# Sales Tax, E-commerce, and Amazon's Fulfillment Center Network*

Jean-Francois Houde
*The Wharton School*
*University of Pennsylvania*

Peter Newberry
*Department of Economics*
*The Pennsylvania State University*

Katja Seim
*The Wharton School*
*University of Pennsylvania*

February 22, 2016

**Abstract**

We estimate the cost savings associated with the expansion of Amazon's fulfillment center network from 2006-2018. We first demonstrate that, in placing a fulfillment center in a new state, Amazon faces a trade-off between the revenue considerations from exposing local customers to sales tax 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 being charged sales tax. We then use the demand estimates and the spatial distribution of demand relative to Amazon's fulfillment centers to produce predicted 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 associated with being closer to customers that render the observed network roll-out optimal.

**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@wharton.upenn.edu, pwnewberry@psu.edu, or kseim@wharton.upenn.edu.

# 1 Introduction

Online retail has grown substantially over the last decade and now makes up nearly 7% of all retail sales.[1] Amazon.com is a key contributor of this growth, with net sales increasing from $8.5 billion in 2005 to $89 billion in 2014.[2] At the same time, Amazon's rise has been accompanied by increasing concentration in online retail, despite initial expectations that the internet would facilitate highly competitive markets where buyers have the option to compare prices and buy from many different platforms and sellers. As the online retail market matures, this no longer appears to be the case; the market today features a few powerful firms. Figure 1 uses data from ComScore to illustrate this trend, focusing on product categories where Amazon.com competes. Over the period 2006-2013, the Herfindahl-Hirshman-index increased seven to ten-fold, depending on whether market shares are based on consumer spending or on transaction counts. In this paper, we explore one possible explanation for this increasing concentration: the competitive advantage that comes with cost savings of a large-scale distribution network.

We investigate the role of economies of scale in distribution using Amazon as a case study. As Amazon's revenue has grown, so has its network of distribution centers (called Fulfillment Centers, denoted as FCs going forward). The number of FCs has grown from 8 centrally located centers in 2006 to 46 FCs spread out across the US by the end of 2015.[3] Amazon has plans to further expand its network to over 90 facilities by 2018, including "Amazon Now" hubs that provide same-day delivery to local customers. The main goal of this paper is to investigate the primary drivers of these expansion decisions.

A larger distribution network has both positive and negative implications for firm profitability. A distinction of online markets lies in the fact that an online retailer is only required to charge its customers sales tax if the firm has a physical presence (or "nexus"), such as a warehouse, in the customer's state of residence.[4] As earlier literature (Einav et al. (2014),Baugh et al. (2014))

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

[2] `www.statista.com/statistics/266282/annual-net-revenue-of-amazoncom/`.

[3] These figures do not include Amazon fresh FCs or sortation centers. See section 2 for details on distribution center types.

[4] In these states, consumers are required to file a use tax return for all of their online purchases. In most cases, consumers do not do this.

shows, demand is sensitive to sales taxes as they raise the effective price of an online transaction. 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 a high-tax state. This incentive would be particularly high in populous states, where in addition, the fixed costs from operating a warehouse are large. Such costs of building a warehouse in a new state are potentially offset by two benefits from proximity to the consumer: first, the firm may be able to make faster deliveries; consumers could derive independent value from such convenience. By cutting the distance between the FC and the consumer, the firm may also be able to derive significant cost saving from shorter shipping routes.

Shipping comprises a significant share of Amazon's operating costs: the financial statements report the costs (net of shipping revenue) are between 3-5% of net sales. These costs totaled nearly $4.3 billion in 2014.[5] We posit that by expanding its network of FCs, Amazon reduces its shipping cost through one of two ways. First, for deliveries that are fulfilled by one of Amazon's contracted shippers (UPS, FedEx, US Postal Service), additional FCs reduce the initial distance a package travels from Amazon's FC to the shippers' sorting center, even if the delivery distance from the sorting center to the customer is unaffected. This reduces the shippers' cost, providing Amazon with bargaining power in negotiations with shippers over the per-package rates they charge Amazon. Second, the select placement of FCs near population hubs has in recent years also allowed Amazon to enter the last-leg delivery market itself, reducing its reliance on the shippers. Both of these mechanisms imply that the amount of cost saving is related to the distance a package must travel from the FC to the customer. Our goal is to estimate this relationship.

We quantify empirically the trade-off between higher tax-inclusive prices and fixed cost, on the one hand, and additional convenience and shipping cost savings on the other. To accomplish this, we use detailed information on internet purchase behavior at Amazon and other online competitors by a large representative consumer sample, combined with detailed information on the location and characteristics of Amazon's FCs over time. We combine these data with a revealed preference approach similar to that of Holmes (2011). Specifically, we use the fact that entrance of a FC

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

into some states results in lost revenue due to new sales tax liabilities. Therefore, assuming profit maximizing behavior on the part of Amazon implies that entry into a state must also result in either increased willingness-to-pay due to faster deliveries or a reduction of shipping costs. This allows us to form a measure of cost saving by relating the combined change in revenue from tax and convenience effects to the reduction in shipping distance through moment inequalities.

The first step in this approach is to estimate a model of shopping mode demand in order to pin down consumer tax sensitivity and response to shipping times. We specify a CES demand model in which 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., Wal-Mart). Importantly, the amount of sales tax charged to the consumer varies both across modes and across time. While some of this variation is a result of the expansion of the FC network, there is also rich variation in tax rates across counties within a state and within counties across time. Therefore, the covariation in the sales tax and expenditures across shopping modes and across time identifies consumers' sensitivity to sales tax.

We estimate the demand model using household level data from the comScore Web Behavior Database, which includes details about the online transactions of between 30,000 and 50,000 households from 2006-2013. We supplement these with survey data from Forrester Research, which provide us with information on demographics of consumers choosing not to buy anything online and with Census data on aggregate retail spending. Using these sources, we calculate the expenditures for a representative consumer in each county across the four modes of shopping in each year from 2006 to 2013. In addition to these expenditures, we observe the sales tax rate for each county and year, and are able to account for county-level consumer demographics and county-level measures of offline retail competition via data from the census. This model is very much in the same spirit of the one found in Einav et al. (2014).

The results of the demand estimation suggest that consumers are sensitive to taxes and that their tax elasticity is around -1.4. This is similar to the estimates from Einav et al. (2014), -1.8, and Baugh et al. (2014), -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 on Amazon of around 9%. The reduction is even more pronounced for high tax states such as New York (12% reduction in sales) and Illinois (11.5% reduction in sales).[6] These examples are especially pertinent because both states have witnessed a tax regime shift in the recent years.[7] In general, the results suggest that the tax effect must be taken into consideration when Amazon, or other online firms, choose where and when to build their fulfillment center network.

We find little evidence that the entrance of a FC leads to an increase in demand for Amazon from increased convenience. The lack of effect of shipping time on demand is likely due to the fact that Amazon's shipping options and prices did not drastically change between 2006-2013. That is, if a customer wanted a package to arrive in two days, they could receive it in two days no matter where they live in relation to a FC.[8]

These estimates imply that there are overall negative revenue effects of opening a fulfillment center in a taxed state. To quantify this effect, we compare Amazon's revenue from 2007 to 2018 under the observed fulfillment center network and tax laws to the revenue if the network and tax laws had remained as they were in 2006. We use our model and the planned locations of future FCs in order to predict revenue for Amazon from 2015 to 2018. We find that expansion has lead to a loss in revenue of nearly $9 billion, or 1.6% of total revenue.

We further explore the tradeoff between the revenue effects of charging sales tax to customers in highly populated states and the shipping to these customers. For example, Amazon's FCs in 2006 were an average of 252 miles from the closest highly populated metropolitan area, with 7 out their 8 FCs being in a different state than this population hub. If they would have separately moved each FC to the population hub, they would have lost an average of $138 in revenue due to the effect

[6]This is calculated using the average sales tax rate across counties in these states as of January 1, 2014. See `http://taxfoundation.org/article/state-and-local-sales-tax-rates-2014`.

[7]In 2008, New York state passed a law that requires Amazon to charge sales tax to its customers, even though Amazon did not have a 'nexus' in the state. The above results suggest that this had an large impact on sales in New York post 2008. Additionally, Amazon has had a long battle with Illinois regarding sales tax, but finally agreed to charge its customers sales tax in the beginning of 2015. Again, the results suggest that Amazon's sales are negatively affected by this shift. However, Amazon has announced plans to build distribution facilities in Illinois to serve the greater Chicago area. See `http://www.chicagotribune.com/business/breaking/ct-amazon-sales-tax-illinois-0124-biz-2-20150123-story.html#page=1` for details.

[8]Of course this has changed with the arrival of same-day shipping. We plan to estimate this effect in future iterations of the paper by observing the entry of FCs on the outskirts of large metropolitan areas.

of charging sales tax.

Next, we quantify the shipping cost savings from the larger FC network size that offset these revenue declines by specifying a profit function where the variable cost of providing goods to consumers in a county is a function the shipping distance from the FC to that county. We assume that the observed network placement is optimal so that the profit under this structure must be at least as large as under any perturbation of the network. A perturbed network configuration affects revenue through the potentially differing sales taxes, fixed costs through the wages and rents paid in the counties where Amazon FCs are located, and variable costs through changes in the total shipping distance across customers. We compute perturbations similar to that of Holmes (2011): we swap the observed opening dates of two FCs.

The set of perturbations allows us to form moment inequalities from which we can estimate the effect of distance on variable cost using the methods in Pakes, Porter, Ho, and Ishii (Pakes et al.). Preliminary estimates suggest that Amazon saves about 0.7% of revenue per 100 miles of shipping distance. Overall, this translates into average shipping cost of somewhere in the range of 1-2% of sales, depending on year, in line with the above figures from Amazon's financial reports. Relative to holding its network size fixed at the 2006 configuration, our estimates suggest that over 2007-2018, Amazon saves over $5.6 billion in shipping costs through its FC network expansion, evidence of significant economies of scale in distribution in online markets. Accounting for the above associated revenue decline due to tax effects, the network expansion results in an overall profit gain of over $3 billion.

This paper is related to several strands of the literature. First, there are a number of papers that examine the response of consumers to sales taxes in offline markets. Agarwal et al. (2013) and Asplund et al. (2007) demonstrate that people substitute their shopping dollars across state and country borders when there exists differences in sales tax rates, while Agarwal et al. (2013) finds that consumers respond to 'tax-holidays' by increasing their spending. Chetty et al. (2009) estimates how the saliency of tax rates affects consumer demand and finds that consumers reduce their purchasing when the taxes are included in the posted price.

There is a large and growing literature examining the effect of sales tax on online purchasing. In

early papers, Goolsbee (2000a, 2000b) used variation in tax rates across municipalities and found that consumers were highly sensitive to sales tax. 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, but do so focusing on a single product category.

Two papers that are similar to the demand estimation in the current paper in many respects are Einav et al. (2014) and Baugh et al. (2014). Einav et al. (2014) estimate the 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 a consumer's sensitivity to distance because they observe the location of both buyers and sellers. One limitation of this paper is that they use only one online outlet, eBay, to estimate tax-sensitivity. That is, they do not consider that eBay buyers can substitute to other online stores, which may or may not be taxed. Baugh et al. (2014) uses a differences-in-differences approach to estimate the effect of the 'Amazon tax', or changes in the laws that forced Amazon to charge sales tax in various states during 2013-2015. Similar to Einav et al. (2014), they do not consider substitution to other taxed or non-taxed online outlets. We thus contribute to this literature by expanding our analysis of the tax sensitivity beyond a single product category and/or one online firm to a large number of product categories and online firms.

A second contribution to demand-side studies of the implications of Amazon's FC network choices lies in our analysis of possible convenience effects associated with a broader distribution network that cuts down on delivery times, which consumers may value independently of possible tax implications that come with Amazon's FC entry into a state. Our detailed information on the distance and delivery time between the closest FC to each consumer and their location allows us to investigate whether expenditures on Amazon respond to shorter shipping distances and faster delivery speeds, in addition to the effect of taxes. Our finding that the latter far dominate any willingness to pay response due to increased delivery convenience rules out such demand-side benefits

from proximity as a consideration for Amazon's network formation.

Our analysis is also related to recent economics literature examining the interplay of brick-and-mortar retail store locations choices and the retailer's distribution network. Holmes (2011) estimates the savings in distribution costs associated with clustering stores near a fulfillment center, while Zheng (2014) uses the proximity of a rival's fulfillment center as an instrument for expected entry in the future. Apart from a small operations research literature that looks at the management of distribution networks for online firms (see Agatz et al. (2008) for an overview of this literature), little work to date has studied such classic industrial organization questions in the context of online markets.

The remainder of the paper is organized as follows. The next section discusses the background of the sales tax laws and Amazon's FC network. Section 3 introduces the main data sources and Section 4 presents the demand side analysis and results. Section 5 provides preliminary 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 a supply-chain consulting company called MWPVL, International.[9] 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 item 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). In the case of Amazon, other types of distribution centers include Amazon Fresh FCs, which supply Amazon's grocery delivery service, 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 after they have shipped from a FC, and 'Prime Now Hubs', which handle same-day delivery for the local markets. For the majority of the analysis, we

[9] `http://www.mwpvl.com/`.

focus on the three primary types of FCs and the 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 46 by the end of 2015 and has plans to expand further to over 90 FCs by 2018. Table 1 and figures 2 and 3 demonstrate this expansion. There are a few things to note to about the expansion until 2010. First, Amazon placed FCs in relatively low populated states that were close to highly populated areas. For example by 2010, Amazon had two FCs in Nevada, both of which were on the border close to California's major cities. Second, they also placed FCs in states with relatively low sales tax, but close to highly populated areas. For example, there were FCs in New Hampshire and Delaware, which are both close to major east coast cities and have zero sales tax. This reflects that sales tax rates are positively correlated with population (across states, correlation of between 0.35 and 0.4 across years), so that entry into a small state near a large state has 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 focused on states with a relatively low sales tax rates (e.g., Pennsylvania with an average tax rate of 6.14% in 2014, compared to 6.94% across the top 20 US states in terms of population). For comparison purposes, we include a map of Wal-Mart's distribution centers, which are spread across the country more evenly than Amazon's FCs, at least in the initial stages of Amazon's distribution network.

Given these patterns, it is clear that the early strategy of this expansion depended on sales tax laws that allow e-commerce firms to avoid charging their customers sales tax, which can be significant, reaching close to 10% in some states. In 1992, the United States Supreme court ruled that the Commerce Clause in the US Constitution "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." Essentially, the law states that an online retailer does not have to charge a consumer sales tax unless it has a physical presence, or a 'nexus', in that consumer's state of residence. It is the duty of consumers to file a 'use-tax' return every year, which includes purchases from out-of-state vendors. However, very few individuals actually comply with this rule.[10]

[10]Baugh et al. (2014) quote that 0.2% of people living in Rhode Island, 0.3% of people living in California and New Jersey, 7.9% of people living in Vermont, and 9.8% of people living in Maine report use taxes.

For most online firms, a physical presence would come in the form of an office headquarters or a fulfillment center, implying that these firms likely do not have to charge sales in many states. As the popularity of e-commerce has grown, policy makers have started to suggest that these laws may be giving online firms an unfair advantage over their brick-and-mortar competitors. Further, states are likely losing out on millions of dollars of tax-revenue (see Bruce et al. (2009)). Because of their success and growing market share, Amazon.com has become the focus of politicians' complaints about these sales tax laws.

In the late 2000s, states began to introduce legislation that involved expanding the definition of a nexus to include the existence of 'affiliates', or websites that allow retailers, such as Amazon, to advertise on their site. For example, if a blogger based in Illinois has a link to Amazon on her site, then Amazon must charge sales tax to all the residents of Illinois. Not surprisingly, Amazon and other big retailers responded to this by shutting down their affiliate programs in states that passed these laws.[11,12] This certainly suggests that Amazon wanted to avoid charging sales tax to its customers. They even say as much in their 2008 annual report[13]:

> *A successful assertion by one or more states or foreign countries that we should collect sales or other taxes on the sale of merchandise or services could result in substantial tax liabilities for past sales, decrease our ability to compete with traditional retailers, and otherwise harm our business.*

However, it is also clear that as Amazon revenue grew in scale, the network of FCs expanded, 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 with a tax rate of 9.46% in 2013 compared to a nationwide average of 6.62% . Finally, by 2018, Amazon plans to have operating FCs in Illinois, Georgia, Ohio and North Carolina. Table 1 displays the number of FCs, the numbers of states that have a FC and the number of taxed states by year. Again, we see

[11] http://techcrunch.com/2011/06/10/amazon-shuts-down-associates-affiliate-program-in-connecticut-over-online-sales-tax/.

[12] http://www.kansascity.com/news/local/article325412/Amazon-shuts-down-Missouri-associates-program-over-sales-tax-dispute.html.

[13] http://media.corporate-ir.net/media_files/irol/97/97664/2007AR.pdf.

that as the network expands, Amazon is entering new states and being taxed in a greater number of states. Overall, the pattern of expansion, both in location and in the dynamics, 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 entry of a fulfillment center does not necessarily mean that Amazon charges sales tax immediately. For example, Amazon first built a FC in Pennsylvania in 2006, but did not begin to charge sales tax until 2011. This 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 they do 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 entry (e.g., Connecticut). The latter explain the, at times, significant discrepancy between the number of states where Amazon's customers pay sales tax and the number of states where Amazon has a FC. Given this, we make various assumptions on how Amazon perceives the relationship between entry and sales tax when making its network location decisions.

# 3 Data

### Consumer Online 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 sample of between 50,000 and 100,000 internet users (households) per year. The users give comScore explicit permission to monitor their activity. There are two primary databases, one that records each browsing session regardless of whether or not a purchase was made, and another that records transactions. Because our focus is on buying behavior, we focus on the latter.[14]

For each transaction, we observe a unique household 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', which is the total price of the transaction including

[14] ComScore is an internet analytics firm that provides data to Fortune 500 companies and large media organizations. See De los Santos et al. (2012) for deeper discussion of comScore's services.

shipping and taxes. In addition, we observe demographic characteristics for each household such as income, age of head of household, and racial background. Table 2 displays information about the reach of the sample in each year. The sample has gradually shrunk over the years from 86,000 households in 2006 to just 46,000 in 2013. Additionally, at least 75% of US counties and every US state plus the District of Columbia (minus Hawaii and Alaska) are represented in the data each year.

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. All of these facts are likely because the sample is drawn from internet users, who are not perfectly representative of the US population. De los Santos et al. (2012) compare the sample of ComScore users they use (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. However, to account for the possibility that ComScore may be over or under sampling certain demographic groups, we adjust our just our data using sampling weights from the Census. We bin each ComScore household into categories based on income, age, and racial makeup, and calculate relative sampling weights based on the prevalence of each category in the ComScore data relative to its prevalence in the county in the 2010 Census.

Table 3 is a first look at the purchasing patterns in the comScore sample. The first and second columns display the yearly average online expenditures and transactions per household and the third column shows the percentage of households with zero transactions. Note that we have limited the data to only transactions in product categories that Amazon sells. Therefore, we omit purchases in categories such as travel and online dating. We further adjust yearly expenditures for the number of weeks we observe the household in browsing data.[15]

Interestingly, we see decreasing average expenditures and transactions over time. Column 3 indicates that this is mostly due to an increasing percentage of households with zero online pur-

[15]We calculate the average weekly expenditures over the number of weeks we observe the household with any browsing activity and then multiply this by 52 weeks.

chases. This is not in line with anecdotal evidence about the take-up of online shopping.[16] It is possible that the sampling procedure is the cause for this, but a more plausible explanation is that comScore is not recording the transactions for some households. This could be due to the household deactivating the comScore behavior monitor or using a second computer (e.g., at work) or a mobile device for their online shopping. Because we do not know which households truly have zero expenditures, rather than using other devices for purchases, we choose to exclude them from the sample. In order to account for the extensive margin, we supplement the ComScore data with survey data from Forrester Research.

Forrester Research conducts annual surveys of the online shopping behavior of a representative sample of the US population, including whether or not the respondents made any online purchases in the last three months, with the survey being conducted between March and July, depending on the year.[17] In the final column on Table 3, we see that there is an increasing number of households shopping online, according to the survey. We match the Forrester and ComScore data based on demographic groupings, and calculate expected expenditures for a household in the ComScore sample that accounts for the propensity of not making any online purchases in their demographic category in the Forrester data. See Appendix A for details. Therefore, after adjusting the comScore data with Forrester's extensive margin, we see a pattern that is more in line with the anecdotal evidence: expenditures and transactions are both increasing over time. Figure 5 further illustrates Amazon's growing market share in expected expenditures over time, growing from approximately 10% in 2006 to 40% in 2013. Note that since we focus only on expenditures at retailers that compete in categories Amazon covers, this market share overstates Amazon's market share in online retail in total, which is estimated at approximately 20% in 2013.

A second limitation of the comScore data is that they do not record Amazon Marketplace purchases separately from Amazon purchases. A worry may be that consumers find a tax-free seller on Amazon Marketplace in a state in which Amazon itself must charge sales tax. This implies that we may underestimate the effect of sales tax because we would attribute a lack of

[16] `http://www.statista.com/statistics/183755/number-of-us-internet-shoppers-since-2009/`.

[17] Note that the survey was only available for 2006-2007 and 2010-2013. We interpolate linearly for the intervening years to construct predicted propensities of buying online.

response in states where Amazon charges taxes to a low sensitivity, when it is actually coming from consumers purchasing from a non-taxed vendor on Amazon Marketplace. However, most laws dictate that even marketplace vendors have to charge sales tax to customers in the states where their goods are housed.[18] To the extent that the vendors employ the 'Fulfilled by Amazon' program, then it is reasonable to assume that the sales tax implications of the network are the same for these vendors and Amazon. We perform a robustness check of the demand model that accounts for this possibility.

### Amazon Fulfillment Center Network

We acquire information about each of the FCs from MWPVL, International. From the opening date and the location of the FC, we calculate the straight-line distance between each FC's street address and the population weighted centroid of every county in the US, and take the minimum of this by county and year. The distance serves as a measure of how far Amazon must ship the good between its FCs and the consumers. It is apparent from Figures 2 and 3 that Amazon often opens FCs that are close to an existing FC, which may be to increase capacity or to have FCs of different types near one another. Because of this, we define a FC cluster as a group of FCs that 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 the observed opening date and shipping distance are associated with the first FC built in the cluster. Note that this does not affect the demand analysis because the tax implications are associated with the first FC built with or without clustering.

Because straight-line distance may not be a perfect measure of shipping speeds, we also obtain the US Postal Service's shipping times between each three digit US zip code and find the minimum shipping time between each county and each FC three-digit zip code.[19] The shipping times are 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. For a detailed description, see usps.com.[20]

[18] `http://www.avalara.com/learn/whitepapers/fba-sellers-guide-sales-tax/`.

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

[20] `http://pe.usps.com/businessmail101/classes/welcome.htm`.

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 USPS.

MWPVL also provides information about the size of every FC in square feet and the number of employees for about half of them. To fill in the missing information about the number of employees, 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 FCs which are open in a given year. There is a general pattern of an increase in size 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.

We also obtain information on the extent to which a given fulfillment center can be used to satisfy same day shipping orders. There are two sources for this information. First, we found the date of implementation for Amazon's early version of same day shipping, called 'Local Express Delivery', from various news sources. Amazon began this service, for which all customers are charged a fee, as early as October, 2009. Second, using a tool on Amazon.com, we find the markets which have free same day shipping available to Prime users as of the Fall of 2015.[21] This information allows us to construct a measure of the radius from a FC in which local customers have access to same day delivery. In future work, we plan to use this data in estimating the demand-side preference for same-day shipping as a shifter of county-level expenditures at Amazon.

### Other Sources of Data

In addition to the above sources, we obtained 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 local taxes can vary within a county and may change mid year. Table 4 displays the average and standard deviation of the tax rate across counties, as well as the average standard deviation of tax rates across counties within a state. The table demonstrates that there is not only variation in tax rates across states, but also substantial variation in rates across counties within a state. In

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

addition, tax rates vary across time, as between 30 and 65% of counties experienced tax changes each year and nearly 78% of the counties experienced at least one tax change over our sample period.

We collected county level wage information from the BLS's Quarterly Census of Employment and Wages and used the annual wage of a 'Retail Trade' employee. While this is not the exact industry in which the Amazon FC employees work, we believe the wages are likely to be similar in a given county. This is evidenced by the fact that the overall average hourly wage of a 'Retail Trade' employee is close to that of an Amazon Associate as reported by glassdoor.com. The land rents were computed by converting the average residential property value in a county to a yearly payment. The property values were obtained from the United State census via `http://metrocosm.com/get-the-data/`. Finally, we collected county level demographics from the 2010 Census, along with information on the number of small and large retailers from the 2012 County Business Patterns.

Table 5 shows the average of demographics for counties where Amazon has an FC and the averages across the US. Overall, there is no clear pattern in which Amazon is placing FCs in counties with lower wages, property values, or population density. This may be due to the location of these counties in relationship to large markets where Amazon is shipping its products. Additionally, the large increase in population density and properties values after 2014 comes from the fact that Amazon opened or is opening facilities close to large metropolitan areas to facilitate same-day delivery.

Finally, we use Amazon's annual reports to provide additional information about Amazon's finances. That is, for each year, we obtain the aggregate sales in relevant categories, the cost of goods sold, which includes inbound and outbound shipping costs, and the aggregate outbound shipping cost.[22] This allows us to formulate the gross profit margin for Amazon, net of outbound shipping cost, which we discuss in more detail in section 5.

[22] `http://phx.corporate-ir.net/phoenix.zhtml?c=97664&p=irol-reportsannual`.

# 4 Demand Analysis

We begin the demand analysis by providing reduced form evidence that consumers are sensitive to sales tax. First, we estimate the transaction level effect of taxes on the likelihood a consumer purchases from Amazon. That is, we run the following linear probability regression:

$$Pr(A_{hijt} = 1) = \beta_0 + \alpha \ln(1 + \tau_{it}) + \lambda_i + \lambda_j + \lambda_t + \epsilon_{hijt}$$

where each observation is a purchase occasion $h$, from a consumer in county $i$, in year $t$, for a product in category $j$. The dependent variable is a dummy variable equaling one if the consumer purchased the product from Amazon.com and zero otherwise. The tax rate is given by $\tau_{it}$; it varies depending on the county tax rate and whether or not Amazon has to charge sales tax to customers in that county in a given year. We do not include the item's price or the item's price at Amazon relative to the price charged by an outside seller because we observe only the price on the platform where the transaction occurs, but not on alternative platforms.[23] However, we include product category and time fixed effects to account for average price differences across online retailers at the category level and across time.

The result in the first column of table 6 indicates that consumer are sensitive to taxes and that they reduce their purchases from Amazon by nearly 16% for a 1% increase in sales tax. Note that we cannot equate this directly to a tax elasticity because the 16% reduction in Amazon's transactions may be substituted towards other taxed outlets (e.g., walmart.com or Wal-Mart).

In columns 2 through 7, we include various measures of shipping speeds to account for the fact that consumers may value faster shipping times from Amazon. The measures are the log of the distance between the centroid of the consumer's county and the closest FC and the shipping times from USPS for the four different classes of mail: priority, first, standard, and package. Since the variation in shipping days is limited, we aggregate shipping times into two categories, fast and slow shipping speed between two three-digit zip codes, excluding the slow shipping speed indicator from our specifications. We find that the tax effect is robust to including the variety of

[23] Nevertheless, including a variable that equals to Amazon's price for Amazon purchases and zero otherwise results in similar estimates.

proxies for shipping time. The log of the shipping distance is significant but has a positive affect on the propensity to purchase from Amazon. This is also the case for most mail classes, as the probability of purchasing from Amazon actually increases for customers in zip code locations with a slower shipping speeds, all else equal. The only parameter estimates that indicate that people significantly prefer faster shipping times is comparing fast priority shipping (1 or 2 day delivery time) to slow (3 day) priority shipping.

There are a number of explanations for the mixed 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 have tastes for faster shipping, or have very nonlinear preferences for shipping times given the recent expansion of Amazon to same-day shipping. Finally, it could be that the expansion of the network of fulfillment centers did not result in faster shipping times to consumers, or if it did, it was a small number of consumers who were affected. We believe this to be the most plausible explanation.

Next, we perform a differences-in-differences analysis of the response in Amazon expenditures to a change in the tax status of a county. We summarize a county's tax status in an indicator variable, $\mathbf{1}_{it}^{taxable}$, that is one if a purchase from Amazon.com by a household in county $i$ in year $t$ is subject to sales tax. 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 the initial agreement. The regression model is given by:

$$e_{it} = \beta_0 + \beta_1 \mathbf{1}_{it}^{taxable} + \lambda_t + \lambda_i + \epsilon_{it}$$

where $e_{it}$ is the log of aggregate expenditures on Amazon from county $i$ in year $t$. Aggregate expenditures are the sum of expenditures on Amazon in year $t$ for the ComScore households who live in county $i$. Results in table 7 suggest that a change in the tax status of a county results in somewhere between a 10 and 12% reduction in expenditures on Amazon, with this effect being significant at the 10% level in all specifications and at the 5% level for specifications (4) and (6). This exercise is essentially the same as the one performed in Baugh et al. (2014) with different data, and they estimate a tax effect of about 9.5%. 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.

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, we specify and estimate a model of demand for retail goods across all modes of online and offline shopping in the following sections.

## Empirical Model

We specify a model of demand where a consumer chooses how much money to spend on retail goods in 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. Specifically, a representative consumer from county $i$ solves the following problem in year $t$:

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

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

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 goods via shopping mode $j$, $\zeta_{ijt}$ is the taste for mode $j$, and $w_{it}$ is total dollars spent on retail goods. The elasticity of substitution between the four modes is given by $\sigma$. Solving for the optimal amount of expenditures in each online mode results in:

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

Dividing this by the expenditures for $j = 0$ and taking logs gives:

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

where the price of each mode can be written:

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

The tax rate in county $i$ and year $t$ is given by $\tau_{it}$, meaning $\tilde{p}_{ijt}$ is the tax-exclusive price of buying goods, which can include shipping charges. Note that $\tau$ may vary across $j$ as the non-taxed competitor never charges sales tax and Amazon does not charge sales tax in a number of states and years. We indicate whether mode $j$ entails a sales tax liability for the customer at time $t$ through the indicator $\mathbf{1}_{ijt}^{taxable}$; purchasing from an offline competitor always implies that the customer has to pay sales tax. 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 mode $j \in \{1, 2, 3\}$ as:

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

and assume that the preference for the offline shopping mode equals $\zeta_{i0t} = \exp(\epsilon_{i0t})^{\frac{1}{1-\sigma}}$. The term $\xi_{it}^{o}$ represents the preference for online shopping in county $i$ at time $t$. We therefore decompose $\xi_{it}^{o}$ into a county-level fixed taste for online shopping, relative to offline, and a time trend for online shopping that we allow to vary at the Census district level. That is, $\xi_{it}^{o} = \xi_{i}^{o} + \xi_{rt}^{o}$, where $r$ denotes the Census district, of which there are nine in the continental US. Similarly, $\xi_{jt}$ captures the time varying preference for shopping mode $j$; $Z_{it}$ is a vector of county level demographics; and $C_{it}$ is a vector of variables measuring the level of offline competition.

Notice that we allow the effects of demographics and competition to vary across shopping modes. For example, having a brick-and-mortar Best Buy in a county may affect the purchases through taxed online competitors (including bestbuy.com) differently than purchases from Amazon.

Motivated by our descriptive evidence above that suggested that distance to the closest FC – or shipping time – did not affect the propensity of purchasing from Amazon, we do not include it in the base model, but do so in various alternative specifications.

The demand shocks are at the county, time, and mode level, and are assumed to be iid across these delineations. Under these assumptions, equation 2 becomes:

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

We make the assumption that the prices for the online shopping modes do not vary across counties, or that $\tilde{p}_{ijt} = \tilde{p}_{jt}$, and that the price of the offline shopping option remains constant over time, or $\tilde{p}_{i0t} = \tilde{p}_{i0}$. While some online firms appear to price discriminate based on a consumer's location, there is limited evidence that this is widespread.[24] Further, Amazon attempted to implement price discrimination in 2000, and quickly abolished it after a backlash from customers.[25,26] Finally, we also assume that the price of outside good does not vary systematically over time. With this, we can re-write equation 3:

$$\begin{aligned}\underbrace{\ln(e_{ijt}) - \ln(e_{i0t})}_{\tilde{e}_{ijt}} =&(1-\sigma)\underbrace{(\ln(1+\tau_{it}\mathbf{1}_{ijt}^{taxable}) - \ln(1+\tau_{it}))}_{\tilde{\tau}_{ijt}} \\ &+ \beta_j Z_{it} + \lambda_j C_{it} + \xi_{rt}^o + \underbrace{\xi_{jt} + (1-\sigma)\ln(\tilde{p}_{jt})}_{\tilde{\xi}_{jt}} \\ &+ \underbrace{\xi_i^o - (1-\sigma)\ln(\tilde{p}_{i0})}_{a_i} + \underbrace{\epsilon_{ijt} - \epsilon_{i0t}}_{\tilde{\epsilon}_{ijt}}\end{aligned} \tag{4}$$

resulting in the following equation for the differences in expenditures on mode $j$ and the outside

[24] http://www.wsj.com/articles/SB10001424127887323777204578189391813881534.
[25] http://www.bizjournals.com/seattle/stories/2000/09/25/daily21.html.
[26] http://news.cnet.com/2100-1017-240700.html.

mode:

$$\tilde{e}_{ijt} = (1-\sigma)\tilde{\tau}_{ijt} + \lambda_j C_{it} + \beta_j Z_{it} + \xi^o_{rt} + \bar{\xi}_{jt} + a_i + \tilde{\epsilon}_{ijt} \tag{5}$$

This equation says that the difference in expenditures between mode $j$ and the offline mode is a function of the difference in sales tax between mode $j$ and the offline mode, a time varying mode level effect that includes the price level of mode $j$, the effect of demographics and local competition on mode $j$, a region-specific time trend, a county effect, and an iid demand shock.

### County Level Expenditures

To form the expenditures for the representative consumer ($\tilde{e}$), we calculate the population weighted yearly average expenditures on each shopping mode for each county. We define a taxed online competitor as one that has a large offline presence such as gap.com, walmart.com and target.com 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 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 8 displays the top ten online stores in each of the taxed and non-taxed categories. Overall, we classify about 34% of the websites that compete in product categories that Amazon carries as taxed competitors. We also exclude any expenditures in product categories that Amazon does not cover and exclude any websites that sell only in these categories. Examples of excluded sites are dating websites (e.g., Match.com), travel websites (e.g., orbtiz.com) and food delivery sites (e.g., dominos.com).

Under these restrictions, we calculate the yearly expenditures for the three online shopping modes for each individual household. The level of household expenditures is then adjusted using the Forrester Data as discussed above. The average expenditures for a county are calculated using demographic weights from the 2010 census. Finally, expenditure for the offline mode is calculated by subtracting online expenditures from the total amount of expenditures on retail goods, assuming

that total retail spending equals 10% of consumer's income.[27] See the appendix for details on how expenditures are calculated.

We examine patterns of spending across the shopping modes in Figure 5. The figure displays four different charts, each with the market share of the shopping modes on the left y-axis. Note that this is the market share of online purchases only. Amazon's market share has steadily increased from under 10% in 2006 to around 40% in 2013, while both the taxed competitors' and non-taxed competitors' market share has declined from around 45% to 30%. At the same time, there has been an increase in the number of fulfillment centers, the number of states with fulfillment centers and the number of states that are taxed. Finally, the average distance from the FC to the consumer has declined from around 300 miles to 150 in 2015 and is projected to further decrease to about 100 miles by 2018.

## Estimation

We estimate equation 5 using OLS with county, mode/year, and region/time level fixed effects. Included in $C_{it}$ is the total number of retail establishments and the number of establishments with more than 50 employees (i.e., large retailers) in county $i$ in year $t$ to account for the fact that large offline retailers (e.g., Wal-Mart) may have a different impact on online demand than small local retailers. County level demographic included in $Z_{it}$ are median income and the distribution of ethnicity.

Identification of the tax sensitivity parameter comes from several sources of variation. First, there is variation 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 $\sigma$. This is similar to the traditional identification argument in difference-in-differences models. Second, within a county and year, some modes are taxed and some are not. The extent to which people's expenditures vary across modes because of the sales tax charged also helps to identify tax sensitivity. Therefore, in order to estimate the

[27] Using data on the relative share of online retail out of total retail from `https://www.census.gov/retail/mrts/www/data/pdf/ec_current.pdf` in combination with total online spending in the comScore data, we calculate average total expenditures on retail goods to be approximately 10%.

tax sensitivity, we make the reasonable assumption that the changes in tax rates due to either Amazon's expansion and/or local laws are exogenous to the unobserved local demand shocks. This reflects that the Amazon FC entry decisions are likely made at a higher geographic level than the county and not made purely in response to a local annual demand shock in a specific county. 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 the above specification with county/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 since we observe only three modes and at most two tax levels across the 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, expansion of the network results in shorter routes and faster shipping times for some consumers. Further, because the tax effects of the expansion are at the state level, there are some consumers who live near state borders who see a decrease in delivery times, but not a change in tax rates. Note that we do not observe the locations and thus distance to Amazon's competitors' fulfillment centers, but we assume that the time-varying preferences for other modes, $\xi_{jt}$, 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 (e.g., 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 distances without variation in taxes and vice versa.

### Results

Results of 6 different specifications are presented in Table 9, where each specification varies by the measure of shipping speed included. Specification (1) does not include any measure, specifications (2) includes the log of the distance to the shortest FC and specifications (3)-(6) include variables for shipping speed in four different classes of delivery. In the latter, the longest delivery time is

the excluded dummy variable. The estimated value of $\sigma$, or tax sensitivity, is around 1.35 and is significant at the 1% level in all specifications. These estimates are similar to the ones reported by both Einav et al. (2014) and Baugh et al. (2014).

The estimates imply that going from not charging sales tax to charging the average sales tax of 6.5% would reduce expenditures on the given mode by around 9 percent. 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 there revenue to decrease by up to 9%, all else equal. In 2008, New York state passed a law that required Amazon to collect sales tax. Our estimates imply that this reduced Amazon's revenue in New York by around 12%.

Once again, we do not find 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 quality effect of expanding the network is approximately zero.

The estimates of the time varying mode fixed effects are presented in Table 10, with the excluded effect being mode 3 in 2006. Amazon's quality is increasing over time, which is not surprising considering the increase in market share over this time period. The taxed online competitors' quality increases at first and then decreases, while the quality of the non-taxed competitors is relatively constant. Finally, the estimates of the parameters on $C$ and $Z$ are presented in 11. This table shows that all types of firms are affected by the number of offline retailers and this effect becomes stronger as the number of large offline retailers increases. However, none of these effects are significant. Many of the demographic coefficients are not significant. Note that because we do not have the demographic variables varying over time, we exclude mode 3's effects.

Table 12 provide some robustness for these results. 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. Overall, the results of these robustness checks are consistent with the base results with the estimated tax sensitivity being between -1.27 and -1.88. This implies that, if anything, we are likely underestimating the tax sensitivity with our base estimates.

### Calculating Total Revenue

Using the estimates of the model, we calculate total revenue in the US for Amazon for the years 2006-2013 and also project revenues in 2014-2018. To do this, we first form predicted revenue for the representative household for every county and year that we observe in the comScore data. Then, for the counties that we never observe, we assume that county level fixed effect is equal to that of the closest county, which allows us to form a predicted amount expenditures on Amazon for the representative household for every county in the US from 2006-2013. We multiply the expenditures of the representative household by the number of households in the county and sum this over all counties for each year.

We compare this revenue to the equivalent sales data reported in Amazon's financial statements.[28] The predicted revenue is less than the revenue reported because, as discussed above, we are missing purchases from other computers or devices and missing purchases from people who are able to turn off the monitoring device. Therefore, we calculate a 'multiplier' of the reported sales divided by the predicted revenue and use it to create an adjusted expenditure for each household across all modes.[29] The implied multiplier is in line with reported shares of purchases that online shoppers make via a mobile device (24.6% in 2014) and the incidence of at-work-online shopping (47% in 2014).[30] The assumption here is that if we are missing purchases, we are missing them

[28] We use the figures of sales from North America in the "Media" and "Electronics and Other General Merchandise" categories. This excludes the "Other" category, which is revenue from "Non-Retail Activity" such as Amazon Web Services (AWS).

[29] The multipliers are 1.05, 1.44, 1.37, 1.40, 1.54, 1.50, 1.46, and 1.71 for the years 2006-2013. The increasing rate is likely due to the fact that mobile shopping has increased over this time period.

[30] See `https://www.internetretailer.com/2015/08/18/mobile-commerce-now-30-all-us-e-commerce` and `http://www.careerbuilder.com/share/aboutus/pressreleasesdetail.aspx?sd=12%2F1%2F2014&id=pr854&ed=12%2F31%2F2014`.

equally from all modes. We then re-run the model and perform the same procedure as above to form 'predicted' revenue, that now equals the sales in the financial report.[31]

To calculate revenues for 2014-2018, we need to obtain future values of the mode specific year effects and the region time effects. We do this by using the growth (or decline) of these effects from 2006-2013 to predict future effects.[32] Table 13 reports the total predicted revenue from 2006-2018. Because we observe Amazon sales in 2014 from their financial report, we can assess how well this procedure does in predicting sales in 2014. We find that we over predict revenue (52.8 billion versus 50.8 billion in the reports), but that the prediction is reasonably close. We also predict that expenditures on mode 2 fall from $28 billion to $23 billion and expenditures on mode 3 decline by approximately $6 billion over the next 4 years. Also displayed are the average shipping distance from the FCs to the customers, which decreases over time, and the number of counties that are taxed, which increases with the FC rollout into new states.

### Evidence of a Tax/Distance Trade Off

The demand estimates suggest that expanding the network has a negative effect on revenues through the tax effect. In order to quantify this effect, we perform a simulation exercise using the estimates of the base model. We calculate counterfactual revenue under a different network of FCs and different tax laws, and assume everything else remains constant. Specifically, we calculate the total expected revenue if the network of fulfillment centers had remained fixed since 2006.

The difference between the counterfactual revenue and observed revenue is presented the first column of Table 14 along with the difference between the average shipping distance. Results indicate that expansion leads to a reduction in expected revenue of around $8.9 billion, or a 1.5% reduction. However, this reduction in revenue is accompanied by a reduction in shipping distance as evidence by the 2nd column of Table 14. This gives a first glimpse at the trade-off that Amazon faces when expanding the network.

To further investigate this trade off, we perform a few additional exercises which can be found

[31] Re-running the model only shifts the mode level effects but does not change the estimates of tax sensitivity.

[32] Specifically, we run 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 estimate of the effect from the 2006-2013 data. Using the estimates of $\gamma$ we predict future values of $m_t$.

in table 15. First, we note that there is correlation between the population of a state and the sales tax rate, implying that moving an FC to a highly populated state results in lower revenue, but likely results in shorter shipping routes.

Second, we analyze the trade off of having an FC near a densely populated area in terms of distance and lost revenue. Specifically, for each FC, we find the county with the highest population density for which it is the closest FC (i.e., the FC which ships to that county). We then calculate the distance between the FC and that county along with the lost revenue if the FC operated in the state where the county is located. Table 16 provides some examples of these 'moves'. The tradeoff is clear for the most part: not being located in the highly populated counties results in an increase in revenue. However, there are some instances where moving closer to the population actually results in an increase in revenues. This is mainly due to the size of the county or the level taxes relative to the size(e.g., DC) and/or taxes (e.g., TN) of the state where the county is located. Figure 5 displays the entire collection of these moves and Table 15 summarizes them. Over time, more and more of FCs are being built near highly populated areas resulting in the effect of a move being 0: the FC moves within its own state. By 2018, over 50% of the FCs are in the same state as the closest highly populated area. The majority of the moves result in lower distance and lower revenue, with both average distance and the average revenue lost declining over time.

# 5 Quantifying Cost Savings from Distribution Network

We now turn to estimating the cost savings of expanding the network of FCs. We posit that these cost savings come the fact that, after expansion, the longest leg of shipping comes from the distributor to the FC rather than from the FC to the customer. This leg is not time sensitive in the sense that it does not occur after the customer has made their order and therefore, the distributor and/or Amazon can use the cheapest delivery option (e.g., semi-trailer truck) for the cross country trip rather than a plane. Additionally, this implies that Amazon relies less on the shipping companies which gives them more bargaining power when negotiating rates.

### 5.1 Profit Function

We formulate Amazon's profit as function of the network of FCs in year $t$, which we denote $a_t^0$, in a similar manner as Holmes (2011). In county $i$, Amazon generates $R_{it}(a_t^0)$ in revenue, which is a function of $a_t^0$ due to sales tax implications and/or the shipping speeds. Note that $R_{it}(a_t^0)$ can include the revenue from shipping charges. Variable profit in year $t$ from county $i$ is given by:

$$\pi_{it}(n_t;\theta) = \mu R_{it}(a_t^0) - \theta d_{it}(a_t^0) R_{it}(a_t^0)$$

where $\mu$ is the markup net of shipping costs and variable shipping cost is given by $\theta d_{it}(a_t^0)$. The parameter $\theta$ measure the shipping cost of per dollar of goods sold per mile and $d_{it}(a_t^0)$ is the distance from consumers in county $i$ to the closest FC. The biggest difference with Holmes (2011) is that we allow variable costs to vary across network location decisions. We assume that this variation comes from differences in the shipping distance from a FC to consumers in county $i$. Therefore, total profit for Amazon, starting in 2006 is given by:

$$\Pi(a^o;\theta) = \sum_{t=2006}^{\infty} \beta^t \sum_i \mu R_{it}(a_t^0) - \theta d_{it}(a_t^0) R_{it}(a_t^0) - F_t(a_t^0) \tag{6}$$

where $a^0$ represents the 'rollout' of the observed network and $\beta$ is the discount factor. The fixed costs of the observed network are given by $F(a_t^0)$. The fixed cost of the network are assumed to be made up of wages paid to employees of the FCs in county $i$, called 'Associates' and denoted $A_{it}$, and the land rents for the total square footage of space in county $i$, $L_{it}$:

$$F_t(a_t^0) = \sum_{i \in I(a_t^0)} w_{it} A_{it} + r_{it} L_{it}$$

The sum is over all the counties where the FCs are located under the observed network, $I(a_t^0)$.

We assume that Amazon chooses it's optimal prices independently of $a_t^0$, so the profit maximization problem, in terms of network formation, can be written as:

$$\max_a \Pi(a;\theta) \tag{7}$$

The fact that Amazon does not price discriminate across locations provides some evidence that prices are set independently of the network. Given the maximization problem in 7, we know that the profit under the observed network rollout ($a^o$) must be greater than any deviation from this ($a$), or:

$$\Pi(a^o;\theta) \geq \Pi(a;\theta) \forall a \neq a^0$$

This allows us to formulate a set of moment inequalities that are used to estimate $\theta$.

### 5.2 Moment Inequalities

As in Holmes, we construct moment inequalities by comparing the discounted profits of the observed sequence of openings with counter-factual roll-outs scenarios. We focus in particular on counter-factual roll-outs that *swap* the opening dates of two FCs. By swapping two opening dates (e.g. $t < t'$), we ensure that the difference in the discounted stream of profit cancels out beyond period $t'$. This allows us to impose optimality conditions without observing the complete sequence of openings.

Consider two opening sequences: observed sequence $a^0$, and counter-factual sequence $a$. If profits were measured without error, profit maximization imply the following linear profit difference inequality:

$$y_a - x_a\theta \geq 0 \tag{8}$$

where $y_a = \sum_{t=2006}^{\infty} \beta^t \left[\mu(R_t(a_t^0) - R_t(a_t)) - (F_t(a_t^0) - F_t(a_t))\right]$ is the discounted profit differential net of shipping costs, $x_a = \sum_{t=2006}^{\infty} \beta^t \sum_i \left[d_{it}(a_t^0)R_{it} - d_{it}(a_t)R_{it}\right]$ is the discounted differences in shipping distance, and $\theta$ is the shipping cost per miles.

The cost savings associated with the optimal roll-out of fulfillment centers are identified by revealed preference. Intuitively, the magnitude of $\theta$ is such that the predicted revenue loss relative to counter-factual sequence $a$ (i.e. $y_a < 0$) is offset by a reduction in shipping distance (i.e. $x_a < 0$). This gives a upper bound on the distance cost. Similarly, $\theta$ is such that a counter-factual sequence associated with an increase in shipping distance (i.e. $x_a > 0$) is offset by a revenue gain from tax avoidance (i.e. $y_a > 0$). This provides an lower bound on the distance cost.

Although this tradeoff is clear when comparing $y_a$ and $x_a$ for alternative paths, we measure Amazon's profits and shipping distance with an error. In particular, we observe $\tilde{y}_a$ and $\tilde{x}_a$:

$$\begin{aligned} \tilde{y}_a &= y_a + \eta_a, \\ \tilde{x}_a &= x_a + \nu_a. \end{aligned}$$

These errors originate from a variety of sources: (i) measurement error in the operating costs of FCs, (ii) estimation error in Amazon's revenue, (iii) omitted variables, and (iv) difference in tax implementation (i.e. Amazon's beliefs). We assume that Both errors are conditionally independent of a vector of non-negative instruments $z_a$:

$$E(\eta_a|z_a) = E(\nu_a|z_a) = 0. \tag{9}$$

Applying these two conditional moment restrictions to the "measured" inequality condition (analog from equation 8) gives:

$$E\left(\tilde{y}_a - \theta\tilde{x}_a|z_a\right) = E(y_a - \theta x_a|z_a) + \underbrace{E(\eta_a - \theta\nu_a|z_a)}_{=0} \geq 0 \tag{10}$$

Our estimator is based on a set of unconditional moment conditions consistent with the previous inequality:

$$E\left[z_a(\tilde{y}_a - \theta\tilde{x}_a)\right] \geq 0. \tag{11}$$

This leads to the following sample moments inequalities:

$$\frac{1}{M}\sum_a z_{a,k}(\tilde{y}_a - \theta^0\tilde{x}_a) = \tilde{m}_k(\theta^0) \geq 0, \quad \forall k = 1, \ldots, K. \tag{12}$$

where $M$ denotes the number of inequalities. The objective function therefore becomes:

$$\hat{\Theta} = \arg\min_\theta \sum_k \min\{0, \tilde{m}_k(\theta)\}^2 \tag{13}$$

Following Holmes, we break the choice of instruments into two parts: (i) basic moments, and (ii) interactions with state variables. The basic moments correspond to a group of indicator variables associated with various degrees of shipping distance increase or decrease. In Holmes, this corresponds to change in store density, which maps one-to-one in the cannibalization vs. economies of density tradeoff. In our case, an increase distance does not directly lead to tax saving. We therefore need to select inequalities based on both elements.

The additional complication relative to Holmes is that we cannot group inequalities based on predicted proximity changes (i.e. $\tilde{x}_a$), since those changes are measured with error. To get around this problem we need to construct an "exogenous" distance change variable that is uncorrelated with this error. Consider for instance a counter-factual rollout plan $a$:

$$\hat{x}_a = \sum_t \sum_i \beta^t \left[ d_i(a_t^0)\hat{R}_{it}(a_t^0) - d_i(a_t)\hat{R}_{it}(a_t) \right] \tag{14}$$

where $\hat{R}_{it}(a_t)$ is a predicted revenue in county $i$ estimated based on pre-determined variables (e.g. population, income, offline retail structure, density, etc.). This measure can be estimated in a "first-stage" for each counter-factual assignments. The "exogenous" net revenue differences, denoted by $\hat{y}_a$, can be obtained similarly. We use both measures to construct the instruments.

The role of the "basic" instruments is to define groups of counter-factual roll-outs that are related to the lower and upper bound cost-savings. We select two types of swapping experiments, reflecting the revenue/proximity tradeoff. For instance, observing Amazing choosing to enter relatively early into a high-tax state when a low-tax (but further away) option was available identifies an upper bound of the shipping cost saving parameter. To identify scenarios that identify this upper bound, we look for counter-factual roll-outs that *increase the proximity* of the network to larger markets and result in higher revenues. Similarly, counter-factual roll-outs that *decrease the proximity* of the network to large markets identify an lower bound on the cost savings. Within each category, we also group counter-factual scenarios that hold more or less constant the magnitude of the shipping distance change. We consider three subgroups, which gives us six basic instruments.

1. **Proximity increase experiment:** Let $t_j$ the chosen opening date of FC $j$.

- Let $S_1$ denote the set of FCs opened at dates $t_{j'} > t_j$, in states that would lead to higher net revenue $\hat{y}_{a(j,j')} > 0$
- Define $a(j, j')$ denote the counter-factual sequence that swaps $j$ with $j'$.
- Group 1: $j' \in S$ such that $\kappa_0 < \hat{x}_{a(j,j')} \leq \kappa_1$
- Group 2: $j' \in S$ such that $\kappa_1 < \hat{x}_{a(j,j')} \leq \kappa_2$
- Group 3: $j' \in S$ such that $\kappa_2 < \hat{x}_{a(j,j')}$

2. **Proximity decrease experiment:** Let $t_j$ the chosen opening date of FC $j$.

- Let $S_2$ denote the set of FCs opened a dates $t_{j'} > t_j$, in states that would lead to higher net revenue $\hat{y}_{a(j,j')} < 0$
- Define $a(j, j')$ denote the counter-factual sequence that swaps $j$ with $j'$.
- Group 1: $j' \in S$ such that $-\kappa_0 \geq \hat{x}_{a(j,j')} > -\kappa_1$
- Group 2: $j' \in S$ such that $-\kappa_1 \geq \hat{x}_{a(j,j')} > -\kappa_2$
- Group 3: $j' \in S$ such that $-\kappa_2 > \hat{x}_{a(j,j')}$

These groups can be written as a set of 6 dummy variables can constitute our "basic instruments". The rest of the instruments are interacting these instruments with $\hat{x}_{a(j,j')}$ and $\hat{x}^2_{a(j,j')}$.

The previous grouping uses indicator variables to construct the instruments. An alternative approach is to use smooth transformations of the indicator variables:

$$\text{Proximity increase:} \quad z^1_a = \Big\{(1-\Phi(\hat{y}_a/\sigma))(1-\Phi(\hat{x}_a/\sigma)), (1-\Phi(\hat{y}_a/\sigma))(1-\Phi(\hat{x}_a/\sigma)) \times \hat{x}^+_a,$$
$$(1-\Phi(\hat{y}_a/\sigma))(1-\Phi(\hat{x}_a/\sigma)) \times \hat{y}^+_a, (1-\Phi(\hat{y}_a/\sigma))(1-\Phi(\hat{x}_a/\sigma)) \times \hat{y}^+_a \hat{x}^+_a \Big\}$$
$$\text{Proximity decrease:} \quad z^2_a = \Big\{\Phi(\hat{y}_a/\sigma)\Phi(\hat{x}_a/\sigma), \Phi(\hat{y}_a/\sigma)\Phi(\hat{x}_a/\sigma) \times \hat{x}^+_a,$$
$$\Phi(\hat{y}_a/\sigma)\Phi(\hat{x}_a/\sigma) \times \hat{y}^+_a, \Phi(\hat{y}_a/\sigma)\Phi(\hat{x}_a/\sigma) \times \hat{y}^+_a \hat{x}^+_a \Big\}$$

The parameter $\sigma$ is a smoothing parameter. As $\sigma \to 0$, the product of the two CDFs is equivalent to create the dummy variables above. The advantage of using the smooth instruments is to use all available inequalities, rather than arbitrarily dropping those that do not satisfy the conditions.

### Preliminary Results

We set $\mu = 0.23$ and $\beta = 0.95$ and estimate $\theta$ for various definitions of $\hat{x}$ and $\hat{y}$. Preliminary results suggest that Amazon saves about 0.7% per dollar of revenue by reducing shipping distance by 100 miles, meaning shipping a package the average distance (250 miles) costs about \$1.75 for every \$100 in goods shipped. Overall, this implies that Amazon saved nearly \$5.6 billion in shipping costs by expanding the network from 2006 to 2018.

## 6 Conclusion

We study the trade-off Amazon faces when choosing the location of a new fulfillment center. Amazon benefits from their 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 they have a physical presence. By raising the tax-inclusive price the consumer faces, this reduces consumers' willingness-to-pay for Amazon's services. Being close to significant population clusters also raises the fixed cost of operating fulfillment centers.

We find that Amazon indeed faces this tradeoff: consumers dislike paying taxes, meaning there be must gains of network expansion due to faster shipping and/or reduced delivery costs. 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 actually 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 2006-2018 resulted in a loss in revenue of around \$8.9 billion dollars for Amazon, but at the same time reduced the average shipping distance from FC to consumer by around 178 miles by the end of 2018. We use a moment inequalities approach, together with the assumption that Amazon's network expansion path is optimal, to infer shipping cost savings from the observed fulfillment center network relative to alternative configurations.

## 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.

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.

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.

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., 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. Moment inequalities and their application. *Econometrica 83*(1), 315–334.

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.

# A Tables and Figures

## A.1 Tables

Table 1: Expansion of fulfillment center network

| Year | # of Centers | # States with with Fulfillment Center | # States with Sales Tax Liability | Ave Size of FCs (1,000 Sq ft) | Ave Employees of FCs (1,000) |
|---|---|---|---|---|---|
| 2006 | 8 | 6 | 4 | 541 | 4.3 |
| 2007 | 9 | 7 | 4 | 810 | 6.5 |
| 2008 | 12 | 10 | 5 | 685 | 5.5 |
| 2009 | 17 | 10 | 5 | 721 | 5.8 |
| 2010 | 17 | 10 | 5 | 721 | 5.8 |
| 2011 | 24 | 10 | 5 | 775 | 6.2 |
| 2012 | 42 | 12 | 8 | 788 | 6.3 |
| 2013 | 40 | 14 | 16 | 820 | 6.6 |
| 2014 | 47 | 14 | 23 | 748 | 6.0 |
| 2015 | 50 | 16 | 26 | 738 | 5.9 |
| 2016 | 89 | 25 | 27 | 611 | 4.9 |
| 2017 | 94 | 26 | 28 | 614 | 4.9 |
| 2018 | 97 | 26 | 28 | 620 | 4.9 |

Notes: The number of states where Amazon purchases are subject to sales tax exceeds the number of states with a FC due to states negotiating for sales taxes being collected immediately after agreeing to let Amazon build a FC in the state, even if there is a delay between the time of agreement and the actual opening of the warehouse.

Table 2: comScore Sample

| Year | Households | States (%) | Counties (%) |
|---|---|---|---|
| 2006 | 86,405 | 49 (100) | 2,872 (91) |
| 2007 | 89,412 | 49 (100) | 2,865 (91) |
| 2008 | 56,722 | 49 (100) | 2,725 (87) |
| 2009 | 55,585 | 49 (100) | 2,661 (85) |
| 2010 | 53,788 | 49 (100) | 2,621 (83) |
| 2011 | 62,849 | 49 (100) | 2,695 (86) |
| 2012 | 54,794 | 49 (100) | 2,623 (83) |
| 2013 | 46,216 | 49 (100) | 2,437 (78) |

Notes: .

Table 3: Household Purchasing

| Year | Online Expenditure | Online Transactions | % Zero Expenditure | Adjusted Expenditure | Adjusted Transactions | % Offline Shoppers Only |
|---|---|---|---|---|---|---|
| 2006 | $239 | 2.4 | 51.8% | $242 | 2.4 | 55.7% |
| 2007 | $254 | 2.5 | 52.0% | $242 | 2.4 | 60.8% |
| 2008 | $196 | 2.0 | 60.0% | $258 | 2.6 | - |
| 2009 | $141 | 1.4 | 67.9% | $268 | 2.7 | - |
| 2010 | $125 | 1.4 | 68.6% | $275 | 2.9 | 32.1% |
| 2011 | $131 | 1.4 | 69.7% | $327 | 3.4 | 23.0% |
| 2012 | $151 | 1.8 | 64.0% | $311 | 3.6 | 23.9% |
| 2013 | $120 | 1.7 | 65.2% | $294 | 4.0 | 15.5% |

Notes: The percent offline shoppers 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.

Table 4: Tax Rate Variation

| Year | Ave Tax Across All Counties | St. Dev Tax Across All Counties | Ave St. Dev Tax Across Counties within State | % of Counties with Tax Change |
|---|---|---|---|---|
| 2006 | 6.28% | 1.58% | 0.33% | 32% |
| 2007 | 6.30% | 1.59% | 0.33% | 28% |
| 2008 | 6.34% | 1.58% | 0.32% | 61% |
| 2009 | 6.33% | 1.55% | 0.31% | 65% |
| 2010 | 6.41% | 1.61% | 0.31% | 57% |
| 2011 | 6.44% | 1.60% | 0.33% | 52% |
| 2012 | 6.44% | 1.60% | 0.34% | 52% |
| 2013 | 6.47% | 1.60% | 0.34% | 33% |
| 2014 | 6.50% | 1.61% | 0.35% | 30% |

Notes: .

Table 5: Characteristics of Counties with FCs

| FCs Open by | Households (1,000) | Pop Density (per Sq Mile) | Median Income ($1,000) | Wage (annual) | Property Value ($ per Sq Ft) |
|---|---|---|---|---|---|
| 2006 | 189.1 | 494 | 51.5 | 23.2 | 9.2 |
| 2010 | 33.7 | 631 | 54.2 | 22.2 | 9.6 |
| 2014 | 24.6 | 547 | 53.7 | 21.1 | 7.9 |
| 2018 | 439.5 | 2,795 | 57.5 | 22.3 | 144.2 |
| All Counties | 37.1 | 245 | 46.6 | 19.7 | 5.9 |

Notes: .

Table 6: Transaction Level Demand Estimates

| Variable name | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Tax Elasticity | -0.16** | -0.16** | -0.16** | -0.17** | -0.16** | -0.17** |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) |
| Log Distance | | 0.00* | | | | |
| | | (0.00) | | | | |
| 1 or 2 Day Priority | | | 0.01** | | | |
| | | | (0.00) | | | |
| 1, 2, or 3 Day Package | | | | -0.00** | | |
| | | | | (0.00) | | |
| 1 Day First Class | | | | | -0.00** | |
| | | | | | (0.00) | |
| 1, 2, or 3 Day Standard | | | | | | -0.00** |
| | | | | | | (0.00) |
| Obs | 2,153,810 | 2,153,810 | 2,153,810 | 2,153,810 | 2,153,810 | 2,153,810 |
| R-Sq | 0.36 | 0.36 | 0.36 | 0.36 | 0.36 | 0.36 |

** 1% * 5%.

Table 7: Diff-in-Diff Demand Estimates

| Variable name | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Taxed Dummy | -0.10 | -0.11 | -0.10 | -0.12* | -0.11 | -0.12* |
| | (0.06) | (0.06) | (0.06) | (0.06) | (0.06) | (0.06) |
| Log Distance | | 0.03 | | | | |
| | | (0.03) | | | | |
| 1 or 2 Day Priority | | | 0.03 | | | |
| | | | (0.09) | | | |
| 1, 2, or 3 Day Package | | | | -0.08 | | |
| | | | | (0.04) | | |
| 1 Day First Class | | | | | -0.06 | |
| | | | | | (0.04) | |
| 1, 2, or 3 Day Standard | | | | | | -0.08 |
| | | | | | | (0.04) |
| Obs | 12,825 | 12,825 | 12,825 | 12,825 | 12,825 | 12,825 |
| R-Sq | 0.62 | 0.62 | 0.62 | 0.62 | 0.62 | 0.62 |

** 1% * 5%.

Table 8: 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 9: CES Demand Estimates

| Variable name | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Tax Elasticity | -1.34** | -1.34** | -1.35** | -1.34** | -1.37** | -1.34** |
| | (0.45) | (0.45) | (0.45) | (0.45) | (0.45) | (0.45) |
| Log Distance | | -0.00 | | | | |
| | | (0.02) | | | | |
| 1 or 2 Day Priority | | | -0.09 | | | |
| | | | (0.05) | | | |
| 1, 2, or 3 Day Package | | | | -0.00 | | |
| | | | | (0.03) | | |
| 1 Day First Class | | | | | -0.08** | |
| | | | | | (0.03) | |
| 1, 2, or 3 Day Standard | | | | | | -0.00 |
| | | | | | | (0.03) |
| Obs | 43,117 | 43,117 | 43,117 | 43,117 | 43,117 | 43,117 |
| R-Sq | 0.15 | 0.15 | 0.15 | 0.15 | 0.15 | 0.15 |

** 1% * 5%.

Table 10: Demand Estimates (Mode Fixed Effects)

| Year | Amazon | Other Taxed Online | Non-taxed Online |
|---|---|---|---|
| 2006 | -1.60* | 0.54 | |
| | (0.71) | (0.68) | |
| 2007 | -1.58* | 0.54 | -0.07 |
| | (0.71) | (0.68) | (0.08) |
| 2008 | -1.04 | 0.90 | 0.16 |
| | (0.72) | (0.69) | (0.08) |
| 2009 | -0.90 | 0.85 | 0.07 |
| | (0.72) | (0.69) | (0.09) |
| 2010 | -0.78 | 0.69 | -0.09 |
| | (0.72) | (0.69) | (0.09) |
| 2011 | -0.40 | 0.86 | -0.12 |
| | (0.72) | (0.69) | (0.09) |
| 2012 | -0.02 | 0.71 | -0.31** |
| | (0.72) | (0.69) | (0.09) |
| 2013 | 0.23 | 0.57 | -0.21* |
| | (0.72) | (0.69) | (0.09) |

** 1% * 5%.

Table 11: Demand Estimates (Demographics)

| Varaible | Amazon | Mode 2 | Mode 3 |
|---|---|---|---|
| Total Offline Competitors | -0.00 | -0.00 | -0.00 |
| | (0.00) | (0.00) | (0.00) |
| Large Offline Competitors | -0.00 | -0.00 | -0.00 |
| | (0.00) | (0.00) | (0.00) |
| Income | -0.00 | -0.10 | |
| | (0.07) | (0.06) | |
| % Pop Black | -0.34 | 0.35 | |
| | (0.25) | (0.22) | |
| % Pop White | 0.06 | 0.24 | |
| | (0.24) | (0.23) | |
| % Pop Asian | 1.73* | 0.04 | |
| | (0.78) | (0.77) | |

** 1% * 5%.

Table 12: Alternative Demand Models

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Variable name | | | | | | |
| Tax Elasticity | -1.52** | -1.72** | -1.86** | | -1.34** | -1.37** |
| | (0.27) | (0.58) | (0.54) | | (0.45) | (0.39) |
| Tax Elasticity (Amazon) | | | | -1.24** | | |
| | | | | (0.48) | | |
| Tax Elasticity (Mode 3) | | | | -1.43 | | |
| | | | | (0.80) | | |
| Obs | 43,117 | 53,914 | 29,584 | 43,117 | 43,298 | 43,117 |
| R-Sq | 0.30 | 0.15 | 0.08 | 0.15 | 0.15 | 0.19 |
| Regression | A-Weights | Zeros | 2008-2013 | Individual Tax Effect | No Forr Adjust-ment | No Pop Weights |

** 1% * 5%.

Table 13: Predicted Revenue

| Year | Amazon ($B) | Mode 2 ($B) | Mode 3 ($B) | Ave Distance (miles) | # Counties Taxed |
|---|---|---|---|---|---|
| 2006 | 5.60 | 20.40 | 24.13 | 296.61 | 317 |
| 2007 | 7.77 | 27.74 | 30.57 | 292.06 | 317 |
| 2008 | 9.78 | 28.77 | 30.48 | 236.45 | 379 |
| 2009 | 12.27 | 29.90 | 30.35 | 226.96 | 379 |
| 2010 | 17.87 | 32.88 | 33.41 | 244.73 | 380 |
| 2011 | 25.27 | 37.91 | 31.35 | 238.53 | 380 |
| 2012 | 32.45 | 28.94 | 22.95 | 225.04 | 755 |
| 2013 | 40.78 | 25.09 | 23.41 | 209.50 | 1194 |
| 2014 | 52.82 | 28.57 | 22.02 | 177.81 | 1670 |
| 2015 | 65.08 | 27.11 | 20.42 | 154.20 | 1943 |
| 2016 | 79.41 | 25.71 | 18.98 | 124.34 | 1989 |
| 2017 | 95.88 | 24.38 | 17.69 | 118.85 | 2053 |
| 2018 | 114.73 | 23.13 | 16.53 | 118.22 | 2053 |
| Total | 559.72 | 360.54 | 322.28 | | |

Notes:

Table 14: Revenue Effects of Expansion

| Year | Δ Revenue ($M) | Δ Ave Distance (miles) |
|---|---|---|
| 2007 | 6.42 | 4.55 |
| 2008 | 30.94 | 60.16 |
| 2009 | 36.75 | 69.65 |
| 2010 | 57.85 | 51.88 |
| 2011 | 85.71 | 58.08 |
| 2012 | 305.42 | 71.57 |
| 2013 | 558.87 | 87.12 |
| 2014 | 912.33 | 118.80 |
| 2015 | 1,256.75 | 142.42 |
| 2016 | 1,552.41 | 172.27 |
| 2017 | 1,874.66 | 177.76 |
| 2018 | 2,242.83 | 178.39 |
| Total | 8,925.85 | 1,092.64 |

Notes:.

Table 15: Effects of Moving FC to Highly Populated Area

| FCs Built by | Correlation of Tax and Population | # of FCs Moved to New State (%) | Ave Distance Moved (Miles) | Ave Revenue Lost from Moving ($M) |
|---|---|---|---|---|
| 2006 | 0.35 | 7 (88) | 252 | 138 |
| 2007 | 0.34 | 8 (89) | 165 | 156 |
| 2008 | 0.34 | 11 (92) | 153 | 76 |
| 2009 | 0.38 | 13 (93) | 172 | 28 |
| 2010 | 0.38 | 13 (93) | 172 | 28 |
| 2011 | 0.35 | 16 (94) | 155 | 59 |
| 2012 | 0.33 | 17 (90) | 157 | 61 |
| 2013 | 0.34 | 21 (81) | 146 | 53 |
| 2014 | 0.34 | 21 (70) | 138 | 47 |
| 2015 | - | 21 (66) | 125 | 41 |
| 2016 | - | 25 (47) | 115 | 38 |
| 2017 | - | 27 (49) | 130 | 35 |
| 2018 | - | 27 (48) | 132 | 32 |

Notes:.

Table 16: Examples of Moving FC to Highly Populated Area

| Year Built | Observed Location | Highest Populated Area | Distance | Revenue Change ($M) |
|---|---|---|---|---|
| 2006 | New Castle, DE | New York, NY | 115 | -495 |
| 2006 | Lyon, NV | Orange, CA | 412 | -318 |
| 2006 | Fayette, KY | Cook, IL | 311 | -212 |
| 2008 | Hillsburough, NH | Suffolk, MA | 39 | -124 |
| 2008 | Lake, IN | Cook, IL | 23 | -124 |
| 2011 | York, PA | DC | 87 | 411 |
| 2012 | Lexington, SC | Pinellas, FL | 427 | -43 |
| 2012 | Hamilton, TN | DeKalb, GA | 102 | 168 |
| 2016 | Pierce, WA | Multnomah, OR | 110 | 225 |
| 2016 | Baltimore City, MD | DC | 35 | 110 |

Notes:.

## A.2 Figures

Figure 1: Hirshman-Herfindahl Indices in Online Retail, 2006-2013

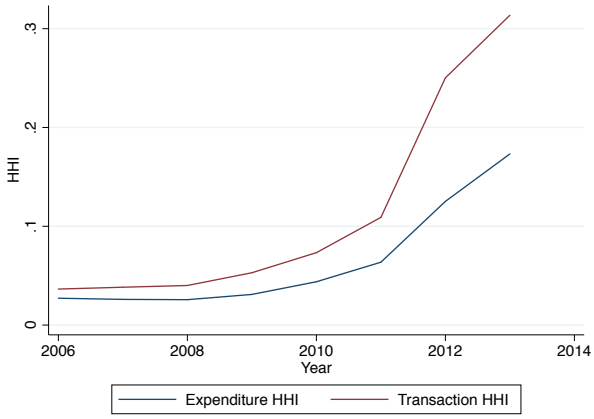
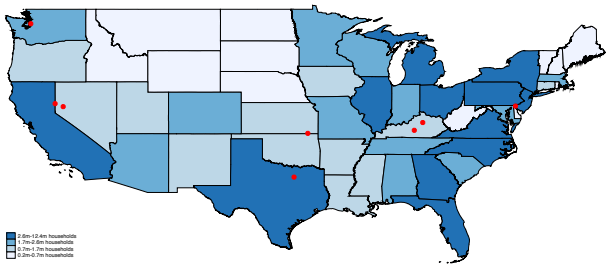
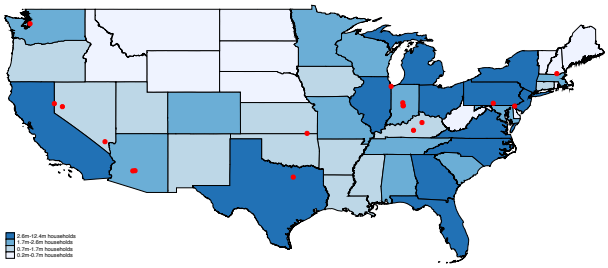


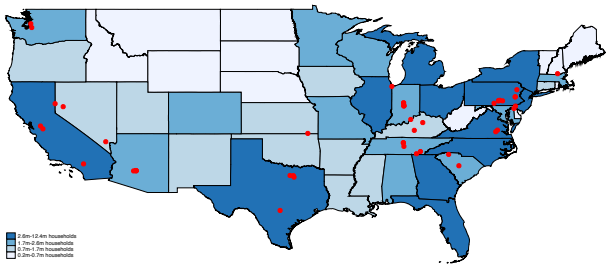
Figure 2: Amazon's fulfillment center Network and Number of Households



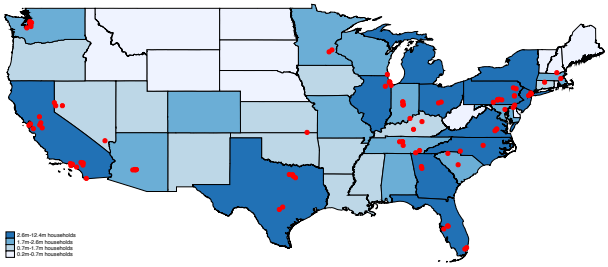
(a) 2006



(b) 2010

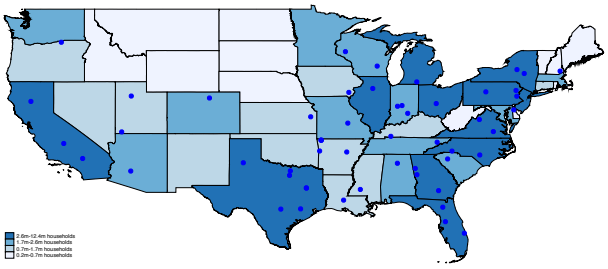


(c) 2014



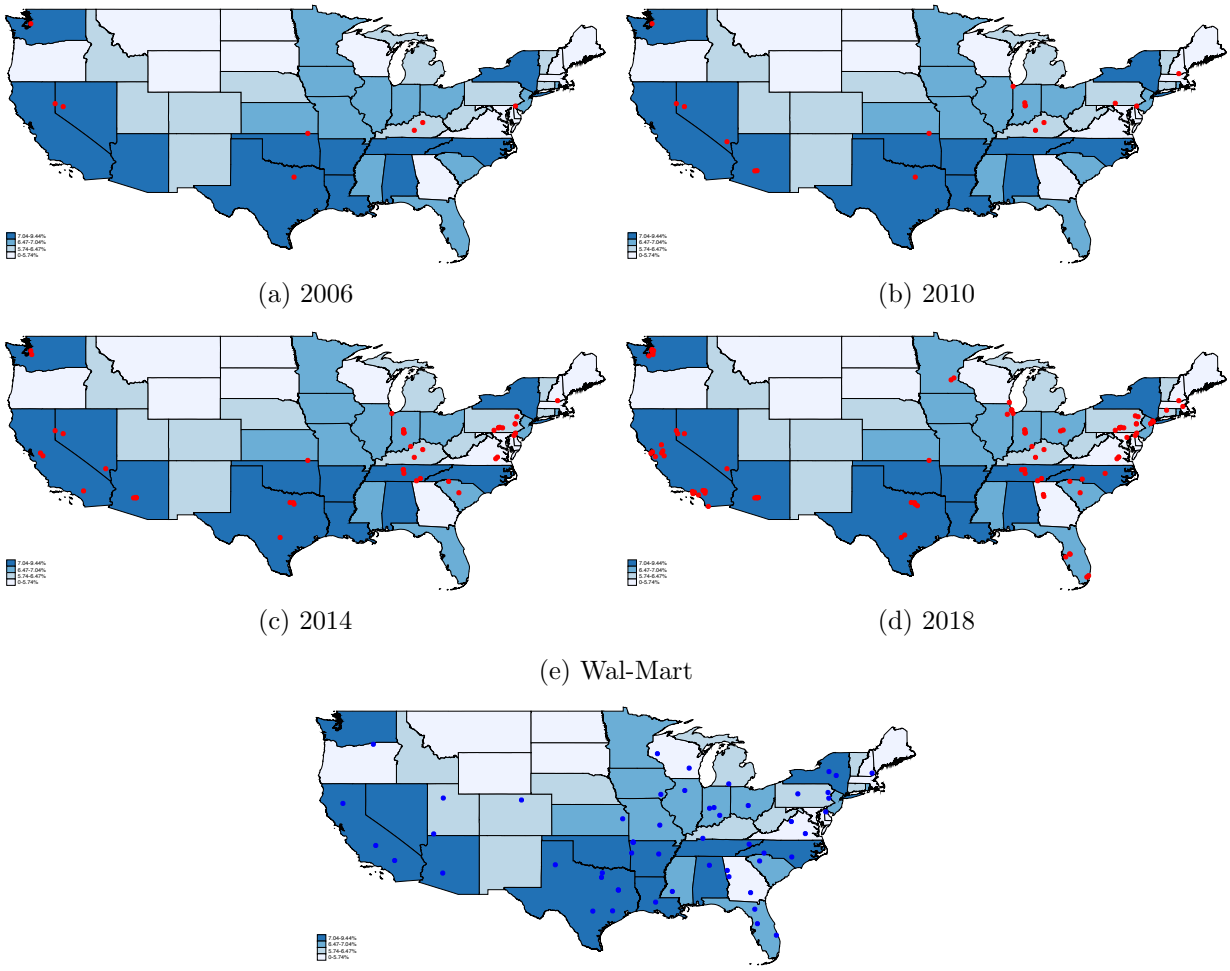
(d) 2018

(e) Wal-Mart



Notes:.

Figure 3: Amazon's fulfillment center Network and Average Tax Rates



(a) 2006

(b) 2010

(c) 2014

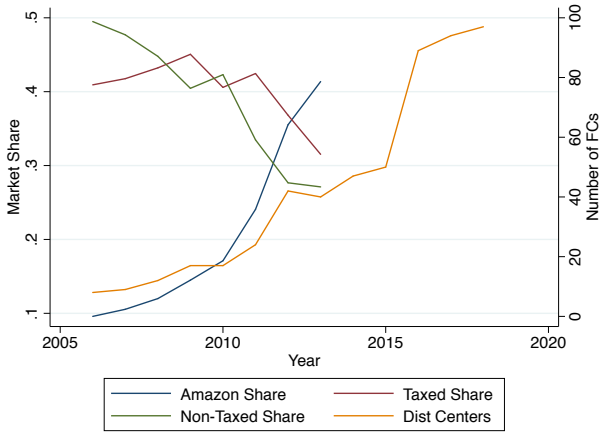
(d) 2018

(e) Wal-Mart

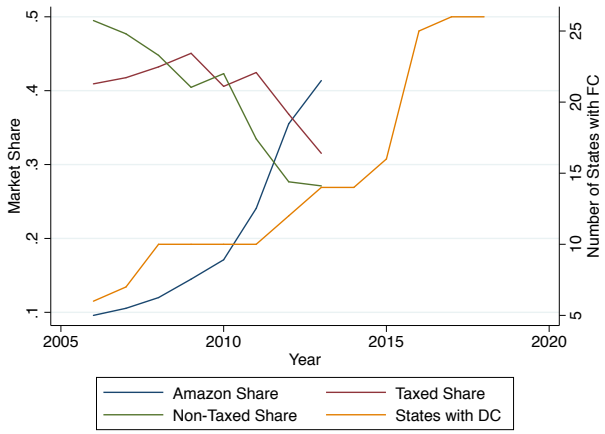
Notes:.

Figure 4: Market Share Dynamics

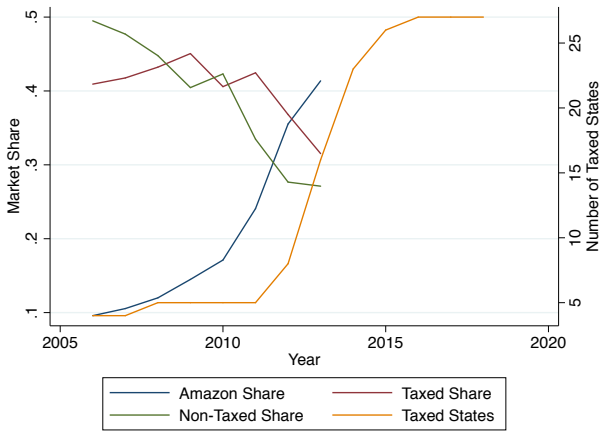
(a) Number of DCs




(b) Number of States with DCs




(c) Number of Taxed States




(d) Distances to DC

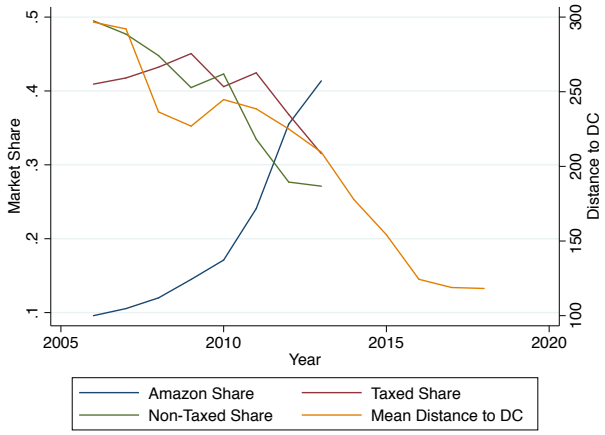
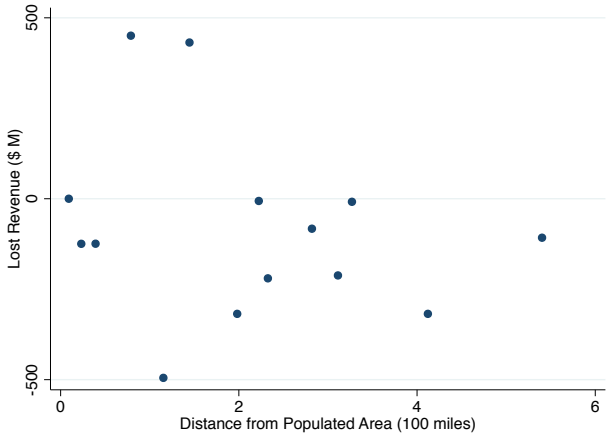
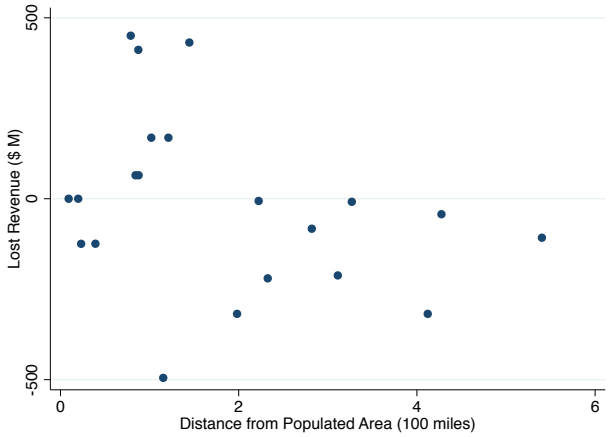


Figure 5: Effects of Moving FC to High Populated Area

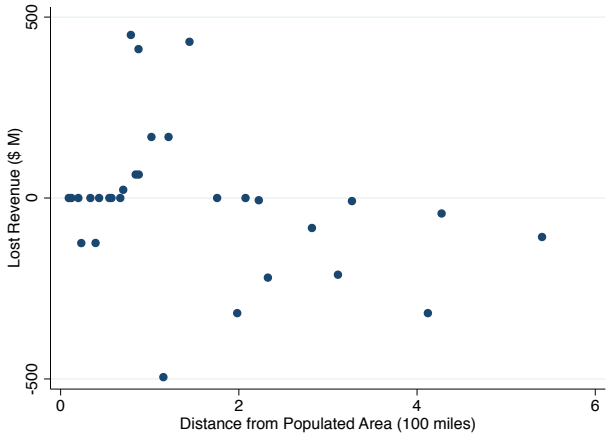
(a) FCs Built by 2010




(b) FCs Built by 2012




(c) FCs Built by 2014




(d) FCs Built by 2016

