The University of Nevada, Reno

Three Empirical Essays in Public Economics

A dissertation submitted in partial fulfillment of the

requirements for the degree of Doctor of Philosophy in

Economics

By

Randall M. Chicola

Dr. Frank M. Fossen / Dissertation Advisor

May 2024

THE GRADUATE SCHOOL

We recommend that the dissertation
prepared under our supervision by

RANDALL M. CHICOLA

entitled

Three Empirical Essays in Public Economics

be accepted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

Frank M. Fossen, Ph.D., Co-Advisor
Advisor

Anna Sokolova, Ph.D., Co-Advisor
Co-advisor

Thomas Harris, Ph.D.
Committee Member

Dilek Uz, Ph.D.
Committee Member

Dana Edberg, Ph.D., MBA
Graduate School Representative

Markus Kemmelmeier, Ph.D., Dean
Graduate School

May, 2024

1 Abstract

This dissertation uses applied microeconometric methods to investigate research questions in public economics (first two chapters) and to explore the use of deep learning to construct data useful in public economics and other fields (third chapter). The second and third chapters focus on Nevada. The first essay examines the elasticity response of consumers to the imposition of online sales taxes. The second essay examines the effect of the Nevadan labor force on Nevada county budgets with a particular focus on employees who commute out of their place of residence to work. The third essay uses Clark County Nevada home and property attribute data to infer precinct-level income.

The first paper: *Implications of Sales Tax Enforcement on E-commerce: Evidence from Nielsen Consumer Panel Data* is a solo paper that investigates how sensitive are e-commerce purchases to state sales taxes. It accomplishes this by aggregating Nielsen Consumer Panel data to determine participating panelists' expenditure shares in each panel year and applying a structural regression model underpinned by microeconomic foundations. Sales taxes were not always applied to e-commerce until court cases settled the legal notion of *physical nexus*. As this legal issue was resolved, states started adopting online sales tax policies for e-commerce at different times. This paper exploits timing variation in the imposition of state sales tax onto online purchases to estimate demand elasticities for online and traditional brick-and-mortar retail shopping. The large elasticities imply that collecting sales taxes from online retailers partially shifts consumption back to brick-and-mortar retailers. The policy effects for sales taxes in Nevada are then compared to other Western states.

The second paper: *Joint Prediction and Simulation of Labor Force and Fiscal Conditions of Nevada Counties* uses Nevada county labor and budget data to examine the dynamic relationships between workers who live in their place of work employment and those who commute outside their county of residence for work to determine the fiscal impact on

Nevada county budgets. After the model is specified through a system of equations with interlinked variables for the estimation process, a simulation is performed to assess the impact of changes in exogenous employment on the labor and fiscal status of each Nevada county.

The third chapter: *Constructing Precinct Level Income Variables Using Deep Learning* uses property and home feature data scrapped from the Clark County Assessor's office as a basis for a neural network model that predicts income shares at the precinct level. The contribution of this work is a method for training a neural network model at a given level of geographic granularity, Census blocks, to extrapolate income share estimates at a different granularity at the precinct level. This method is then compared to standard Ordinary Least Squares (OLS) regression by comparing the mean squared error (MSE) of the respective prediction methods.

Dedicated to my parents Linda & Timothy Chicola and little sister Melissa Chicola.

2 Acknowledgements

I would like to thank the faculty at the University of Nevada, Reno who provided exemplary and persistent support in often challenging circumstances for the benefit of my education and so many others. In particular, I would like to thank Frank M. Fossen, Dilek Uz, Thomas Harris, Anna Sokolova, Dana Edberg, Thomas Quint, and so many others in the Economics Department as well as the Department of Mathematics and Statistics at the University of Nevada, Reno. With academic rigor and passion, they tend to the flames of curiosity and knowledge that are too often squelched elsewhere in the world.

Contents

3 List of Tables

4 List of Figures

5 Implications of Sales Tax Enforcement on E-commerce: Evidence from Nielsen Consumer Panel Data

Randall M. Chicola

The University of Nevada, Reno

5.1 Abstract

How sensitive are e-commerce purchases to state sales taxes? Understanding the interplay of online consumer behavior, state sales tax policy changes, and competitive dynamics of online retail consumption relies on empirical elasticity estimates. This paper exploits timing variation in the imposition of state sales tax onto online purchases to estimate demand elasticities for online and traditional brick-and-mortar retail shopping. Elasticities of demand for goods sold online with respect to prices for online goods of 1.2 and 1.23 were found for uncompensated and compensated demands, respectively, when evaluated at their sample mean for the unconstrained preferred demand estimation specification. The large elasticities imply that collecting sales taxes from online retailers shifts consumption partially back to brick-and-mortar retailers. For example, if California did not impose sales on online retailers, online panelists' expenditure shares would be 9.47% larger in this state.[*]

[*]Researcher's own analyses based in part on data from The Nielsen Company (US), LLC and marketing databases provided through the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are those of the researcher and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

5.2 Introduction

5.2.1 Research Question and Motivation

How do online state sales tax changes affect consumer shopping decisions? This research endeavors to better understand retail consumer behavior dynamics under policy shocks to state sales tax policy. More broadly this research may help to better compare state-level shocks to online retail sales. The motivation for pursuing this research question is the intersection of the competing interests of consumer welfare, tax regime application consistency, tax revenues, and competitive fairness in the retail shopping sector. Often products sold online might be the same as at a brick-and-mortar retail store, but the nature of the shopping experience is quite different with online systems offering convenient and streamlined services for pricing, ordering, and delivery. Additionally, overall cost considerations favored online goods which had no sales taxes imposed initially. The question of whether the success of online retailers was due to a fundamentally better shopping experience or an unfair sales tax disadvantage of brick-and-mortar retailers became an issue of public debate. Consequently, determining empirical estimates of elasticities for online goods is an important step in helping parse the sales tax component's contribution to these consumer trends from other competitive factors.

5.2.2 Overview

While e-commerce is currently a mainstay of facilitating everyday transactions, growth to its present state as a principal source of economic activity may be related to utilizing legal technicalities that stem from a once opaque legal notion of *physical nexus* which led to a situation where no sales taxes were collected for online sales. Previous mail-order catalog retailers had touched on the legal issue, but the economic impact of the legal complications was minimal until the meteoric rise of online retail marketplaces. Eventually, states developed the technological infrastructure required for sales tax collection and individually legislated sales tax laws governing online sales between 2008 and 2017 before the Wayfair (2018) case taken up by the Supreme Court in 2018 settled the issue. In this paper, these state adoption dates along with the respective sales tax rates are used together with household purchases with the exception of grocery items to estimate price elasticities of demand for online goods. Non-grocery items are focused on because states often have grocery exemptions to reduce the tax burden on lower-income households. Sales tax-modified prices of non-grocery goods are analyzed using methods developed by Blundell and Robin (1999) and Lecocq and Robin (2015) that expand the capabilities of the demand systems developed by Deaton and Muellbauer (1980) and Banks et al. (1997).

5.2.3 Legislative Background

From the Commerce Clause in Article I of the United States' Constitution, the Dormant Commerce Clause is a legal doctrine established to prevent interstate protectionism and promote interstate commerce. For example, the Dormant Commerce Clause prevents a state from unlawfully adopting legislation governing product safety that applies only to products imported from other states and countries while exempting producers within the state. States attempting to pass legislation discriminating against interstate or international commerce must overcome a significant legal burden of proof to justify a legitimate local purpose for which no other recourse is available. A legal *nexus* describes the type and extent of a connection between a business and a taxing jurisdiction.

HessvIllinois (1967) in 1967 was one of the first cases to address the application of *nexus* to state use taxes on mail-order retailers, with the Supreme Court ruling remittance of sales or use taxes were not required of a seller without any physical presence in that state. The physical presence component of *nexus* was again tested with QuillvND (1992) in 1992 just as fundamental internet infrastructure for e-commerce was beginning to take shape, so it would have been difficult to have sufficient legal foresight that could anticipate the future growth of e-commerce. The Supreme Court's decision forbade states from compelling remote sellers to remit sales tax collections unless they had a physical presence in the state as well as highlighting two pragmatic enforcement challenges, the undue compliance burden on remote sellers among different state laws and technological infrastructure to remit sales taxes data. States organized to address this with the Streamlined Sales and Use Tax Agreement (SSUTA) in 2005.

Retail e-commerce growth continued concurrently with the absence of online sales tax collections. Traditional brick-and-mortar establishments argued that online retailers had an unfair competitive advantage, with small business owners being especially disadvantaged while jurisdictions considered the effects on the labor force as well as forgone sales and property tax revenues. Other legal cases followed until the Wayfair (2018) case where the Supreme Court overturned the physical presence requirement for establishing a jurisdictional nexus in 2018. By this time many online retailers had expanded their physical nexus via distribution centers rendering the issue less consequential. In this paper, to estimate price elasticities of demand for online goods, the variation in the timing of states passing online sales tax and use policies throughout the period from 2001 to 2017 is exogenous to consumers and is utilized in concert with household consumer panel data for the continental United States.

5.3 Literature Review

There are three areas of literature that are particularly relevant. The first relates to the economic analysis of sales tax policy. This literature is reviewed to determine the scope of research involving online consumption and sales tax policy. The second area of literature relates to demand systems and their implementation in applied work. This is reviewed to better understand the econometric framework of demand systems and how underlying structural equations can be estimated with different approaches to compare elasticity estimates for relevant categories of goods. Additionally, demand system literature serves as a platform to discuss how the underlying assumptions of microeconomic theory might be tested by such structural models with recent improvements in software and computing capability. The third area of literature pertains to the intersection of demand systems with a policy such as changes in sales tax policy. Since the emergence of online retail sales is a relatively new change in the nature of economic transactions, even for early adopters of online shopping, most applicable sales-tax-related literature has been published after the year 2000.

Often due to historical necessity, prior demand estimation literature depended on various aggregation levels and collection methods that were available such as expenditure surveys, industry reporting, or aggregated government records which may have been less granular and precise compared to transaction-level household panel data. This analysis contributes to the existing body of literature by using transaction-level direct purchase panel data in conjunction with sales tax data to determine the expenditure shares and price elasticities of household panelists for online goods.

5.3.1 Sales Tax and E-commerce Literature

Much of the prior literature focuses on particular e-commerce vendors such as Einav et al. (2014) who examine eBay data to estimate retail purchase sensitivity. Their work differs in that they employ an item-level empirical method that looks at eBay specifically and uses the “tax table” feature from eBay’s website. While this tax data is relevant for their application, the Nielsen data used in this research necessitates a tax data source that can be applied more generally to the purchases of many online vendors. Similarly, Alm et al. (2010) looks exclusively at eBay seller tax compliance for specific commodities.

Often research is quite product-specific. Ellison and Ellison (2006) examines memory device sales online where they try to analyze internet and traditional commerce channels for substitution effects which would be important for future competition between retail channels. They also consider the de facto sales tax advantage for online retail as well as geographical considerations. The data from Pricewatch.com used in their analysis does not capture all purchase observations, but only from two website listings so traditional retail purchases are unknown.

Other authors such as Goolsbee (2000) estimate the lower bound of sales tax revenue losses due to e-commerce sales from both retail business-to-consumer (B2C) and business-to-business (B2B) e-commerce. Their data comes from the U.S. Census 2006 Annual Retail Trade Report. Since this survey data is aggregated, their fiscal impact is calculated from an estimate of e-commerce taxes due to fewer taxes collected. They use a complex but indirect method of calculating e-commerce sales via regressing e-commerce shipments on GDP and GDP growth. While a laudable approach, it is more advantageous to be able to measure e-commerce expenditures directly, as is possible with the Nielsen consumer panel data.

Ballard and Lee (2007) also empirically tests the effect of retail sales tax rates in local and neighboring counties on Internet purchases from the Current Population Survey between 1997 and 2001. At the time this research was done the nexus issue was of relevance, but this was before the 2018 Wayfair ruling. Ballard and Lee (2007) use a binary dependent variable for whether or not individuals engage in online purchases with sales tax rates as the explanatory variable. The present study differs in that it addresses what Ballard and Lee (2007) wrote would be better in footnote 8, using data that reports how much consumers spent on online purchases. Furthermore, this analysis can look at the expenditure share of online goods relative to the remainder of consumer purchases.

In the legal literature space, there has been a focus on examining larger retailers such as Amazon and on the further examination of the legal paradigm. Gordon (2009), for example, examines "entity isolation" tactics leveraged by online retailers to avoid being forced to collect sales taxes as well as other legal cases and state responses to tax affiliates that were part of entity isolation tactics.

Baugh et al. (2014) investigates the effects of online sales taxes on purchases in comparison to main street retailers for a number of states. This analysis is done midway through the process of states adopting laws and as shown in the first column of Table 1, a large portion of states did not adopt until after their NBER working paper was published. They employ a differences-in-differences approach using household-level data. Their data is transaction-level data from an unnamed financial institution. They note the danger of bias generated by parallel trend assumption, caused by sellers engaging in price discrimination geographically in response to state sales tax initiatives. It may be difficult to determine whether the data from the financial institution is a representative sample of the United States population. Nielsen Consumer Panel data is nationally representative when using their projection factor as a sample weight.

Alm and Melnik (2005) use the special supplement in the Current Population Survey (CPS) published by the U.S. Bureau of Labor Statistics (BLS) as a more robust and representative dataset in comparison to the Forrester Research data used in Goolsbee (2000). They obtain the same qualitative result, higher sales tax yielding a higher probability of an online purchase although the impact elasticity of online purchase probability is quantitative smaller (0.52) compared to (2.3) in Goolsbee (2000). The elasticities estimated in this analysis (approximately 1.2) fall within the range of these two publications.

Other literature focuses more on state revenue losses from e-commerce sales prior to legislative amelioration for brick-and-mortar establishments. Bruce and Fox (2001) estimate revenue for 2001, 2006, and 2011. They predict a $16 billion revenue loss for state and local governments due to the impact of a narrowing tax base. The policy implications for sales tax rates would include future increases that would need to be imposed by state and local governments to supplement budgetary shortfalls from lost brick-and-mortar retail sales tax revenues.

5.3.2 Demand Estimation Literature

The primary variable of interest is the household expenditure share of online shopping goods. Since the data available includes the consumer's state of residence, the hypothesis is that for panelist households residing in a particular state that collects sales tax for online purchases, the expenditure share of online spending as a proportion of their total recorded expenditure would decrease as consumers adjust to the sales tax regime being enforced onto their online purchases. The demand estimation literature is foundational to this analysis, and this literature spans a wide breadth of economic disciplines such as agriculture economics, energy, and industrial organization.

This literature often uses demand estimation techniques because it is amenable for numerous policy applications. For example, demand estimation can address challenges of consumption inequality presented by Attanasio and Pistaferri (2016) as done by DeDad (2019). Among the various demand estimation methods, the first to be proposed was the Almost Ideal Demand System (A.I.D.S) by Deaton and Muellbauer (1980). This demand estimation framework has been used in numerous applied areas and to numerous product categories, such as the Japanese (Jing et al., 2004) and the United States' (Gallet, 2010) meat markets, Italian tobacco (Jones and Mazzi, 1996), Saudi honey imports (Alnafissa and Alderiny, 2020) and groups of consumer goods Hausman et al. (1994).

After the inception of (A.I.D.S), a number of different modifications have been proposed. The first of which extends the robustness of the (A.I.D.S) by including demographic characteristics and was first proposed by Deaton and Muellbauer (1980) though data limitations prevented them from implementing the idea at that time and was later pursued in Subramanian and Deaton (1996). A commonly used modification is the addition of a quadratic term, proposed by Banks et al. (1997) which is known as Quadratic Almost Ideal

Demand System (Q.U.A.I.D.S) and will be implemented for this analysis.

Another modification includes using consumer taste preferences among different regions as aptly demonstrated by Atkin (2013) for regional tastes in India as well as by De Sousa et al. (2018) for regional consumption convergence in French dairy markets. Other modifications incorporate ideological differences as measured by tastes as in DeDad (2019). Finally, some literature involved the technical programming aspects of (A.I.D.S) such as Poi (2002), Poi (2008), Poi (2012), and Lecocq and Robin (2015).

5.3.3 Demand Systems Intersected with Policy

Some tax literature connects policy analysis to consumer demand systems. Brännlund and Nordström (2004) uses demand estimation to analyze consumer responses after environmental policy changes. Tax policy is also of particular interest in application to using demand system estimation methods. Madden (1996) examines the sensitive nature of the demand system that underlies a tax proposal by comparing four different demand estimation methods. Similarly, using Belgian consumer survey and national account data, Decoster and Schokkaert (1990) compare the effects of indirect taxes using price elasticities generated by different demand systems.

Walls and Ashenfarb (2022) also look at policy effects using a two-stage QUAIDS for recreation goods and then simulating a hypothetical 5% tax increase to determine potential tax revenues that can help fund the use of public lands. Taxable items that broaden this base such as hunting and fishing fees and taxes on related sporting goods are some of the potential tax channels that could better match public goods use to the industries that most directly benefit from their use. This work is related in that it examines tax policy using a demand system but differs in that its primary focus is on the public land that is complementary to the particular products being taxed.

5.4 Data

5.4.1 Overview

The primary sources of data are the Nielsen Consumer Panel data and the sales tax data from the Tax Foundation and applicable state departments of revenue. The period under examination is from 2011 to 2017 for the continental United States because 2011 is the earliest year available for the combined tax rate data. The consumer panel data does not include Hawaii or Alaska. Overall, in a panel year over the observation period, in the Nielson data, there are about 60,000 panelists who make approximately 9 million trips yielding 50 to 60 million purchase item observations. About 40% of the panelists in a particular panel year had made at least one online trip. The analysis will focus on the intensive margin of online consumption by these consumers who will be referred to as "Online Shoppers" and exclude common sales tax-exempt items such as groceries when possible.

5.4.2 Nielsen Consumer Panel Dataset

Household non-grocery purchases are used from the Consumer Panel Dataset by the Nielsen Company made available by the University of Chicago's Kilts Center Archive Nielsen (2019).

A smartphone application or scanner is provided by Nielsen to household participants in order to record their transactions. The dataset is nationally representative of the demographic composition of U.S. Census estimates, once sample weights provided in the data are used. Universal Product Code (UPC) level purchase data of trips taken by participating panelists are aggregated over the panel year to calculate $w_{i,h,t}$, expenditure shares for i (online and traditional) aggregated goods of household h in a given year t. Additional household demographic characteristics, product attributes, and hierarchical levels of product categorical granularity, the broadest of which are called "departments", are features of the dataset. The type of retail channel is recorded for trip purchases and is important for this analysis as one such retail channel denotes online purchases.

5.4.3 Sales Tax Data

Sales tax data is compiled for each state and procured from the Department of Revenue of applicable states. Additionally, combined sales tax rates which are defined as the state plus average local sales tax rates, available for the 2011 to 2020 period were provided by the Tax Foundation as with Fritts (2020) and are presented in Table 26 within the appendix. The associated combined state and local sales tax rate is applied to the household purchases if sold in a state that collects sales tax and if the purchase was a traditional brick-and-mortar purchase.

State and local sales tax rate figures are applied to online prices according to whether the date of purchase occurred after states adopted online sales tax policies and began collecting online sales taxes. Thus, the legislated adoption year is used for applying sales taxes to applicable purchases from the consumer panel data. Though it may be argued that consumers may anticipate incoming sales taxes upon first public knowledge of legislative intent, legislative outcomes were uncertain as well as when the information was known to all consumers. As the legal case history for many states elucidates, it is also likely knowledgeable consumers would postpone their consumption response until a final court determination was made.

There are many details that vary by state as to the types of goods and services that are taxable or exempt organizations such as charities. While it seems to be reasonable to assume none of the panelists as households qualify for such exemptions, the different products under a state sales tax regime are more troublesome. Prescription medication, non-prescription medication, and groceries are the most common exempt categories, where 32 of the states exempt groceries outright, and of those that tax groceries, some tax at a lower rate or have exemption exclusions stipulated for certain products like candy and soda. To avoid an intractable application of state-level grocery sales tax conditions to millions of product purchases, grocery items were excluded from the demand system.

5.5 Descriptive Statistics

Figure 1 uses U.S. Census data via FRED to give external figures for the share of e-commerce in total retail sales. This figure documents a very strong growth in online retail sales over the period of analysis.

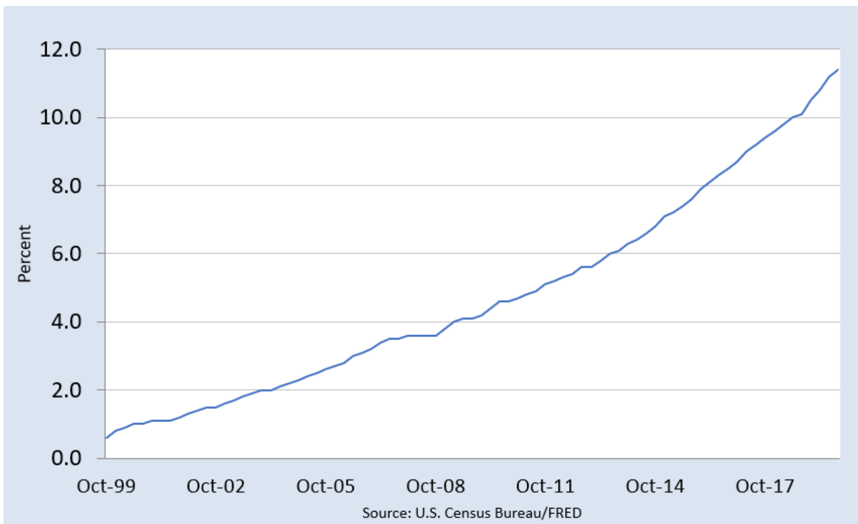


Figure 1: E-Commerce Retail Sales as a Percent of Total Sales

5.5.1 Variation in Sales Tax Rates

There appears to be promising variation for elasticity estimation purposes in sales tax rates. A key source of variation is the adoption year of sales taxes for online purchases that vary across states. Table 1 displays the year in which a state adopted an online sales tax regime. There is considerable variation in the adoption years, and importantly, the adoption years do not follow a clear geographical or otherwise systematic pattern. This variation likely contributes to the state trends in online expenditure shares. For example, in California, the online expenditure share increased by 11% from 4.77% in 2011 to 5.28% in 2017. Meanwhile, its neighbor Nevada had a mean online expenditure share of 4.14% in 2011 which increased by only 4% to about 4.32% by 2017. The smaller increase in the online sales share in Nevada may be partially due to the adoption of sales taxes for online sales in Nevada in 2014, whereas the regime change was led by California in 2012. Moreover, there is variation in general sales tax rates across states and time. The primary source of this variation is standard legislative changes in the level of both the state and local sales tax rates over time. As shown in the appendix, these changes include not only different magnitude increases in sales tax rates but also some situations where one state decreased its sales tax rate while its neighbor increased its sales tax rate. For example, in 2011, California had a combined sales tax rate of 9.08% which decreased to 8.48% by 2017. In contrast, Nevada's combined sales tax rate increased from 7.96% in 2011 to 8.14% by 2017.

Given the general increase in the online expenditure share shown in Figure 1 during the same period when states also implemented online sales taxes, one might suspect that online sales are not very price elastic. However, to separate shifting consumer preferences for online consumption from the effects of sales taxes, an econometric analysis is needed. The tax elasticities can be estimated by exploiting the variation of sales tax rates between states in concert with their online sales tax policy adoption years.

5.5.2 Panelist Household Profile

With some descriptive statistics, we can outline the characteristics that describe the Nielsen panelist households. These descriptive statistics are partitioned by variables relating to the econometric structure of QUAIDS demand system, head of household characteristics, and general household characteristics. Table 2 displays the price and expenditure variables that are canonical components of QUAIDS demand systems as well as the nominal expenditures for both categories and total expenditures. Mean online expenditures are $249.50 per year

Table 1: State Online Sales Tax Adoption Year

State	Adoption Year	State	Adoption Year	State	Adoption Year	State	Adoption Year
AL	2016	KY	2005	NJ	2013	VA	2017
AR	2011	LA	2017	NM	2017	VT	2013
AZ	2013	MA	2013	NV	2014	WA	2013
CA	2012	MD	2013	NY	2017	WI	2013
CO	2016	ME	2017	OH	2015	WV	2008
CT	2013	MI	2015	OK	2017	WY	2017
DE	None	MN	2014	OR	None		
FL	2014	MO	2017	PA	2012		
GA	2013	MS	2017	RI	2017		
IA	2017	MT	None	SC	2016		
ID	2017	NC	2014	SD	2017		
IL	2015	ND	2001	TN	2014		
IN	2014	NE	2017	TX	2012		
KS	2005	NH	None	UT	2017		

for the period which corresponds to about 4.5% share of goods purchased.

Table 2: QUAIDS Demand System Variable Descriptive Statistics

Variable	Mean	Std.Dev.
Online Expend. Share	4.5%	8.9%
Trad. Expend. Share	95.5%	8.8%
Med. Online PUC w/ Tax	$4.57	$1.24
Med. Trad. PUC w/ Tax	$2.70	$0.18
Online Expenditures	$250	$574
Trad. Expenditures	$5,976	$3,703
Total Expenditures	$6,224	$3,792

*PUC is the Per Unit Cost. Median state PUC has state sales tax rates applied. Expenditures are US dollars per year in prices of online and traditional goods.

It is important to note that these figures are aggregated from prices of individual items which were adjusted for their tax treatment into a state-level median. The sample means displayed are consequently a mean of state median prices. Household characteristics are captured in Table 3 which display the mean for the demographic binary categories used in the demand system. The categories were re-coded to a more general categorization.

This analysis uses the male head of household figures for age, education, and occupation variables unless not present. To then better delineate whether the primary head of the household was male or female, a “Head of Household is Male” binary was added.

Additionally, to better display the distributions within the sample, the categorical variable descriptions except for age are presented as binaries to display percentages of each variable group. The Table 27 displays the original groupings of categories.

It can be seen in Table 3 that the sample tends to lean toward a two-member household (42.5%) with a white (79.82%) male (73.11%) head of household in a one-family domicile (78.4%) making between 45 to 70 thousand dollars (24.54%) with no children under 18 (78.41%). While most household heads are employed over 35 hours a week, there is a significant portion not employed for pay (35.66%). These descriptive statistics seem to indicate a strong case for controlling for these variables in the regressions in order to obtain consistent estimates.

Table 3: Household (HH) Demographic Variables Descriptive Statistics (In Percent Except for Age)

Variable	Mean
Household Size	
One Member	25.44
Two Members	42.50
Three Members	14.25
Four Members	10.96
Five or More Members	6.85
Domicile/Residence Type	
One Family Home	78.40
Two Family Home	4.07
Three or More Family Home	13.41
Mobile Home or Trailer	4.12
Income Bracket	
Under 15k	5.96
15k - 29.9k	13.91
30k - 44.9k	17.04
45k - 69.9k	24.54
70k - 99.9k	20.68
100k+	17.87
Age and Presence of Children	
Under 6 Only	3.48
A Least One Child 6 to 17	18.11
No children under 18	78.41
Head of Household Employment Hours	

Under 35 hrs/week	11.31
Over 35	53.03
Not Employed for Pay	35.66
Head of Household Occupation	
White Collar job categories	41.78
Blue Collar job categories	21.58
Military;Student, employed; Other	36.64
Race	
White	79.82
Black	11.57
Asian	3.93
Other	4.67
Ethnicity Hispanic	6.27
Head of Household is Male	73.11
Married	62.69
Head of Household Age	56.84
148,363 Observations	

Table 4 displays the distribution of age ranges for heads of household. It shows many households are within the prime working age although there is a heavier representation in the 55 to 64 years and over 65 years old brackets. This may be indicative of the self-selection of panelists who may be retired and more inclined to participate in using the Nielsen scanner. For the final specification of the regression, the integer age variable was used rather than the age brackets to reduce multicollinearity issues.

In Table 5 we see the distribution of hours worked by the male and female head of

Table 4: Male and Female Head of Household Age

	Female Head Age										
Male Head Age	No Head	Under 25	25-29	30-34	35-39	40-44	45-49	50-54	55-64	65+	Total
No Head	-	190	755	1,563	2,249	2,669	3,684	5,181	12,207	11,390	39,888
Under 25	64	114	45	20	8	10	9	11	12	5	298
25-29	305	172	1,100	331	67	21	14	16	38	10	2,074
30-34	581	41	1,008	2,596	626	110	46	35	45	20	5,108
35-39	764	10	241	1,829	3,016	771	194	77	85	28	7,015
40-44	1,078	10	96	548	2,139	3,437	1,093	320	146	93	8,960
45-49	1,679	7	43	182	713	2,465	4,014	1,308	522	138	11,071
50-54	2,289	6	29	85	257	925	3,075	5,135	2,023	219	14,043
55-64	4,754	13	37	59	149	425	1,600	5,385	16,991	1,755	31,168
65+	4,039	13	6	32	49	76	230	770	7,366	16,157	28,738
Total	15,553	576	3,360	7,245	9,273	10,909	13,959	18,238	39,435	29,815	148,363

household. We can see from the totals that many panelists are not employed for pay which may allow a head of household more time to participate as a panelist. Of the heads of household who do work, working over 35 hours is the most prevalent. Employment hours categories are included in the regression as control variables.

Table 5: Head of Household Hours of Employment

	Female Head Employment					
Male Head Employment	No Head	Under 30 hours	30-34 hours	35+ hours	Not Employed for Pay	Total
No Head	-	4,462	1,790	16,884	16,752	39,888
Under 30 hours	1,289	1,261	290	1,928	2,178	6,946
30-34 hours	610	503	333	1,066	1,074	3,586
35+ hours	7,514	8,904	3,071	22,171	20,128	61,788
Not Employed for Pay	6,140	2,656	907	6,393	20,059	36,155
Total	15,553	17,786	6,391	48,442	60,191	148,363

5.5.3 Trends in Online Consumption

Figure 2: Nielsen Consumer Panel Extensive Margin Percentage of 1 Minimum Online Shopping Trip Per Year

The extensive margin criterion, having at least one online shopping trip in the panel year, is important to parse the relevant data for further analysis. Figure 2 displays the percentage of panelists who had a minimum of one online shopping trip per year. It can be seen the growth trend in online shopping has been modest but positive. Just because panelists made at least one online trip, it does not tell us how often online shoppers make online trips. Consequently, another question that might be asked is: "How often do panelists go on online shopping trips compared to traditional brick-and-mortar establishments per year?"

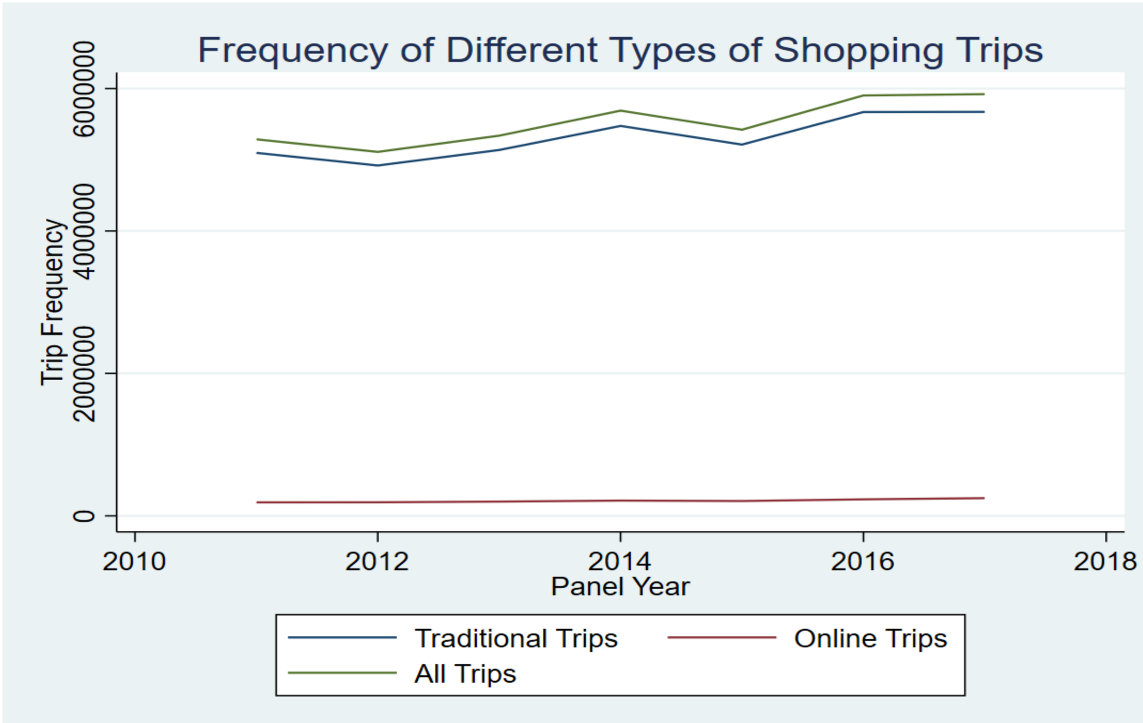


Figure 3: Nielsen Consumer Panel Trip Frequencies Per Year

Figure 3 shows that online shopping is a relatively small proportion of shopping trips, which is due to a majority of the trips going to grocery stores and discount stores. Figure 4 shows the year-over-year (YoY) changes in the number of shopping trips by type. It demonstrates that online shopping trips overall have a higher growth rate, or as in 2015, incur a milder downturn than their traditional counterparts.

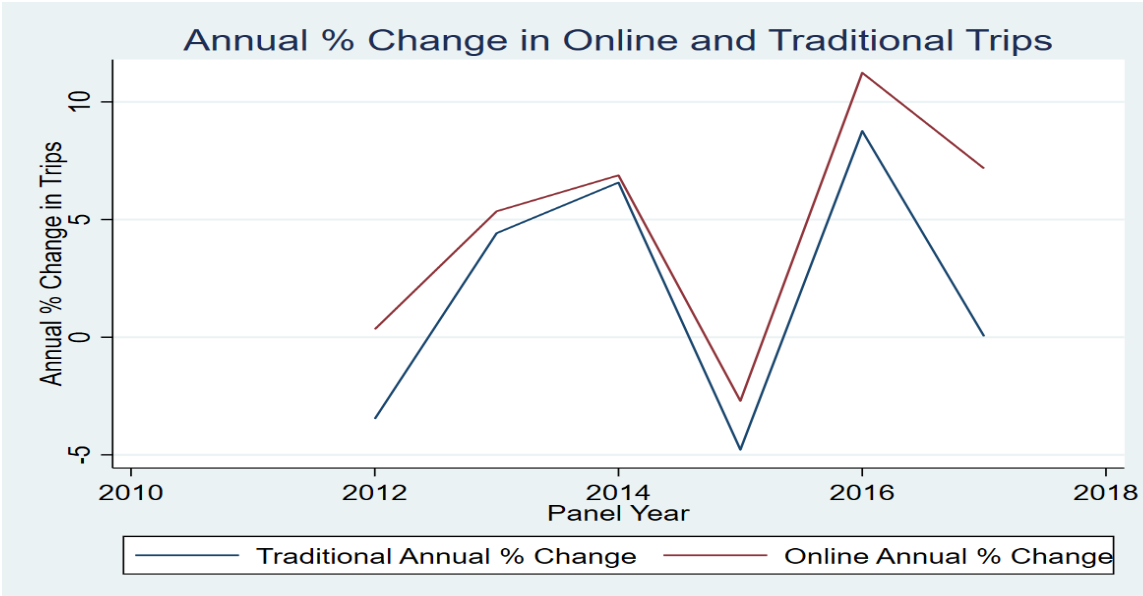


Figure 4: Nielsen Consumer Panel Shopping Trip Year-over-year Change

Trips while useful, do not tell us everything we might like to know about online shopping trends. The YoY change in the value of products sold online is provided in Figure 5. Figure 5 demonstrates similar growth patterns in online expenditures as Figure 4, including the resilience of online sales to expenditure contraction in 2015.

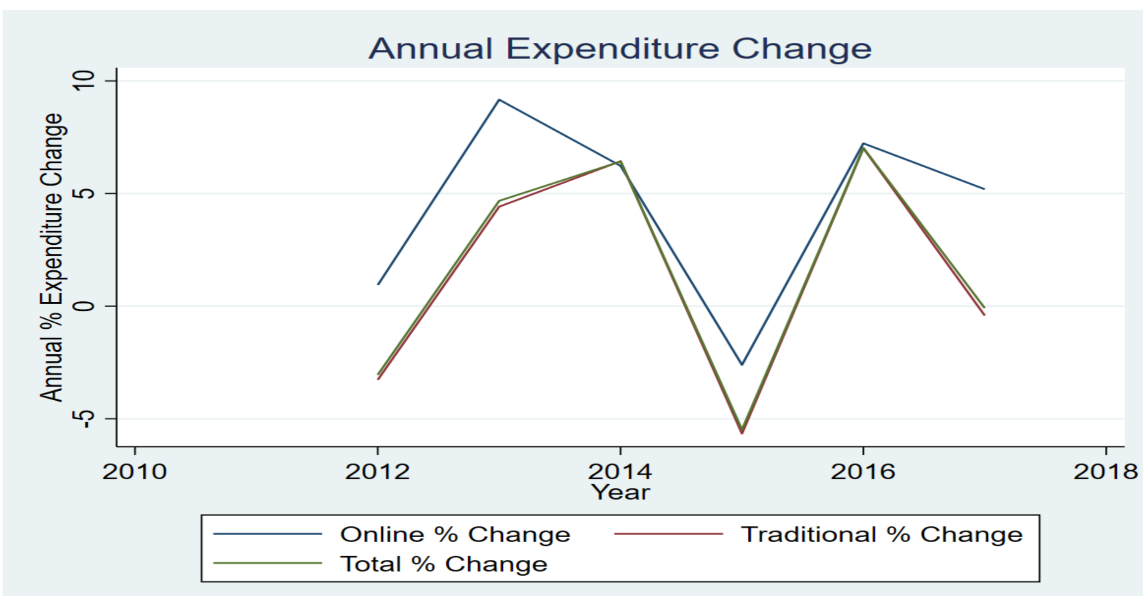


Figure 5: Nielsen Consumer Panel Online and Traditional Expenditure Change

Finally, Table 6 shows for the 2011 to 2017 period the number and percent of purchased items bought by panelists who made none or at least one online trip in the given panel year. It can be seen that for this set of panelists, online purchased items increases over time as one might expect with the increasing prevalence of online shopping over the period with a similar trend as when looking at trips.

Table 6: Panelists Who Made None or at Least One Online Purchase in Panel Year

Panel Year	None	One or More	Total	None (%)	One or More (%)
2011	44,544,026	21,777,822	66,321,848	67.16%	32.84%
2012	41,461,998	21,490,273	62,952,271	65.86%	34.14%
2013	40,077,286	23,065,539	63,142,825	63.47%	36.53%
2014	39,884,696	24,832,424	64,717,120	61.63%	38.37%
2015	39,678,860	24,111,887	63,790,747	62.20%	37.80%
2016	41,326,261	26,441,125	67,767,386	60.98%	39.02%
2017	39,415,929	27,038,277	66,454,206	59.31%	40.69%

5.6 Conceptual Framework and Methodology

Following Lecocq and Robin (2015), the analysis uses Stone's Price index (Stone (1954)), proper linearized, and quadratic variants of the Almost Ideal Demand System under unconstrained as well as price parameter homogeneity constrained circumstances. The quadratic variant is the preferred specification. Chi-squared testing for Slutsky symmetry was performed after the demand estimation. The demand system's approach assumes the household optimizes its resource allocation between online and traditional goods. The expenditure shares calculated from the panel data are akin to disposable expenditures as not all household expenditures are recorded in the data (e.g., rent & utilities).

Expenditure shares $w_{i,h,t}$, are defined as the amount spent on a good relative to the expenditure budget $x_{h,t}$ available $w_{i,h,t} = \frac{p_{i,h,t} * q_{i,h,t}}{x_{h,t}}$, where $p_{i,h,t}$ and $q_{i,h,t}$ are the price and quantity of good i purchased by household h in year t and are estimated following Banks et al. (1997). For expenditure period t, household h's expenditure share function for good i is specified as:

$$w_{i,h,t} = \alpha_{i,h,t} + \gamma_{\mathbf{i,t}}^{'}\mathbf{p}_{s,t} + \beta_{i,t}\{x_{h,t} - a(\mathbf{p}_{s,t},\theta)\} + \lambda_{i,t}\frac{\{x_{h,t} - a(\mathbf{p}_s,\theta)\}^2}{b(\mathbf{p}_s,\theta)} + u_{i,h,t} \tag{1}$$

where $\mathbf{p}_{s,t}$ is a log median price N-vector indexed by $i \in 1,2$ for online and traditional goods which includes sales taxes where applicable. That is $\mathbf{p}_{s,t}$ is comprised of $p_{i,s,t}$ which is the log median price of the good in state jurisdiction s for year t, otherwise denoted $\mathbf{p}_{s,t} \in \{p_{on,s,t}\ ,\ p_{tr,s,t}\}$. θ is the set of all parameters while $x_{h,t}$ is log total expenditure of household h in year t, and $u_{i,h,t}$ is an error term.

Additionally, nonlinear price aggregators are:

$$a(\mathbf{p}_{s,t},\theta) = \alpha_0 + \alpha^{'}\mathbf{p}_{s,t} + \frac{1}{2}\mathbf{p'}_{s,t}\mathbf{\Gamma}\mathbf{p}_{s,t} \tag{2}$$

$$b(\mathbf{p}_{s,t},\theta) = \exp(\beta^{'}\mathbf{p}_{s,t}) \tag{3}$$

where $\alpha_t = (\alpha_1,...,\alpha_N)_t^{'}, \beta_t = (\beta_1,...,\beta_N)_t^{'}, \mathbf{\Gamma}_t = (\gamma_1,...,\gamma_N)_t$. α_t, β_t, and $\mathbf{\Gamma}_t$ must satisfy the additivity ($\sum_{i=1}^n \alpha_i = 1$ and $\sum_{i=1}^n \beta_i = 0$), homogeneity ($\sum_{i=1}^n \gamma_{i,j} = 0$ and $\sum_{i=1}^n \lambda_i = 0$), and Slutsky symmetry ($\gamma_{i,j} = \gamma_{j,i}$) conditions for a canonical almost ideal demand system, though the unconstrained model is also examined. Demographic variables use the translating approach presented first by Pollak and Wales (1981) by entering the demand system via the α's whereby $\alpha_{h,t} = \alpha_{\mathbf{i,t}}^{'}\mathbf{s}_{h,t}$ and $\mathbf{s}_{h,t}$ is the set of household demographic variables observed in the panel data.

5.7 Empirical Results

5.7.1 Unconstrained and Homogeneity Constrained Regression Results

To begin the analysis, pooled results from the panel period are used to examine whether the timing of a state's implementation of sales tax enforcement affected the price (including sales tax) and thereby the number of goods sold online. After initial unconstrained results are obtained, the homogeneity constraint is imposed on the model. If the results are robust when imposing these conditions, then the chi-squared test results are used to determine whether symmetry holds for the dataset.

Table 7 presents the Q.U.A.I.D.S results for the expenditure share of online goods. The residual expenditure share is spent on goods bought at traditional brick and mortar retail[†]. The γ coefficients for both price parameters are statistically significant with signs that would be expected. An increase in the log median price of online goods corresponds to a decrease in online expenditure shares while an increase in the log median price of traditional goods corresponds to an increase in the expenditure share of online goods as expected for traditional and online goods baskets that are substitutable.

Table 7: Unconstrained and Homogeneity Constrained QUAIDS for Online Goods

Dep. Var. Online Share $(w_{2,h,t})$	**Unconstrained**	**Homogeneity**
$\gamma_{Ln\,Median\,Online\,PUC\,w/tax}$	-0.122***	-0.121***
	(0.00573)	(0.00572)
$\gamma_{Ln\,Median\,Trad\,PUC\,w/tax}$	0.134***	0.121***
	(0.00699)	(0.00572)
β_{lnx}	-0.175***	-0.175***

Continued on next page

[†]Using the traditional expenditure share as the dependent variable leads to identical estimates with flipped signs of the coefficients.

Table 7 – continued from previous page

	Unconstrained	**Homogeneity**
	(0.00457)	(0.00457)
λ_{lnx^2}	0.0104***	0.0104***
	(0.000284)	(0.000284)
Head of Household is Male (Y/N)	0.00392***	0.00392***
	(0.000745)	(0.000745)
Married (Y/N)	-0.00745***	-0.00753***
	(0.000838)	(0.000837)
Head of Household Age	0.000120***	0.000124***
	(0.0000221)	(0.0000221)
Household size (omitted base category: Four members):		
One Member	0.00969***	0.00963***
	(0.00123)	(0.00123)
Two Members	0.000457	0.000417
	(0.000971)	(0.000971)
Three Members	0.000287	0.000276
	(0.000939)	(0.000939)
Five or More Members	-0.00161	-0.00163
	(0.00110)	(0.00110)
Domicile Type (omitted base category: One Family Domicile):		
Two Family Domicile	0.00405***	0.00418***
	(0.00114)	(0.00114)
Three or more Family Domicile	0.00855***	0.00871***
	(0.000695)	(0.000694)

Continued on next page

Table 7 – continued from previous page

	Unconstrained	**Homogeneity**
Mobile,Trailer, or Unreported	0.00420***	0.00429***
	(0.00115)	(0.00115)
Income Bracket (omitted base category: 45k to 69.6k) :		
Under 15k	-0.00112	-0.00125
	(0.00109)	(0.00109)
15k - 29.9k	-0.000553	-0.000658
	(0.000793)	(0.000792)
30k - 44.9k	-0.00185**	-0.00190**
	(0.000710)	(0.000710)
70k - 99.9k	-0.00116	-0.00110
	(0.000672)	(0.000671)
100k+	0.00311***	0.00328***
	(0.000727)	(0.000725)
Age and Presence of Children (omitted base category: Under 6 Only):		
A Least One Child 6 to 17	-0.00213	-0.00215
	(0.00133)	(0.00133)
No children under 18	0.00222	0.00223
	(0.00139)	(0.00139)
Employment Hours (omitted base category: Over 35):		
Under 35 hrs/week	0.000167	0.000222
	(0.000771)	(0.000770)
Not Employed for Pay	0.0111***	0.0112***
	(0.00237)	(0.00237)

Continued on next page

Table 7 – continued from previous page

	Unconstrained	Homogeneity
M/F HH Education (omitted base category: Some college):		
HS or less	-0.00125*	-0.00127*
	(0.000613)	(0.000613)
College graduate or post-graduate	-0.000279	-0.000304
	(0.000553)	(0.000553)
M/F Occupation (omitted base category: White Collar job categories):		
Blue Collar job categories	-0.000711	-0.000721
	(0.000640)	(0.000640)
Military or Student, or Non-employed	-0.00579*	-0.00585*
	(0.00230)	(0.00230)
Race (omitted base category: White):		
Black	0.00306***	0.00310***
	(0.000713)	(0.000713)
Asian	0.00723***	0.00768***
	(0.00118)	(0.00117)
Other	0.00432***	0.00447***
	(0.00114)	(0.00114)
Hispanic (ethnicity)	0.000381	0.000656
	(0.00100)	(0.001000)
Constant	0.750***	0.761***
	(0.0187)	(0.0184)

‡*,**,*** significant at the 10%,5%,1% level

5.7.2 Uncompensated and Compensated Own-price, Cross-price, and Budget Elasticities

The QUAIDS demand system connects microeconomic foundations of utility maximization that connect the cost or expenditure function that describes the cost of subsistence $a(\mathbf{p_s}, \theta)$ or bliss $b(\mathbf{p_s}, \theta)$. While the gamma price coefficients found within the budget share equation are informative, they are complemented by their respective elasticity calculations. The budget or expenditure elasticity, as well as uncompensated and compensated price elasticities for two comparison goods i and j, are derived by differentiating the household budget share equation with respect to x_h and $p_{j,s,t}$ giving:

$$\mu_{i,h,t} = \beta_{i,t} + 2\lambda_{i,t}\frac{x_{h,t} - a(\mathbf{p_s}, \theta)}{b(\mathbf{p_s}, \theta)} \tag{4}$$

$$\mu_{i,j,h,t} = \gamma_{i,j,t} - \mu_{i,h,t}(\alpha_{\mathbf{j,h,t}} + \gamma_{\mathbf{j,t}}\mathbf{p}) - \lambda_{i,t}\beta_{i,t}\frac{x_{h,t} - a(\mathbf{p_s}, \theta)^2}{b(\mathbf{p_s}, \theta)} \tag{5}$$

In turn, from these partial derivatives, expenditure, uncompensated, and compensated elasticities are given by:

$$e_{i,h,t} = \frac{\mu_{i,h,t}}{w_{i,h,t}} + 1 \tag{6}$$

$$e^u_{i,j,h,t} = \frac{\mu_{i,j,h,t}}{w_{i,h,t}} - \delta_{i,j,h} \tag{7}$$

$$e^c_{i,j,h,t} = e^u_{i,j,h,t} + e_i w_{j,h,t} \tag{8}$$

where $\delta_{i,j,h}$ is the Kronecker delta.

This specification for elasticity follows the QUAIDS model of Banks et al. (1997). Furthermore, it employs the computationally advantageous methodology described by Blundell and Robin (1999) by utilizing commands later developed by Lecocq and Robin (2015) for the almost ideal demand system iterated linear least-squares (AIDSILLS). It allows regression results generated by the expenditure share equation to calculate elasticities in the typical interpretation of the percentage change in quantity with respect to a percentage change in price.

Table 8 presents the uncompensated and compensated price elasticities for the unconstrained model. The rows denote the prices of online and traditional goods, $\gamma_{\mathbf{p_{Online,t}}}$ and $\gamma_{\mathbf{p_{Trad,t}}}$, respectively. Columns one and three display the corresponding own and cross-price elasticities derived from the associated expenditure shares. Columns two and four report the associated standard errors. The absolute magnitude of the uncompensated own-price elasticity for online goods is 1.260 for the unconstrained model (Table 8). The compensated own-price elasticity for online goods for the unconstrained model is 1.231. These elasticities indicate online consumer behavior among Nielsen panelists is quite sensitive to price changes in online goods that may be induced by the imposition of state sales taxes. For the homogeneity-constrained model, elasticities presented in (Table 9) demonstrate similar results with uncompensated and compensated elasticities of 1.236 and 1.207, respectively.

Table 8: Unconstrained Own and Cross-Price Elasticities

	w_{Online}	(se)	$w_{Traditional}$	(se)
Uncompensated				
Online	-1.260***	0.025	0.011***	0.001
Traditional	0.830***	0.102	-1.035***	0.004
Compensated				
Online	-1.231***	0.025	0.052***	0.001
Traditional	1.518***	0.103	-0.065***	0.004

* p<0.05,** p<0.01, ***p<0.001

Median per unit cost with state sales tax used. Columns are the elasticity of the quantity of online goods derived from expenditure shares. The corresponding standard errors are reported in the row for the respective price

Table 9: Homogeneity Constrained Own and Cross-Price Elasticities

	w_{Online}	(se)	$w_{Traditional}$	(se)
Uncompensated				
Online	-1.236***	0.023	0.010***	0.001
Traditional	0.522***	0.025	-1.022***	0.001
Compensated				
Online	-1.207***	0.023	-0.051***	0.001
Traditional	1.207***	0.023	-0.052***	0.001

* $p<0.05$,** $p<0.01$, ***$p<0.001$

Median per unit cost with state sales tax used. Columns are the elasticity of the quantity of online goods derived from expenditure shares. The corresponding standard errors are reported in the row for the respective price

Traditional goods are less elastic than online goods. The estimated unconstrained own-price elasticity ranges between magnitudes of 1.035 for uncompensated to 0.065 for compensated. The homogeneity-constrained own-price elasticity for traditional goods ranges from magnitudes of 1.022 for uncompensated to .052 for compensated. Notably, unlike the estimate for online goods, the estimate for traditional goods differs between uncompensated and compensated elasticities. It goes from slightly elastic for uncompensated to quite an inelastic range of 0.065 to 0.052 in the unconstrained and homogeneity-constrained cases for compensated own-price elasticity.

Additionally, the tables present cross-price elasticities. This indicates whether an online goods basket may be sensitive to price changes in a traditional brick-and-mortar goods basket as indicative of a substitute goods basket. The cross-price elasticity for online goods with respect to traditional goods basket prices ranges between 0.830 with uncompensated and 1.518 with compensated demand in the unconstrained model. For the homogeneity-constrained model, the estimate varies between 0.522 and 1.207 for uncompensated and compensated respectively. Traditional expenditure shares on the other hand are consistently inelastic with respect to changes in online prices.

The budget elasticities for the unconstrained and homogeneity-constrained models are presented in Table 10 below. The expenditure share of online goods is inelastic with respect to the consumer budget in both cases at 0.717 and 0.714, respectively. If a 10% increase in budget corresponds to about a 7% increase in demand for online goods, this elucidates the differential between the uncompensated and compensated demand price elasticities for online expenditure shares in Table 8 and Table 9. In both the unconstrained and homogeneity-constrained tables the compensated online own-price elasticity is less elastic than the uncompensated elasticity.

Table 10: Budget Elasticities

	Unconstrained	Homogeneity Constrained
w_{Online}	0.717***	0.714***
(se)	0.010	0.010
$w_{Traditional}$	1.012***	1.012***
(se)	0.000	0.000

* p<0.05,** p<0.01, ***p<0.001

As previously noted, when comparing the unconstrained model to the homogeneity-constrained model, a chi-squared test can be applied in concert with the unconstrained model to determine whether or not the homogeneity constraint holds. For the unconstrained model, the calculated chi-squared value of 9.67 is greater than all conventional critical value thresholds, so homogeneity is rejected. Similarly, when homogeneity is imposed on the model, a χ^2 test for Slutsky symmetry is calculated to be 14,706 so symmetry is rejected at typical confidence levels as well. Finally, the predicted online and traditional expenditure shares were 4.1% and 95.9% in both the unconstrained and homogeneity-constrained models.

5.7.3 Sales Tax Policy Implications of Elasticities

Using estimated elasticities, it is possible to estimate how different expenditure shares would be if states had never implemented sales taxes on online retail goods. With sales taxes remaining on traditional brick-and-mortar goods, we would expect that the online expenditure share would be greater.

Examining the elasticities for selected states of online goods we can not only compare the differences in uncompensated and compensated elasticities between states but the elasticity value can be multiplied by the prevailing sales tax rate in the state to determine an overall effect of sales tax enforcement for online purchases on the expenditure share of online goods in each state. The figures presented are likely an upper bound of policy potential as the estimated sample is indicative of the intensive margin of online shoppers and not retail consumption overall. The results appear in Table 11.

Table 11: Elasticities of Online Expenditure Share and Sales Tax Impact by State

State	Budget Elas.	Uncomp.	Comp.	Sales Tax	Uncomp*Sales Tax
CA	0.691***	-1.306***	-1.274***	7.25%	-9.47%
	0.009	0.023	0.023		
OR	0.722***	-1.260***	-1.232***	0.00%	0.000%
	0.011	0.026	0.026		
WA	0.737***	-1.290***	-1.259***	6.50%	-8.39%
	0.010	0.024	0.024		
NV	0.795***	-1.303***	-1.271***	6.85%	-8.93%
	0.012	0.027	0.026		
TX	0.794***	-1.350***	-1.320***	6.25%	-8.44%
	0.013	0.030	0.030		

5.7.4 Variance Inflation Factors

We ascertain the presence and extent of multicollinearity among expenditure terms, and the models' variance inflation factors (VIFs) are used for this purpose. Table 12 presents the VIF results based on the pooled model and uses the traditional rule of thumb of a VIF greater than 10 being indicative of multicollinearity. For the unconstrained case, it is readily apparent that both the total expenditure term and quadratic total expenditure term at 176.71 and 161.91 respectively, exhibit strong multicollinearity. This is not surprising as the quadratic term is correlated with the level by design in the quadratic AIDS framework, so this is not a cause of concern. Other variables that reach above the threshold of 10 are the log median online price at 10.7 and a few demographic variables. The VIF figure for log median online price is close to VIF figures obtained from other readily available commodity data sets Lecocq and Robin (2015). The demographic variables exhibiting multicollinearity are not ”employed for pay” at 26.05 and ”military, student or non-employed” which is explained by the fact that when employment hours worked reach zero, the chances of respondents being in the occupation category ”military, student, or non-employed” increases.

Table 12: Variance Inflation Factors (VIFs) for Unconstrained and Homogeneity Constrained Models

	Unconstrained	**Homogeneity**
$\gamma_{Ln\,Median\,Online\,PUC\,w/tax}$	10.71	10.67
$\gamma_{Ln\,Median\,Trad\,PUC\,w/tax}$	1.46	1.46
β_{lnx}	176.71	176.74
λ_{lnx^2}	161.91	161.78
Head of Household is Male	2.23	2.23
Married	3.33	3.33
Head of Household Age	1.76	1.76
Household size (omitted base category: Four members):		
One Member	5.85	5.85
Two Members	4.66	4.66
Three Members	2.18	2.18
Five or More Members	1.55	1.55
Domicile Type (omitted base category: One Family Domicile)		
Two Family Domicile	1.03	1.03
Three or more Family Domicile	1.14	1.14
Mobile, Trailer, or Unreported	1.06	1.06
Income Bracket (omitted base category: 45k to 69.6k)		
Under 15k	1.36	1.36
15k - 29.9k	1.53	1.53
30k - 44.9k	1.45	1.45
70k - 99.9k	1.5	1.5
100k+	1.57	1.57
Age and Presence of Children (omitted base category: Under 6 Only)		
A Least One Child 6 to 17	5.34	5.34
No children under 18	6.64	6.64
Employment Hours (omitted base category: Over 35)		
Under 35 hrs/week	1.21	1.21
Not Employed for Pay	26.05	26.05
M/F HH Education (omitted base category: Some college)		
HS or less	1.45	1.45
College Graduate or Post-graduate	1.53	1.53
M/F Occupation (omitted base category: White Collar Job Categories)		
Blue Collar Job Categories	1.4	1.4
Military or Student, or Non-employed	24.84	24.84
Race (omitted base category: White)		
Black	1.06	1.06
Asian	1.07	1.07
Other	1.18	1.18
Hispanic (ethnicity)	1.2	1.2

5.7.5 Individual Panel Year Results

In this extension, we ask whether consumers became more or less price-sensitive over time in their online shopping behavior, as the overall volume of online sales increased strongly (Figure 1). We estimate separate models for each year. The quadratic unconstrained estimates are presented in Table 13. The estimated γ coefficient always has the anticipated negative sign that indicates an online price increase leads to a lower online expenditure share. The magnitude of the γ coefficient for online shopping has a slightly positive trend over the 2011 to 2017 period with the greatest uptick in 2017, where γ reached -0.152. This indicates online shoppers have become more price-sensitive as online purchases have increased over time.

Male head of household, marriage, domicile type, and household size are among several demographic variables of interest that were statistically significant over most periods. Other demographic variables such as race, household size, and household income have varying levels of statistical significance in some but not all years in the period.

Along with the output from the unconstrained model the AIDS ILLS command reports a joint χ^2 test for homogeneity. The results are mixed with four out of seven years in the period rejecting the null at conventional levels.

Table 13: Online Expenditure Shares of Unconstrained Quadratic AIDS for Panel Years

Online Share ($w_{1,h,t}$)	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**
$\gamma_{Ln\ Median\ Online\ PUC\ w/tax}$	-0.126***	-0.112***	-0.114***	-0.126***	-0.132***	-0.121***	-0.152***
	(0.0190)	(0.0170)	(0.0162)	(0.0158)	(0.0158)	(0.0130)	(0.0153)
$\gamma_{Ln\ Median\ Trad\ PUC\ w/tax}$	0.0943***	0.107***	0.107***	0.132***	0.151***	0.138***	0.161***
	(0.0224)	(0.0219)	(0.0207)	(0.0197)	(0.0191)	(0.0156)	(0.0173)
β_{lnx}	-0.179***	-0.170***	-0.166***	-0.173***	-0.172***	-0.165***	-0.194***
	(0.0144)	(0.0137)	(0.0131)	(0.0125)	(0.0127)	(0.0111)	(0.0113)
λ_{lnx^2}	0.0104***	0.0010***	0.0098***	0.0104***	0.0103***	0.0099***	0.0117***
	(0.0009)	(0.0009)	(0.0008)	(0.0008)	(0.0008)	(0.0007)	(0.0007)
Head of Household is Male	0.0110***	0.0105***	0.0067**	0.0041*	0.0018	0.0001	-0.0009
	(0.0024)	(0.0022)	(0.0020)	(0.0019)	(0.0019)	(0.0017)	(0.0018)
Married	-0.0063**	-0.0105***	-0.0088***	-0.0079***	-0.0057**	-0.0085***	-0.0060**
	(0.0024)	(0.0025)	(0.0023)	(0.0022)	(0.0022)	(0.0020)	(0.0021)
Head of Household Age	0.0003***	0.0003***	0.0003***	0.0002**	0.0000	0.0000	-0.0001

Continued on next page

Table 13 – continued from previous page

Online Share $(w_{1,h,t})$	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0000)	(0.0001)
Household size (omitted base category: Four members):							
One Member	0.0082*	0.0072*	0.0067*	0.0080*	0.0103***	0.0125***	0.0143***
	(0.0038)	(0.0035)	(0.0034)	(0.0032)	(0.0031)	(0.0029)	(0.0031)
Two Members	-0.0013	-0.0018	-0.0006	0.0023	0.0003	0.0019	0.0018
	(0.0030)	(0.0028)	(0.0026)	(0.0026)	(0.0025)	(0.0022)	(0.0024)
Three Members	0.0026	-0.0038	0.0008	0.0022	-0.0007	-0.0001	0.0006
	(0.0029)	(0.0027)	(0.0026)	(0.0025)	(0.0024)	(0.0022)	(0.0023)
Five or More Members	0.0010	-0.0034	-0.0018	-0.0003	-0.0004	-0.0031	-0.0028
	(0.0034)	(0.0031)	(0.0031)	(0.0029)	(0.0028)	(0.0025)	(0.0027)
Domicile Type (omitted base category: One Family Domicile):							
Two Family Domicile	0.0011	-0.0025	0.0125***	0.0067*	0.0029	0.0020	0.0041
	(0.0036)	(0.0033)	(0.0032)	(0.0029)	(0.0029)	(0.0026)	(0.0028)
Three or more Family Domicile	0.0050*	0.0061**	0.0078***	0.0089***	0.0075***	0.0091***	0.0127***
	(0.0022)	(0.0021)	(0.0019)	(0.0018)	(0.0018)	(0.0016)	(0.0017)
Mobile,Trailer, or Unreported	0.0090*	0.0064	0.0045	0.0046	0.0044	0.0032	0.0005
	(0.0035)	(0.0035)	(0.0031)	(0.0029)	(0.0029)	(0.0027)	(0.0029)
Income Bracket (omitted base category: 45k to 69.6k):							
Under 15k	0.0064	0.0004	-0.0050	-0.0000	-0.0013	-0.0050*	-0.0016
	(0.0035)	(0.0032)	(0.0030)	(0.0028)	(0.0028)	(0.0025)	(0.0027)
15k - 29.9k	0.0018	-0.0013	0.0005	-0.0003	-0.0030	0.0007	-0.0017
	(0.0024)	(0.0023)	(0.0021)	(0.0020)	(0.0020)	(0.0019)	(0.0020)
30k - 44.9k	0.0009	-0.0032	-0.0034	-0.0019	-0.0001	-0.0038*	-0.0008
	(0.0022)	(0.0021)	(0.0019)	(0.0018)	(0.0018)	(0.0017)	(0.0018)
70k - 99.9k	-0.0012	-0.0031	-0.0026	-0.0018	-0.0017	0.0001	0.0014
	(0.0021)	(0.0020)	(0.0018)	(0.0017)	(0.0017)	(0.0015)	(0.0016)
100k+	0.0013	0.0026	0.0009	0.0018	0.0031	0.0054**	0.0054**
	(0.0023)	(0.0022)	(0.0020)	(0.0019)	(0.0019)	(0.0017)	(0.0017)
Age and Presence of Children (omitted base category: Under 6 Only)							
A Least One Child 6 to 17	-0.0059	-0.0030	-0.0000	-0.0001	-0.0030	0.0016	-0.0026
	(0.0041)	(0.0038)	(0.0040)	(0.0036)	(0.0035)	(0.0029)	(0.0031)
No children under 18	0.0013	0.0007	0.0054	0.0035	0.0004	0.0045	0.0018
	(0.0043)	(0.0040)	(0.0041)	(0.0038)	(0.0037)	(0.0031)	(0.0032)
Employment Hours (omitted base category: Over 35)							
Under 35 hrs/week	-0.0031	0.0026	-0.0009	0.0003	0.0002	0.0009	0.0014
	(0.0024)	(0.0023)	(0.0021)	(0.0020)	(0.0019)	(0.0018)	(0.0019)
Not Employed for Pay	0.0118	0.0137*	0.0087	0.0132*	0.0108	0.0089	0.0064
	(0.0076)	(0.0066)	(0.0061)	(0.0057)	(0.0064)	(0.0059)	(0.0061)
M/F HH Education (omitted base category: Some college)							
HS or less	0.00028	-0.0015	-0.0027	-0.0035*	-0.0026	-0.00037	0.0006
	(0.0020)	(0.0018)	(0.0017)	(0.0016)	(0.0016)	(0.0014)	(0.0015)
College grad or post-grad	0.0011	-0.0018	-0.0014	-0.0017	-0.0013	0.0019	0.0008
	(0.0017)	(0.0016)	(0.0015)	(0.0014)	(0.0014)	(0.0013)	(0.0014)
M/F Occupation (omitted base category: White Collar job categories)							
Blue Collar job categories	0.0009	-0.0002	-0.0003	-0.0027	-0.0028	-0.0006	0.0000
	(0.0020)	(0.0019)	(0.0018)	(0.0017)	(0.0016)	(0.0015)	(0.0016)
Military/Student/Non-employed	-0.0085	-0.0098	-0.0052	-0.0097	-0.0052	-0.0021	0.0031
	(0.0074)	(0.0064)	(0.0059)	(0.0055)	(0.0063)	(0.0057)	(0.0059)
Race (omitted base category: White)							
Black	0.0033	0.0027	0.0054**	0.0037*	0.0013	0.0024	0.0033

Continued on next page

Table 13 – continued from previous page

Online Share $(w_{1,h,t})$	2011	2012	2013	2014	2015	2016	2017
	(0.0024)	(0.0022)	(0.0020)	(0.0018)	(0.0018)	(0.0016)	(0.0017)
Asian	0.0073	0.0106**	0.0135***	0.0073*	0.0074*	0.0038	0.0054*
	(0.0038)	(0.0035)	(0.0034)	(0.0031)	(0.0031)	(0.0026)	(0.0027)
Other	0.0004	0.0008	0.0084**	0.0079**	0.0057	0.0033	0.0034
	(0.0037)	(0.0034)	(0.0032)	(0.0030)	(0.0029)	(0.0026)	(0.0026)
Hispanic (ethnicity)	0.0050	0.0010	-0.0026	-0.0012	0.0005	0.0000	0.0019
	(0.0033)	(0.0031)	(0.0029)	(0.0027)	(0.0025)	(0.0022)	(0.0023)
Constant	0.806***	0.742***	0.724***	0.743***	0.743***	0.710***	0.838***
	(0.0572)	(0.0551)	(0.0534)	(0.0511)	(0.0524)	(0.0466)	(0.0480)

*Standard errors in parentheses $*p < 0.05 **p < 0.01 ***p < 0.001$

5.8 Conclusion

The demand estimation results for the Nielsen Consumer Panel data set demonstrate that the predicted expenditure shares are within comparable shares of online sales from external data sources. Additionally, own-price elasticity for online goods is consistently price elastic at elasticity values around 1.2 to 1.3, indicating that consumers are price-sensitive in regard to their online shopping patterns. In particular, these results indicate that consumers are sensitive to changes in state sales tax policy. For example, if California had not introduced sales taxation for online goods, the expenditure share of online goods for consumers in California would be 9.5% higher. This suggests that sales tax policy toward online retailers can substantially shift consumption between online and traditional retailers and can therefore be an effective instrument if policymakers intend to strengthen small local brick-and-mortar businesses.

The results of this analysis show that traditional goods are less elastic than online goods. One possibility for why this might be expected is that traditional brick-and-mortar goods are a default for consumers. If a good can not be found online at a competitive price or conveniently delivered, familiarity with existing brick-and-mortar retailers gives consumers one last fallback option that is less elastic as a final procurement source. Since e-commerce has evolved to the point where small brick-and-mortar businesses can market their goods online just as larger online retailers, the delineation between what constitutes traditional or online goods can be further tested. Additionally, if states continue to modify the level of their sales tax rate, the effectiveness of such an instrument might be further examined using these methods. The methods used in this paper can be applied in future work with a further granularity of goods categories. Specifically, instead of examining broad categories of online or traditional goods, product categories within these two broad categories might be explored such as apparel or electronics categories.

6 Joint Prediction and Simulation of Labor Force and Fiscal Conditions of Nevada Counties

Randall Chicola (University of Nevada, Reno)
Dr. Man-Keun Kim (Utah State University)
Dr. Wuyang Hu (The Ohio State University)
Dr. Thomas Harris (University of Nevada, Reno)

6.1 Abstract

This study provides a system to jointly modeled labor and fiscal conditions of Nevada counties. Each is specified as a system of equations that are linked by allowing variables from the labor module to enter the fiscal module. Following the identification of parameters in the two modules, a simulation is analyzed to account for the effects of changes in exogenous employment on the labor and fiscal status of each Nevada county in addition to dynamic relationships of county budgets. In particular, uncertainties and noise in the estimation process are explicitly considered which allows the simulation process to produce confidence intervals rather than single point most likely solutions. This model may serve as a basis for understanding the ramifications of COVID-19 on the state of Nevada as county budget data becomes available in the future.

6.2 Introduction

County planning decisions involve onerous complexity in regard to matching tax base revenues with an appropriate allocation of resources to the amenities and services most valued by its residents. Rural counties in particular face uniquely challenging obstacles in making the fiscal decisions that best enhance economic welfare and growth in a constrained resource environment. Often rural counties that are not diverse and rely on one or a few economic sectors find themselves vulnerable to being fiscally impacted by an idiosyncratic shock to a singular economic sector.

Building upon previous studies, the ongoing applications of the Community Policy Analysis Network (CPAN) framework are applied in understanding the fiscal and labor sectors of the Nevada counties' economy. The empirical approach outlined by Harris et al. (2000) and Yeo and Holland (2004) was followed with significant modifications. Compared with these previous studies, this current analysis addresses explicitly the issue of uncertainties in impact analysis and consequently policy decision-making. Given the complexity of this process, a great deal of uncertainties is expected (Harris, 1995).

As Chalmers and Anderson (1977) state, "Uncertainty is the essence of the planning problem and the public is not well served by a strategy that simply plans for the most likely future." Without an appropriate mechanism to incorporate these uncertainties into the decision-making process, results may be significantly biased, and this makes it difficult to provide balanced and effective policy guidance to the county authorities. Some previous studies in county-level decision-making notice some of the factors involved (e.g. Swenson and Eathington, 1998) but have not presented a clear approach to operationalize the analysis.

The primary objective of this paper is to develop a system that can link the labor and fiscal sides of county-level decision-making while considering the uncertainties involved. First, a system of labor force models is established to generate the interrelationship between total labor supply, in- and out-commuting labor, unemployment, and county population growth. This procedure is critical given the uniqueness of Nevada counties. Nevada has a very low overall population density and yet the majority of Nevada's population is concentrated in three metropolitan statistical areas (MSA's): Las Vegas-Henderson-Paradise, Reno-Sparks, and Carson City. Even a slight labor or population change may introduce significant impacts to some counties (Swenson and Otto, 1998). Second, following Harris et al. (2000) the fiscal conditions of Nevada counties are analyzed through the analysis of each county's revenues and expenditures. A new approach is applied to model the potential correlations between these two fiscal aspects. Third, following Yeo and Holland (2004), the link between the labor force and fiscal stages of the county-level decision-making is introduced by allowing variables describing each county's labor force characteristics to enter the models representing revenue and expenditure functions of a county government. Finally, the uncertainties involved with the decision-making process are addressed. This is achieved by taking appropriate consideration of the errors involved in various models.

6.3 Literature Review

Efficient local economic growth initiatives should consider all fiscal and labor factors which will lead to the need for creating a comprehensive approach by incorporating these factors in rural decision-making (Harris et al., 2000). Following this need and with the assistance of relevant input-output analysis, a group of studies has focused on the impacts of regional policy changes and economic development on regional or industry sector fiscal conditions (e.g., Beemiller, 1989; Song et al., 1992). However, not until the launch of the CPAN (Community Policy Analysis Network) framework in 1995, has impact analysis been formalized and presented in a systematic manner. The CPAN effort has brought impact analysis to a position with higher priority and drawn increasing attention from both researchers and government policymakers.

The CPAN framework recommends studying regional economies in a comprehensive manner with a focus on empirical module construction, appropriate estimation technique selection, sound interpretation, and feasible extension of results. Deller (1995) provides an overview of several characteristics of this framework and the early history involved in developing CPAN. Johnson and Scott (1996) and Swenson (1996) both offer applications of the CPAN framework addressing local economic issues and the impact on regional decision-making from changes in economic conditions. Swenson and Otto (1998) and Swenson and Eathington (1998) further summarize the development of methods based on the CPAN framework and impact analysis in general. Since then, empirical impact analyses within the framework of CPAN using more advanced techniques or with extended application areas have grown rapidly. Harris et al. (2000) and Yeo and Holland (2004) studied regional fiscal conditions and labor or population growth by building systems of models that reflect county-level decision-making. Shields and Deller (2003) extract techniques from the broad impact analysis to improve the communication between regional authorities responsible for economic development and the general public. They also placed a special focus on the pros and cons of the method. Bangsund et al. (2004) studied the impact of conservation policies on local agricultural and recreational activities. Evans and Stallmann (2006) extended the basic CPAN framework to create a customized system referred to as the TEX$AFE to address local situations in Texas.

A comprehensive approach is needed that considers all factors affecting the balance between economic growth and a county government's resource constraints (Harris et al., 2000). Additionally, as shown in Brückner and Pappa (2012), a comprehensive approach can capture the responses of labor force participation, employment, and unemployment from government expenditure shocks generated by fiscal policy changes. Using the relevant input-output has been in previous work for regional policy applications for economic development (Beemiller, 1989) and fiscal conditions of industries (Song et al., 1992).

Impact analysis took additional steps forward with CPAN (Community Policy Analysis Network) framework in 1995, by systematically formalizing and presenting data-driven analysis for regional and community development with a position with higher priority and has drawn increasing attention from both researchers and government policymakers.

Further advancement of impact analysis with structural vector autoregressions (SVARs) by Brückner and Pappa (2012) incorporating the search and matching model of Diamond-

Mortensen-Pissarides with insider and outsider labor market participant searching dynamics. Petrović et al. (2021) found that for Central and East European EU economies, the effects of public investment on output are strong and persistent, but with shorter and weaker persistence for public consumption. Similar dynamics that they studied at an international scope are important for a regional scope in regards to local public investment and joint investment between localities. Cardi et al. (2020) analyzes sectoral fiscal impacts of government spending, finding that the negative wealth effect is created from a spending shock that causes households to provision more labor, increasing real GDP for 16 O.E.C.D. countries. Their application of sectoral fiscal multiplier estimates and sectoral share response to a government spending shock and partitioning of the effects between traded and non-traded sectors.

6.4 Conceptual Framework

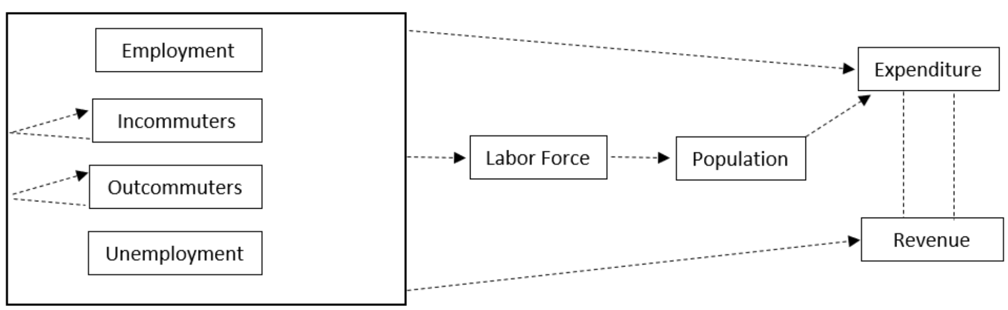


Figure 6: Conceptual Framework Diagram

The labor sector and the fiscal sector of a county are considered jointly and can be conceptualized by Figure 6. This diagram can be viewed as two sectors that are linked by some factors. The two sectors can be referred to as the labor module (the large box with a solid border at the very left of the diagram and the two boxes in the middle of the diagram) and the fiscal module (the two smaller boxes at the very right edge of the diagram). In particular, the population also functions as a link between these two modules. This conceptual framework is discussed from left to right, i.e., from the labor module to the fiscal module. The labor module considers four factors in determining the supply of labor that is included in the large solid box at the left of the diagram. These factors are employment, incommuting labor to the county known as incommuters, outcommuting labor to the county known as outcommuters, and the unemployed. There certainly are other factors that may be considered important in this relationship. Some of these factors are considered in the specific models introduced in the next section. Nevertheless, the four factors in the box of Figure 6 represent those most commonly seen in the literature (e.g., Yeo and Holland, 2004).

After defining these four factors in the labor module, the total labor force can be defined as a function that is affected by the four factors, which is captured by the arrow from the large box to the labor force. Following Yeo and Holland (2004), the labor force in turn determines the population growth in a county. This relationship is explained by the arrow from the labor force to the population in Figure 6. The labor module then includes the labor force and the population as a system of simultaneous equations. It is noteworthy however that within the four factors used to explain the labor force, incommuters and outcommuters are likely to be endogenous. Any potential individual may decide to travel out of their resident county for work and this endogeneity has to be properly addressed in the modeling process. These two endogenous variables are reflected in Figure 6 by the dashed arrows directing to themselves. An appropriate modeling approach is to include equations explaining incommuters and outcommuters together with the two equations for the labor force and population growth. Thus, these four equations complete the labor module.

For the fiscal module, the conceptual framework assumes that it is distinctively different from the labor sector but yet closely related (Swenson and Eathington, 1998). The distinctiveness is reflected by the fact that although the two individual models associated with the fiscal sector (expenditure and revenue) are estimated jointly to form the fiscal module, they are not directly included in the labor module. The close relationship between these two modules is created by the fact that many factors in the labor module are assumed to have direct impacts on the two aspects of the fiscal module. First of all, the population is an important factor for counties to determine their expenditure. The increase in population may increase the overall expenditure but due to the potential higher concentration of population, the per capita expenditure by the county governments may decrease (Yeo and Holland, 2004). These effects can be captured by incorporating the population variable into the expenditure model.

The second source of impacts to the fiscal module may come from the factors that affect the supply of labor; i.e., the four factors identified in the labor module and included in the large box with a solid border at the left of the diagram. These factors may also affect revenue in addition to expenditure. These relationships are represented by the dashed arrows linking these factors to the two equations of the fiscal module. Finally, the two fiscal models are incorporated into one complete system where they are modeled jointly. This is represented by the double dashed line connecting these two measures in Figure 6. In this structure, factors affecting one aspect of the fiscal module (either expenditure or revenue) will also indirectly affect the other.

The conceptual framework presented in Figure 6 demonstrates a comprehensive system that accounts for both the labor and the fiscal sides of potential impacts to a county introduced by a change in economic conditions. It provides a more complete picture of the county-level impacts than considering only one side or the other (Harris et al., 2000). Given its completeness, the system is also relatively simple to implement. One can evaluate the labor and fiscal modules separately and the links between these two sectors are naturally created by the factors that appear in both modules, such as the population. Besides simplicity, the framework described in Figure 6 also differs from other existing systems that address the labor and fiscal conditions simultaneously (Yeo and Holland, 2004).

This is reflected by its capacity to allow uncertainty. No analysis can capture all impacts that may affect the system (Swenson and Eathington, 1998). The dotted arrows in the figure reflect this fact. Uncertainties or noise in the system may be introduced from two aspects. First are the omitted factors. Due to the complexity of the issues involved, models with a limited number of variables may in fact only be viewed as incomplete. Second, errors may be introduced in the estimation process. These uncertainties and errors make it necessary to allow the system to support results with upper and lower bounds with a certain level of confidence. These uncertainties or errors can be introduced in all stages of the framework outlined in Figure 6 and therefore can easily accommodate this requirement.

6.5 Data

Data used in this study are obtained and merged from two major sources for the period 2004 to 2016. The labor module uses data procured from the St. Louis Federal Reserve FRED economic data. For the labor module, FRED sources labor force from the U.S. Bureau of Labor Statistics (B.L.S.), persons in the civilian labor force for each county and year. Unemployment figures are also from B.L.S. and the place of work employment $POWEMP_{i,t}$ is sourced from BLS Quarterly Census of Employment and Wages. For in- and out-commuting activities in each county, the U.S. CENSUS ”OntheMap” application is used. Additionally, U.S. Census Bureau resident population figures for each county and year are used.

External labor force and external employment are denoted $XLF_{i,t}$ and $XMP_{i,t}$. Data for $XLF_{i,t}$ and $XMP_{i,t}$ are from B.L.S. employed persons and civilian labor data but were aggregated from the contiguous counties adjacent to each of the seventeen Nevada counties. Distance figures were determined by the shortest time driving distance between the county seats for each pair of adjacent counties using Google maps.

It is noticeable, however, that since Nevada only has seventeen counties, it determines that the labor module has a limited degree of freedom. To reduce this problem, (Harris et al., 2000) integrated bordering counties from nearby states based on the BEA classification. In this analysis, we choose to focus on Nevada counties only because this may offer a more direct description of conditions that would be of interest to the state. In addition, although not applied in this study, more advanced statistical methods, such as the Bayesian approach, may assist the analysis with limited data. This remains an interesting future research avenue for regional impact studies since in many cases, these studies are troubled by the lack of observations.

In the fiscal module, annual county general fund revenue and expenditure data for all 17 Nevada counties from the Nevada Department of Taxation were compiled into a panel dataset (Nevada Department of Taxation, various issues). The use of cross-sectional data for the labor module and panel data for the fiscal module does not impose a problem in this study. This is because the two modules are estimated separately but joined by the common factors in both modules. The parameter estimates are those that are useful in interpretation and follow-up analysis.

6.6 Model and Empirical Methodology

6.6.1 Labor Model

Following the discussion on the aforementioned conceptual framework and (Swenson and Eathington, 1998), the labor module is comprised of four equations simultaneously explaining four quantities: labor force, incommuters, outcommuters, and the population. Based on the cross-sectional nature of the data used under the labor module and using subscript i to denote counties, subscript t to denote the year, and superscript to denote the equation number, the four equations can be expressed as:

$$LF_{i,t} = \beta_0^1 + \beta_1^1 POWEMP_{i,t} + \beta_2^1 INCOMM_{i,t} + \beta_3^1 OUTCOMM_{i,t} + \beta_4^1 UNEMP_{i,t} + \epsilon_{i,t}^1 \tag{9}$$

$$INCOMM_{i,t} = \beta_0^2 + \beta_1^2 POWEMP_{i,t} + \beta_2^2 XLF_{i,t} + \beta_3^2 UNEMP_{i,t} + \epsilon_{i,t}^2 \tag{10}$$

$$OUTCOMM_{i,t} = \beta_0^3 + \beta_1^3 POWEMP_{i,t} + \beta_2^3 XEMP_{i,t} + \beta_3^3 UNEMP_{i,t} + \epsilon_{i,t}^3 \tag{11}$$

$$POP_{i,t} = \beta_0^4 + \beta_1^4 LF_{i,t} + \epsilon_{i,t}^4 \tag{12}$$

where $LF_{i,t}$ is the total civilian labor force, $POWEMP_{i,t}$ is the place of work employment for the number of employees in the covered area for all industries, $INCOMM_{i,t}$ is the total number of incommuters who are employed in the study area but living outside the study area, $OUTCOMM_{i,t}$ is the total number of outcommuters who are living in the study area but employed outside the study area, $XLF_{i,t}$ external labor force, $XEMP_{i,t}$ is external employment, $POP_{i,t}$ is total population, $XLF_{i,t}$ and $XEMP_{i,t}$ are created following Yeo and Holland (2004) and Swenson and Otto (1998) where:

$$XLF_{i,t} = \sum_j \frac{Contiguous\ Labor\ Force_{j,t}}{distance_{ji}^2} \tag{13}$$

$$XEMP_{i,t} = \sum_{j} \frac{Contiguous\ Employment_{j,t}}{distance_{ji}^2} \tag{14}$$

Subscript j represents the adjacent counties to county i and $distance_{ji}$ represents the distance between county seats of county j to i.

Given the construct of the four equations for the labor force models, the system is simultaneous which requires using the three-stage least squares (3SLS) estimation approach. The first group of candidates for instrument variables is the exogenous variables present in the equations. These variables are: $POWEMP_{i,t}$, $XLF_{i,t}$, and $XEMP_{i,t}$. Given the system, to ensure that it is identifiable, at least four instruments are required. Although more instrument variables can be used, an exact instrument may often reduce the complexity of estimating and validating the system (Kennedy, 2003). After several trials, the variable describing contiguous employment is selected as the additional instrument variable as the incorporation of this variable generates the highest overall model fit in comparison to several other competing options. As to the functional form, there is no explicit theoretical guidance. The selection is rather on a case by case. Given the most commonly used linear or logarithm forms, we have tested different models with these specifications and the linear models appear to have the best model fit.

6.6.2 Fiscal Model

The fiscal impact module considers the expenditure and revenue of each of the seventeen counties in Nevada. Denote $PCEXP_{i,t}$ and $PCREV_{i,t}$ as the per capita expenditure and revenue for each county and subscript t for time, the following expressions can be specified:

$$PCEXP_{i,t} = f(POWEMP_{i,t}, UNEMP_{i,t}, POPDEN_{i,t}, PCB_{i,t-1}) \tag{15}$$

$$PCREV_{i,t} = f(POWEMP_{i,t}, UNEMP_{i,t}, , PCB_{i,t-1}) \tag{16}$$

where $POPDEN_{i,t}$ is the population density of county i in year t and $PCB_{i,t-1}$ is the fiscal balance per capita in county i in year t. $PCB_{i,t-1}$ is used instead of other income variables (such as household incomes) and this is because the county-level fiscal conditions are of interest in this study. Per capita fiscal balance enters the models in a one-period lagged form since it is only expected that the previous term's balance, which is observed at the end of the previous term, will affect the current period's financial decisions of the government, but not the current term's balance.

Also, the lagged per capita balance variable is supported by Holtz-Eakin et al. (1987) whose research suggested that standard regressions that examine only the contemporaneous relationships of county budget fiscal variables are inappropriate. There exist inter-temporal relationships that need to be incorporated into county government fiscal models. In this application, the two models take a linear functional form, which is the most commonly seen in the literature. The panel nature of the data is addressed by incorporating one-way fixed effects into the two models to capture the differences between counties.

Similarly, variables in the fiscal models may take either their original or the log form in the literature (Harris et al., 2000). After some investigation, models with the linear dependent variable and log independent variables seem to have the best fit. This structure is taken as the final model specification. Variable $PCB_{i,t-1}$ however was not transformed into log format due to the fact that some observations of the balance terms are negative Specifically, the expenditure and revenue models can be written as:

$$\begin{aligned} PCEXP_{i,t} = \alpha_{0,i} + \alpha_1 LNPOWEMP_{i,t} + \alpha_2 LNUNEMP_{i,t} \\ + \alpha_3 LGPCB_{i,t} + \alpha_4 LNPOPDEN_{i,t} + \epsilon_{i,t} \end{aligned} \tag{17}$$

$$PCREV_{i,t} = b_{0,i} + b_1 LNPOWEMP_{i,t} + b_2 LNUNEMP_{i,t} + b_3 LGPCB_{i,t} + \nu_{i,t} \quad (18)$$

where LNPOWEMP, LNUNEMP, and LNPOPDEN are the log transformation of variables POWEMP, UNEMP, and POPDEN. Variable LGPCB is the lagged term of variable PCB while $\epsilon_{i,t}$ and $\nu_{i,t}$ are error terms. It is obvious that the above models could be consistently estimated individually as separate models, but since government expenditure and revenue often result from one system of decisions, these two quantities are expected to be correlated. In other words, error terms $\epsilon_{i,t}$ and $\nu_{i,t}$ are expected to be correlated as well. This gives a system of equations that can be estimated by the seemingly unrelated regression (SUR) approach. The definition and descriptive statistics of variables used in this study are summarized in Table 14.

Table 14: Variable Definition and Sample Statistics

VARIABLE	DEFINITION	MEAN	STD. DEV.
LABOR FORCE MODELS:			
LF	total labor force	62,414.73	183,525.20
POWEMP	place of work employment	55,636.73	162,960.39
INCOM	total number of incommuters	14,727.48	45,105.30
OUTCOM	total number of outcommuters	18,359.53	60,278.58
UNEMP	total number of unemployment	5,529.53	17,895.07
XLF	external labor force	22,074.65	137,508.91
XEMP	external employment	315.68	1,253.44
POP	total population	148,228.82	424,782.71
CONEMP	contiguous employment	626,374.95	1,355,917.83
FISCAL MODELS:			
PCEXP	per capita government expenditure (inflation adjusted)	1,616.23	2,011.59
PCREV	per capita government revenue (inflation adjusted)	1,705.09	2,107.96
POWEMP	place of work employment	55,636.73	162,960.39
UNEMP	total number of unemployment	5,529.53	17,895.07
POPDEN	population density (person/ sq. mile)	36.32	73.96
LNPOWEMP	log of place of work employment	9.13	1.72
LNUNEMP	log of total number of unemployment	6.71	1.72
LNPOPDEN	log of population density (person/ sq. mile)	1.84	1.91
LGPCB	lagged per capita balance	1,084.91	2,048.27

6.7 Estimation Results

The three-stage least squared (3SLS) estimation results of the labor force models are presented in Table 15. The overall adjusted R squared values are high indicating a good model fit. Signs of model coefficients are consistent with previous literature. Place of work employment has positive impacts in all three models: labor force, incommuters, and outcommuters. For the labor force model, while incommuters decrease the labor force, outcommuters have a positive impact. Although this result is difficult to interpret, it is consistent with Yeo and Holland (2004). Further investigation of this issue may be warranted.

Table 15: 3SLS Estimation Results of Labor Force Models (No Fixed Effects)

Variable	Labor Force		Incommuters		Outcommuters		Population Growth	
	Coeff.	Std Error	Coeff.	Std Error	Coeff.	Std Error	Coeff.	Std Error
Constant	607.575	781.911	3943.595	1509.271	2408.419	512.258	6516.913	3964.707
POWEMP	1.182	0.016	0.188	0.009	0.0512	0.003		
INCOMM	-1.905	0.4375						
OUTCOMM	1.311	0.296						
XLF			0.013	0.004				
XEMP					41.475	0.421		
LF							2.270	0.020
R^2	0.992		0.466		0.963			

Variable XLF (external labor force) has a positive coefficient in the incommuters model, indicating that the increase in the external labor force will lead to an increase in the number of incommuters. Similarly, an increase in external employment (variable XEMP) will increase the number of outcommuters. Finally, the total labor force has a direct positive impact on the growth of the population as reflected by the positive coefficient associated with variable LF in the population model. Since the labor module is closely related to the fiscal module and the fiscal module also extends the labor module, we interpret the fiscal module in more detail.

The single equation and SUR estimation results of the fiscal models are presented in

Table 16. Most of the fixed effects constant terms are significant in both models under either the single equation or the SUR estimation procedures. These constants however are omitted from this table for simplicity. The adjusted R^2 value indicates that the SUR model has a better fit than the single equation estimation with regards to the expenditure model but worse than the fit with regards to the single equation revenue model. Signs of coefficients are consistent across the two estimation procedures. Place of work employment has a positive impact on both government expenditure and revenue. This can be interpreted as employment contributing to the increase in the volume of the local economy. Based on the formulation of the models, the marginal effect of employment on government expenditure evaluated at the average employment level across Nevada is given as the ratio of the coefficient associated with variable LNPOWEMP and the average employment: $\frac{b_{LNPOWEMP}}{average(POWEMP)}$.

Table 16: Single Equation and SUR Estimation Results of the Fiscal Models

	Expenditure			Revenue	
Variable	Single Equation	SUR Model	Variable	Single Equation	SUR Model
LNPOWEMP	0.0229	0.0306	LNPOWEMP	0.0621	0.0621
Std. Err.	0.0218	0.0201	Std. Err.	0.0203	0.0201
LNUNEMP	-0.0700	-0.0675	LNUNEMP	-0.0895	-0.0895
Std. Err.	0.0201	0.0197	Std. Err.	0.0212	0.0210
LNPOPDEN	0.0360	0.0249	LGPCB	0.0000437	0.0000437
Std. Err.	0.0155	0.0104	Std. Err.	0.0000107	0.0000106
LGPCB	0.0000581	0.000056			
Std. Err.	0.0000105	0.0000101			
Constant	1.4229	1.3561	Constant	1.2694	1.2694
Std. Err.	0.1386	0.1192	Std. Err.	0.1097	0.1087
R^2	0.1924	0.2052	R$\hat{2}$	0.1831	0.1942

The implied marginal effect is 0.00121 and 0.012285 under the single equation and SUR estimation respectively. These indicate that the one additional employee will generate approximately \$0.001 to \$0.012 in expenditure per capita for Nevada counties on average. Similarly, one additional employee will contribute \$0.0023 and \$0.01624 in per capita revenues of the counties under the single equation and SUR methods. Therefore, if only the public fiscal conditions are considered, new employment leads to more revenue than expenditure to the county governments.

LNUNEMP is also significant under the revenue model and unemployment will lead to less expenditure for the counties. Based on a similar calculation as above, a one-person decrease in the number of unemployed will yield an additional $-0.004122 in per capita expenditure to the county in the single equation approach and $0.004465 in per capita expenditure in the SUR approach. It is clear from these results that an additional employee can stimulate the economy and generate economic activity for Nevada counties due to the coefficient's sign in the single equation and SUR methods. However, unemployment only produces net government expenditure and accordingly impacts fiscal balances for Nevada country governments. Furthermore, the marginal effect of unemployment (evaluated at the average unemployment level) on county expenditures is close to four times greater than that from employment. Therefore, any further degradation of federal unemployment assistance programs to county governments without accompanying financial assistance will negatively impact county government fiscal balances.

County population density is only used to explain county expenditures. Previous studies show that a higher population density yields lower per capita government expenditure (Harris et al., 2000). This is consistent with the result in the paper. The marginal effect indicates that one additional person per square mile in Nevada can reduce the government expenditure by $98.22 per capita in the single equations case or $441.17 per capita in the SUR model, as seen in Table 17 and Table 16 respectively. These values are considerably large in part because many rural counties in Nevada are the most sparsely populated in the nation. Ensuring the offering and quality of many federally and state-mandated public facilities and services to all Nevadans can be a costly proposition, especially for the sparsely populated rural areas.

Table 17: Marginal Effects Single Equations and SUR

	Single Equation	SUR
LNPOWEMP	67.21	128.1
	(70.01)	(61.61)
LNUNEMP	-229.0***	-216.9***
	(64.53)	(64.42)
LNPOPDEN	98.22*	
	(49.62)	
LGPCB	0.859***	0.886***
	(0.0336)	(0.0327)
_cons	1429.9**	1018.7**
	(444.3)	(333.7)
N	221	221

The lagged per capita county balance variable (LGPCB) is significant and positive in both fiscal models, except for the SUR revenue model. These positive coefficients also confirm the findings of Holtz-Eakin et al. (1987), that inter-temporal relationships need to be incorporated in county government models. Since variable LGPCB was included in the models in its linear form, the marginal effect is the coefficient associated with the variable in the models. Each additional per capita dollar from the last fiscal year can generate an additional \$0.86 in per capita expenditure (based on the single equation model) and approximately \$0.89 (consistent in two estimation procedures) in increased per capita revenue. An alternative way to explain this effect is to first calculate net revenue or loss, the difference between the expenditure and revenue implied in the period, then this calculated amount may be interpreted as the result of the balance increase or dissipation from the beginning of the period to the end of the period. In this case, holding other factors constant, a one-dollar per capita balance entering the fiscal year can generate \$0.06 in per capita balance at the end of the fiscal year.

6.8 Simulation

Given the estimation results of the labor force and the fiscal impact models, it is feasible to investigate the impact of an external shock on the local Nevada labor and fiscal conditions. Following Swenson and Eathington (1998), simulations of model predictions are conducted by assuming the place of work employment as an exogenous shock. In community models, employment opportunities can often be assumed to come from external investment such as the location or relocation of a manufacturing or service business. Jobs are created by these external investments and therefore stimulate both the labor force and fiscal balances in a community. A simulation model usually begins with the labor force models. Since the labor force models are simultaneous, the assumed exogenous shock (such as a change in place of work employment in variable POWEMP) will not give a unique solution to the system. A solution however can be achieved by converting the labor force models into the reduced form through substituting. After substituting, the labor force model can be rewritten as:

$$LF_i = (\beta_0^1 + \beta_2^1\beta_0^2 + \beta_3^1\beta_0^3) + (\beta_1^1 + \beta_2^1\beta_1^2 + \beta_3^1\beta_1^3)POWEMP + \beta_2^1\beta_2^2 XLF + \beta_3^1\beta_2^3 XEMP \tag{19}$$

The reduced form labor force equation, the population growth equation, and the fiscal models can be further written by substituting the corresponding estimated coefficients:

$$POP_i = \beta_0^4 + \beta_1^4 LF \tag{20}$$

$$PCEXP_{i,t} = \alpha_{0,i} + \alpha_1 LNPOWEMP_{i,t}\alpha_2 LNUNEMP_{i,t} + \alpha_3 LNPOPDEN_{i,t} + \alpha_4 LGPCB_{i,t} \quad (21)$$

$$PCREV_{i,t}b_{o,i} + b_1 LNPOWEMP_{i,t}b_2 LNUNEMP_{i,t} + b_3 LGPCB_{i,t} \quad (22)$$

To incorporate the errors involved in estimating the labor force and fiscal impact equation systems, estimated coefficients are drawn using the approach specified by Krinsky and Robb (1986). The coefficients are assumed to be drawn from a multivariate normal distribution with mean and covariance matrix given by the estimation results in the two systems, respectively.

Table 18 and Table 19 summarize the results after assuming the place of work employment changes by 1% for each county within the state of Nevada. In both tables, various quantities measured before and after the change are reported. The purpose of presenting the values before the change is to verify the robustness of the simulation process by comparing the simulated measures with the actual data. Simple comparisons show that the simulations indeed produce reliable results. The mean estimates and their associated standard deviations reported in Table 18 and Table 19 are obtained after 1,000 simulation replications. All results are significantly based on the relative magnitude of the mean estimates to their corresponding standard deviations. Table 18 gives the labor module simulation results for the overall labor force and population. For each county, the increase of 1% place of work employment will yield approximately 1,083 person increase in the overall labor force. This is because the coefficient of variable POWEMP is translated to approximately one in its relationship to variable LF (labor force).

Table 18: Simulation Results of the Labor Force Module

Labor Force						
Counties	Base	EMP 1%	Change (after-before)	% Change	T-Test stat	P-value
NEVADA	**1,072,687**	**1,083,183**	**10,497**	**0.98%**	**1.42**	**0.16**
Carson City	38,125	38,364	240	0.63%	1.70	0.09
Churchill	6,407	6,474	67	1.05%	1.70	0.09
Clark	782,158	789,948	7,789	1.00%	1.16	0.25
Douglas	19,111	19,264	153	0.80%	1.00	0.32
Elko	16,778	16,979	202	1.20%	1.78	0.08
Esmeralda						
Eureka	672	714	42	6.29%	0.98	0.33
Humboldt	4,142	4,215	72	1.74%	1.69	0.09
Lander	364	391	28	7.60%	0.64	0.52
Lincoln						
Lyon	770	780	11	1.37%	0.18	0.86
Mineral						
Nye	7,225	7,321	96	1.32%	1.05	0.30
Pershing						
Storey	25,513	25,570	57	0.22%	0.40	0.69
Washoe	171,135	172,841	1,705	1.00%	1.14	0.26
White Pine	287	322	35	12.35%	1.32	0.19

Population						
Counties	Base	EMP 1%	Change (after-before)	% Change	T-Test stat	P-value
NEVADA	**2,531,422**	**2,555,252**	**28,830**	**0.94%**	**1.52**	**0.13**
Carson City	92,798	93,342	544	0.59%	1.80	0.07
Churchill	20,943	21,095	152	0.73%	1.97	0.05
Clark	1,771,484	1,789,168	17,684	1.00%	1.21	0.23
Douglas	49,664	50,012	348	0.70%	1.05	0.29
Elko	44,345	44,803	458	1.30%	2.00	0.05
Esmeralda	6,513	6,513	0	0.00%		
Eureka	8,032	8,128	96	1.19%	0.99	0.33
Humboldt	15,837	16,001	164	1.04%	1.79	0.07
Lander	7,314	7,377	63	0.86%	0.65	0.52
Lincoln	6,491	6,491	0	0.00%		
Lyon	8,000	8,024	24	0.30%	0.27	0.79
Mineral	6,495	6,495	0	0.00%		
Nye	22,703	22,919	217	0.96%	1.22	0.22
Pershing	6,484	6,484	0	0.00%		
Storey	64,417	64,547	130	0.02%	0.39	0.69
Washoe	392,781	396,652	3,871	0.99%	1.18	0.24
White Pine	7,119	7,200	80	1.13%	1.43	0.15

The standard deviation associated with each county's labor force gives an interval range for which the estimated labor force may fall after the change of place of work employment. For example, for the county of Carson City, the predicted increase in the labor force after an increase in place of work employment of 1% is approximately 38,364 with a t-test statistic of 1.7 which coincides with a p-value of 9%. Therefore, when analyzing the impact of the increased place of work employment on the labor force, county authorities or county economic development professionals should use the interval for economic and fiscal impact analysis rather than a simple mean estimate. The same principle applies to other aspects in both the labor module and the fiscal module. However, calculated standard deviation estimates for sparsely populated rural Nevada counties such as Esmeralda County are quite large compared with the mean estimates. Predictions in these counties can be impacted greatly by outliers. For population, the change in each county is approximately 1%. This can also be explained by the coefficient of variable LF in the model of population, which is approximately 2. Overall, the labor module indicates a magnifying effect of place of work employment on population growth. In other words, one unit increase in place-of-work employment will create a one-unit increase in the labor force and subsequently cause the population to increase by two units.

Table 19: Simulation Results of Fiscal Model and Chances of a Negative Fiscal

	Per Capita Fiscal Balance (dollars)					Probability of a Negative Per Capita Fiscal Balance		
Counties	Base Avg.	EMP 1%	Change	T-Test stat	P-value	Base Avg.	EMP 1%	Change
NEVADA	88	148	60	2.31	0.02	19%	22%	-7%
Carson City	-14	289	303	15.66	0.00	54%	3%	-51%
Churchill	6	13	7	1.15	0.25	44%	39%	-6%
Clark	-4	-117	-113	-7.66	0.00	53%	86%	33%
Douglas	3	111	108	9.63	0.00	48%	10%	-39%
Elko	-47	78	125	12.19	0.00	73%	15%	-59%
Esmeralda	7	9	2	0.02	0.99	45%	45%	0%
Eureka	1,238	1,974	736	1.79	0.07	30%	22%	-8%
Humboldt	41	132	91	6.95	0.00	34%	8%	-26%
Lander	497	597	100	1.63	0.10	14%	6%	-8%
Lincoln	-237	-49	188	9.55	0.00	96%	66%	-30%
Lyon	-21	-43	-22	-2.55	0.01	68%	78%	10%
Mineral	-121	-121	0	-0.01	0.99	79%	78%	0%
Nye	-6	-15	-9	-0.91	0.37	53%	55%	2%
Pershing	17	70	53	2.62	0.01	45%	26%	-19%
Storey	106	118	12	0.17	0.87	35%	35%	0%
Washoe	26	-34	-60	-5.88	0.00	34%	67%	32%
White Pine	6	-187	-193	-3.94	0.00	53%	37100%	18%

Similarly, the fiscal module is simulated in accordance with the increase in place of work employment. From the mean estimate, the increase of POWEMP increases per capita expenditure in most counties, except for Carson City, Eureka, Storey, and Washoe counties. This can be explained: although an increase in POWEMP will lead to an increase in population, which by theory and by the model coefficient in this study, will decrease the per capita expenditure, the direct impact of POWEMP and the induced impact of unemployment act positively on per capita expenditures. Depending on the relative strength of these variables and the magnitude of other variables in the expenditure model, the effects on per capita expenditures may either be positive or negative. Nevertheless, these comparisons are only based on the mean estimates of expenditures before and after the change. When the standard deviations are considered, the interval of per capita expenditure may overlap (such as in Storey county), which in turn indicates that for these counties based on a certain level of statistical confidence, the change in per capita expenditures may be either positive or negative. For the per capita revenue, results based on the mean estimates show that all changes are positive. When standard deviations associated with each mean estimate are considered, these differences are also not fixed. For example, a 1% POWEMP increase in Clark County may increase the county's per capita revenues by 0.23 cents based on mean estimates but the same result does not necessarily hold when standard deviations are incorporated. Model results that incorporated both mean and standard deviation estimates affirm the concerns of Chambers and Anderson (1977) that uncertainty is the essence of the planning problem and that the public may not be well served by fiscal analysis that incorporates only the most likely future.

Table 19 presents the simulated results of a 1% change from the base level in POWEMP on the probability of a negative fiscal balance. Overall, the average Nevada county reduces the probability of a negative fiscal balance by 7%. The change in probability from the base level values differs widely when examining individual counties. Washoe County for example has a 32% increase in the probability of a negative fiscal balance. This result aligns with results from Table 19 where the fiscal module simulation for Washoe indicates a 19.01% increase in per capita expenditures but only a 9.31% increase in per capita revenues.

6.9 Conclusions

In this article a system that jointly considers both the labor and fiscal aspects of the county-level decision-making process are analyzed. In the labor module, it is found that place of work employment has positive impacts on the overall labor force, incommuters, and outcommuters. The endogenous variables incommuters and outcommuters also have important implications on the overall labor force. The growth of the county population can be explained relatively well by the labor force. Model results indicate a one-person labor force increase yields a two-person increase in county population. The fiscal module is closely related to the labor module since some key factors in the labor sector are significant in explaining the fiscal status of a county, and that increased population density tends to lower county per capita expenditure. Increasing place of work employment may boost the size of the counties' economy. The fiscal balance from the previous year also has important implications on the current year's, cut government expenditures and revenues. Incorporating previous year balances enhances the county fiscal model by incorporating dynamic relationships in county budgets, which is often ignored.

The effects of linking the labor and the fiscal modules are further demonstrated by a series of simulations. Changes are assumed to be introduced exogenously to the place of work employment. The simulation results provide not only a replication of current situations but also predictions under the exogenous changes. Most importantly, the results show that if only mean estimates are considered, county governments' decisions may be biased in terms of reflecting the magnitude and even the direction of the impacts of changes. The statistical confidence intervals provided in this study may help the decision makers to incorporate uncertainty in their planning process by deriving potential impacts other than the most likely future. One way of future modification may be using the approach outlined in Harris et al. (2000) where they decompose the labor force and employment into sectors according to their nature, such as tradable labor and non-tradable labor. Finally, such a fiscal and labor model might be adapted to other exogenous shocks such as the COVID-19 pandemic.

7 Constructing Precinct Level Income Variables Using Deep Learning

Randall Chicola (University of Nevada, Reno)

Dr. Dilek Uz (University of Nevada, Reno)

7.1 Abstract

While precinct-level voting analyses are common in political science, there is a dearth of readily available income data with a matching geographical granularity. This work uses a sequential class of deep learning neural network (NN), county assessor data, including assessor home price data, and census block group level income data from Clark County, the most populous county of Nevada, to construct income variables at the precinct level. The model is trained for each median income and percentage of people in various income brackets at the Census block group level, and the resulting fitted model is then applied to comparable data at the precinct level. The neural network model provided a 6% reduction in the MSE compared to a regression model. The constructed income variables will subsequently be used in the analysis of the voting behavior of Nevadans during the 2016 and 2018 elections.

7.2 Introduction

In the United States, a precinct is a defined geographic area, often within a municipality or county, used as a unit for organizing and conducting elections. It typically represents a local voting district and is administered by election officials to facilitate the voting process.

Precinct-level voting analyses are highly prevalent in political science and economics. Cantoni (2020), Broxterman and Jin (2022), and Finan et al. (2021) are recent examples in which controlling for income at the precinct level could be beneficial for analysis. While precincts are often delineated based on population size and geographical boundaries, they do not perfectly overlap with census boundaries. Therefore, there is no straightforward way to control for demographic variables at the precinct level. The goal of this chapter is to contribute to the research methodology by developing and implementing a machine-learning algorithm to construct income variables at the precinct level.

In order to accomplish this, the attributes of the home parcels were scraped from the Clark County Assessor website. This data was then separately aggregated to the Census block group and precinct levels. The Census block group-level data was used to train a deep-learning model. The trained model was then applied to the precinct-level data. For robustness, the model was compared to linear regression, and mean squared error (MSE) was then compared to determine the relative performance of the machine learning method.

Previously, other researchers worked on this issue. Amos et al. (2017) examine four spatial methods to infer political behavior. While previous approaches utilize spatial extrapolation methods, our work employs a machine learning approach and incorporates detailed housing information, separating it from existing research. We will then use the constructed income variables in analyzing Nevadans' voting behavior in the 2016 and 2018 referendums which involves a very specific sequence of events.

The rest of the chapter continues as follows: the next section provides an overview of the events that unfolded around the time of the referenda. Section 2 gives a brief literature review. Section 3 explains the variables. Section 4 summarizes the machine learning method. Section 5 presents the results and section 6 offers concluding remarks.

7.3 Background

7.3.1 Net Metering Debate in Nevada

Distributed generation (DG) refers to the generation of electricity from small-scale power sources located near the point of use, typically close to where the electricity will be consumed. These sources can include photovoltaic (PV) solar panels, wind turbines, fuel cells, microturbines, and small natural gas generators. Distributed generation contrasts with centralized power generation, where electricity is produced by large remote power plants and then transmitted over long distances through the electrical grid.

The DG capacity in the United States has experienced rapid growth over the past couple of decades. Accompanying the technological advances of increasing solar energy capacity are changes in electric utility rate tariff structures to account for these distributed systems being added to the grid. A household's surplus solar power from DG systems pushed to the grid offsets the times that it draws electricity from a utility, the billing process of which is referred to as "net energy metering" (NEM). Renewable portfolio standards provided additional impetus for distributed solar generation growth.

Under NEM, customers with rooftop PV systems receive credits for each kWh of solar electricity they generate at the retail electricity rate. However, due to the fundamental nature of the electric utility business, a significant portion of the revenues via retail rates goes towards covering fixed costs, such as transmission and distribution infrastructure, and public initiatives such as energy efficiency programs. Consequently, as shown by Darghouth et al. (2011), crediting NEM customers at the retail rate greatly exceeds the social value of solar power, even when considering benefits such as reduced air pollution, greenhouse gas emissions, and transmission losses. Essentially, net metering subsidizes PV adopters by transferring more fixed costs to the bills of the rest of the ratepayers, which is referred to as cost shifting.

The Public Utilities Commission of Nevada (PUCN) initially established a quota of 3% of peak demand, which quickly filled in summer of 2015. Requests to increase or lift the quota led to concerns of unfair cost shifting. An investigative study Energy and Environmental Economics Incorporated (2016) which investigated this concern found rooftop solar users shift $36 million in costs to non-solar ratepayers annually. Utilities charge a per kilowatt (kWh) volumetric charge and solar users were receiving a rebate funded by ratepayers while simultaneously buying less from the utility and paying lower variable volumetric prices. Utility revenue requirements would then be shifted to the remaining customers. Typically, homeowners who were able to afford solar installation were more affluent, so an equity issue arose driven by this cost-shifting mechanism that led the PUNC to consider modifications.

Modifications began in December 2015 when PUNC decided to gradually reduce the net metering credit from 11c/kwh to 2c/kwh and increase their monthly fixed rate from $13 to $38 for net metered customers. Additionally, the PUCN decided not to grandfather in the existing net metering arrangements. Public outrage and lawsuits ensued and the solar lobby launched an aggressive communications campaign critical of the utility company NV Energy.

The cost-shifting dynamics of the study were not generally well understood by the public, so the lawsuits and referenda that ensued became a public indictment of NV Energy. The public perception was that NV Energy did not want to lose its monopoly status or share the electricity revenues with the citizens. In September 2016, a Nevada court overturned higher fixed charges for existing rooftop solar customers but upheld the net metering and fixed charge changes for new rooftop solar customers. In September 2016 NV Energy filed with Nevada an agreement to grandfather 32,000 existing residential rooftop solar users to credit them at the retail rate.

7.3.2 Debate Around the Casinos Leaving the Nevada Electric Utilities

The net metering issue was used as a proxy issue by the casinos that were simultaneously embroiled with NV Energy in an exit fee dispute, which was unrelated to the net metering debate. The casinos wished to lower their bill by buying electricity from 3rd party providers or generating it on their own. NV Energy objected that irreversible capacity and transmission investments were made to serve these casinos per their obligation as the regulated utility. NV Energy argued that the casinos need to pay an exit fee if they want to terminate their contract, otherwise, there would be a large cost-shifting to the remaining Nevada rate-payers.

Casino owners funded a campaign called "Nevadans for Affordable, Clean Energy Choices". In Nevada, anyone can add a question to the ballot during the elections if enough signatures are collected. "Question 3" was added to the ballot for the 2016 elections and passed with 72.36%. In Nevada, initiated constitutional amendments need approval in two successive general elections, so Question 3 was also in the 2018 election. This time, however, Nevadans voted not to approve with a 67.05%, a sharp contrast to 2016 results.

The text of Question 3 is as follows:

"Shall Article 1 of the Nevada Constitution be amended to require the Legislature to provide by law for the establishment of an open, competitive retail electric energy market that prohibits the granting of monopolies and exclusive franchises for the generation of electricity?"

These referenda provided an opportunity for an event study approach to demonstrate whether voting behavior might be affected by solar ownership by exploiting the unique sequence of events and subsequent referenda in Nevada. To accomplish this larger endeavor, however, this required precinct-level income data which was not readily available as with other geographic regions. This motivates this work, which aims at inferring a reasonable proxy income and other demographic variables related to Nevada homeowners that can be combined with solar panel installation data.

7.4 Literature Review

The relevant literature pertains to the technical necessities of understanding the mathematical structure of machine learning computation and applied methods in economics, its practical application to relevant areas for this research question, and the solar energy research question applications that this work is a stepping stone towards.

7.4.1 Machine Learning Mathematics and Methods in Economics

While the algorithmic design of machine learning has been established for quite some time, its popularity, prominence, and potential were not fully appreciated until an inflection point between hardware capability and software frameworks developed to where a critical density of researchers could leverage those capabilities. Designating the best moment to begin the discussion of these advances is up for deliberation. While the neural event logic proposed by McCulloch and Pitts (1943) or the perceptron by Rosenblatt (1958) might arguably be reasonable starting points of machine learning advances, a full history is not the goal. Rather a reasonable branching point in the story-line is the goal to give a flavor of the methods and advancements that lead to the methods used here.

Mathew et al. (2021) provides a concise summary of the evolutionary timeline of the machine learning field. Also, the rationale for using deep learning methods over traditional machine learning is made; the performance threshold continues to improve with deep learning as more data is brought to bear, while traditional method performance flattens as more data is added. Additionally, Mathew et al. (2021) describes the use cases for supervised learning models and guidance as to why the methods would be appropriate for this work. With a supervised learning model that maps a function of data to an output variable, the goal is to use the errors generated from the training dataset to correct the output. Stochastic gradient descent obtains an optimum value with a convex function to prevent a solution from being trapped in a local minimum. Other literature such as Berner et al. (2021) goes further into mathematical detail of kernel regimes, challenges of deep learning architectures with a high number of input dimensions, and topology of loss landscapes.

Nosratabadi et al. (2020) also provides an exhaustive review of deep learning methods and their application domains in economics. They demonstrate the increasing prominence of machine learning methods in economics growing exponentially in 2017 as measured by the documents in Thomson Reuters Web-of-Science. In particular, it emphasizes the advantages of methods like long short-term memory (LSTM) networks and deep learning for certain use cases like stock market prediction.

Optimization paradigms for stochastic programming are a crucial element in deep learning models and are highlighted in the Process Systems Engineering (PSE) community by Ning and You (2019), comparing the alternatives of chance-constrained optimization and robust optimization. They demonstrate that the challenges of employing large datasets with stochastic elements using chance-constrained paradigms often lead to computational constraints because it involves multivariate integration and additionally the feasible region is not convex (LeCun et al., 2015).

7.4.2 Precinct-level Income

The addition of precinct-level income data has the potential to complement the existing body of research in political science, public economics, and related fields. Quite often public data from The Census Bureau reports at the county or Census tract level. Public policies are usually directly affected by voter input such as referenda or state constitutional amendments. Additionally, political analysis and policy research institutes often focus on precinct-level voter activity. Such research and analysis could be augmented with income data at this geographical unit.

For example, in public economics, Kaplan et al. (2022) measured spatial sorting by creating a variance-based index that can decompose partisanship heterogeneity into differences between and among communities. Additionally, their index satisfies seven axiomatic properties that identify the mathematical structure of ideological partisanship as being analogous to index measures of inequality. Their axiomatic approach specifies that this heterogeneity can be decomposed into different income groups at various levels of aggregation. Continuing research at the precinct level of aggregation would benefit from precinct-level income data provided by the machine learning estimation method we contribute, as their work focuses on county-level variance.

Newman et al. (2021) investigate gendered language in political initiatives and specifically the effect greater female control over economic resources has on popular support for the gender-neutral language implemented in various state and local governments in the United States including California and Lincoln, Nebraska. Along with other metrics, they employ the use of a male-to-female income ratio in multivariate regression for the various cities being analyzed. Their full model results would be bolstered by precinct-level income as well.

Studies of geographic inequality and land use such as those investigated by Trounstine (2020) would also benefit from precinct-level income data. The inescapable connection between income, property values, and public amenities is highlighted by their work. Their findings of rigid land use regulations being more prevalent in whiter communities at the precinct level in California would have greater precision if their work had precinct-level data rather than specification using the share of the city above the 90th percentile in income.

Investigating precinct-level income for the larger purpose of investigating rooftop solar referenda voting patterns is a classification problem. Deep learning has been used to address other economic classification tasks. For macroeconomic tasks, Maliar and Maliar (2022) create a deep learning classification (DLC) method for continuous or discrete labor choices of hours worked and binary labor choice. They utilize the sigmoid activation function for its favorable derivative and unit interval mapping properties for their purposes and logistic regression for their discrete choice model. Utilizing deep learning models for the classification of income brackets in applied microeconomics does not appear to have been done and is the contribution of this work.

7.4.3 Precinct-level Voting in Political Science and Economics

Cantoni (2020) studies the effect of polling distances on voting costs in Minnesota and Massachusetts, focusing on the two channels of voter inactivity purging registered voters off the rolls and dissuasion of registration. By looking at parcel resident vote casting rather than registered voter turnout percentages, he finds that about a quarter-mile distance reduced cast ballots by 2 to 5% in the four elections analyzed. His analysis at the parcel level to avoid endogeneity is comparable as our income prediction uses parcel and home characteristics to estimate income. Future analyses such as his could better control for income in addition to using distance from the parcel to the polling place as the explanatory variable which would support efforts to address sources of bias for suburban and urban voters.

Similarly, Broxterman and Jin (2022) examines the effect of a natural experiment that implemented an appointed financial control board on housing prices to determine whether voters of the incumbent administration that was replaced were not voting in their economic self-interest. They use zip code-level Census data to estimate the appreciation rate of home prices using a triple differences (DDD) model. Having precinct-level income estimates would complement their DDD modeling approach.

Finan et al. (2021) also examines block-level data to examine evidence of neighbors' effect on voting behavior in Mexico. A similar study in U.S. precincts could similarly identify the close neighbor's influence on voting while controlling for income. Furthermore, this methodology could use home price data in Mexico to control for income in the block and precinct-level study Finan et al. (2021) perform.

7.5 Data

7.5.1 Overview of Sources and Pre-Processing

The property features used are from address-level housing data from the Clark County assessors in Nevada.[§] This data was then geocoded and merged with corresponding census block group-level income and precinct data. The list of properties requested directly from Clark County Nevada contained the Assessor's Parcel Number (APN). The precinct number associated with each APN was procured from the Legislative Counsel Bureau Research Division. The web-scraping script used this APN list to iterate over the property search tool on the county assessor's website and then tried to pull property feature data elements with either object IDs or Cascading Style Sheets (CSS) selectors. This data collection occurred in early 2023 to collect tax-assessed values for both 2021 and 2022.

The web-scraped APNs then required some modification for more amenable use in the census block group and precinct-level analysis. This included address cleaning, re-categorization of some variables, and creation of descriptive statistic variables for those geographic units of granularity. The latitude, longitude, and income bracket data were obtained from Geocodio. The income data attached was the Census Bureau's 2019 American Community Survey (ACS) household income data. Finally, this income data was merged into the individual property features along with a list of APNs and their associated precinct numbers so that we could separately collapse at the precinct level as well as the census block group level.

7.5.2 County Assessor Property Data

Typical data features available include numerous property features, some of which are common elements in general real estate while others are more specific to an individual county.

[§]This chapter for neural network modeling focuses on Clark County but can be extended to Washoe County later as well.

To develop a model that can be applied to multiple counties, common features were chosen. Taxable assessed values of the properties, improvements, any prior sale information, square footage, bedrooms, bathrooms, and similar home attributes were commonly available.

Some of these common county property features needed to be filtered further to focus on properties that relate to household income. For example, land use features would exclude observations related to irrelevant land use categories such as commercial, industrial, churches, hotels, golf courses, and government facilities. Similarly, some categorical variables such as construction type, roof type, home style, and the sale type of the property's most recent sale were re-categorized to make the variable more amendable to modeling. For example, multiple types of foreclosures were rolled up into one foreclosure category. Since this household-level data would be aggregated to either the Census block group, Census block, or precinct level, these variables were translated to the mean, median, percentile, and standard deviation at the Census block group and precinct levels.

7.5.3 Geo-location

Adding latitude and longitude data can benefit the flexibility for the constructed data to be used for other purposes. Using Geocodio allowed an amenable way to merge the income data needed but also to merge latitude and longitude data so that solar panel data installation information could then be easily merged for later analyzing the effects of predicted income and solar installation on voter referendum results related to Nevada's energy policies. The geo-location process for the home feature was done in batches of 2,000 over time to comply with the site's free feature functionality.

7.5.4 Income Brackets

Once household-level home price data features are obtained, income data must be merged. Ideally, household-level income data would be the most accurate, but instead of directly knowing each household's income as well as their home features, Census income data at some geographic level must be merged with the scrapped home features for the modeling to proceed. This means there is a research choice to be made as there is a natural trade-off in modeling with different granularity levels such as Census block groups or Census blocks. Census blocks provide a larger set for the neural network to train on which is advantageous. However, a larger granularity such as block groups is more statistically accurate as each block group contains between 600 and 3,000 people. Census blocks are known to be less reliable statistically since there are fewer people in a geographic area that is partitioned further.

There are sixteen income brackets as well as median income that can be estimated where the explanatory income bracket is expressed as a percentage of people in that income bracket at some geographic level such as block groups. Initially, the neural network models each bracket separately unconstrained by the condition for all brackets to sum to one to get a sense of how accurate the method is without any a priori knowledge of how the income brackets are related to one another. Then the sigmoid function is applied to bound the predicted income bracket percentages between zero and one. Further variants can predict median income dollar values rather than percentages or simultaneously solve all sixteen income brackets.

7.5.5 Clark County Nevada Descriptive Statistics

Table 20 shows the property characteristics for Clark County.

Table 20: Clark County Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Total Taxable Value	519,513	$257,400.30	$568,507.30	$500	$73,700,000
Acres	519,517	0.15	0.35	0	78
Total Square Footage	519,517	1,825.72	908.38	0	43,557
Total Garage (Sqr feet)	519,517	385.23	270.38	0	9,160
Bedrooms	519,517	3.13	1.08	0	71
Full Bath	519,517	2.14	0.72	0	30
Half Bath	519,517	0.33	0.49	0	22
Fireplace	519,517	0.57	0.65	0	12

7.6 Methodology

7.6.1 Overview: Deep Learning Neural Network Model

The deep learning neural network or deep neural network (D.N.N.) utilizes many so-called "Hidden Layers" or "Dense Layers" that receive information from either the input layer or other hidden layers that have already been trained. Each layer applies kernels called a "filter bank" to generate feature maps. Each kernel is a matrix or tensor that intertwines the weights learned in the training process with the input data itself. The deep learning neural network then adjusts the parameters of the filter bank kernels to minimize a selected loss function to perform the user's end task, such as the classification of income brackets based on property features.

As illustrated by Nosratabadi et al. (2020), the D.N.N. can be generally represented as:

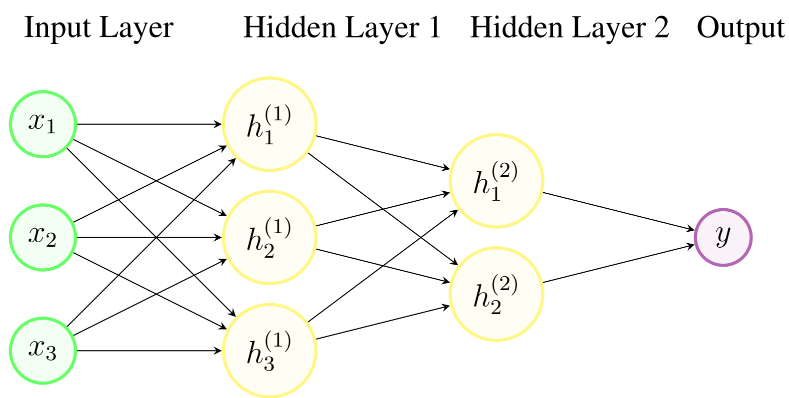
$$f(X) = f_k(...f_2(f_1(X; w_1); w_2)..., w_k) \tag{23}$$

For this work, two hidden layers ($k = 2$) are used and might be represented as:

$$f(X) = f_2(f_1(X; w_1); w_2) \tag{24}$$

Figure 7 is a visual representation of the model layers, albeit with fewer neural nodes for brevity. In actuality, the first layer of the model has 64 neurons while the second has 32 neurons. Several architecture design decisions were made to focus on the most integral advantages of machine learning while minimizing computational intensity.

Figure 7: Deep Learning Neural Network



7.6.2 Imputer and Standardization

After the model was split into an 80% training and a 20% test data set, the next step was addressing missing values with some imputing strategy. Since machine learning algorithms such as neural networks are unable to deal with missing values directly without errors or biased results, imputation becomes a necessary step in data processing. Mean was chosen for the benchmark imputation strategy and future work would include using alternative strategies such as median, mode, k-nearest neighbors (KNN), or regression-based imputation to further check for robustness. Once the dataset is complete following the imputation process, standardization is applied independently to each data feature which makes compiling the neural network architecture quicker and more efficient. These pre-processing steps help ensure consistency across the training and test datasets as well as ensure that the model can be generalized.

7.6.3 Dense Layers and Rectified Linear Unit (ReLU)

The Rectified Linear Unit (ReLU) is defined as $f(x) = max(0, x)$ as described by Hara et al. (2015). ReLU is computationally less burdensome than the alternatives of the sigmoid or tanh activation functions because it uses such a simple threshold operation with a straightforward derivative of either zero or one. Consequently, compared to tanh or sigmoid, ReLU is less susceptible to the vanishing gradient problem where backpropagation gradients become too small, which slows the learning rate. ReLU can mitigate this because as long as positive inputs are received, ReLU yields a gradient constant of one. While the non-zero gradient is maintained, errors can then be backpropagated so that the weights can be updated. Preventing this in the early layers is particularly crucial for preventing slow learning rates. ReLU is less likely to result in an overfitting model because when it receives a negative input value, it outputs zero which contributes to more sparsely activated neurons. There is an additional benefit, as it allows for a simple way to model complex non-linear relationships. When an input is positive, the resulting output is directly proportional, but when the input is negative the output is zero. This abrupt change from a linear relationship to zero, or vice versa, introduces non-linearity. This mechanism makes ReLU more robust by allowing for complex nonlinear features to be captured more accurately than by linear transformations alone.

7.6.4 Adam: Stochastic Gradient Descent Optimizer

The Adaptive Moment Estimation (Adam) optimization algorithm for stochastic gradient descent (SGD) amalgamates components of Root Mean Square Propagation (RMSprop) and Momentum optimization methods. With each parameter being optimized in RMSprop, the average of recent gradient computations is used to adjust the learning rate. As detailed by Kingma and Ba (2014), these adjusted learning rates are then divided by the square root of the exponentially decaying average of the squared gradients. To expedite the gradient descent process by adding inertia to the process, momentum optimization adds a fraction of the prior update to the current update to navigate computational issues such as flat regions, small gradients, or narrow valleys. Adam retains the mean (first moment) and the uncentered variance (second moment) to utilize both moving averages, thereby combining the two concepts that underpin RMSprop and Momentum.

The Adam process is as follows:

Step 1: First-moment estimate bias updated.

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \tag{25}$$

Step 2: Second-moment estimate updated.

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2)(g_t)^2 \tag{26}$$

Step 3: First-moment bias correction.

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \tag{27}$$

Step 4: Second-moment bias correction.

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \tag{28}$$

Step 5: Parameters updated.

$$\theta_{t+1} = \theta_t - \frac{\alpha}{1 - \sqrt{\hat{v}_t + \epsilon}} \hat{m}_t \tag{29}$$

where: α is the learning rate, β_1 and β_2 are the rates of exponential decay for the moment estimates such that $\beta_1, \beta_2 \in [0, 1)$, m_t is the first-moment vector, v_t is the second-moment vector, g_t is the gradient, ϵ is a small constant to prevent an undefined fraction. During step 1, the weights and biases (model parameters) often utilize a randomization scheme such as Xavier or He initialization. The first and second-moment vectors are initialized at zero typically before training the model.

The iterative process of Adam begins by taking initial values for the parameters θ, the first-moment vector $m = 0$, the second moment $v = 0$, and the time step $t = 0$. Using a fixed number of training observations that is a subset of the actual dataset or so-called "mini-batch", with every t iteration, the current mini-batch gradient g_t is computed. Adam incrementally modifies α depending on the gradients of the first and second moment exponential decay and corrects for the moments' biases.

7.7 Results

7.7.1 Unconstrained Income Bracket Percentages

Results from the 1,028 precincts estimated from training on 17,267 blocks show interesting results. Depending on the adjustments in a run, the prediction results were mixed. Some predicted blocks would sum up fairly close to one when summing the predicted income groups while others were more varied. This is shown in both training predictions on blocks in Table 21 as well as precincts in Table 22. The sixteen income groups are numbered lowest to highest brackets and the income ranges are defined in Table 28 in the appendix and numbered for the heat maps beginning with Figure 12 in the appendix. The mean squared error from the training of the model is presented in Table 23. Additionally, the histogram displayed in Figure 8 shows that the summation of the income shares predicted by the neural has a center of mass around one when looking at all the observations together.

After the neural network runs for the training data at the block levels, the predicted model results from the block training level data are used on the precinct level collapsed data to estimate the income. The precinct results presented in Figure 9 had a center of mass around 93% rather than around 100% as with the block training estimation which might be expected with negative values.

Table 21: Sample of Training Block Prediction Results

block ID	**283110121**	**535210161**	**583410081**	**790020082**	**491810041**	**533530103**
lt 10K pct	6.21%	2.88%	-0.46%	16.06%	3.68%	2.01%
btw 10k 15k pct	2.78%	1.64%	5.62%	7.59%	1.75%	2.40%
btw 15k 20k pct	2.78%	1.64%	5.62%	7.59%	1.75%	2.40%
btw 20k 25k pct	3.30%	2.95%	2.60%	8.75%	2.52%	2.62%
btw 25k 30k pct	3.36%	2.78%	1.56%	5.73%	2.99%	2.88%
btw 30k 35k pct	1.11%	2.42%	-0.68%	9.17%	2.03%	2.40%
btw 35k 40k pct	4.23%	2.32%	3.43%	3.73%	3.67%	2.05%
btw 40k 45k pct	6.85%	1.69%	4.57%	7.33%	3.91%	3.79%
btw 45k 50k pct	2.69%	4.69%	5.45%	2.83%	3.64%	5.48%
btw 50k 60k pct	9.65%	3.84%	6.87%	10.62%	10.63%	6.72%
btw 60k 75k pct	9.91%	6.06%	7.58%	7.83%	8.60%	15.20%
btw 75k 100k pct	12.69%	13.69%	21.40%	9.81%	23.55%	19.16%
btw 100k 125k pct	9.73%	13.72%	9.90%	8.19%	12.40%	6.82%
btw 125k 150k pct	7.98%	10.00%	8.67%	3.68%	6.65%	7.78%
btw 150k 200k pct	5.32%	7.82%	9.69%	3.23%	15.22%	14.94%
gt 200k pct	6.21%	10.84%	11.68%	6.85%	3.31%	7.20%
Summation	94.77%	89.00%	103.49%	118.98%	106.31%	103.85%

Table 22: Sample of Precinct Prediction Results

Precinct	**1038**	**1039**	**1063**	**1067**
lt 10K pct	8.7%	12.3%	8.7%	11.8%
btw 10k 15k pct	0.8%	1.5%	4.7%	7.0%
btw 15k 20k pct	4.0%	2.1%	3.2%	11.1%
btw 20k 25k pct	2.4%	3.0%	8.0%	12.4%
btw 25k 30k pct	-2.6%	3.0%	4.7%	7.1%
btw 30k 35k pct	2.5%	5.8%	5.8%	4.3%
btw 35k 40k pct	2.6%	6.2%	6.5%	7.0%
btw 40k 45k pct	2.1%	2.6%	4.1%	8.6%
btw 45k 50k pct	3.4%	5.8%	7.3%	4.0%
btw 50k 60k pct	4.3%	5.5%	8.3%	6.5%
btw 60k 75k pct	2.5%	5.8%	5.8%	4.3%
btw 75k 100k pct	13.0%	17.2%	16.0%	7.4%
btw 100k 125k pct	10.6%	9.6%	7.4%	4.1%
btw 125k 150k pct	4.7%	6.4%	4.5%	2.0%
btw 150k 200k pct	9.1%	11.3%	3.7%	2.1%
gt 200k pct	17.0%	21.3%	-1.1%	2.2%
Summation	84.9%	119.6%	97.5%	102.0%

Table 23: Mean Squared Error of the Training Blocks

Income Bracket	MSE
lt 10K pct	0.002971389
btw 10k 15k pct	0.002031843
btw 15k 20k pct	0.001618739
btw 20k 25k pct	0.001530719
btw 25k 30k pct	0.00149093
btw 30k 35k pct	0.001667177
btw 35k 40k pct	0.00218023
btw 40k 45k pct	0.00156843
btw 45k 50k pct	0.001515624
btw 50k 60k pct	0.002596316
btw 60k 75k pct	0.003324356
btw 75k 100k pct	0.004732762
btw 100k 125k pct	0.003588215
btw 125k 150k pct	0.002356644
btw 150k 200k pct	0.003144305
gt 200k pct	0.004480462

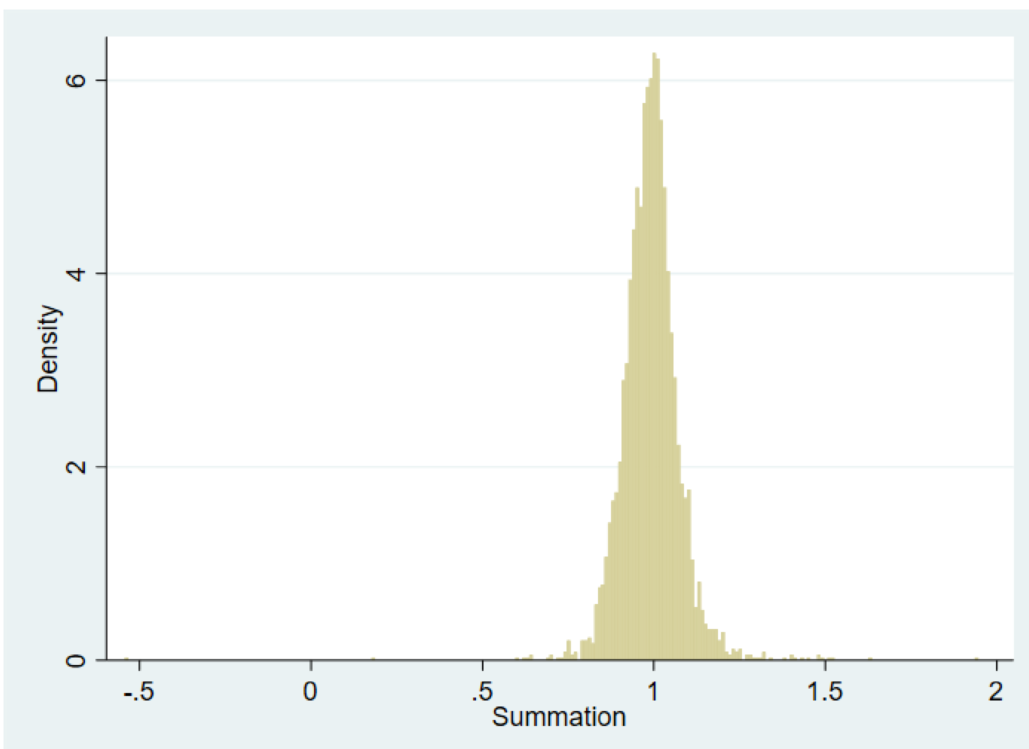


Figure 8: Distribution of all 16 Income Group Summations of Block Training Data

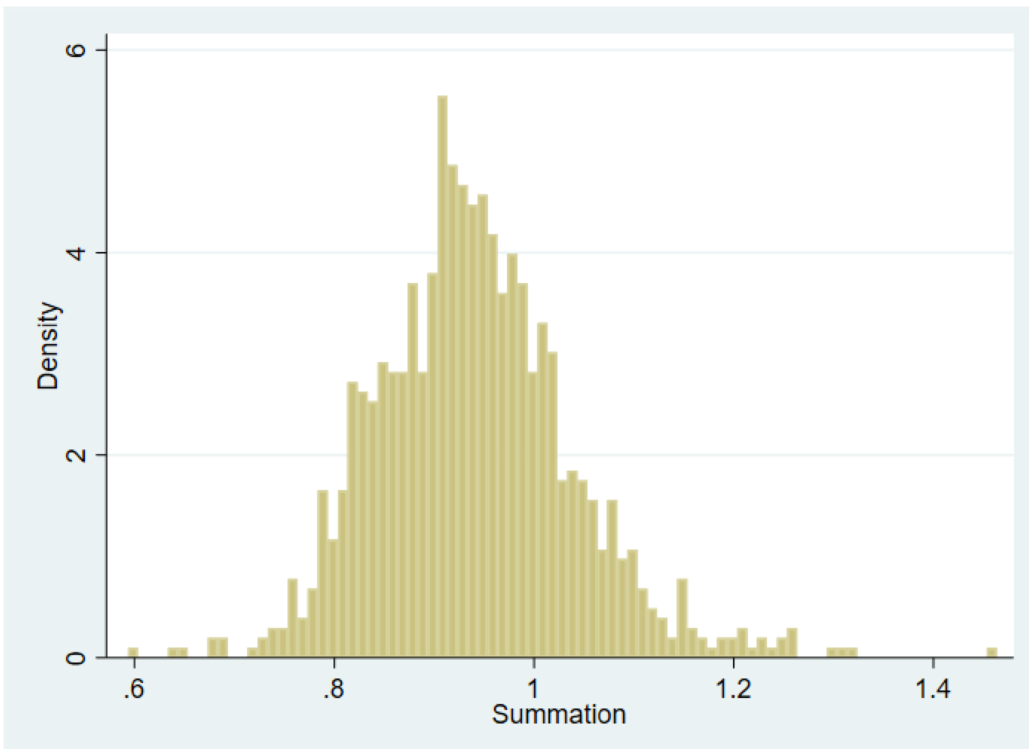


Figure 9: Distribution of all 16 Income Group Summations of Precinct Estimation Data

7.7.2 Constrained Using Sigmoid Activation

Figure 10 shows the sigmoid version of the training model using the Census block group data. The means squared error (MSE) for the sigmoid version of the model is shown in Table 24. When the sigmoid version of the model is applied to the precinct data, we get results that no longer contain negative values but at the trade-off of overestimating the mean summation of income group percentages. Rather than the center mass being approximately one it is above one as can be seen in Figure 11.

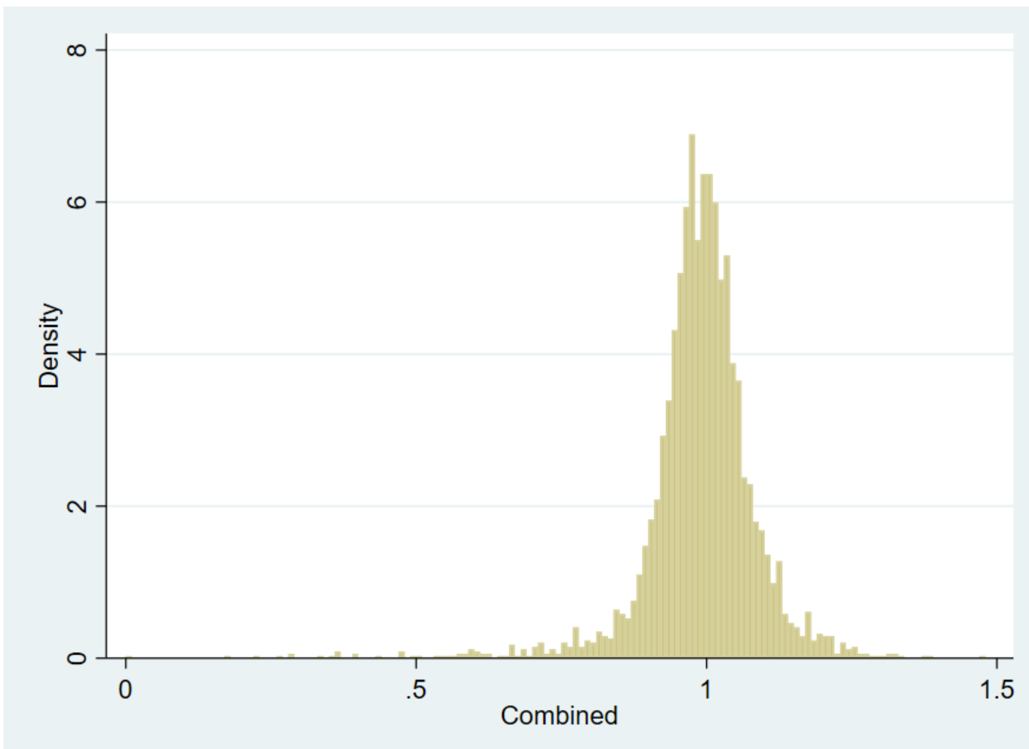


Figure 10: Distribution of all 16 Income Group Summations of BlockGroup Sigmoid Estimation Data

Table 24: Mean Squared Error (MSE) of the Sigmoid NN Block Training model

Income Group	**MSE**
lt 10K pct	0.003054427
btw 10k 15k pct	0.001787868
btw 15k 20k pct	0.001694766
btw 20k 25k pct	0.001571496
btw 25k 30k pct	0.001498637
btw 30k 35k pct	0.001627517
btw 35k 40k pct	0.002121855
btw 40k 45k pct	0.002003519
btw 45k 50k pct	0.001572871
btw 50k 60k pct	0.002717982
btw 60k 75k pct	0.003332802
btw 75k 100k pct	0.005096013
btw 100k 125k pct	0.003499941
btw 125k 150k pct	0.002157687
btw 150k 200k pct	0.003040435
gt 200k pct	0.004258048

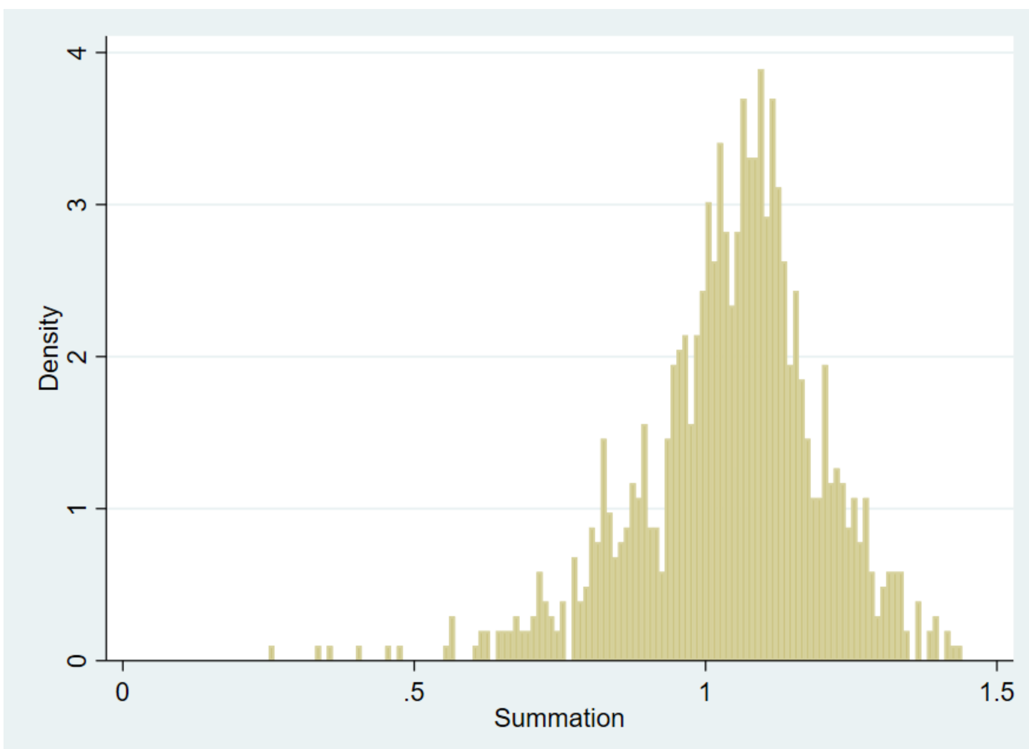


Figure 11: Distribution of all 16 Income Group Summations of Precinct Sigmoid Estimation Data

7.7.3 Comparison to Regression

When we compare the performance of the individual sigmoid neural network models to individual regressions and compare the MSEs of the two methods, we get mixed results. As seen in Table 25, the neural network outperforms in terms of MSE for 11 out of 16 income groups which tend to be the lower and middle-income groups while regression outperforms the neural network for 5 out of 16 income groups which tend to be for the more affluent income groups. The full regression results are contained inTable 29 and Table 30 in the appendix.

Table 25: Mean Squared Error (MSE) Comparison of the Sigmoid NN Block Training model to Regression

Income Group	**MSE**	**MSE Regression**	**Outperform**
lt 10K pct	0.003054427	0.0051	NN
btw 10k 15k pct	0.001787868	0.0021	NN
btw 15k 20k pct	0.001694766	0.0021	NN
btw 20k 25k pct	0.001571496	0.0022	NN
btw 25k 30k pct	0.001498637	0.0018	NN
btw 30k 35k pct	0.001627517	0.0021	NN
btw 35k 40k pct	0.002121855	0.0024	NN
btw 40k 45k pct	0.002003519	0.0018	Regression
btw 45k 50k pct	0.001572871	0.0017	NN
btw 50k 60k pct	0.002717982	0.0034	NN
btw 60k 75k pct	0.003332802	0.0036	NN
btw 75k 100k pct	0.005096013	0.0050	Regression
btw 100k 125k pct	0.003499941	0.0036	NN
btw 125k 150k pct	0.002157687	0.0020	Regression
btw 150k 200k pct	0.003040435	0.0021	Regression
gt 200k pct	0.004258048	0.0026	Regression

7.8 Conclusion

In conclusion, we find that the neural network at its current preferred Sigmoid specification is associated with an average MSE of 0.00257 compared to an average MSE of 0.00273 using standard regression methods, a 6% overall improvement. The neural network appears to outperform at lower and middle-income groups while regression still appears to outperform at higher income groups. Our proposed method of using deep learning to predict income and demographic variables at the precinct level will be useful to a wide range of future studies in the areas of political science and economics, which need to combine voting data at the precinct level with income and demographic variables only available for different geographic partitions.

In our particular application, future work would include two primary additions. The first would be a simultaneous estimation of the income groups both for the neural network and multivariate multiple regression. The second would be using the estimated income values to incorporate solar panel data from NV Energy as well as voter referendum data to infer solar energy referenda voting preferences based on solar panel installation and estimated income by precinct.

8 Bibliography

Alm, J. and M. I. Melnik (2005). Sales taxes and the decision to purchase online. *Public Finance Review 33*(2), 184–212.

Alm, J., M. I. Melnik, et al. (2010). Do ebay sellers comply with state sales taxes? *National Tax Journal 63*(2), 215–236.

Alnafissa, M. and M. Alderiny (2020). Analysis of saudi demand for imported honey using an almost ideal demand system (aids). *Journal of the Saudi Society of Agricultural Sciences 19*(4), 293–298.

Amos, B., M. P. McDonald, and R. Watkins (2017, 04). When Boundaries Collide: Constructing a National Database of Demographic and Voting Statistics. *Public Opinion Quarterly 81*(S1), 385–400.

Atkin, D. (2013). Trade, tastes, and nutrition in India. *American Economic Review 103*(5), 1629–63.

Attanasio, O. P. and L. Pistaferri (2016). Consumption inequality. *Journal of Economic Perspectives 30*(2), 3–28.

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

Bangsund, D. A., N. M. Hodur, and F. L. Leistritz (2004). Agricultural and recreational impacts of the conservation reserve program in rural north dakota, usa. *Journal of Environmental Management 71*(4), 293–303.

Banks, J., R. Blundell, and A. Lewbel (1997). Quadratic engel curves and consumer demand. *Review of Economics and statistics 79*(4), 527–539.

Baugh, B., I. Ben-David, H. Park, et al. (2014). *The" Amazon Tax": Empirical Evidence from Amazon and Main Street Retailers*. Citeseer.

Beemiller, R. M. (1989). Local area fiscal impact analysis using the regional input—output modeling system (rims ii). *Socio-Economic Planning Sciences 23*(5), 261–269.

Berner, J., P. Grohs, G. Kutyniok, and P. Petersen (2021). The modern mathematics of deep learning. *arXiv preprint arXiv:2105.04026*, 86–114.

Blundell, R. and J. M. Robin (1999). Estimation in large and disaggregated demand systems: An estimator for conditionally linear systems. *Journal of Applied Econometrics 14*(3), 209–232.

Brännlund, R. and J. Nordström (2004). Carbon tax simulations using a household demand model. *European Economic Review 48*(1), 211–233.

Broxterman, D. A. and T. C. Jin (2022). House prices, government quality, and voting behavior. *The Journal of Real Estate Finance and Economics 64*(2), 179–209.

Bruce, D. J. and W. F. Fox (2001). *State and local sales tax revenue losses from e-commerce: Updated estimates*. Center for Business and Economic Research Knoxville, Tennessee.

Brückner, M. and E. Pappa (2012). Fiscal expansions, unemployment, and labor force participation: Theory and evidence. *International Economic Review 53*(4), 1205–1228.

Cantoni, E. (2020). A precinct too far: Turnout and voting costs. *American Economic Journal: Applied Economics 12*(1), 61–85.

Cardi, O., R. Restout, and P. Claeys (2020). Imperfect mobility of labor across sectors and fiscal transmission. *Journal of Economic Dynamics and Control 111*, 103815.

Chalmers, J. A. and E. Anderson (1977). Economic/demographic assessment manual: Current practices, procedural recommendations, and a test case. *Prepared for Bureau of Reclamation, Engineering and Research Center, Denver, Colo. November 1977. 300*.

Darghouth, N. R., G. Barbose, and R. Wiser (2011). The impact of rate design and net metering on the bill savings from distributed pv for residential customers in california. *Energy Policy 39*(9), 5243–5253.

De Sousa, J., T. Mayer, and E. Sihra (2018). Market integration and convergence in consumption patterns.

Deaton, A. and J. Muellbauer (1980). An almost ideal demand system. *The American Economic Review 70*(3), 312–326.

Decoster, A. and E. Schokkaert (1990). Tax reform results with different demand systems. *Journal of Public Economics 41*(3), 277–296.

DeDad, M. (2019). Tastes in the united states: Convergence or divergence? Technical report, Working Paper: Indiana University.

Deller, S. C. (1995). Preliminary estimates of basic cpan model.

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

Ellison, G. and S. F. Ellison (2006). Internet retail demand: Taxes, geography, and online-offline competition. Technical report, National Bureau of Economic Research.

Energy and I. Environmental Economics Incorporated (2016, August). Nevada net energy metering impacts evaluation 2016 update. Accessed April 29, 2024.

Evans, G. K. and J. I. Stallmann (2006). Safesim: the small area fiscal estimation simulator. *Community policy analysis modeling*, 167–180.

Finan, F., E. Seira, and A. Simpser (2021). Voting with one's neighbors: Evidence from migration within mexico. *Journal of Public Economics 202*, 104495.

Fritts, J. (2020). State and local sales tax rates, 2020. *Tax Foundation*. `https://taxfoundation.org/2020-sales-taxes/` (visited: 2020-11-17).

Gallet, C. A. (2010). Meat meets meta: a quantitative review of the price elasticity of meat. *American Journal of Agricultural Economics 92*(1), 258–272.

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

Gordon, M. R. (2009). Up the Amazon without a paddle: Examining sales taxes, entity isolation, and the affiliate tax. *NCJL & Tech. 11*, 299.

Hara, K., D. Saito, and H. Shouno (2015). Analysis of function of rectified linear unit used in deep learning. In *2015 international joint conference on neural networks (IJCNN)*, pp. 1–8. IEEE.

Harris, T. R. (1995). An approach to incorporate risk in the estimation of fiscal impacts from industry location in a rural community. *working paper*.

Harris, T. R., J. S. Shonkwiler, G. E. Ebai, and P. Janson (2000). Application of bea economic areas in the development of the great basin fiscal impact model. *Journal of Regional Analysis & Policy 30*(1), 75–92.

Hausman, J., G. Leonard, and J. D. Zona (1994). Competitive analysis with differenciated products. *Annales d'Economie et de Statistique*, 159–180.

HessvIllinois (1967). National bellas hess, inc. v. department of revenue of the state of illinois. Washington D.C. Supreme Court of United States.

Holtz-Eakin, D., W. K. Newey, and H. S. Rosen (1987). The revenues-expenditures nexus: evidence from local government data. Technical report, National Bureau of Economic Research.

Jing, X., R. Mittelhammer, and T. Heckelei (2004). A quaids model of japanese meat demand. *Selected Paper American Agricultural Economics Association Denver*.

Johnson, T., J. M. and J. Scott (1996). The conceptual basis of the show me models.

Jones, A. and M. G. Mazzi (1996). Tobacco consumption and taxation in Italy: an application of the quaids model. *Applied Economics 28*(5), 595–603.

Kaplan, E., J. L. Spenkuch, and R. Sullivan (2022). Partisan spatial sorting in the United States: A theoretical and empirical overview. *Journal of Public Economics 211*, 104668.

Kennedy, P. (2003). *A guide to econometrics*. MIT press.

Kingma, D. P. and J. Ba (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Krinsky, I. and A. L. Robb (1986). On approximating the statistical properties of elasticities. *The Review of Economics and Statistics*, 715–719.

Lecocq, S. and J.-M. Robin (2015). Estimating almost-ideal demand systems with endogenous regressors. *The Stata Journal 15*(2), 554–573.

LeCun, Y., Y. Bengio, and G. Hinton (2015). Deep learning. *Nature 521*(7553), 436–444.

Madden, D. (1996). Marginal tax reform and the specification of consumer demand systems. *Oxford Economic Papers 48*(4), 556–567.

Maliar, L. and S. Maliar (2022). Deep learning classification: Modeling discrete labor choice. *Journal of Economic Dynamics and Control 135*, 104295.

Mathew, A., P. Amudha, and S. Sivakumari (2021). Deep learning techniques: an overview. *Advanced Machine Learning Technologies and Applications: Proceedings of AMLTA 2020*, 599–608.

McCulloch, W. S. and W. Pitts (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics 5*, 115–133.

Newman, B. J., S. L. DeMora, and T. T. Reny (2021). Female empowerment and the politics of language: Evidence using gender-neutral amendments to subnational constitutions. *British Journal of Political Science 51*(4), 1761–1772.

Nielsen (2019). Consumer panel dataset manual: The kilts center archive of the nielsen company. Technical report, Kilts Center for Marketing.

Ning, C. and F. You (2019). Optimization under uncertainty in the era of big data and deep learning: When machine learning meets mathematical programming. *Computers & Chemical Engineering 125*, 434–448.

Nosratabadi, S., A. Mosavi, P. Duan, P. Ghamisi, F. Filip, S. S. Band, U. Reuter, J. Gama, and A. H. Gandomi (2020). Data science in economics: comprehensive review of advanced machine learning and deep learning methods. *Mathematics 8*(10), 1799.

Petrović, P., M. Arsić, and A. Nojković (2021). Increasing public investment can be an effective policy in bad times: Evidence from emerging eu economies. *Economic Modelling 94*, 580–597.

Poi, B. P. (2002). From the help desk: Demand system estimation. *The Stata Journal 2*(4), 403–410.

Poi, B. P. (2008). Demand-system estimation: Update. *The Stata Journal 8*(4), 554–556.

Poi, B. P. (2012). Easy demand-system estimation with quaids. *The Stata Journal 12*(3), 433–446.

Pollak, R. A. and T. J. Wales (1981). Demographic variables in demand analysis. *Econometrica: Journal of the Econometric Society*, 1533–1551.

QuillvND (1992). Quill corp. v. north dakota. Washington D.C. Supreme Court of United States.

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review 65*(6), 386.

Shields, M. and S. Deller (2003). Using economic impact models as an educational tool in community economic development programming: Lessons from pennsylvania and wisconsin. *Journal of Extension 41*(3), 1–7.

Song, B., M. D. Woods, G. A. Doeksen, and D. Schreiner (1992). Multiplier analysis for agriculture and other industries. *OSU extension facts (USA)*.

Stone, R. (1954). Linear expenditure systems and demand analysis: an application to the pattern of british demand. *The Economic Journal 64*(255), 511–527.

Subramanian, S. and A. Deaton (1996). The demand for food and calories. *Journal of Political Economy 104*(1), 133–162.

Swenson, D. (1996). The iowa economic/fiscal impact modeling system: An overview. *Department of Economics, Iowa State University*.

Swenson, D. and L. Eathington (1998). A manual for community and fiscal impact modeling systems.

Swenson, D. and D. M. Otto (1998). The iowa economic/fiscal impact modeling system. *Journal of Regional Analysis and Policy 28*(1100-2016-89749), 64–75.

Trounstine, J. (2020). The geography of inequality: How land use regulation produces segregation. *American Political Science Review 114*(2), 443–455.

Walls, M. and M. Ashenfarb (2022). Efficiency and equity of an outdoor recreation equipment tax to fund public lands. *Land Economics 98*(3), 520–536.

Wayfair (2018). South dakota v. wayfair, inc. Washington D.C. Supreme Court of United States.

Yeo, J. and D. Holland (2004). Economic growth in washington: An examination of migration response and a test of model accuracy. *International Regional Science Review 27*(2), 205–237.

9 Appendix

Table 26: Combined State and Average Local Sales Tax Rates (Percentage) from the Tax Foundation

State & Year	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Alabama	8.03	8.33	8.45	8.85	8.91	8.97	9.03	9.1	9.14	9.22
Alaska	1.12	1.77	1.69	1.69	1.76	1.78	1.76	1.76	1.43	1.76
Arizona	9.01	9.12	9.16	8.16	8.17	8.25	8.32	8.33	8.37	8.4
Arkansas	8.1	8.58	8.61	9.24	9.26	9.3	9.34	9.41	9.43	9.47
California	9.08	8.11	8.38	8.44	8.44	8.48	8.48	8.54	8.56	8.66
Colorado	6.98	7.44	7.39	7.41	7.44	7.5	7.5	7.52	7.63	7.65
Connecticut	6	6.35	6.35	6.35	6.35	6.35	6.35	6.35	6.35	6.35
Delaware	0	0	0	0	0	0	0	0	0	0
Florida	6.99	6.62	6.62	6.63	6.65	6.66	6.8	6.8	7.05	7.05
Georgia	6.95	6.84	6.99	6.99	6.96	7	7.15	7.15	7.29	7.31
Hawaii	4.35	4.35	4.35	4.35	4.35	4.35	4.35	4.35	4.41	4.44
Idaho	6.03	6.02	6.02	6.03	6.01	6.03	6.03	6.03	6.03	6.03
Illinois	8.22	8.2	8.13	8.18	8.19	8.65	8.69	8.7	8.74	9.08
Indiana	7	7	7	7	7	7	7	7	7	7
Iowa	6.84	6.81	6.82	6.78	6.78	6.8	6.8	6.8	6.82	6.94
Kansas	7.95	8.26	8.25	8.19	8.2	8.61	8.68	8.68	8.67	8.68
Kentucky	6	6	6	6	6	6	6	6	6	6
Louisiana	8.69	8.85	8.87	8.91	8.91	9.98	10.02	10.02	9.45	9.52
Maine	5	5	5	5.5	5.5	5.5	5.5	5.5	5.5	5.5
Maryland	6	6	6	6	6	6	6	6	6	6
Massachusetts	6.25	6.25	6.25	6.25	6.25	6.25	6.25	6.25	6.25	6.25
Michigan	6	6	6	6	6	6	6	6	6	6

Minnesota	7.14	7.18	7.16	7.19	7.2	7.31	7.29	7.42	7.43	7.46
Mississippi	7	7	7	7.07	7.07	7.07	7.07	7.07	7.07	7.07
Missouri	7.46	7.49	7.46	7.78	7.81	7.87	7.97	8.03	8.13	8.18
Montana	0	0	0	0	0	0	0	0	0	0
Nebraska	6.39	6.77	6.78	6.79	6.8	6.87	6.9	6.89	6.85	6.93
Nevada	7.96	7.93	7.93	7.94	7.94	7.98	8.14	8.14	8.14	8.32
New Hampshire	0	0	0	0	0	0	0	0	0	0
New Jersey	7	6.97	6.97	6.97	6.97	6.97	6.85	6.6	6.6	6.6
New Mexico	7.14	7.24	7.26	7.3	7.35	7.55	7.63	7.66	7.82	7.82
New York	8.52	8.48	8.48	8.48	8.48	8.49	8.49	8.49	8.49	8.52
North Carolina	7.82	6.85	6.87	6.9	6.9	6.9	6.95	6.95	6.97	6.97
North Dakota	5.57	6.39	6.52	6.62	6.56	6.78	6.79	6.8	6.85	6.86
Ohio	6.78	6.75	6.8	7.11	7.1	7.14	7.14	7.15	7.17	7.17
Oklahoma	8.33	8.66	8.67	8.76	8.77	8.85	8.86	8.91	8.92	8.94
Oregon	0	0	0	0	0	0	0	0	0	0
Pennsylvania	6.34	6.34	6.34	6.34	6.34	6.34	6.34	6.34	6.34	6.34
Rhode Island	7	7	7	7	7	7	7	7	7	7
South Carolina	7.25	7.13	7.08	7.13	7.13	7.23	7.37	7.37	7.43	7.46
South Dakota	5.22	5.39	5.82	5.83	5.83	6.34	6.4	6.4	6.4	6.4
Tennessee	9.44	9.45	9.44	9.45	9.45	9.45	9.45	9.46	9.47	9.53
Texas	7.61	8.14	8.14	8.16	8.05	8.17	8.17	8.17	8.19	8.19
Utah	6.62	6.68	6.67	6.68	6.68	6.76	6.77	6.77	6.94	7.18
Vermont	6	6.14	6.14	6.14	6.14	6.17	6.18	6.18	6.18	6.22
Virginia	5	5	5	5.63	5.63	5.63	5.63	5.63	5.65	5.65
Washington	8.64	8.8	8.86	8.88	8.89	8.92	9.2	9.18	9.17	9.21
West Virginia	6	6	6.04	6.07	6.07	6.29	6.29	6.37	6.39	6.41

Wisconsin	5.42	5.43	5.43	5.43	5.43	5.41	5.42	5.42	5.44	5.46
Wyoming	5.3	5.34	5.34	5.49	5.47	5.42	5.26	5.46	5.36	5.34

Table 27: Demographic Category Codebook

Variable Name	**Generalized Category**
Household Size	
Single Member	0
Two Members	1
Three Members	2
Four Members	3
Five Members	4
Six Members	5
Seven Members	6
Eight Members	7
Nine+ Members	8
Type of Residence	
One Family House	0
One Family House (Condo/Coop)	1
Two Family	2
Two Family House (Condo/Coop)	3
Three+ Family House	4
Three+ Family House (Condo/Coop)	5
Mobile Home or Trailer	6

Unreported	7

Household Income

Under $5000	0
$5000-$7999	1
$8000-$9999	2
$10,000-$11,999	3
$12,000-$14,999	4
$15,000-$19,999	5
$20,000-$24,999	6
$25,000-$29,999	7
$30,000-$34,999	8
$35,000-$39,999	9
$40,000-$44,999	10
$45,000-$49,999	11
$50,000-$59,999	12
$60,000-$69,999	13
$70,000-$99,999	14
$100,000 +	15

*Total household income range for the full year that is 2 years prior to the Panel Year.

Household Composition

Married	0
Female Head Living with Others Related	1
Male Head Living with Others Related	2

Female Living Alone	3
Female Living with Non-Related	4
Male Living Alone	5
Male Living with Non-Related	6

Age – Male & Female Head of Household

Under 25 Years	0
25-29 Years	1
30-34 Years	2
35-39 Years	3
40-44 Years	4
45-49 Years	5
50-54 Years	6
55-64 Years	7
65+ Years	8
No Male/Female Head	9

Age and Presence of Children

Under 6 only	0
6-12 only	1
13-17 only	2
Under 6 & 6-12	3
Under 6 & 13-17	4
6-12 & 13-17	5
Under 6 & 6-12 & 13-17	6
No Children Under 18	7

Male/Female Head Employment

Under 30 hours	0
30-34 hours	1
35+ hours	2
Not Employed for Pay	3
No Male/Female Head	4

Male/Female Head Education

Grade School	0
Some High School	1
Graduated High School	2
Some College	3
Graduated College	4
Post College Grad	5
No Male/Female Head or Unknown	6

Marital Status

Married	0
Widowed	1
Divorced/Separated	2
Single	3
Unknown	4

Male/Female Head Occupation

No male/female head of household	0

Whitecollar 1	1
Whitecollar 2	2
Bluecollar 1	3
Whitecollar 3	4
Whitecollar 4	5
Bluecollar 2	6
Bluecollar 3	7
Military	8
Bluecollar 4	9
Farmer	10
Students Employed ¡30 hours	11
Bluecollar 5	12
Housewife Retired Unable to work Unemployed/Laid off	13

Race

White/Caucasian	0
Black/African American	1
Asian	2
Other	3

Table 28: Income Bracket Group Numbers Crosswalk

Income Bracket	**Income Group Number**
lt 10K pct	incomegroup1
btw 10k 15k pct	incomegroup2
btw 15k 20k pct	incomegroup3
btw 20k 25k pct	incomegroup4
btw 25k 30k pct	incomegroup5
btw 30k 35k pct	incomegroup6
btw 35k 40k pct	incomegroup7
btw 40k 45k pct	incomegroup8
btw 45k 50k pct	incomegroup9
btw 50k 60k pct	incomegroup10
btw 60k 75k pct	incomegroup11
btw 75k 100k pct	incomegroup12
btw 100k 125k pct	incomegroup13
btw 125k 150k pct	incomegroup14
btw 150k 200k pct	incomegroup15
gt 200k pct	incomegroup16

Table 29: Individual Regression Comparison for 1st 8 Income Bracket Percentages

Variable	**lt 10K**	**10-15k**	**15-20k**	**20-25k**	**25-30k**	**30-35k**	**35-40k**	**40-45k**
Tot_SF	-0.0001	0.0000	-0.0001	-0.0001	-0.0001	0.0000	0.0000	0.0000
	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
EstateAcres	-0.0566	-0.0821	0.0455	-0.0637	-0.0135	0.1018	0.0370	0.0784
	0.0509	0.0329	0.0327	0.0334	0.0306	0.0327	0.0348	0.0307
beds	-0.0145	-0.0209	0.0075	-0.0061	-0.0014	0.0085	0.0075	-0.0055
	0.0172	0.0111	0.0111	0.0113	0.0104	0.0111	0.0118	0.0104
Fullbath	0.0382	0.0003	-0.0065	0.0192	0.0057	-0.0020	0.0056	0.0057
	0.0244	0.0158	0.0156	0.0160	0.0146	0.0157	0.0167	0.0147
Halfbath	0.0057	-0.0245	0.0346	-0.0018	-0.0116	0.0139	-0.0364	-0.0158
	0.0370	0.0239	0.0237	0.0243	0.0223	0.0238	0.0253	0.0223

Continued on next page

Table 29 – continued from previous page

Variable	lt 10K	10-15k	15-20k	20-25k	25-30k	30-35k	35-40k	40-45k
Fireplace	0.0161	0.0254	-0.0247	0.0187	0.0298	0.0128	0.0102	-0.0164
	0.0237	0.0153	0.0152	0.0155	0.0143	0.0152	0.0162	0.0143
GarageSF	-0.0001	-0.0001	-0.0002	-0.0001	0.0000	-0.0001	0.0000	-0.0001
	0.0001	0.0001	0.0001	0.0001	0.0000	0.0001	0.0001	0.0000
Basement	0.0000	0.0002	0.0000	0.0001	0.0001	-0.0001	0.0000	0.0000
	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
TaxVal21to22	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p25bldgsf	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p25land	0.0742	0.0889	0.0018	0.0520	0.0090	-0.0577	-0.0311	-0.0424
	0.0374	0.0242	0.0240	0.0245	0.0225	0.0240	0.0256	0.0226
p25beds	0.0099	-0.0039	-0.0019	0.0059	0.0006	0.0007	0.0053	-0.0023
	0.0079	0.0051	0.0051	0.0052	0.0048	0.0051	0.0054	0.0048
p25fbath	-0.0290	-0.0008	0.0015	-0.0074	0.0054	0.0088	-0.0044	0.0008
	0.0113	0.0073	0.0072	0.0074	0.0068	0.0072	0.0077	0.0068
p25hbath	-0.0006	0.0107	-0.0010	-0.0064	0.0028	-0.0162	0.0181	-0.0057
	0.0183	0.0119	0.0118	0.0120	0.0110	0.0118	0.0125	0.0111
p25fire	-0.0022	-0.0161	0.0051	-0.0070	-0.0095	-0.0066	-0.0092	0.0090
	0.0109	0.0070	0.0070	0.0071	0.0066	0.0070	0.0075	0.0066
p25garage	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p25bsmtarea	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Continued on next page

Table 29 – continued from previous page

Variable	lt 10K	10-15k	15-20k	20-25k	25-30k	30-35k	35-40k	40-45k
	.	.	.	.	.	.	.	.
p25taxable	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p75bldgsf	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p75land	0.0408	0.0424	0.0007	0.0388	-0.0217	-0.0456	-0.0257	-0.0371
	0.0247	0.0160	0.0159	0.0162	0.0149	0.0159	0.0169	0.0149
p75beds	0.0002	0.0088	-0.0011	-0.0038	0.0005	-0.0087	-0.0127	0.0025
	0.0060	0.0039	0.0039	0.0040	0.0036	0.0039	0.0041	0.0036
p75fbath	-0.0006	0.0044	0.0031	0.0013	-0.0030	0.0035	0.0071	0.0022
	0.0077	0.0050	0.0050	0.0051	0.0046	0.0050	0.0053	0.0047
p75hbath	-0.0043	-0.0011	-0.0096	0.0000	-0.0007	0.0006	0.0090	0.0051
	0.0090	0.0058	0.0058	0.0059	0.0054	0.0058	0.0062	0.0054
p75fire	0.0003	-0.0072	0.0053	-0.0057	-0.0080	0.0004	0.0009	0.0051
	0.0087	0.0056	0.0056	0.0057	0.0052	0.0056	0.0060	0.0053
p75gararea	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p75bsmtarea	-0.0001	-0.0001	0.0000	-0.0001	0.0000	0.0000	-0.0001	0.0000
	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
p75taxable	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
medbldgsf	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Continued on next page

Table 29 – continued from previous page

Variable	lt 10K	10-15k	15-20k	20-25k	25-30k	30-35k	35-40k	40-45k
medlandsize	-0.0699	-0.0383	-0.0432	-0.0239	0.0277	0.0073	0.0228	0.0044
	0.0365	0.0236	0.0234	0.0239	0.0220	0.0235	0.0250	0.0220
medbeds	-0.0089	0.0100	-0.0052	0.0033	-0.0056	0.0006	-0.0067	0.0023
	0.0072	0.0047	0.0046	0.0047	0.0044	0.0047	0.0050	0.0044
medfullbaths	0.0012	-0.0083	0.0000	0.0023	0.0081	0.0062	0.0096	0.0011
	0.0096	0.0062	0.0062	0.0063	0.0058	0.0062	0.0066	0.0058
medhalfbaths	-0.0020	0.0003	-0.0122	-0.0030	0.0016	0.0022	0.0109	0.0037
	0.0109	0.0070	0.0070	0.0071	0.0065	0.0070	0.0074	0.0066
medfire	-0.0074	-0.0066	0.0111	-0.0046	-0.0079	-0.0059	0.0046	0.0008
	0.0091	0.0059	0.0058	0.0060	0.0055	0.0059	0.0062	0.0055
medgararea	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
medbsmtarea	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	.	.	.	.	.	.	.	.
medtaxable	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
sdbldgsf	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
sdlandsize	0.0213	0.0301	0.0153	0.0187	0.0084	-0.0120	-0.0061	-0.0092
	0.0110	0.0071	0.0070	0.0072	0.0066	0.0070	0.0075	0.0066
sdbeds	0.0449	0.0115	-0.0067	0.0128	0.0071	0.0096	0.0277	0.0091
	0.0116	0.0075	0.0074	0.0076	0.0070	0.0074	0.0079	0.0070
sdfullbaths	-0.0184	0.0038	0.0367	0.0017	-0.0033	0.0067	-0.0200	-0.0116

Continued on next page

Table 29 – continued from previous page

Variable	lt 10K	10-15k	15-20k	20-25k	25-30k	30-35k	35-40k	40-45k
	0.0171	0.0110	0.0109	0.0112	0.0103	0.0110	0.0117	0.0103
sdhalfbaths	-0.0129	-0.0074	-0.0035	-0.0102	0.0046	-0.0141	0.0121	0.0029
	0.0243	0.0157	0.0156	0.0159	0.0146	0.0156	0.0166	0.0147
sdfireplaces	-0.0396	-0.0227	-0.0046	-0.0030	-0.0101	-0.0070	-0.0065	0.0060
	0.0186	0.0120	0.0119	0.0122	0.0112	0.0120	0.0127	0.0112
sdgararea	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
sdbsmtarea	0.0001	0.0000	0.0000	0.0000	-0.0001	0.0001	0.0000	0.0000
	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
sdtaxable	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
cons	0.1266	0.0767	0.0756	0.0575	0.0595	0.0686	0.0675	0.0521
	0.0117	0.0075	0.0075	0.0076	0.0070	0.0075	0.0080	0.0070
rmse	0.0712	0.0460	0.0456	0.0466	0.0428	0.0457	0.0487	0.0429
r2	0.2914	0.2707	0.2720	0.2037	0.1753	0.1717	0.1075	0.0831
MSE	0.0051	0.0021	0.0021	0.0022	0.0018	0.0021	0.0024	0.0018

Table 30: Individual Regression Comparison for 2nd 8 Income Bracket Percentages

Variable	45-50k	50-60k	60-75k	75-100k	100-125k	125-150k	150-200k	gt 200k
Tot_SF	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001
	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000

Continued on next page

Table 30 – continued from previous page

Variable	45-50k	50-60k	60-75k	75-100k	100-125k	125-150k	150-200k	gt 200k
EstateAcres	-0.0421	0.0102	-0.0240	0.0168	0.0024	0.0350	0.0277	-0.0728
	0.0294	0.0414	0.0432	0.0504	0.0431	0.0323	0.0331	0.0363
beds	-0.0022	0.0265	0.0018	0.0193	0.0159	-0.0036	-0.0064	-0.0261
	0.0099	0.0140	0.0146	0.0171	0.0146	0.0109	0.0112	0.0123
Fullbath	0.0204	-0.0083	-0.0010	-0.0086	0.0114	-0.0066	-0.0326	-0.0411
	0.0141	0.0198	0.0207	0.0241	0.0206	0.0155	0.0159	0.0174
Halfbath	-0.0160	-0.0033	0.0277	0.0556	-0.0312	-0.0182	0.0050	0.0157
	0.0214	0.0301	0.0314	0.0367	0.0314	0.0235	0.0241	0.0264
Fireplace	0.0009	-0.0019	-0.0399	-0.0232	0.0123	0.0096	-0.0242	-0.0055
	0.0137	0.0193	0.0201	0.0235	0.0201	0.0151	0.0154	0.0169
GarageSF	0.0001	0.0000	0.0000	0.0003	0.0002	0.0001	0.0001	0.0000
	0.0000	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
Basement	0.0001	0.0001	-0.0002	-0.0001	-0.0002	0.0000	-0.0002	0.0001
	0.0001	0.0001	0.0001	0.0002	0.0001	0.0001	0.0001	0.0001
TaxVal21to22	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p25bldgsf	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p25land	0.0117	-0.0001	-0.0272	-0.0587	-0.0258	-0.0222	-0.0118	0.0390
	0.0216	0.0304	0.0317	0.0371	0.0317	0.0237	0.0244	0.0267
p25beds	0.0059	-0.0051	0.0081	-0.0040	0.0015	0.0021	-0.0100	-0.0128
	0.0046	0.0065	0.0067	0.0079	0.0067	0.0050	0.0052	0.0057
p25fbath	-0.0025	0.0065	0.0046	0.0216	0.0055	-0.0065	-0.0059	0.0017

Continued on next page

Table 30 – continued from previous page

Variable	45-50k	50-60k	60-75k	75-100k	100-125k	125-150k	150-200k	gt 200k
	0.0065	0.0092	0.0096	0.0112	0.0095	0.0071	0.0073	0.0080
p25hbath	0.0141	0.0134	-0.0073	-0.0190	-0.0005	0.0034	-0.0159	0.0101
	0.0106	0.0149	0.0156	0.0182	0.0155	0.0116	0.0119	0.0131
p25fire	-0.0002	0.0065	0.0160	0.0109	-0.0097	-0.0108	0.0077	0.0161
	0.0063	0.0089	0.0092	0.0108	0.0092	0.0069	0.0071	0.0078
p25garage	0.0000	0.0000	0.0000	-0.0001	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p25bsmtarea	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	.	.	.	.	.	.	.	.
p25taxable	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p75bldgsf	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p75land	0.0222	0.0173	-0.0154	-0.0204	0.0032	-0.0039	-0.0066	0.0112
	0.0143	0.0201	0.0210	0.0245	0.0210	0.0157	0.0161	0.0176
p75beds	0.0024	0.0033	0.0041	0.0047	-0.0055	0.0021	0.0015	0.0015
	0.0035	0.0049	0.0051	0.0060	0.0051	0.0038	0.0039	0.0043
p75fbath	-0.0001	-0.0013	-0.0071	-0.0014	-0.0189	-0.0036	0.0095	0.0049
	0.0045	0.0063	0.0065	0.0076	0.0065	0.0049	0.0050	0.0055
p75hbath	0.0053	-0.0048	0.0017	-0.0084	0.0044	0.0078	0.0008	-0.0058
	0.0052	0.0073	0.0076	0.0089	0.0076	0.0057	0.0059	0.0064
p75fire	-0.0080	-0.0008	0.0088	0.0085	-0.0006	0.0004	0.0045	-0.0040
	0.0050	0.0071	0.0074	0.0086	0.0074	0.0055	0.0057	0.0062

Continued on next page

Table 30 – continued from previous page

Variable	45-50k	50-60k	60-75k	75-100k	100-125k	125-150k	150-200k	gt 200k
p75gararea	0.0000	0.0000	0.0000	-0.0001	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p75bsmtarea	0.0000	0.0000	0.0001	0.0002	0.0000	0.0000	0.0000	-0.0001
	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
p75taxable	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
medbldgsf	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
medlandsize	0.0053	-0.0119	0.0628	0.0526	0.0086	-0.0167	-0.0121	0.0247
	0.0211	0.0297	0.0310	0.0362	0.0309	0.0232	0.0238	0.0260
medbeds	-0.0037	-0.0097	0.0026	0.0040	0.0048	0.0018	0.0081	0.0021
	0.0042	0.0059	0.0061	0.0072	0.0061	0.0046	0.0047	0.0052
medfullbaths	-0.0016	0.0113	0.0062	-0.0113	-0.0278	-0.0075	-0.0060	0.0163
	0.0056	0.0079	0.0082	0.0096	0.0082	0.0061	0.0063	0.0069
medhalfbaths	0.0052	0.0080	-0.0037	-0.0110	0.0115	0.0013	-0.0088	-0.0036
	0.0063	0.0089	0.0092	0.0108	0.0092	0.0069	0.0071	0.0078
medfire	-0.0014	-0.0053	0.0085	0.0107	-0.0046	-0.0028	0.0057	0.0053
	0.0053	0.0074	0.0077	0.0090	0.0077	0.0058	0.0059	0.0065
medgararea	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
medbsmtarea	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	.	.	.	.	.	.	.	.
medtaxable	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Continued on next page

Table 30 – continued from previous page

Variable	45-50k	50-60k	60-75k	75-100k	100-125k	125-150k	150-200k	gt 200k
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
sdbldgsf	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
sdlandsize	0.0021	-0.0169	-0.0051	-0.0239	-0.0177	-0.0153	-0.0011	0.0111
	0.0063	0.0089	0.0093	0.0109	0.0093	0.0070	0.0071	0.0078
sdbeds	-0.0017	-0.0087	-0.0135	-0.0270	-0.0163	-0.0091	-0.0125	-0.0271
	0.0067	0.0094	0.0098	0.0115	0.0098	0.0074	0.0075	0.0083
sdfullbaths	-0.0042	0.0049	-0.0062	-0.0122	0.0110	-0.0015	0.0048	0.0075
	0.0098	0.0139	0.0145	0.0169	0.0144	0.0108	0.0111	0.0122
sdhalfbaths	0.0147	0.0185	0.0052	-0.0038	0.0079	-0.0042	-0.0013	-0.0078
	0.0140	0.0198	0.0206	0.0241	0.0206	0.0154	0.0158	0.0173
sdfireplaces	0.0107	0.0290	0.0320	0.0126	-0.0041	-0.0202	0.0064	0.0213
	0.0107	0.0151	0.0158	0.0184	0.0158	0.0118	0.0121	0.0133
sdgararea	0.0000	0.0000	0.0000	0.0000	-0.0001	-0.0001	-0.0001	-0.0001
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
sdbsmtarea	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000
sdtaxable	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
cons	0.0462	0.0804	0.1002	0.0915	0.0533	0.0369	0.0107	-0.0034
	0.0067	0.0095	0.0099	0.0115	0.0099	0.0074	0.0076	0.0083
rmse	0.0410	0.0579	0.0604	0.0705	0.0603	0.0451	0.0463	0.0507
r2	0.0842	0.0930	0.1150	0.2346	0.2581	0.2522	0.4202	0.6196

Continued on next page

Table 30 – continued from previous page

Variable	45-50k	50-60k	60-75k	75-100k	100-125k	125-150k	150-200k	gt 200k
MSE	0.0017	0.0034	0.0036	0.0050	0.0036	0.0020	0.0021	0.0026

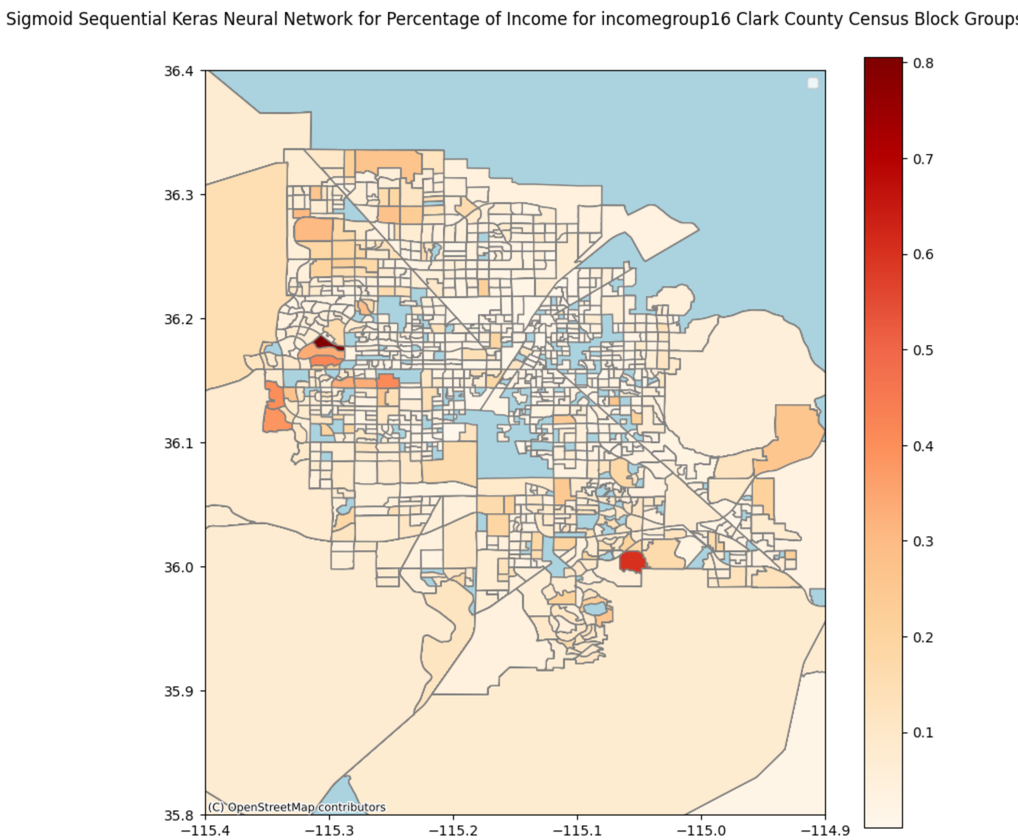


Figure 12: Block Level NN estimate of percentage in 200k+ Income Bracket

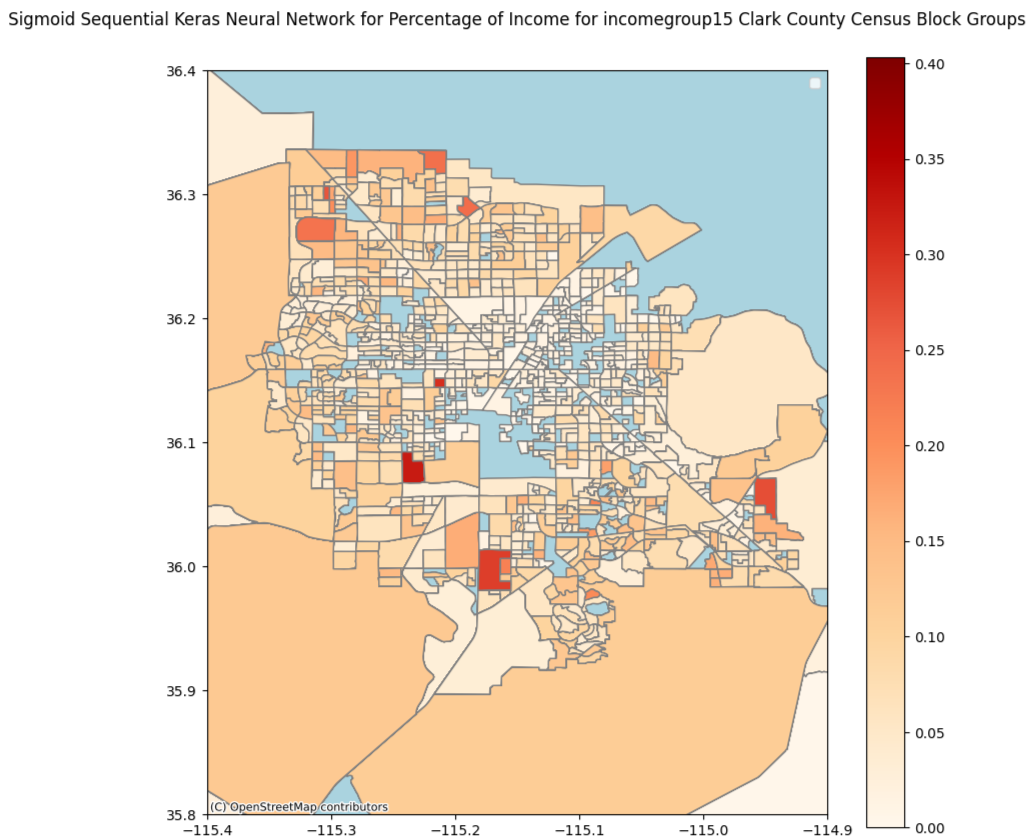


Figure 13: Block Level NN estimate of percentage in 150k to 200k Income Bracket

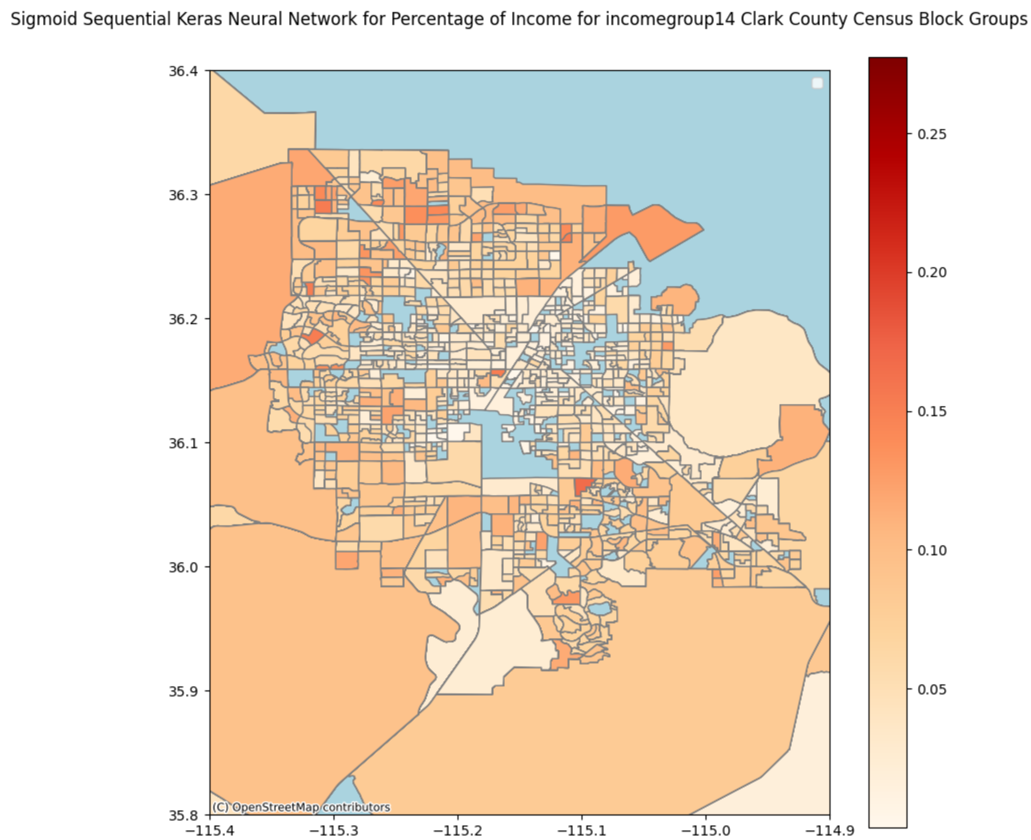


Figure 14: Block Level NN estimate of percentage in 125k to 150k Income Bracket

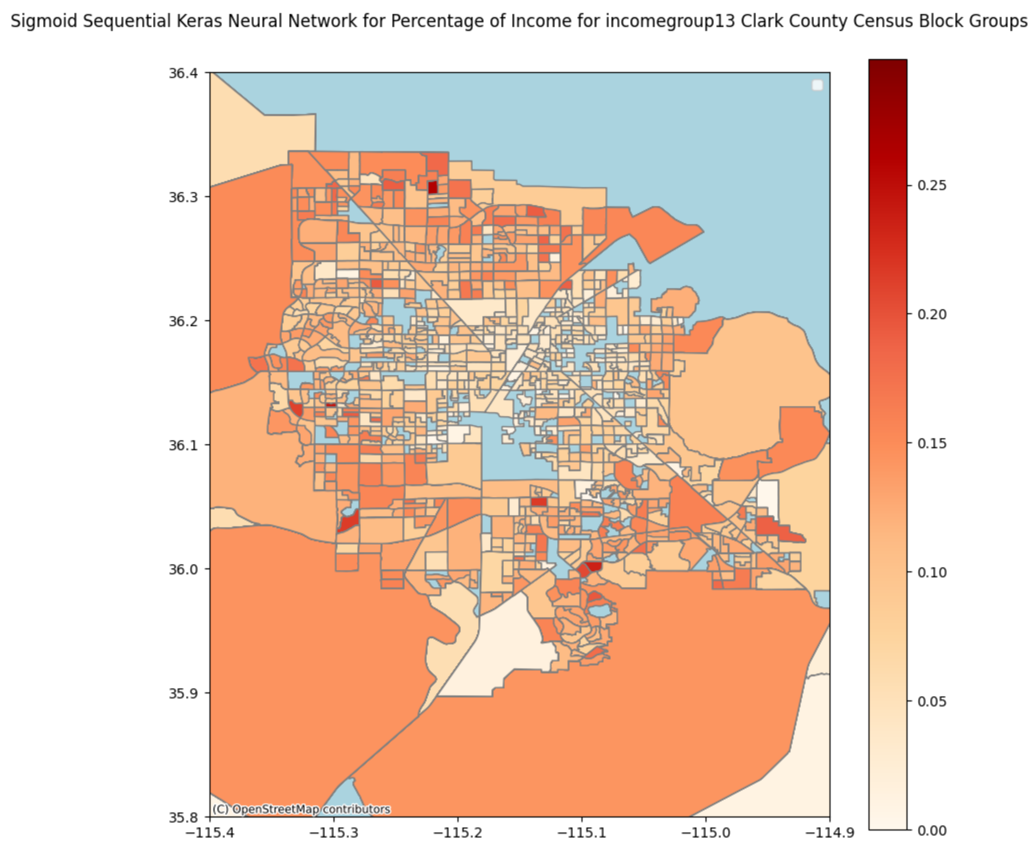


Figure 15: Block Level NN estimate of percentage in 100k to 125k Income Bracket

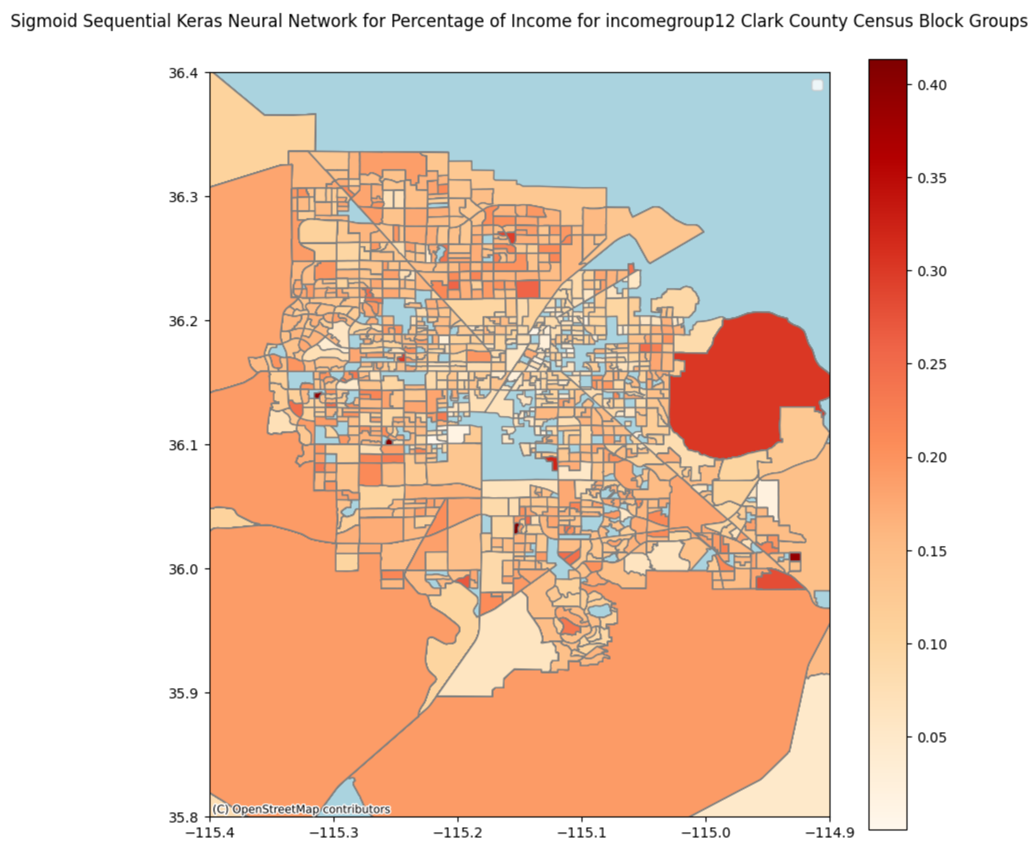


Figure 16: Block Level NN estimate of percentage in 75k to 100k Income Bracket

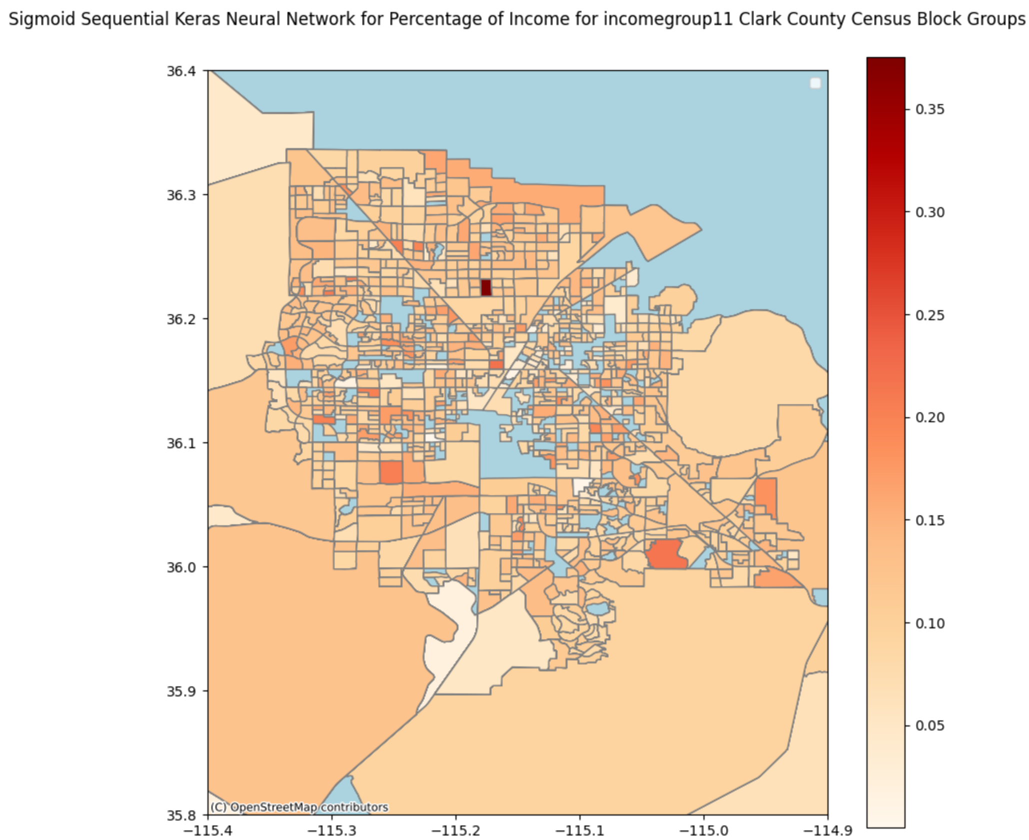


Figure 17: Block Level NN estimate of percentage in 60k to 75k Income Bracket

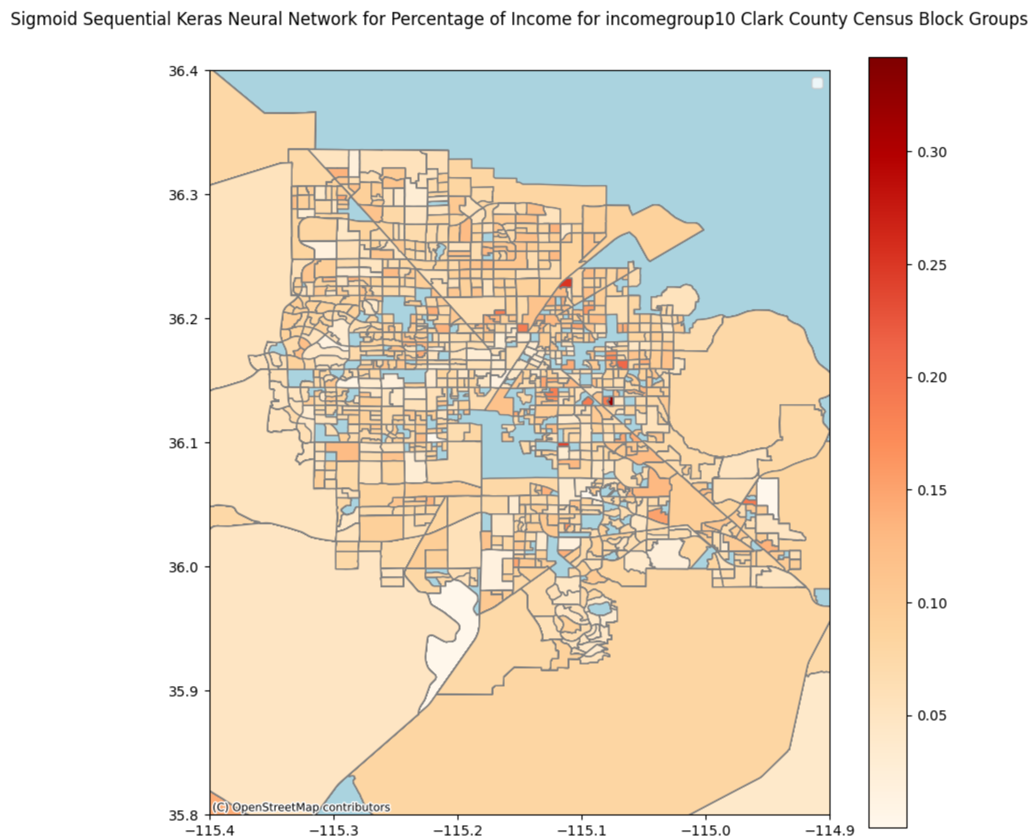


Figure 18: Block Level NN estimate of percentage in 50k to 60k Income Bracket

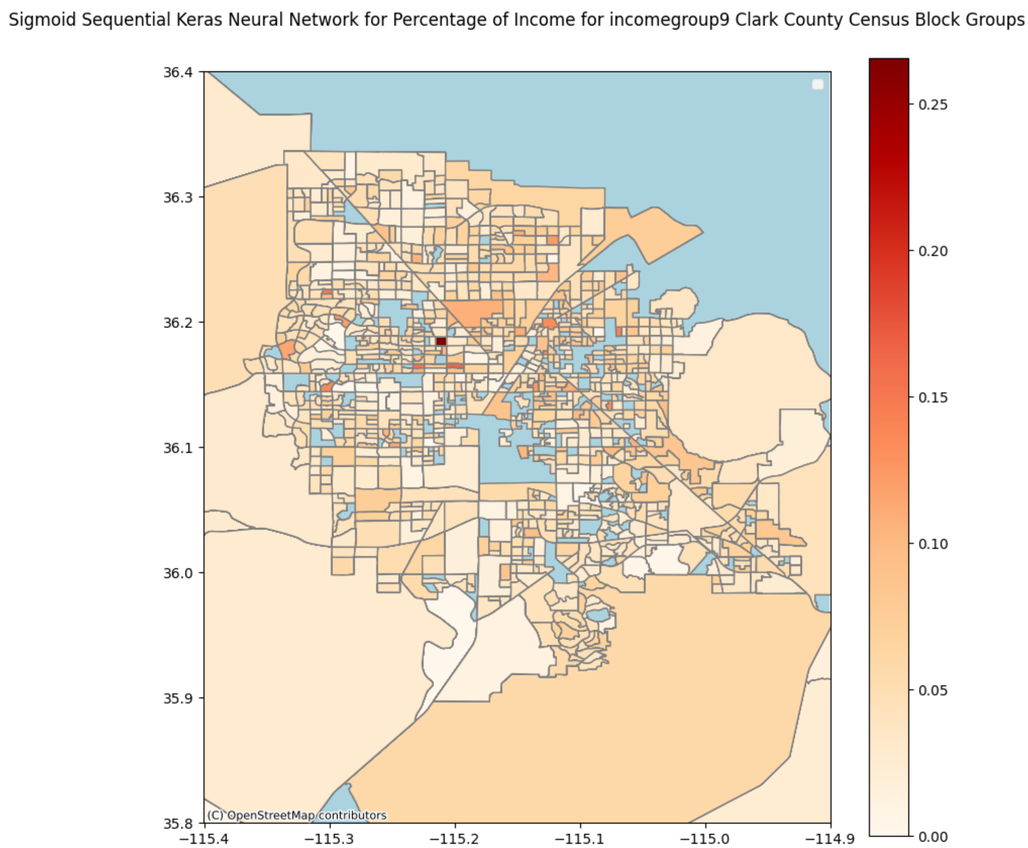


Figure 19: Block Level NN estimate of percentage in 45k to 50k Income Bracket

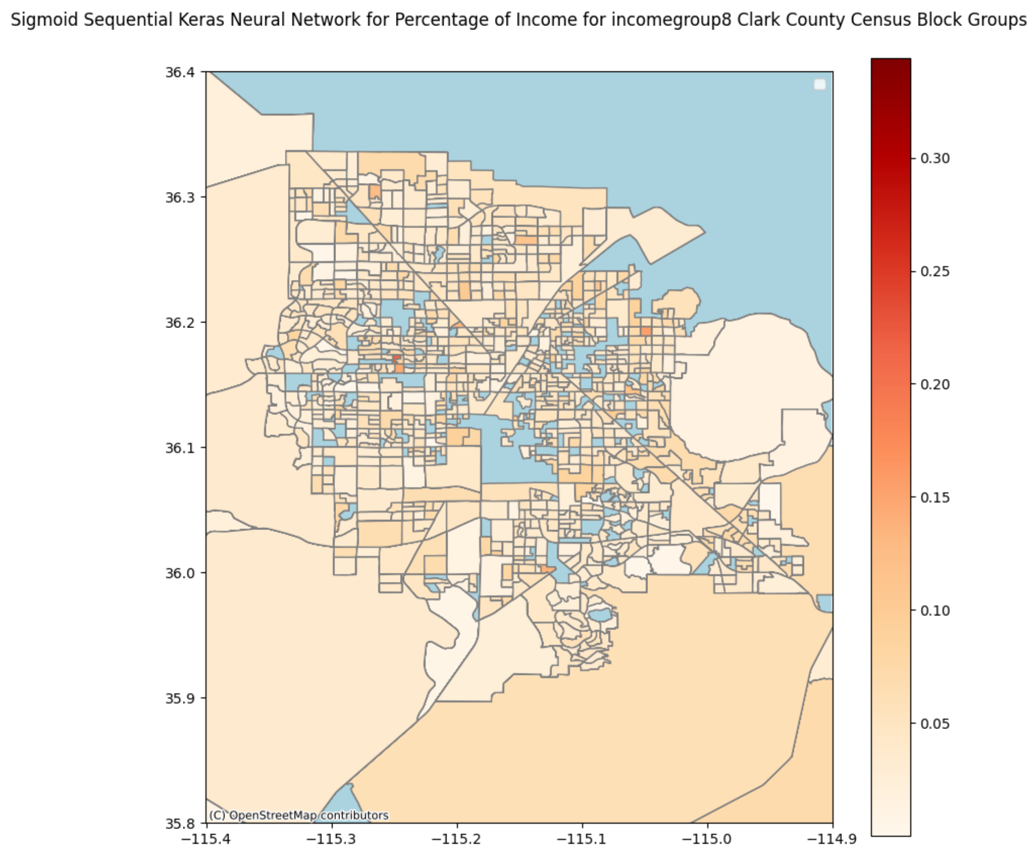


Figure 20: Block Level NN estimate of percentage in 40k to 45k Income Bracket

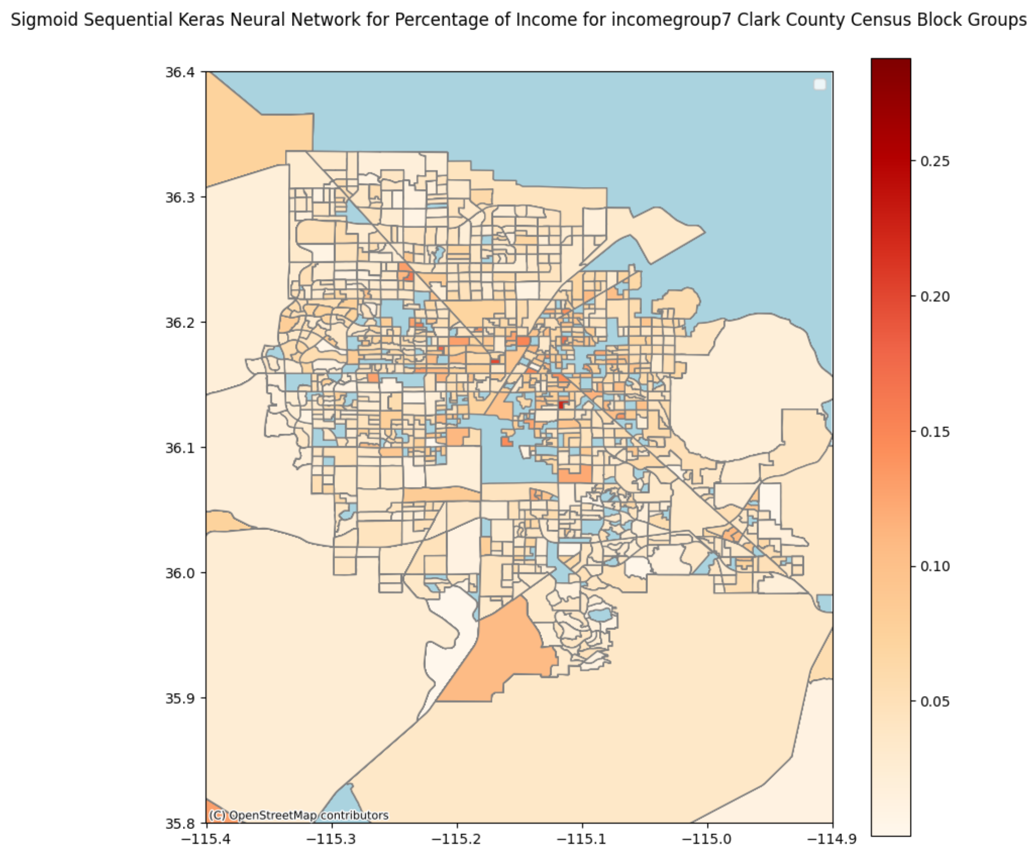


Figure 21: Block Level NN estimate of percentage in 35k to 40k Income Bracket

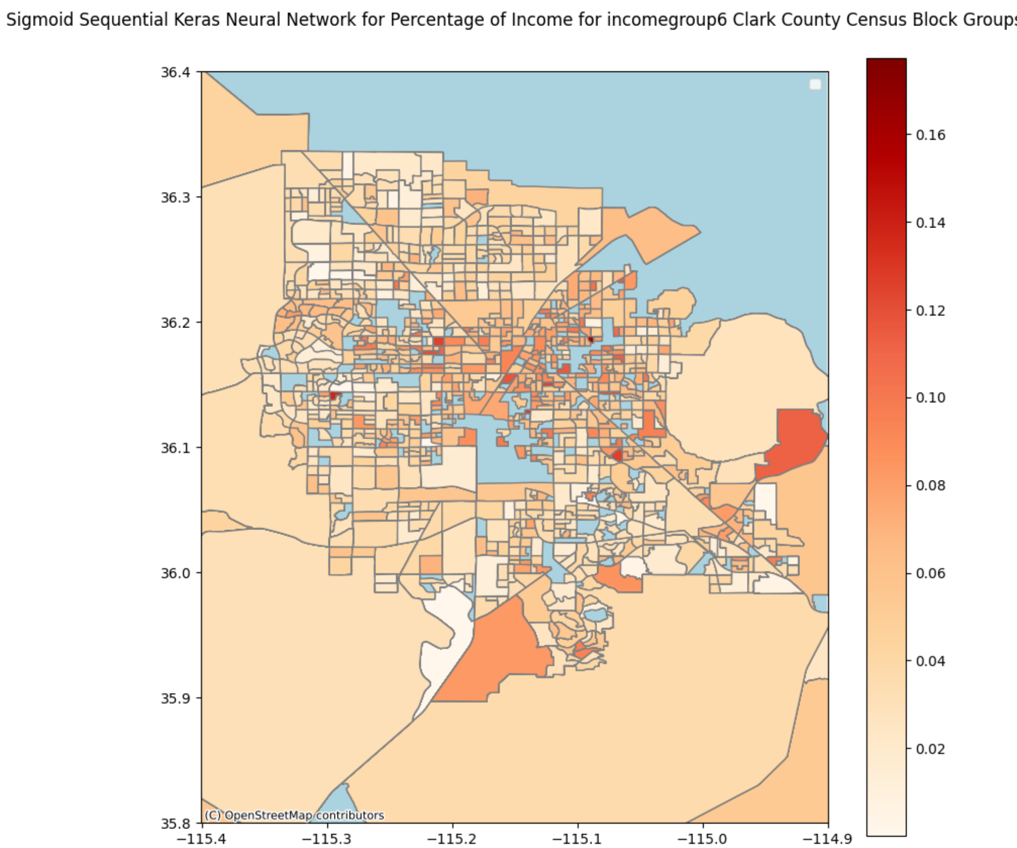


Figure 22: Block Level NN estimate of percentage in 30k to 35k Income Bracket

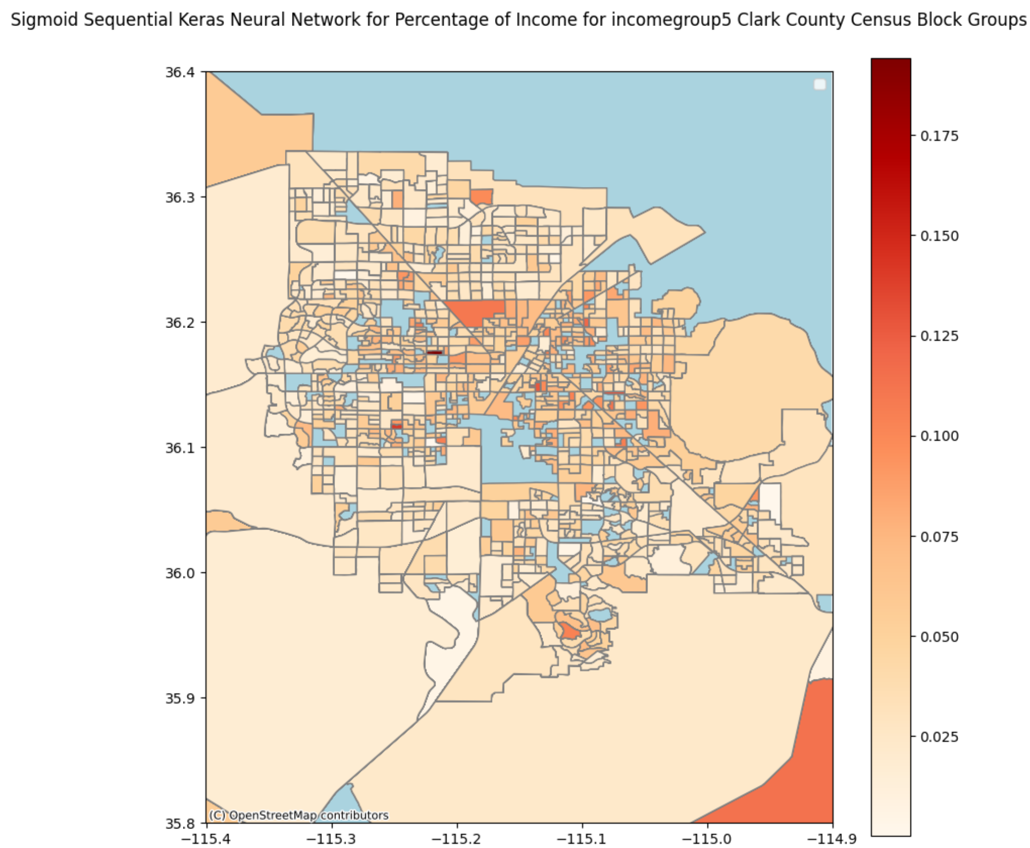


Figure 23: Block Level NN estimate of percentage in 25k to 30k Income Bracket

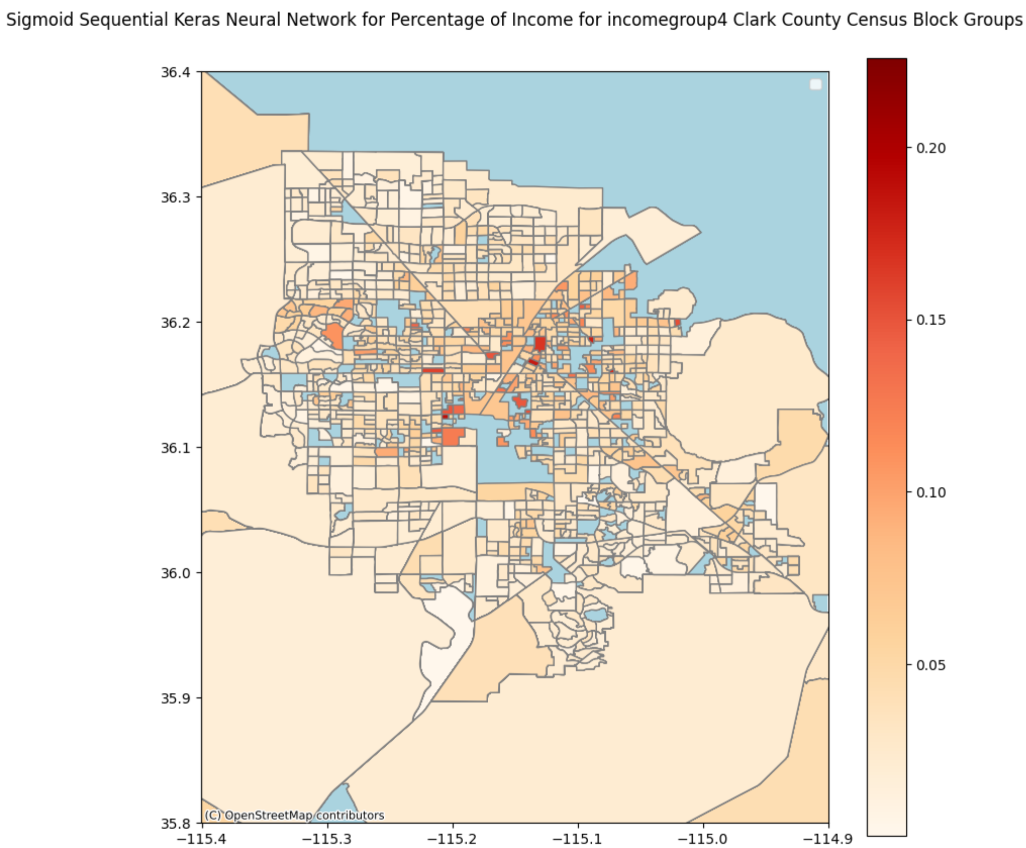


Figure 24: Block Level NN estimate of percentage in 20k to 25k Income Bracket

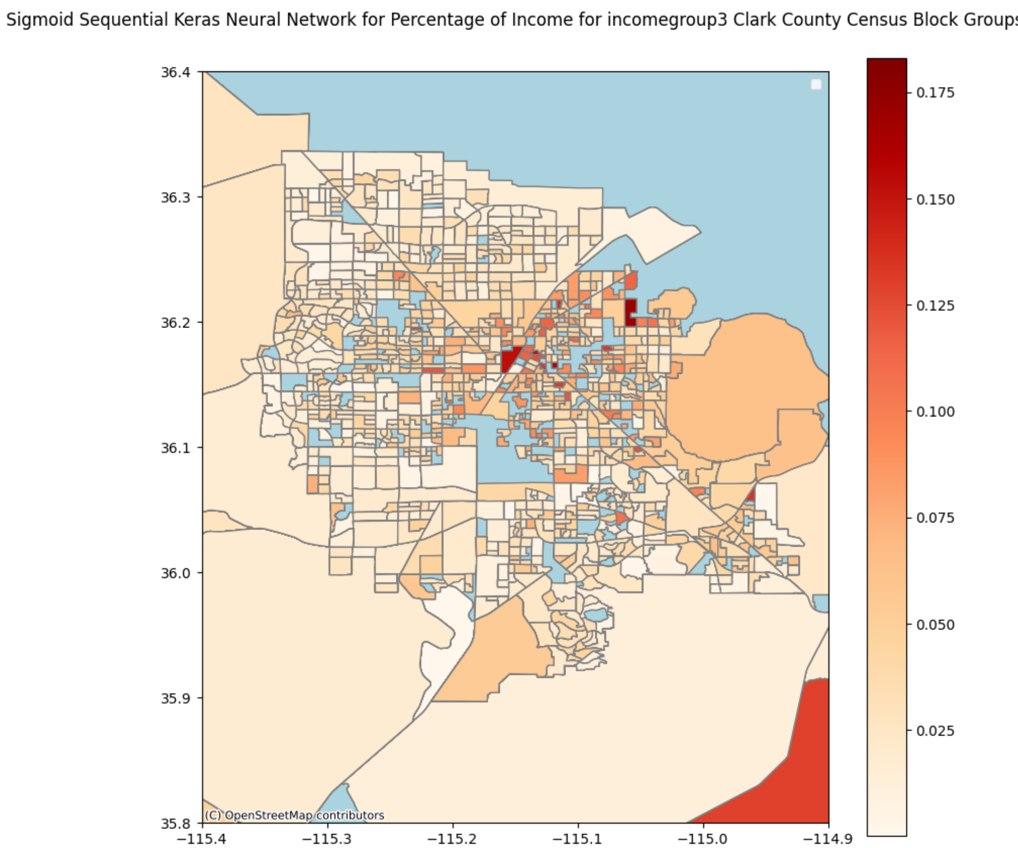


Figure 25: Block Level NN estimate of percentage in 15k to 20k Income Bracket

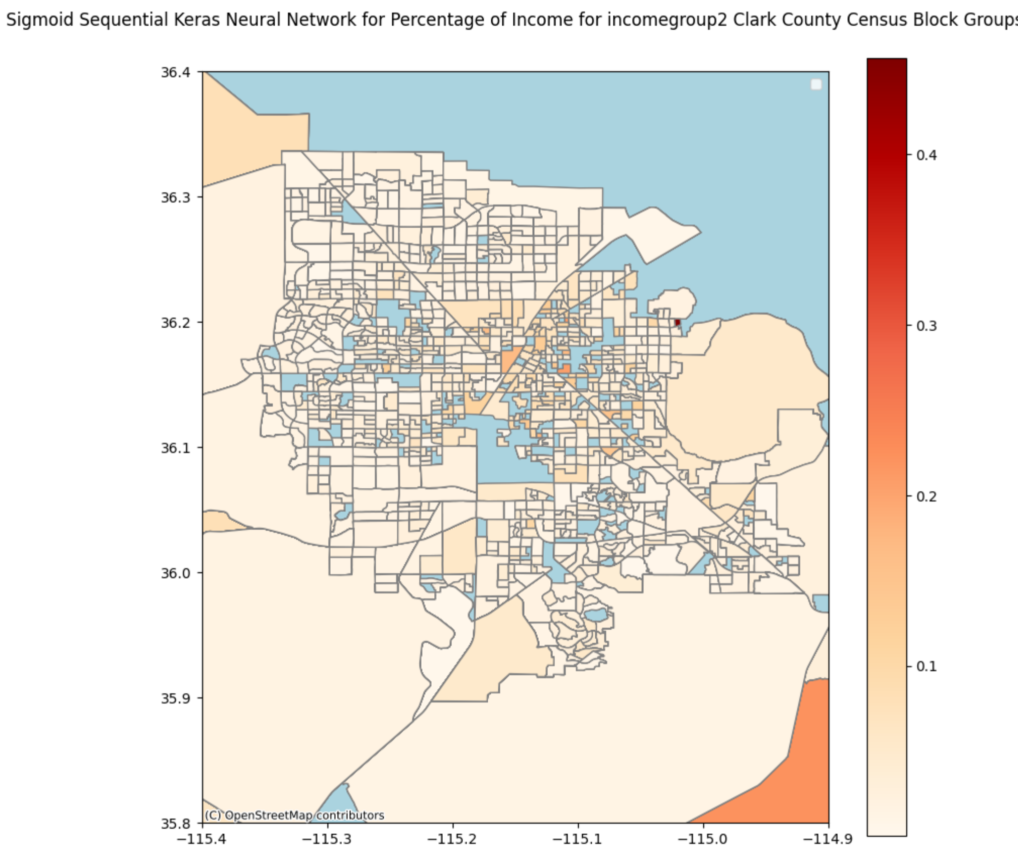


Figure 26: Block Level NN estimate of percentage in 10k to 15k Income Bracket

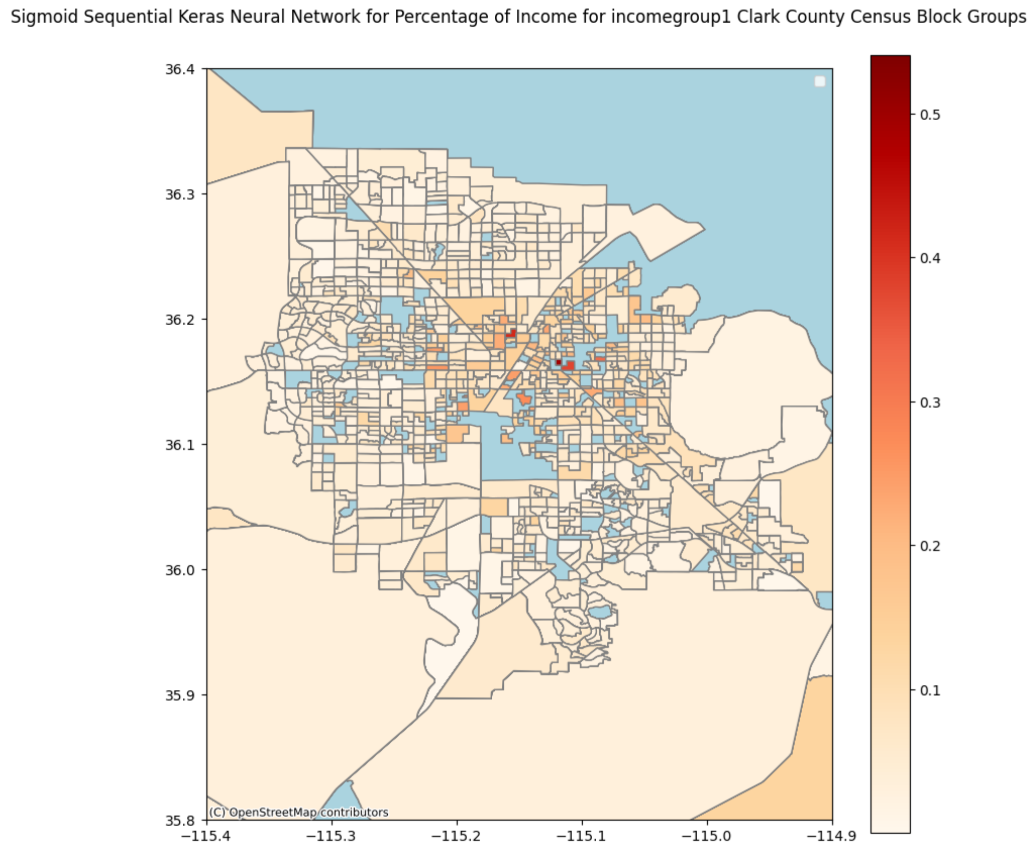


Figure 27: Block Level NN estimate of percentage in less than 10k Income Bracket

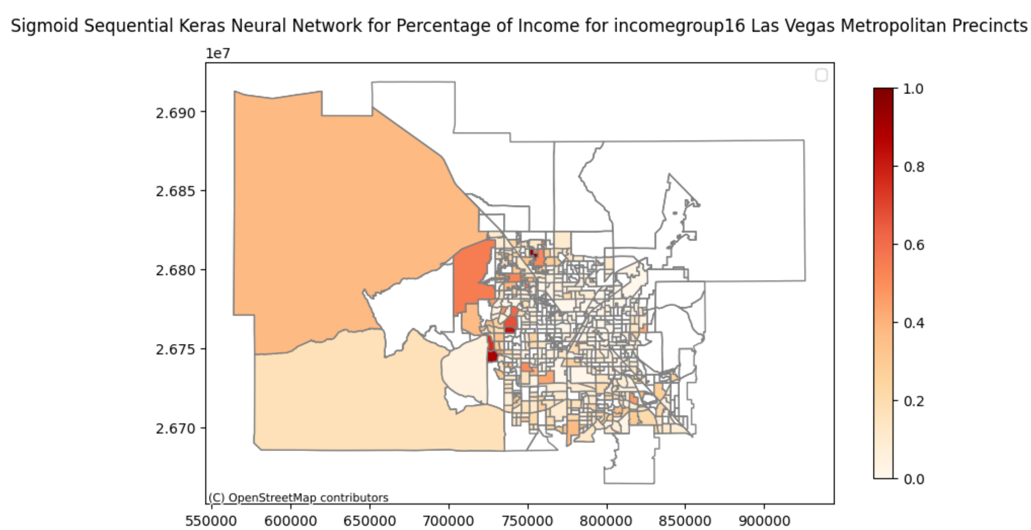


Figure 28: Precinct Level NN estimate of percentage in 200k+ Income Bracket

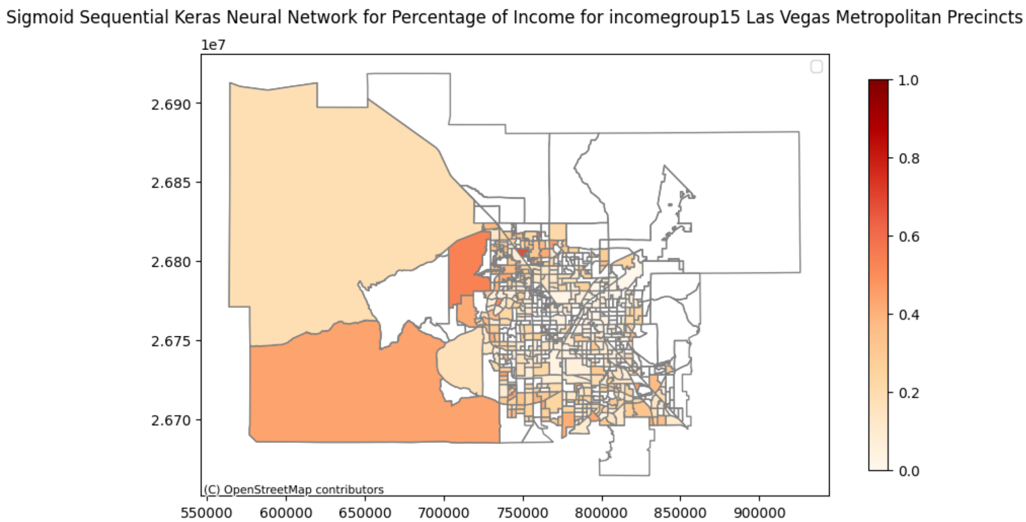


Figure 29: Precinct Level NN estimate of percentage in 150k to 200k Income Bracket

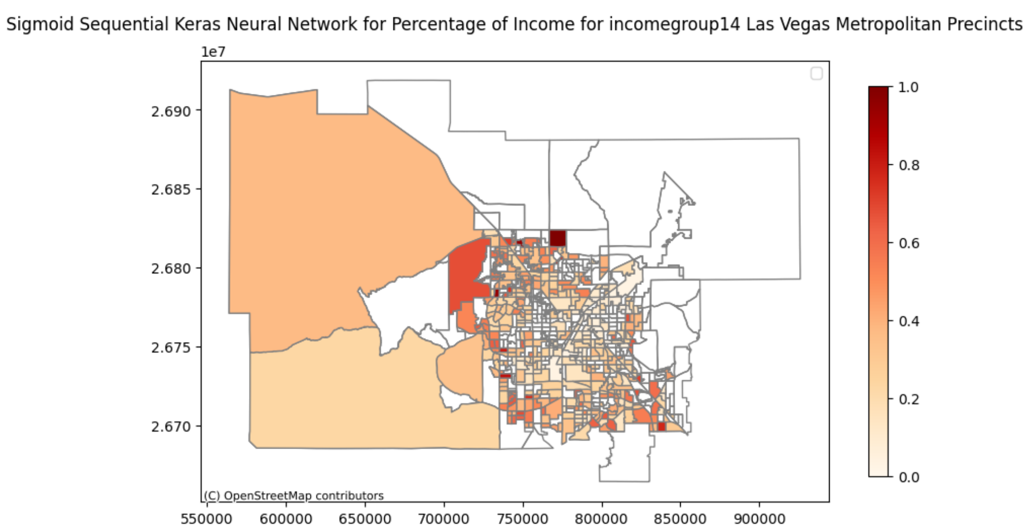


Figure 30: Precinct Level NN estimate of percentage in 125k to 150k Income Bracket

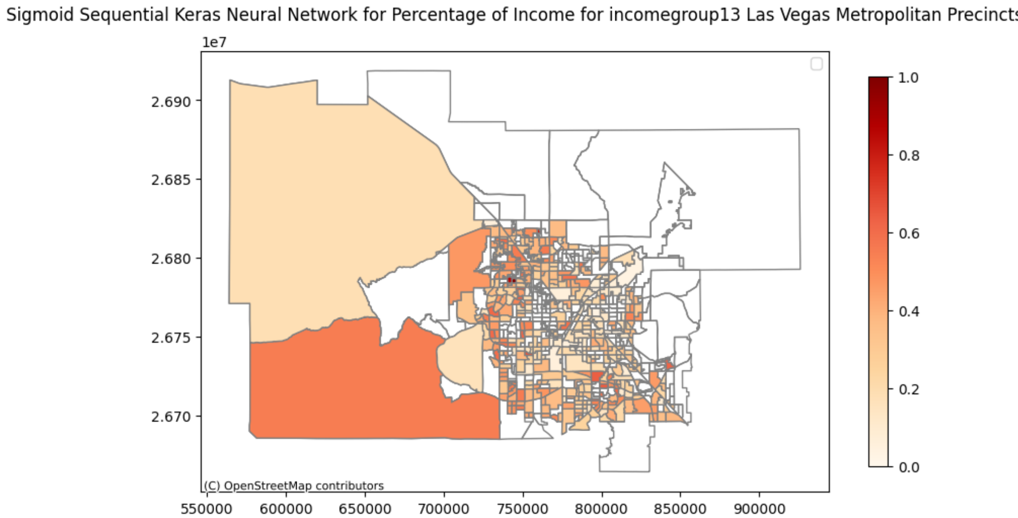


Figure 31: Precinct Level NN estimate of percentage in 100k to 125k Income Bracket

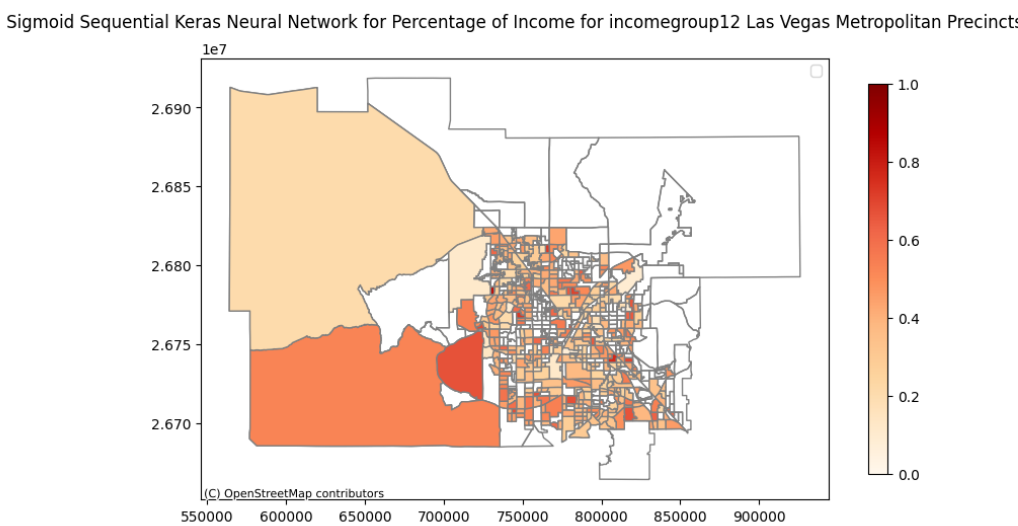


Figure 32: Precinct Level NN estimate of percentage in 75k to 100k Income Bracket

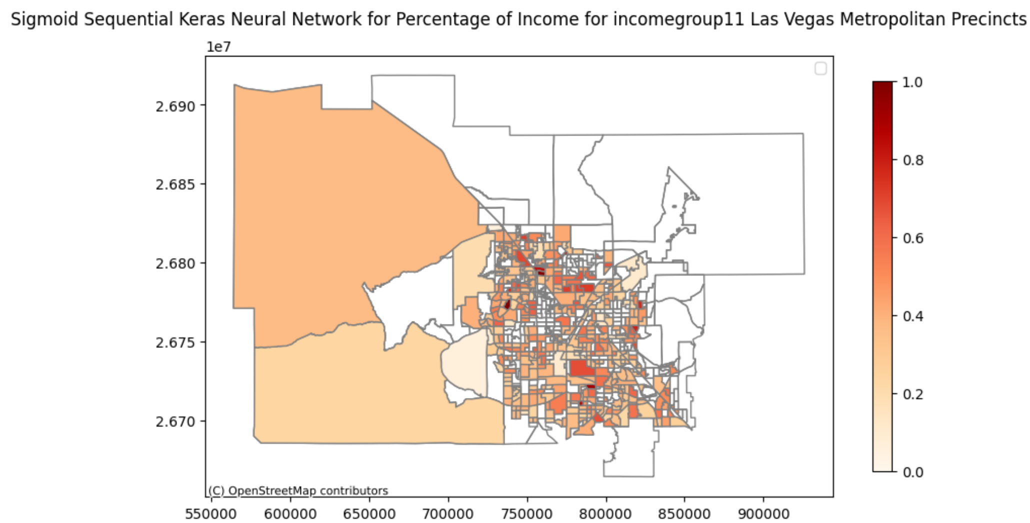


Figure 33: Precinct Level NN estimate of percentage in 60k to 75k Income Bracket

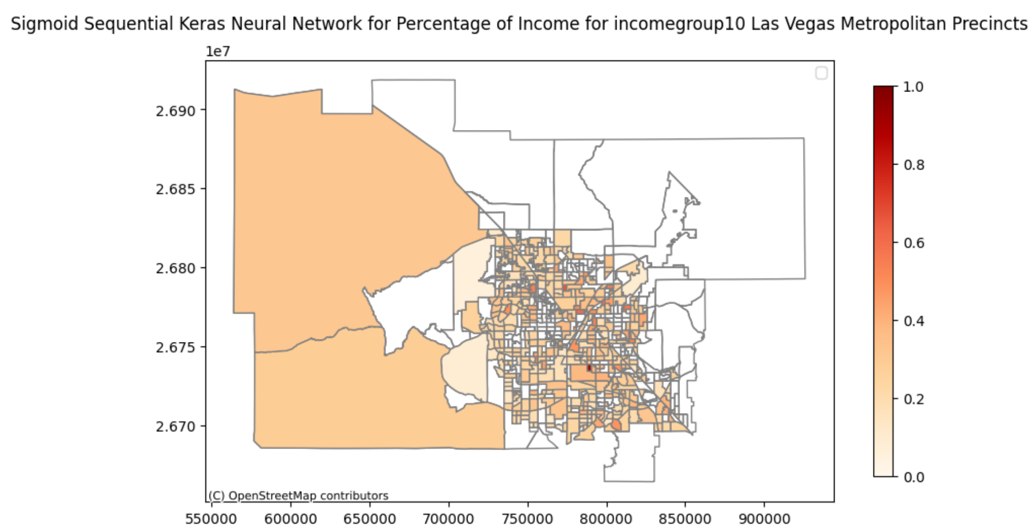


Figure 34: Precinct Level NN estimate of percentage in 50k to 60k Income Bracket

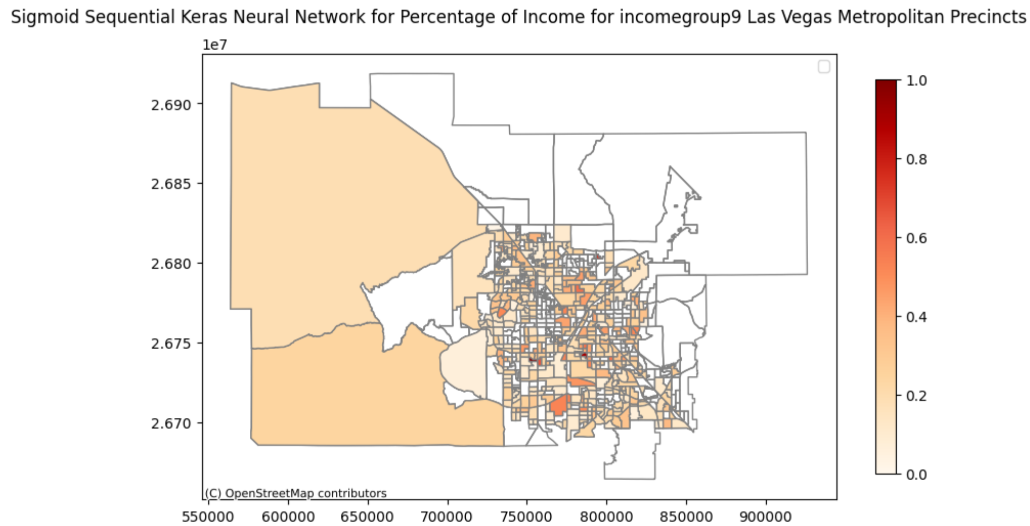


Figure 35: Precinct Level NN estimate of percentage in 45k to 50k Income Bracket

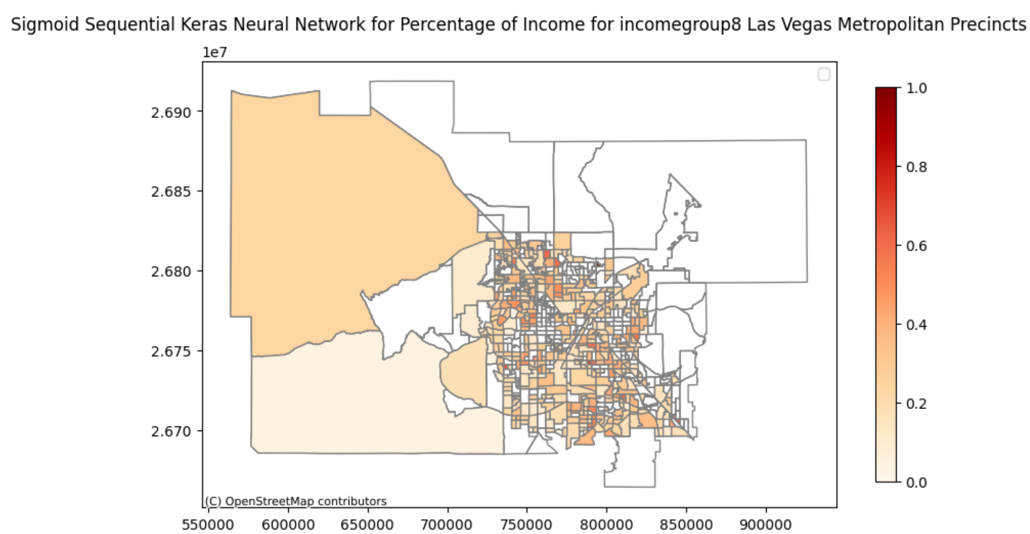


Figure 36: Precinct Level NN estimate of percentage in 40k to 45k Income Bracket

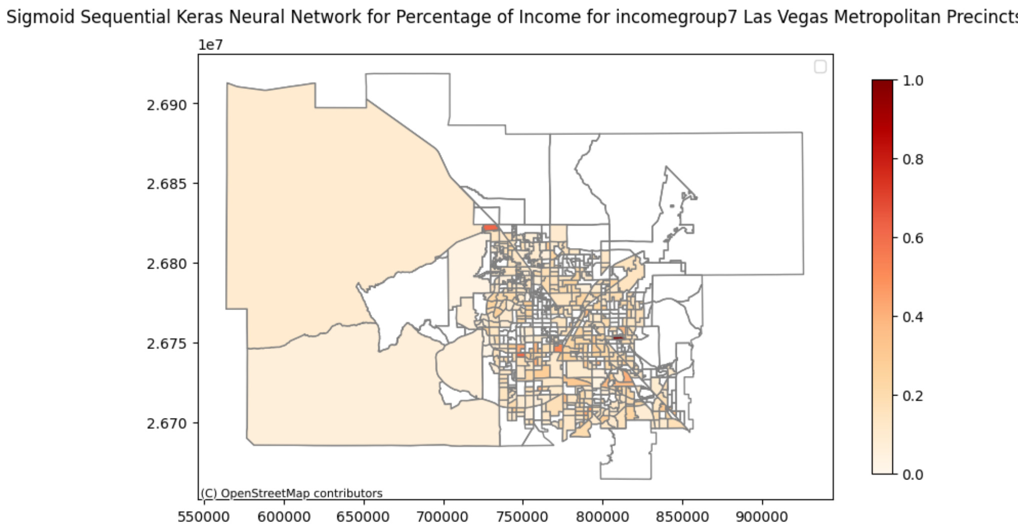


Figure 37: Precinct Level NN estimate of percentage in 35k to 40k Income Bracket

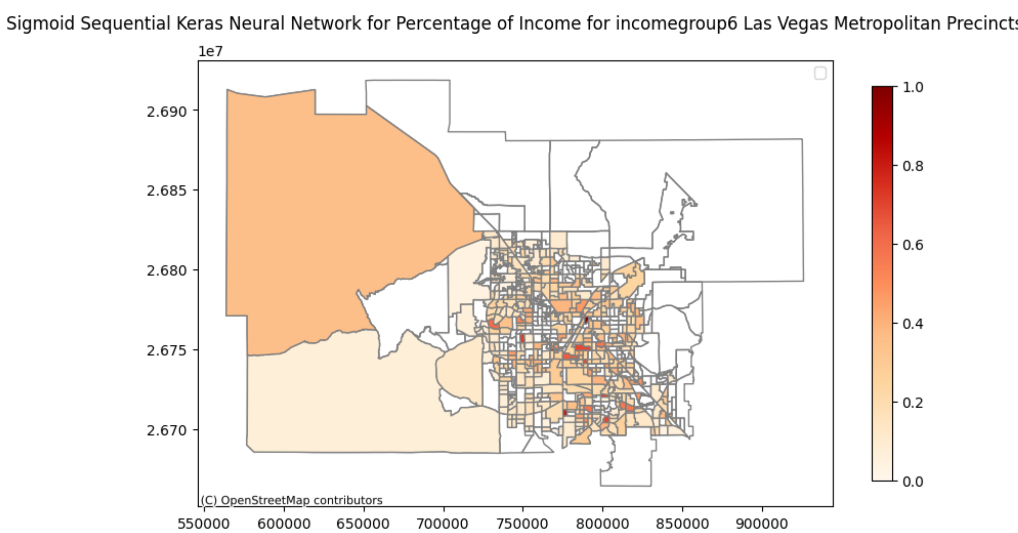


Figure 38: Precinct Level NN estimate of percentage in 30k to 35k Income Bracket

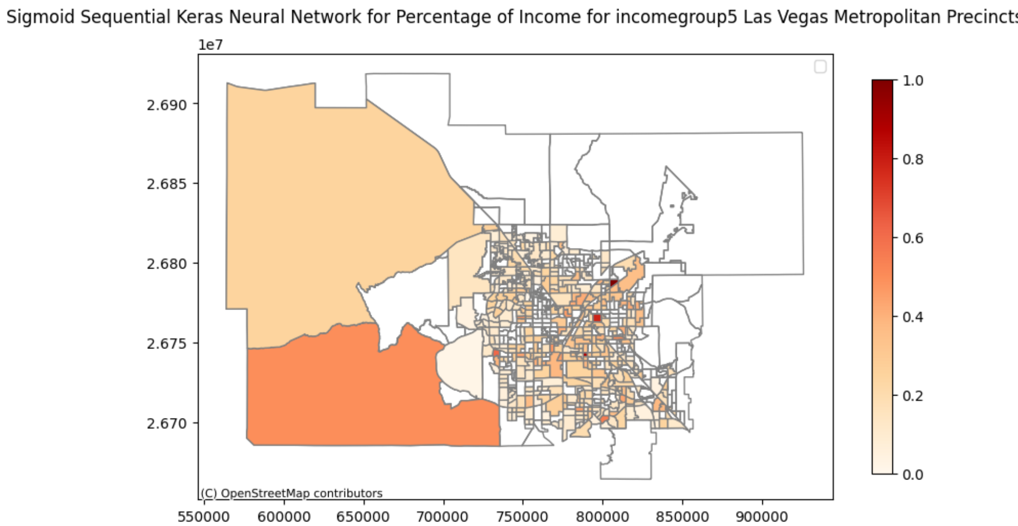


Figure 39: Precinct Level NN estimate of percentage in 25k to 30k Income Bracket

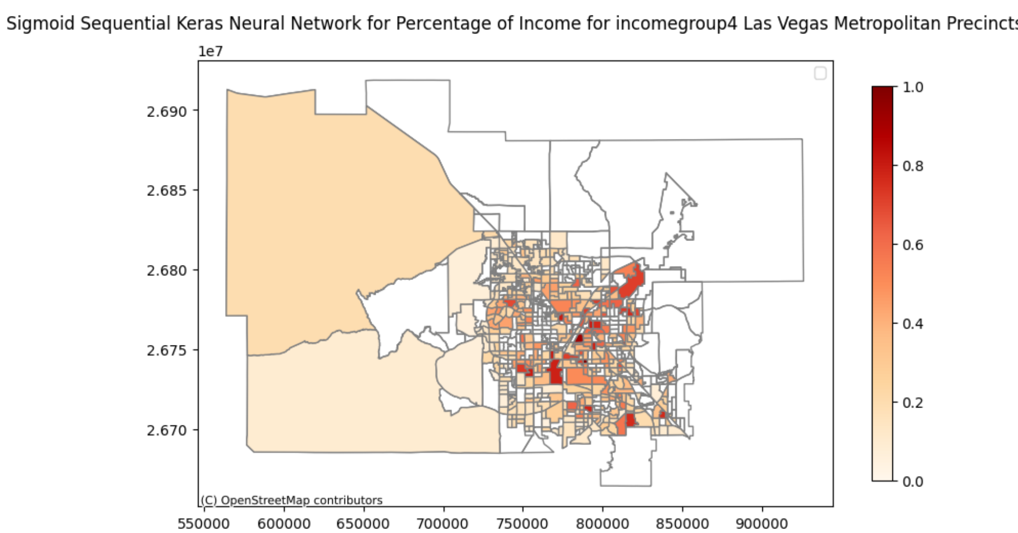


Figure 40: Precinct Level NN estimate of percentage in 20k to 25k Income Bracket

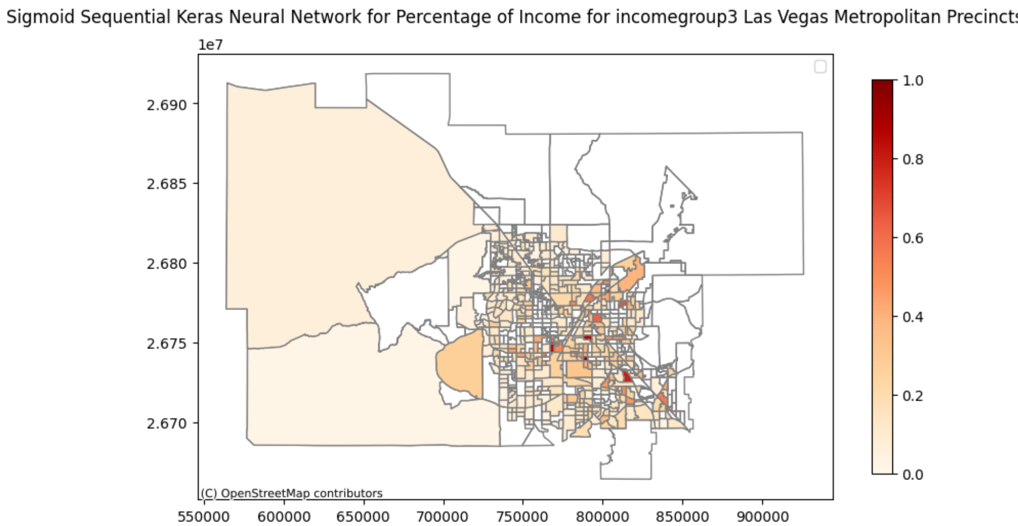


Figure 41: Precinct Level NN estimate of percentage in 15k to 20k Income Bracket

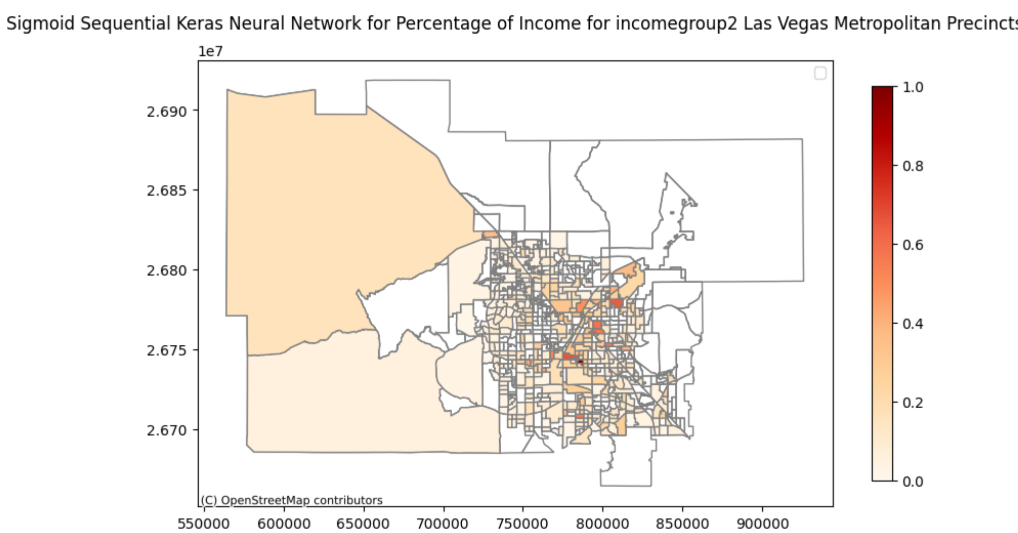


Figure 42: Precinct Level NN estimate of percentage in 10k to 15k Income Bracket

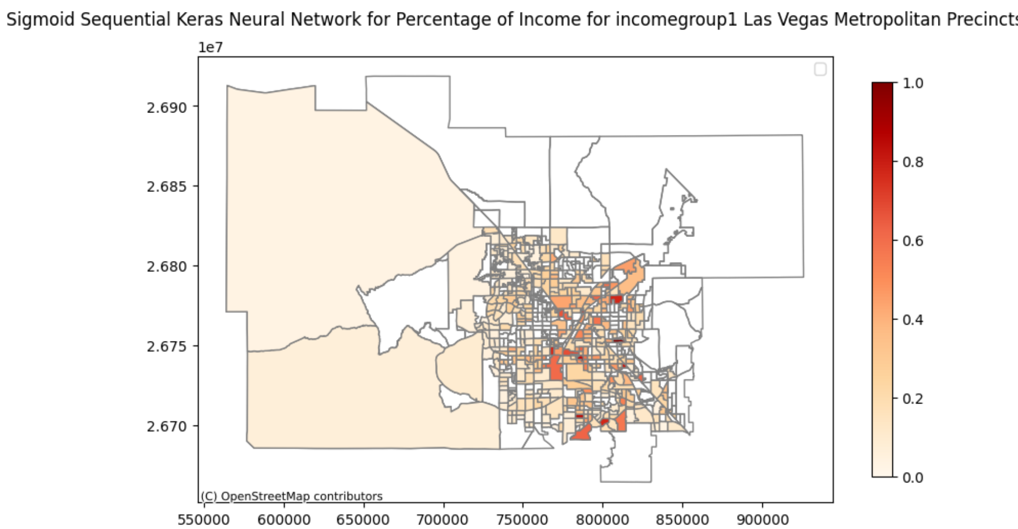


Figure 43: Precinct Level NN estimate of percentage in less than 10k Income Bracket