party transactions, μ^{3PS}, which we discuss below.

Amazon also reports the total 'Cost of Goods Sold', which is comprised of wholesale costs and shipping costs. We compute the cost of goods sold for North America by multiplying the total cost by the ratio of sales from North America to total sales. Because we estimate shipping costs, we need to subtract these costs from the margin. Therefore, we net out the 'Shipping Costs' that are observed in the annual report (e.g., 2016 p. 25). Again, we adjust these by the share of sales that are from North America.

The "Other Costs" variable should not include any land or labor costs that we are going to estimate. So we collect the "Fullfillment Cost" from the annual report, which "consist of those costs incurred in operating and staffing our North America and International fulfillment and customer service centers and payment processing costs" (e.g., 2016 p. 27), and net out preliminary estimates of the costs of labor, land, and density from our model. We adjust this by the share of North American sales and the share of third-party sales in order to get 'Other Costs' for US Amazon transactions only.

The calculated value of μ^{own} varies from year to year, with a high of about 0.2 in the early years and a low of about 0.09 in 2011. We use the average across all years, approximately 0.15, and assume that this is the margin that Amazon realizes on direct transactions, net of any costs we estimate.

Next we turn to the margin Amazon earns on third party transactions, which includes both a royalty rate and seller fees such as membership fees and stocking fees. It is not necessary for us to observe these separately, as we only care about the overall margin on third-party sales. In order to determine μ^{3PS}, we use the amount of revenue Amazon earned from third party sellers from 2014-2018, which, along with the Amazon's total revenue and the share of third-party sales, allows us to back out μ^{3PS} for these years. The revenue from third-party retailers is observed starting in the 2016 annual report, which reports this measure 2 years retroactively (p. 68). The backed out margin ranges from 0.32 to 0.35 during this time frame, so a reasonable estimate may be 0.335. However, because it is likely more costly for Amazon to receive, manage, and ship inventory from a third-party retailer, we use a conservative estimate of $\mu^{3PS} = 0.3$.

[43] According to S & P Capital Platform's segment analysis of Amazon, nearly 98% of North American revenue came from the United States in 2017.

Therefore, the weighted average margin is given by:

$$\bar{\mu}_t = s_t^{3PS} 0.30 + (1 - s_t^{3PS}) 0.15$$

B Appendix: Estimation

B.1 Projecting Spending

Here, we describe the steps to project demand outside of our sample using additional data. First, we collect data on the variables in C_{it} and Z_{it} for the years and counties outside of our sample from the US Census Bureau. Second, to predict the county fixed effects we use the estimates of a linear regression of the estimated in-sample fixed effects on a large number of county characteristics. This auxiliary regression fits the data well with an R-squared of 0.89. We also use a regression with time-trend covariates to predict the census division-year fixed effects. We use the predictions of the county fixed effect and the census division-year regressions for both in-sample and out-of-sample data, which smooths the spending data to account for possible measurement error.

Finally, in order to predict the mode-year fixed effects outside of our sample, we bring in aggregate spending data from Amazon's annual reports and the US Department of Commerce. Specially, we can use the estimates of the demand model and the projections discussed above to predict total yearly spending for a given set of fixed effects. So we find the values of the fixed-effects such that these predictions match the information from the aggregate spending data.

B.2 Instrument Construction

Here, we describe how we construct the instruments. Recall, that the instruments must be orthogonal to the measurement error from the demand model, but correlated with the components of the profit difference. Therefore, we construct measures of these components that are not a function of the estimated demand model.

The first variable, a shifter of the shipping distance component, is the total weighted shipping distance difference, where instead of using the number of orders predicted by the demand model as the weight, we use the population of the county:

$$\hat{X}_d^{j,j'} = \sum_{t=t(j)}^{t(j')} \beta^t \left(\sum_{i=1}^{I} \sum_{l=1}^{L} \mathrm{Pop}_{it} \hat{\Omega}_{i,l}(N_t | \mathbf{a}^0) d_{il} - \mathrm{Pop}_{it} \hat{\Omega}_{i,l}(N_t | \mathbf{a}^{j,j'}) d_{il} \right)$$

The function $\hat{\Omega}_{i,l}(N_t|\mathbf{a})$ represents the estimated O-D matrix under a network strategy $\mathbf{a}$.

The second, a shifter of the vertically integrated orders component, is the weighted number of vertically integrated orders, where again we use county population instead of the estimated number of orders:

$$\hat{X}_{vi}^{j,j'} = \sum_{t=t(j)}^{t(j')} \beta^t \left(\sum_{i=1}^{I} \sum_{l=1}^{L} \mathrm{Pop}_{it} \hat{\Omega}_{i,l}^{sc}(N_t | \mathbf{a}^0) - \mathrm{Pop}_{it} \hat{\Omega}_{i,l}(N_t | \mathbf{a}^{j,j'}) \right)$$

Shifters of gross profit differences include the differences in the two average input prices (wages and rent):

Table B-1: Definition of moment conditions

Trade-offs	Variables	Nb. Swaps: $Z_r^{j,j'} = 1$	
Distance & Taxes	$\hat{X}_d^{j,j'}$ & $\widehat{\Delta \text{Tax}}^{j,j'}$	83 (-/+)	1857 (+/-)
Distance & Cost	$\hat{X}_d^{j,j'}$ & $\widehat{\Delta \text{Input prices}}^{j,j'}$	648 (-/+)	2282 (+/-)
VI & Taxes	$\hat{X}_{vi}^{j,j'}$ & $\widehat{\Delta \text{Tax}}^{j,j'}$	133 (+/+)	967 (-/-)
VI & Cost	$\hat{X}_{vi}^{j,j'}$ & $\widehat{\Delta \text{Input prices}}^{j,j'}$	1037 (+/+)	1307 (-/-)
Distance & VI	$\hat{X}_d^{j,j'}$ & $\hat{X}_{vi}^{j,j'}$	606 (-/-)	1883 (+/+)
Distance & Density	$\hat{X}_d^{j,j'}$ & $\hat{X}_p^{j,j'}$	566 (-/+)	2100 (+/-)
Taxes alone	$\widehat{\Delta \text{Tax}}^{j,j'}$ + $\lvert \hat{X}_d^{j,j'} \rvert < \sigma_{vi}$ & $\lvert \hat{X}_{vi}^{j,j'} \rvert < \sigma_d$	119 (+)	989 (-)
Fraction of swaps		0.79	

$$\Delta \text{Input prices}^{j,j'} = \sum_{t=t(j)}^{t=t(j')} \beta^t \frac{1}{n_t^0} \left(IC_t(N_t|\mathbf{a}^0) - IC(N_t|\mathbf{a}^0) \right)$$

where $IC_t(N_t|\mathbf{a})$ is the sum of the input cost (either wage or rent) across active clusters under strategy $\mathbf{a}$. Note the input cost of a cluster is the average across the facilities in that cluster, if those facilities are in different counties. When all facilities within a cluster are in the same county, then it is just the input cost in that county.

The shifter of the density cost is difference in average population density:

$$\Delta \text{Density}^{j,j'} = \sum_{t=t(j)}^{t=t(j')} \beta^t \frac{1}{n_t^0} \left(\bar{Dens}_t(N_t|\mathbf{a}^0) - \bar{Dens}(N_t|\mathbf{a}^0) \right)$$

where $\bar{Dens}_t(N_t|\mathbf{a})$ is the sum of the population density across active locations.

Finally, as an additional shifter of the gross profit difference, we construct the difference in the population weighted average tax rate between strategies $\mathbf{a^0}$ and $\mathbf{a}^{j,j'}$:

$$\Delta \text{Tax}^{j,j'} = \sum_{t=t(j)}^{t=t(j')} \beta^t \left(\sum_{i=1}^{I} \frac{\text{Pop}_{it}}{\text{Total pop}_t} \cdot \left(\tau_{it}\left(\mathbf{a}^0\right) - \tau_{it}\left(\mathbf{a}^{j,j'}\right) \right) \right).$$

where $\tau_{it}(\mathbf{a})$ is the tax rate charged on Amazon in county i under strategy $\mathbf{a}$.

Table B-1 formally defines the indicator variables that we use as instruments in the estimation, as well as the number of permutations associated with each one. The last two columns report the number of swaps used for each moment. Overall, we use 79% of all available swaps in the estimation. The (+) and (-) signs beside each entry indicate the sign of the first and second variables generating the trade-offs. Recall that differences are expressed relative to the rejected option of opening facility j' early (and delaying j). A positive sign indicates that the chosen option leads to an increase in the variable of interest.

The first row demonstrates the trade-off between distance and tax. We observe 83 swaps leading to negative distance and positive tax changes, and 1,857 swaps leading to an increase in distance and a tax decrease. Note that we observe a larger number of distance increase/tax decrease swaps because most new entries in dense areas take place late in the sample, and generate little tax changes. In contrast we observe a large number of swaps generating trade-offs between cost and distance (second row).

The trade-offs associated with the expansion of the sortation network correspond to swaps leading to positive differences in vertical integration and tax (or cost), and negative differences in vertical integration and tax or cost. These swaps correspond to cases in which Amazon chose to increase the number of vertically integrated transactions at the expense of lower revenue or higher input prices.

The next two trade-offs capture the joint variation in distance and vertical integration, and distance and population density. Instruments associated with choices that increase both distance and vertical integration impose restrictions on the ratio of θ_d over θ_{vi}. The trade-off between distance and population density similarly restricts the ratio of θ_d and the effect of density on fixed-costs (κ).

Finally, we include a pair of moments associated with variation in taxes, and minimal changes in distance and vertical integration (i.e. difference is smaller than the inter-quartile range in both variables). These additional moments produce positive and negative changes in gross profits, which impose restrictions on the scale of the fixed-cost parameter (κ).

B.3 Model Fit

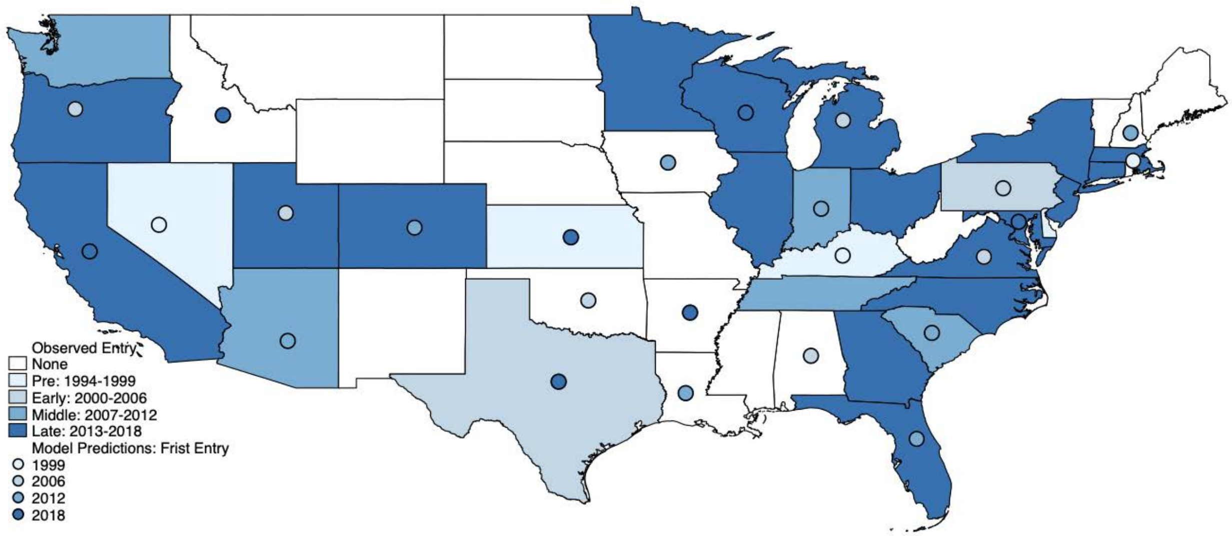
We present the fit of the model's predictions of the roll-out in Figure B-1, which displays information about the timing of observed first entry into a state (shade) and the corresponding predictions of our model (points).

For the low-demand state (1999), the model accurately predicts the regional distribution of active fulfillment centers, as the predicted optimal configuration includes one location each in the west (Nevada), the center (Kentucky) and the east (Massachusetts; note however that the observed network's location is in Delaware). As demand increases (2006 and 2012), the density and number of facilities expand rapidly. In 2006, the model predicts a slightly more dense network than the actual network, with locations in additional states in the mid-west (Michigan), the southeast (Alabama), the mid-Atlantic (Virginia), and the west (Oregon). Though, the model accurately predicts entry on the east-cost (Pennsylvania) and south (Oklahoma). However, the observed entry in the south was in Texas rather than Oklahoma.

The 2012 network is similar to the observed one, with expansions in the mid-west (Indiana), south (South Carolina) and west (Arizona). The model also predicts entry into Florida and Colorado during this time, which is slightly earlier than what we observe.

For 2018, further entry is predicted in the west (California), mid-Atlantic (Maryland) and mid-west (Wisconsin), which are all entries that are observed. However, there are some states which the model misses in the mid-west (Illinois), east (New York), and southeast (Georgia).

Figure B-1: Observed (shade) and Model Predicted (dot) First Entry into State



Notes: The dots are places at the centroid of the state and not at a particular location.