Economies of Density in E-Commerce: A Study of Amazon's Fulfillment Center Network *

Jean-François Houde
Department of Economics
Cornell University

Peter Newberry
Department of Economics
The Pennsylvania State University

Katja Seim
The Wharton School
University of Pennsylvania

April 18, 2017

Abstract

We examine the economies of density associated with the expansion of Amazon's distribution network from 2006 to 2018. We demonstrate that, in placing a fulfillment center in a new state, Amazon faces a trade-off between the revenue implications of exposing local customers to sales tax on their purchases and the cost savings from reducing the shipping distance to those customers. Using detailed data on online transactions, we estimate a model of demand for retail goods and show that consumers' online shopping is sensitive to sales taxes. We then use the demand estimates and the spatial distribution of consumers relative to Amazon's fulfillment centers to predict revenues and shipping distances under the observed fulfillment center roll-out and under counterfactual roll-outs over this time period. Using a moment inequalities approach, we infer the cost savings from being closer to customers that render the observed network roll-out optimal. We find that Amazon saves between \$0.17 and \$0.47 for every 100 mile reduction in the distance of shipping goods worth \$30. In the context of its distribution network expansion, this estimate implies that Amazon has reduced its total shipping cost by over 50% and increased its profit margin by between 5 and 14% since 2006. Separately, we demonstrate that prices on Amazon have fallen by approximately 40% over the same period, suggesting that a significant share of the cost savings have been passed on to consumers.

Keywords: e-commerce, sales tax, Amazon, distribution, logistics, shipping
JEL Codes: H71, L11, L81

*All correspondence may be addressed to the authors via e-mail at `houde@cornell.edu`, `pwnewberry@psu.edu`, or `kseim@wharton.upenn.edu`.

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, and the Dean's Research Fund for financial support. 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.

1 Introduction

The Internet might have made it possible for the smallest of businesses to sell their wares, but the price of delivering cardboard boxes highlights just how powerful an advantage size can be in the Internet economy.[1]

Online retail has grown substantially over the last decade and makes up nearly 8.5% of all retail sales as of 2016.[2] Amazon.com, henceforth Amazon, is a key contributor of this growth, with net sales increasing from $10.7 billion in 2006 to $107 billion in 2015.[3] Amazon's rise has been accompanied by a substantial increase in concentration in online retail, as Figure 1 illustrates. This trend counters initial expectations that the internet would facilitate highly competitive markets due to lower search costs, lower fixed costs on the part of sellers whose physical locations are less relevant to serving customers, and an increasing reliance on an already existing network of distribution and shipping companies whose services could help to partially replace the offerings of brick-and-mortar stores. In this paper, we explore the possible sources for the increase in concentration with a focus on the competitive advantage that comes with the cost savings of a large-scale distribution network.

We investigate the economies of scale in distribution using Amazon as a case study. As Amazon's revenue has grown, so has its network of distribution centers, so called fulfillment centers, which we denote as FCs going forward. The number of FCs has grown from 8 centrally located centers in 2006 to 90 spread out across the US by the end of 2016.[4] Amazon will further expand its network to over 100 facilities by 2018, including "Amazon Now" hubs that provide same-day delivery to local customers. A larger distribution network has implications for firm profitability on both the demand and supply sides.

The expansion of the network has the potential to affect consumer willingness-to-pay through two opposing demand-side forces. First, if consumers value the convenience of receiving packages faster and expansion leads to shorter delivery times, then such convenience effects constitute a local quality attribute of Amazon's service that may in part be responsible for the firm's increasing market share. Second, under a 1992 U.S. Supreme Court decision (Quill Corp. vs. North Dakota 504 U.S. 298), a mail-order retailer, such as an online seller, is required to collect sales tax on the customer's behalf if the retailer has a physical presence or "nexus", including a warehouse, in the customer's state of residence. Fulfillment center additions, if they entail expansion into a new state, may thus increase sales tax inclusive prices on Amazon. As earlier literature (Einav et al., 2014; Baugh et al., 2014) shows, demand is sensitive to sales taxes as they raise the effective price of an online transaction.

[1] See "FedExs Price Rise Is a Blessing in Disguise for Amazon," New York Times 5/9/2014, `http://nyti.ms/1fUabBh`, accessed on 1/17/2017.

[2] See `www.census.gov/retail/mrts/www/data/pdf/ec_current.pdf`, accessed on 1/12/2017.

[3] See annual reports at `http://phx.corporate-ir.net/phoenix.zhtml?c=97664&p=irol-reportsannual`.

[4] These figures do not include Amazon Fresh FCs or sortation centers. See Section 2 for details on distribution center types.

To limit the revenue implications of exposing customers to new sales tax liabilities, a supply-side response would thus be not to establish a presence in a high-tax state, especially if it were populous. In addition, a new fulfillment center adds fixed costs of opening and operating the facility, relative to a smaller network. Adding a distribution center also makes the existing distribution network more dispersed, shortening the distance to a subset of customers. This can translate into shipping cost savings in one of two ways. First, Amazon's distribution facilities may be closer, on average, to its shippers' local sorting facilities, reducing the nodes through which packages travel in the shipment process. This reduces the shippers' cost, providing Amazon with bargaining power in negotiations with shippers over the per-package rates they charge. Second, by being closer to its customers, Amazon has increasingly been able to enter into the local delivery market itself, replacing traditional shippers in last-leg deliveries in large urban markets near its distribution centers. These economies of density significantly affect Amazon's overall profitability as shipping comprises a large share of Amazon's operating costs.[5,6] On net, then, the implicit tax-related costs of building a warehouse in a new state, together with fixed cost differences, offset the benefits of expansion due to proximity to the consumer.

We empirically quantify this trade-off between higher tax-inclusive prices and fixed costs, on the one hand, and additional convenience and shipping cost savings on the other, while controlling for a potential increase in platform quality through, for example, an expansion of the product assortment. To accomplish this, we combine data on the purchase behavior at Amazon and other online competitors by a large, representative consumer sample with detailed information on the location and characteristics of Amazon's fulfillment centers over time. We take a revealed preference approach similar to that of Holmes (2011) that uses, as one source of identification, the fact that expansion of the FC network into a new state results in lost revenue due to new sales tax liabilities. Therefore, observing such an expansion in the data must mean that it also results in either revenue offsets due to increased willingness-to-pay for faster deliveries, or a reduction of shipping costs sufficient to outweigh the tax-driven revenue declines.

The first step in this approach is to estimate a model of demand in order to pin down consumer tax sensitivity and response to shipping times. We use data on both online and offline purchasing of retail goods that we construct by combining the comScore Web Behavior database, Forrester Research surveys, and household spending data from Environmental Systems Research Institute (ESRI). In the model, a representative consumer in a county chooses her yearly expenditures on four different modes of shopping: (1) Amazon.com, (2) a taxed online competitor (e.g., walmart.com), (3) a non-taxed online competitor (e.g., overstock.com), and (4) a taxed offline competitor (e.g.,

[5] Amazon's financial statements report net shipping costs between 3-5% of net sales for the period 2006 to 2015. These costs totaled over $5 billion in 2015. See annual reports at `http://phx.corporate-ir.net/phoenix.zhtml?c=97664&p=irol-reportsannual`.

[6] Expansion also affects the distance of "inbound" shipments from the goods suppliers to Amazon's FCs. As we discuss below, suppliers are shipping goods to brick and mortar retail outlets all around the country; as a result, we assume that these costs do not change significantly as Amazon expands its FC network.

Wal-Mart). The amount of sales tax charged to the consumer varies both across time and county due to FC entry and local sales tax laws, implying that variation in spending across these dimensions identifies the consumers' sensitivity to sales tax. Shipping times vary separately from sales taxes due to the variation in the locations of consumers relative to the new FC, allowing for identification of the convenience effect. Similar to identification in a difference-in-difference approach, we use fixed effects to control for time varying mode preferences and county-specific determinants of spending.

The results of the demand estimation suggest that consumers are sensitive to taxes and that their tax elasticity is around -1.4, similar to the estimates from Einav et al. (2014) at -1.8 and Baugh et al. (2014) at -1.5. The magnitude of this estimate implies that moving from a non-taxed regime to a taxed regime at the average sales tax rate of 6.5%, all else equal, results in reduction of expenditures at Amazon of 9.3%. We find little evidence that the entrance of a FC leads to an increase in demand for Amazon due to increased convenience. The lack of effect of shipping time on demand is likely due to the fact that Amazon's local shipping options and prices did not drastically change between 2006 and 2013, with any nationwide changes being accounted for through the mode-year fixed effects. Importantly, together these estimates imply that overall, there are negative revenue effects of opening a fulfillment center in a new state, which must be offset by a decrease in costs. We find that mean preferences for Amazon, relative to the offline shopping mode, increase steadily from 2006 to 2013, but are relatively constant for the other two online shopping modes, suggesting that price reductions and/or improvements in platform quality have made Amazon more attractive over this time period.

In the second step, we quantify the shipping cost savings from the larger FC network. We specify a profit function where the variable cost of shipping goods to consumers in a county is proportional to the shipping distance from the FC to that county, scaled by the revenue value of the shipment. We assume that the observed sequence of fulfillment center placements is optimal so that the discounted profit stream under the observed network roll-out must be at least as large as under any alternative sequencing of fulfillment center openings. The network configuration affects revenue through the sales tax channel, fixed costs through the wages and rents paid in the counties where Amazon FCs are located, and variable costs through the total shipping distance across customers. We compute perturbations to the observed network roll-out similar to those of Holmes (2011): to construct alternative network roll-outs, we rely on swaps of fulfillment center opening dates that have offsetting effects on profitability through revenue and shipping cost effects. For example, moving up the opening date of a FC in a low-populated, low-tax state and in turn delaying the opening of a FC in a more populous, high-tax state implies both higher revenue streams due to lower exposure to sales tax early on, but longer shipping routes and thus higher cost.

Relying on revealed preference, a profit comparison of the actual and such counterfactual roll-outs allows us to impose bounds on the firm's shipping cost per dollar per mile such that the observed roll-out dominates its perturbations. We estimate these bounds on the effect of shipping distance on variable cost using a moment inequalities estimator based on the methods of Pakes et al. (2015) and Andrews and Soares (2010). In doing so, we rely on similar instruments to Holmes' (2011), adjusted for the nature of our problem where any measurement error in revenue affects both the difference in pre-shipping profit and in revenue-scaled shipping distance and thus total shipping cost under the actual and perturbed FC roll-outs. We use the bounds estimates to compare cost under the enlarged and original distribution networks, estimating the cost saving of network expansion without having to solving the dynamic problem that Amazon faces.

Our estimates imply that it costs Amazon between \$0.17 and \$0.47 to ship \$30 of goods, the median order value in our data, 100 miles, net of shipping revenue. These estimates are robust to alternative choices of instruments, various proxies for shipping cost, and the inclusion of preferences for same-day shipping. We use the estimates to perform three exercises. First, we quantify the effects of the distribution network expansion on revenue, fixed costs, and shipping costs from 2006 to 2018. We find that expansion led to a \$9.6 billion (\$2.2 billion) decrease in revenue (profit net of shipping cost) due to consumers' increased tax exposure and a \$3.3 billion increase in the fixed costs of operating the network of FCs. However, Amazon saved between \$5 billion and \$13 billion in net shipping costs by decentralizing its distribution network, implying a net profit increase of between 500 million and \$8 billion over this time period. Second, we investigate the likely long-run effects of expansion, assessing the effect of shipping cost savings on the firm's profit margin at the end of the planned network expansion that we observe in the data. Our results suggest by 2018, the firm's margin will be between 5 and 14% higher than what it would have been had the distribution network remained at its 2006 configuration until that point. Finally, our bounds estimates imply that the average total net shipping-related cost per order, including both the fixed cost of operating the distribution center network and the variable shipping costs, decreased from between \$2.69 and \$4.67 per order in 2006 to between \$1.09 and \$2.14 in 2013.

To relate the estimated shipping cost savings to the observed increase in concentration, we decompose the estimated time-varying aggregate preferences for Amazon into a price component and a quality component. We find that, while Amazon's relative quality increased over our sample period, its relative prices fell by approximately 45%, mirroring the estimated 55% decline in shipping costs. While our model is not formulated to predict the precise price adjustment in response to a decline in cost – or an increase in quality – this provides suggestive evidence that Amazon's expansion has granted the firm a competitive advantage through a more efficient network. Correlating the quality contribution to Amazon's mean preferences and the variety of product categories that the platform covers over time further suggests that the increase in Amazon's quality went hand-in-hand with an expansion of the variety of goods that the platform sells, which many cite as an explanation for Amazon's success.[7]

This paper is related to several strands of literature. An increasing body of work focuses on the

[7]See http://www.nytimes.com/2009/09/20/business/20amazon.html, accessed on 1/12/2017.

estimation of demand in online retail markets, with the studies focusing on the effects of sales tax and the gains from variety being the most relevant to our study. 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. They are also able to estimate the consumer sensitivity to distance from observing the locations of both buyers and sellers. Baugh et al. (2014) uses a differences-in-differences approach to estimate the effect of the 'Amazon tax', or changes in various states' laws between 2013 and 2015 that forced Amazon to collect sales tax. In estimating the tax sensitivity, neither set of authors is able to fully consider substitution to other taxed or non-taxed online outlets. We thus contribute to this literature by expanding our analysis of the tax sensitivity and the estimation of demand for retail goods beyond a single online firm to a large number of online and offline retailers.

Our focus on product variety as a dimension of Amazon's quality is motivated by work examining the gains from variety in e-commerce, including Brynjolfsson et al. (2003) and Quan and Williams (2016). Brynjolfsson et al. (2003) uses data from online bookstores to demonstrate that consumer welfare gains from increased product variety far outweigh the also sizable gains due to lower prices and increased competition. Findings in Quan and Williams (2016) suggest that the extent of such gains varies significantly across local markets, depending on the availability of competing, localized, assortment in brick-and-mortar stores. This highlights the importance of controlling for localized variation in the attractiveness of the offline – and other online – shopping modes in estimating demand, as we do in assessing the contribution of quality differences across shopping channels to demand, which we show to largely reflect variation in product variety that provides large online retailers such as Amazon with a clear advantage over their smaller competitors. To our knowledge, we are the first to investigate the possible convenience effects associated with a broader distribution network that reduces delivery times, a quality attribute that consumers may value independently of variety differences across shopping modes.

Our analysis is also related to recent operations research and economics literature on management of distribution networks for online firms (see Agatz et al. (2008) for an overview) and on the relationship between a brick-and-mortar retailer's store location choices and its distribution network. Zheng (2014) relates the proximity of a rival's fulfillment center to the chain's expected future entry. Building on Barwick (2008) who estimates the scale economies associated with operating multiple stores in close geographic proximity, without decomposing the sources of such scale economies, Holmes (2011) directly estimates the savings in distribution costs associated with clustering stores near a fulfillment center. These studies take the configuration of the distribution network as given and use variation in the distances from a fulfillment center to potential store locations to identify the parameters of interest. Instead, we study the development of the network of FCs 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 Amazon's FC network and summarizes relevant aspects of sales tax laws. Section 3 introduces the main data sources and Section 4 presents the demand side analysis and results. Section 5 provides estimates of distribution cost savings and Section 6 concludes.

2 Amazon's FC Network and Sales Tax

We obtain information about Amazon's fulfillment centers from the supply-chain consulting company MWPVL, International (`http://www.mwpvl.com/`). MWPVL provides information on the location, size, opening date, and closing date of each FC. Also observable for a subset of the locations is the fulfillment center 'type'. The type of FC is usually defined by the size of the items being shipped and/or the speed of delivery. The primary types are centers that focus on large items that cannot be sent combined with any other products (non-sortable), small items that can be combined in one package (small sortable) and large items that can be combined into one package (large sortable). Other types of distribution centers include Amazon Fresh FCs, which supply Amazon's grocery delivery orders, return centers, redistribution centers for third-party distribution, and centers for select specialty items such as jewelry. Starting in 2014, Amazon started to build 'sortation' centers, which are used to sort packages by zip code areas after they have shipped from a FC, and 'Prime Now Hubs', which handle same-day delivery for select local markets. For the majority of the analysis, we focus on the three primary types of FCs and Prime Now Hubs, as these are the centers that ship non-grocery items directly to consumers.

For these types, Amazon has expanded from 8 FCs in 2006 to 90 by the end of 2016 and will further to over 100 FCs by 2018. Table 1 and Figure 2 demonstrate this expansion. Table 1 displays the number of FCs, the numbers of states that have a FC, the number of taxed states and associated counties by year, and the average distance between the closest FC to a county, weighted by county's population. In Figure 2 illustrates the trends in the table, shading states by numbers of households and displaying the population-weighted average sales tax in 2013, the last year of our online demand data. A few aspects of the FC expansion until 2010 are noteworthy. First, Amazon placed FCs in relatively low population states that were close to highly populated areas. For example, Amazon opened two FCs in Nevada, both of which were on the California border close to that state's major cities. Second, they also placed FCs in states with relatively low sales tax. For example, the company opened a FC in New Hampshire, in addition to the one it already operated in Delaware, both of whom are close to major East Coast cities and have zero sales tax. Such choices of states are also advantageous since sales tax rates are positively correlated with population (across states, correlation of between 0.35 and 0.4 across 2006 to 2013), so that entry into a small state near a large state has limited tax implications for only a small population, while allowing the firm to serve both states' populations more efficiently. Third, when Amazon did expand to highly populated states, they initially focused on states with a relatively low sales tax rates (e.g., Pennsylvania with an average tax rate of 6.1% in 2013, compared to 6.9% across the top 20 US states in terms of population and a maximum in those states of 9.5% (TN)). For comparison purposes, we include in the last panel of Figure 2 a map of Wal-Mart's distribution centers as of 2014, which are spread across the country more evenly than Amazon's FCs, at least in the early years.

These patterns suggest that the early strategy of Amazon's FC expansion exploited cross-state variation in sales taxes combined with the fact that legally, e-commerce firms are only obligated to collect sales tax in certain instances. A 1992 U.S. Supreme Court decision concluded that the U.S. Constitution's Commerce Clause "prohibits a State from imposing the duty of use tax collection and payment upon a seller whose only connection with customers in the State is by common carrier or by mail." The court ruled that a mail-order retailer, including an e-commerce firm, does not have to collect sales tax from consumers unless it has a physical presence, or a 'nexus', in that consumer's state of residence. In such instances, it is instead the duty of consumers to file a 'use-tax' return every year, which calculates tax liability on purchases from out-of-state vendors. Few individuals comply with this rule.[8]

A large and growing literature examining the effect of sales tax on online purchasing has found significant consumer responses to sales tax rates using, for example, variation in tax rates across municipalities (Goolsbee, 2000a; Goolsbee, 2000b). Both Alm and Melnik (2005) and Ballard and Lee (2007) find small but significant effects of sales tax on the decision of whether or not to shop online, while Scanlan (2007) finds that this sensitivity is heterogeneous across the level of tax rates. Ellison and Ellison (2009), Smith and Brynjolfsson (2001), Anderson et al. (2010), and Goolsbee et al. (2010) all find that online shoppers are sensitive to sales tax, typically focusing on purchases in a single product category.[9]

For most online firms, a physical presence would take the form of an office headquarter or a fulfillment center, implying that these firms likely do not have to collect sales taxes in many states. As the popularity of e-commerce has grown, policy makers have suggested that the differential tax collection responsibilities may be giving online firms, and in particular Amazon as the dominant online retailer, an unfair advantage over their brick-and-mortar competitors. Given the low rates of out-of-state purchase reporting for use tax compliance by consumers, states are also likely losing out on millions of dollars of tax revenue (see Bruce et al., 2009).

[8] In the 45 states that have a use tax, only about 1.6 percent of taxpayers paid it in 2013. See NPR, "Most People Are Supposed To Pay This Tax. Almost Nobody Actually Pays It." April 2013, `http://www.npr.org/sections/money/2013/04/16/177384487/most-people-are-supposed-to-pay-this-tax`, accessed on 1/12/2017. This compares to an estimated 60.3 percent of the US population who made online purchases in 2013 according to eMarketer 2016, albeit not necessarily all from out-of-state retailers.

[9] There is also evidence on the response of consumers to sales taxes in offline markets, analyzing the substitution of shopping expenditures across state and country borders (Agarwal, Chomsisengphet, Ho, and Qian, 2013; Asplund, Friberg, and Wilander, 2007) and over time due to tax holidays (Agarwal, Marwell, and McGranahan, 2013). Chetty, Looney, and Kroft (2009) estimate how the saliency of tax rates affects consumer demand.

As early as 2008, states began passing legislation that expanded the definition of 'nexus' to apply to locations of the firm's 'affiliates', typically websites that allow retailers, such as Amazon, to advertise on their site. For example, if an Illinois blogger had a link to Amazon on her site, then Amazon would have to charge sales tax to all Illinois residents. Not surprisingly, Amazon and other big retailers responded to such laws by shutting down their affiliate programs in the relevant states.[10] Amazon acknowledged in its 2008 annual report, page 16, 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 ... decrease our ability to compete with traditional retailers, and otherwise harm our business."

As Amazon grew in scale, however, the network of FCs expanded beyond low-tax low-population states, 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 highly populated states, such as California and Virginia, and high tax states, such as Tennessee (9.5% tax rate in 2013). Finally, by 2018, Amazon will have added FCs in Georgia, Illinois, North Carolina, and Ohio. Overall, the pattern of expansion and location choices imply that there exists a trade-off between being close to customers and charging them sales tax. Our main goal is to empirically investigate this trade-off.

It is important to note that the opening of a fulfillment center does not necessarily imply that Amazon begins charging sales tax immediately thereafter. For example, Amazon first built a FC in Pennsylvania in 2006, but did not begin to charge sales tax until 2011. Such a delay is often due to legal battles with the state government as to what constitutes a nexus. On the other hand, sometimes Amazon charges sales tax even when the company does not have a FC in that state. This can be due to changes in state laws (e.g., New York) or because of legal agreements with the state to begin charging sales tax ahead of the opening of a FC (e.g., Connecticut). The latter explain the, at times, significant discrepancy between the number of states where Amazon's customers pay sales taxes and the number of states where Amazon has a FC. In our analysis below, we consider various assumptions on Amazon's perception of the relationship between FC entry and sales tax collection when making its network location decisions.[11]

[10] See e.g., `http://techcrunch.com/2011/06/10/amazon-shuts-down-associates-affiliate-program-in-connecticut-over-online-sales-tax/`, and `http://www.kansascity.com/news/local/article325412/Amazon-shuts-down-Missouri-associates-program-over-sales-tax-dispute.html` accessed on 1/12/2017.

[11] In addition to negotiating with states over the effectiveness of its tax collection obligation, Amazon at times receives other forms of government financial assistance when opening a FC in a new locality. The online government subsidy data base `Goodjobfirst.org` lists 39 different economic development programs that Amazon benefitted from during 2006 and 2013. The vast majority consist of ongoing tax credits and training cost reimbursements, which affect the firm's variable profit. Six consist of grants or reduced interest loans. Since we do not observe Amazon's ultimate take-up of such measures, we incorporate their effect on Amazon's choice of when to enter a location via a measurement error component to variable profit.

3 Data

Consumer Purchases

The primary data source for the estimation of the demand model is the comScore Web Behavior Database. ComScore tracks the online purchasing and browsing activity of a random sample of internet "users". The users give comScore explicit permission to monitor their browsing and purchasing activity. Each user is associated with a machine (computer) and any activity on that machine, including activity from other users, is recorded. Therefore, we assume that each user is equivalent to a household.

For each purchase transaction, we observe a unique machine identifier, the time of the purchase, the product category and price for each individual item in the basket, the name of the domain where the transaction occurred, and a 'basket total', or the total cost to the consumer for the transaction including shipping and taxes. In addition, we observe demographic characteristics for each household such as age and race of the head of household and household income, together with the zip code of their residence.[12]

Below, we use the comScore data to predict revenue that Amazon, and competing online and offline retailers, derive from each location in the country in each year. We relegate the details of how we construct household-level expenditures in each location to Appendix A. Here, we only discuss the primary data patterns that guided our approach.

The comScore sample begins with 86,000 households in 2006, but shrinks over the years to 46,000 in 2013. Nevertheless, every state in the continental US, the District of Columbia, and at least 78% of US counties are represented in the data each year. See columns 1 and 2 of Table 3. Because the coverage of finer geographic areas is limited – for example, at most 42% of zip codes are covered in any given year – we focus on counties as our geographic delineation for the majority of the analysis.

We construct household-level online and offline expenditures for a representative household in each county by supplementing the comScore data with survey data from Forrester Research and data from the US Census. The Forrester data allow us to account for households who have zero online expenditures (i.e., the extensive margin), while the Census data provide us with weights to correct for the possibility that the comScore sample is not representative of a given county's population.

We consider three online shopping channels, which we denote as 'modes': Amazon, taxed online competitors of Amazon, and non-taxed online competitors of Amazon. We restrict online retail to

[12]The comScore data contain a categorical income variable. To simplify the regression analysis below and facilitate matching the comScore data to auxiliary data sources, we convert the categorical income data to a continuous income variable. We do so by fitting county-specific log-normal distributions to the categorical income distribution by race and age of the head of the household from the 2010 US Census. We then use the county, race, and age-specific fitted income distributions to calculate expected income levels for each comScore income category and assign to each household the expected income within its reported income bin.

product categories that Amazon carries, dropping spending at, for example, dating websites (e.g., match.com), travel websites (e.g., orbtiz.com) and food delivery sites (e.g., dominos.com). Within the resulting universe of competing websites, we define a taxed online competitor as one that has a large offline presence such as gap.com, walmart.com and target.com, and thus already collects sales tax in most states, and a non-taxed competitor as one without a national offline presence, such as overstock.com. We note that neither of these groups are perfectly defined, as some sites that we classify as having a national offline presence may in fact not have a store in every state and sites that we classify as not having an offline presence surely have a headquarters and/or a fulfillment center located in at least one state. Table 2 displays the top ten online stores in each of the taxed and non-taxed categories. Overall, we classify about 34% of websites as taxed competitors.

When adjusted by the extensive margins in the Forrester data, the average comScore household spends \$250 annually on online retail, rising to \$374 by 2013 (see Table A.1). Aggregating across households to the representative household in each county and making use of the census weights results in the average household spending of \$217 in 2006 and \$273 in 2013 (see Table 3). The differences between this and the average spending reported in Table A.1 are likely due to ComScore oversampling households that are younger and have higher income. The increase in household spending on Amazon from \$23 in 2006 to \$100 in 2013 mirrors the general increase in the household's total online expenditures from 2006 to 2013, while spending on the other two modes is relatively steady throughout.

These spending patterns result in the evolution of market shares for the different shopping modes that we display in Figure 3. The market shares are based on household-weighted sums of representative household expenditures in each county. We illustrate that Amazon's market share of online expenditures grew from around 10% in 2006 to almost 40% in 2013, eating into the market shares of both the taxed and non-taxed online competitors.

There are three primary limitations of our constructed sample on consumer spending. First, comScore does not do not record Amazon Marketplace purchases separately from Amazon purchases. This has implications for both the estimation of tax sensitivity, as we cannot separate out expenditures at non-taxed Marketplace retailers, and the calculation of Amazon's total revenue. Second, many households use their computers at work and/or mobile devices for making some or all of their online purchases, so that we may miss a non-trivial amount of online activity. Third, both the comScore sample and the Forrester sample (except for 2006-2007) are drawn from the pool of consumers who are internet users. While the percentage of US households with internet access during our sample ranges from 63% to 75%,[13] we do not observe internet use by county and are unable to account for this margin when calculating the expenditures for the representative household in a county. We may therefore overestimate e-commerce expenditures by assigning to the representative household in each county the weighted average expenditure of households in the

[13] See `http://www.census.gov/topics/population/computer-internet/data/tables.ALL.html`, accessed on 1/17/2017.

comScore sample. We discuss these issues and how we deal with them in greater detail below.

In order to construct the amount of household spending at offline retailers for the types of goods in our sample, we use a combination of 2012 county-level data from ESRI and SimplyMap. ESRI provides a measure of average household spending on all retail goods, including food, while the SimplyMap provides household expenditures on food. Because we exclude food from our online sample, we subtract food expenditures from the ESRI total to calculate our measure of total retail spending. We then subtract the average household's online expenditures from total retail spending to calculate the amount of expenditures at offline retailers. The average spending on offline retail across households is between $17,000 and $18,000 per year. Since the ESRI and SimplyMap data are only available for one of the years in our sample – 2012 – we flexibly control for spending time trends in our empirical work below.

Amazon Fulfillment Center Network

We acquire information about each of the FCs from MWPVL, International. Based on the location of the FC, we calculate the great-circle distance between each FC's street address and the population-weighted centroid of every county in the US, and exploit the opening dates of the FCs to calculate the minimum distance by county and year over the set of operating centers in a given year. This distance serves as a measure of how far Amazon must ship the good between its FCs and consumers.

It is apparent from Figure 2 that as part of its FC roll-out, Amazon often opens FCs that are close to an existing FC, possibly to increase capacity at the location or to cluster FCs of different types near one another. Specifically, nearly 70% of the FCs have at least one other FC that is within 20 miles. Therefore, we define a FC cluster as a group of FCs whose locations are in the same state and within 20 miles of each other. In the analysis below, we assume that Amazon decides where and when to build a cluster of FCs, rather than each individual FC in the cluster and use the opening date and shipping distance to the first FC built in the cluster, reducing the number of FC locations from over 100 to 60. This does not affect the demand analysis because any tax collection obligations are triggered by the first FC built in a state.

A downside to assuming that Amazon serves each customer from the physically closest fulfillment center is that it does not recognize, for example, that access to different modes of transportation across fulfillment centers might mean that it is cheaper for Amazon to service a particular county location from a fulfillment center other than its closest. We investigate this possibility by predicting which FC would be the lowest-cost supplier for each county under the assumption that the per-pound distribution cost for Amazon is proportional to the average per pound cost for shipments in the Census' 2012 Commodity Flow Survey (CFS).

Using the CFS data on all individual shipments of commodities in the relevant products categories, together with information from the Department of Transportation on unit shipping cost by mode of transportation, we estimate the realized per-pound cost of a shipment for the average package between any origin and destination county. Under the assumption that this per-pound cost is proportional to Amazon's shipping cost, we use it to identify the lowest-cost fulfillment center for each county and year. See Appendix A for details. Not surprisingly, the closest FC is also the lowest-cost, based on this measure, for 73% of county-year observations. If we use the lowest-cost FC to calculate Amazon's shipping distance to each county, the household-weighted average shipping distance falls from 301 miles in 2006 to 162 miles by 2018, compared to 297 and 116 miles, respectively, based on assigning each county to the closest FC (see Table 1). The correlation between the two alternative measures of shipping distance is 0.98, suggesting that the assumption that Amazon relies on the closest FC does not materially affect our estimated shipping distances relative to the predicted lowest-cost option.

Because great-circle distance may not be a perfect proxy for shipping speeds, we also obtain the US Postal Service's shipping times between each three-digit zip code zone and find the minimum shipping time between each county and the three-digit zip code zone of each FC.[14] The Postal Service reports shipping times for four different classes of mail: first, priority, standard, and package. These classes differ slightly in the size and number of packages allowed in the delivery.[15] Since the variation in shipping days is limited, we transform shipping time into a categorical variable – slow and fast shipping speed, where the definition of slow depends on the class of shipping.[16] We were unable to find shipping times for the other shipping services used by Amazon – UPS and FedEx – but they are likely very similar to the Postal Service's.

MWPVL also provides information about the FC size in square feet and, for 30 of the FCs, the number of employees, which we employ as fixed cost shifters in our analysis below. To fill in the missing workforce information, we assume that the number of employees per square foot is the same for FCs of the same type. The last two columns of Table 1 provide the average number of employees and size for operating FCs in a given year. There is a general pattern of an increase in average size and employees early in the sample period, followed by a decrease as Amazon began to build the smaller Prime Now hubs starting in 2014. Note that we do not observe any information about the FCs for Amazon's competitors, so we cannot speak to competition among online firms in terms of their FC network choices.

Finally, we obtain information on the extent to which a given fulfillment center can be used to satisfy same day shipping orders during our sample period. We rely on various news sources to identify the counties and the corresponding date of implementation of Amazon's early version of

[14] We choose one three digit zip code per county after verifying that shipping times do not vary significantly within a given county.

[15] For a detailed description, see `http://pe.usps.com/BusinessMail101?ViewName=ClassesOfMail`, accessed on 1/17/2017. Amazon uses First Class and Priority mail for standard shipments. See `https://www.amazon.com/gp/aw/sp.html?mp=&oid=&s=A1PVXLBETRT7XG&t=shipping`, accessed on 1/17/2017.

[16] Slow is defined as more than 2 days for "Priority" mail, more than 1 day for "First Class" mail, and more than 3 days for both "Package" and "Standard" mail.

same day shipping, 'Local Express Delivery'. Amazon introduced this service, for which customers are charged a fee, in select markets as early as October, 2009.

Other Sources of Data

We obtain state, county, and local sales tax rates from Tax Data Systems, now part of Thomas Reuters. For each year and county, we calculate the average tax rate, as tax rates can vary within a county and may change mid year. The average sales tax rate is 6.5%, with an average standard deviation of 1.6% across counties in every year and an average standard deviation of 0.3% across counties within a state and year. In addition, tax rates vary across time, as between 30 and 65% of counties experience tax rate changes from year to year and nearly 78% of counties experience at least one tax rate change over our sample period.

We use county level demographics from the 2010 Census and annual information on the number brick-and-mortar retailers and the number of large retailers, which we define as those with more than 50 employees, from 2006-2013 County Business Patterns. Further, we collect county-level wage information as the average annual wage of a 'Retail Trade' employee across the BLS's 2012 Quarterly Censuses of Employment and Wages. While this is not an exact match to the distribution and warehousing jobs of Amazon FC employees, we believe retail trade to be the closest match; the average hourly wage of a 'Retail Trade' employee is similar to that of an Amazon Associate as reported by `glassdoor.com`.

Next, we obtain a measure of the cost of renting a square foot of warehousing space using data from SNL Financial's real estate research. According to Amazon's 2013 financial statement, more than 99% of its fulfillment space is leased and not owned, so that the fixed costs of operating the distribution network include significant rental costs. From SNL, we observe a May 2016 snapshot of the rent paid and square footage of properties classified as "Warehouses or Distribution Centers". Using these, we calculate the average rent per square foot for all counties appearing in the SNL data. Some Amazon FCs are in counties that do not appear in the SNL data (26 out of 60). We estimate rent for those counties as follows. For each county represented in SNL, we calculate the ratio of commercial rent to residential property value and average across counties within a state, where we obtained residential property values from the 2013 American Community Survey.[17] For the counties not in the SNL data, we then estimate the rent per square foot as their residential property value times this ratio.

Finally, we rely on Amazon's annual reports for additional information about Amazon's finances. For each year, we observe the North American sales figures in the "Media" and "Electronics and Other General Merchandise" categories. In order to match the comScore data, we exclude the "Other" category, which is revenue from "Non-Retail Activity" such as Amazon Web Services. We also collect information on the cost of goods sold, which includes inbound and outbound shipping

[17] The data was compiled by `http://metrocosm.com/get-the-data/`.

costs, and the aggregate outbound shipping cost. This allows us to estimate Amazon's revenue margin net of outbound shipping cost, which we discuss in more detail in Section 5.

In summary, the supplementary data that varies over our sample period includes the sales tax rate and the number of offline retailers, while the demographics, wages, and land rents are observed at a single point in time.

4 Demand Analysis

In the primary demand analysis below, we estimate the effect of taxes and shipping speed on the yearly expenditures of the representative consumer in a county. In order to demonstrate preliminary evidence of the direction and importance of these effects, we begin with the comScore micro-data and estimate the transaction-level effect of taxes and shipping speed on the likelihood that a consumer purchases from Amazon. We run the following linear probability regression:

$$Pr(A_{ohijt} = 1) = \beta_0 + \alpha \ln(1 + \tau_{it}\mathbf{1}_{it}^{taxable}) + \gamma d_{it} + \beta_1 C_h + \beta_2 Z_{it} + \lambda_{ijt} + \epsilon_{ohijt}$$

where each observation is a purchase occasion o, from household h, in county i, in year t, for a product in category j. The dependent variable is an indicator whether the household purchases the product from Amazon.com. We summarize a household's tax status in an indicator variable, $\mathbf{1}_{it}^{taxable}$, that is one if purchases from Amazon by households in county i and year t trigger sales tax collection by Amazon. The tax status variable changes over time due to the entry of a FC and/or due to the future entry of a FC where the state insisted on collecting sales taxes immediately following an agreement with Amazon. As a result, the tax rate Amazon exposes the household to is zero if Amazon does not collect sales tax from households in county i and year t and equal to the local sales tax rate τ_{it} otherwise. The measure of the shipping speed from Amazon to the household is given by d_{it}, which varies depending on the specification. Also included are controls for household income and race, C_h, and the number of small and large offline retailers in the household's county, Z_{it}. Finally, λ_{ijt} is a collection of county, year, and product category fixed effects. We do not include the item's price because we observe the price only at the website where the purchase occurs, but not on alternative sites.[18] However, the product category and year fixed effects capture the average price differences across online retailers by category and across time.

When discussing the results, we focus on the effect of taxes and shipping speed and leave the analysis of the other controls for the main demand specification. The result in the first column of Table 4, which excludes any measure of shipping speed, indicates that consumers are sensitive to taxes and that the probability of an Amazon purchase falls by 15 percentage points for a one percentage point increase in sales tax, a result which is statistically significant at the 5% level.[19]

[18] A specification that includes Amazon's price for Amazon purchases results in similar estimates.

[19] In MN, NJ, PA, RI, and VT, the majority of clothing purchases is sales tax exempt. To verify that such exemptions

In columns 2 through 7, we include various proxies of shipping speeds to account for the fact that consumers may value faster delivery times from Amazon. These are whether or not the county in which the household lives had access to "local express delivery", the log of the distance between the centroid of the household's county and the closest FC, and whether or not the household's county can be serviced with fast delivery times by the Postal Service for each of its four classes of mail. We find that the tax effect is robust to including the variety of proxies for delivery time. Neither the local express delivery indicator, the log of the shipping distance, nor the fast shipping speed indicator across mail classes, have a statistically significant effect on the propensity of purchasing from Amazon.

There are a number of explanations for the results on the shipping speed coefficients. It could be that our measures do not accurately reflect shipping times. Another explanation is that consumers simply do not value faster shipping sufficiently to affect purchase probabilities, or have very heterogeneous preferences for shipping times given Amazon's recent expansion of same-day shipping. Finally, it could be that the expansion of the network of fulfillment centers did not result in faster shipping times to local consumers, or if it did, it was a small number of consumers who were affected. In this scenario, the first order effect of building more FCs is that it results in lower costs for Amazon to ship packages in a given time-frame (i.e., 2-day shipping). We believe the latter to be the most plausible explanation.

We now move from the comScore micro-data to our constructed average spending measure. We provide further evidence of the effect of taxes and shipping speed by estimating a differences-in-differences regression of the annual expenditures on Amazon for the representative household in a county on the 'treatment' variables of whether or not Amazon collects sales tax on its purchases and the measure of shipping speed. In addition, this exercise demonstrates the intuition behind the identification of the effect of sales tax and shipping speed in the main demand specification below. The regression model is given by:

$$ExpAm_{it} = \beta_0 + \sigma \mathbf{1}_{it}^{taxable} + \gamma d_{it} + \beta_1 C_{it} + \beta_2 Z_{it} + \lambda_{it} + \epsilon_{it}$$

where $ExpAm_{it}$ is the log of Amazon expenditures by the representative household in county i in year t. The covariates $\mathbf{1}_{it}^{taxable}$, d_{it}, and Z_{it} are the same as in the transaction level regression. The demographic variables, C_{it}, are the income of the representative household and the weighted share of the population in each race group in county i.[20] Finally, λ_{it} is made up of a county fixed effect and a year fixed effect. Results in Table 5 suggest that a change in the tax status of a county

do not drive the estimated tax responsiveness, we conduct two robustness checks. Since Amazon's "Apparel" may not accurately define products that are sales tax exempt, we first remove households from the six relevant states. Second, we remove any purchases that fall into the Apparel category. We find that the tax coefficient remains significant and of similar magnitude as the presented results.

[20]We assign to the representative household the Census weighted average of the income and race of the comScore households in county i and year t.

results in somewhere between a 9.6 and 10.8% reduction in expenditures on Amazon, with this effect being significant at the 10% level for a majority of the specifications. Dividing this effect by the average tax rate of 6.5% suggests that an increase in the tax of 1 percentage point (an increase in the tax inclusive price of 1%) leads to a 1.6% reduction in spending on Amazon. This elasticity is in line with those of Baugh et al. (2014) and Einav et al. (2014). We once again include measures of shipping speeds in this demand estimation and find little evidence that they significantly affect the purchasing behavior of consumers.

Empirical Model

A weakness shared by both of the above exercises is that we do not consider the substitution between Amazon and other taxed and non-taxed shopping options. Because of this, any reduction in Amazon's transactions due to higher taxes does not equate directly to a tax elasticity since these transactions may be substituted towards other taxed outlets (e.g., walmart.com or Wal-Mart). Therefore, we now specify and estimate a model of demand for retail goods across all modes of online and offline shopping that allows us to determine the effect on overall transactions of higher sales taxes. Another advantage of this model is that it is a micro-founded model that allows us to predict total revenue for Amazon, which we will use in the estimation of the supply side model.

We specify a model where a consumer chooses his retail spending amount on each of four different modes of shopping. The modes are Amazon, $j = 1$, taxed online competitors, $j = 2$, non-taxed online competitors, $j = 3$, and offline competitors, $j = 0$.

We follow Einav et al. (2014) and specify a CES utility function. A representative consumer from county i solves the following problem in year t:

$$\max_{q_{i0t},..q_{i3t}} \quad \left(\sum_{j=0}^{3} \left(\frac{q_{ijt}}{\zeta_{ijt}} \right)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$$

$$\text{s.t.} \quad \sum_{j=0}^{3} p_{ijt} q_{ijt} \leq w_{it}$$

where q_{ijt} represents the quantity purchased via shopping mode j in time period t, p_{ijt} is the price of purchasing one unit of retail products via shopping mode j, ζ_{ijt} is the taste for mode j, and w_{it} is household i's budget for retail goods. The elasticity of substitution between the four modes is given by σ. Solving for the optimal expenditure amount on each online mode, e_{ijt}, results in:

$$e_{ijt} = \frac{(p_{ijt}\zeta_{ijt})^{1-\sigma}}{P_{it}^{1-\sigma}} w_{it},$$

where P_{it} denotes a weighted-average price index across all four modes. Dividing this by expenditure on the offline option, $j = 0$, and taking logs gives:

$$\ln(e_{ijt}) - \ln(e_{i0t}) = (1-\sigma)(\ln(p_{ijt}) - \ln(p_{i0})) + (1-\sigma)(\ln(\zeta_{ijt}) - \ln(\zeta_{i0t})) \tag{1}$$

where the price of each mode can be written as:

$$p_{ijt} = (1 + \tau_{it}\mathbf{1}_{ijt}^{taxable})\tilde{p}_{ijt},$$

denoting as $\tilde{p}_{ijt}$ the tax-exclusive price of buying goods, which may include shipping charges. Note that the tax liability and hence $\mathbf{1}_{ijt}^{taxable}$ may vary across shopping modes j as the non-taxed online competitor never charges sales tax and Amazon does not charge sales tax in a number of states and years. With this, Equation (1) becomes:

$$\begin{aligned}\ln(e_{ijt}) - \ln(e_{i0t}) = &(1-\sigma)(\ln(1 + \tau_{it}\mathbf{1}_{ijt}^{taxable}) - \ln(1+\tau_{it})) \\ &+ (1-\sigma)(\ln(\tilde{p}_{ijt}) - \ln(\tilde{p}_{i0t})) \\ &+ (1-\sigma)(\ln(\zeta_{ijt}) - \ln(\zeta_{i0t}))\end{aligned} \tag{2}$$

We model the taste for the each online mode $j \in \{1, 2, 3\}$ as:

$$\zeta_{ijt} = \exp(\xi_{jt} + \gamma d_{it}\mathbf{1}_j^{Amazon} + \beta_j Z_{it} + \lambda_j C_{it} + \epsilon_{ijt})^{\frac{1}{1-\sigma}}$$

and assume that the preference for the offline shopping mode equals:

$$\zeta_{i0t} = \exp(\Delta\xi_{i0t} + \epsilon_{i0t})^{\frac{1}{1-\sigma}}$$

The term $\Delta\xi_{i0t}$ represents time-varying preferences for online shopping, and we assume that it consists of a county-level, time-invariant, preference for online shopping, ξ_{i0}, and a time trend that captures changes in preferences for online shopping at the level of the county's Census Division, $\Delta\xi_{r0t}$, or $\Delta\xi_{i0t} = \xi_{i0} + \Delta\xi_{r0t}$.[21]

The mode-year effect, ξ_{jt}, accounts for time varying preferences for shopping mode j and can be thought of as the mode's quality that doesn't vary across locations. Aspects such as product variety, return policy, and customer service likely play a major role in determining ξ_{jt}. For Amazon specifically, the availability of its Prime service, which all customers have access to regardless of location, would be included in ξ_{jt}.

We allow mode preferences to reflect observable variation in demographics C_{it}, such as median income and the county's racial makeup, the level of offline competition Z_{it} proxied as the total number and the number of large retail establishments in the county, and, for Amazon, the above

[21]There are nine Census Divisions: Pacific, Mountain, West North Central, West South Central, East North Central, East South Central, Middle Atlantic, South Atlantic, and New England.

measures of shipping speed, d_{it}, with the effects of demographics and competition possibly varying across shopping modes.

The demand shocks vary at the county, time, and mode level, and we assume that they are $i.i.d.$ across these delineations. Under these assumptions, Equation (2) becomes:

$$\begin{aligned} \ln(e_{ijt}) - \ln(e_{i0t}) =& (1-\sigma)\left(\ln(1+\tau_{it}\mathbf{1}_{ijt}^{taxable}) - \ln(1+\tau_{it})\right) \\ &+ \gamma d_{it}\mathbf{1}_j^{Amazon} + (1-\sigma)\left(\ln(\tilde{p}_{ijt}) - \ln(\tilde{p}_{i0t})\right) \\ &+ \beta_j Z_{it} + \lambda_j C_{it} + \xi_i^o + \xi_{rt}^o + \xi_{jt} + \epsilon_{ijt} - \epsilon_{i0t} \end{aligned} \tag{3}$$

We make the assumption that the prices for the online shopping modes do not vary across counties, or that $\tilde{p}_{ijt} = \tilde{p}_{jt}$. While some online firms appear to price discriminate based on a consumer's location, there is limited evidence that this is widespread.[22] Further, Amazon attempted to implement price discrimination in 2000, and quickly abolished it after a backlash from customers.[23] Finally, we assume that the base price of the offline shopping option remains constant over time (i.e., $\tilde{p}_{i0t} = \tilde{p}_{i0}$) and that any variation can be captured through Z_{it}. With this, we can re-write Equation (3):

$$\begin{aligned} \underbrace{\ln(e_{ijt}) - \ln(e_{i0t})}_{\tilde{e}_{ijt}} =& (1-\sigma)\underbrace{\left(\ln(1+\tau_{it}\mathbf{1}_{ijt}^{taxable}) - \ln(1+\tau_{it})\right)}_{\text{Price variation } (\tilde{\tau}_{ijt})} \\ &+ \gamma d_{it}\mathbf{1}_j^{Amazon} + \beta_j Z_{it} + \lambda_j C_{it} + \underbrace{\xi_{jt} + (1-\sigma)\ln(\tilde{p}_{jt})}_{\text{Mode-year FE } (\tilde{\xi}_{jt})} \\ &- \underbrace{\Delta\xi_{r0t}}_{\text{Region-year FE}} - \underbrace{\bar{\xi}_{i0} + (1-\sigma)\ln(\tilde{p}_{i0})}_{\text{County FE}} + \underbrace{\epsilon_{ijt} - \epsilon_{i0t}}_{\text{Residual}} \end{aligned} \tag{4}$$

This final estimating equation attributes the difference in expenditures between mode j and the offline mode to the difference in sales tax between mode j and the offline mode, the convenience effect due to lower shipping times for mode 1 (Amazon), a time varying mode level effect that includes the price of mode j, the effect of demographics and local competition on mode j, a region-specific time trend, a county fixed effect, and an iid demand shock.

Estimation

We estimate Equation (4) using OLS with county, mode/year, and region/time level fixed effects. The procedure to calculate the expenditures for each mode is detailed in Section 3. Identification of the tax sensitivity parameter comes from two primary sources of variation. First, there is variation

[22] See e.g., `http://www.wsj.com/articles/SB10001424127887323777204578189391813881534`, accessed on 1/17/2017.

[23] See `http://www.bizjournals.com/seattle/stories/2000/09/25/daily21.html` and `http://news.cnet.com/2100-1017-240700.html`, accessed on 1/17/2017.

in tax rates in a county across time due to changes in local laws and/or entry of an Amazon fulfillment center. Therefore, changes in expenditures between taxed and non-taxed modes as a result of these changes helps to identify σ. This is similar to the traditional identification argument in difference-in-differences models. Second, there is variation in tax rates across counties within a state, which allows us to exploit covariation between the level of the response to changes in the tax status and the level of the tax rate.

Given this level of variation, we make the assumption that changes in tax rates due to either Amazon's expansion and/or local laws are exogenous to unobserved local demand shocks. This reflects that Amazon's entry decisions are made at a higher geographic level than the county and not made purely in response to a local annual demand shock. To investigate the sensitivity of our results to this assumption and, similarly, to the assumption above that the price of the offline option does not vary over time, we have also estimated Equation (4) with county/year and state/year fixed effects. We find similar results, but choose to proceed with the main specification to avoid identifying the tax effect purely from variation in taxes across modes within a county or state, since we observe only three modes and at most two tax levels across modes.

As in the reduced form regressions above, we explore whether consumers are sensitive to shipping speeds. We identify this effect using changes in the distance and shipping times resulting from the expansion of the FC network. Specifically, the effect of a new FC on shipping times varies both within and across state borders, as it depends on the location of consumers. Note that we do not observe the locations of and thus distance to Amazon's competitors' fulfillment centers, but we assume that the time-varying preferences, ξ_{jt}, capture any overall effect of changes in shipping times for a given mode. Therefore, the effects of shipping speeds are Amazon specific.

Finally, a nice feature of the data is that entry of a fulfillment center into a state does not always lead to Amazon having to charge sales tax (recall the Pennsylvania example from above). Amazon may also have to charge sales tax in a state without having a fulfillment center. Therefore, even within a state, we sometimes see variation in shipping distances over time without similar variation in the county's tax status and vice versa.

Matching Total Sales

One of the goals of the demand model is to be able to predict Amazon's total sales under different configurations of the network. However, because of three data limitations, the model presented above may not accurately predict total revenue of either Amazon or the online retail segment in total. First, we are missing sales made through other channels such as mobile devices or second computers at home or work. This issue exists across all three online modes of shopping, suggesting that we underestimate revenue for all modes. Second, the comScore sample consist of online users only. As a result, we potentially overestimate online spending by the average household. This counters, but unlikely offsets, the underreporting introduced by missing sales due to the increased prevalence of online shopping away from the home computer.[24] Third, we cannot separate Amazon purchases from Amazon Marketplace purchases, implying that our model predicts the sum of Amazon's own sales and its sales from the sellers on Marketplace. This issue exists only for Amazon and results in an overestimate of its sales.

To address these data limitations, we supplement the county-level expenditure data with information on Amazon's yearly revenue obtained from their financial reports and information on annual total online revenue from the US Department of Commerce.[25] This information allows us to scale our demand model predictions to match, when summing across modes, overall online retail expenditures, as well Amazon expenditures in particular. We do so by including two additional sets of annual 'multiplier' parameters in estimation that we identify by matching the model's aggregate predictions to these external estimates.

The Amazon specific multipliers, α_{t1}, are annual parameters that represent the combination of the missing sales and the extent of Marketplace purchases in the comScore data. For example, if $\alpha_{t1} = 1.2$, it implies that we need to scale up our predictions by 20% in order to match the revenue reported in the financial statements. The value of the Amazon multipliers may be greater or less than one depending on the relative impact of missing sales versus Marketplace purchases. The multipliers that apply to modes 2 and 3 and have a similar interpretation as the Amazon multipliers. However, because these multipliers represent the missing sales in the comScore data, they will take values greater than one; we restrict them to be the same for the two modes since we do not observe aggregate online revenues separately for the two channels, $\alpha_{t2} = \alpha_{t3}$.

We estimate these parameters by iterating over the OLS regression of Equation (4) for different values of the multipliers until the predicted yearly revenue for Amazon matches the revenue observed in their financial reports and the yearly total online revenue matches the revenue reported by the US Department of Commerce. We begin by estimating the model in Equation (4) without any multipliers. We then predict spending for each of the three modes for the representative household in each county and, scaled by the number of households, for the county in total. For counties not covered by the comScore data, we use their attributes to predict spending, assuming that the county fixed effect takes on the estimated fixed effect of the geographically closest county. These allow us to create yearly total predicted revenue for each mode and to generate starting values for

[24] eMarketer estimates that the share of online shoppers who made at least one purchase on a mobile device increased from 10.8% in 2010 to 53% by 2013. A FindLaw.com survey finds that 35% of the population shopped online while at work in 2014. Lastly, Nakono estimates that the average number of desktops and laptops per household increased from 2.12 in 2006 to 3 in 2012. See `http://company.findlaw.com/press-center/2014/thirty-five-percent-of-americans-shop-online-while-at-work.html` and `https://www.nakono.com/tekcarta/databank/personal-computers-per-household/`, accessed on 1/22/2017.

[25] Overall online sales are available at `https://www.census.gov/retail/ecommerce/historic_releases.html`. We adjust overall sales to represent the total online sales for the relevant product categories only by multiplying online sales by the annual spending share of these categories in the comScore data. The assumption is that the share of the relevant product categories for the comScore sample is equal to the share across the population.

the multipliers:

$$\alpha_{t1}(0) = \frac{R_{1t}(0)}{\bar{R}_{1t}} \tag{5}$$
$$\alpha_{t2}(0) = \alpha_{t3}(0) = \frac{R_{2t}(0) + R_{3t}(0)}{\bar{R}_{2,3t}}$$

where $R_{jt}(0)$ represents the predicted revenue for mode j, $\bar{R}_{1t}$ represents Amazon's officially reported revenue, and $\bar{R}_{2,3t}$ represents the expenditures on modes 2 and 3 calculated by subtracting Amazon's revenues from the total online revenue reported by the Department of Commerce. We then multiply each representative household's observed expenditure on mode j in year t by the value of the corresponding multiplier:

$$\bar{e}_{ijt} = \alpha_{tj}(0) \times e_{ijt} \tag{6}$$

and re-estimate the OLS regression model of Equation (4) with the adjusted dependent variable, $\tilde{e}_{ijt} = \ln(\bar{e}_{ijt}) - \ln(e_{i0t})$. Using the estimates, we again predict the total expenditure for each mode and calculate updated multipliers according to Equation (5). We iterate over this procedure until predicted and actual revenue match.[26]

The implicit assumption in this procedure is that the shares of missing transactions and Amazon Marketplace expenditures do not vary across counties. Primarily, adjusting expenditures in this way affects the estimates of the mode-year fixed effects; the other parameters remain largely unchanged when estimating the model without the multipliers. Importantly, there is very little change in the estimated tax sensitivity and the convenience effect. We confirm this in Appendix C, which displays the estimates of the demand model without any multipliers.

Results

Results under six different specifications are presented in Table 6, where each specification uses a different measure of shipping speed. Specification (1) starts with only sales taxes and demographic controls; specification (2) adds the indicator of local express delivery availability; specification (3) instead includes the log of the distance to the closest FC, and specifications (3)-(6) include indicators for fast shipping speeds in the four different classes of delivery. Since sales tax enters with a coefficient of $(1-\sigma)$ in Equation (4), we display in Table 6 the transformed coefficient that corresponds only to the elasticity of substitution $(-\sigma)$ only. The estimated elasticity of substitution is approximately -1.43 and is significant at the 5% level in all specifications. In CES specifications,

[26] In practice, we stop iterating when the multipliers converge:

$$\|\alpha(k) - \alpha(k+1)\| < \epsilon = 1.00E-06$$

where $\alpha(k)$ is the vector of multipliers in iteration k and $\|\cdot\|$ is the Euclidian norm.

the elasticity of substitution is approximately equal to the price elasticity for mode j when the spending share for mode j is very small. The expenditure shares are under 0.02 for each of the online modes, meaning that the elasticity is approximately -1.4. Thus, a one percentage point increase in the tax rate (i.e., a 1% increase in the price) leads to a decrease in demand of 1.4%. These effects are similar to the ones reported by both Einav et al. (2014) and Baugh et al. (2014).

An alternative interpretation of the estimates is that going from not charging sales tax to charging the average sales tax of 6.5% would reduce expenditures on each online mode by 9.1 percent (6.5×1.4). That is, if Amazon agrees to charge sales tax in a state through either an agreement with the state government or because they build a fulfillment center, they can expect their revenue in the state to decrease by up to 9.1%, all else equal. In 2008, New York state passed a law that required Amazon to collect sales tax. Given the state average sales tax of 8.5% that year, our estimates imply that the law reduced Amazon's revenue in New York by approximately 11.9%.

One possible issue with these estimates is the fact that a consumer in a state where Amazon itself collects taxes may be able to buy from a Marketplace seller who does not have to charge sales tax. In that case, our estimate would be a lower bound of the responsiveness to taxes: we would attribute a lack of a combined Amazon spending response to a low tax sensitivity. Most state laws dictate that Marketplace vendors collect sales tax from customers in the states where their goods are stored, and most vendors store their goods in Amazon's fulfillment centers.[27] As a result, their sales tax obligations are identical to Amazon's itself.

In line with the above estimates, Table 6 does not provide strong evidence that our measures of shipping times drive substitution between Amazon and other shopping outlets. This is further evidence that the expansion of the network did not significantly change the shipping times from Amazon, or in other words, the convenience effect of expanding the network is approximately zero over the period of our sample. Going forward, we therefore rely on Specification (1) as our primarily empirical model.

We present the estimates of β_j and λ_j, the mode-specific spending effects of demographics and offline competition, for Specification (1) in Table 7. A larger number of offline choices drives down expenditures for all modes of online shopping, with the effect of large offline competitors (e.g., Wal-Mart or Target) being stronger. However, the effect is only significant at the 5% level for taxed online competitors, implying that online firms with an offline presence face more competition from large offline retailers than online only firms. We also find that higher income households spend more via online channels. Relative expenditures do not correlate in a statistically significant way with the remaining demographic variables. Note that because we do not have time-varying data

[27]See `http://www.webretailer.com/lean-commerce/amazon-sellers-survey-2016/`, accessed on 1/17/2017, indicating that 79% of all sellers and 88% of sellers with revenue upward of $1 million use Amazon fullfilment centers, and `http://www.avalara.com/learn/whitepapers/fba-sellers-guide-sales-tax/`, accessed on 1/17/2017, for an overview of sales tax collection obligations of such vendors.

on the racial makeup of counties, we estimate their effect relative to the offline mode, mode 3.

The estimates of the time varying mode fixed effects are presented in Table 8, with the excluded effect being mode 3 in 2006. We use them to construct a measure of average mode preferences:

$$\bar{\xi}_{jt} = \hat{\tilde{\xi}}_{jt} + \hat{\beta}_j \bar{Z}_t + \hat{\lambda}_j \bar{C}_t \tag{7}$$

where $\hat{\tilde{\xi}}_{jt}$ is the estimated mode fixed effect and $\bar{Z}$ and $\bar{C}$ are the averages of the demographic and competition variables. We display $\bar{\xi}$ in Figure 4. Not surprisingly, given Amazon's increase in market share, mean preferences for Amazon are increasing over time. Mean preferences for both the taxed online competitors and the non-taxed competitors are also increasing over time, but at a slower rate than Amazon's. Due to the way we specified demand via $\tilde{\xi}$, this could be due to a relative decrease in Amazon's prices and/or a relative increase in its platform quality.

Finally, we turn to the estimates of the multipliers. The Amazon-specific multiplier are increasing over time, from 1.03 in 2006 to 1.69 in 2013.[28] This implies that the comScore data understates Amazon's 2013 revenue by about 41% ($\frac{0.69}{1.69}$). The multipliers for modes 2 and 3 are also generally increasing over time, rising from 1.02 in 2006 to 2.93 in 2013, a number that implies that we are missing about 66% ($\frac{1.93}{2.93}$) of 2013 sales for these modes. The increasing rate is likely due to the growing prevalence of mobile shopping and at-work-online shopping, as well as multiple desktop and laptop computers in the home, not all of whom are tracked by comScore (see footnote 24).

Recall that the Amazon multipliers are a combination of missing sales and Marketplace revenue, so they are equivalent to the product of a Marketplace multiplier (less than 1) and a missing sales multiplier (greater than 1). If the missing sales multiplier for Amazon were the same as the other two modes', then the Marketplace multiplier equals α_{t1}/α_{t2}, or 0.57 for 2013. This means that 43% of the revenue that we observe is from Marketplace sellers, a number that is in line with the estimates of the share of third party sales on Amazon.[29]

Predicting Total Revenue

In order to estimate the supply side model that follows, we form predictions of Amazon's total revenue from 2006-2018 under various assumptions about the FC network rollout and the timing of tax changes due to the rollout. In this section, we explain this procedure for the observed rollout of FCs and the observed timing in changes in the tax status, and examine patterns in the estimated spending across the three modes of online shopping.

We use the demand estimates of specification (1) in Table 6 to predict annual expenditures of a representative household in county i for each mode over the period 2006-2013. We then scale

[28] The multipliers are 1.03, 1.39, 1.50, 1.55, 1.52, 1.50, 1.01, and 1.69 for the years 2006-2013.

[29] In the second quarter of 2013, 40% of overall units shipped through the platform were from third-party sellers according to `https://www.statista.com/statistics/259782/third-party-seller-share-of-amazon-platform/`, accessed on 1/17/2017.

up the resulting figures by the number of households in each county to derive our estimate of total revenue for each mode. Results are displayed in the upper panel of Table 9. Recall that the estimation procedure implies that our predicted revenue for Amazon is exactly equal to the revenue reported in their financial statements and the total revenue across the three modes is to exactly equal to the revenue reported by the Department of Commerce. Not surprisingly, the table indicates general increase in spending across all modes, with Amazon's sales increasing faster than the other two modes'. Using these predictions, we also construct the spending share for each online mode and display them in Figure 5. This differs from Figure 3 for two reasons. First, the shares are calculated across all modes, including the offline mode, resulting in shares below 0.02 for the three online modes. Second, the shares are calculated using the predictions of the model, rather than the raw data. Figure 5 indicates that, while the share of expenditures bought online is increasing over the sample period, Amazon's share is growing at a faster rate compared to modes 2 and 3.

We now turn to the projections of revenue for 2014-2018, which requires out-of-sample values of the mode-specific and region-specific time trends. We predict these by projecting out from the estimated effects given their growth over 2006 to 2013.[30] The lower panel of Table 9 reports the total predicted revenue from 2014 to 2018. Because we observe Amazon 2014 and 2015 sales from their financial reports, we can assess how well this procedure predicts sales in these years. We overpredict revenue by only around 5% in both years (a predicted \$53.6 and \$65.9 billion versus an actual \$50.8 and \$62.9 billion in 2014 and 2015, respectively). We also predict that in aggregate, mode 2 expenditures rise from \$62 billion in 2014 to \$80 billion in 2018 and mode 3 expenditures rise from \$49 billion to \$61 billion.

5 Quantifying Cost Savings

We now turn to estimating the cost savings of expanding the network of FCs. We posit that these cost savings arise from the FC network growth leading to a shortening of the outbound shipping distance to customers from their assigned FC. A shorter shipping distance can benefit Amazon in two ways. The company relies primarily on shipping companies to carry out its deliveries and negotiates with shippers over fixed per-package rates. With more localized FCs, the outbound shipping distance from the FC to the shipping service's sorting center falls. Second, the expansion may lead to elimination of the leg from the FC to the sorting facility for some locations, allowing the shipping service to deliver the package directly from the FC to the customer. Under either scenario, the shipper's costs are lower, giving Amazon more bargaining power when negotiating their per package shipping rates.[31]

[30] For each mode-specific effect $\tilde{\xi}_{jt}$, $j = 1, 2,$ or 3, and each region trend $\Delta\xi_{r0t}$, we perform the following regression: $\hat{m}_t = \gamma_0 + \gamma_1 \ln(t) + \gamma_2 \ln(t)^2 + \epsilon_t$, where $\hat{m}_t$ represents the estimated 2006-2013 time trend under consideration. Using the estimates of γ, we predict future values of m_t.

[31] See `http://nyti.ms/1fUabBh`, accessed on 1/17/2017, for a discussion of Amazon's bargaining leverage in negotiating with delivery companies.

Profit Function

To quantify potential economies of density in distribution, we formulate a profit function for Amazon that depends on the rollout of FCs in a similar manner as the rollout of Wal-Mart stores studied in Holmes (2011). Relative to Wal-Mart's sequence of store openings, Amazon's rollout of FCs features far fewer openings. We therefore assume a relatively parsimonious structure for Amazon's decision process. The profit function represents the expected stream of discounted annual profits starting from the begging of our sample (2006) and it is made up of three primary components. The components are the annual variable profit, which depends on the locations of the FCs, the annual fixed costs necessary to operate the network of FCs, and the one time sunk costs necessary to open the collection of FCs. We assume that, in 2006, Amazon has perfect foresight in predicting these components and chooses the rollout that maximizes their expected discounted stream of profits.

We now describe these three components in more detail. We summarize the set of FCs that Amazon operates across counties in the vector a_t, with $a_{it} = 1$ if there is a FC in county i in year t. The collection of a_t over time, which we denote as a below, represents the 'rollout' of the network of FCs. We assume that variable profit in year t from county i is given by:

$$\pi_{it}(a_t;\theta) = \mu R_{it}(a_t) - \theta d_{it}(a_t) R_{it}(a_t) \tag{8}$$

where $R_{it}(a_t)$ is the revenue generated in county i, which is a function of a_t due to sales tax implications. As we don't find strong support for convenience effects in Section 4, we assume that the revenue is independent of any of our measures of shipping speed. In practice, we use the estimates of Specification 1 of Table 6 as the primitives defining the demand model. The term μ is the share of revenue that Amazon receives as profit excluding net shipping costs. Largely for convenience, we assume μ is constant across all periods and set it to 0.23 based on information obtained from Amazon's financial reports, as we discuss further in Appendix B.

We denote variable net shipping costs by $\theta d_{it}(a_t) R_{it}(a_t)$, where the parameter θ measures the net shipping cost per mile per dollar of goods sold and $d_{it}(a_t)$ is the distance from consumers in county i to the FC Amazon uses for them. The reason these costs are interpreted as the *net* shipping cost is because we allow μ to account for any shipping costs which are recouped through shipping revenue. Additionally, in our primary specification, we assume that d_{it} is the minimum of the great-circle distance between the population weighted centroid of county i the network FCs open in year t.

The annual fixed costs of operating a FC that is L_i square feet and employs A_i workers in county i are given by F_i:

$$F_i = w_i A_i + r_i L_i \tag{9}$$

where w_i is the county-level wage obtained from the Bureau of Labor Statistics and r_i is rental rate per square foot calculated from the SNL data. According to our measures, these fixed costs increase from \$1.2 billion in 2006 to \$3.7 billion in 2013 due leasing more FC space and hiring additional employees to run the additional FCs.

We assume that there are sunk costs associated with opening a FC in year t, denoted S_t, that do not vary by county. Such sunk costs include machinery and other materials used in the construction of an FC. Given these three components, the discounted sum of Amazon's total profit across counties starting in 2006 is:

$$\Pi(a;\theta) = \sum_{t=2006}^{\infty} \beta^{t-2006} \sum_i \pi_{it}(a_t;\theta) - F_{it}a_{it} - S_t(a_{it} - a_{it-1}) \tag{10}$$

where β is the discount factor, which we set to 0.95. The profit maximization problem in 2006, in terms of the network rollout, can then be written as:

$$\max_a \Pi(a;\theta) \tag{11}$$

By revealed preference, we know that profit under the observed network rollout (a^o) must be greater than any alternative rollout (a), or:

$$\Pi(a^o;\theta) \geq \Pi(a;\theta) \forall a \neq a^0 \tag{12}$$

This allows us to formulate a set of moment inequalities that we use to estimate the value of θ.

Moment Inequalities

As in Holmes (2011), we construct moment inequalities by comparing the discounted profits of the observed sequence of openings with counter-factual roll-out scenarios. We focus on counter-factual roll-outs that *swap* the opening dates of two FC clusters. Each swap consists of moving the opening date of the earliest FC in the cluster. The total number of unique perturbations, after removing swaps of FCs that opened in the same year, is 1,575, where we include planned FC openings through 2018, as listed in MWPVL, to maximize the number of possible perturbations.

Relative to perturbations that consider, e.g., alternative locations the company could have chosen but did not, simply swapping the opening dates of two FCs, t and t', has the advantage that the comparison of the firm's profit streams under the perturbed and the actual FC roll outs does not require knowledge of post-sample continuation values. The latter cancel out, allowing us to impose optimality conditions without observing – or predicting as the equilibrium outcome of maximizing the profit from network formation in Equation (11) – the complete sequence of future FC openings. Since we assume that sunk costs vary only over time, but not by location, the firm incurs the same total sunk costs across FCs in any given year t under the actual network and the perturbed networks we consider. Sunk costs thus do not affect the relative profit comparison of alternative network roll outs; our estimation procedure does not identify their magnitudes.

Consider two FC opening sequences: observed sequence a^0, and a counter-factual opening sequence. While the following exposition applies to any alternative network roll-out, we focus on those alternative sequences we consider in estimation. Thus, denote as $a^{j,k}$ the roll-out of FCs that results from switching FC j's opening date with that of FC k, holding fixed the opening dates of all remaining FCs as observed in the data. We order FCs by their opening date, so that $t_j < t_{j+1}$, $j = 1, ..., J$, denoting as t_j the opening date of FC j and as J the total number of FC cluster openings during our sample period; $J = 60$. If profits were measured without error, revealed preference combined with profit maximization implies the following linear profit difference inequality:

$$y^{j,k} - x^{j,k}\theta \geq 0, \text{where} \tag{13}$$

$$y^{j,k} = \sum_{t=2006}^{2018} \beta^{t-2006} \left[\sum_i \mu \left(R_{it}(a_t^0) - R_{it}(a_t^{j,k}) \right) - \left(F_{it} a_{it}^0 - F_{it} a_{it}^{j,k} \right) \right]$$

$$x^{j,k} = \sum_{t=2006}^{2018} \beta^{t-2006} \sum_i \left[d_{it}(a_t^0) R_{it}(a_t^0) - d_{it}(a_t^{j,k}) R_{it}(a_t^{j,k}) \right]$$

where $y^{j,k}$ is the discounted profit differential net of shipping costs and $x^{j,k}$ is the discounted difference in revenue-weighted shipping distance from comparing the observed opening sequence to one that has swapped the opening dates of FCs j and k.

In order to form the values of $y^{j,k}$ and $x^{j,k}$, we calculate Amazon's total revenue under both a^0 and $a^{j,k}$. Any revenue differences come from changes in the timing of sales tax collection when the opening date of FC j is swapped with the opening date of FC k. That is, when computing $R_{it}(a_t^0)$ and $R_{it}(a_t^{j,k})$, we assume that all other aspects affecting revenue, including the prices and behavior of competitors, are fixed at the levels estimated in the demand model.

Due to the importance of the timing of sales tax collection, we have to make an assumption regarding Amazon's expectation at the time of committing to a FC about when sales tax collection liability commences given FC entry into a state. A concrete example illustrates the challenge: consider a counterfactual where we swap the opening dates of a Pennsylvania FC that opened in 2008 and resulted in PA sales tax collection starting in 2011 with an Indiana FC that opened in 2007 but resulted in sales tax collection only in 2014. Several alternatives present themselves. In our main specification, we assume agnostically that Amazon expects sales tax collection to begin in a state at the date of the earliest opening of an FC in that state, and thus, that the observed network is optimal under this expectation. In the specific example, this would imply that Amazon's profit under the observed sequence, where Amazon begins collecting sales taxes in Pennsylvania at its observed FC opening date in 2008 and in Indiana at its FC opening date of 2007 is greater than the alternative profit when the FC opening date and onset of tax collection move to 2007 in Pennsylvania and to 2008 in Indiana. In the Appendix, we present estimation results for a perfect foresight assumption, where Amazon has correct expectations about the outcome of negotiations with each state government regarding the timing of sales tax collections and assume that this state-specific time lag remains the same when we change the opening date of an FC in a perturbation. While Amazon's true expectations likely lie between these two extremes, the results are similar under this assumption.

The remaining inputs in $y^{j,k}$ and $x^{j,k}$ are the differences in revenue weighted shipping distance and fixed costs. The shipping distance to consumers in county i and in year t changes under the perturbed network if the swap of opening dates results in a different FC being the closest to county i in year t. For example, the closest FC to a consumer living in central PA in 2007 under the observed network was one located in Delaware. The swap with the IN FC results in an FC being open in PA in 2007, implying a shorter shipping distance to consumers in central PA starting in 2007. Finally, in calculating discounted profit under a perturbed network, we assume that the FC's square footage and labor force are unaffected by an opening date swap. As a result, the counterfactual Pennsylvania FC that opens in 2007 instead of 2008 continues to have the fixed cost that we assign to it based on its own MWPVL information.

The parameter of interest, θ, is the net shipping cost per mile per dollar of revenue.[32] Its magnitude is bounded by two types of profitability comparisons. First, observing Amazon entering relatively early into a high-tax state when a low-tax (but further away) option was available, but chosen later, identifies a lower bound of the shipping cost parameter. For these comparisons, we predict a negative pre-shipping profit differential of the observed network roll out relative to the relevant counterfactual roll out sequences, $a^{j,k}$ (i.e. $y^{j,k} < 0$), either due to relative revenue losses or fixed cost increases. For the observed roll out to be preferred nevertheless, such a loss has to be more than offset by a reduction in revenue-weighted shipping distance (i.e. $x^{j,k} < 0$). This gives a lower bound on the shipping cost so that the decrease in weighted shipping distance, when scaled by θ, exactly makes up for the loss in pre-shipping profit.

Second, for other perturbations, we predict higher pre-shipping profit ($y^{j,k} > 0$), but also an increased shipping distance ($x^{j,k} > 0$), for the observed configuration relative to the alternative. These correspond to instances where Amazon chose to enter a FC relatively early into a low-tax state, when a high-tax option that was closer to its customer base was available but chosen later. For the observed roll out sequence to be optimal, the profit impact of the increased weighted

[32] Note that this interpretation of θ implicitly assumes that the expansion of the FC network does not affect inbound shipping routes and costs from the manufacturers to the FC. If a share of inbound shipping costs were passed on to Amazon, its wholesale price would depend on its FC network, which would confound the effect of its own shipping costs on profitability through θ. The fact that Amazon's suppliers already deliver goods to many other retailers across the US and/or have their own widespread distribution network alleviates this concern. For example, Barnes & Noble requires book publishers to deliver inventory directly to stores, while Walmart requires shipment to distribution centers, which are increasingly close to Amazon's FCs with Amazon's expansion. (The average distance from a FC to the closest Walmart fulfillment center falls from 92.2 miles in 2006 to 65.4 miles in 2013.)

shipping distance cannot exceed the pre-shipping profit differential, yielding an upper bound on the shipping cost.

While this discussion compares exact pre-shipping profitability and shipping cost for alternative paths, in practice we measure both quantities with error, for a number of reasons. These include: (i) measurement error in or omitted contributions to the fixed operating costs of FCs, (ii) estimation error in the firm's revenue that would affect both $y^{j,k}$ and $x^{j,k}$, or (iii) a mis-representation of the role of the geopolitical environment on Amazon's profit and decision making. The political environment affects firm profitability in several ways, including through potential public incentives that reduce the ongoing cost of a FC being located in a particular jurisdiction and through the timing of any change in sales tax collection triggered by the FC entry. Predicted revenue would not correspond to the firm's expected revenue at the time of the FC roll-out if Amazon's beliefs of the onset of its tax collection obligation and outcomes of negotiations with local governments differed from our assumption about their beliefs, or if changing the opening date of an FC from its actual opening date, as we do in our perturbations, would have affected the company's sales tax treatment differently from what we assume in calculating the profitability of the perturbed network.

We thus observe $\tilde{y}^{j,k}$ and $\tilde{x}^{j,k}$. We assume that all sources of error in $y^{j,k}$ and $x^{j,k}$ enter additively, resulting in:

$$\begin{aligned} \tilde{y}^{j,k} &= y^{j,k} + \eta^{j,k}, \\ \tilde{x}^{j,k} &= x^{j,k} + \nu^{j,k} \end{aligned} \tag{14}$$

and that $\eta^{j,k}$ and $\nu^{j,k}$ are mean-zero disturbances that are independent of an L vector of non-negative instruments $z^{j,k}$:

$$E(z^{j,k}\eta^{j,k}) = E(z^{j,k}\nu^{j,k}) = 0. \tag{15}$$

Interacting the empirical equivalent of the moment inequality condition in Equation (13) with the instruments $z^{j,k}$ implies that:

$$E\left[z^{j,k}(\tilde{y}^{j,k} - \theta\tilde{x}^{j,k})\right] = E[z^{j,k}(y^{j,k} - \theta x^{j,k})] + \underbrace{E[z^{j,k}(\eta^{j,k} - \theta\nu^{j,k})]}_{=0} \geq 0, \tag{16}$$

This leads to the following set of l sample moment inequalities as in Pakes et al. (2015):

$$\tilde{m}_l(\theta) = \frac{1}{M}\sum_{j=1}^{J}\sum_{k=j+1}^{J} z_l^{j,k}(\tilde{y}^{j,k} - \theta\tilde{x}^{j,k}) \geq 0, \quad \forall l = 1, \ldots, L. \tag{17}$$

where M equals to the number of perturbations: $M = \sum_{j=1,\ldots,J}\sum_{k=j+1,\ldots,J} 1$, or $M = 1,575$.

Instruments

We use similar sets of instruments to the ones employed in Holmes (2011). The first set consists of indicator variables that identify perturbations that capture the interaction between the effect of sales taxes and shipping distance on the discounted profit differential net of shipping costs ($\tilde{y}^{j,k}$) and the discounted difference in weighted shipping distance ($\tilde{x}^{j,k}$). Recalling the discussion from above on roll-out comparisons that inform the lower and upper bounds on the shipping cost per mile per dollar, we select two types of "experiments." To identify the lower bound, we look for counterfactual roll-outs that both *increase the average shipping distance* and result in higher revenues (Group 1). For the upper bound, we consider alternative roll-outs that *decrease the shipping distance* of the network, while also reducing discounted revenue streams due to sales taxes being collected on a sufficiently large number of customers early on. (Group 2).

1. *Group 1 Perturbations: Increase in shipping distance.* The indicator variable $z_1^{j,k}$ equals to one for all perturbations (j,k) such that (a) $\tilde{y}^{j,k} < 0$, and (b) $\tilde{x}^{j,k} < 0$, zero otherwise, $j = 1, ..., J,\ k = j+1, ..., J$.

2. *Group 2 Perturbations: Decrease in shipping distance.* The indicator variable $z_2^{j,k}$ equals to one for all perturbations (j,k) such that (a) $\tilde{y}^{j,k} > 0$, and (b) $\tilde{x}^{j,k} > 0$, zero otherwise, $j = 1, ..., J,\ k = j+1, ..., J$.

These indicator variables thus identify counterfactual sequence that qualify for experiments 1 and 2. We also experimented with delineating groups 1 and 2 into subgroups defined by the magnitude of $\tilde{x}^{j,k}$, similar to Holmes (2011), but did not find the results to be significantly different when taking this approach.

The second set of instruments consists of interactions between $z_1^{j,k}$ and $z_2^{j,k}$ and positive transformations of the change in revenue-weighted shipping distance:[33]

$$\tilde{x}_+^{j,k} = \tilde{x}^{j,k} - \min_{\substack{j'=1,...,J, \\ k'=j'+1,...,J}} \tilde{x}^{j',k'} \tag{18}$$

A complication in our setting relative to Holmes (2011) is that we assume that shipping costs are variable, rather than fixed costs, so that we specify the profit function to depend on revenue-weighted shipping distances, instead of a revenue-independent measure of shipping speed. Thus, measurement error in revenue propagates to both $\tilde{y}^{j,k}$ and $\tilde{x}^{j,k}$; as a result, $z_1^{j,k}$, $z_2^{j,k}$, and $\tilde{x}_+^{j,k}$ are not valid instruments.[34]

[33] In alternative specifications in Appendix C, we include the transformed shipping distance separately, without interacting them with $z_1^{j,k}$ and $z_2^{j,k}$. This allows us to use all the perturbations in estimation rather than limiting them to the Group 1 and 2 perturbations. Results are robust to this adjustment

[34] In contrast, Holmes (2011) derives instruments based on exogenous profit differences including store density and shipping distances, which are assumed to be fixed cost components of store profit.

We therefore construct an alternative set of instruments by projecting $\tilde{y}$ and $\tilde{x}$ on exogenous determinants of the pre-shipping profit and weighted shipping distance differential for perturbation (j,k). We run the regressions

$$\begin{aligned} \tilde{y}^{j,k} &= \beta_y Z^{j,k} + \epsilon_y^{j,k} \\ \tilde{x}^{j,k} &= \beta_x Z^{j,k} + \epsilon_x^{j,k} \end{aligned} \tag{19}$$

where $Z^{j,k}$ are exogenous shifters of $\tilde{y}^{j,k}$ and $\tilde{x}^{j,k}$. We employ the discounted sum of differences in household weighted shipping distance, household weighted sales tax rates, and fixed costs, along with indicators of instances where either FC j of FC k, or both, are not the first FC openings in their state. We then use the estimates of β to form the predicted values of $\tilde{y}$ and $\tilde{x}$, which we call $\hat{y}$ and $\hat{x}$. The latter are then used to re-calculate exogenous versions of the instruments, $\hat{z}_1^{j,k}$, $\hat{z}_2^{j,k}$, and $\hat{x}_+^{j,k} \times \hat{z}_i^{j,k}, i = 1, 2$. For reference, we present results from estimating the shipping cost with both the original version of the instruments, together with this second version, below.

Estimation

Estimation is based on the moment inequalities of Equation 17. Our goal is to find the maximum and minimum values of θ such that all moments are satisfied and to provide inference on the values of these bounds. Due to the moments being informative about bounds on θ, rather than a point estimate, inference under moment inequalities requires methods that differ from the methods traditionally used in a GMM setting. Therefore, we follow Andrews and Soares (2010) (AS hereafter) in constructing confidence sets in which the true value of θ falls 95% of the time and report these sets as our estimates. Intuitively, the AS method involves estimating a test statistic for a given value of θ and rejecting the value of θ if the test statistic is less than the 95th percentile of the test statistic distribution computed from a number of "bootstrapped" samples of perturbations. The confidence set is the set of θs that are not rejected.

For a candidate value of the shipping cost parameter θ^0, we calculate the L sample moments $\tilde{m}(\theta^0)$ and the $L \times L$ sample variance covariance matrix of the empirical moments $\hat{\Sigma}(\theta^0)$. We estimate $\hat{\Sigma}(\theta^0)$ using the method in Holmes (2011) to account for both error in the demand model and correlation across perturbations. See appendix B for details. We then calculate the test statistic:

$$T(\theta^0) = \min_{t \geq 0} \left(M^{\frac{1}{2}} \tilde{m}(\theta^0) - t \right) \hat{\Sigma}(\theta^0)^{-1} \left(M^{\frac{1}{2}} \tilde{m}(\theta^0) - t \right) \tag{20}$$

where t is a k by 1 vector. Here, the term $(M^{\frac{1}{2}} \tilde{m}(\theta^0) - t)$ is equal to 0 when the moment is satisfied and $M^{\frac{1}{2}} \tilde{m}(\theta^0)$ when it is not. Therefore, the test statistic equivalent to a standard GMM objective function after replacing any positive moment with 0. In order to calculate the critical value for the test statistic, we draw R bootstrap samples of perturbations. We sample J FCs from the population with replacement, where J is the observed number of FC clusters, to form a set which we denote J_r for a given bootstrap sample r. We then construct a set of perturbations a_r from swapping the opening dates of the FCs contained in J_r. Note that this could result in swaps appearing multiple times and/or swaps eliminated from the bootstrap sample. For each bootstrap sample r, we form:

$$T_r(\theta^0) = \min_{t \geq 0} \left(\tilde{m}_r^{**}(\theta^0) - t\right) \hat{\Sigma}_r^{**}(\theta^0)^{-1} \left(\tilde{m}_r^{**}(\theta^0) - t\right) \tag{21}$$

where $\tilde{m}_r^{**}$ and $\hat{\Sigma}_r^{**}$ are calculated using the bootstrap sample of perturbations and the procedure in appendix B. The sample of R bootstrapped values of T_r allow us to find a critical value for the original estimation sample's value of T as the $(1-\alpha)^{\text{th}}$ sample quantile. Denote this critical value as $\hat{c}(\theta^0, 1-\alpha)$. If $T(\theta^0) > \hat{c}(\theta^0, 1-\alpha)$, we reject the hypothesis that $\theta = \theta^0$ and eliminate θ^0 as a candidate value.

AS suggest using a grid search in order to determine the confidence set. However, because we only have one parameter to estimate, we begin with one value of θ^0 in the confidence set (we employ θ^0 that minimizes Equation (20)), together with guesses at low and high values of θ^0 outside the confidence set, and use the bisection-method to find the maximum and the minimum value of θ^0 such that $T(\theta^0) \leq \hat{c}(\theta^0, 1-\alpha)$. We set $R = 2000$ and $\alpha = 0.95$.

Shipping Distance, Sales Tax, and Fixed Cost Tradeoff

To provide some intuition for the variation in the data that identifies the per dollar per mile net shipping cost, we provide a graphical description of the perturbations that illustrate the trade-off Amazon faces between charging sales tax, shorter shipping distances and the fixed costs of operating FCs.

Figure 6 displays four scatter plots, where each point represents data from one of the 1,575 perturbations. The y-axis in each plot represents the difference in the household level average shipping distance during 2006-2018 between the observed network and the perturbed network. For example, a value of -5 implies that the average shipping distance during 2006-2018 under the perturbed network is 5 miles less than the average shipping distance under the observed network. This varies from -10 to 20 miles; most perturbations are between -5 and 5 miles. Each x-axis represents the difference in one or more components of Amazon's profit function between the observed network and the perturb network: Figure 6a uses the difference in the discounted sum of pre-shipping profit ($\tilde{y}$) during 2006-2018 in millions of dollars, Figure 6b the difference in the discounted sum of pre-shipping variable profit ($\mu \times R$) during 2006-2018 in millions of dollars, Figure 6c the difference in the total the number of households taxed during 2006-2018 in millions of households, and Figure 6d the difference in the discounted sum of fixed costs (F) during 2006-2018 in millions of dollars. In each plot, we display two lines of best fit, one for all perturbations (red) and one for the subset of perturbations that qualify for Experiments 1 and 2 (yellow).

Recall that the identification of upper bound of the shipping cost parameter comes from situations where Amazon opened an FC in a low-population/low tax state early. This would result in a situation where both the difference in pre-shipping profit and the difference in shipping distance are positive. The lower bound is identified from situations where Amazon opened an FC in a high-population/high tax state early, which would result in both the difference in pre-shipping profit and shipping distance being negative. Therefore, we expect a positive relationship between the differences in pre-shipping profit and the average shipping distance. Figure 6b shows that this positive relationship exists across all perturbations. Of course, this relationship becomes stronger when we limit the perturbations to the ones that qualify for experiments 1 and 2.

In Figures 6a, 6c, and 6d we examine the sources of this positive relationship by decomposing pre-shipping profit into the variable profits and the fixed costs. Figure 6a shows that there is a positive correlation between the differences in pre-shipping variable profit and the difference in the average shipping distance, a relationship that is due to the fact that the difference in the average shipping distance is negatively correlated with the difference in the number of households who are charged sales tax, seen in Figure 6c. Figure 6d suggests that there is actually a positive correlation between the difference in the average shipping distance and the difference in total fixed costs across all perturbations. This implies that FCs that are close to highly populated areas have lower fixed costs, something that appears to be driven by a few outliers in the upper right and lower left quadrant of the figure. However, in the perturbations that qualify for our experiments, this relationship is negative.

Together the plots in Figure 6 suggest that Amazon faces a trade-off between charging sales tax and shipping distance, with fixed costs also playing an important role. Therefore, identification of the shipping cost parameter θ comes from both the tax/distance and the fixed cost/distance trade-off.

Results

We present the 95% confidence sets for θ under various definitions of the instruments in the first two panels of Table 10. We also show the number of perturbations that qualify for each experiment.

We first focus on Panel I, where we use the data to construct the instruments. The first row displays results when using only the indicators of group membership, $z_1^{j,k}$ and $z_2^{j,k}$, while the second row adds the interaction between the indicators and $\tilde{x}_+^{j,k}$. We have explored including higher order functions of $\tilde{x}_+^{j,k}$, with little change in the estimated bounds. The estimates imply that the shipping cost per dollar per mile is between 4.43E-05 and 1.04E-04, with the addition of the continuous distance-based instruments leading to a tighter interval, which we find to be true across specifications.

Next, we turn to the second panel, where we display results when using a first-stage regression to construct the instruments. Interestingly, this results in an upward shift of the confidence interval, suggesting that correlation in the error terms across perturbations leads to a downward bias in the estimates. The 95% confidence interval has a lower bound of around 5.50E-05 and an upper bound of up to 2.26E-04. None of the confidence intervals include zero, suggesting that – as expected – shipping cost and thus any cost savings from expansion are significantly different from 0.

The baseline results relied on the assumption that Amazon serves each customer from the closest FC to the customer's location, which as discussed above, does not recognize that the nearest fulfillment center might not necessarily be the cheapest. In Panel 3, we therefore present shipping cost estimates under the assumption that Amazon serves a customer from the FC with the lowest transportation cost, where we predict the identity of the lowest-cost FC based on average delivery flows between county locations (see Section 3 and Appendix A). Under this model the correlation between the shipping distance and the sales tax effects of expansion is not as strong, which results in wider confidence intervals. The fact that the latter do not include zero provides further evidence, however, that expansion has led to significant cost savings.

For the remainder of the analysis, we focus on the estimates in the second row of Panel 2, where we employ both $\hat{z}_1^{j,k}$ and $\hat{z}_2^{j,k}$ and interactions with $\hat{x}_+^{j,k}$ as instruments. These results imply that it costs Amazon between \$0.17 and \$0.47 to ship \$30 of goods, the median order in the raw comScore data, 100 miles, net of shipping revenue. Therefore, net of shipping revenue, it costs Amazon somewhere between \$0.52 to \$1.40 for a \$30 shipment covering 300 miles, the average shipping distance in 2006, and between \$0.20 and \$0.53 in 2018, when the average shipping distance is only 115 miles, a significant amount of savings due to the expansion of the network.

Robustness

We compute three different sets of robustness estimates, which are presented in detail in Appendix C. First, in calculating revenue under a perturbed FC opening sequence, we assume that there is a state-specific difference between the opening date of a FC and the onset of Amazon's tax collection obligations, and that this lag equals to the observed lag, even when the opening date of the particular FC differs under the perturbation. This assumption would reflect that the lag in the change in tax liability might be due to Amazon negotiating with (or challenging) the state government regarding their obligation to collect taxes. This assumption would imply that Amazon correctly predicted the amount of time they would be able to avoid collecting sales tax in a given state, either through negotiation or legal challenges, at the time they made their FC roll-out decision.

Second, we allow consumers to value the availability of same day shipping. Because same day shipping was largely implemented after our sample period, it may be the case that consumers value this service and that Amazon placed FCs taking this into consideration, which would be outside of our model. The opening of small FCs outside of major metropolitan areas is evidence that such concerns may be material. As we cannot easily estimate this demand side effect with our data, we instead impose one, holding the remaining demand side parameters fixed at the estimated levels, and consider the effect of varying its strength on the estimated supply side.

Third, we allow for a more sophisticated logistic model. The true shipping logistics model that Amazon employs is likely significantly more complicated than the ones assumed above, where a consumer is always served by e.g., the closest FC. In this robustness check, we allow shipments to county i to come from multiple FCs, where the amount coming from each FC is a function of the average shipping cost estimate from the Commodity Flow Survey and the capacity of the FC, approximated by its size.

None of these adjustments change the results to a large degree, as we find a lower bound of the confidence interval for the shipping cost per dollar per mile as low as 1.06E-05 and an upper bound as high as 2.63E-04. This provides confidence that our model captures the important features of the data generating process and that our estimates are not systematically biased due to our assumptions.

Implications

We perform three exercises assuming that the shipping costs fall between 5.75E-05 and 1.55E-04 as under specification 2 in Panel 2. First, we compute the total revenue, the total shipping costs and the total fixed costs over our sample period (2006-2018) under the observed network and under the 2006 network, to isolate the contribution of the FC expansion to the firm's profitability over this time frame while holding other features of the market, such as prices and demand growth, fixed. The results are reported in the first four rows of Table 11. Predictably, Amazon loses revenue due to the tax effect and increases fixed operating costs as it expands its network, but it saves a significant amount in shipping costs. Specifically, the expansion led to a reduction in revenue of nearly \$9.6 billion, or a loss of \$2.2 billion in profit net of shipping cost, and an increase in fixed operating cost of \$3.3 billion. This is offset by between a \$5.0 and \$13.3 billion decrease shipping costs, which represents a 51% reduction.[35]

One weakness of this exercise is that it does not consider that Amazon might have difficulty fulfilling all orders in the later half of our sample with a distribution network held fixed at 2006 levels due to capacity constraints. Assuming that Amazon could adjust its short-run capacity at a cost in order to fulfill orders, this implies that we are underestimating the total fulfillment costs under the 2006 network and that our results are likely a lower bound on the total cost savings from expansion.

While this exercise considers cost savings during the FC roll-out, one might also be interested

[35]The linearity of the shipping cost function results in the same percentage reduction for the upper and lower bounds

in the long-run effects of expansion on the firm's variable cost and market position. While our model is challenged to address this directly, we form a measure of the long-term effect of expansion by comparing average profit margins including shipping costs in 2018, relative to what they would have been had the distribution network not grown. We compute weighted average profit margins in 2018 under the observed network of FCs and under the network in 2006, where the weights are based on the expenditures from each county:

$$\sum_i (\mu - \theta d_{i,2018}(a))\omega_{i,2018}$$

where $\omega_{i,2018}$ is county i's share of aggregate 2018 revenue and μ, as above, is 0.23. For the actual network, a^0, the overall margin is 22.4 cents for every dollar sold under the lower bound of the confidence interval and 21.5 cents under the upper bound. Under the 2006 network, a^{2006}, the margin falls to between 21.3 cents and 18.4 cents per dollar, a reduction of between 1.2 and 3.1 cents for every dollar, or between 5 and 14%. As above, these numbers are likely lower bounds when considering that expansion also eases the costs of adjusting short term capacity.

Finally, we estimate how much the expansion has reduced the average net cost of shipping an order over the sample period. We calculate the average net shipping cost per order by summing all distribution-related costs: the yearly fixed costs and net shipping costs across all counties based on our estimates of θ. We then divide these total costs by the number of orders in each year. As our model does not predict the number of orders, we rely on the raw comScore data to create an annual measure of average expenditure per order for the sample shoppers and use the predicted expenditures from the model to calculate the expected number of orders per year. Figure 7 shows that the net average shipping cost falls from between \$2.67 and \$4.69 in 2006 to between \$1.09 and \$2.14 in 2013. Our model projects that the cost per order will continue to fall to approximately \$1 by 2018.

Put together, these three exercises demonstrate that Amazon's expansion has resulted in significant cost decreases as a result of the FC expansion.

Discussion

Using the results of both the demand and supply models, we conclude with some suggestive evidence of the sources of the increase in concentration in online retail over the past decade that motivated the analysis. Two sources of increased concentration present themselves: first, the quality of the leading online platforms could have increased over time; second, due to economies of scale and density in distribution, the large online retailers increasingly benefit from cost advantages. Here, we attempt to evaluate, in a reduced form, the gains in Amazon's market position that can be attributed to each of these two factors.

To do so, we decompose mode level preferences plotted in Figure 4 into a price effect and a quality effect. Recall, that the fixed effect, $\hat{\xi}_{jt}$, includes the effect of prices, meaning that we can decompose the mode preferences into a price component, $\bar{p}_{jt}$, and a quality component, Ξ_{jt}:

$$\bar{\xi}_{jt} = \overbrace{\underbrace{(1-\hat{\sigma})\ln(p_{jt})}_{\bar{p}_{jt}} + \tilde{\xi}_{jt}}^{\hat{\xi}_{jt}} \underbrace{\phantom{\tilde{\xi}_{jt}} + \hat{\beta}_j \bar{Z}_t + \hat{\lambda}_j \bar{C}_t}_{\Xi_{jt}}$$

Our model does not predict price levels; we instead assume constant margins over time. We therefore go to the raw data in order to form a 'price-index' for each mode and year.[36] Figure 8a plots these price indices in absolute terms. For reference, we also translate the absolute price index into a contribution to mode preferences, $(1-\hat{\sigma})\ln(p_{jt})$, on the right axis. Amazon's prices steadily decline from 2006-2013, while the prices of non-taxed online competitors (mode 3) stay relatively constant. The prices of other taxed online competitors (mode 2) follow Amazon's early on, but level off from 2010 to 2013. Returning to Figure 7 and relying on the mid-points of the confidence interval, Amazon's costs decrease by approximately 55% over this time period, while prices decline by approximately 47%. While we cannot causally relate the price decreases to the cost declines using our model, this provides suggestive evidence that Amazon passed some of the cost savings realized through the expansion of the network onto consumers, resulting in a competitive advantage.

Figure 8b displays the quality contribution to mode preferences (Ξ_{jt}), which again shows an increase on Amazon compared to the other two modes. While there are many different components that could make up this quality, we examine one which we can measure with the raw data: product assortment. We use the comScore data aggregated to the category level and calculate the HHI using expenditure shares across product categories for each mode and year and plot it in Figure 8c. Amazon's assortment HHI is slightly higher than mode 2 in 2006, implying that Amazon sold fewer product varieties at that time than all the taxed sites put together. Mode 3 sells an even larger variety, which is not surprising considering that the category is made up of the large collection of non-taxed online competitors. Over time, we see that Amazon has increased its assortment significantly, reaching approximately the levels of mode 3 by 2013. Both modes 2 and 3 have kept their assortment relatively constant over time. The patterns in these two graphs suggest that at least some of the increase in Amazon's share is due to an increase in assortment.

[36]We run the following OLS regression:

$$p_{ocjt} = \zeta_{jt} + \lambda_c + \epsilon_{ocjt}$$

where the dependent variable is the tax and shipping inclusive price on purchase occasion o, in product category c, on mode j, and in year t. In practice, in order to calculate the dependent variable, we subtract the sum of the tax exclusive price of the individual products in a basket from the basket total and then divide this by the total number of products in the basket. This provides a measure of the average of the shipping charges and taxes per product within a basket. We then add this average to the price of the individual product to form the tax inclusive price of the given product. We use the estimated value of ζ_{jt} as the price index of mode j in year t.

6 Conclusion

We examine sources for the increase in concentration of online retail over the last decade, with a focus on the economies of density associated with Amazon's large scale distribution network. Amazon benefits from its fulfillment centers being close to consumers for two potential reasons: first, the customer herself may value faster shipping, and second, it saves on delivery costs. At the same time, state laws dictate that Amazon must charge sales tax to consumers in most states where the company has a physical presence. By raising the tax-inclusive price the consumer faces, this reduces consumers' willingness-to-pay for Amazon's services. Further, locating close to significant population clusters also comes with higher fixed costs of operating fulfillment centers. In order to analyze the connection between observed expansion and the increase in concentration, we estimate the contribution of these different effects of expansion to firm profitability.

Our evidence is consistent with a trade-off between taxes and delivery costs savings: consumers dislike paying taxes, meaning the company must have benefited in other ways – faster shipping and/or reduced delivery costs – from FC network expansion. Our demand estimates indicate that consumer demand does not respond to our measures of shipping times, most likely due to the fact that expansion did not result in faster shipping speeds for the vast majority of consumers apart from the possibility of one-day shipping. Therefore, we find that the network expansion from 2006 to 2018 resulted in a loss in revenue of around \$9.6 billion dollars for Amazon, but at the same time reduced the average shipping distance from the FC to the consumer by around 180 miles by the end of 2018. We use a moment inequalities approach, together with the assumption that Amazon's observed network expansion path is optimal, to infer shipping cost savings from the observed fulfillment center network relative to alternative configurations. Results suggest that Amazon spends between \$0.17 and \$0.47 per 100 miles for every \$30 dollars of goods shipped, and that the reduction in shipping distance associated with the expansion of its distribution network has resulted in between \$5 and \$13.3 billion in savings on shipping costs and an increase in profit margins of up to 14%.

These results imply that economies of scale in e-commerce are an important driver of the channel's growth. We find that Amazon expanded its network of FCs in order to reduce its average handling cost by a significant margin. These variable cost savings were realized through important sunk and fixed cost investments, most of whom we do not account for in our estimation strategy that exploits the timing of individual openings, rather the magnitude of the geographic expansion. Therefore, this investment strategy is profitable only to the extent that Amazon expects to maintain a dominant position as an online retail platform.

While our model does not allow us to link these estimated cost savings and the observed concentration in online retail over the sample period directly, we show that the decrease in costs is mirrored by a decrease in prices relative to other online retailers thus providing suggestive evidence of the importance of economies of density. Further, we show that improvements in the quality of

Amazon's platform is another driver of the increase in concentration, with an increase in product assortment likely being a source for this improvement.

References

Agarwal, S., S. Chomsisengphet, T. Ho, and W. Qian (2013). Cross border shopping: Do consumers respond to taxes or prices? Working Paper.

Agarwal, S., N. Marwell, and L. McGranahan (2013). Consumption responses to temporary tax incentives: Evidence from state sales holidays. Working Paper.

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*(2), 339–356.

Alm, J. and M. I. Melnik (2005). Sales taxes and the decision to purchase online. *Public Finance Review 33*(2), 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*(2), 229–239.

Andrews, D. W. and G. Soares (2010). Inference for parameters defined by moment inequalities using generalized moment selection. *Econometrica 78*(1), 119–157.

Asplund, M., R. Friberg, and F. Wilander (2007). Demand and distance: evidence on cross-border shopping. *Journal of Public Economics 91*(1), 141–157.

Ballard, C. L. and J. Lee (2007). Internet purchases, cross-border shopping, and sales taxes. *National Tax Journal 60*(4), 711–725.

Barwick, P. J. (2008). What happens when wal-mart comes to town: An empirical analysis of the discount retailing industry. *Econometrica 76*(6), 1263–1316.

Baugh, B., I. Ben-David, and H. Park (2014). The "Amazon Tax": Empirical Evidence from Amazon and Main Street Retailers. Technical report, National Bureau of Economic Research.

Bruce, D., W. F. Fox, and L. Luna (2009). State and local government sales tax revenue losses from electronic commerce. *State Tax Notes 52*(7), 537–558.

Brynjolfsson, E., Y. J. Hu, and M. D. Smith (2003). Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers. *Managment Science 49*(11), 1580–1596.

Chetty, R., A. Looney, and K. Kroft (2009). Salience and taxation: Theory and evidence. *The American Economic Review 99*(4), 1145–1177.

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*(6), 2955–2980.

Einav, L., J. Levin, and N. Sundaresan (2014). Sales taxes and internet commerce. *The American Economic Review 104*(1), 1–26.

Ellison, G. and S. F. Ellison (2009). Tax sensitivity and home state preferences in internet purchasing. *American Economic Journal: Economic Policy 1*(2), 53–71.

Goolsbee, A. (2000a). In a world without borders: The impact of taxes on internet commerce. *The Quarterly Journal of Economics 115*(2), 561–576.

Goolsbee, A. (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*(1), 131–154.

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

Pakes, A., J. Porter, K. Ho, and J. Ishii (2015). Moment inequalities and their application. *Econometrica 83*(1), 315–334.

Quan, T. and K. Williams (2016). Product Variety, Across-Market Demand Heterogeneity, and the Value of Online Retail. Working Paper.

Scanlan, M. A. (2007). Tax Sensitivity in Electronic Commerce. *Fiscal Studies 28*(4), 417–436.

Smith, M. D. and E. Brynjolfsson (2001). Consumer decision-making at an internet shopbot: Brand still matters. *Journal of Industrial Economics 49*(4), 541–558.

Zheng, F. (2014). Spatial Competition and Preemptive Entry in the Discount Retail Industry. Working Paper.

Tables and Figures

Table 1: Expansion of Fulfillment Center Network

Year	Count of FCs	States with FC	States Taxed	Counties Taxed	Avg Distance (m)	Avg FC Size (000 sqft)	Avg FC Employees
2006	8	6	4	317	297	544	504
2007	9	7	4	317	292	522	484
2008	12	10	5	379	236	487	452
2009	17	10	5	379	227	570	529
2010	17	10	5	379	243	570	529
2011	24	10	5	379	236	657	610
2012	32	12	8	752	223	718	666
2013	41	14	16	1,189	207	765	709
2014	48	14	23	1,644	175	707	656
2015	54	16	26	1,937	152	665	616
2016	90	27	27	1,983	123	562	521
2017	101	28	28	2,047	116	552	511
2018	104	28	28	2,047	116	558	517

Notes: The number of states where purchases from Amazon are subject to sales tax exceeds the number of states with a fulfillment center due to states negotiating for sales tax collection to begin immediately upon entering an agreement with Amazon for a fulfillment center to be build in the state, even if the center opening occurs later. Average distance is the weighted average distance in miles from each county centroid to the nearest fulfillment center, where the weights are based on county number of households.

Table 2: Taxed and Non-Taxed Competitors

Sales Rank	Taxed	Non-Taxed
1	walmart.com	dell.com
2	jcpenney.com	qvc.com
3	staples.com	yahoo.net
4	victoriassecret.com	hsn.com
5	officedepot.com	yahoo.com
6	bestbuy.com	quillcorp.com
7	apple.com	overstock.com
8	target.com	ebay.com
9	sears.com	orientaltrading.com
10	costco.com	zappos.com
Total (%)	192 (34)	375 (66)

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

Table 3: Household Online Purchasing

Year	Households (000)	Counties (%)	Online Expenditures ($/year)	Amazon ($/year)	Taxed Competitors ($/year)	Non-Taxed Competitors ($/year)
2006	87.1	92	217	23	91	104
2007	90.0	92	215	25	92	98
2008	57.0	87	232	32	102	99
2009	55.9	85	226	37	96	94
2010	54.1	84	243	51	95	98
2011	63.4	87	301	78	121	102
2012	55.2	84	283	108	99	77
2013	46.5	78	273	100	91	83

Notes: Households refers to the number of households in the comScore data, together with the percentage of counties they represent. Expenditures are the unweighted average spending by the representative household in each county.

Table 4: Regression Models of Propensity of Purchasing from Amazon

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Tax Elasticity	-0.156**	-0.142*	-0.158**	-0.152**	-0.162**	-0.164**	-0.162**
	(0.074)	(0.074)	(0.075)	(0.074)	(0.075)	(0.075)	(0.075)
Local Express Delivery		-0.019*					
		(0.010)					
Log Distance			0.001				
			(0.002)				
1 or 2 Day Priority				0.019*			
				(0.011)			
1, 2, or 3 Day Package					-0.004		
					(0.004)		
1 Day First Class						-0.005	
						(0.004)	
1, 2, or 3 Day Standard							-0.004
							(0.004)
Obs	2,291,291	2,291,291	2,291,291	2,291,291	2,291,291	2,291,291	2,291,291
R-Sq	0.355	0.355	0.355	0.355	0.355	0.355	0.355

*** 1% ** 5% * 10%. Notes: Depicted are linear probability models of a household's likelihood of choosing Amazon on a given purchase occasion. Standard errors are clustered at the county/year level. Product category and year fixed effects included in all models along with consumer demographics and measures of offline competition. Local Express Delivery indicates whether same day shipping is available in the household's county at the time of the purchase occasion. Log Distance is the log of the great-circle from the centroid of the household's county to the nearest fulfillment center. For Specifications (4)-(7), shipping times are grouped into "long" and "short," with "long" being the excluded category.

Table 5: Regression Models of Annual Amazon Expenditures per County

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Amazon Purchase Taxed	-0.105*	-0.105*	-0.104*	-0.108*	-0.096	-0.104*	-0.096
	(0.060)	(0.060)	(0.061)	(0.061)	(0.061)	(0.061)	(0.061)
Local Express Delivery		0.147					
		(0.185)					
Log Distance			-0.002				
			(0.035)				
1 or 2 Day Priority				-0.057			
				(0.107)			
1, 2, or 3 Day Package					0.033		
					(0.042)		
1 Day First Class						0.004	
						(0.052)	
1, 2, or 3 Day Standard							0.033
							(0.042)
Obs	12,486	12,486	12,486	12,486	12,486	12,486	12,486
R-Sq	0.448	0.448	0.448	0.448	0.448	0.448	0.448

*** 1% ** 5% * 10%. Notes: Depicted are regression models of the log of annual Amazon expenditures per county. Year and county fixed effects included in all models along with consumer demographics and measures of offline competition. Amazon Purchase Taxed is an indicator variable that a purchase from Amazon is subject to sales tax in the county in a given year. Local Express Delivery indicates whether same day shipping is available in the county in that year. Log Distance is the log of the great-circle distance from the country centroid to the nearest fulfillment center. For Specifications (4)-(7), shipping times are grouped into “long” and “short,” with “long” being the excluded category.

Table 6: CES Demand Estimates: Effects of Taxes and Shipping Speed

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Tax Elasticity	-1.430***	-1.454***	-1.471***	-1.440***	-1.400***	-1.429***	-1.400***
	(0.486)	(0.486)	(0.490)	(0.486)	(0.486)	(0.486)	(0.486)
Local Express Delivery		0.186					
		(0.149)					
Log Distance			-0.011				
			(0.016)				
1 or 2 Day Priority				-0.058			
				(0.058)			
1, 2, or 3 Day Package					0.040		
					(0.031)		
1 Day First Class						0.002	
						(0.034)	
1, 2, or 3 Day Standard							0.040
							(0.031)
Obs	42,390	42,390	42,390	42,390	42,390	42,390	42,390
R-Sq	0.236	0.236	0.236	0.236	0.236	0.236	0.236

*** 1% ** 5% * 10%. Notes: Presented are the estimated tax elasticity and the effect of shipping speeds. Regressions include county, mode-year, and region-year fixed effects along with mode-level effects of local demographics and competition. Shipping speed proxies described in notes to Table 5.

Table 7: CES Demand Estimates: Demographics

	Amazon	Mode 2	Mode 3
Total Offline Competitors	-0.002	-0.002	-0.003
	(0.016)	(0.016)	(0.016)
Large Offline Competitors	-0.010	-0.016***	-0.010*
	(0.006)	(0.006)	(0.006)
Income	0.655***	0.638***	0.629***
	(0.020)	(0.018)	(0.018)
% Pop Black	-0.309	0.328	
	(0.271)	(0.241)	
% Pop White	-0.130	0.442*	
	(0.258)	(0.253)	
% Pop Asian	1.660**	-0.386	
	(0.821)	(0.800)	

*** 1% ** 5% * 10%. Notes: Presented are the estimated demographic effects from Specification (1) in Table 6. Total Offline Competitors denotes the number of retail establishments in the county, measured in hundreds of establishments, and Large Offline Competitors the number of retail establishments with more than 50 employees. Mode 2 denotes taxed online competitors and mode 3 denotes non-taxed online competitors.

Table 8: CES Demand Estimates: Mode Fixed Effects

Year	Amazon	Mode 2	Mode 3
2006	-1.777***	-0.597*	
	(0.343)	(0.322)	
2007	-1.421***	-0.303	0.191**
	(0.352)	(0.332)	(0.084)
2008	-0.850**	0.088	0.461***
	(0.354)	(0.334)	(0.091)
2009	-0.630*	0.123	0.464***
	(0.355)	(0.334)	(0.095)
2010	-0.578	-0.020	0.326***
	(0.356)	(0.335)	(0.098)
2011	-0.071	0.338	0.496***
	(0.354)	(0.334)	(0.095)
2012	0.220	0.536	0.684***
	(0.352)	(0.333)	(0.094)
2013	0.663*	0.711**	0.934***
	(0.354)	(0.335)	(0.102)

*** 1% ** 5% * 10%. Notes: Presented are the estimated mode-year fixed effects from Specification (1) in Table 6.

Table 9: Predicted Revenue ($Billion)

Year	Amazon	Mode 2	Mode 3	Total
2006	5.61	21.21	25.98	52.79
2007	7.77	27.76	30.65	66.18
2008	9.78	27.81	29.01	66.60
2009	12.28	29.06	29.37	70.71
2010	17.88	34.91	35.38	88.17
2011	25.27	42.56	35.73	103.57
2012	32.46	50.22	41.82	124.50
2013	40.79	51.23	43.22	135.24
2014	53.59	61.70	49.14	164.44
2015	65.85	66.57	52.33	184.74
2016	80.04	71.30	55.43	206.78
2017	96.09	75.90	58.44	230.43
2018	114.11	80.33	61.34	255.78

Notes: Depicted are the revenue predictions based on the model estimates. Out-of-sample revenues from 2014-2018 are predicted based on projections of the mode-year and region-year fixed effects. See Section 5 for details. Mode 2 denotes taxed online competitors and mode 3 denotes non-taxed online competitors.

Table 10: Shipping Cost Estimates

	Instruments	95% Confidence Set	# Perturbations Group 1	# Perturbations Group 2
		I. Closest FC: No First Stage		
(1)	$\tilde{z}^{j,k}$	[3.43E-05,1.04E-04]	332	496
(2)	$\tilde{z}^{j,k}\tilde{z}^{j,k} \times \tilde{x}^{j,k}_{+}$	[3.60E-05,9.29E-05]	332	496
		II. Closest FC: First Stage		
(1)	$\hat{z}^{j,k}$	[5.45E-05,2.26E-04]	263	564
(2)	$\hat{z}^{j,k}, \hat{z}^{j,k} \times \hat{x}^{j,k}_{+}$	[5.75E-05,1.55E-04]	263	564
		III. Lowest Cost FC: First Stage		
(1)	$\hat{z}^{j,k}$	[3.90E-05,1.20E-04]	180	557
(2)	$\hat{z}^{j,k}, \hat{z}^{j,k} \times \hat{x}^{j,k}_{+}$	[3.50E-05,4.56E-04]	180	557

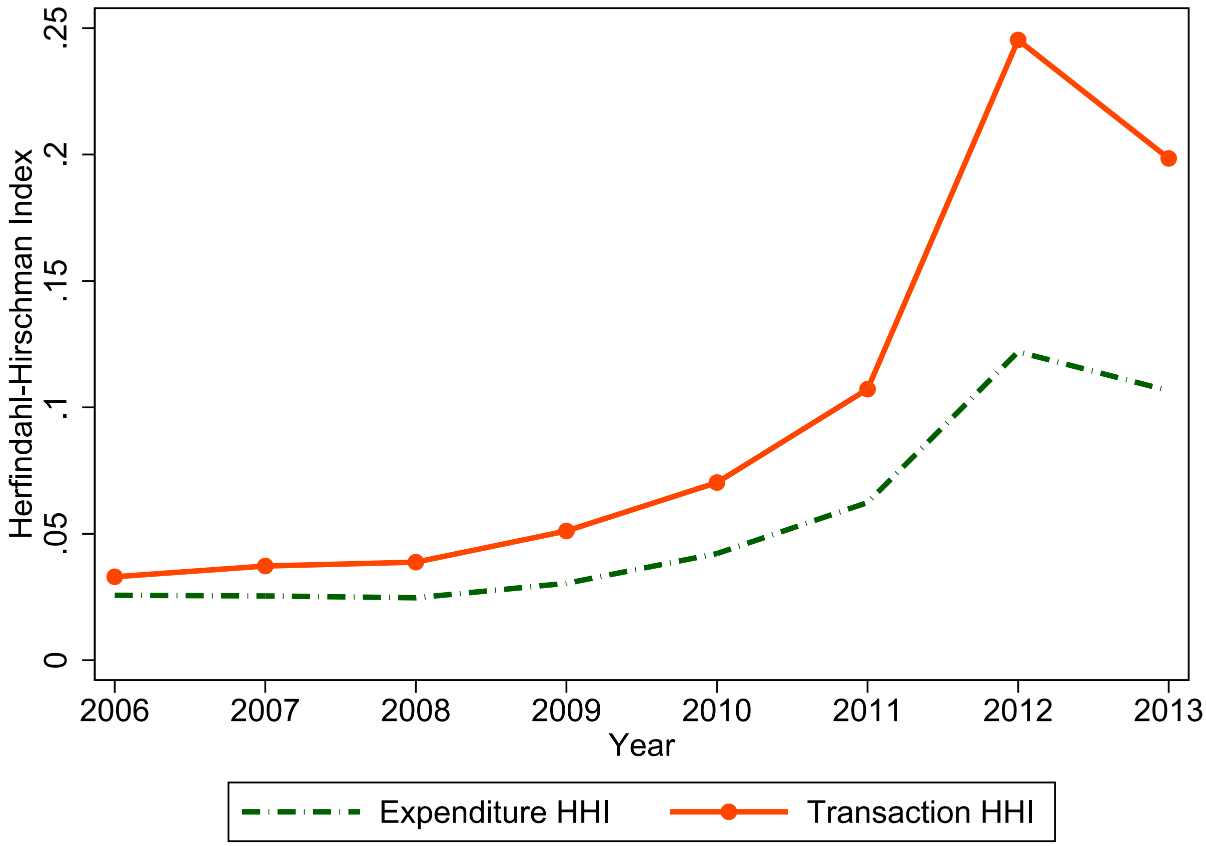
Notes: We depict confidence sets for Amazon's net shipping cost per dollar of goods sold per mile constructed using the methods of Andrews and Soares (2010). Panel I uses in Specification (1) two indicators for informative perturbations as instruments $\tilde{z}_a$, and adds in Specification (2) these indicators interacted with the shifted change in revenue-scaled shipping distance $\tilde{x}_a^+$. Panels II and III use indicators of informative perturbations $\hat{z}$ constructed from exogenous contributions to pre-shipping profit and revenue-scaled shipping distance. Panel III assumes that Amazon uses the lowest-cost FC to serve the customers in a county. The last two columns indicate the number of perturbations that belong to each experiment.

Table 11: Effect of Fulfillment Center Expansion

	Observed Network	2006 Network	Change	Percent Change (%)
$\mu\times$Revenue ($B)	129.15	131.36	-2.21	-1.71
Labor Cost ($B)	4.30	1.41	2.90	67.34
Land Cost ($B)	0.57	0.14	0.43	74.94
Shipping Cost ($B)	[4.66 , 12.55]	[9.62 , 25.91]	[-4.96 , -13.36]	-51.56
2018 Margin (%)	[22.4 , 21.5]	[21.3 , 18.4]	[1.2 , 3.1]	[5.13 , 14.43]

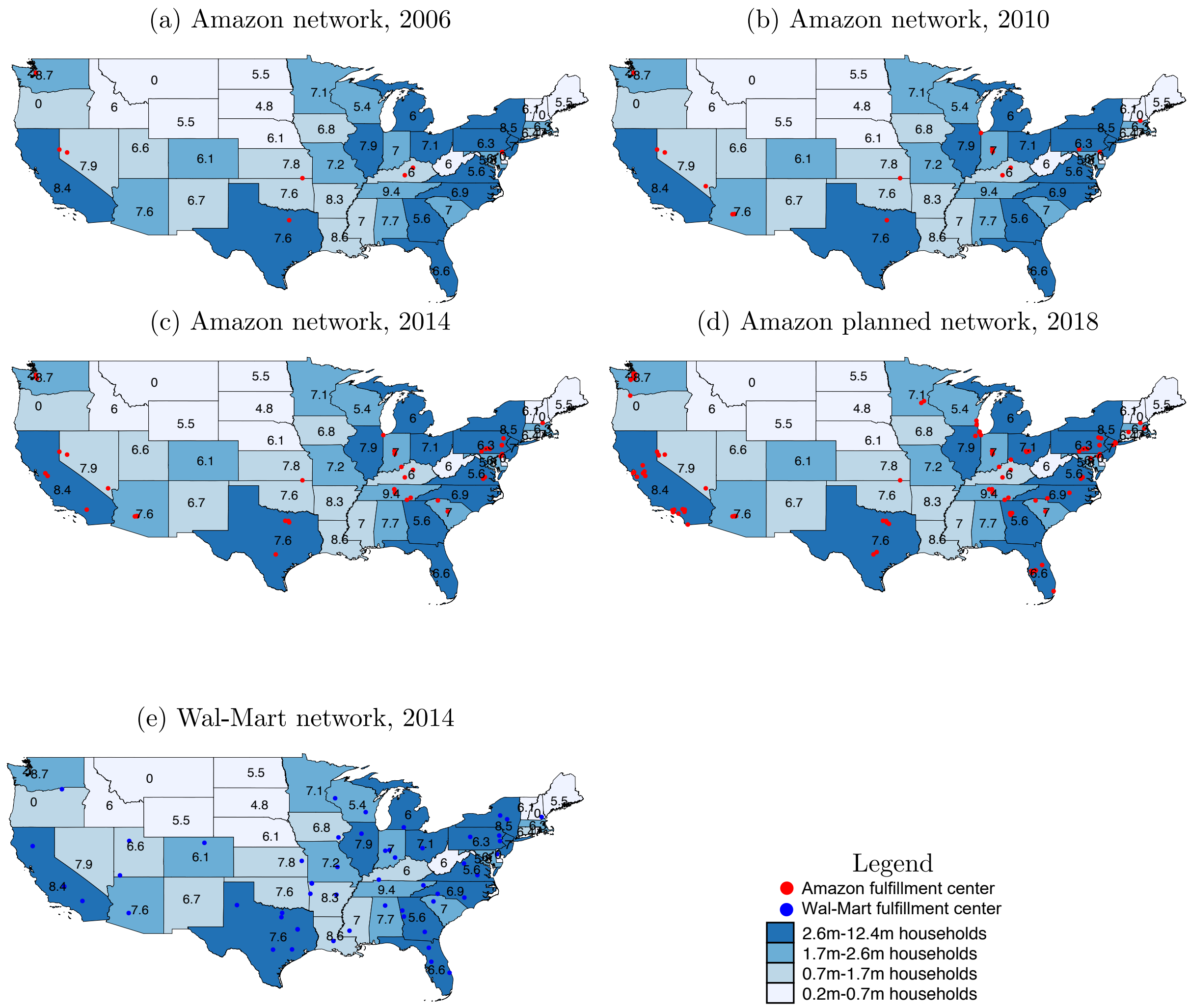
Notes: Table compares outcomes under observed fulfillment center roll-out to outcomes holding the fulfillment center network fixed as in 2006. For this counterfactual fulfillment center network, we predict revenue assuming that the tax status of a county does not change from 2006 onwards and that the shipping distance remains fixed. The total shipping costs are calculated using the 95% confidence set for shipping cost per dollar of goods sold per mile under Specification (2) from Panel II in Table 10.

Figure 1: Concentration in Online Retail, 2006-2013



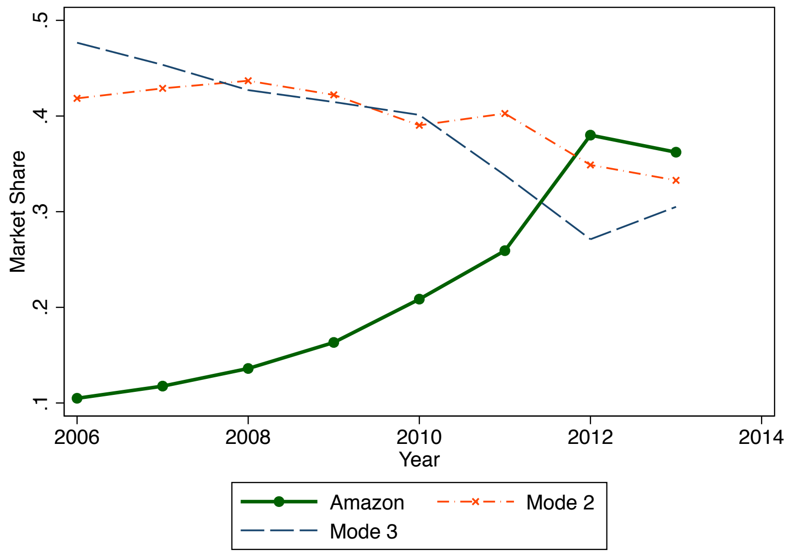
Notes: Authors' calculations using comScore Web Behavior database restricted to categories in which Amazon has a presence. "Expenditure HHI" and "Transaction HHI" denote Herfindahl-Hirshman indices based on market shares using the raw expenditures and number of transactions, respectively, at each retailer in each year.

Figure 2: Amazon's Fulfillment Center Expansion



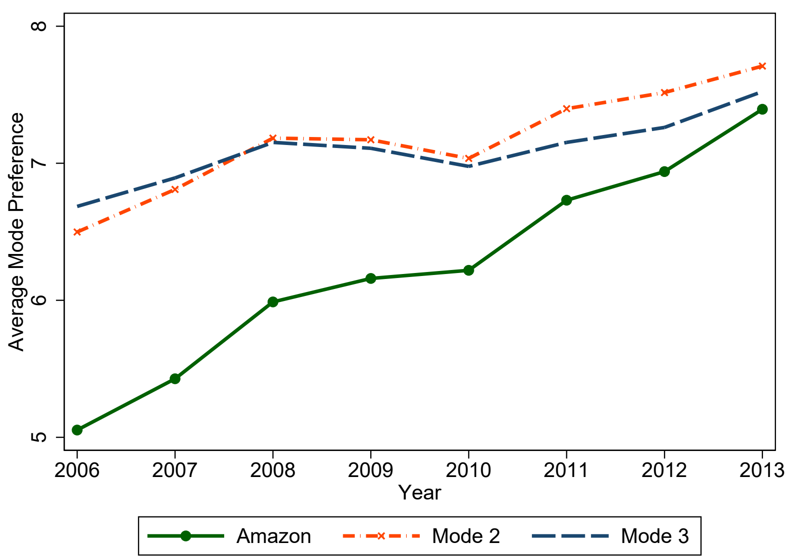
Notes: Maps exclude Amazon Fresh, Returns, and Sortation Centers. Shading represents number of households per state. We display the state-level population-weighted average 2013 sales tax rate in each map

Figure 3: Online Market Shares, 2006-2013



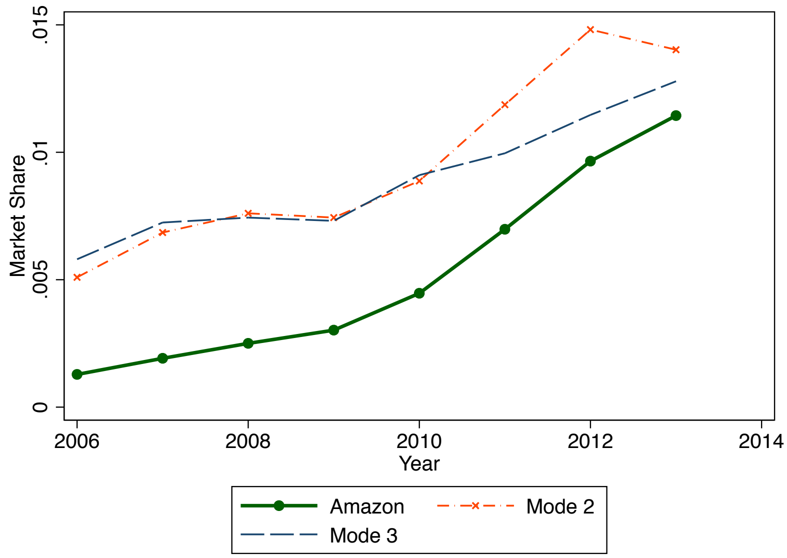
Notes: Shares are calculated using aggregate revenue derived from the adjusted expenditures of the representative household. Mode 2 denotes taxed online competitors and mode 3 denotes non-taxed online competitors.

Figure 4: Average Mode Preference, 2006-2013



Notes: We display the time series of the mean of mode preferences, derived from the estimated mode fixed effects in Table 8 and demographic interactions in Table 7 evaluated at the mean of the demographic attributes. Mode 2 denotes taxed online competitors and mode 3 denotes non-taxed online competitors.

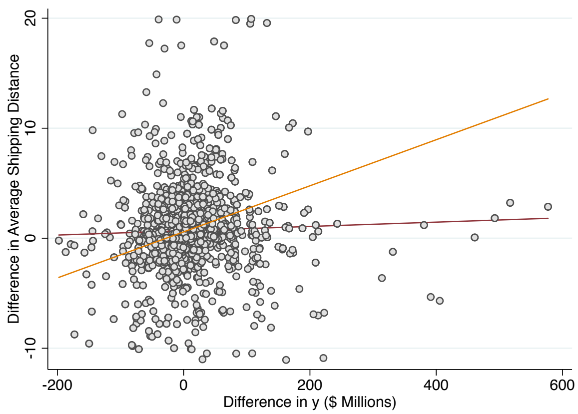
Figure 5: Estimated Retail Market Shares of Online Modes, 2006-2013



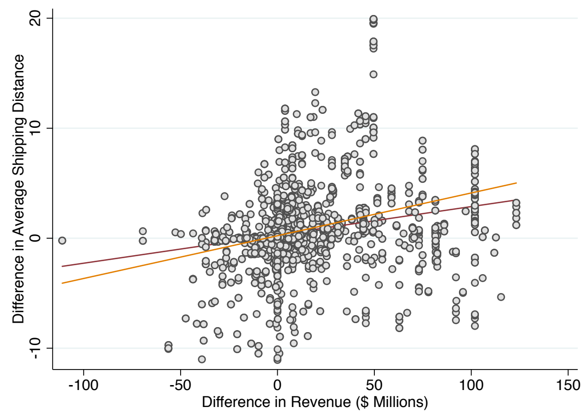
Notes: Market shares are calculated using aggregate spending derived from predicted annual household level expenditures for each county under the estimates of specification (1) in Table 6. Mode 2 denotes taxed online competitors and mode 3 denotes non-taxed online competitors.

Figure 6: Comparison of Outcomes under Perturbed and Actual Fulfillment Center Networks

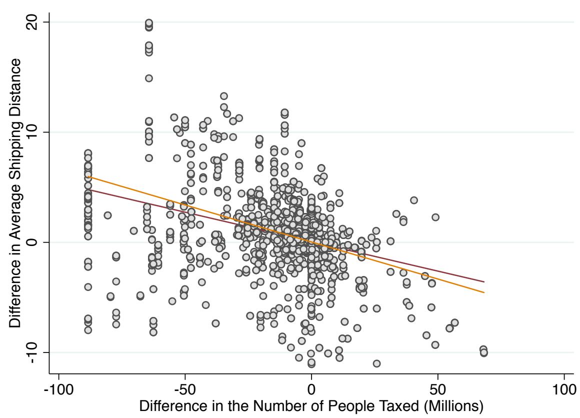
(a) Change in discounted sum of pre-shipping profit ($\mu \times R - F$) 2006-2018



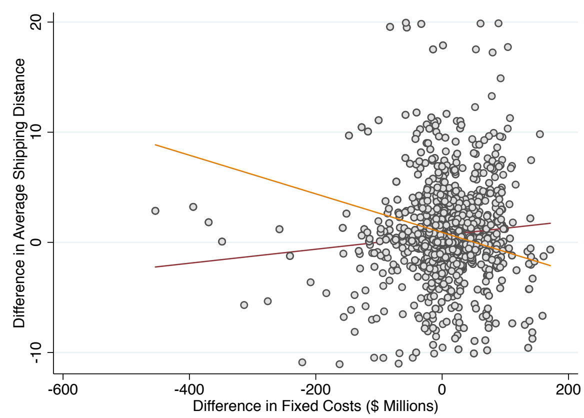
(b) Change in discounted sum of pre-shipping variable profit ($\mu \times R$) 2006-2018



(c) Difference in number of households taxed 2006-2018

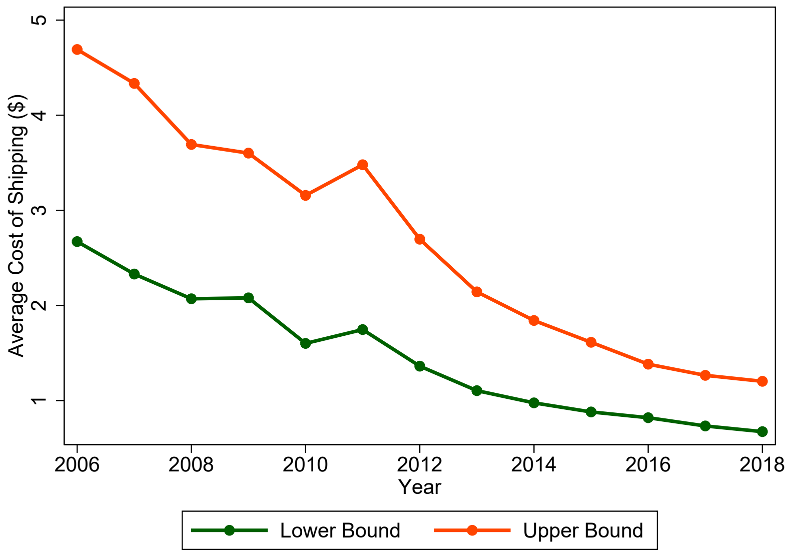


(d) Difference in discounted sum of fixed costs 2006-2018



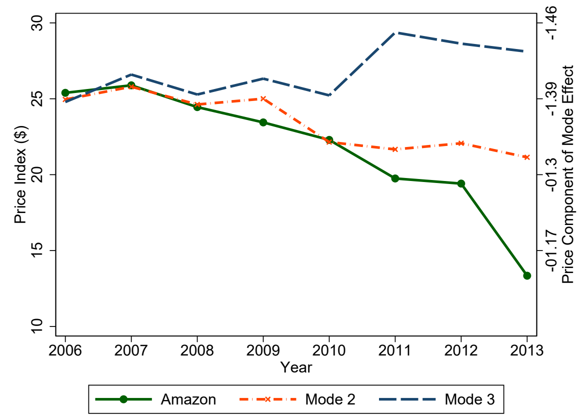
Notes: Each point represents one of the 1,575 perturbations. The y-axis in all four plots is the difference in the household level average shipping distance between the observed network and the perturbation from 2006-2018. The red and yellow lines depict the lines of best fit for all perturbations and for the subset of perturbations that qualify for experiments 1 or 2, respectively.

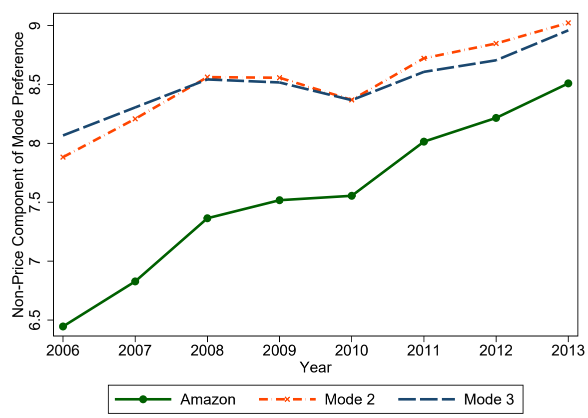
Figure 7: Average Shipping Cost per Order



Notes: The figure displays the 95% confidence interval for the average net cost of shipping one order using the estimates from the fourth row in Table 10. We construct the net shipping cost per order by dividing total annual fixed and shipping costs predicted by the model by the annual number of orders, estimated using a combination of the raw data and predictions of the model.

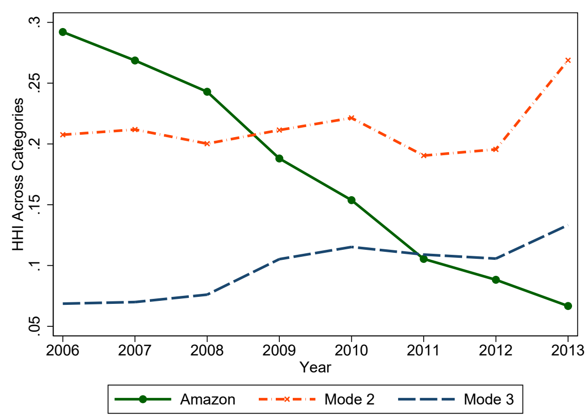
Figure 8: Sources of Concentration





(a) Prices

(b) Quality



(c) Assortment

Notes: Price indices calculated using raw comScore Web Behavior data for categories in which Amazon has a presence, netting out mean category-level prices. Quality is the residual mode preference, after subtracting the price component. Assortment denotes the HHI at the product category level, summing the squared revenue share of each product category. Mode 2 denotes taxed online competitors and mode 3 denotes non-taxed online competitors.

A Appendix: Data

Calculating Expenditures

In what follows, we provide a description of the procedure to calculate the expenditures for the representative household in county i and year t on the three different modes of shopping. First, we annualize each household's online expenditures by grossing up the observed weekly expenditures by 52 weeks, where the observed weekly expenditures are the total expenditures divided by the total number of weeks we observe any browsing behavior by the household. We do this because the browsing data suggests that many households are only in the sample for a short time period. Denote the expenditures for household h in year t as e_{ht}.

We present the annual average of household online expenditures in the first column of Table A.1. The second column displays the average number of transactions calculated in a similar manner to expenditures, while the third column shows the percentage of households with zero online expenditures. Interestingly, we see decreasing average expenditures and transactions over time. Column 3 indicates that this is due to an increasing percentage of comScore households with zero online purchases. This not in line with outside evidence about the take-up of online purchasing; eMarketer estimates the share of such internet users who do not make online purchases is 32% in 2009.

While comScore's sampling procedure could be causing this pattern, a more plausible explanation is that comScore does not record a household's transactions because the household deactivated the behavior monitor or because the household uses a second computer (e.g., at work) or a mobile device for all of their online shopping. Because we do not know which households truly do not purchase online and have zero expenditures, we choose to exclude all households with zero transactions from the sample. In order to account for the extensive margin of becoming an online buyer, we supplement the comScore data with survey data from Forrester Research, Inc. (`www.forrester.com`).

The survey data comes from Forrester's "North American Technographics Online Benchmark Survey", which was conducted from 2006-2007 by mail and 2010-2014 online. It surveyed between 30,000 and 60,000 households in the United States and Canada. The exact content of the survey questions varies by year, but generally the most pertinent question for us is the one which asks the user "Have you bought anything online in the past three months?". The survey also asks for information about the age, income, race, and zip code of the respondent. Documentation for the survey is available from the authors upon request. Patterns in the Forrester responses are in line with anecdotal evidence about the use of the take up of e-commerce (see the last column of Table A.1).

We first use the Forrester data to estimate the extent of online shopping for a given demographic group by running the following linear probability regression:

$$Pr(D_{ht} = 1) = \beta_0 + \beta_1 Z_h + \beta_2 I_h + \gamma_t + \epsilon_{ht}$$

where the dependent variable, D_{ht}, indicates whether the household made any online purchases in the past three months. The explanatory variables Z denote a set of dummy variables indicating the characteristics of respondent household h (i.e., race, age and census region), I is a continuous measure of their income, and γ_t are yearly dummy variables. We then use the estimates of the model to predict the probability that a household in the comScore data made any online purchase in year t, $\hat{p}_{ht}$, 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 these predicted purchase probabilities, we calculate the Forrester adjusted expected online expenditures from data for comScore households with positive expenditures, $\tilde{e}_{ht} = \hat{p}_{ht} e_{ht}$. We report raw averages of online expenditures for the Forrester adjusted data, along with transactions that are adjusted in the same way as expenditures, in Table A.1. Household expenditures increase from \$250 per year in 2006 to \$374 in 2013.

Table A.1: Household Online Purchasing

	Raw comScore Data					Forrester
Year	Expenditures ($)	Transactions	Offline Only (%)	Adjusted Expenditures ($)	Adjusted Transactions	Offline Only (%)
2006	248	2.5	49.8	250	2.5	55.5
2007	260	2.6	51.2	248	2.4	60.8
2008	201	2.1	59.3	274	2.8	-
2009	148	1.5	66.7	282	2.9	-
2010	129	1.4	67.9	288	3.1	32.1
2011	134	1.4	69.3	342	3.5	23.0
2012	155	1.8	63.6	327	3.8	23.9
2013	164	2.2	60.1	374	5.0	15.5

Notes: Expenditures and Transactions are the average amount of per-household spending and number of online purchase transactions. Offline Only denotes the share of households with no online expenditures in the comScore data. Forrester 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.

Next, we aggregate across households to derive the expenditure of a representative household in county i in year t on online retail. We start by calculating the average online expenditures for demographic group z in county i and year t, in aggregate and for mode j:

$$\bar{e}_{it}^{z} = \frac{1}{N_{it}(z)} \sum_{h \in H_{it}(z)} \tilde{e}_{ht} \text{ and}$$

where $H_{it}(z)$ is the set of comScore households in demographic group z in county i and $N_{it}(z)$ is the size of this set.

Each demographic group is defined by age of the head of the household, the household income level, and the race of the head of the household. We find the comScore sample of households to be generally representative of the Unites States population according the 2010 Census, with three exceptions: (1) the head of the household is younger, (2) a higher percentage of the households are white, and (3) the household income is higher. De los Santos et al. (2012) similarly 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. We account for the possibility that comScore may be over- or undersampling certain demographic groups using sampling weights. We classify each comScore household by income and race and age of the head of household, and calculate a household-specific sampling weight based on the relative prevalence of each demographic category in the comScore data and the 2010 Census at the county level.

Using the resulting weights for demographic group z, denoted w_i^z, and the average expenditures from the comScore data, we create the total online expenditures for a representative consumer in a given county:

$$e_{it} = \sum_z w_i^z \bar{e}_{it}^z$$

and the modes' representative shares and expenditures:

$$s_{ijt} = \sum_z w_i^z s_{ijt}^z \quad \text{and} \quad e_{itj} = s_{ijt} e_{it}.$$

Here, s_{ijt}^z are demographic group level shares for each shopping mode in each county calculated from the raw ComScore data:

$$s_{ijt}^z = \frac{\bar{e}_{ijt}^z}{\bar{e}_{it}^z}$$

$\bar{e}_{ijt}^z$ represents the annual average expenditures on mode j for county i across all households in demographic group z and $\bar{e}_{it}^z$ is the sum of $\bar{e}_{ijt}^z$ across all three modes.

Shipping Costs

In order to identify the lowest cost FC for a given county, we estimate a shipment cost function that predicts the average per pound cost of shipping a package between a given origin and destination county using data from the 2012 Commodity Flow Survey (CFS). Assuming that this cost is proportional to the shipping cost for Amazon allows us to predict which FC is the lowest cost FC for each county.

The CFS consists of individual commodity shipments that are described by type, origin and destination, value, weight, mode of transportation, distance shipped, and ton-miles of commodities shipped. We select shipments from the CFS based on the modes of transportation that Amazon uses (e.g., air) and the commodity type Amazon sells (e.g., electronic merchandise).

We first calculate the costs per pound for each of these shipments, which are approximated by multiplying the unit cost (per-pound per-mile) by the distance shipped. We do not observe the unit cost, so for the shipment modes besides 'stepwise-pricing parcel' we assume the unit costs are equal to the 2012 'Average Freight Revenue Per Ton-mile', which we obtain from the Department of Transportation.[37,38] For parcel service shipments, we calculate the postage costs of a subsample of 131 shipments in the CFS using the Postage Price Calculator on USPS website, and then run a regression of the postage price on an interaction of the shipping weight and shipping distance. We then use the estimated coefficient as our unit cost per ton-mile and predict total parcel service costs per pound for the remainder of the data.

Next, we specify and estimate the following shipment total cost function:

$$C_i = \beta_0 + \beta_1 d_i + O_i + D_i + \epsilon_i \tag{A.1}$$

where C_i is the cost per pound of shipment i, d_i is the great-circle distance between the origin and the destination, and O_i and D_i are origin and destination county level fixed effects. Importantly, while the shipping costs were calculated using the 'distance shipped', d_i is measured as the great-circle distance, allowing us to predict a measure of costs in our data. In practice, we estimate a number of different specifications which include interactions and/or non-linear functions of distance and choose the model with the best fit. Using the estimates of $\hat{\beta}$, we are able to predict the average shipping cost per pound for each FC (f) and county (i) pair in our data, which we denote as C_{if}. The lowest cost FC is then the FC which solves:

$$f_i^* = \underset{f}{\operatorname{argmin}}\, C_{if}$$

B Appendix: Estimation

Calculating μ

Amazon reports the total amount of revenue in the "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. Therefore, these sales are accounted for through the 'multiplier' that we use to match the predictions of the model in the financial reports. See Section 4 for a discussion of these multipliers. They also report the "Cost of Goods Sold" for all of their sales. 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. This

[37] `https://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/national_transportation_statistics/html/table_03_21.html`

[38] Due to missing data for the 'Barge' and 'Truck' modes, we predict 2012 unit costs using the estimates of an AR-1 regression.

provides us with a value for the "gross margin". However, as Amazon states, the reported cost of goods sold includes both outbound shipping costs and inbound shipping costs (through wholesale prices). Recall that we assume that inbound shipping costs from suppliers to Amazon do not vary with the network of FCs. As we are estimating the outbound shipping costs, we exclude them from the gross margin by adding in the reported "Net Shipping Cost as a Percentage of Revenue" to the gross margin. This value grows over time, but we suspect this is due to an increase in Amazon's non-shipping related activities (i.e. Amazon Web Services and digital goods). We therefore set $\mu = 0.23$.

Estimation of Variance-Covariance Matrix

In our problem, we cannot simply calculate $\hat{\Sigma}(\theta^0)$ using equation 3.2 in AS because (1) there is first stage error in our revenue equation and (2) there exists correlation across a because the same FCs are used in many perturbations. We take the approach described in Section 8.1 of Holmes (2011) in order to account for these two issues.

Specifically, we draw a subsample of FCs of size $b << J$, where J is the total number of FCs. We keep the perturbations in which both the swapped FCs are in the subsample, thereby forming the 'deviation subsample' indexed by s. We then take a draw of the tax parameter $\hat{\sigma}_s$ from a normal distribution centered at $\hat{\sigma}$ with variance $\frac{J}{b}\hat{\Sigma}^{\sigma}$, where $\hat{\Sigma}^{\sigma}$ is the estimated variance from the first demand model. Note that in order to account for first stage error, we only draw the tax parameter and assume that all other features of the demand model approximately cancel out when calculating the perturbation. This assumption is made for simplification because for each draw of the parameters, we must compute all the perturbations. There are over 90 parameters of the demand model, meaning we would have to draw a large number of vectors of parameters in order to have variation across the full joint distribution. Instead, we take 100 draws of the tax parameters and calculate the set of perturbations for 100 times. While the non-linearity of the demand model imply that the other components of demand won't exactly cancel out, we believe that it is a good approximation and that the results will not be significantly affected.

We calculate $\tilde{y}_s^{j,k}$ and $\tilde{x}_s^{j,k}$ with the deviation subsample and the new tax parameter σ_s to form:

$$\tilde{m}_{l,s}(\theta^0) = \frac{1}{M}\sum_{j=1}^{J}\sum_{k=J+1}^{J} z_{l,s}^{j,k}[\tilde{y}_s^{j,k} - \theta^0\tilde{x}_s^{j,k}]$$

resulting in the l $\times$ 1 vector $\rightarrow; dem_{l,s}(\theta^0))$. The variance-covariance matrix is then estimated from the S different deviation subsamples:

$$\hat{\Sigma}(\theta^0) = \frac{b}{J}\text{varcov}(\tilde{m}(\theta^0))$$

where $\tilde{m}(\theta^0)$ is the S $\times$ l matrix of moments.

In practice, we set $b = \frac{J}{3}$ and $S = 200$ and use the diagonal matrix equivalent of $\hat{\Sigma}(\theta^0)$ as our weight matrix in Equation (20). The reason for the latter is that we ran into a number of instances where the full matrix was not invertible.

Bootstrap Moments and Variance Covariance Matrix

We follow the procedure is described in Section 4.2 of Andrews and Soares (2010). From the bootstrap sample of R perturbations, we calculate $\{\tilde{m}^*_r(\theta^0), \hat{\Sigma}^*_r(\theta^0)\}$ where:

$$\tilde{m}^*_r(\theta^0) = M_r^{\frac{1}{2}}(\hat{D}^*(\theta^0))^{-\frac{1}{2}}(\tilde{m}_r(\theta^0) - \tilde{m}(\theta^0))$$

and:

$$\hat{\Sigma}^*_r(\theta^0) = (\hat{D}^*_r(\theta^0))^{-\frac{1}{2}}\hat{\Sigma}^*_r(\theta^0)(\hat{D}^*_r(\theta^0))^{-\frac{1}{2}}$$

The elements of $\tilde{m}_r(\theta^0)$ are given by:

$$\tilde{m}^*_{l,r}(\theta^0) = \frac{1}{M_r}\sum_{j=1}^{J_r}\sum_{k=J_r+1}^{J} z^{j,k}_{l,r}[\tilde{y}^{j,k}_r - \theta^0\tilde{x}^{j,k}_r]$$

and:

$$\hat{D}^*_r(\theta^0) = Diag(\hat{\Sigma}^*_r(\theta^0))$$

$Diag(X)$ represents a vector containing the diagonal elements of X. Here, $\hat{\Sigma}^*_r(\theta^0)$ is computed using the same subsampling method described above, but replacing the the 'population' of FCs from which we draw the deviation subsample with J_r. We then eliminate the elements of $\{\tilde{m}^*_r(\theta^0), \hat{\Sigma}^*_r(\theta^0)\}$ where:

$$M^{\frac{1}{2}}\frac{\tilde{m}_l(\theta^0)}{\hat{\nu}_l(\theta^0)} > \kappa = (ln(M))^{\frac{1}{2}}$$

where $\tilde{m}_l(\theta^0)$ is the lth element of $\tilde{m}(\theta^0)$ and $\hat{\nu}_l(\theta^0)$ is the corresponding standard deviation taken from $\hat{\Sigma}(\theta^0)$. This results in a weakly smaller set of moments and their variance-covariance matrix:

$$\{\tilde{m}^{**}_r(\theta^0), \hat{\Sigma}^{**}_r(\theta^0)\}$$

C Appendix: Robustness

Demand

In what follows, we provide robustness to the CES demand estimates reported in the body of the paper. First, we provide the 'starting' estimates of the CES demand model, or the estimates without the multipliers, in Tables C.1 and C.2. While the estimates of the the tax effect and the convenience effect are similar to the estimation with the multipliers, the mode level fixed effects show some differences. This is especially true for mode 3, which has significantly larger fixed effects when the multipliers are employed.

Table C.1: CES Demand Estimates of Tax Elasticity: No Multipliers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Tax Elasticity	-1.493***	-1.504***	-1.516***	-1.504***	-1.469***	-1.499***	-1.469***
	(0.480)	(0.480)	(0.483)	(0.480)	(0.480)	(0.480)	(0.480)
Local Express Delivery		0.076					
		(0.131)					
Log Distance			-0.006				
			(0.016)				
1 or 2 Day Priority				-0.069			
				(0.057)			
1, 2, or 3 Day Package					0.031		
					(0.031)		
1 Day First Class						-0.013	
						(0.034)	
1, 2, or 3 Day Standard							0.031
							(0.031)
Obs	42,399	42,399	42,399	42,399	42,399	42,399	42,399
R-Sq	0.186	0.186	0.186	0.186	0.186	0.186	0.186

*** 1% ** 5% * 10%. Notes: Presented are the estimated tax elasticity and the effect of shipping speeds estimated without the multipliers. Regressions include county, mode-year, and region-year fixed effects along with mode-level effects of local demographics and competition. Shipping speed proxies described in notes to Table 5.

Table C.2: CES Demand Estimates of Mode Fixed Effects: No Multipliers

Year	Amazon	Mode 2	Mode 3
2006	-1.566***	-0.608*	
	(0.339)	(0.318)	
2007	-1.511***	-0.602*	-0.097
	(0.348)	(0.327)	(0.083)
2008	-0.990***	-0.220	0.161*
	(0.349)	(0.330)	(0.090)
2009	-0.809**	-0.230	0.121
	(0.350)	(0.330)	(0.094)
2010	-0.808**	-0.551*	-0.193**
	(0.351)	(0.331)	(0.097)
2011	-0.329	-0.251	-0.086
	(0.349)	(0.330)	(0.094)
2012	0.040	-0.453	-0.285***
	(0.347)	(0.328)	(0.093)
2013	0.175	-0.353	-0.121
	(0.350)	(0.331)	(0.100)

*** 1% ** 5% * 10%. Notes: Presented are the estimated mode-year fixed effects from Specification (1) in Table C.1.

Tables C.3, C.4, and C.5 provide results for a number of different variants of the demand model and results when adjusting the sample used in estimation, while Table C.6 displays results under alternative measures of the shipping distance. These estimates do not account for the multipliers.

In Table C.3, specification (1) uses weighted OLS, where the weights are analytical weights based on the number of observations used to calculate the average expenditures in a county, specification (2) incorporates the zeros by changing any mode-level expenditures that are equal to \$0 to \$1, specification (3) uses data from 2008 and beyond because the number of households in the comScore sample shifts after 2007, specification (5) does not use the Forrester correction to adjust expenditures, and specification (6) does not use population weights to create the expenditures. Specification (4) estimates the tax sensitive separately for each mode, as a way of accounting for Amazon Marketplace. That is, it allows for an Amazon specific tax effect that may be lower than the effect for offline retailers and mode 3 because not all customers on Amazon are charged sales tax.

In Table C.4, the first two specifications include a state-year level fixed effect (1) or a county-year level fixed effect (2) rather than a region-year level effect in order to account for the possibility that the unobservable factors which may be correlated with the tax variable are more local compared to our main specification. Specification (3) includes dummy variables indicating whether or not a FC entered the state in which the county is located one, two, three, or more than three years ago. The is intended to capture any sort of delayed marketing push or push to get users to adopt Amazon Prime which is associated with the entry of a FC. Finally, specification (4) includes an interaction between these dummy variables and the tax variable in order capture any possible learning which may happen over time. That is, it may take consumers time in order to learn that they are actually being charged higher prices than the were before, meaning the response to FC entry may be delayed. While we don't display the results here, we have also explored a learning effect with regards to shipping speeds, but again fail to find any evidence that our measures of shipping speeds affect consumer decisions. In Table C.5, we include an interaction between the income of the representative household in county i and the measure of shipping speed for the different specifications.

Finally, in Table C.6, we adjust the measure of shipping speed by allowing the consumer from county i to receive shipments from different FCs, where the amount of spending coming from each FC is a function of the estimated shipping cost from that FC and the capacity. See the Cost section of this Appendix for details on how we derive the distribution of goods coming from different FCs.

Overall, the results of these robustness checks are consistent with the base results with the estimated tax sensitivity being between -1.2 and -1.9 and the convenience effect not being significant. The only exception is the interaction between the priority shipping dummy variable and the log of income in Table C.5. This indicates that there is no effect of faster shipping at the average level of income, but that there may be for higher income counties. However, because this effect does not appear in any of the other specifications, we believe that the convenience effects are not strong in our sample.

Table C.3: CES Demand Estimates of Tax Elasticity: Robustness to Sample Construction

	(1)	(2)	(3)	(4)	(5)	(6)
Tax Elasticity	-1.307***	-1.687***	-1.867***		-1.203*	-1.199***
	(0.286)	(0.585)	(0.575)		(0.625)	(0.401)
Tax Elasticity (Amazon)				-1.166**		
				(0.515)		
Tax Elasticity (Mode 3)				-2.900***		
				(0.847)		
Obs	42,399	52,617	29,053	42,399	43,811	42,400
R-Sq	0.315	0.162	0.135	0.185	0.137	0.199
Regression	A-Weights	Zeros	2008-2013	Individual Tax Effect	No Forrester Adjustment	No Population Weights

*** 1% ** 5% * 10%. Notes: Presented are alternative CES demand model specifications. See text for model descriptions. Regressions include county, mode-year, and region-year fixed effects along with mode-level effects of local demographics and competition. Models that include shipping speed proxies in addition yield similar tax elasticities.

Table C.4: CES Demand Estimates of Tax Elasticity: Robustness to Controls (1)

Variable name	(1)	(2)	(3)	(4)
Tax Elasticity	-1.325***	-1.215***	-1.401***	-1.644***
	(0.498)	(0.443)	(0.481)	(0.525)
Entry Year 0			-0.034	
			(0.050)	
Entry Year -1			-0.015	
			(0.061)	
Entry Year -2			0.093	
			(0.072)	
Entry Year -3			0.046	
			(0.073)	
Entry Year $¡$ -3			-0.078	
			(0.058)	
Tax*(Entry Year 0)				-0.620
				(0.759)
Tax*(Entry Year -1)				-0.045
				(0.845)
Tax*(Entry Year -2)				-0.719
				(1.061)
Tax*(Entry Year -3)				-1.104
				(1.082)
Tax*(Entry Year $¡$ -3)				0.551
				(0.406)
Obs	42,399	42,399	42,399	42,399
R-Sq	0.195	0.240	0.186	0.186
Fixed Effects	Year-State, County	Year-County	County	County

*** 1% ** 5% * 10%. Notes: Presented are alternative CES demand model specifications. Specifications (1) and (2) include different delineations of the fixed effects, while (3) includes indicators of the time since the fulfillment center opened in a consumer's state and (4) includes interactions between these dummies and the tax rate charged. Regressions include additional mode-year fixed effects and mode-level effects of local demographics and competition. Models that include shipping speed proxies in addition yield similar tax elasticities.

Table C.5: CES Demand Estimates of Tax Elasticity: Robustness to Controls (2)

	(1)	(2)	(3)	(4)	(5)	(6)
Tax Elasticity	-1.504***	-1.515***	-1.501***	-1.467***	-1.496***	-1.467***
	(0.480)	(0.483)	(0.480)	(0.480)	(0.480)	(0.480)
Shipping Speed*Log Income	-0.023	-0.019	0.134***	-0.016	-0.024	-0.016
	(0.198)	(0.022)	(0.030)	(0.023)	(0.024)	(0.023)
Local Express Delivery	0.330					
	(2.193)					
Log Distance		0.201				
		(0.240)				
1 or 2 Day Priority			-1.499***			
			(0.324)			
1, 2, or 3 Day Package				0.202		
				(0.245)		
1 Day First Class					0.243	
					(0.258)	
1, 2, or 3 Day Standard						0.202
						(0.245)
Obs	42,399	42,399	42,399	42,399	42,399	42,399
R-Sq	0.186	0.186	0.186	0.186	0.186	0.186

*** 1% ** 5% * 10%. Notes: Presented are alternative CES demand model specifications. Each regression includes and interaction between household income and the measure of shipping speed. Regressions include county, mode-year, and region-year fixed effects along with mode-level effects of local demographics and competition. Models that include shipping speed proxies in addition yield similar tax elasticities.

Table C.6: CES Demand Estimates of Tax Elasticity: Robustness to Shipping Distance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Tax Elasticity	-1.469***	-1.497***	-1.468***	-1.497***	-1.508***	-1.531***	-1.474***
	(0.481)	(0.480)	(0.480)	(0.480)	(0.494)	(0.487)	(0.480)
Log Distance	-0.031	0.025	-0.028	0.016	-0.015	-0.022	-0.020
	(0.040)	(0.025)	(0.030)	(0.026)	(0.025)	(0.023)	(0.024)
Obs	42,399	42,399	42,399	42,399	40,755	41,695	42,399
R-Sq	0.186	0.186	0.186	0.186	0.185	0.186	0.186
$\{\beta_1, \beta_2\}$	$\{10, 0\}$	$\{10, 20\}$	$\{20, 0\}$	$\{20, 20\}$	$\{200, 0\}$	$\{200, 20\}$	Low Cost

*** 1% ** 5% * 10%. Notes: Presented are alternative CES demand model specifications that differ in their measure of shipping speed, proxied by various weighted average distances to the full set of fulfillment centers. For county i, the weight attributed to fulfillment center f is $\omega_{ift} = \exp(-\beta_1 cost_{if} + \beta_2 cap_f) \,/\, \sum_{g \in F_t} \exp(-\beta_1 cost_{ig} + \beta_2 cap_g)$, where the last row, labeled $\{\beta_1, \beta_2\}$, indicates the coefficients used in a given specification, *cost* denotes the estimated cost of shipping from i to f (see page 69), and *cap* denotes the square footage of the fulfillment center. "Low Cost" assigns all of the weight to the fulfillment center with the lowest logistics model shipping cost. Regressions include mode-year and region-year fixed effects and mode-level effects of local demographics and competition.

Cost: State Level Lag

We allow for the assumption about when the tax laws will change upon entry to be a function of the observed lag between entry and the change in the law as described in Section 5. For example, if we swap a FC that opened in Pennsylvania in 2008 and resulted in sales tax liabilities starting in 2011 with one that opened in Nevada in 1999 but resulted in sales tax liabilities only in 2014, then we would assume that the FC in Pennsylvania opens in 1999 and sales tax is collected beginning in 2002 (i.e., there is a 3 year lag as there was with the actual FC that opened in Pennsylvania in 2008).

The results in Table C.7 are mostly in line with the no-lag assumption, with slightly wider confidence intervals for the base model and tighter intervals for the low-cost model. Importantly, non of the confidence sets include zero.

We have also estimated the model under a opening date specific lag. In the example above, this assumption would imply that the lag would "follow" the swapped FCs. That is, taxes would begin being collected in Nevada in 2011 and in Pennsylvania in 2014. The results under this specification do not vary significantly from the estimates of the no-lag model.

Table C.7: Shipping Cost Estimates: Fixed Lag between FC Opening and Sales Tax in Effect

	Instruments	95% Confidence Set	# Perturbations Group 1	# Perturbations Group 2
		I. Closest FC: No First Stage		
(1)	$\tilde{z}_a$	[5.39E-05,1.67E-04]	313	505
(2)	$\tilde{z}_a, \tilde{z}_a \times \tilde{x}_a^+$	[6.11E-05,1.31E-04]	313	505
		II. Closest FC: First Stage		
(1)	$\hat{z}_a$	[4.28E-05,2.63E-04]	272	588
(2)	$\hat{z}_a, \hat{z}_a \times \hat{x}_a^+$	[3.89E-05,1.61E-04]	272	588
		III. Lowest Cost FC: First Stage		
(1)	$\hat{z}_a$	[5.46E-05,9.36E-05]	220	576
(2)	$\hat{z}_a, \hat{z}_a \times \hat{x}_a^+$	[5.76E-05,8.43E-05]	220	576

Notes: We depict confidence sets for Amazon's shipping cost per dollar of goods sold per mile constructed using the methods of Andrews and Soares (2010). We assume that for a perturbed fulfilment center opening sequence, the delay between an opening and sales taxes going into effect is the same as the one observed under the actual opening sequence. For a description of the specifications, see notes to Table 10.

Cost: Same Day Shipping

As discussed in the body of the paper, the lack of a convenience effect could be because our data period does not include significant changes to shipping times. However, anecdotal evidence and Figure 2 indicate that Amazon started to place FCs in locations close to large markets in anticipation of introducing same-day delivery. In ignoring this effect of FC expansion, we may be incorrectly attributing an expected boost in demand due to same-day shipping to cost savings from being close consumers. Intuitively, this would lead to an overestimate of the shipping cost parameter.

In order to address this issue, we add a preference for same-day shipping after 2014 for counties which have access to this service and then re-estimate the supply model. In order to determine the counties which would have same-day delivery, we use a tool on Amazon.com to find the markets which have same day shipping as as of the Fall of 2015.[39] This information allows us to construct a measure of the radius from a FC in which local customers have access to same day delivery, which is about 50 miles. We then add the term $\gamma^{same}\mathbf{1}_{ijt}^{same}$ to ζ_{ijt}, where $\mathbf{1}_{ijt}^{same}$ is equal to one for counties within 50 miles of a FC in year $t \geq 2014$. Because we cannot estimate γ^{same} directly, we choose reasonable values for it and check if/how this changes the results. We choose $\gamma^{same} \in \{0.10, 0.20\}$, which increase the expenditures on Amazon due to same-day shipping by approximately 10 and

[39] `http://www.amazon.com/b?node=8729023011`.

20%, respectively. These values are in line with the results in Table 6, which suggest that the effect of 'Local-Express Delivery', the early version of Same-Day delivery, is 0.186.

Results from these robustness estimates under the no lag assumption are in rows four and five of Table C.9. We focus on the first column, where the instruments are the indicator variables based on the first stage regression. The first three rows display the baseline estimates, where the direct comparison is to the second row. The downward shift in the confidence set is due to the fact that FC entry now results in both a demand and a supply side effect. That is, the loss in revenue due to the tax effect and an increase in fixed costs is offset by both a reduction in shipping cost and an increase in demand due to same-day shipping.

However, the fact that the confidence sets do not include zero provides reassurance that even if there exists a same day shipping effect, there is still significant cost savings associated with expansion. Therefore, it is likely that economies of density are a key component in determining FC locations.

Cost: Logistics Model

In the primary estimates, we assume that the closest fulfillment center to county i handles all of the shipping to that county. This may be incorrect for a few reasons. First, the lowest cost FC may not be the same as the closest FC. For example, it may be cheaper to ship items to Minneapolis from the FC in Indiana than from the FC in Wisconsin. Presumably Amazon would take this into account when making their distribution decisions. Second, there may be capacity constraints implying that not all shipments can come from a small local FC. Third, it may be the case that the universe of products is not available across all FCs. This last concern is somewhat mitigated by the fact that FCs of different types are often clustered.

In order to address this, we estimate the supply side of the model considering that the shipment of products to county i may be distributed across the FCs in year t and assume that this distribution is a function of both the cost of shipping between the FC and county i and the capacity of the FC relative to the total capacity. The relative capacity captures the fact that a single FC may not be able to satisfy all the local demand and the fact that not all types of products will be available at all FCs. The distribution function determines the percentage of expenditures by county i in year t that are shipped from fulfillment center f:

$$\omega_{ift} = \frac{\exp(-\beta_1 C_{if} + \beta_2 cap_f)}{\sum_{g \in F_t} \exp(-\beta_1 cost_{ig} + \beta_2 cap_g)}$$

where C_{if} is the estimated cost of shipping an item between county i and FC f, cap_f is the capacity of FC f in millions of square feet, and F_t is the set of FCs that are open in year t. The method used to calculate C_{if} is described in Appendix A.

Table C.8: Shipping Distances Implied by Alternative Logistics Models

$\{\beta_1, \beta_2\}$	[Mean Shipping Distance (StDev), Mean ω for Closest FC] 2006	2010	2014	2018
Closest	[315 (224), 1.00]	[287 (221), 1.00]	[240 (225), 1.00]	[200 (207), 1.00]
$\{10, 0\}$	[366 (231), 0.06]	[357 (228), 0.05]	[352 (233), 0.03]	[338 (219), 0.02]
$\{10, 20\}$	[443 (235), 0.04]	[457 (245), 0.03]	[436 (304), 0.02]	[405 (268), 0.01]
$\{20, 0\}$	[335 (228), 0.07]	[318 (225), 0.07]	[290 (227), 0.05]	[271 (214), 0.04]
$\{20, 20\}$	[397 (234), 0.05]	[400 (238), 0.05]	[347 (259), 0.03]	[311 (232), 0.02]
$\{200, 0\}$	[320 (226), 0.84]	[295 (221), 0.84]	[252 (224), 0.74]	[241 (222), 0.52]
$\{200, 20\}$	[324 (226), 0.78]	[295 (220), 0.84]	[255 (222), 0.70]	[237 (216), 0.42]

Notes: The table summarizes shipping distances under different assumptions about Amazon's choice of fulfillment center to serve each county. For county i, the share of revenue satisfied by fulfillment center f is $\omega_{ift} = \exp(-\beta_1 cost_{if} + \beta_2 cap_f) \,/\, \sum_{g \in F_t} \exp(-\beta_1 cost_{ig} + \beta_2 cap_g)$. Column (1), labeled $\{\beta_1, \beta_2\}$, lists the weight placed on the cost of shipping from from i to f, $cost_{if}$, and the capacity of f, cap_f. Each entry in the table contains the average and standard deviation in shipping distance across US counties under the particular choice of β_1 and β_2, together with the average share of revenue satisfied by the closest fulfillment center for each logistic model.

Given this distribution, Equation (10) becomes:

$$\Pi(a;\theta) = \sum_{t=2006}^{\infty} \beta^{t-2006} \sum_i \mu R_{it}(a_t) - \left(\sum_{f \in F_t} \theta d_{ift}(a_t) \times \omega_{ift} \times R_{it}(a_t) \right) - F_{it} a_{ti} - S_t(a_{it} - a_{it-1}) \tag{C.1}$$

where d_{ift} is the distance between FC f and county i. Here, the total shipping distance is equivalent to a weighted average of the distance from the FCs in the network, where the weights are based on the percentage of goods coming from each FC.

Again, because we cannot estimate this distribution directly, we choose reasonable values of β_1 and β_2 and check if/how this affects the results. In practice, we choose three different values for β_1 and two different values for β_2, resulting in six additional specifications. Table C.8 displays the summary statistics for the shipping distances and the ωs for the closes FC under each combination of the parameters. The values for the base model (closest FC) are in the first row for comparison. A low value of β_1 puts less emphasis on the shipping cost from each FC, which results in shipping being more uniformly distributed around the country. This increases the average shipping distance and the amount of goods being shipped from the closest FC. As β_1 increases, deliveries get more concentrated to the lowest cost FC, which lowers the shipping distance and increases the amount of good being shipped from the closest FC. Note that when $\beta_1 = 200$ and $\beta_2 = 0$, the average distance is close to that of the base model and the majority of shipments are coming from the closest FC.

Increasing the value of β_2 increases the amount of goods that are going to come from the biggest

FCs. Note that a value of $\beta_2 = 20$ and $\beta_1 = 0$ results in the distribution of goods shipped from FC f is approximately equal to FC f's capacity relative to total capacity. Therefore, we can think of the model when $\beta_2 = 0$ as a model where capacity does not play a role and the model when $\beta_2 = 20$ as when the amount of goods shipped from FC f is proportional to its capacity. We can see that increasing β_2 increase the average shipping distance and decreases the amount of goods coming from the closest FC.

Overall, under any specification, it is clear that expansion of the network leads to a decrease in the shipping distance to the consumers, implying that there is still a tradeoff between the tax and shipping effects of entering a new state. However, it is not clear how these alternative assumptions will affect the estimates of the supply side model: it depends on the relationship between the change in shipping distance, fixed costs, and revenue as a result of specific FC entry, which may vary across these different models.

Results from these models appear in the last 6 rows of C.9. Again focusing on column 1, we see that the interval shifts downwards, which is likely because there is not as strong as a connection between the tax implications and the cost savings of FC entry. However, because the results do not vary significantly from the base model, we believe that the simple model of logistics is a good approximation of the true data generating process.

Cost: Continuous Instruments

In our base model, the instruments are defined base on whether or not perturbations fall into experiment 1 or experiment 2, implying that there are a number of perturbations that receive 0 weight and, thus, are not used. Therefore, in the second column of Table C.9 we present results for our primary models along with the robustness models when including the continuous measure of distance as an instrument. In these specifications, all 1,575 perturbations have weight in the objective function. The results indicated by a star are ones in which the method of Andrews and Soares (2010) results in an empty set. Therefore, we display the value of θ which minimizes the objective function.

The confidence sets shift down because there is now positive weight on perturbations which do not fall into experiments 1 or 2. However, even under the most conservative of estimates, there is still a positive level of cost savings, although they are quite small.

Table C.9: Shipping Cost Estimates: Robustness to Model Assumptions

Model	Confidence Set $\hat{z}_a$	Confidence Set $\hat{z}_a, \hat{x}_a^+$
Base Specification		
Closest FC: No First Stage	[3.43E-05,1.04E-04]	[2.34E-05,7.05E-05]
Closest FC: First Stage	[5.45E-05,2.26E-04]	[4.65E-05,1.02E-04]
Lowest Cost FC: First Stage	[3.90E-05,1.20E-04]	[4.29E-05,4.29E-05]*
Introduction of Same-Day Shipping Demand Effect		
Small Same-Day Effect ($\beta = 0.10$)	[1.06E-05,1.38E-04]	[1.35E-08,6.15E-05]
Large Same-Day Effect ($\beta = 0.20$)	[2.28E-05,1.28E-04]	[1.36E-08,5.42E-05]
Alternative Logistics Models		
Cost: low, capacity: no $\{10, 0\}$	[5.73E-05,9.99E-05]	[4.46E-05,4.46E-05]*
Cost: low, capacity: yes $\{10, 20\}$	[2.50E-05,1.47E-04]	[2.15E-05,1.87E-04]
Cost: medium, capacity: no $\{20, 0\}$	[4.08E-05,1.52E-04]	[3.49E-05,3.49E-05]*
Cost: medium, capacity: yes $\{20, 20\}$	[3.28E-05,9.53E-05]	[2.96E-05,7.39E-05]
Cost: high, capacity: no $\{200, 0\}$	[3.92E-05,1.14E-04]	[4.25E-05,4.25E-05]*
Cost: high, capacity: yes $\{200, 20\}$	[4.35E-05,1.08E-04]	[4.53E-05,4.53E-05]*

Notes: Displayed are the estimated 95% confidence sets under alternative supply and demand side and instrumenting assumptions. A * indicates that the methods of Andrews and Soares (2010) resulted in an empty confidence set. Instead we display the value of θ that minimizes the objective function. The first column uses indicators for informative perturbations as instruments, as in Specifications (1) in Table 10. The second column adds continuous instruments. The same-day specifications assume that a county is served by its closest fulfillment center; the alternative logistics model specifications assign a revenue share to each center based on cost and capacity considerations, described above as [weight on cost, inclusion of capacity effect, $\{\beta_1, \beta_2\}$]. See notes to Table C.8. Same-day shipping and alternative logistics model estimates use informative perturbations predicted from exogenous contributions to variable profit and shipping cost deviations as instruments.