Economics Bulletin

Volume 40, Issue 1

Estimating the impact of e-commerce on retail exit and entry using Google Trends

David C Vitt
Department of Economics, Farmingdale State College

Abstract

I address the degree to which variation in exposure to e-commerce is associated with establishment entry and exit in the retail industry at the county level. To measure exposure to e-commerce, I rely on within-state variation in relative search frequency for the phrase “amazon prime” as reported by Google Trends. To generate exogenous variation in this e-commerce exposure measure, I use within state variation in the relative search frequency for “porn” and “cat videos”. Fixed effects instrumental variable estimates suggest at least 10 of the 27 retail industry groups experience net exit with increasing e-commerce exposure, while at least 6 experience net entry. To address endogeneity concerns about my instruments, particularly that they are driven by a notion of “hipster-ness”, I conduct a robustness check to show that my results fail to replicate in consideration of a strategy to tease out this identification threat.

I would like to thank two anonymous referees, John Conley, and Alexander McQuoid for their valuable feedback, your help has significantly improved my manuscript, thank you.
Citation: David C Vitt, (2020) "Estimating the impact of e-commerce on retail exit and entry using Google Trends", *Economics Bulletin*, Volume 40, Issue 1, pages 679-688
Contact: David C Vitt - vittd@farmingdale.edu.
Submitted: October 15, 2018. **Published:** February 29, 2020.

1. Introduction

In this article, I estimate the impact of variation in consumer's tastes for e-commerce on entry and exit in brick and mortar retail. There is a small literature focused on estimating the impact of e-commerce on retail structure (Goldmanis et al. 2009; Bakos et al. 2005); as well as the impact of e-commerce on prices and price dispersion (Brynjolfsson and Smith 2000; Clay, Krishnan, and Wolff 2001). The primary limitation of this literature is the use of survey data to measure e-commerce use. Furthermore, empiricists have a reasonable endogeneity concern that appears be unaddressed: measures of e-commerce use and measures of retail industry market structure are likely simultaneously determined. The contribution of this note is to demonstrate the utility of Google Trends in addressing these limitations and further exploring the impact of e-commerce on entry and exit in brick and mortar retail.

Google Trends reports the relative search frequency for a desired search phrase at the Google search engine. With exception for the least searched phrases, relative search frequency data is available at the country, state and MSA levels over any desired period of time since 2004. Given the known limitations of methods based on stated preferences (Vosen and Schmidt 2011), Google Trends offers a significant advantage by measuring the revealed intensity of consumer preference or interests vis-à-vis the within-unit variation in relative search frequency.

The use of Google Trends in economic research is recent. (Choi and Varian 2012) survey the system and its many advantages, and demonstrate its utility in predicting motor vehicle sales and unemployment claims. Other specific examples of using Google Trends include creating real time uncertainty indices (Castelnuovo and Tran 2017), predicting cinema sales (Hand and Judge 2012), as an instrument for firearm sales (Vitt et al. 2018), explaining interest in cryptocurrency (Yelowitz et al. 2015; Vitt 2017b), and investigating the impact of e-commerce on retail employment (Vitt 2017a).

My empirical strategy relies on within-state variation in consumer's propensity to use the Internet for tasks associated with the search phrases “amazon prime”, “porn”, and “cat videos”. Searches for “amazon prime” are used as a proxy for consumer's propensity to use e-commerce, and can be similarly regarded as brick and mortar's exposure to e-commerce competition. It is reasonable to believe that the propensity to use e-commerce is a function of the ease of reaching a brick and mortar retailer, along with the product variety and markups therein. To address this concern, it is necessary to find an instrument for the variation in e-commerce use.

Below, I argue that consumers who readily use the internet for tasks associated with the keywords “porn” and “cat videos” are also likely to use it to shop online. Similarly, I make the assertion that searches for these keywords have no direct impact on retail entry or exit except through how they determine the intensity of e-commerce use, and specifically that the urges to go online for these tasks are effectively random.

Using plausibly exogenous within-state variation in various consumer “urges”, I instrument for variation in e-commerce use and trace out any subsequent impact on county level retail establishment counts disaggregated by industry group. My preferred specification allows for across industry heterogeneity in the slope parameter associated with e-commerce exposure. This flexibility allows me to investigate differences in the impact of e-commerce across retail industry groups. I find evidence that e-commerce is a technology that induces net exit in at least 10 of the 27 retail industry groups, while it induces net entry in at least 6 of the 27 retail industry groups. I briefly try to characterize this heterogeneity and lay the foundation for the future work on the mechanisms generating the heterogeneous responses.

The rest of this paper is organized as follows. Section 2 describes my methods, first discussing sources of data and then on the empirical strategy, including how the Google Trends instruments are constructed. Section 3 describes features of my estimated results, while section 4 entertains a robustness exercise to address instrument validity. Section 5 concludes by suggesting future avenues of exploration in this topic.

2. Methods

2.1 Data Description

For observations of retail establishment counts, I rely on the County Business Patterns, available from the U.S. Census Bureau. This data presents establishment counts for every county and year at various levels of industry aggregation according to the North American Industry Classification System (NAICS). I focus on the 4 digit industry group level of aggregation for retail (e.g. all NAICS codes 44XX-45XX) for the period 2007-2015.

My variable of interest and my instruments, both of which are discussed below, are the relative search frequency for particular terms as reported by Google Trends. Relative search frequency for any search phrase is defined as the proportion of total searches dedicated to that phrase in the particular area and timeframe, which is then scaled on a range of 0 to 100. I construct this variable for the period 2007-2015, with further details on construction discussed in the next section.

All other control variables are sourced from the U.S. Census Bureau. Summary statistics for all variables are presented in Table 1.

Table 1. Summary Statistics

	mean	Within-state std. dev.	min	max
Establishments	17.05	4.83	1	3742
Median Income	51926.8	2124.9	37173	75920
Population	125300.8	6939.9	442	9962789
Amazon Search Intensity	42.0	10.5	10.8	68.3
Porn Search Intensity	56.7	17.4	21.9	90.1
Cat Video Search Intensity	32.6	14.1	2.2	88.1
Beard Oil Search Intensity	49.4	14.2	9.3	89.7
Farm to Table Search Intensity	37.3	11.1	11.5	91.9

Construction of the final 3 variables in this table is reviewed in detail in the "Empirical Strategy" section of this paper.

2.2 Empirical Strategy

The causal path motivating my strategy starts with exogenous changes in consumer's propensity to use the internet for pornography or cat videos, and how this variation in Internet use intensity causes variation in the intensity of e-commerce use. This line of reasoning is associated with the first stage regression. The second stage regression reflects a path by which variation in the intensity of e-commerce use is associated with variation in the demand for products at brick and mortar retailers, and the subsequent exit or entry that changes establishment counts.

My preferred first stage specification is given by:

$$
\begin{aligned}
e - commerce\ exposure_{st} \\
&= \underbrace{\pi_1\ pornography\ search\ intensity_{st} + \pi_2\ cat\ video\ search\ intensity_{st}}_{\text{exogenous sources of internet use intensity}} \\
&+ \underbrace{X_{st}\Pi + \mu_s + \mu_t}_{\text{controls and fixed effects}} + v_{st}
\end{aligned}
\tag{1}
$$

where $e - commerce\ exposure_{st}$ is the logarithm of relative search frequency for the phrase "amazon prime" for state s in year t as reported by Google Trends. Similarly, the instruments $pornography\ search\ intensity_{st}$ and $cat\ video\ intensity_{st}$ are the relative search frequency for the keywords "porn" and "cat videos" at Google for state s in year t, as reported by Google Trends.

These search intensity variables are constructed by visiting Google Trends, searching for the aforementioned phrases, and in sequence downloading the monthly data for each state and averaging the relative search frequency for each year. Included as controls in X_{st} are the median income in state s for year t, as well as state fixed effects and year fixed effects. Idiosyncratic shocks to e-commerce exposure from unobserved sources are given by v_{st}.

It's useful to be reminded of what constitutes an ideal instrument in this context. I use relative search frequency for particular phrases as the instrument, so appropriate choice of phrasing is central to my results. Instruments must be relevant – the phrases I use should be strongly correlated with my measure of e-commerce use (searches for "amazon prime"). The instrument must also be excludible- one that is unrelated to brick and mortar entry and exit, except through how variation in searches for this phrase could mean variation in consumer's e-commerce use.

Regarding the former criteria, my chosen keywords to use as instruments are highly relevant; individuals who use the Internet for pornography or cat videos are likely to also use the Internet for e-commerce. This is confirmed by the joint significance of the coefficients on these instruments in the first stage given at the bottom of Table 2.

The exclusion restriction also seems reasonable: variation in the use of the Internet for pornography or cat videos should have no impact on retail establishment counts except through generating variation in consumer use of e-commerce. Variation through time in searches for these phrases is largely determined by random variation in consumer tastes and attitudes towards the subjects they are useful for. Both capture the behavior of large and distinct cross-sections of the population. Perhaps for stores associated with the adult industry, most of which are included in "Other Misc. Stores" (NAICS 4539), the "porn" keyword may proxy latent demand for adult industry products. To the degree this is true, results for this particular industry group should be discounted by the reader.

My preferred second stage specification is given by:

$$establishments_{ict} = \beta_{1i} \widehat{e-commerce\ exposure}_{st} + \underbrace{X_{ct}\Gamma + \phi_i\,\mathrm{t} + \mu_{ic} + \mu_t}_{\substack{\text{controls,}\\ \text{industry trend}\\ \text{panel and year fixed effects}}} + u_{ict} \quad (2)$$

where $establishments_{ict}$ is the establishment count in retail industry group i in county c for year t. The effect of interest is β_{1i}, the coefficient on the fitted value of e-commerce intensity. Included as controls are the population and median income within the state. An industry specific time trend is included as a linear control for secular variation specific to the industry group. The panel fixed effect, μ_{ic} performs the within industry-county demeaning, and controls for time invariant confounding effects as well as for mean differences in the effect of interest and controls. The year fixed effect, μ_t, controls for shocks common to all counties, like those arriving from changes in macroeconomic conditions or other secular variation.

On account of the different levels of aggregation used in the second stage, and to allow for the idiosyncratic disturbances to be correlated, as reviewed in (Bertrand, Duflo, and Mullainathan 2004), I attempt to mitigate any bias in my standard error estimates by clustering at the state level.

3. Results

Table 2. OLS and FEIV regression of retail establishment counts on e-commerce exposure

	(1) Log Establishments	(2) Log Establishments	(3) Log Establishments	(4) Log Establishments	(5) Log Establishments
E-Commerce Exposure (all retail industry groups)	0.792**** (0.0158)	0.00548 (0.00407)			
Log Population		0.315**** (0.0346)	0.319**** (0.0345)	0.365**** (0.0328)	0.382**** (0.0331)
Log Median Income		0.0472**** (0.0123)	0.0464**** (0.0123)	0.0438**** (0.0124)	0.0527**** (0.0131)
Automobile Dealers			0.0917**** (0.0128)	-0.175**** (0.0526)	-0.00359 (0.0546)
Other Motor Vehicle Dealers			0.0834**** (0.0182)	0.215**** (0.0586)	-0.258** (0.107)
Automotive Parts, Tire			-0.0583**** (0.00986)	-0.0426 (0.0411)	-0.0426 (0.0440)
Furniture Stores			0.127**** (0.0170)	-0.150** (0.0604)	-0.101 (0.0799)
Home Furnishings Stores			0.110**** (0.0173)	0.0488 (0.0638)	0.0199 (0.0800)

Electronics and Appliance			-0.0240	-0.460****	-0.264
			(0.0148)	(0.0644)	(0.880)
Building Material and Supplies			0.0317***	0.350****	0.0202
			(0.0104)	(0.0458)	(0.0495)
Lawn and Garden Equipment			0.0936****	0.333****	0.022
			(0.0171)	(0.0695)	(0.0655)
Grocery			-0.0112	-0.0302	-0.0522
			(0.0147)	(0.0444)	(0.0494)
Specialty Food			0.0472**	-0.238**	-0.0307
			(0.0225)	(0.102)	(0.0987)
Alcohol (Beer Wine Liquor)			-0.121****	0.00732	0.244****
			(0.0160)	(0.0684)	(0.0681)
Health and Personal Care			-0.0595****	-0.165****	-0.0151
			(0.0109)	(0.0458)	(0.0466)
Gasoline			-0.0716****	-0.0820**	-0.0891
			(0.00881)	(0.0339)	(0.368)
Clothing			-0.00782	-0.0570	0.0151
			(0.0162)	(0.0515)	(0.0705)
Shoe			-0.0449***	0.0892	0.0966
			(0.0165)	(0.0760)	(0.0936)
Jewelry Luggage Leather			0.0661****	-0.451****	0.141*
			(0.0170)	(0.0742)	(0.0800)
Sporting and Hobby			0.0262	-0.337****	0.0293
			(0.0201)	(0.0790)	(0.0889)
Books Periodicals Music			0.483****	1.226****	0.294
			(0.0334)	(0.143)	(0.173)
Department Stores			0.0174	-0.725****	-0.0527
			(0.0146)	(0.102)	(0.0868)
Other General Merch.			-0.144****	-0.104**	0.0635
			(0.0120)	(0.0452)	(0.0434)
Florists			0.172****	0.0343	0.132**
			(0.0190)	(0.0675)	(0.0630)
Office Supplies			0.177****	0.0256	-0.0270
			(0.0184)	(0.0750)	(0.0824)
Used Merch			-0.201****	-0.353****	-0.117
			(0.0215)	(0.0944)	(0.0969)
Other Misc. Stores			0.0215	-0.212***	-0.112
			(0.0168)	(0.0819)	(0.0930)
Electronic Shopping			-0.365****	0.928****	-0.196*
			(0.0295)	(0.161)	(0.107)
Vending Machine Operators			0.0802**	-0.0300	0.130
			(0.0320)	(0.123)	(0.171)
Direct Selling			-0.111****	0.544****	-0.112
			(0.0172)	(0.0727)	(0.0783)
Instruments	None	None	None	Search Intensity: Porn, Cat Videos	Search Intensity: Farm to Table , Beard Oil

Fixed Effects	None	Industry-County	Industry-County, Year	Industry-County, Year	Industry-County, Year
Stage 1 F (KP)				30.42	2.16
Trend	None	None	None	Industry Specific	Industry Specific
Std. Error	State Cluster	State Cluster	State Cluster	State Cluster	State Cluster
Observations	341,994	341,994	341,994	341,994	341,994

Standard errors in parentheses, Industry-County refers to a fixed effect that indicates industry group i present in county c.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

In column 2, I estimate a simple fixed effects (FE) model by adding industry-county fixed effects along with controls for market size and median income. I find that the effect of interest, e-commerce exposure, is very close to zero (relative to column 1), and statistically insignificant. I proceed by incrementally adding features to reach my preferred specification.

The naïve specifications in columns 1-2 impose homogeneity of the slope parameter on the variable of interest: $\beta_{1i} = \beta$ for all industries. A homogeneity restriction of this sort could easily mask economic significance, and does not realize the full potential of the data. Imposing this homogeneity makes the slope parameter represent a weighted average of effects across retail industries. It's reasonable to believe that some industries have benefitted from e-commerce, while others have likely not fared well from being closer to their competition, and these effects could be missed when aggregating to a single estimated across industry effect.

Column 3 presents estimation of (2) using OLS with the added flexibility of allowing industry-specific slope parameters for the variable of interest. Notice that the need for this flexibility is validated by the heterogeneity in effects across industries: the impact of e-commerce exposure is different across retail industries. This specification still suffers from endogeneity concerns, so I pay little attention to it other than to note the heterogeneity across industry groups.

Column 4 utilizes all the features of my identification strategy with a fixed effects instrumental variable (FEIV) estimator. The first stage F statistic, presented at the bottom of the table, suggests my estimates do not suffer from excessive bias due to weak instruments according to the conventional rule of thumb. Additionally, in a Hansen J test of overidentifying restrictions, I fail to reject the null hypothesis of valid instruments.

Interpretation of the results in column 4 goes as follows: for a 1% increase in e-commerce exposure, there is an expected β_{1i} % change in retail establishment counts for retail industry i in county c. For example, for a 10% increase in e-commerce exposure, there is an expected 7.25% decrease in department stores, 4.51% decrease in jewelry and luggage stores, 4.6% decrease in electronics and appliance stores.

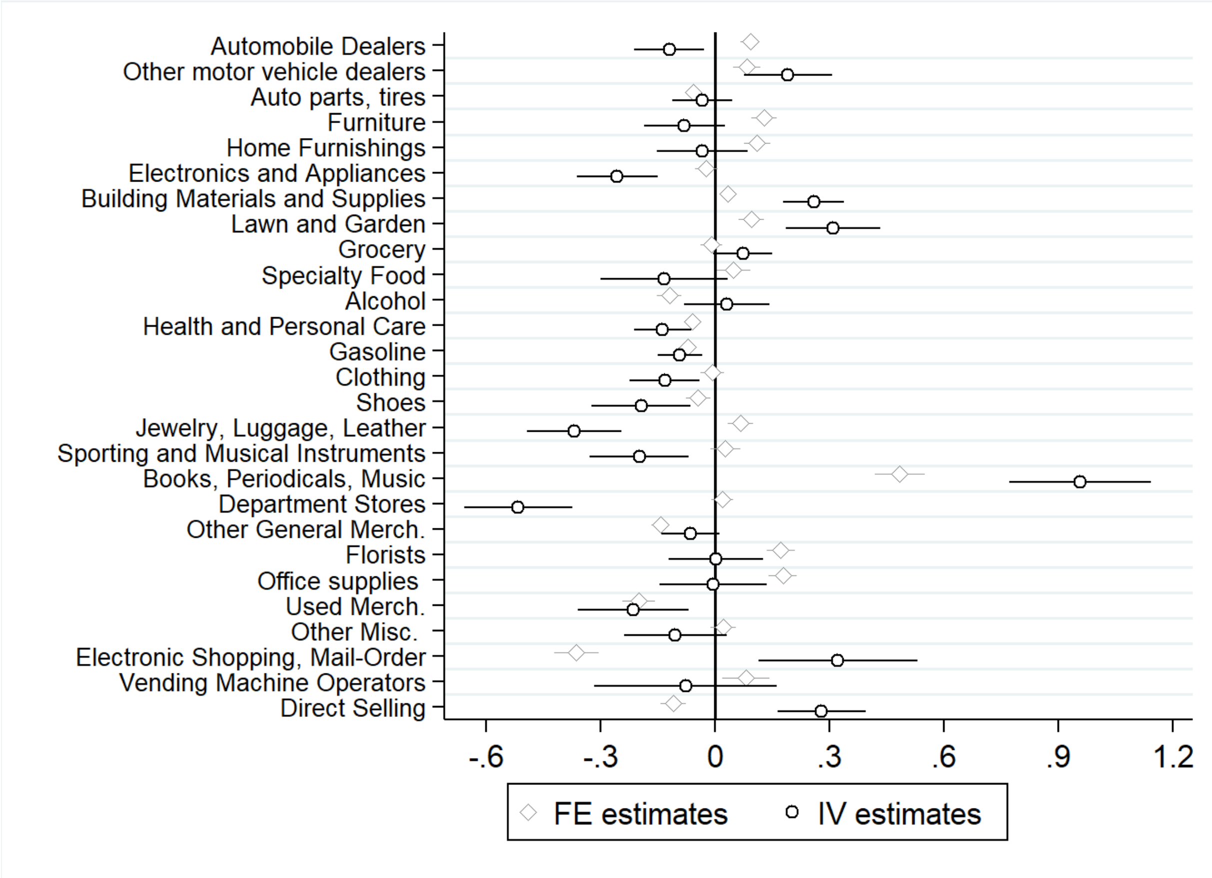

Figure 1. OLS and FEIV point estimates indicated by circles and diamonds respectively, with the horizontal lines representing the confidence interval of the point estimate. Confidence intervals based on standard errors estimated with clustering at the state level.

As an alternative to exhaustively going through the list of estimates, I present standard coefficient plots in Figure 1. Industry groups are labeled on the vertical axis. FEIV point estimates from column 4 are indicated by hollow circles in black, with confidence intervals indicated by the horizontal lines originating from the point estimate. Similarly, I present FE point estimates from column 3 as grey hollow diamonds with similar confidence interval bands.

Among the interesting FEIV results to highlight is in the electronic shopping and mail order industry group (NAICS 4541), for a 10% increase in e-commerce exposure, county level firm counts in this retail group would be expected to increase 9.28%. This is intuitive, as consumers in the state reveal their preference for shopping online, firms attempt to capture these customers, and must enter to do so, explaining the increase in county level counts in this group.

A comparison of the FE and FEIV estimates either in Figure 1 or in column 3 and 4 of Table 2, yields a non-parametric way to demonstrate robustness of my result. Of the 27 NAICS retail industry groups, 16 have e-commerce exposure effects significant at the 95% or greater confidence level. Of the 16 statistically significant effects, 7 carry a different sign than the corresponding FE estimate in column 3. For 10 of the represented retail industry groups with a significant e-commerce exposure effect, the estimates suggest that e-commerce is a technology that causes net exit within the industry. An informal binomial sign test under the hypothesis of equal probability of negative and positive estimates suggests that the probability of 10 estimates with negative signs in 16 trials is 22%. For the remaining 6 industries with statistically significant and positive e-commerce exposure effects, empirical evidence suggests that e-commerce is a technology that increases expected profits and induces net entry.

4. Robustness Exercise

One curious result deserving attention on robustness and validity of the instruments is the simultaneous large positive impact of e-commerce use on the Book, Periodical and Music industry group, and large negative impact on Gasoline sales. The gas result can be intuitively explained by the substitution for e-commerce decreasing demand for trips to brick and mortar retail, and therefore decreasing demand for gasoline.

The positive coefficient on Books, Periodicals, and Music stands in contrast to the (Goldmanis et al. 2009) result, which suggested net exit as a result of increased e-commerce competition. In considering threats to identification in this particular context, one might think that some notion of "hipster-ness" could be driving the result. The more hip a place is, the more intensely it uses the internet for e-commerce, the less it relies on personal vehicles, the greater its preference for quirky bookstores, etc. I find this mechanism plausible, but unlikely to confound my estimates. While I maintain that any residual time varying hipster-ness not captured by the industry-group trends and year fixed effects is likely not related with the urges that drive the searches my instruments measure, I nonetheless defend this position with a brief robustness exercise.

If the hipster-ness mechanism is driving my result, then similar results should replicate with a different set of instruments that are clear measures of the intensity of hipsters in the population. I measure the intensity of hipsters with searches for "farm to table" and "beard oil". While this choice represents a gross generalization of the habits of hipsters (that they have beards and prefer to eat at restaurants that know where food is sourced), I argue they proxy the intensity of the hipster subpopulation. Further, I argue that searches for these terms have no immediate impact on retail establishment counts for most industry groups, except through how they measure the presence of a subpopulation that may have differential propensity to shop online.

Replacing these as the instruments in equation (1) and then running the preferred specification in (2), I arrive at the results in column 5 of Table 2. Comparing my preferred specification in column 4 to this robustness check in column 5, I note that the statistical significance of many of the estimates in column 4 does not replicate in column 5. I also note that all significant industry group effects in column 5 differ either in sign or in significance status compared to my preferred specification results in column 4. In interpret the totality of the evidence as suggesting the hipster-ness mechanism does not drive my result.

5. Conclusion

In this paper, I introduced an identification strategy that relies on Google Trends as a more credible source of variation in e-commerce use than the current practice relying on survey data. Google Trends provides an easy route to direct measurement of consumer's revealed preference for using the Internet for various tasks by reporting the relative search frequency for desired phrases within a state over time. This single database allows me to collect information on e-commerce usage, but also to motivate appropriate instruments to address endogeneity concerns. Specifically, I measured intensity of e-commerce usage through the relative search frequency for the phrase "amazon prime". I instrumented for variation in this search phrase using variation in searches for "porn" and "cat videos", both of which are determined by random variation in consumer tastes and urges, and are likely candidates to satisfy the exclusion restriction.

I regressed county level establishment counts on instrumented e-commerce exposure using a fixed effects instrumental variable estimator that allows for across industry heterogeneity in the slope parameters on the variable of interest. I find that the impact of plausibly exogenous increases in e-commerce exposure leads to net exit in 10 of the 16 retail industries with statistically significant e-commerce exposure effects. Empirically, this also suggests that e-commerce is a technology that increases profits and induces net entry in the remaining 6 retail industries with statistically significant estimates.

Future work in this area should focus in two areas. First, further effort to develop and validate the use of Google Trends as a source of instruments for economic questions. Additionally, it will be useful to explore the mechanisms that generate this heterogeneity in effects across industry groups. To what degree do industry groups adjust to e-commerce on other intensive margins like employment? Having a better understanding of the alternative margins of adjustment will be useful in analyzing the mechanisms generating this heterogeneity and will improve our understanding of the organization of this sector of our real economy.

References

Bakos, Yannis, Henry C. Lucas, Wonseok Oh, Gary Simon, Siva Viswanathan, and Bruce W. Weber. 2005. "The Impact of E-Commerce on Competition in the Retail Brokerage Industry." *Information Systems Research* 16 (4): 352–71. https://doi.org/10.1287/isre.1050.0064.

Bertrand, M, E Duflo, and S Mullainathan. 2004. "How Much Should We Trust Differences-In-Differences Estimates?" *The Quarterly Journal of Economics* 119 (1): 249–75. https://doi.org/10.1162/003355304772839588.

Brynjolfsson, Erik, and Michael D. Smith. 2000. "Frictionless Commerce? A Comparison of Internet and Conventional Retailers." *Management Science* 46 (4): 563–85. https://doi.org/10.1287/mnsc.46.4.563.12061.

Castelnuovo, Efrem, and Trung Duc Tran. 2017. "Google It Up! A Google Trends-Based Uncertainty Index for the United States and Australia." *Economics Letters*. https://doi.org/10.1016/j.econlet.2017.09.032.

Choi, Hyunyoung, and Hal Varian. 2012. "Predicting the Present with Google Trends." *Economic Record* 88 (SUPPL.1): 2–9. https://doi.org/10.1111/j.1475-4932.2012.00809.x.

Clay, Karen, Ramayya Krishnan, and Eric Wolff. 2001. "Prices and Price Dispersion on the Web:

Evidence from the Online Book Industry." *The Journal of Industrial Economics* 49 (4): 521–539. https://doi.org/10.1111/1467-6451.00161.

Goldmanis, Maris, Maris Goldmanis, Chad Syverson, Chad Syverson, Putnam Investments, and Putnam Investments. 2009. "E-Commerce and the Market Structure of Retail Industries." *Economic Journal* 120: 651–82. https://doi.org/10.1111/j.1468-0297.2009.02310.x.

Hand, Chris, and Guy Judge. 2012. "Searching for the Picture: Forecasting UK Cinema Admissions Using Google Trends Data." *Applied Economics Letters* 19 (11): 1051–55. https://doi.org/10.1080/13504851.2011.613744.

Homburg, Christian, Nicole Koschate, and Wayne D. Hoyer. 2005. "Do Satisfied Customers Really Pay More? A Study of the Relationship Between Customer Satisfaction and Willingness to Pay." *Journal of Marketing* 69 (2): 84–96. https://doi.org/10.1509/jmkg.69.2.84.60760.

Pine, B. Joseph, and James H. Gilmore. 1999. "The Experience Economy: Work Is Theatre & Every Business a Stage." *Business*. https://doi.org/10.1080/02642069700000028.

Vitt, David C., McQuoid, Alexander F., Moore, C., & Sawyer, S. (2018). "Trigger Warning: The Causal Impact of Gun Ownership on Suicide". *Applied Economics*, *50*(53), 5747–5765. https://doi.org/10.1080/00036846.2018.1488064

Vitt, David C. 2017a. "Did Internet Kill the Retail Trade Star ? Using Google Trends to Estimate the Impact of E-Commerce on Retail Employment." https://ssrn.com/abstract=3074132.

———. 2017b. "Does Fiat-to-Bitcoin Exchange Activity Lead to Increased User- to-User Bitcoin Transaction Activity?" *Academy of Economics and Finance Journal* 8: 71–76.

Vosen, Simeon, and Torsten Schmidt. 2011. "Forecasting Private Comsumption: Survey Based Indicators vs Google Trends." *Journal of Forecasting* 578 (January): 565–78.

Yelowitz, Aaron, Matthew Wilson, Aaron Yelowitz, and Matthew Wilson. 2015. "Characteristics of Bitcoin Users : An Analysis of Google Search Data." *Applied Economics Letters* 22 (13): 1030–36. https://doi.org/10.1080/13504851.2014.995359.