Labor Market Senior Project
Department of Economics

"Hedonic Pricing Model: The Value of Fiber-Optic Internet Access"

Trevon Burgins
Spring 2018

Instructor: *Dr. Francesco Renna*

Abstract

High-speed internet demand is on the rise around the world. In 2015, the City of Hudson launched a years-long project that would make Hudson the only city in Ohio to offer a true gigabit internet service (Mackinnon, 2015). The project has the estimated potential to take five years to complete. The focus of the city-wide project is on serving the demand of businesses in the area. The city sees high-speed internet as a way to attract and retain employers, business, and firms alike. While their project aims at how businesses value fiber-optic internet service, this study evaluates how consumers in Ohio value that same service.

Table of Contents Page

I. Introduction

High-speed internet demand is on the rise around the world. High-speed internet experienced subscription growth of 93 million in 2017, which represents a growth of almost 8% in all countries in the Organisation for Economic Cooperation and Development (OECD, 2018). A federal budget requested by the Trump administration includes a $1.6 billion increase in 2018. The request, if approved, would result in a budget of $95.7 billion for federal information technology (I.T.) (Miller, 2017), while most government agencies are expected to either decline or remain the same. This figure confirms and emphasizes the importance of high-speed technology in the modern world. In more local news, in 2015, the City of Hudson launched a years-long project that would make Hudson the only city in Ohio to offer a true gigabit internet service (Mackinnon, 2015). The project has the estimated potential to take five years to complete. The focus of the city-wide project is on serving the demand of businesses in the area. The city sees high-speed internet as a way to attract and retain employers, business, and firms alike. While their project aims at how businesses value fiber-optic internet service, this study evaluates how consumers value that same service. Specifically, this study seeks to evaluate how consumers in the cities around Hudson, Ohio value fiber-optic internet access.

Hudson has shown an interest in expanding its resource to other cities in the state, such as Cleveland and Akron. However, a major set-back to implementing fiber-optic connections in certain areas is the cost requirements, especially when other servers such as digital subscriber lines (DSL) and cable are still being used in the area (Anderson, 2017). This cost is a major factor for why there is only 25% fiber coverage nationwide (Anderson, 2017). The problem in the United States is the internet providers are burdened with the costs of physically setting up a fiber network. This issue places the United States in 18th place of OECD countries in terms of coverage. However, in Japan, the leader of fiber-optic coverage, the government provides tax reliefs to companies who set up fiber networks (Frucci, 2007). As a result, Japan is the world's fastest commercially available internet service, while the United States is among some of the worst developed countries in terms of speed and price (Murnane, 2017). The outcome of the model may even influence government policy, specifically in Ohio, to provide financial incentives or relief to those providers setting up a fiber network, if the service has a significant effect on housing values.

Section II is a review of significant literature and study that has already been done on the topic and/or the methodology involving valuation of a housing characteristic. Section III identifies the testable hypothesis and the theoretical model that is to be used. This explanation is critical as no matter how "statistically significant" the results may be, there must be an economic theory that provides validity to the study. Section IV shows the empirical examination of the study. The empirical examination consists of a description of the data used and an analysis of the regression results. Section V concludes the study. The conclusion will summarize the results, identify limitations to the model, explore possible policy implications, and determine how the research can be expanded and/or utilized to further understand the effects.

II. Literature Review

There is limited relevant literature on this specific topic of fiber-optic value to the consumer, but there is research on the value of high-speed networks to consumers. Majority of the papers whose topic is centered around consumer value uses the hedonic pricing method. The hedonic method is essential when evaluating housing values as it allows for easy interpretation of how much the consumer is willing to pay on housing when certain characteristics or amenities exist.

The first literature is one that lays the foundation for an appropriate valuation using hedonic pricing models. It explains that after the housing crisis of 2008, many real estate owners and investment fund managers were interested in the value of real estate assets (Monson, 2009). Monson continues by saying that “market price is a function of each tangible & intangible building characteristic and other outside influencing factors.” The study found that variables such as square footage, number of bedrooms, bathrooms, units within the structure, and age all are appropriate when using the hedonic pricing method on housing values. For this reason, the models used in my study uses each of these key variables. The paper by Monson concludes that hedonic models are important tools in determining the correlation between building characteristics and price. He ends the paper with the notion that hedonic model results may aid in development decisions that generate the highest value of land.

A case study showed that there were significant differences across residential property that is not captured by other approaches (Donaldson, 2009). The researcher used a combination of 31 independent variables to determine which factors influence the price of homes in the Gulf Coast region. The results of the first model were found to be statistically significant but also counterintuitive. For example, the variable ‘bedroom’ was yielding a parameter estimate of -24,240 which means that for each additional bedroom, the price of the house decreases by $24,240. There is an obvious fault in the result as bedrooms are a huge selling point in the real estate market. After further analysis, the problem lied in the multicollinearity of the model. After removing the variables that were highly correlated with each other, the results showed more promise. All the variables had parameter estimates that made logical sense and had higher t-values. The results showed that factors like size, bay-front property, the existence of a pool or fireplace, had significant value to how much consumers are willing to pay for that characteristic. Many of the characteristics from Donaldson's study will be used for this paper's model as well.

Another study estimates the consumer's willingness to pay for internet information usage. The model looks at the valuation customers place on internet information usage based on characteristics of the internet service, whether internet is being used or not, and the level of usage (Lee, 2003). This paper provides an interesting application of the hedonic pricing model by using what is called the Box-Cox non-linear function method. This method normalizes the data in terms of normal distribution and constant variance. This method proved effective as the model shows that variables such as how many register and the speed of internet transmission have a significant value to consumers. The findings show a model of Price = 3.17 + 0.0061SPEED - 0.0042CIRCLE + 0.00075E-SHOP + 0.00064SEARCH – 0.0021HEALTH. The main takeaway for this study is that customers are willing to pay more for higher internet speed networks.

The last piece of literature creates a hedonic price index based on quality and compares the results to the Consumer Price Index (CPI) internet access results. The author understands that a hedonic price index for internet access declines much more than an index that does not account for quality change. The motivation for this specific study was due to the author's observation that most hedonic price index models in the past had observations that consisted largely, or entirely, of dial-up access. The author set out to create an index that accounted for changes in quality of internet access, such as improved bandwidth and length of service contracts (Williams, 2008). The model used in this study was, like the previous, a group of hedonic regression models, including a Box-Cox function. The models produced significant results. The result that relates to this paper's study is the confirmation that the items in the sample show a trend of improvement in service quality in the form of increased bandwidth rates (Williams, 2008).

In this section, the relevant literature has been provided. Using these scholarly journals and case study, I can now produce a solid theoretical model that is backed by economics theory and results. In Section III, the theoretical model is outlined.

III. Theoretical Model

This study attempts to quantify the value of fiber-optic internet access by examining housing values. That is, I test that consumers have a higher willingness to pay for housing that has access to fiber-optic networks. Previous studies have determined that the most appropriate model for testing this hypothesis is the hedonic pricing method. This method will show the effects that fiber-optic access, as well as other physical characteristics, has on consumer housing values. The theoretical model is shown below.

$$\text{Housing Value} = f(\text{Fiber-Optic Access, Physical Characteristics})$$

Most of the physical characteristics variables are considered in the model due to previous literature verifying their significance. The niche of this study lies in the variable of fiber-optic access as a characteristic to housing values. The expected value for this characteristic is positive indicating that access to high-speed networks represent a positive value for consumers when assessing housing value. In Table 1, all the variables used in the study are shown. In the next section, a description of the data used and the applied regression analysis is recorded.

Table 1

Variables	
Yi = housing value	
X1 = fiber-optic service	X10 = plumbing
X2 = house acreage	X11 = hot water
X3 = trailer status	X12 = shower
X4 = farm status	X13 = toilet
X5 = commercial use	X14 = age of structure, decade
X6 = kitchen	X15 = unit sctructure
X7 = sink	X16 = bedrooms
X8 = stove	X17 = trailer status
X9 = rooms	

IV. Empirical Examination

This section provides an in-depth explanation of the data and their respective source. Appropriate and accurate data are necessary to conduct a legitimate analysis. Also, the regression model is provided and explained.

A. Data Description

All the data included in this study was made available by Integrated Public Use Microdata Series USA (IPUMS-USA). IPUMS is the world's largest individual-level population database. IPUMS-USA consists of microdata samples from the United States (usa.ipums.org). This study uses microdata for the year 2015 as that is the last year that the American Community Survey specified the questionnaire to only fiber-optic internet, as opposed to high-speed internet. The survey results from the cities of Akron and Cleveland only are then observed in this study as they are the only cities in Ohio that had complete data. All observations that were reported missing or not available were deleted from the dataset. Dummy variables were created for fiber-optic service, farm status, commercial use, kitchen, stove, sink, plumbing, shower, toilet, and metropolitan status, which gives the value of 1 if the observation has the characteristic, and a value of 0 if not. It is important to note that hedonic pricing models that focus on housing characteristics typically have many variables that are highly correlated to each other. It is important to identify these variables as removing them may improve the model. The descriptive statistics indicate that all of the data is normal and therefore can be used for an accurate model. This data was used in the regression analysis in the next section of this paper, which is why you can also find the expected sign of the variable in the tables shown below. Tables 2 and 3 show the descriptive statistics for Akron and Cleveland alike.

TABLE 2
City of Akron
Descriptive Statistics

Variables	Description	N	Expected Sign	Mean	Minumum	Maximum	Std Dev
VALUEH	value of housing units	775	N/A	124206.58	1000	1469000	123296.93
HOUSEAGE	the age of the structure	775	+	58.67	0	80	20.78
CIFIBER	if home is using a fiber-optic service plan	775	+	0.09	0	1	0.28
FARM	if household is of farm status	775	-	0.002	0	1	0.05
COMUSE	if household is of commercial use	775	+	0.009	0	1	0.09
ACREHOUS	if home located on 10+ acres	775	+	0.005	0	1	0.07
KITCHEN	if household has a kitchen	775	+	0.99	0	1	0.03
SINK	if household has a sink with faucet	775	+	1	0	1	0
STOVE	if household has a stove	775	+	0.99	0	1	0.03
ROOMS	number of rooms in the home	775	+	7.14	1	19	1.88
UNITSSTR	# of housing units within the structure	775	+	3.02	1	4	0.19
BEDROOMS	number of bedrooms in the home	775	+	4.18	0	9	0.92
TRAILER	if household is of trailer status	775	-	0.001	0	1	0.03

TABLE 3
City of Cleveland
Descriptive Statistics

Variables	Description	N	Expected Sign	Mean	Minumum	Maximum	Std Dev
VALUEH	value of housing units	1341	N/A	95338.33	1000	1469000	126645
HOUSEAGE	the age of the structure	1341	+	67.11	0	80	18.45
CIFIBER	if home is using a fiber-optic service plan	1341	+	0.1	0	1	0.3
COMUSE	if household is of commercial use	1341	+	0.01	0	1	0.1
ACREHOUS	if home located on 10+ acres	1341	+	0.002	0	1	0.05
KITCHEN	if household has a kitchen	1341	+	0.99	0	1	0.06
SINK	if household has a sink with faucet	1341	+	0.99	0	1	0.03
STOVE	if household has a stove	1341	+	0.99	0	1	0.06
ROOMS	number of rooms in the home	1341	+	6.97	1	19	2.04
UNITSSTR	# of housing units within the structure	1341	+	3.05	1	4	0.19
BEDROOMS	number of bedrooms in the home	1341	+	4.22	0	9	0.92
TRAILER	if household is of trailer status	1341	-	0.001	0	1	0.03

B. Regression Analysis and Results

The hedonic price regression is identified below. This regression applies the theoretical model to the data made available by IPUMS-USA.

$$Log(VALUEH_{i,c,t}) = \beta_0 + \beta_1 CIFIBER_{i,c,t} + \beta_2 X_{2i,c,t} + \beta_3 X_{3i,c,t} + \beta_4 X_{4i,c,t} \ldots \beta_{17} X_{17i,c,t.} + \varepsilon_i$$

The variable VALUEH represents the housing value of "i" individual's home in city "c". The subscript "t" represents the year of which the survey was conducted. This model looks only at the year 2015. A log-linear regression model is appropriate as it estimates a percentage change in the value of the dependent variable if one unit is added, or exists, to the independent variables. This method is widely accepted to be the best method when evaluating housing values. The log-linear model is the most easily interpreted model in this context. This model fits the typical methodology in terms of structure when using a hedonic price function. The model shows that VALUEH is a combination of the physical characteristics of the home with the variable fiber-optic being the focus of the model.

Tables 2 & 3 below present the regression output for the log-linear hedonic models consisting of the variables fiber-optic service, house acreage, farm status, commercial use, kitchen, sink, stove, rooms, plumbing, hot water, shower, toilet, structure age, unit structure, bedrooms, and trailer status. Some variables may not be present throughout the model due to high multicollinearity of the missing variables.

Table 2

Akron Regression Results

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	12.38	0.43	28.16	<.0001
HOUSEAGE	1	-0.01	0.001	-14.01	<.0001
CIFIBER	1	0.1	0.07	1.43	0.15
FARM	1	0.99	0.6	1.63	0.10
COMUSE	1	-0.09	0.22	-0.41	0.68
ACREHOUS	1	0.01	0.43	-0.99	0.32
ROOMS	1	0.1	0.01	7.1	<.0001
UNITSSTR	1	-0.23	0.13	-1.78	0.07
BEDROOMS	1	-0.02	0.03	-0.87	0.38
TRAILER	1	-1.74	0.65	-2.67	0.00
F Value		37.47	<.0001		
Adj R-Squared		0.3			

Table 3

Cleveland Regression Results

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	11.37	0.27	41.26	<.0001
HOUSEAGE	1	-0.01	0.0009	-11.37	<.0001
CIFIBER	1	0.08	0.05	1.4	0.16
COMUSE	1	0.37	0.16	2.24	0.02
ACREHOUS	1	-0.35	0.32	-1.08	0.28
ROOMS	1	0.07	0.01	6.09	<.0001
UNITSSTR	1	0.06	0.07	0.85	0.39
BEDROOMS	1	-0.04	0.02	-1.58	0.11
TRAILER	1	-2.14	0.49	-4.32	<.0001
F Value		24.55	<.0001		
Adj R-Squared		0.1			

The results of the models show that fiber-optic internet service does serve a role in how consumers value housing that has said service. The results found that a household in Akron with access to fiber-optic internet service has an increased housing value of 10%, and an increase of 8% with houses in Cleveland. With F values of 37.47 and 24.55 at the 5% significance level, this model confirms that the joint effects of all the variables are different from zero. However, due to the low t values and significance levels, we do not have enough confidence to reject the null hypothesis completely. The models in this study show no confidence level over 65%. When running the regression, the variables plumbing, hot water, shower, and toilet were too correlated to accurately evaluate, therefore they were removed from the model. This influenced me to ensure that no other variables could negatively affect the model. I decided to find the variance inflation factor for all of the explanatory variables. This factor determines whether multicollinearity is inflating the standard errors (Donaldson, 2009). To find this factor I used the following equation where R^2 is the r-squared of an independent variable measured as a function of the rest of the independent variables.

$$\text{VIF}\ (\beta_{i,n}) = \frac{1}{1 - R^2}$$

The resulting value is then square-rooted to determine the inflation of the standard error. Using Donaldson's method, a value over 5 would indicate a multicollinearity problem, resulting in a re-evaluation of the model. The VIFs and the standard error relative to their variables are shown in Table 4. As shown, none of the variables remaining in the model exceed the maximum limit for multicollinearity specified above, which means the model remains valid as is. Also, as shown in Tables 2 and 3, 30% and 12% of the variances in housing values are accounted for by the models. These low values would normally be a cause for concern. However, these values make sense as there is a myriad of factors that influence house values, such as environmental amenities or unemployment rates. The major result is the identification of a proper model and insignificant focus variable.

Table 4

Variable	R-Squared	VIF	$\sqrt{VIF}$
fiber-optic service	0.0079	1.0079629	1.003974
age of structure	0.0384	1.0399334	1.019771
house acreage	0.0562	1.0595465	1.029343
farm status	0.055	1.0582011	1.028689
commercial use	0.0014	1.001402	1.000701
kitchen	0.7808	4.5620438	2.135894
sink	0.2465	1.32714	1.152016
stove	0.747	3.9525692	1.988107
trailer	0.3724	1.5933716	1.262288
units in sctructure	0.3689	1.5845349	1.258783
bedrooms	0.4255	1.740644	1.319335
rooms	0.4138	1.7059024	1.306102

V. Conclusions

This study can confirm, with a very low degree of confidence, that consumers in Akron and Cleveland value a house with fiber-optic internet service higher than a house without it. The findings of this model are somewhat inconclusive and cannot yet be used as a definitive key indicator to what consumers value when a service provider is determining where to locate or install their networks. Other physical characteristics were confirmed to have significant effects on housing value that agrees with previous studies, such as age of structure and trailer status, which adds to the already existing literature, thereby confirming the results. The goal of this study was to help the local and city governments of Ohio understand what their consumers value in terms of development. The assumption is that these governments want the area to have access to high-speed internet but do not want to bear the burden of the costs associated. This leaves the cost to the internet service providers who may be unwilling to create a fiber network in the area with no financial support. Understanding that fiber-optic internet is an amenity to housing may influence local governments to provide financial incentives to these service providers for installing a network in their area. The result is the city has higher housing values has higher personal consumption which stimulates the economy of the area. Also, businesses may flourish in terms of technology and consumers will be better off having higher internet speeds and home equity.

Even though the model has some significant results with respect to fiber optic internet, it may be improved in the future by addressing the limitations and providing guidance to future research. One significant limitation is the use of survey data. The dependent variable is the value that a homeowner believes the selling price would be if sold on that day. Typical hedonic models specify the necessity for actual house sales as opposed to housing prices. Another limitation is the general application of hedonic price regressions. For example, the hedonic model and results are typically only valid if the consumer has a perceived understanding of the difference in internet service and their effects on house values. So, if the individual taking the survey is unaware of the positive benefits of fiber-optic internet then the model may not capture the valuation of that amenity. The last major limitation is the fiber-optic variable. In this study, I used fiber-optic access synonymously with the number of people with subscriptions. So, based on this assumption we can't accurately differentiate who has access and who simply doesn't subscribe to it.

This study could be expanded by using more independent variables associated with housing prices such as cost of labor, CPI, GDP, construction costs, etc. Doing this would increase the accuracy of the model and may identify a stronger relationship between fiber-optic internet and housing values. Also, I believe a spatial analysis may be appropriate when considering access to fiber-optic networks. This would require extensive understanding of the fiber-optic mechanism. Things such as maximum distance a network can reach a consumer and the actual number of people within an area who has access to the network would be imperative to a spatial analysis. On one last note, if more observations were observed, maybe over time, then the robustness of the model could possibly show a clearer significance.

References

Miller, Jason. “Federal Spending on Technology Could Reach $95 Billion in 2018.” FederalNewsRadio.com, FederalNewsRadio.com, 26 Dec. 2017, federalnewsradio.com/budget/2017/05/federal-spending-on-technology-could-reach-95-billion-in-2018/.

OECD. “OECD Broadband Statistics Update.” OECD Broadband Statistics Update – OECD, Feb. 2018, www.oecd.org/internet/broadband/broadband-statistics-update.htm

Mackinnon, Jim. “Hudson, Ohio, to Wire Entire City with High-Speed Fiber.” Government Technology: State & Local Government News Articles, Akron Beacon Journal, 23 July 2015, www.govtech.com/dc/articles/Hudson-Ohio-to-Wire-Entire-City-with-High-Speed-Fiber.html.

Anderson, Duane. “Fiber-Optic Internet In the United States at a Glance.” *Broadband Now*, 14 Nov. 2017, broadbandnow.com/Fiber.

Frucci, Adam. “Why Does Japan Get All The Super-Fast Fiber Optic Love?” Gizmodo, Gizmodo.com, 4 Oct. 2007, gizmodo.com/307203/why-does-japan-get-all-the-super-fast-fiber-optic-love.

Murnane, Kevin. “Speedtest Ranks Internet Access Speed In More Than 100 Countries.” Forbes, Forbes Magazine, 12 Sept. 2017, www.forbes.com/sites/kevinmurnane/2017/08/14/speedtest-ranks-internet-access-speed-in-more-than-100-countries/#2825213065b5.

Donaldson, Jeffrey R., and Speros Margetis. “Case Study: Hedonic Price Estimation for Residential Property: Gulf Coast Properties, Inc.” Journal of Financial Education, vol. 35, 2009, pp. 143–159. JSTOR, JSTOR, www.jstor.org/stable/41948630.
Williams, Brendan. “A hedonic model for Internet access service in the Consumer Price Index.” (2008).

Monson, M. (2009). Valuation using hedonic pricing models. *Cornell Real Estate Review*, *7*, 62-73.

Lee, Hyung Seok, Kwangtae Park, and Sang Yong Kim. "Estimation of information value on the Internet: application of hedonic price model." *Electronic Commerce Research and Applications* 2.1 (2003): 73-80.