Comparing Tax Elasticity Estimates

Mallick Hossain

4/18/2020

Introduction

This writeup compares my tax elasticity estimates with those in Houde, Newberry, and Seim (2017). I estimate that Amazon sales decrease by about 4% after sales tax collection, while HNS estimates about a 10% fall in Amazon sales (from Table 5 in HNS). What is driving this difference? Possible candidates that I will examine are the following:

- Data cleaning and organization
- Sample selection
- Aggregation

Model Setup

To ensure that estimates are comparable, I recreate the analysis done in HNS, which examines changes in annual county-level Amazon expenditures and Amazon collects sales tax. The regression equation is as follows:

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

where $ExpAm_{it}$ is the log of Amazon expenditures by the representative household in county i in year t. $1_{it}^{taxable}$ indicates whether Amazon collects sales tax in county i in year t. This takes a value of 1 in the year Amazon begins collecting sales tax.[1] d_{it} is the shipping speed from Amazon to the county. C_{it} includes demographic variables of the representative household in the county. Z_{it} is the number of small and large offline retailers in the county. Finally, λ_{it} includes a county fixed effect and a year fixed effect (not county-year fixed effect).

Since I do not have county-level demographics, shipping speeds, or retailer counts, I omit them and only estimate the following:

$$ExpAm_{it} = \beta_0 + \sigma 1_{it}^{taxable} + \lambda_{it} + \epsilon_{it}, \tag{2}$$

Data Cleaning and Organization

One possible candidate is that choices made during sample construction affect the estimates. This does not appear to be the case. For completeness, some notable differences in my sample construction are the following:

- Removal of items > $500

[1] This assumption may introduce some measurement error since Amazon may not collect sales tax until later in the year. For example, collecting sales tax in January and December are treated the same way under this assumption. Disaggregating to finer time intervals like quarters or months would address this concern.

Table 1: comScore Transactions

Step	Transactions
Starting Transactions:	$4,934,867$
Unduplicated Transactions:	$3,956,424$
Amazon Categories:	$2,478,115$
Invalid Prices:	$2,269,680$
Invalid Domains:	$2,021,800$

- Removal of purchases on domains that are unlikely to compete with Amazon
- Less precise adjustments to annual spending compared to HNS

I will not present all permutations of the data organization here, but starting with the likely sample used in HNS and then making the modifications above does not meaningfully change the estimates. I obtain a σ of about -0.105 across all specifications. It changes slightly, but only in the third decimal place.

The table below summarizes the steps in my data cleaning. Compared to HNS, only the "Invalid Prices" and "Invalid Domains" are more restrictive than in HNS, but as stated above, these restrictions do not meaningfully impact the estimation results.

Sample Selection

Another possibility is the sample selection procedure. The main sample selection differences are that I extend my sample through 2016 and that I restrict my sample to households that have made a purchase on Amazon. This is similar to the selection criterion of Baugh, Ben-David, and Park (2018) in which they restrict their analysis to households that spent more than $200 on Amazon in 2011. This does not appear to substantially affect the results either.

To create comparable estimates with HNS, I start with two samples. The first sample ("Full Sample") is the full set of transactions (post-cleaning). With this sample, I then compute average monthly spending by county and then multiply by 12 to get the annual spending. This is a rough approximation for the more precise adjustment done in HNS, which adjusts based on the number of months a household is present in the sample and then inflates that total spending to an annual amount (e.g., if a household was only active for six months, then their total spending is doubled).

The second sample restricts the Full Sample to only households that have made purchases on Amazon, which is about 18% of households. Then, I aggregate these expenditures to the annual county level as before.

For each sample, I run two estimates. In the first, I include all county-years, including counties that have zero spending on Amazon. One risk is that I impute zeros for counties that do not have a household in the comScore data. However, given that over 78% of counties are represented each year (and often over 90%), the noise introduced is small. Furthermore, given that I come close to the estimates in HNS, I am not too worried about this. For comparison of my estimates with HNS, the estimate from Table 5 in HNS is -0.105.

Columns (2) and (3) report estimates that are closest to those in HNS. The similarity of these two estimates is not surprising since Column (2) restricts analysis to only county-years with positive spending while Column (3) restricts analysis to counties with households that have made purchases on Amazon. These two groups are likely to overlap substantially. Column (4) is also in the ballpark, but lower (in magnitude) than Column (3). This is because the restriction to only positive county-months drops about 1000 observations which were zeros.

The most noticeable outlier is Column (1). This estimate is substantially lower in magnitude and this is because it includes over 8,000 county-years with zero spending on Amazon. Given that HNS uses 12,486 observations, it is likely that their estimate is generated from the restriction to county-years with positive spending, seen in Column (2).

Table 2: Elasticity Estimates (County-Year Level)

	Full Sample		Amazon Sample	
	(1)	(2)	(3)	(4)
Collect	0.032	-0.100^{**}	-0.103^{*}	-0.078^{*}
	(0.059)	(0.047)	(0.063)	(0.045)
County FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Positive County-Years	N	Y	N	Y
Observations	21,906	13,296	14,259	13,278
Adjusted R^2	0.268	0.214	0.145	0.110
Note:	*p<0.1; **p<0.05; ***p<0.01			

Overall, it appears that the restriction to households that have made purchases on Amazon (column 3) generates a similar estimate as that obtained in HNS.

Aggregation

The final analysis examines how aggregation may affect the estimates. HNS uses county-year observations while I opt for household-month observations. The previous sections have demonstrated that data cleaning and sample selection choices are unlikely to generate the disparate estimates obtained. Therefore, the remaining explanation must be the level of aggregation.

Using household-month observations, I estimate the following regression:

$$ExpAm_{ht} = \beta_0 + \sigma 1_{ht}^{taxable} + \lambda_{ht} + \epsilon_{ht}, \tag{3}$$

where $ExpAm_{ht}$ is the log of Amazon expenditures by household h in month-year t.[2] $1_{it}^{taxable}$ indicates whether Amazon collects sales tax for household h in month-year t. Finally, λ_{ht} includes a household fixed effect and a month-year fixed effect (not household-month-year fixed effect). The table below reports the results for both my Amazon sample and for the Full sample.

Table 3: Elasticity Estimates (Household-Month Level)

	Full Sample	Amazon Sample
	(1)	(2)
Collect	-0.011^{**}	-0.035^{**}
	(0.005)	(0.014)
Household FE	Y	Y
Month-Year FE	Y	Y
Observations	5,107,014	1,276,856
Adjusted R^2	0.188	0.097
Note:	*p<0.1; **p<0.05; ***p<0.01	

Column (1) is substantially smaller in magnitude than the estimate in column (2). This is likely due to the inclusion of many household-months with no Amazon spending. Column (2) is virtually the same estimate as

[2] Since I allow for household-months with zero spending, I use $\log(1 + \mathit{Spending})$.

what I already obtain. Converting these to elasticities by dividing by the average tax rate of 6.8% gives an elasticity range of -0.16 to -0.51.

Conclusion

Overall, it appears that the choice of aggregation (county-year versus household-month) is the primary driver of the differences in elasticity estimates between my analysis and HNS and that this is not due to sample selection or organization.