Unified Price Indices for Spatial Comparisons

by

Adam Hall

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Statistics)
in The University of Michigan
2019

Doctoral Committee:

Professor Roderick Little, Co-Chair
Professor Matthew Shapiro, Co-Chair
Professor Michael Elliot
Professor Johann Gagnon-Bartsch
Professor Ben Hansen

Adam Hall

hallac@umich.edu

ORCID iD 0000-0002-9102-8441

ACKNOWLEDGEMENTS

I'd like to express my sincerest gratitude to my advisors Roderick Little and Matthew Shapiro. Without your advice and guidance, this project would not have been possible. I owe much to Bettina Aten and Eric Figueroa, who first introduced me to the problem that would inspire this dissertation thesis. I'd also like to thank my wife Ruffa for supporting me (and putting up with me) throughout the years it took to complete this project. Finally, I'd like to thank my mother, whose patient sacrifices have made it possible for me to get to this point.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

Appendix

LIST OF ABBREVIATIONS

ACCRA American Chamber of Commerce Reseachers Association

ACS American Community Survey

BEA Bureau of Economic Analysis

BLS Bureau of Labor Statistics

CES Constant Elasticity of Substitution

CCER Council for Community and Economic Research

CGPI Common Goods Price Index

COLI Cost Of Living Index

CPI Consumer Price Index

ELI Entry Level Item

EPI Exact Price Index

FIPS Federal Information Processing Standards

MSA Metropolitan Statistical Area

RPP Regional Price Parities

SADJ Spread Adjustment

SUPI Spatial Unified Price Index

UPC Universal Product Code

UPI Unified Price Index

VADJ Variety Adjustment

ABSTRACT

This thesis proposes a method to measure fine-grained spatial differences in purchasing power. This method exploits the recent availability of computer-generated retail scanner data sets to compensate for the absence of sufficiently detailed pricing data in the national accounts. To correct for regional differences in product availability and quality, it extends the theoretical Unified Price Index (UPI) proposed in Redding and Weinstein (2016)[34] from the temporal to the spatial context. In this formulation, differences in product availability in different places are treated as analogous to differences in product availability at different times due to the "birth" and "death" of products across time.

It provides an example of how to apply this method, by estimating differences in food prices between Michigan counties from information in the Nielsen Retail Scanner Dataset.[1] The estimation of these indices can be divided into three steps. First, spatial UPIs are estimated comparing the cost of living between each pairing of Michigan counties for the 554 different categories of food included in the data. For example, one of these indices might compare the price of bacon in each county, while another might compare prices for fresh fruit, etc. Next, the GEKS method is applied to impose transitivity on each set of comparisons. These product level indices are aggregated into an index reflecting the cost of living for all food using a weighted geometric mean. The weights for each product in this process are the share

[1] Researcher(s) own analyses calculated (or derived) 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(s) 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.

of total food expenditure that the product accounts for. Finally, superpopulation estimates of the geometric variance of these indices are produced using a cluster bootstrap method. The indices we estimate suggest that the raw prices of food goods are similar between counties in Michigan, with a cross-county standard deviation of about 0.02. When differences in available product varieties are taken into account however, the estimated cost of living in rural areas is consistently higher than the estimated cost of living in more populous counties.

CHAPTER I

Introduction

Most people know that the value of a unit of currency varies based on the country in which it is spent, or across time. The same phenomenon holds for different regions of the United States. Economists and statistical bureaus have sophisticated techniques for adjusting nominal values into real terms before making comparisons between different countries or time periods, but comparisons within the United States are often still made on the basis of nominal dollar values. This is especially true for research that makes comparisons at low levels of geographic aggregation, such as counties.

Because the price indices required to adjust nominal figures at such low levels of aggregation are not available, researchers are often forced to ignore regional differences in purchasing power when making these comparisons. For example, studies may use nominal incomes to gauge differences in the purchasing power of consumers between different counties, even in cases where researchers might prefer real incomes for this purpose. These nominal figures do not take into account differences in the prevailing price levels faced by people in the regions being compared, nor do they consider differences in the kinds of products that are available in the areas being compared. As a result, the levels of consumption available to the people being researched may not be accurately represented. This can affect the outcomes of any study that has to contend with significant regional price variation.

For example, one could imagine a world in which the price of avocados is lower in states such as California and Florida, where avocados are commonly produced. States like Nebraska that are distant from the farms where avocados are grown might have higher prices for avocados, due to transportation costs.

In this hypothetical, the consumer's "real" purchasing power for avocados is higher in California and Florida than it is in Nebraska. Any researcher attempting to compare the number of avocados that such a consumer could purchase in California and Nebraska would therefore be misled if they relied only on the consumer's nominal wealth as their indicator. This kind of effect can add a bias of unpredictable direction to comparisons made on the basis of personal income, which could be corrected if the prevailing regional price levels were known.

The preceding example is contrived, but social scientists frequently make analogous comparisons using income or similar statistics in order to research topics such as public health, income inequality, and economic development. For example, Murray (2006)[32] used (race specific) county per capita income levels as one of the criteria defining the counties to be included in each of his paper's titular "Eight Americas," between which he found significant disparities in estimated life expectancy, mortality risk, and health care utilization. Murray's "Americas" are defined partially based the relationship of per capita nominal incomes to the poverty line. As a result, their scope could potentially change if the poverty line comparison were made in real rather than nominal terms.

Along similar lines, Kovandzic & Sloan (2002)[25] use per capita personal income as a control variable in their study of the effect of the level of policing within Florida counties on crime rates. Kovandzic & Vieraitis (2006)[26] use average county income and the percentage of within-county incomes below the federal poverty line as control variables in a study of the relationship between incarceration and crime rates. If the price of goods in one county is much higher than in another, studies that use nominal figures as independent variables are implicitly grouping together people with significant differences in potential consumption. The percentage of people classified as impoverished could change if real incomes were used as the basis for the poverty line calculation. This chance could have an impact on regression estimates, depending on the magnitude and direction of the differences between real and nominal incomes. If nominal incomes are noisy enough indicators of real incomes, the resulting spread between the real and measured values of the control variables could affect the outcome of statistical tests. In the worst cases, this could lead researchers to reject promising lines of inquiry, or to make spurious conclusions.

Studies relying on statistics derived from nominal incomes, such as Gini coefficients, may also be impacted. Muramatsu (2003)[31] found a relationship between within-county income inequalities, as measured by Gini coefficients computed based on nominal household incomes, and the rate of depression among elderly Americans. If there is significant price variation, there may also be a divergence between the nominal and real Gini coefficients computed in those counties. This might complicate the relationship between the predictor and outcome variables in unforseen ways. These are just a small sample of the broad range of studies that rely on the kinds of comparisons that might be improved by the cost of living indices developed in this proposal.

CHAPTER II

Literature Review

One reason that cost of living adjustments are often ignored when comparing small areas such as counties is that the information required to make the necessary adjustments is either unavailable, or prohibitively expensive. Producing price indices at such low levels of aggregation poses both practical and theoretical difficulties. For most goods and services, government statistical agencies do not collect the data necessary to estimate small-scale spatial price indices straightforwardly. Hence many of the indices making this kind of comparison are restricted in the regions that they can compare, rely on inference from samples in adjacent or "similar" areas to impute values for regions with no data in-sample, or require the participation of large numbers of volunteers to contribute economic data about the cost of items selected based on their individual preferences.

In Section 2.1 we briefly review the econometric literature about cost of living indices. We discuss the economic approach to price indices and the concepts that underlie it, as well as the various properties we might desire from any set of price indices comparing the cost of living across time and space. In Section 2.2, we review literature relating to two of the highest profile efforts at producing regional cost of living indices within the United States. Finally, in Section 2.3 we review several efforts to use bar code scanner databases to improve the measurement of cost of living indices.

2.1 Background on the Theory of Price Indices

Before we discuss the various attempts that have been made at measuring differences in the cost of living, we briefly summarize the economic literature undergirding the concept of the "cost of living." Generally, the cost of living is conceptualized in terms of differences in the level of utility available to a representative consumer across the time periods or spatial areas being compared. It is assumed that the consumer has some utility function $F(\vec{q})$ that relates the quantities $\vec{q}$ of each type of good they could consume to the level of satisfaction they would derive from that consumption.

The basic economic approach to price comparisons, described in Diewert (1976) [12] and Diewert (1979) [13], is to compare the maximum level of utility that is accessible to the representative consumer in the two comparison circumstances for some fixed budget. The function that relates the lowest possible cost at which the consumer can obtain u units of utility to the vector of relevant product prices $\vec{p}_i$ and quantities consumed $\vec{q}_i$ in circumstance i is referred to as the consumer's *expenditure function*:

$$C(\vec{p}_i, \vec{q}_i, u) \tag{2.1}$$

A price index comparing consumer welfare across any two circumstances i and j can be constructed as the ratio of the expenditure functions in each circumstance [13]:

$$P(\vec{p}_i, \vec{q}_i, \vec{p}_j, \vec{q}_j, u) = \frac{C(\vec{p}_j, \vec{q}_j, u)}{C(\vec{p}_i, \vec{q}_i, u)} \tag{2.2}$$

Typically, it is assumed that the consumer's expenditure function is homogenous in u, so that

$$C(\vec{p}_i, \vec{q}_i, u) = u \times C(\vec{p}_i, \vec{q}_i, 1) \tag{2.3}$$

In this case, we can represent $P(\vec{p}_i, \vec{q}_i, \vec{p}_j, \vec{q}_j, u)$ as

$$P(\vec{p}_i, \vec{q}_i, \vec{p}_j, \vec{q}_j) = \frac{C(\vec{p}_j, \vec{q}_j, 1)}{C(\vec{p}_i, \vec{q}_i, 1)} \tag{2.4}$$

the ratio of the minimum expenditure required to obtain one unit of utility in each circumstance, referred to as the consumer's *unit expenditure functions.* [13]

The indices i and j commonly represent different years in which the consumer lived, or different countries that the consumer might reside in. In the first case, $P(\vec{p}_i, \vec{q}_i, \vec{p}_j, \vec{q}_j)$ would compare inflation rates across time, while in the second they would compare the cost of living in each of the comparison countries. Most price indices comparing the cost of living across temporal or spatial units can be considered attempts at approximating this "true" cost of living difference under some set of assumptions about the consumer's utility, and index numbers can be evaluated based on the extent to which they are able to do so.

If an index is equal to the ratio of the consumer's expenditure functions for all reasonable[1] price and quantity vectors under some assumption about the consumer's utility function, we call that index an *exact price index.* If a price index is "exact" for a utility function that is flexible enough to approximate any twice continuously differentiable utility function in value, the index is referred to as a *superlative price index.* [13] Because of this flexibility, a superlative index number is robust to some degree of misspecification in the utility function. Many popular price index formulas, such as the Fisher, Tornqvist, and Walsh indices, are superlative in this sense, although the utility functions that they are exact for may differ.[16] Other price index formulas in common use, such as the Paasche, Laspeyres, and Jevons indices are exact, but for utility functions that are not sufficiently flexible for them to qualify as superlative indices.[16]

[1] The precise conditions required of the price and quantity vectors are listed in Diewert (1979)[13]

2.2 Existing Spatial Cost of Living Indices

In this section, we discuss the methods of two prior efforts to measure regional differences in cost of living within the United States, and the advantages and disadvantages associated with them. In Section 2.2.1 we discuss the Bureau of Economic Analysis' Regional Price Parities, which estimates differences in price between different states and metropolitan areas of the United States on the basis of data gathered for the Consumer Price Index. In Section 2.2.2 we discuss the Cost of Living Index (COLI) published by The Council for Community and Economic Research, which is based on data produced by a network of volunteers that price products across the country.

2.2.1 Regional Price Parities

The most widely available measures of the differences in price level within the United States are the Regional Price Parities (RPPs) estimated by the Bureau of Economic Analysis. Presently, RPPs are freely available from the Bureau of Economic Analysis at the state and metropolitan area level, but not at any finer level of aggregation. The same is true of inflation rates, which are available nationally and within some broad regional groupings, but not at the level of counties. But there are good reasons to believe that national, state, or metropolitan area price level estimates do not fully capture all the important variation between counties. For example, it seems unlikely that rental price levels in poor regions of Appalachia should be the same as rental price levels in the United States as a whole, or that the prevailing food prices in rural Texas communities should be the same as food prices in Dallas.

The first major obstacle to estimating RPPs is the scarcity of appropriate data. No comprehensive government survey exists that is designed to estimate price differences between regions. The primary source of the price data the BEA uses in the computation of the RPPs is the Consumer Price Index (CPI) survey, which is designed to compare overall United States price levels across time. For this reason, we briefly summarize the methods by which the Bureau of Labor Statistics (BLS) obtains this information.

The RPP is based on price quotations from stores in geographical units that include the majority of the US population, but do not conform to meaningful state or local boundaries. Instead, average prices are estimated within 38 CPI index areas, most of which roughly correspond to highly populated metropolitan areas and their suburbs. These index areas are the only levels at which the sampling weights necessary for estimation procedures are available, and thus the lowest level of spatial aggregation at which RPPs can be produced without resorting to some kind of inference to fill in the missing regional information. No rural areas are included within these areas, and hence the RPPs for predominantly rural states are inferred from prices gathered in the least populated suburbs of urban areas.[4]

Within each sampling area, the CPI collects the prices of items within 16 expenditure classes, each of which constitutes a broad class of goods and services. These expenditure classes include many essential categories of consumer expenditure, such as food, apparel, transportation, and housing.[4] To estimate the cost of living associated with each of these expenditure classes, a market basket of representative goods and services is chosen. For example, the price of food is represented by the prices of a list of particular types of goods, such as bread, ground beef, apples, etc.[29] The BLS collects data about 211 of these good types, which are referred to as "item strata." Many item strata can be divided into substrata based on some set of additional distinguishing characteristics. For example, the item stratum "cakes, cupcakes, and cookies" is divided into two substrata: "cakes and cupcakes" and "cookies."[29] These substrata are referred to as "entry level items," or ELIs.

The item strata that constitute the CPI market basket and their composition are chosen based on the frequency with which they are purchased in the Consumer

Expenditure Survey.[28] Because items are selected on this basis, indices based on the CPI basket may not fully represent the impact of regional or niche items on the cost of living. Additionally, because the Consumer Expenditure Survey takes time to conduct, the basket for each year is selected based on survey results from previous years. As a result, CPI data for a given year will not include any information about items that appeared between its publication and the date of the previous Consumer Expenditure Survey.[28]

In each sampled store, a single product variety is randomly sampled from the store's stock to represent the price level of its associated ELI.[29] For example, a 2 liter bottle of Coca-Cola might be chosen to represent all carbonated beverages. This product will be repeatedly priced on a monthly or bimonthly basis for the remainder of the year.[29] The aggregate of these price quotations, along with notes about the characteristics of each sampled product variety and the location of the store in which it was sampled, constitute the raw data from which the BEA estimates RPP indices.

This estimation is done in two broad stages. In the first stage, RPPs are estimated for each of the 16 expenditure categories within the 38 CPI index areas. To estimate these RPPs, the BEA requires estimates of the average price $p_{il}^{\mathbb{IL}}$ for each item stratum i in $\mathbb{I}$ and CPI index area l in $\mathbb{L}$, and the total dollar expenditure $e_{il}^{\mathbb{IR}}$ for each item strata i in $\mathbb{I}$ and index area l in $\mathbb{L}$.

To estimate the prices $p_{il}^{\mathbb{IL}}$, data on various characteristics of the included goods are used in what is known as a "hedonic regression model" to estimate a "quality adjusted" price for each kind of good or service.[4][2] This is necessary to mitigate potential biases caused by differences in the quality of goods and services sampled in different areas. Such biases can occur when price differences due to quality differences in the sampled items are conflated with the geographic price differences we want to measure. For example, imagine that the price of a bag of Honeycrisp apples sold in Saginaw is higher than the price of a bag of Red Delicious apples sold in Lansing.

We might not want to infer from this that Saginaw has a higher cost of living than Lansing without adjusting for the fact that Honeycrisp apples are considered a higher quality product than Red Delicious apples.

The BEA attempts to make the prices between those areas more comparable by fitting a regression model to estimate the prices for apples in each area holding apple variety, organic certification, and size, constant. Concretely, if $\mathbb{V}$ is the set of all observed apple varieties (omitting a single reference variety), $variety(j)$ is the apple variety of observation j, $\delta(variety(j) = v)$ is a dummy variable for whether observation j is of variety v, $organic_j$ is a dummy variable for whether observation j is organic, and $size_j$ is the apple size of observation j, then the BEA fits the following hedonic regression model for apples[2]:

$$\begin{aligned} log(p^{\mathbb{IL}}_{apples,l})_j = &\alpha_l + \sum_{v \in \mathbb{V}} \beta^{VAR_v} \delta(variety(j) = v) + \\ &\beta^{SIZE} size_j + \beta^{ORG} organic_j + \varepsilon_j \end{aligned} \tag{2.5}$$

Here α_l is the effect on log apple prices of being sold in index area l. β^{VAR_v}, β^{SIZE} and β^{ORG} are the effects on log apple prices of being apple variety v, apple size, and organic status respectively. ε is a normally distributed noise term, assumed to have mean zero and constant error variance.

More generally, for each item i in the top 75 item strata in terms of total expenditure, the BEA estimates ELI specific hedonic regressions based on a set of measured characteristics $\mathbb{K}_i$. These regressions are of the form:

$$log(p^{\mathbb{IL}}_{il})_j = \alpha_l + \sum_{k \in \mathbb{K}_i} \beta^k X_{kj} + \varepsilon_j \tag{2.6}$$

where $log(p^{\mathbb{IL}}_{il})_j$ is the log price of the jth observation within item i, and X_{kj} is a measurement of characteristic k taken on observation j.[2] Once these models are fit, hedonically adjusted prices $p_{il}^{\mathbb{IL}}$ can be estimated as the exponentiated weighted least squares estimates of the area coefficients:[2]

$$\hat{p}_{il}^{\mathbb{IL}} = exp(\hat{\alpha}_l) \tag{2.7}$$

Because specifying separate regression models for every ELI would be very labor intensive, the average prices of the remaining products are estimated using the "Weighted Country Product Dummy" method.[4] This involves pooling data from the remaining ELIs and solving for the expenditure weighted least squares estimates $\hat{\alpha}_l$ and $\hat{\beta}_i$ from a population model of the form

$$log(p_{il}^{\mathbb{IL}}) = \alpha_l + \beta_i + \varepsilon_{il} \tag{2.8}$$

The resulting fitted values

$$\hat{p}_{il}^{\mathbb{IL}} = exp(\hat{\alpha}_l + \hat{\beta}_i) \tag{2.9}$$

are used as the estimated price relatives for these items. When there are no missing observations, this is equivalent to taking the geometric mean of the price observations within each item.[4]

The estimates $\hat{e}_{il}^{\mathbb{IL}}$ of the total dollar expenditures on item i in CPI index area l are then used in combination with the $\hat{p}_{il}^{\mathbb{IL}}$ to estimate "notional quantities" $q_{il}^{\mathbb{IL}}$:

$$\hat{q}_{il}^{\mathbb{IL}} = \frac{\hat{e}_{il}^{\mathbb{IL}}}{\hat{p}_{il}^{\mathbb{IL}}} \tag{2.10}$$

These quantities can be interpreted as estimates of the number of units sold within the period under consideration.[4]

[2] For model identification reasons, one of the index areas in $\mathbb{L}$ is arbitrarily chosen as a reference area, and its value of α is set equal to zero. Reference levels are also chosen for all categorical characteristic variables that may appear in the hedonic model.

The BEA uses these data to estimate RPP indices for each of the 16 expenditure categories within the 38 CPI index areas. This is accomplished using the Geary multiliateral price index[21], adapted from its initial use for comparing international purchasing power parities in the Penn World Table[27] by treating regions of the US analogously to countries. If we define $\mathbb{I}_g$ as the market basket of items i within each expenditure category g in the set of expenditure categories $\mathbb{G}$, this method solves the following system of equations:

$$RPP_{gl}^{\mathbb{GL}} = \frac{\sum_{i \in \mathbb{I}_g} \hat{p}_{il}^{\mathbb{IL}} \hat{q}_{il}^{\mathbb{IL}}}{\sum_{i \in \mathbb{I}_g} \gamma_i^{\mathbb{IL}} \hat{q}_{il}^{\mathbb{IL}}} \tag{2.11}$$

where

$$\gamma_i^{\mathbb{IL}} = \sum_{l \in \mathbb{L}} \frac{\hat{p}_{il}^{\mathbb{IL}} \hat{q}_{il}^{\mathbb{IL}}}{RPP_{gl}^{\mathbb{GL}} \sum_{l \in \mathbb{L}} \hat{q}_{il}^{\mathbb{IL}}} \tag{2.12}$$

The $\gamma_i^{\mathbb{IL}}$s represent the average price of ELI i across all areas, and are therefore referred to as "international prices" in the context of the Penn World Table, which estimates the cost of living between countries. In this application they are more accurately interpreted as "national prices," since RPPs compare regions within the United States.

The RPP for expenditure category g is thus the ratio of total expenditures priced in nominal dollars to total expenditures priced in national prices in each area to be our price index, as in Equation 2.11. [4] This results in RPP estimates like 0.8 or 1.15, meaning that the price levels for g in an area are 20% lower than or 15% higher than the US average respectively. Geary indices are not themselves superlative in the sense discussed in Section 2.1, but they closely approximate indices that are.[11] They also have the advantage of additivity, which means that indices at different levels of aggregation can be easily produced, and that these indices will always be consistent with the overall index.[4]

In the second stage, the RPP estimates for the CPI index areas are used to estimate indices for states and/or metropolitan areas using the following method. Estimates $\hat{m}_{ic}$ of the total money income earned by all of the households within each county are gathered from the American Community Survey (ACS). These are used to ratio allocate the expenditures within each category and CPI sampling area down to the counties contained within that area.[4] The expenditure within each county is assumed to be directly proportional to its share of the total income within the sampling area. For example, if a county represents 25% of the total income earned by households in a sampling area, then 25% of that area's total expenditure is allocated to it.

Stated more formally, for each county c in $\mathbb{C}_l^{\mathbb{L}}$, the set of US counties contained in CPI index area l, the BEA estimates the expenditure on expenditure group $g \in \mathbb{G}$ as

$$\hat{e}_{gc}^{\mathbb{GC}} = \frac{\hat{m}_c}{\sum_{k \in \mathbb{C}_l^{\mathbb{L}}} \hat{m}_k} \hat{e}_{gl}^{\mathbb{GL}} \tag{2.13}$$

The average price level in each county is then assumed to be the same as the RPP generated in the sampling area that contains it [4]:

$$RPP_{gc}^{\mathbb{GC}} = RPP_{gl}^{\mathbb{GL}} \tag{2.14}$$

RPPs for the target areas (states or MSAs) in each expenditure class are then estimated by taking a weighted geometric average over the RPP estimates for the counties that compose each target area:[4]

$$RPP_{ga}^{\mathbb{GA}} = \prod_{c \in \mathbb{C}_a^{\mathbb{A}}} \left(RPP_{gc}^{\mathbb{GC}} \right)^{\omega_{gc}} \tag{2.15}$$

where the weights ω_{gc} are are the shares of the total within-area expenditure accounted for by each county:

$$\omega_{gc} = \frac{\hat{e}_{gc}^{\mathbb{GC}}}{\sum_{k \in \mathbb{C}_a^{\mathbb{A}}} \hat{e}_{gk}^{\mathbb{GC}}} \tag{2.16}$$

This process is repeated for the five years that are closest to the target year for our estimate.[4] For example, an estimate of the RPPs in 2011 would draw in data from years in the range of 2009-2013. [3] Based on these estimates, the weighted geometric mean $\overline{RPP}_{ga}^{\mathbb{GA}}$ of the expenditure category RPPs across this five-year window is taken, in order to improve the numerical stability of the resulting estimates.[3] These averages, along with estimates of the target area a level expenditures $e_{ga}^{\mathbb{GA}}$ in expenditure group g, are then used to estimate notional quantities:

$$\hat{q}_{ga}^{\mathbb{GA}} = \frac{\hat{e}_{ga}^{\mathbb{GA}}}{\overline{RPP}_{ga}^{\mathbb{GA}}} \tag{2.17}$$

Based on these estimates, a second Geary multilateral index is used to the 16 expenditure category RPPs to an all-items RPP for each target area $a \in \mathbb{A}$ by solving the following system of equations:

$$RPP_a^{\mathbb{A}} = \frac{\sum_{g \in \mathbb{G}} \overline{RPP}_{ga}^{\mathbb{GA}} \hat{q}_{ga}^{\mathbb{GA}}}{\sum_{g \in \mathbb{G}} \gamma_g^{\mathbb{GA}} \hat{q}_{ga}^{\mathbb{GA}}} \tag{2.18}$$

where

$$\gamma_g^{\mathbb{GA}} = \sum_{g \in \mathbb{G}} \frac{\overline{RPP}_{ga}^{\mathbb{GA}} \hat{q}_{ga}^{\mathbb{GA}}}{RPP_a^{\mathbb{A}} \sum_{g \in \mathbb{G}} \hat{q}_{ga}^{\mathbb{GA}}} \tag{2.19}$$

The method sketched[3] above was designed for the estimation of RPPs in larger regions, such as states and metropolitan areas, that can be defined as collections of counties. In those cases, the model would average together several of the county level estimates into the final price estimate based on a list of the counties that comprise that area. Thus some level of imprecision in the county level price estimates can be smoothed out in the final result. If each county itself is the final target of estimation however, this method is inadequate, because it will produce estimates that are

[3]For the sake of brevity, several details about the CPI sampling methodology and the estimation of the RPPs have been omitted from this review. For more detail about CPI sampling methods, consult the Consumer Price Index Handbook of Methods[29]. For more detail about estimating the RPPs, see Aten (2005)[2], Aten (2011)[4], and Aten (2017)[3].

essentially identical to the estimate for sampling area which contains that county.

The reason for this is that the imputation of the prices and expenditures from the CPI areas to the counties is extremely blunt. It's not clear how much these imputations can be improved without additional data at the county level, either to use directly in the estimation or to validate our assumptions. Otherwise we must attempt to infer characteristics about the price and expenditure levels of individual areas from aggregate data, which would require strong and untestable assumptions. Additionally, the RPPs are based exclusively on data gathered for goods in the CPI market basket. Because this market basket is constructed on the basis of estimates from the previous year's Consumer Expenditure Survey, the contribution of newer or regional niche items with low nationwide annual expenditures may be ignored. Further, price estimates from hedonic regression models may conceal significant internal variation in quality when there is heterogeneity in the product varieties sold across different counties.

2.2.2 CCER Cost of Living Index

Another measure of the local cost of living is the cost of living index (COLI), produced by the Council for Community and Economic Research (CCER). This index is sometimes called the ACCRA cost of living index, as the CCER was formerly known as the American Chamber of Commerce Reseachers Association (ACCRA). It provides estimates of the cost of living in large cities within the United States for various classes of commodities, and is available for purchase each year for a cost of between \$50 and \$550, depending on the particular information desired.[8] The COLI is based on data gathered by volunteers associated with regional groups, such as chambers of commerce, economic development agencies, or universities.[9]

CCER's primary concern when estimating the COLI is comparing the cost of living of cities and urban areas. As such, volunteers are directed to gather data exclusively from places within federally designated metropolitan statistical areas (MSAs), or cities of at least 35,000 people within counties with a minimum population of 50,000. Within such restrictions, volunteers are advised that "as a practical matter, you should price the urbanized portion of your metro area or place" in order to prevent the accidental inclusion of information from more sparsely populated locations.[8] Areas that do not meet these qualifications, or in which no volunteers contribute price information, are excluded.

The comparisons made by the COLI are based on six broad categories of goods and services: groceries, housing, utilities, transportation, health care, misc. goods and services. [8] Each of these categories is assigned a weight reflecting the importance of that category. The price level for each category is represented by the prices of a fixed basket of goods and services, and each category is associated with its own specific guidelines about data collection for the goods and services in those baskets.

For example, the grocery category is based on a basket of 26 different goods, such as Coke or lettuce, chosen for their broad availability. Volunteers pricing the goods in this basket are generally directed to gather the prices of these goods from a minimum of five different retail stores, three times per year.[8] They are expected to use their own judgement about what individual varieties of good to price (e.g. what brand of lettuce, what flavor of Coke, etc.). Because the target consumers of the COLI are members of "moderately affluent professional and managerial households,"[8] these volunteers aim to price product varieties that are important to that subset of the population in each area. With some restrictions, such as the requirement that prices should be obtained exclusively from chain supermarkets, they are also empowered to choose which stores they visit to price these goods. [8]

Given the average prices $p_{im}^{\mathbb{I}\mathbb{M}}$ for each of the 59 items i in the market basket $\mathbb{I}$ within each element m in the set $\mathbb{M}$ of metropolitan statistical areas (MSAs), along with expenditure shares ω_i for each good in the basket sourced from the BLS Consumer

Expenditure Survey, the COLI comparing the cost of living in area m to the national average cost of living is calculated as

$$COLI_m = \sum_{i \in \mathbb{I}} \omega_i \left(\frac{p_{im}^{\mathbb{IM}}}{\bar{p}_{im}^{\mathbb{I}}} \right) \tag{2.20}$$

where $\bar{p}_i^{\mathbb{I}}$ is the national average price for item i, $\bar{q}_i^{\mathbb{I}}$ is the national average expenditure for item i, and

$$\omega_i = \frac{\bar{p}_i^{\mathbb{I}} \bar{q}_i^{\mathbb{I}}}{\sum_{j \in \mathbb{I}} \bar{p}_j^{\mathbb{I}} \bar{q}_j^{\mathbb{I}}} \tag{2.21}$$

As noted by Koo (2000)[24], this expression simplifies to an index similar to a Laspeyres price index, with the national average prices and quantities as the "comparison area":

$$COLI_m = \frac{\sum_{i \in \mathbb{I}} p_{im}^{\mathbb{IM}} \bar{q}_i^{\mathbb{I}}}{\sum_{i \in \mathbb{I}} \bar{p}_i^{\mathbb{I}} \bar{q}^{\mathbb{I}} i} \tag{2.22}$$

CCER also publishes county level indices based on an econometric model using these $|\mathbb{M}|$ MSA level indices. CCER characterizes the model that they employ as follows: "By utilizing ordinary least squares regression analysis, we tested various combinations of the independent variables to identify the best model for use with the data for the 300 areas around the country for which Cost of Living Index data exist. To allow for nonlinear relationships, we also tested squared versions of appropriate independent variables in the model. The variables used in the model are population, population density, income per capita, growth rates for both population and income per capita, government cost, unemployment rate, and C2ER defined regions. Criteria for inclusion of a variable included statistical significance (typically at the 5% level or better), intercorrelation with other variables, impact on the adjusted coefficient of determination (R2), and economic logic." [7] Thus values for the county level COLI are inferred from the MSA level COLI using county level data from the BEA and the Census, rather than directly measured.

Given the strong focus on collecting price data relevant to upper income professionals living in urban environments, these estimates are more useful in characterizing the cost of living in urban counties near large metropolitan areas than they are for smaller or more rural places. Even in this context, there is a good amount of subjectivity involved in these estimates due to the autonomy enjoyed by CCER's volunteers.

2.3 Price Indices based on Scanner Data

The indices described in Sections 2.2.2 and 2.2.1 have limited data coverage within the target areas. For example, the expense of monitoring prices across a large number of retail outlets motivates the CPI to sample stores exclusively within urban areas, and subsequently to select a single product variety to represent the price of an ELI within each sampled store. This necessitates the estimation of the price level for each ELI over relatively large areas, which makes it difficult for the BEA to produce RPPs for areas with low population density, or at low levels of spatial aggregation. Similar cost constraints force CCER to base their COLI indices on price quotations gathered by volunteers, and there is minimal control over product selections.

To gain more detailed data than is available through traditional sampling approaches, researchers and some statistical agencies have begun to experiment with the estimation of price indices from data generated by barcode scanners or other point of purchase devices.[14] Feenstra and Shapiro (2003)[18] includes early research papers exploring the potential impact of scanner data on economic measurement, as well as some of the practical and theoretical challenges associated with their use. Ehrlich et. al. (2019)[15] provide a more recent summary of the state of this research, including a comparison of the indices we discuss in Section 2.3.2 to alternative approaches, such as large scale hedonic adjustment. J De Haan et al (2016)[10] notes that between the publication of those two papers, at least six European countries have begun using scanner data in the estimation of their consumer price indices, and that more seem likely to follow. Outside of Europe, the Australian Bureau of Statistics has been using scanner data from grocery chains as a component of their CPI since 2014 [36][30]. Though the United States has not yet adopted scanner data as an official component of its CPI, economists at the Bureau of Labor Statistics have published papers comparing inflation indices produced using Nielsen's Retail Scanner Database to the government CPI estimates within various regions of the US, with the assumption that the indices based on scanner data are likely the more accurate of the two.[20] These data have several advantages over data generated through more traditional means.

First, because scanner data are automatically generated during routine economic activity, the cost of collecting data from such databases can be dramatically lower than the costs associated with manually collecting price data through repeated survey samples, phone calls, and/or direct visits to large numbers of stores. The marginal cost of collecting an additional price quotation from these databases is so low, in fact, that retail scanner databases are often able to provide the census of all transactions that took place within a store.[18][1] Further, bar code scanners can also generate data at a much higher frequency than would be feasible using any other sampling method, because they record each transaction at the exact time that it occurs.[18][1] Thus retail scanner databases can easily contain data on the number of units sold of every available product within each individual week. As a result, there is almost no lag between the introduction of new products to the market and the inclusion of their price and expenditure measurements in scanner databases. This is a significant advantage over data produced with traditional sampling methods, which often sample within a basket of goods chosen based on consumer expenditure data that is years out of date.

Further, scanner data generally provide simultaneous measurements of the price and number of units sold for each product variety within the stores. This level of detail enables the estimation of cost of living indices that more fully account for the impact of substitution effects, i.e. consumers' ability to substitute expensive items for less expensive ones, on their cost of living.[18][1] This enables price indices to account for the impact of changes in the available products in more sophisticated ways.[15]

Ehrlich et. al. (2019)[15] cite two broad approaches that accomplish this task. The first approach, exemplified by Bajari and Benkard (2005)[5] uses detailed transaction data to improve the estimation of hedonic regression models. The second framework is exemplified by Feenstra (1994)[17], which used a constant elasticity of substitution (CES) utility model to account for the welfare impact of changing product availability across time. Redding and Weinstein (2019)'s[35] "Unified Price Index" generalizes Feenstra (1994)'s[17] indices to account for changes in consumer preferences across time.

Both approaches can be used with scanner data to quantify the impact of changes in product availability on consumer welfare. It is likely that both approaches could also be applied in a spatial context. We base our approach to spatial indices upon the second framework for two main reasons. First, considerable time would be required to specify separate hedonic regressions for hundreds of different products. Second, our indices are based on the Nielsen retail scanner data, which do not always contain enough information about product attributes for hedonic approaches to be effective. For these reasons, the remainder of this review will focus on indices derived from the CES model, rather than these alternative approaches. We discuss Feenstra indices in detail in Section 2.3.1, and Unified Price Indices in Section 2.3.2.

2.3.1 Feenstra Price Indices

The kinds of products carried by stores change over time. For example, consumer electronics stores in the 1930's did not sell iPods or Thinkpad laptops, and new rotary phones and punch cards are not sold at the local electronics store in 2019. This fact complicates the comparison of the price level of consumer electronics in the 1930's to the price level in 2019. Differences in product mix across time make it impossible to directly compare the price of product varieties between these two time periods.

Perhaps the simplest response to this difficulty is to calculate an index based on a basket of goods that are available in both time periods, ignoring any newly introduced or recently discontinued products. Henceforth we refer to an index that does this as a "common goods price index" (CGPI). Using common goods indices as approximations to the "true" cost of living index can make sense in contexts where the set of available products to compare doesn't change very much across time. However, this kind of index ignores changes in consumption patterns over time, and thus introduces biases of increasing magnitude into our price index as the market composition of period t_2 diverges from that of period t_1. For example, it would be misleading to compare the cost of telephones between 1920 and 2019 based entirely on the price of rotary phones, as a common goods index based in 1920 might necessitate. Different kinds of telephone have been invented since that time, and a common goods price index cannot reflect their prices or qualities.

Feenstra (1994)[17] develops an approach to correct the biases introduced by product turnover based on the assumption that consumers behave in accordance with a constant elasticity of substitution (CES) utility model. This model posits the existence of "taste" parameters b_{ut} that measure the utility that a consumer will derive from consuming a unit of product variety u. Consuming one unit of a product associated with a high b_{ut} value contributes more towards consumer utility than consuming one unit of a product with a lower value. If some product is prohibitively expensive or otherwise becomes unavailable, the consumer can substitute different product varieties for the unavailable ones. The consumer's propensity to do so is modeled by the "elasticity of substitution," an unknown parameter that is assumed to be constant within a set of substitutable product varieties. This framework makes it possible to formally model the impact of the creation and destruction of goods across time on consumer welfare.

Specifically, assume that during time period t there are $|\Omega_t|$ product varieties included in the set of product varieties Ω_t. Each product variety $u \in \Omega_t$ has an associated price p_{ut} and quantity consumed q_{ut} in each time period. We stack these product prices and quantities over u, and obtain the vectors

$$\vec{p}_t = \begin{bmatrix} p_{1t} \\ \vdots \\ p_{|\Omega_t|t} \end{bmatrix} \qquad \text{and} \qquad \vec{q}_t = \begin{bmatrix} q_{1a} \\ \vdots \\ q_{|\Omega_t|t} \end{bmatrix} \tag{2.23}$$

According to the CES model, the utility $U(\vec{q}_t)$ that a consumer derives from a given consumption bundle $\vec{q}_t$ is modeled as

$$U(\vec{q}_t) = \Big(\sum_{u\in\Omega_t} b_{ut} q_{ut}^{\rho}\Big)^{1/\rho} \tag{2.24}$$

where ρ is an unknown positive and real valued parameter.

Based on this model, for a fixed vector of prices $\vec{p}_t$, Varian (1978)[37] describes how to derive the associated unit expenditure function:

$$C(\vec{p}_t, \vec{q}_t) = \Big(\sum_{u\in\Omega_t} d_{ut} p_{ut}^{1-\sigma}\Big)^{\frac{1}{1-\sigma}} \tag{2.25}$$

In these expressions σ is the "elasticity of substitution," that expresses the extent to which a consumer is willing to shift his consumption from one type of good to another in response to changes in their relative prices. When $\sigma > 1$, then the consumer sees the good varieties as substitutes. When $\sigma < 1$, the goods are instead considered complements. As $\sigma \to \infty$, consumers become more willing to substitute between different product types, and as $\sigma \to 1$ consumers attach more value to consuming particular product varieties. The parameter ρ in Equation 2.24 is related to the elasticity of substitution σ by the equation

$$\rho = \frac{\sigma - 1}{\sigma} \tag{2.26}$$

Similarly, d_{ut} is a parameter monotonically related to b_{ut} by the equation

$$d_{ut} = b_{ut}^{\sigma} \tag{2.27}$$

This value can also be interpreted as expressing the degree of consumer preference for product u.

As discussed in Section 2.1, an exact price index for the CES utility function comparing time period t_1 to time period t_2 can be constructed by taking the ratio of the unit expenditure functions from Equation 2.25:

$$EPI_{t_1 t_2} \coloneqq \frac{C(\vec{p}_{t_2}, \vec{q}_{t_2})}{C(\vec{p}_{t_1}, \vec{q}_{t_1})} \tag{2.28}$$

Define the set $\Omega_{t_1 t_2}$ as

$$\Omega_{t_1 t_2} \coloneqq \Omega_{t_1} \cap \Omega_{t_2} \tag{2.29}$$

If we assume that consumer taste parameters d_{ut} remain constant across the time periods being compared so that

$$d_{ut_1} = d_{ut_2}, \ \forall u \in \Omega_{t_1 t_2} \tag{2.30}$$

and that the market baskets in t_1 and t_2 are identical, so that

$$\Omega_{t_1 t_2} = \Omega_{t_1} = \Omega_{t_2} \tag{2.31}$$

then we can construct this index as follows.

First, define the expenditure shares s_{ut} of product u in time period t as

$$s_{ut} = \frac{p_{ut}q_{ut}}{\sum_{v\in\Omega_t} p_{vt}q_{vt}} \tag{2.32}$$

and weights $\omega_{ut_1t_2}$ as

$$\omega_{ut_1t_2} = \frac{\left(\frac{s_{ut_2}-s_{ut_1}}{log(s_{ut_2})-log(s_{ut_1})}\right)}{\sum_{v\in\Omega_{t_1t_2}}\left(\frac{s_{vt_2}-s_{vt_1}}{log(s_{vt_2})-log(s_{vt_1})}\right)} \tag{2.33}$$

Then the common goods price index that is exact for the CES utility function, known as the "Sato-Vartia price index," can be written as[17]

$$SV_{t_1t_2} = \prod_{v\in\Omega_{t_1t_2}} \left(\frac{p_{vt_2}}{p_{vt_1}}\right)^{\omega_{vt_1t_2}} \tag{2.34}$$

Feenstra (1994)[17] shows that the Sato-Vartia index in Equation 2.34 can be generalized to produce an exact cost of living index when different (but overlapping) sets of goods are available in each time period, i.e. in situations where $\Omega_{t_1} \neq \Omega_{t_2}$ but $|\Omega_{t_1t_2}| > 0$. Further, assuming that consumer preferences remain constant over time, the value of this index can be estimated without knowing the values of the taste parameters d_{ut}. [17]

More specifically, let λ_t be the share of expenditure at time t on goods that are available in both times t_1 and t_2, i.e.

$$\lambda_t = \frac{\sum_{v\in\Omega_{t_1t_2}} p_{vt}q_{vt}}{\sum_{v\in\Omega_t} p_{vt}q_{vt}} \tag{2.35}$$

for $t \in \{t_1, t_2\}$, and define the "variety adjustment" term as

$$VADJ_{t_1t_2} = \left(\frac{\lambda_{t_2}}{\lambda_{t_1}}\right)^{\frac{1}{1-\sigma}} \tag{2.36}$$

for a given value of the elasticity of substitution σ. Then Feenstra (1994)'s[17] index is:

$$F_{t_1t_2} = SV_{t_1t_2} \times VADJ_{t_1t_2} \tag{2.37}$$

The variety adjustment term in Equation 2.37 corrects the bias in the Sato-Vartia index caused by the differences between the baskets of goods available in periods t_1 and t_2. Intuitively, the λ_t terms capture the extent to which consumers value product varieties that are exclusively available in period t relative to varieties available in both comparison periods. If consumers value goods that are available in both periods highly relative to goods that are only available in time t, then consumer expenditure on these common goods will be a high proportion of the total expenditure, and hence λ_t will be high. If the reverse is true, then λ_t will be low. If λ_{t_2} is smaller than λ_{t_1}, consumers are better off in period t_2 than they were in period t_1, and the cost of living is lower than the Sato-Vartia index would suggest; if the reverse is true, then the cost of living will be higher than the Sato-Vartia index suggests.

A consequence of this is that the CES utility model implies that when a product variety is discontinued, it becomes more expensive for a consumer to gain a set level of utility. Conversely, the introduction of new varieties of product makes it less expensive for the consumer to maintain a set standard of living. One way of interpreting this behavior is that an unavailable product variety effectively has a price of ∞. When a new product variety is introduced to the market, its price "declines" from ∞ to a finite number, thereby decreasing the cost of living. Similarly, when a product variety is withdrawn from the market, its price "increases" to ∞. Less abstractly, we can interpret this property as arising from a preference for variety on the part of consumers.

This framework is well-suited to estimating price indices from scanner data. The Feenstra index allows us to avoid directly estimating the quality parameters b_u, as long as we have simultaneous measures of the price and expenditure for each product variety. Scanner data enables the simultaneous measurement of prices and expenditures for a large number of goods categorized by unique bar codes, within which products can be assumed to have essentially identical characteristics. Thus the product varieties in Equation 2.24 can be straightforwardly identified with individual bar codes (or UPC's) in the scanner data.

2.3.2 Unified Price Indices

Redding and Weinstein (2019)[34][35] employ the same constant elasticity of substitution (CES) utility framework as Feenstra (1994)[17] to generalize several common price index formulas, showing that they can all be considered special cases of a single *Unified Price Index* (UPI). In particular, Redding and Weinstein (2019)[34][35]'s index can be considered a generalization of the Feenstra index from Section 2.3.1 to cases in which consumers have different preferences in each of the periods being compared. Under this assumption, we risk biasing our comparison because consumer preferences change over time.

Assuming that consumers have time-varying preferences is intuitively plausible in many situations. For example, a consumer in 1930 might value a rotary telephone highly, seeing it as a vast improvement over alternatives such as the telegraph or hand-written communication. By contrast, a consumer in 2010 who recieved the same phone might be underwhelmed due to the absence of now-common features such as portability, text messaging, and mobile internet browsing. In this hypothetical, the utility the representative consumer would obtain from purchasing one rotary telephone has declined dramatically from 1930 to 2010. Ignoring this fact would introduce what Redding and Weinstein (2019)[34][35] refers to as a “consumer valuation bias” into estimates of the cost of living index. In our phone example, the “consumer valuation bias” is due to the fact that the cost of obtaining a unit of utility from consuming a rotary phone in 2010 is higher than the cost of obtaining a unit of utility from consuming a rotary phone in 1930. Ignoring this fact leads us to estimate incorrect unit expenditure functions, and thus incorrect price indices.

Formally, assume that we have data on product prices and quantities sold within two time periods, t_1 and t_2. Define Ω_t as the set of product varieties that are available in period t, and $\Omega_{t_1 t_2}$ as the set of products available in both periods t_1 and t_2 (i.e., $\Omega_{t_1 t_2} = \Omega_{t_1} \cap \Omega_{t_2}$). At each time period $t \in \{t_1, t_2\}$ we have data on the price p_{ut} and quantity consumed q_{ut} of some number of product varieties $u \in \Omega_t$. We stack these product prices and quantities over u, into the vectors

$$\vec{p_t} = \begin{bmatrix} p_{1t} \\ \vdots \\ p_{|\Omega_t|t} \end{bmatrix} \qquad \text{and} \qquad \vec{q_t} = \begin{bmatrix} q_{1t} \\ \vdots \\ q_{|\Omega_t|t} \end{bmatrix} \tag{2.38}$$

for both periods t_1 and t_2. We want to use this information to produce an index measuring the differences in price level between the two time periods.

The utility $U(\vec{q_t})$ that a consumer derives from a given consumption bundle $\vec{q_t}$ is modeled as

$$U(\vec{q_t}) = \Big(\sum_{u\Omega_t}^{U} b_{ut} q_{ut}^{\rho} \Big)^{1/\rho} \tag{2.39}$$

where ρ is positive and real valued. As discussed in Section 2.3.1, this utility function implies that the consumer's unit expenditure function is

$$C(\vec{p_t}, \vec{q_t}) = \Big(\sum_{u \in \Omega_t} d_{ut} p_{ut}^{1-\sigma} \Big)^{\frac{1}{1-\sigma}} \tag{2.40}$$

where

$$d_{ut} = b_{ut}^{\sigma}, \ \forall u \in \Omega_t \text{ and } t \in \{t_1, t_2\} \tag{2.41}$$

for some constant elasticity of substitution parameter σ.

The major difference between the utility model used by Feenstra (1994)[17] and this framework is the assumption that

$$d_{ut_1} = d_{ut_2} \tag{2.42}$$

Both approaches assume that consumers maximize a CES utility function, and that the available goods can change across time. However Redding and Weinstein (2019)[34][35] *do not* assume that Equation 2.42 is true. Consumer preferences need not remain constant across the time periods being compared.

This is why the Unified Price Index can be considered a generalization of the Feenstra index described in Section 2.3.1.

Accordingly, the Unified Price Index has a similar structure to the Feenstra index. In particular, define λ_t as the share of total expenditure at t on all items that are sold in both periods t_1 and t_2 so that

$$\lambda_t := \frac{\sum_{v \in \Omega_{t_1 t_2}} p_{vt} q_{vt}}{\sum_{v \in \Omega_t} p_{vt} q_{vt}} \tag{2.43}$$

and $r_{ut}^{\Omega_{t_1 t_2}}$ as the share of expenditure on goods common to both periods accounted for by product variety u in period t:

$$r_{ut}^{\Omega_{t_1 t_2}} := \frac{p_{ut} q_{ut}}{\sum_{v \in \Omega_{t_1 t_2}} p_{vt} q_{vt}} \tag{2.44}$$

Then the Unified Price Index can be factored[34][35] into a “Common Goods Price Index” (CGPI) that depends only on products sold in both time periods, which we calculate as

$$CGPI_{t_1t_2}(\vec{p}_{t_1}, \vec{p}_{t_2}, \vec{q}_{t_1}, \vec{q}_{t_2}, \sigma) = \prod_{v \in \Omega_{t_1t_2}} \left(\frac{p_{vt_2}}{p_{vt_1}}\right)^{\frac{1}{|\Omega_{t_1t_2}|}} \left[\prod_{v \in \Omega_{t_1t_2}} \left(\frac{r_{vt_2}^{\Omega_{t_1t_2}}}{r_{vt_1}^{\Omega_{t_1t_2}}}\right)^{\frac{1}{|\Omega_{t_1t_2}|}}\right]^{\frac{1}{\sigma-1}} \tag{2.45}$$

and a "Variety Adjustment" term that captures the impact of the creation and destruction of products over time, which is calculated as

$$VADJ_{t_1t_2}(\vec{p}_{t_1}, \vec{p}_{t_2}, \vec{q}_{t_1}, \vec{q}_{t_2}, \sigma) = \left(\frac{\lambda_{t_2}}{\lambda_{t_1}}\right)^{\frac{1}{\sigma-1}} \tag{2.46}$$

The $CGPI$ term in Equation 2.45 can be further factored into the geometric mean of the relative prices of each good, commonly known as the Jevons index

$$Jevons_{t_1t_2}(\vec{p}_{t_1}, \vec{p}_{t_2}) = \prod_{u \in \Omega_{t_1t_2}} \left(\frac{p_{ut_2}}{p_{ut_1}}\right)^{\frac{1}{|\Omega_{t_1t_2}|}} \tag{2.47}$$

and the "Spread Adjustment" (SADJ) term

$$SADJ_{t_1t_2}(\vec{p}_{t_1}, \vec{p}_{t_2}, \vec{q}_{t_1}, \vec{q}_{t_2}, \sigma) = \left[\prod_{v \in \Omega_{t_1t_2}} \left(\frac{r_{vt_2}^{\Omega_{t_1t_2}}}{r_{vt_1}^{\Omega_{t_1t_2}}}\right)^{\frac{1}{|\Omega_{t_1t_2}|}}\right]^{\frac{1}{\sigma-1}} \tag{2.48}$$

Thus we can write the Unified Price Index as

$$UPI_{t_1t_2}(\vec{p_{t_1}}, \vec{p_{t_2}}, \vec{q_{t_1}}, \vec{q_{t_2}}, \sigma) = Jevons_{t_1t_2} \times SADJ_{t_1t_2} \times VADJ_{t_1t_2} \tag{2.49}$$

Each of the factors in Equation 2.49 has an economically meaningful interpretation. The Jevons index in Equation 2.47 measures the average difference in prices between product varieties that are sold in both period t_1 and t_2. This is the only term in the UPI whose value is unaffected by the elasticity of substitution σ, since it aims to capture only the impact of the raw price differences between t_1 and t_2 on consumer welfare.

The SADJ term in Equation 2.48 accounts for changes in consumer preferences across time. It measures the impact of differences in the "spread" of the quality-adjusted prices $\frac{p_{ut}}{d_{ut}}$ on the cost of living in each time period $t \in \{t_1, t_2\}$. To see this, note that for each $t \in \{t_1, t_2\}$ the geometric mean

$$\prod_{v \in \Omega_{t_1 t_2}} (r_{vt}^{\Omega_{t_1 t_2}})^{\frac{1}{|\Omega_{t_1 t_2}|}} \tag{2.50}$$

is maximized when

$$r_{ut}^{\Omega_{t_1 t_2}} = \frac{1}{|\Omega_{t_1 t_2}|}, \forall u \in \Omega_{t_1 t_2} \tag{2.51}$$

i.e. when each common product variety has equal market share. This implies that the value of $SADJ$ will be larger, and hence the cost of living will be higher, when the market shares in t_2 are more evenly dispersed than those in t_1. The intuition behind this result is that the quality-adjusted prices of each product variety are reflected in its market share. If the quality-adjusted prices of product varieties in period t are similar, then the market shares of the varieties should also be similar. Consumers benefit from having access to a thick market, in which they can choose to substitute consumption towards less expensive product varieties and away from more expensive ones.

Finally, the VADJ term in Equation 2.46 accounts for changes in the set of product varieties available between periods t_1 and t_2. Mathematically, this term is identical to the $VADJ$ term associated with the Feenstra index from in Section 2.3.1, and its interpretation does not change.

Because the elasticity of substitution σ appears in both the CGPI and VADJ terms, the degree to which consumers are able to substitute between product varieties has a powerful impact on how the UPI responds to the creation and destruction of product varieties, or to changes in consumer preferences over time. The degree to which the $SADJ$ and $VADJ$ terms matter depends on the value of the elasticity of substitution parameter σ in an intuitive way. The impact of the $SADJ$ term is mediated by the elasticity of substitution σ, because only consumers that substitute one product variety for another are able to shift their spending towards lower cost options. If consumers don't see the product varieties in Ω_t as substitutes, i.e. if $\sigma \leq 1$, then they cannot benefit from substituting one variety for another. The impact of the $VADJ$ term is related to σ for a similar reason. If consumers are more willing to substitute one product variety for another, the gains or losses in welfare due to the creation or destruction of product varieties across time should be mitigated.

In both cases, as $\sigma \to \infty$, consumers consider the available products to be closer and closer substitutes for the missing ones. At the limit, neither adjustment term has an impact on the cost of living because consumers consider the products to be perfect substitutes. In this circumstance, differences between the product varieties lose all meaning, and only differences in the relative price matter. Thus the UPI reduces to the Jevons index, which measures those differences exclusively. On the other hand, as $\sigma \to 1$, the consumer instead considers the available products to be poorer and poorer substitutes for the missing ones, and consequently the impact of the adjustment terms increases.

The framework discussed in this section is particularly important for our project, since the Unified Price Index put forward by Redding and Weinstein (2019) [34][35] forms the foundation of the index we will use to compare the cost of living spatially. We discuss how our approach relates to the Unified Price Index in more detail in Chapter III.

CHAPTER III

Spatial Unified Price Indices

In this section, we develop the price index formula that we will use to compare the cost of living between different spatial locations. We construct our cost of living indices as ratios of the minimum expenditure required to obtain one unit of utility in each of the comparison areas, as discussed in Section 2.1. This approach requires us to model consumer utility, and compare the outcomes associated with this model across the set of comparison areas. Because we plan to base these comparisons on retail scanner data of the type discussed in Section 2.3, we encounter issues that are analogous to the issues faced by researchers attempting to construct scanner data based inflation indices. In particular, some of the biases inherent to traditional common good price indices (CGPIs) used for spatial comparisons are conceptually similar to the biases in CGPIs for inflation.

For example, the kinds of products carried in stores can vary depending on location. Thirsty shoppers might expect to find Cactus Cooler in convenience shops in California, but not in the corner stores of Michigan. Similarly, consumers are more likely to find Faygo soda routinely stocked in Detroit than it would be in Los Angeles. This fact presents an obstacle to comparing the cost of living differences due to carbonated beverages between Los Angeles County and Wayne County. Differences in products sold across these counties make it impossible to directly compare the price of product varieties between them. This is analogous to the problem of product turnover across time discussed by Feenstra (1994)[17] and Redding and Weinstein (2016)[34], in that it prevents us from directly comparing products across regions. From the standpoint of a consumer, a regional product such as Faygo soda, which is widely available in the Detroit metropolitan area but not elsewhere is "discontinued" when a consumer moves from Michigan to California, and is "introduced to the market" when the reverse occurs. Thus to the extent that consumers have strong preferences for them, the local availability (or lack thereof) of favored product varieties can affect consumer well-being in analogous ways to the creation and destruction of product varieties across time.

Further, consumers in different locations have different consumption preferences. In fact, the differences in market composition from place to place are often directly related to this fact. For example, consumers in Michigan may place a higher value on all varieties of warm winter coats, seeing them as a necessity of life due to the state's cold winters. In contrast, consumers in California might view warm coats as more of a luxury item, useful for skiing trips to the mountains or trips out of state, but not a day-to-day necessity. Thus consumer valuation of warm coats over other clothing goods can be expected to differ from place to place. Retail outlets respond to such differences in consumer preferences by changing the kinds of goods that they carry. In our example, clothing stores in Michigan might find it profitable to stock a broader variety of warm winter coats than similar outlets in California. Thus any effort to compare the cost of living between California and Michigan needs to consider the possibility that consumer preferences vary between the two areas, lest they risk introducing a "consumer valuation bias" analogous to the one identified by Redding and Weinstein (2016)[34] discussed in Section 2.3.2.

In order to address these obstacles, we propose an index based on Redding and Weinstein (2016)[34]'s theoretical "Unified Price Index" (UPI) to measure differences in the cost of living between different spatial areas. We elaborate on the construction and theoretical commitments of this index more formally in Sections 3.1, 3.2, and 3.3 below.

3.1 Bilateral Spatial Comparisons with the UPI

In this section, we describe more concretely the process by which we apply the UPI as developed by Redding and Weinstein (2016)[34] to compare spatial cost of living between two areas. By treating spatial heterogeneity in available product varieties as analogous to temporal heterogeneity, we can adapt [34]'s Unified Price Index to a spatial context.

Suppose that we have data on product prices and quantities sold within areas a_1 and a_2 Define Ω_{a_1} as the set of product varieties that are available in area a_1, and $\Omega_{a_1 a_2}$ as the set of product varieties that are available in both a_1 and a_2 (i.e., $\Omega_{a_1 a_2} = \Omega_{a_1} \cap \Omega_{a_2}$). In each area $a \in \mathbb{A}$, we have data on the price p_{ua} and quantity consumed q_{ua} of some number of product varieties $u \in \Omega_a$. For each comparison area a, we stack these product prices and quantities over u, into the vectors

$$\vec{p}_a = \begin{bmatrix} p_{1a} \\ \vdots \\ p_{|\Omega_a|a} \end{bmatrix} \qquad \text{and} \qquad \vec{q}_a = \begin{bmatrix} q_{1a} \\ \vdots \\ q_{|\Omega_a|a} \end{bmatrix} \tag{3.1}$$

for areas a_1 and a_2. Then we can calculate λ_a, the share of total expenditure in area a on all items that are sold in both areas a_1 and a_2, as

$$\lambda_a := \frac{\sum_{v \in \Omega_{a_1 a_2}} p_{va} q_{va}}{\sum_{v \in \Omega_a} p_{va} q_{va}} \tag{3.2}$$

and $r_{ua}^{\Omega_{a_1 a_2}}$, the share of expenditure on goods common to both areas accounted for by product variety u in area a, as

$$r_{ua}^{\Omega_{a_1 a_2}} := \frac{p_{ua} q_{ua}}{\sum_{v \in \Omega_{a_1 a_2}} p_{va} q_{va}} \tag{3.3}$$

for comparison areas a_1 and a_2. We want to use this information to produce spatial indices measuring the differences in price level between areas a_1 and a_2.

We can then compute the UPI between areas a_1, a_2 as the following product:

$$UPI_{a_1 a_2} = Jevons_{a_1 a_2} \times SADJ_{a_1 a_2} \times VADJ_{a_1 a_2} \tag{3.4}$$

where

$$Jevons_{a_1 a_2} = \prod_{u \in \Omega_{a_1 a_2}} \left(\frac{p_{ua_2}}{p_{ua_1}} \right)^{\frac{1}{|\Omega_{a_1 a_2}|}} \tag{3.5}$$

is the spatial analogue to the Jevons index described in Section 2.3.2,

$$SADJ_{a_1 a_2} = \left[\prod_{v \in \Omega_{a_1 a_2}} \left(\frac{r_{va_2}^{\Omega_{a_1 a_2}}}{r_{va_1}^{\Omega_{a_1 a_2}}} \right)^{\frac{1}{|\Omega_{a_1 a_2}|}} \right]^{\frac{1}{\sigma - 1}} \tag{3.6}$$

is the spatial "Spread Adjustment" (SADJ), and

$$VADJ_{a_1 a_2} = \left(\frac{\lambda_{a_2}}{\lambda_{a_1}} \right)^{\frac{1}{\sigma - 1}} \tag{3.7}$$

is the spatial "Variety Adjustment" (VADJ).

The interpretation of these terms is analogous to the interpretation of Equations 2.47, 2.48 and 2.46, but with the differences compared across spatial, rather than temporal, units. In the spatial context, the Jevons index in Equation 3.5 captures the average difference in price between goods sold in both areas a_1 and a_2. The SADJ term in Equation 3.6 compares the "spread" of the quality-adjusted prices in each area, in order to measure the degree to which consumers in each area could improve their quality of life by substituting expensive (in quality-adjusted terms) product varieties for relatively inexpensive ones. Finally, the VADJ term in Equation 3.7 measures the extent to which differences in the product varieties available in each location impact consumer satisfaction.

Note that this procedure can fail to estimate valid UPIs even when there are data in an area in the (rare) circumstance that the product varieties sold in some area are entirely disjoint from the product varieties sold in all other areas. We ignore the data from such an area, since it cannot be compared to the data from any other area without collapsing multiple UPCs into broader categories in a way that could introduce undesirable quality differences within product varieties. Intuitively, if Wayne-brand apples are the *only* apples sold in Wayne county, and Wayne-brand apples are sold *nowhere else*, then there's no way to infer the quality of Wayne-brand apples compared to other apples by observing differences in consumer expenditure on Wayne vs. non-Wayne apples.

3.2 Spatial UPI

In Section 3.1, we outlined how to apply Redding and Weinstein (2016)[34]'s UPI to scanner data in order to compare spatial, rather than intertemporal, differences in the cost of living between two areas. In practice however, we are usually interested in comparing cost of living across more than two areas. Thus we wish to extend the index described in Section 3.1, in order to convert it from a from a bilateral to a multilateral index number.

Two difficulties dissuade us from directly applying the UPI from Section 3.1 as an index for multilateral spatial comparisons. First, depending on the levels of spatial and temporal aggregation chosen, the heterogeneity between products sold in different areas within the United States can be greater than the heterogeneity between products sold at different time periods within the United States. The UPI requires at least one product variety to be sold in both areas in order to form the $CGPI$. But some areas may sell completely disjoint sets of products of a particular type, so that there is no overlap in the varieties of good sold. In these circumstances the UPI comparing the two areas cannot be computed.

Second, rather than comparing two time points with an obvious ordinal arrangement, our spatial UPI seeks to compare several unordered areas to each other at once. Formally, if $\mathbb{A} = \{1, \ldots, A\}$ is the set of areas to compare, we want to compare the cost of living in every pair of areas $(a_1, a_2) \in \mathbb{A} \times \mathbb{A}$. We would like to use the matrix of all possible pairwise UPI comparisons

$$\mathbf{UPI} = \begin{bmatrix} UPI_{11} & \ldots & UPI_{1A} \\ \vdots & \ddots & \vdots \\ UPI_{A1} & \ldots & UPI_{AA} \end{bmatrix} \tag{3.8}$$

to make this comparison. Unfortunately, the UPI as calculated using Equation 3.4 is an intransitive index. This means that the system of comparisons based on M can produce inconsistent results when price levels in two areas are compared through an intermediary. For example, a researcher who wants to compare consumer purchasing power across states might encounter the paradoxical implication that a person who moved from Michigan to Ohio would have had more buying power if they had first moved from Michigan to Indiana, and then from Indiana to Ohio. But the order of comparison does not affect the structure of the economy, and hence it should not affect our indices. Transitivity is necessary to ensure consistency between the set of comparisons in $\mathbb{U}$, and avoid these kinds of contradictions.

Formally, an index $I_{a_1a_2}$ comparing base area a_1 to comparison area a_2 is transitive if $I_{a_1a_2} \times I_{a_2a_3} = I_{a_1a_3}, \forall a_1, a_2, a_3 \in \mathbb{A}$ (See [22]). We impose transitivity by employing the Gini-Éltetö-Köves-Szulc (GEKS) method explicated in [33]. The idea behind this is to find the transitive indices

$$\mathbf{SUPI} = \begin{bmatrix} SUPI_{11} & \dots & SUPI_{1A} \\ \vdots & \ddots & \vdots \\ SUPI_{A1} & \dots & SUPI_{AA} \end{bmatrix} \tag{3.9}$$

that are "closest" to the intransitive indices in M according to a least squares loss function. The elements of the resulting S matrix will be our spatial UPI (SUPI) estimates.

Formally, when applying GEKS we are solving the following minimization problem:

$$\begin{aligned} \underset{S}{\text{argmin}} \quad & \sum_{a_1} \sum_{a_2} \left(ln(UPI_{a_1 a_2}) - ln(SUPI_{a_1 a_2}) \right)^2 \\ \text{subject to} \quad & SUPI_{a_1 a_2} \times SUPI_{a_2 a_3} = SUPI_{a_1 a_3}, \forall a_1, a_2, a_3 \in \mathbb{A}. \end{aligned} \tag{3.10}$$

This constraint is simplified by the result, described in [33], that any collection of index numbers $\{I_{a_1 a_2}\}_{a_1, a_2 \in \mathbb{A}}$ is transitive if and only if there exist some numbers $\vec{\pi} = \begin{bmatrix} \pi_1 & \dots & \pi_A \end{bmatrix}$ such that $I_{a_1 a_2} = \frac{\exp \pi_{a_2}}{\exp \pi_{a_1}}, \forall a_1, a_2 \in \mathbb{A}$. If we exploit this relationship, the number of parameters is reduced from A^2 to A, and the problem simplifies to:

$$\underset{\vec{\Pi}}{\text{argmin}} \quad \sum_{a_1} \sum_{a_2} \left(ln(UPI_{a_1 a_2}) - (\pi_{a_2} - \pi_{a_1}) \right)^2 \tag{3.11}$$

For the purposes of identification, one of these π values is set to zero in advance. The area associated with this π value becomes the reference area. The resulting objective function for our least squares minimization problem is convex, and as such can be solved for $\vec{\Pi}$ relatively easily. We can then calculate the value of each of index in \$ as

$$SUPI_{a_1 a_2} = \frac{exp(\pi_{a_2})}{exp(\pi_{a_1})} \tag{3.12}$$

Defining our spatial UPIs in this way has two major benefits. First, the index values still reflect the quality and variety adjustments provided by [34]'s UPI, but they also yield internally consistent sets of spatial comparisons. Second, we can infer spatial UPI values for areas whose product lists do not overlap, by treating the input comparisons $UPI_{a_1a_2}$ as "data" for estimating $\vec{\pi}$, and the incomparable areas as "missing data" in the $\mathbb{U}$ matrix. Once we obtain these $\vec{\pi}$ estimates, we can apply Equation 3.12 to estimate $SUPI_{a_1a_2}$ indices comparing areas that would otherwise be incomparable.

The main drawback of this approach is the potential for a loss of so-called "area characteristicity," described in Kravis (1982)[27] as the degree to which the items being compared are characteristic of the products sold in the areas being compared. Because the π_a estimates are based on the average relationship between the log UPIs comparing area a to all the other areas in $\mathbb{A}$, some information that is specific to each individual comparison is smoothed out. This tradeoff between transitivity and area characteristicity is inherent to most multilateral indices[27], and is difficult to avoid regardless of the transitive multilateral index formula chosen.

Another issue is the fact that the resulting indices are dependent on the set of areas included in the comparison set. This is a significant drawback when computing inflation indices, as it implies that the price indices for every previous year would require annual revision once the current year's inflation estimate is computed.[23] This problem has a minimal impact in our specific application, because it is generally safe to assume that the set of areas to be compared remains constant within a fixed time period.

3.3 Aggregation across Goods

When estimating the spatial UPI, it is advisable to restrict the class of product varieties included in $\Omega = \cup_{a \in \mathbb{A}} \Omega_a$ to items that are similar, so that consumers might substitute any item for another item in that set. For example, it might make sense to include several different kinds of coffee as product varieties and produce a "coffee" index. But it might not make sense to include tea and coffee in the same set, since consumers may not always see these goods as substitutes.

A better approach to compare prices across a broader set of goods is to combine spatial UPIs for multiple narrowly defined individual goods into a single index. For example, perhaps we have price indices for corn, radishes, and carrots, and we take some function of those indices to get a single index comparing the prices of vegetables. We propose one way of doing this below.

Formally, suppose that we have some set of goods $\mathbb{G}$. Assume that we have computed $|\mathbb{G}|$ matrices $\mathbb{S}_g$ of SUPI values comparing the cost of living associated with product g across all areas in some set $\mathbb{A}$, as described in Section 3.2. Based on these indices, we wish to find a matrix C of "category" price indices $CSUPI_{a_1a_2}$ comparing the average price level of all goods in $\mathbb{G}$ across all areas in the set $\mathbb{A}$:

$$C = \begin{bmatrix} CSUPI_{11} & \dots & CSUPI_{1A} \\ \vdots & \ddots & \vdots \\ CSUPI_{A1} & \dots & CSUPI_{AA} \end{bmatrix} \tag{3.13}$$

To accomplish this, we propose to construct the entries of C as follows:

$$CSUPI_{a_1a_2} = \left(\prod_{g \in \mathbb{G}} SUPI_{ga_1a_2}^{\omega_{ga_1a_2}} \right) \tag{3.14}$$

Here, the weights $\omega_{ga_1a_2}$ represent the degree to which the cost of living of product g ought to impact the category level comparison of areas a_1 and a_2.

We wish to pick weights that reflect the contribution of each product according to its importance to the average consumer, while ensuring that the category level indices $C_{a_1a_2}$ retain the transitivity of the individual indices in the matrices $\mathbb{S}_g$. More formally, we wish to pick a set of economically meaningful weights such that

$$CSUPI_{a_1a_2}CSUPI_{a_2a_3} = CSUPI_{a_1a_3} \tag{3.15}$$

for all choices of areas $a_1, a_2, a_3 \in \mathbb{A}$.

We can see which choices for weights will allow this condition to hold more easily by considering Equation 3.15 on the log scale:

$$log(CSUPI_{a_1a_2}) + log(CSUPI_{a_2a_3}) = log(CSUPI_{a_1a_3})$$

Substituting Equation 3.14 into the right hand side of this expression, we can derive the following:

$$\begin{aligned} log(CSUPI_{a_1a_3}) &= log\left(\prod_{g\in\mathbb{G}} SUPI_{ga_1a_3}^{\omega_{ga_1a_2}}\right) \\ &= \sum_{g\in\mathbb{G}} \omega_{ga_1a_3} log(SUPI_{ga_1a_3}) \\ &= \sum_{g\in\mathbb{G}} \omega_{ga_1a_3}(\pi_{a_3} - \pi_{a_1}) \end{aligned} \tag{3.16}$$

Similarly, we can write the left hand side of this expression as

$$\begin{aligned} log(CSUPI_{a_1a_2}CSUPI_{a_2a_3}) &= log\left(\left(\prod_{g\in\mathbb{G}} SUPI_{ga_1a_2}^{\omega_{ga_1a_2}}\right)\left(\prod_{g\in\mathbb{G}} SUPI_{ga_2a_3}^{\omega_{ga_2a_3}}\right)\right) \\ &= \sum_{g\in\mathbb{G}} \omega_{ga_1a_2} log(SUPI_{ga_1a_2}) + \sum_{g\in\mathbb{G}} \omega_{ga_2a_3} log(SUPI_{ga_2a_3}) \\ &= \sum_{g\in\mathbb{G}} \omega_{ga_1a_2}(\pi_{ga_2} - \pi_{ga_1}) + \sum_{g\in\mathbb{G}} \omega_{ga_2a_3}(\pi_{ga_3} - \pi_{ga_2}) \\ &= \sum_{g\in\mathbb{G}} [\omega_{ga_1a_2}\pi_{ga_2} - \omega_{ga_1a_2}\pi_{pi} + \omega_{ga_2a_3}\pi_{ga_3} - \omega_{ga_2a_3}\pi_{ga_2}] \\ &= \sum_{g\in\mathbb{G}} [(\omega_{ga_1a_2} - \omega_{ga_2a_3})\pi_{ga_2} + \omega_{ga_2a_3}\pi_{ga_3} - \omega_{ga_1a_2}\pi_{ga_1}] \end{aligned} \tag{3.17}$$

Thus the transitivity condition expressed in Equation 3.16 is equivalent to requiring weights such that

$$\sum_{g \in \mathbb{G}} \omega_{ga_1a_3}(\pi_{a_3} - \pi_{a_1}) = \sum_{g \in \mathbb{G}} [(\omega_{ga_1a_2} - \omega_{ga_2a_3})\pi_{ga_2} + \omega_{ga_2a_3}\pi_{ga_3} - \omega_{ga_1a_2}\pi_{ga_1}] \quad (3.18)$$

For this condition to hold, it is both necessary and sufficient to pick weights that satisfy

$$\omega_{ga_1a_2} = \omega_{ga_2a_3} = \omega_{ga_1a_3} \quad (3.19)$$

for all goods $g \in \mathbb{G}$ and all choices of areas $a_1, a_2, a_3 \in \mathbb{A}$. Hence to maintain transitivity, our weights can depend on the category of good, but cannot vary based on the set of areas that are being compared by the index.

For this reason, we propose the following weights. Define the total expenditure on product variety u of good g in area a as e_{gva}, and the set of product varieties in product category g that are sold in area a as $\Omega_{ga}^{\mathbb{GA}}$. Then the category weights are

$$\omega_{ga_1a_2} = \omega_g = \frac{\sum_{a \in \mathbb{A}} \sum_{v \in \Omega_{ga}^{\mathbb{GA}}} e_{gva}}{\sum_{h \in \mathbb{G}} \sum_{a \in \mathbb{A}} \sum_{v \in \Omega_{ha}^{\mathbb{GA}}} e_{hva}} \quad (3.20)$$

The resulting indices weight the areal comparisons of each good $g \in \mathbb{G}$ by their share of the total expenditure across all areas on items in $\mathbb{G}$.

In our vegetable example, this would imply that if 75% of consumer spending on vegetables across all areas is on corn, 15% on radishes, and 10% on carrots, the value of the vegetable index will be much more influenced by the value of the corn index than the value of the other two vegetables. This will be true even if in the particular areas being compared, radishes are much more popular than corn.

CHAPTER IV

Estimation Methodology

In this chapter, we outline the methods we use to estimate the price indices described in Chapter III from Nielsen's Retail Scanner Data. The information available from Nielsen is detailed in Section 4.1, alongside a brief discussion of the advantages and disadvantages these data as the basis for our index estimates relative to the alternatives discussed in Chapter II. Section 4.2, describes how this information can be used to estimate SUPIs for each individual product module in the Nielsen data, and subsequently for arbitrary categories of these products. Finally, in Section 4.3 we discuss the use of a cluster bootstrap method to characterize the uncertainty associated with each of these index estimates. This is helpful for diagnosing which SUPI estimates are based on small numbers of observations, or are highly sensitive to the inclusion or exclusion of particular product varieties within the Nielsen sample. We apply the procedure we detail here to estimate SUPI indices comparing the price of food across all the counties in the state of Michigan, and discuss the results in Chapter V.

4.1 Nielsen's Retail Scanner Data

Nielsen's retail scanner data are gathered from participating retail outlets, whose point-of-sale systems automatically record the price of all transactions. As a result, detailed information about the units sold in each store are available from over 35,000 stores across the continental United States, generally on a weekly basis. The files containing these data are organized into three types: stores files, products files, and movement files.

The stores files detail which retail outlets are included in the Nielsen sample each year. Broadly, these data include information about the parent company, retail channel, and geographic location of each of the stores that price quotes were collected from. This enables us to pinpoint the location in which each set of in-sample transactions took place with remarkable specificity. For our purposes, the most valuable information included in these files are the state and county each store is located in, and the store identifier that enables us to match individual transactions to the particular stores in which they occurred.

The products files contain records of all of the products sold in each year, categorized by their Universal Product Code (UPC). Each UPC uniquely identifies a type of product, within which all economically salient characteristics are assumed to be equivalent. As such, the UPC is the most basic unit for which prices and sales data are recorded. Nielsen organizes the approximately 3.2 million unique UPCs included in the data into about 1,075 "product modules," which represent relatively narrow classes of goods such as cell phones or frozen fish. Thus Nielsen's product modules are roughly analogous to the Entry Level Items (ELIs) used by the Bureau of Labor Statistics, as described in Section 2.2.1. Each product module is identified by a unique code and classified as a member of one of ≈ 125 "product groups." A product group is thus a broader category of items, such as light bulbs and electrical goods, that can include differing numbers of individual product modules. Finally, the product groups are organized into 10 "departments." A department is a high level characterization of the items it contains, such as "mass market merchandise" or "dry goods." The products file includes labels corresponding to each of these classifications for each

UPC, along with information such as the amount of product included in each UPC, and the UPC's associated brand.

The data also include the units in which the quantity of good contained each UPC is measured. About 1.4 million of the UPC's included in these data have additional information about product characteristics available from the Nielsen Consumer Panel Data. We ignore these additional characteristics because they are not required to estimate the SUPIs described in Chapter III.

The movement data constitute the overwhelming majority of Nielsen's retail scanner data by size. These consist of several files, one for each combination of product module code and year. These files contain weekly records of the transactions for UPCs included in the indicated product module and year, recorded at each of the stores in the stores file. Specifically, they include information about which UPCs were sold, how many were sold, when, and at what price. As an example, one movement data file might contain the weekly prices and sales volumes for all varieties of ground beef sold in the year 2009. They also include information such as whether a product was being promoted by the store when it was sold.

4.1.1 Data Aggregation

Note that computing the SUPI comparing product g across areas a_1 and a_2 discussed in Chapter III requires knowledge of vectors of area level prices

$$\bar{p}_{ga}^{\mathbb{GA}} = \begin{bmatrix} p_{g1a} \\ \vdots \\ p_{g|\Omega_a|a} \end{bmatrix} \tag{4.1}$$

as well as area level quantities

$$\bar{q}_{ga}^{\mathbb{G}\mathbb{A}} = \begin{bmatrix} q_{g1a} \\ \vdots \\ q_{g|\Omega_a|a} \end{bmatrix} \tag{4.2}$$

for each product g in the set $\mathbb{G}$ of products, and each area a in the set $\mathbb{A}$ of all comparison areas.

The Nielsen data do not furnish these area prices and quantities directly. Instead, they provide us with weekly prices and sales volume for each UPC in each participating store, along with general information about that store's location. In order to estimate the vectors in Equations 4.1 and 4.2, we need to aggregate these data to the appropriate spatiotemporal resolution. Usually we will not need to compare spatial or temporal price differences on a weekly basis, and the indices we produce in this paper specifically will be annual. Hence, for each product within each store, we must aggregate a year's worth of weekly prices and quantities before we can generate our estimates.

Before performing this aggregation however, we must consider the potential for error. Though the Nielsen data contains a wealth of useful information, because it is taken directly from the internal databases of participating retailers it retains the potential for significant measurement errors. For example, many point of sale systems generate transactions for product returns as well as purchases. Because prices in the Nielsen database must be positive, some rows in the data that represent returned products, and hence are associated with negative prices, may be generated with the lowest possible positive price, \$0.01. For similar reasons, promotions such as "buy one get one free" deals may also generate \$0.01 prices rather than registering a \$0.00 price for one of the items. Additionally, even though the data are mostly generated automatically, retailers might occasionally enter sales prices for some transactions manually. This might occur when stores experience technical malfunctions, for example inoperative bar code scanners. Thus there is also the potential for human errors, such as typos, in a portion of the price quotes. For example, a product normally sold at \$1.02 might mistakenly be recorded at a price of \$102.00.

These kinds of errors can have large effects on estimates of the average price level, and hence on our SUPI estimates. Most food products cost significantly more than \$0.01 per unit, and significantly less than \$102.00 per unit. Thus the inclusion of false prices with large magnitudes could cause our estimates of the average price per UPC across time to appear higher or lower than they otherwise would.

For this reason, before aggregation we screen the data for outliers or inaccurate records in the following way. First, we simply ignore any entries within a store with product prices listed as \$0.01 in order to avoid conflating product returns with product sales. Second, we filter the raw movement data to remove entries with outlying prices, i.e. weeks with prices that are extremely far from the median value for that product variety within the store that carries them will be omitted. More specifically, within each store we remove any entries with prices more than 3 interquartile ranges (IQR) above the 75th percentile or below the 25th percentile of that product variety's prices. This screening procedure should enable us to catch the data entry errors that would be most influential on the mean.

Formally, denote the set of stores included in our sample as $\mathbb{S}$, the set of weeks in the year as $\mathbb{W}$, and the set of UPCs as Ω. Let $p_{usw}^{\Omega\mathbb{SW}}$ be the price of product u in store s during week w , and $q_{usw}^{\Omega\mathbb{SW}}$ be the number of units of product u sold in store s in area a and week w. Also define $\mathbb{Q}_{us}^{\alpha}$ as the αth quantile of the weekly prices for UPC u in store s, and IQR_{us} as the interquartile range of the prices for UPC u in store s. Then the weekly movement data entries $p_{usw}^{\Omega\mathbb{SW}}$ and $q_{usw}^{\Omega\mathbb{SW}}$ are ignored when at least one of the following conditions hold:

1. $p_{usw}^{\Omega\mathbb{SW}} = \$0.01$

2. $p_{usw}^{\Omega\$\mathrm{W}} < \mathbb{Q}_{us}^{0.25} - 3 \times IQR_{us}$

3. $p_{usw}^{\Omega\$\mathrm{W}} > \mathbb{Q}_{us}^{0.75} + 3 \times IQR_{us}$

The proportion of observations that this procedure removes within each product module is summarized in Table 4.1 below. In most product modules, between 0 to 6% of the weekly observations are removed by this screen. The modules with high proportions (e.g. 20%+) of removed observations often contain few price quotations, so that filtering a small number of observations removes a large proportion of the data. For example, 1 out of 4 total observations of product module 2687 (frozen beef steaks) were removed in 2009, implying a 25% filtration rate. In a less extreme case, 1,522 out of 6,053 weekly observations of product module 1472 (monosodium glutamate and other flavor enhancers) in 2011 were removed by the outlier screen, which also implies a removal rate of about 25%. Though 6,053 total observations is much larger than 4, it is still small relative to the median of 85,009 weekly observations across all product modules in that year.

Table 4.1: Proportion of Filtered Quotations by Product Module

Year	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
2009	0.00	0.03	0.04	0.05	0.06	0.25
2010	0.00	0.03	0.04	0.05	0.06	0.18
2011	0.00	0.02	0.04	0.05	0.06	0.25
2012	0.00	0.03	0.04	0.05	0.06	0.20

Table 4.2: Number of Weekly Quotations by Product Module

Year	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
2009	4.00	21,716.50	82,226.00	248,229.12	232,911.00	6,838,750.00
2010	5.00	21,892.50	83,159.00	252,985.49	234,918.00	6,985,376.00
2011	2.00	22,029.25	85,009.00	268,340.67	248,478.50	7,456,623.00
2012	1.00	19,041.00	85,388.00	268,359.09	237,616.50	7,670,785.00

Because product modules with outlying removal rates tend to be more sparsely observed, they are often not observed in some counties. As we discuss in Section

4.2.2, our food CSUPIs are based on "common" product modules, i.e. ones that are observed in every in-sample Michigan county. Thus many product modules with lower numbers of observations and higher removal rates do not impact our aggregate indices. The proportion of observations that are removed within the common product modules are summarized in Table 4.3. The distribution of these proportions is similar

Table 4.3: Proportion of Filtered Quotations by Common Product Module

Year	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
2009	0.00	0.03	0.04	0.04	0.05	0.10
2010	0.01	0.04	0.05	0.05	0.06	0.14
2011	0.01	0.03	0.04	0.05	0.06	0.13
2012	0.01	0.04	0.04	0.05	0.06	0.13

to the distribution across all product modules, but has lower maximum values. We also note that the median common product module has far more quotations in-sample than the median product module, as we can see by comparing Tables 4.2 and 4.4.

Table 4.4: Number of Weekly Quotations by Common Product Module

Year	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
2009	13,932.00	167,431.50	287,939.00	681,617.18	646,931.50	6,838,750.00
2010	9,997.00	138,722.75	263,420.50	615,170.67	564,741.25	6,985,376.00
2011	24,216.00	139,550.00	268,990.00	631,164.02	582,416.00	7,456,623.00
2012	9,304.00	144,862.00	267,247.00	641,713.30	567,538.00	7,670,785.00

Given that the data has been filtered as described above, let $e_{usw}^{\Omega\$\mathbb{W}}$ denote the total expenditure on product u in store s and week w, calculated as

$$e_{usw}^{\Omega\$\mathbb{W}} = p_{usw}^{\Omega\$\mathbb{W}} \times q_{usw}^{\Omega\$\mathbb{W}} \tag{4.3}$$

We then calculate $e_{us}^{\Omega\$}$, the total annual expenditure in store s on product variety u, as the simple sum of the expenditure on product variety u sold within the set $\mathbb{W}$ of weeks in target time period:

$$e_{us}^{\Omega\$} = \sum_{w \in \mathbb{W}} e_{usw}^{\Omega\$\mathbb{W}} \tag{4.4}$$

We also calculate the store averaged prices for product u, p_{us}, as

$$p_{us}^{\Omega\mathbb{S}} = \sum_{w \in \mathbb{W}} r_{usw}^{\Omega\mathbb{S}\mathbb{W}} p_{usw}^{\Omega\mathbb{S}\mathbb{W}} \tag{4.5}$$

Here the weights r_{usw} are defined as the share of total store expenditure that occurred in week w:

$$r_{usw}^{\Omega\mathbb{S}\mathbb{W}} = \frac{e_{usw}^{\Omega\mathbb{S}\mathbb{W}}}{\sum_{w \in \mathbb{W}} e_{usw}^{\Omega\mathbb{S}\mathbb{W}}} \tag{4.6}$$

This weighting scheme ensures that short-term price variation does not unduly influence the store price level, unless a large number of units are sold.

After computing these store level average prices, we use a similar weighted averaging process to aggregate the store level data spatially, into county level prices for each product type u. Specifically, we again estimate area level expenditures e_{ua_1} for product u in area a_1 as

$$\hat{e}_{ua_1}^{\Omega\mathbb{A}} = \sum_{s \in S_{a_1}} e_{us}^{\Omega\mathbb{S}} \tag{4.7}$$

where S_{a_1} is the set of stores contained within area a_1.

Similarly, we estimate area prices $\hat{p}_{ua_1}^{\Omega\mathbb{A}}$ for product u and area a_1 as

$$\hat{p}_{ua_1}^{\Omega\mathbb{A}} = \sum_{s \in S_{a_1}} \hat{r}_{us}^{\Omega\mathbb{S}} p_{us}^{\Omega\mathbb{S}} \tag{4.8}$$

where the weights $\hat{r}_{us}^{\Omega\mathbb{S}}$ are defined as the share of the total areal expenditure in our sample that occurred in store s:

$$\hat{r}_{us}^{\Omega\mathbb{S}} = \frac{e_{us}^{\Omega\mathbb{S}}}{\sum_{s \in S_{a_1}} e_{us}^{\Omega\mathbb{S}}} \tag{4.9}$$

As in the temporal aggregation step, these weights assign more weight to price levels within stores whose sales represent large shares of areal expenditure.

Because the Nielsen data includes all price and expenditure data over the course of a year within each store included in the sample, we assume that within each store the average price $p_{us}^{\Omega\$}$ and expenditure $e_{us}^{\Omega\$}$ is known for all product varieties u that are sold within store s. However, not all stores within each area are included in our sample. Thus we regard the county-level quantities $\hat{p}_{ua}^{\Omega\mathbb{A}}$ and $\hat{e}_{ua}^{\Omega\mathbb{A}}$ as estimates of some unknown "true" county prices and expenditures. The uncertainty in these estimates is due to the existence of unobserved stores that may sell observed product varieties at different prices than are observed in sample, or sell product varieties that are not observed in our sample. By treating our estimates in this way, we are assuming that the data is a representative sample of the population in each area. In Section 4.1.2, we discuss several reasons that this assumption is questionable, among other limitations of the Nielsen retail scanner data.

4.1.2 Advantages and Limitations

The information discussed above includes all of the elements required to calculate average prices and demand for each good sold within each store, and to pinpoint the locations where transactions occurred across both time and space. These features make the Nielsen Retail Scanner dataset well-suited to estimating the SUPIs discussed in Chapter III, which require precisely this information. Despite this, there are limitations inherent to working with the Nielsen data.

Most seriously, the process that determined which stores were solicited for inclusion in the dataset is not disclosed to users, and participation was voluntary. The stores that are included in the sample are generally large chain stores, whose parent companies exercise varying degrees of input into which locations are included in the sample. The Nielsen Retail Scanner Dataset Manual explains that "for participating retailers, typically all stores in a retail chain within the 48 contiguous states are included," with the proviso that "in rare instances, a retailer may consider a small number of their stores as confidential and exclude them from the dataset."Because these cases are described as "rare," one might surmise that the impact of the excluded stores is minimal. However there is no way to know whether this is true, particularly when the data are used to measure price and expenditure levels in small areas. It is possible that the confidential stores account for large portions of consumer expenditure in the areas where they are located. Nielsen does not identify which retailers contribute to the retail scanner data, nor do they provide information about when, where or why retailers consider some stores to be confidential. For this reason it is difficult to characterize the impact of the excluded retailers and stores concretely. If retailers that contribute their data to the retail scanner data are different than those that do not, or confidential stores are different than other stores, their exclusion could cause the sample to misrepresent economic conditions in the areas being measured. Further, stores that are not associated with any larger parent company, or that do not use barcode scanners in order to process their transactions, are excluded from the sample entirely. Hence despite the volume of data available to us, it is unlikely that the stores represented in the Nielsen Retail Scanner Data constitute a representative sample of the population of stores within each area.

Another drawback is the lack of availability of prices for any services, or goods that are not associated with bar codes. This poses some limits on the kinds of economic activity that our index will be able to measure. For example, indices produced based on retail scanner data will not be able to address questions about differences in housing prices, or about variation in the cost of restaurant food. This is an area in which there still seems to be an argument for sample survey based data, such as that collected by the Bureau of Labor Statistics for the Consumer Price Index (CPI), as discussed in Section 2.2.1. There are also geographical restrictions on the indices that can be produced with this information. In particular, the Nielsen data does not contain information on any stores that are located outside of the continental United States. As such, indices for US states like Alaska and Hawaii or territories such as Puerto Rico cannot be estimated from Nielsen's sample. Additionally, the detailed characteristics necessary to adjust prices for quality through techniques such as hedonic regression are not consistently available across the items in the Nielsen sample. At best these sorts of characteristics might be known for the 1.4 million UPC's that are also in Nielsen's Consumer Panel Data, a figure which constitutes less than half of the 3.2 million UPC's in the Retail Scanner Data. As a result, methods based on adjusting product prices for quality using regression models will have limited applicability.

Despite these issues, one might argue that for the purposes of estimating spatial price indices for retail goods, the Nielsen data is of a higher quality than the sources of data used by the alternative indices we describe in Sections 2.2.1 and 2.2.2. The Nielsen data contains observations about many more product varieties than the data that the Council for Community and Economic Research (CCER) or the Bureau of Economic Analysis (BEA) can collect using traditional methods, such as sample surveys. These varieties are not chosen randomly according to some scheme as is the case in the CPI data, or by volunteers from a particular socioeconomic stratum as in the case of the CCER's volunteer data gatherers, but instead represent the census within each participating store.

Though the stores are not randomly sampled, Nielsen estimates that the transactions included in its Retail Scanner data constitute over half of the total sales volume from chain food and drug stores, and almost a third of the total sales volume originating from mass market merchandisers. Potential bias is mitigated by the fact that a large portion of the population is included in our sample.

Despite the fact that they tend to be concentrated in urban areas, the stores included in Nielsen's data are fairly widely dispersed. For example, the retail scanner data contains at least one store in 82 out of Michigan's 83 counties. We might expect that prices for the same product varieties will be similar in stores that are close to each other, since proximity makes it easier for consumers to shop at the least expensive outlet. Thus in small areas like counties, we might make an economic case that market competition should cause the between-store price variation to be relatively small within a given UPC. If this assumption is true, then any bias in our estimated area prices due to the participation of an unrepresentative sample of stores is likely to be outweighed by the reduction of variance due to having the census of product varieties, as compared to randomly sampling a single product variety from within each store to represent an entire product module. However even under this assumption, differences between the product varieties included the sample and the population retain the capacity to introduce error in our SUPI estimates. This is primarily due to the impact of the UPI variety adjustment term from Equations 3.7 and 4.14.

4.2 Estimating the SUPI

In this section, we discuss how to estimate the population value of the SUPI from our sample, at a given value of the elasticity of substitution parameter σ. Due to the fact that the elasticity of substitution attempts to quantify how consumers might react in various counterfactual situations, it is difficult to estimate from purely observational data without making strong identifying assumptions. Redding and Weinstein (2016)[34] discuss three approaches for estimating σ in an intertemporal context. The first requires estimating instrumental variables models within each product module. The second requires us to assume that changes to second-differenced supply and demand are orthogonal and heteroskedastic within all product modules. The third estimates only upper and lower bounds for σ, rather than providing a point estimate. This approach is no longer considered in Redding and Weinstein (2019)[35], the more recent revision of Redding and Weinstein (2016)[34]. Conducting experimental research on consumer propensity to substitute might enable us to circumvent some of the identification problems associated with the use of observational data. However, this would be a significant undertaking in its own right, and is beyond the scope of this research.

Consequently, we leave the question of how to estimate this parameter to others, and simply posit values for it. Specifically, we estimate SUPI values when σ is one of the elements of the following set:

$$\sigma \in \{2, 4, 6, 8, 10, 12, \infty\} \tag{4.10}$$

Estimating our indices for each of these values enables us to show how differences in our assumptions about consumer propensity for substitution between product varieties affect our estimated SUPI values.

Given these parameter values for σ, we describe how we estimated SUPIs comparing the county-level cost of living associated with (non-alcohol) food products in the state of Michigan. Michigan is an interesting case to study for several reasons. Michigan is a relatively large state, ranked as the 8th most populous in the 2010 Census. It has a large urban center in Detroit, surrounded by rings of relatively populous suburbs. Despite this, much of northern Michigan and the upper peninsula are predominantly rural. This gives the state a good mix of urban, suburban, and rural counties.

We chose to compare food costs between these counties because of food's ubiquity, importance, and high coverage rates within the Nielsen sample. Non-alcohol food products were manually identified from the description of the product groups in the products file. The identified non-alcohol food product groups are listed in Table B.1. Each of these product groups contains varying numbers of product modules, which are the objects of our lowest level SUPI indices. We would like to produce indices for a total of 583 observed product modules contained within these product groups. We discuss how to estimate the SUPI for a single product module in Section 4.2.1, and for broader categories of goods in 4.2.2. Aside from a few complications, this is accomplished by directly applying the formulas in Chapter III to the area prices and expenditures estimated as described in Section 4.1.1. Further details of this application are spelled out in Sections 4.2.1 and 4.2.2.

Only about a quarter of the modules included in the product groups from Table B.1 are ultimately used in our food indices, due to concerns about missing data. In particular, we exclude product modules that are not observed in every available Michigan county. Though the state of Michigan contains a total of 83 counties, the Nielsen sample contains observations on at least one store that stocks food items within 82 of these counties, listed in Table B.2. Hence there are a total of 82 available counties in our sample, which are listed in Table B.2. The lone excluded county is Keeweenaw, a small county in Michigan's upper peninsula with a 2010 population of around 2,156 people.[6]

The distribution of the number of counties in which each product module is observed is visualized in Figure 4.1. From this histogram, we can see that this distribution in each year appears multimodal, with sharp spikes at around 20 and 82 counties per product module, and a relatively low frequency of other values. We can also see that the number of product modules represented in all or almost all counties increases over time, possibly pointing to an increase in overall data quality between 2009 and 2012.

Table 4.5 shows the five number summary of the number of counties per product module within each year. From this table, we can see that around a quarter of the product modules in each year are generally represented in every county each year.

Table 4.5: Five Number Summaries of # of Counties per Product Module by Year

Year	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
2009	1	23	63	54	81	82
2010	1	26	65	56	82	82
2011	1	27	73	58	82	82
2012	1	29	79	60	82	82

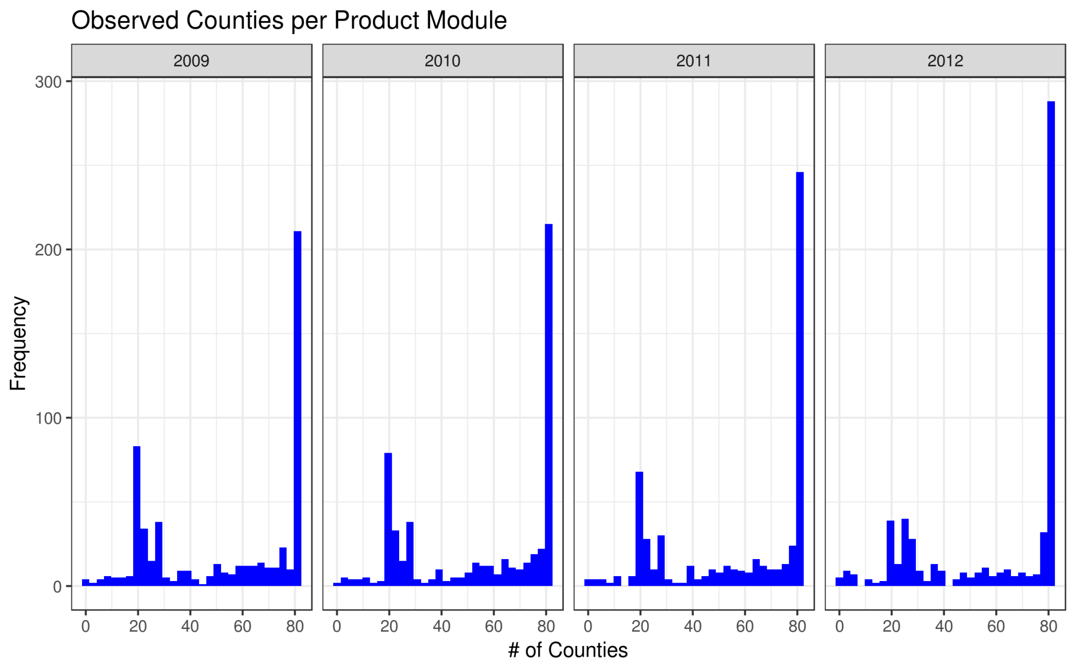

Figure 4.1: Histogram of # of observed counties per product module

The product modules that are observed in every county form the basis of our Food CSUPIs in each year. We discuss the reasoning behind our choice to omit the product modules that are not observed in each of these counties in Section 4.2.2.

4.2.1 Estimating Product Module SUPIs

Let Ω_a^g represent the set of product varieties of a given product module $g \in \mathbb{G}$ sold in area $a \in \mathbb{A}$, and $V = |\Omega_a^g|$, the number of observed product varieties in area a. Within a given a, we can stack the area prices $\hat{p}_{ua}^{\Omega\mathbb{A}}$ and expenditures $\hat{e}_{ua}^{\Omega\mathbb{A}}$ as estimated in Section 4.1.1 over the set of product varieties contained in product module g. In this section we do not use any store level quantities, and hence we suppress the superscripts Ω and $\mathbb{A}$ denoting the level of product and area aggregation respectively. This gives us the vectors

$$\vec{p}_{ga} = \begin{bmatrix} \hat{p}_{1a} \\ \vdots \\ \hat{p}_{Va} \end{bmatrix} \qquad \text{and} \qquad \vec{e}_{ga} = \begin{bmatrix} \hat{e}_{1a} \\ \vdots \\ \hat{e}_{Va} \end{bmatrix} \tag{4.11}$$

for each area $a \in \mathbb{A}$ and each product module $g \in \mathbb{G}$.

Given these vectors, we estimate product module SUPIs using the equations described in Chapter III. Specifically, let $\Omega^g_{a_1a_2}$ be the set of varieties of a product module g that are observed in both areas a_1 and a_2, and $V^c = |\Omega^g_{a_1a_2}|$, the number of common varieties of g. Then the Jevons price index for product module g is estimated from $\vec{p}_{ga_1}$ and $\vec{p}_{ga_2}$ using Equation 3.5:

$$\widehat{Jevons}_{ga_1a_2} = \prod_{u \in \Omega^g_{a_1a_2}} \left(\frac{\hat{p}_{ua_2}}{\hat{p}_{ua_1}} \right)^{\frac{1}{V^c}} \tag{4.12}$$

The area expenditure vectors $\vec{e}_{a_1}$ and $\vec{e}_{a_2}$ are used to estimate the area a common expenditure shares

$$\hat{r}_{ua} = \frac{\hat{e}_{ua}}{\sum_{v \in \Omega^g_{a_1a_2}} \hat{e}_{ua}}$$

for $a \in \{a_1, a_2\}$, as required to calculate the spread adjustment (SADJ) term of the UPI using Equation 3.6:

$$\widehat{SADJ}_{ga_1a_2} = \left[\prod_{v \in \Omega^g_{a_1a_2}} \left(\frac{\hat{r}_{va_2}}{\hat{r}_{va_1}} \right)^{\frac{1}{V^c}} \right]^{\frac{1}{\sigma-1}} \tag{4.13}$$

Finally we estimate the lambda ratios

$$\hat{\lambda}_{ga} = \frac{\sum_{v \in \Omega^g_{a_1a_2}} \hat{e}_{va}}{\sum_{v \in \Omega^g_{a}} \hat{e}_{va}}$$

for $a \in \{a_1, a_2\}$ that are required to estimate the variety adjustment term with Equation 3.7:

$$\widehat{VADJ}_{ga_1a_2} = \left(\frac{\hat{\lambda}_{a_2}}{\hat{\lambda}_{a_1}}\right)^{\frac{1}{\sigma-1}} \tag{4.14}$$

Having estimated these terms, we can estimate the product module g UPI between counties a_1 and a_2 as

$$\widehat{UPI}_{ga_1a_2} = \widehat{Jevons}_{ga_1a_2} \times \widehat{SADJ}_{ga_1a_2} \times \widehat{VADJ}_{ga_1a_2} \tag{4.15}$$

We estimate $\widehat{UPI}_{ga_1a_2}$ for all pairwise combinations of counties in the state of Michigan, and apply the GEKS procedure described in Section 3.2 to these comparisons to estimate the SUPI relative to some chosen reference area.

4.2.2 Estimating Aggregate Indices

Given that we have estimated SUPI values for some collection of individual goods as discussed in Section 4.2.1, we would like to apply the weighted geometric mean discussed in Section 3.3 to estimate indices for a collection of goods $\mathbb{G}$ as

$$\widehat{CSUPI}_{a_1a_2} = \left(\prod_{g\in\mathbb{G}} \widehat{SUPI}_{ga_1a_2}^{\hat{\omega}_g}\right) \tag{4.16}$$

where

$$\hat{\omega}_g = \frac{\sum_{a\in\mathbb{A}} \sum_{v\in\Omega_a^g} \hat{e}_{va}}{\sum_{a\in\mathbb{A}} \sum_{g\in\mathbb{G}} \sum_{v\in\Omega_a^g} \hat{e}_{va}} \tag{4.17}$$

In practice, estimating $\widehat{CSUPI}_{a_1a_2}$ using Equation 4.16 is not possible if there are missing SUPI values for any of the product modules in $\mathbb{G}$.

Although the SUPI accounts for differences in the *product varieties* that are available in each area, our CSUPIs do not account for differences in the *products* available in each area. For this reason, our procedure can cope with missing product varieties better than it copes with missing products. For example, if Swiss Miss brand chocolate pudding is observed in area a_1 but not in area a_2, the SUPI will quantify the impact of this difference on the pudding cost of living. However, it may be the case that our sample does not contain data on *any* varieties of pudding in area a_2. In such a case, we say that pudding is a missing product in area a_2. This missingness means that we cannot estimate $\widehat{SUPI}^{\hat{\omega}_{pudding}}_{pudding,a_1a_2}$. If pudding is one of the goods in $\mathbb{G}$, then this missing SUPI value prevents us from calculating $\widehat{CSUPI}_{a_1a_2}$ straightforwardly.

This problem arises because the weighted geometric mean is effectively a "common goods" approach to aggregating our indices. The construction of the mean assumes that indices for every relevant product are, or at least could be, observed in every area. Under this assumption, the absence of data about pudding is because the stores that sell pudding are not included in our sample. While this is possible, we do not have enough information about the coverage of the Nielsen sample in each area to know whether it is true. It is also possible that some products are not available from any retailers in some areas, and thus the data on these products are structurally missing.

If all grocery goods are assumed to be substitutes, then we could cope with structurally missing products by treating the SUPIs for each product as "prices," and using the UPI and GEKS to create aggregate indices. This approach would account for the impact of product availability on consumer well-being by treating the individual products within a category of products similarly to how the SUPI treats the individual product varieties within each product. However, the extent to which different food products can usefully be considered substitutes is unclear. For example, the idea that bananas, honey, coffee and ground beef are substitutes seems more questionable than the idea that Folgers, Starbucks, and Maxwell House coffees are substitutes. For this reason, we avoid this approach and accept the compromises that the weighted geometric mean requires us to make.

Given this framework, we could deal with missing products in one of three ways. First, we could simply take the weighted geometric average of the available good indices $\widehat{SUPI}_{ga_1a_2}$, omitting the indices for the missing products. This preserves the all of the available information from our initial product level $\widehat{SUPI}_{ga_1a_2}$ comparisons without making any additional assumptions about the values of the missing product comparisons. Unfortunately, the category level indices $\widehat{CSUPI}_{a_1a_2}$ resulting from this procedure will no longer be transitive. To see this, let $\mathbb{G} = \{1, \ldots, G\}$ and assume without loss of generality that good G is missing in area a_1, so that $\widehat{SUPI}_{ga_1a_2}$ and $\widehat{SUPI}_{ga_1a_3}$ cannot be estimated. Then we would calculate

$$
\begin{aligned}
\widehat{CSUPI}_{a_1a_2}\widehat{CSUPI}_{a_2a_3} &= \left(\prod_{g=1}^{G-1} \widehat{SUPI}_{ga_1a_2}^{\hat{\omega}_p}\right)\left(\prod_{g=1}^{G} \widehat{SUPI}_{Ga_2a_3}^{\hat{\omega}_g}\right) \\
&= \widehat{SUPI}_{Ga_2a_3}^{\hat{\omega}_G}\left(\prod_{g=1}^{G-1} \widehat{SUPI}_{ga_1a_2}^{\hat{\omega}_g}\right)\left(\prod_{g=1}^{G-1} \widehat{SUPI}_{ga_2a_3}^{\hat{\omega}_g}\right) \\
&= \widehat{SUPI}_{Ga_2a_3}^{\hat{\omega}_G}\left(\prod_{g=1}^{G-1} \widehat{SUPI}_{ga_1a_3}^{\hat{\omega}_g}\right) \\
&= \widehat{SUPI}_{Ga_2a_3}^{\hat{\omega}_G}\widehat{CSUPI}_{a_1a_3}
\end{aligned}
\tag{4.18}
$$

Thus with even one missing good, the estimated category level indices will be intransitive, unless the estimated cost of living for the missing product is equal across all areas in $\mathbb{A}$.

More seriously, using this method means that some of our comparisons will include information about products that are omitted from other comparisons. This can cause bias in our CSUPIs due to composition effects. For example, if a cluster of product modules with disproportionately high cross-county SUPI variance are the only ones observed in area a_1, then a CSUPI calculated in this way will overestimate the differences between a_1 and other counties. We must also consider that our imposition of transitivity upon the product module level SUPIs makes comparisons within each product dependent. Depending on which counties are missing, this can introduce biases into our CSUPI estimates. For example, imagine that counties with high costs of living are missing at a higher rate than other counties. In this scenario, the SUPIs we estimate for product modules with missing counties might be lower on average than in products where every county was observed. The CSUPIs are a weighted average of these SUPIs, and would thus inherit these biases.

To avoid these issues, we could use some form of imputation to fill in the missing values. If this were done for each missing county in every product module, it might be possible to mitigate the composition bias issues mentioned above. This would also enable us to preserve transitivity. However, it is unclear what information is both widely available, and useful for predicting unobserved values of the SUPI. It is possible that the data required for this imputation would differ based on the product module, effectively requiring us to specify hundreds of separate imputation models. This is an effort that is beyond the scope of our project.

More fundamentally, this approach assumes that the imputed products are not structurally missing, i.e. every individual product $g \in \mathbb{G}$ is sold in every area, and missing data are solely due to the exclusion of stores from our sample. Imputing SUPI values for areas in which the associated product is not even available for purchase could be misleading. Without additional information on the coverage rates of the Nielsen sample in each area, we have no way to know when missing products are structurally missing, and thus whether imputation is appropriate.

For these reasons we prefer a third alternative, which is restricting the product set $\mathbb{G}$ that our category level indices are based on to products that are represented in all areas, rather than attempting to impute any of the missing product indices. This "common goods" approach has the significant downside that it requires us to discard information about all products that are missing in even one area. Only about a quarter of the product modules observed in our sample remain after this restriction has been imposed. However, it does not require us to impute, and ensures that CSUPIs comparing any two areas are based on the same set of products.

4.2.3 Estimation Limitations

In this section, we discuss three limitations of our approach to estimating SUPIs and CSUPIs from the data, and how they might affect our results. The first limitation is that we avoid estimating the unknown elasticity of substitution parameter σ, and instead produce estimates for several arbitrarily chosen σ values within each product module, as definfed by Nielsen. This assumes that product varieties within Nielsen product modules are all equally substitutable. This may make more sense in narrowly defined product modules, such as "mustard", than in more broadly defined modules such as "shelf stable entrees / side dishes." If product varieties that are not substitutes are grouped into the same category, our SUPIs will assume that consumers have more options for consumption of that good than they really do. In this circumstance, we might expect the variety adjustment term from the UPI to be lower than it "should" be. In our estimates, σ values are also assumed to be identical *between* different product modules. This implies, for example, that different varieties of tuna are exactly as substitutable as different varieties of vinegar. Thus the potential impact of between-product σ variation on our CSUPIs is ignored. As we will see in Chapter V, our SUPI estimates depend heavily on the values of σ we posit. For this reason, we can expect that these assumptions will have a strong impact on our results.

Another limitation is that our CSUPIs drop all goods that are not observed in every area under consideration, as discussed in Section 4.2.2. While this enables us to deal with missing data without needing to specify imputation models or distinguish structurally missing products from products that are not included in our sample for other reasons, these conveniences come with a substantial cost. About 75% of the in-sample product modules for each year are ignored as a result of this restriction. These products are not necessarily excluded at random, meaning that our CSUPIs may misrepresent differences in the aggregate cost of living between areas if missing products are systematically different than included ones. The set of products that are observed in all counties can also change from year to year, potentially confounding comparisons of CSUPI values across time.

Finally, a third limitation relates to our procedure for removing outliers. As discussed in Section 4.1.1, this procedure can be more aggressive than alternative methods for removing outliers, such as trimming a fixed percentage of observations from the upper and lower tails of each price distribution. Alternatively, it can be much less aggressive; in some product modules, no observations are removed by this screen. We can see from Table 4.1 that the percentage of filtered observations from common product modules ranges from about 0% to 14% across the four years we consider. A simplier outlier screen that trims a consistent proportion of the observations within each store and product variety may yield more easily interpretable results than the approach that we have chosen.

4.3 Estimation Uncertainty

The Nielsen retail scanner data set provides complete information on annual sales data within all stores included in the sample. However, because it does not include every store, our SUPI estimates will have some degree of error associated with them. Indices comparing areas on the basis of sparse data may have suspiciously large or small values, and for this reason it is useful to have some indication of which comparisons are reliable. We can divide the uncertainty that affects our estimates into two components.

The first component is error due to unobserved stores selling observed product varieties at different prices. For example, the average price of bacon in Washtenaw county may be $3.00 in our sample, but $2.75 when the unobserved stores selling this variety are accounted for. Similarly, the share of expenditure on bacon may be higher in the county as a whole than is reflected in the observed stores.

The second component is error due to unobserved stores selling unobserved product varieties. The SUPI has no way to distinguish between products that are entirely unavailable for purchase in an area and products that are not sold in the stores in our sample. Because the UPI assumes that consumers value variety, areas where the in-sample stores sell fewer product varieties than are actually available can appear more "expensive" relative to areas where the in-sample stores stock a higher percentage of the available product varieties.

The main challenge in accounting for both sources of uncertainty is that we do not have access to a representative random sample of stores in each area, as discussed in Section 4.1.2. Because we have limited information about the factors that motivate each retailer to participate, we are unable to characterize the impact of this non-random selection procedure on our estimates. Instead, we ignore this issue by assuming that stores are missing from our sample at random, and estimate the variance due to each component under this assumption.

Given this assumption, we apply the cluster bootstrap method explicated by Field and Welsh (2007)[19] to estimate the bias and variance of our SUPIs and CSUPIs. In our application of this method, we treat the estimated area prices and expenditures as being nested within product variety (or UPC) clusters. More formally, we rely on the following superpopulation model. Assume that within each product module, we draw $|\Omega|$ different product varieties from an infinite superpopulation of possible varieties that could have been observed. Each product variety u is sold in areas $\mathbb{A}_u = \{a_{u1}, \ldots, a_{u|\mathbb{A}_u|}\}$, within which we estimate the vectors

$$\vec{p}_u^{\Omega} = \begin{bmatrix} \hat{p}_{ua_{u1}}^{\Omega\mathbb{A}} \\ \vdots \\ \hat{p}_{ua_{u|\mathbb{A}_u|}}^{\Omega\mathbb{A}} \end{bmatrix} \quad \text{and} \quad \vec{e}_u^{\Omega} = \begin{bmatrix} \hat{e}_{ua_{u1}}^{\Omega\mathbb{A}} \\ \vdots \\ \hat{e}_{ua_{u|\mathbb{A}_u|}}^{\Omega\mathbb{A}} \end{bmatrix} \tag{4.19}$$

of stacked area price and expenditure estimates. Each element of $\vec{p}_u^{\Omega}$ and $\vec{e}_u^{\Omega}$ are computed from store level prices and expenditures $(p_{us}^{\Omega\$}, e_{us}^{\Omega\$})$ as discussed in Section 4.1.1. These store level prices and expenditures for each product variety u in area a are jointly drawn from an infinite superpopulation of possible stores in each area according to an unknown distribution $\mathbb{F}$ with mean price $\bar{p}_{ua}$ and mean expenditure $\bar{e}_{ua}$:

$$(p_{us}^{\Omega\$}, e_{us}^{\Omega\$}) \sim \mathbb{F}\big((p, e)|\bar{p}_{ua}, \bar{e}_{ua}\big) \tag{4.20}$$

This enables us to generate bootstrap replicates of our SUPIs for each product module by repeatedly sampling with replacement from the clusters (product varieties) within each product module, and applying the procedure discussed in Section 4.2.2 to the resampled data. Section 4.3.1 contains a detailed description of how this is done.

Resampling area prices and expenditures as vectors preserves the empirical correlation between each element of the vectors. This means that the relationship of area-level prices and expenditures within each product variety should be the same in our bootstrap replications as it is in the data, both within and between areas. This is important because the Jevons, SADJ, and VADJ terms in Equations 3.5, 3.6 and 3.7 are ratios of area-level functions of prices and expenditures. Breaking this correlation structure would result in SUPI replicates that do not reflect the economic relationships between the areas being compared. For example, imagine that the prices of common goods in area a_2 are double the prices of common goods in area a_1. The bootstrap distribution of the Jevons index might not be centered around 2 if prices are resampled across different areas without taking this relationship between a_1 and a_2 into account. This would result in a biased bootstrap distribution that may not reflect the appropriate mean and variance for our SUPI estimate. For this reason, it is necessary to resample in clusters that maintain the observed relationship between prices and expenditures in different counties.

Resampling product varieties within products has two major advantages over alternative methods of cluster resampling, such as resampling stores within areas. First, resampling product varieties enables us to estimate bootstrap variances for areas in which small numbers of stores are observed. A substantial number of counties in our sample contain products that are only observed in one store. If we generated our bootstrap distribution by resampling stores within areas, we would not be able to estimate variances for SUPIs or CSUPIs involving these counties. In our chosen framework, we can estimate variances even in counties where a single store is observed as long as there are multiple observed product varieties within that store. Second, resampling product varieties can be done with data aggregated to the area level, which is appreciably smaller in size than the same data aggregated to the store level. Because the required data are of a smaller size, and there is no need to aggregate store level data to county level data on each replication, resampling product varieties is more computationally efficient than resampling stores.

To gain these advantages however, we make a problematic assumption. In our framework prices $p_{us}^{\Omega\$}$ and expenditures $e_{us}^{\Omega\$}$ can be correlated within product variety u. Similarly, product prices $p_{us_1a_1}$ and $p_{us_2a_2}$ can be correlated across arbitrary areas $a_1, a_2 \in \mathbb{A}_u$, as can expenditures $e_{us_1a_1}$ and $e_{us_2a_2}$. However, the prices and expenditures of different product varieties are assumed to be *mutually independent* of each other within each area.

This assumption deserves further discussion, as it is both important to the procedure we outline in this section, and probably untrue. Effectively, this is equivalent to assuming that there are no store level effects that could induce correlations in the prices of the different product varieties sold within that store. A violation of this assumption would occur if, for example, Whole Foods were to sell the exact same varieties of ground beef as other stores, but at systematically higher prices. In such a circumstance, the prices of different product varieties in Whole Foods will be correlated, which in turn will introduce correlations between our area prices and expenditures that contradict our assumption here.

This false assumption can affect the bootstrap variance estimates we describe in Section 4.3.2. If different product varieties within the same store are priced similarly, so that the average prices of different product varieties are positively correlated, then we would expect this procedure to underestimate the variance of the SUPI. However as long as the magnitude of the average correlation between the area prices is relatively small, our estimates should still be serviceable.

4.3.1 Generating Bootstrap Replicates

Given the framework discussed above, we generate bootstrap replications of our food CSUPIs according to a three-step process. In the first step, we generate B bootstrap replications of the observed data within each product module $g \in \mathbb{G}$ by sampling with replacement from the UPCs. Formally, if there are $|\Omega^g|$ UPCs in the set Ω^g, then for each bootstrap replication $b \in \{1, \ldots, B\}$ we draw a sample Ω_{bg} of size $|\Omega^g|$ within each product module g.

The vectors $(\vec{p}_u^{\Omega}, \vec{e}_u^{\Omega})$ from Equation 4.19 that are associated with the resampled product varieties are then used to assemble a bootstrap data set $\mathbb{D}_g^b$ for each product module $g \in \mathbb{G}$ and bootstrap replication $b \in \{1, \ldots, B\}$. For example, say that the set of observed product varieties in product module g is $\Omega^g = \{u_1, u_2, u_3\}$. A resample Ω^{bg} of Ω^g might be $\{u_3, u_3, u_1\}$. In this case, the bootstrap data $\mathbb{D}_g^b$ associated with Ω^{bg} would be $\mathbb{D}_g^b = \left((\vec{p}_g^{\mathbb{G}})^b, (\vec{e}_g^{\mathbb{G}})^b\right)$, where the price and expenditure vectors are defined as

$$(\vec{p}_g^{\mathbb{G}})^b = \begin{bmatrix} \vec{p}_{u_3}^{\Omega} \\ \vec{p}_{u_3}^{\Omega} \\ \vec{p}_{u_1}^{\Omega} \end{bmatrix} \qquad \text{and} \qquad (\vec{e}_g^{\mathbb{G}})^b = \begin{bmatrix} \vec{e}_{u_3}^{\Omega} \\ \vec{e}_{u_3}^{\Omega} \\ \vec{e}_{u_1}^{\Omega} \end{bmatrix} \tag{4.21}$$

respectively.

In the second step, we stack the area prices and expenditures from each bootstrap data set $\mathbb{D}_g^b$ over area (rather than UPC), and use the resulting vectors as described in Section 4.2.1 to estimate the matrix

$$\mathbf{UPI}_g^b = \begin{bmatrix} \widehat{UPI}_{g11}^b & \dots & \widehat{UPI}_{g1A}^b \\ \vdots & \ddots & \vdots \\ \widehat{UPI}_{gA1}^b & \dots & \widehat{UPI}_{gAA}^b \end{bmatrix} \tag{4.22}$$

containing the bth replication of all the UPIs comparing areas in $\mathbb{A} = \{1, \dots, A\}$ within each product module $g \in \mathbb{G}$.

Note that because the area prices and expenditures are resampled within UPC, it is possible that the set of areas for which the UPI is estimable will change from replication to replication. For example, suppose that in our previous example product variety u_2 is the only variety of product module g that is sold in some area $\tilde{a}$. Because our resample Ω^{bg} does not include u_2, there is no way to compute the UPI between area $\tilde{a}$ and any other area for this product module. The proportion of times such cases occur relative to the total number of bootstrap replications is recorded, and treated as an additional diagnostic statistic measuring the sensitivity of each comparison to changes in the sampled product varieties. A high proportion of non-estimable replications typically indicates a "brittle" comparison, where the value of the UPI depends heavily on one or two UPCs. Because these non-estimable UPIs are ignored in subsequent calculations, comparisons associated with a high proportion of non-estimable replications should be distrusted even if variance estimates based on the bootstrap appear to be low.

The bootstrap UPI matrices $\mathbf{UPI}_g^b$ provide us with the necessary information to compute $SUPIs$ on each of the bootstrap price data sets. Because SUPIs are transitive, the set of areal comparisons that they imply will be consistent regardless of the chosen reference area. Therefore in each replication we need only compute SUPIs relative to a fixed reference area a_1, rather than needing to estimate all pairwise

SUPIs. For this reason, we denote the bth bootstrap SUPI comparing area a_1 to a_2 as $\widehat{SUPI}^{b}_{ga_1a_2}$, and compute the vector of bootstrap estimates

$$\overrightarrow{SUPI}^{b}_{g} = \begin{bmatrix} \widehat{SUPI}^{b}_{ga_1 1} \\ \vdots \\ \widehat{SUPI}^{b}_{ga_1 A} \end{bmatrix} \tag{4.23}$$

Thus for each bootstrap replication $b \in \{1, \ldots, B\}$, we estimate $|\mathbb{G}|$ associated UPI matrices $\mathbf{UPI}^{b}_{g}$, one for each product module. These are in turn used to estimate $|\mathbb{G}|$ vectors $\overrightarrow{SUPI}^{b}_{g}$ containing the SUPIs for each product module.

Finally, the third step consists of taking the expenditure weighted geometric means of the $|\mathbb{G}|$ product module indices in order to estimate food-level SUPIs relative to the reference area a_1, as described in Section 4.2.2. This should yield a vector of bootstrap estimates $\widehat{CSUPI}^{b}_{a_1a_2}$ for the estimated category level SUPIs $\widehat{CSUPI}_{a_1a_2}$:

$$\vec{C}^{b} = \begin{bmatrix} \widehat{CSUPI}^{b}_{a_1 1} \\ \vdots \\ \widehat{CSUPI}^{b}_{a_1 A} \end{bmatrix} \tag{4.24}$$

4.3.2 Estimating SUPI Variability

In this section, we discuss how to use the bootstrap replicates of the UPI, SUPI, and category level SUPI indices described in Section 4.3.1 to estimate the variability of each of these price indices. For each of these indices, rather than estimating the arithmetic standard deviation, we instead turn to the geometric standard deviation. The main reason for this is that the geometric standard deviation of $SUPI_{a_1a_2}$ will be the same as the geometric standard deviation of $SUPI_{a_2a_1}$, whereas the arithmetic standard deviation will not.

We estimate geometric standard deviations for our indices as follows. Suppose that we have B bootstrap replicates of the UPI index comparing areas a_1 and a_2 in product module g. Then the geometric standard deviation of this index is estimated as

$$GSD(UPI_{ga_1a_2}) = exp\left(\sqrt{\frac{1}{B}\sum_{b=1}^{B}\left(log(\widehat{UPI}^{b}_{ga_1a_2}) - \frac{1}{B}\sum_{b=1}^{B}log(\widehat{UPI}^{b}_{ga_1a_2})\right)^2}\right) \tag{4.25}$$

Similarly, the geometric standard deviation of the SUPI index comparing areas a_1 and a_2 in product module g is estimated as

$$GSD(SUPI_{ga_1a_2}) = exp\left(\sqrt{\frac{1}{B}\sum_{b=1}^{B}\left(log(\widehat{SUPI}^{b}_{ga_1a_2}) - \frac{1}{B}\sum_{b=1}^{B}log(\widehat{SUPI}^{b}_{ga_1a_2})\right)^2}\right) \tag{4.26}$$

Finally, if $\widehat{CSUPI}_{a_1a_2}$ is the estimated SUPI for a category $\mathbb{G}$ of goods such as food, then the geometric standard deviation for $\widehat{CSUPI}_{a_1a_2}$ is estimated as

$$GSD(CSUPI_{a_1a_2}) = exp\left(\sqrt{\frac{1}{B}\sum_{b=1}^{B}\left(log(\widehat{CSUPI}^{b}_{a_1a_2}) - \frac{1}{B}\sum_{b=1}^{B}log(\widehat{CSUPI}^{b}_{a_1a_2})\right)^2}\right) \tag{4.27}$$

These calculations might yield results like 1.01, 1.2 or 1.5 meaning that the average bootstrap replication was within 1%, 20% or 50% respectively of its mean. Because geometric standard deviations are multiplicative, the variability that each of these values imply depends on the geometric mean of the bootstrap indices. For example, an index with a geometric mean of 1.03 and geometric standard deviation of 1.01 is less variable than an index with a geometric mean of 10.3 and a geometric standard deviation of 1.01.

The lowest possible value for the geometric standard deviation is 1, which would imply that every bootstrap replication had exactly the same estimated index value. This extreme result will typically only occur in cases where there is only one in-sample UPC within a product module for at least one of the comparison areas. As discussed in Section 4.3.1, these cases are often associated with high rates of non-estimable UPI values. When this is the case, they should be interpreted as signifying that there is too little information in our sample to properly estimate the variability associated with that index.

Note that all of the cases in Tables B.4, B.5, B.6 and B.7 are listed with geometric standard deviations rounded to two decimal places. A value of 1.00 in these tables generally implies a geometric standard deviation that is below this rounding threshold. Aside from indices comparing the reference county to itself, all of the estimated food indices have some variability in their bootstrap replicates, and hence none of their geometric standard deviations are equal to 1 before rounding.

CHAPTER V

Results

In this chapter, we discuss the results of estimating the indices described in Chapter III using the data and methods described in Chapter IV to compare the cost of living associated with non-alcohol food items between the 82 counties listed in Table B.2.

In Section 5.1 we explore the product module level SUPI estimates. Because there are 583 separate food product modules, and hence 583 sets of potential SUPI comparisons that we could in principle examine, we focus on a random sample of 3 of the 113 product modules that are observed in all areas across all years. This enables us to visualize the relationships between our estimates of the SUPI, the UPI, and the various UPI subindices discussed in Section 3.1 across time and space without being overwhelmed with information, or encountering "gaps" in our visualizations where the index value could not be estimated. This also ensures that the product modules that we visualize are ones which ultimately contribute to the Food CSUPIs, whose values and interpretation we discuss in Section 5.2.

5.1 Product Module SUPIs

In this section, we discuss estimation results for a randomly selected sample of three product modules, as discussed above. The three product modules that were sampled are listed in Table 5.1 below:

Table 5.1: Sampled Product Modules

Product Module Code	Product Module Description
1188	VINEGAR
1209	SEAFOOD-TUNA-SHELF STABLE
1360	CRACKERS - FLAVORED SNACK

The SUPI estimates for these product modules at each assumed value of σ are summarized in Tables 5.2, 5.3 and 5.4 below:

Table 5.2: Vinegar (1188) SUPI 5 Number Summaries by Year and σ

σ	Year	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
2	2009	0.80	5.72	9.60	9.89	12.64	77.50
2	2010	1.00	8.56	16.17	16.29	20.70	131.17
2	2011	0.82	5.54	9.03	9.57	10.88	96.83
2	2012	1.00	4.05	10.87	14.44	14.18	382.79
4	2009	0.80	1.38	1.85	1.74	2.07	3.70
4	2010	1.00	2.11	2.50	2.36	2.81	5.11
4	2011	0.87	1.67	2.00	1.90	2.15	4.48
4	2012	1.00	1.52	2.17	2.07	2.38	6.61
6	2009	0.76	1.08	1.33	1.27	1.40	2.02
6	2010	1.00	1.54	1.72	1.66	1.86	2.67
6	2011	0.88	1.29	1.49	1.43	1.57	2.42
6	2012	1.00	1.26	1.55	1.51	1.68	2.94
8	2009	0.74	0.99	1.13	1.12	1.23	1.55
8	2010	1.00	1.34	1.48	1.44	1.57	2.02
8	2011	0.89	1.15	1.31	1.28	1.38	1.86
8	2012	1.00	1.17	1.34	1.33	1.46	2.07
10	2009	0.73	0.92	1.04	1.04	1.12	1.41
10	2010	1.00	1.24	1.35	1.33	1.43	1.73
10	2011	0.89	1.08	1.21	1.20	1.29	1.61
10	2012	1.00	1.13	1.24	1.24	1.35	1.71
12	2009	0.72	0.88	1.00	1.00	1.06	1.34
12	2010	1.00	1.19	1.28	1.26	1.34	1.57
12	2011	0.89	1.06	1.16	1.15	1.23	1.47
12	2012	1.00	1.10	1.18	1.19	1.28	1.51
Inf	2009	0.62	0.75	0.79	0.82	0.91	1.07
Inf	2010	0.92	0.99	1.02	1.02	1.04	1.10
Inf	2011	0.84	0.92	0.98	0.97	1.01	1.09
Inf	2012	0.87	0.94	0.98	0.98	1.01	1.08

Reference County: Washtenaw (FIPS 26161)

Table 5.3: Tuna (1209) SUPI 5 Number Summaries by σ

σ	Year	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
2	2009	1.00	3.38	6.89	8.65	9.05	37.92
2	2010	1.00	2.90	7.07	7.41	9.36	27.28
2	2011	1.00	2.72	4.94	5.24	6.37	24.98
2	2012	1.00	1.75	2.85	2.95	3.16	26.43
4	2009	1.00	1.59	2.00	1.96	2.24	3.57
4	2010	1.00	1.51	1.98	1.88	2.17	3.18
4	2011	1.00	1.45	1.75	1.69	1.92	2.68
4	2012	1.00	1.29	1.50	1.44	1.55	2.92
6	2009	1.00	1.37	1.58	1.52	1.68	2.26
6	2010	1.00	1.28	1.51	1.47	1.61	2.07
6	2011	1.00	1.28	1.42	1.38	1.50	1.81
6	2012	1.00	1.21	1.32	1.28	1.35	1.88
8	2009	1.00	1.26	1.41	1.36	1.48	1.85
8	2010	1.00	1.19	1.35	1.32	1.42	1.72
8	2011	1.00	1.20	1.30	1.27	1.35	1.58
8	2012	1.00	1.17	1.25	1.21	1.27	1.56
10	2009	1.00	1.19	1.32	1.29	1.38	1.66
10	2010	1.00	1.16	1.27	1.25	1.33	1.56
10	2011	1.00	1.15	1.24	1.21	1.28	1.46
10	2012	1.00	1.15	1.21	1.18	1.23	1.40
12	2009	1.00	1.15	1.27	1.24	1.31	1.55
12	2010	1.00	1.13	1.22	1.21	1.28	1.46
12	2011	0.99	1.12	1.20	1.18	1.23	1.39
12	2012	1.00	1.13	1.18	1.16	1.20	1.31
Inf	2009	0.97	1.02	1.06	1.06	1.10	1.16
Inf	2010	0.90	1.01	1.04	1.03	1.07	1.18
Inf	2011	0.88	1.01	1.04	1.03	1.05	1.12
Inf	2012	0.97	1.04	1.07	1.07	1.09	1.18

Reference County: Washtenaw (FIPS 26161)

Table 5.4: Flavored Snack Crackers (1360) SUPI 5 Number Summaries by σ

σ	Year	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
2	2009	1.00	2.58	5.97	6.49	9.09	16.65
2	2010	1.00	2.84	7.83	8.97	14.68	20.32
2	2011	1.00	2.99	8.45	9.92	14.60	33.54
2	2012	1.00	2.47	5.42	5.78	7.67	24.15
4	2009	1.00	1.44	1.82	1.79	2.10	2.63
4	2010	1.00	1.46	2.03	1.96	2.45	2.82
4	2011	1.00	1.47	2.09	2.01	2.39	3.40
4	2012	1.00	1.37	1.77	1.69	1.93	2.52
6	2009	1.00	1.28	1.43	1.42	1.55	1.86
6	2010	1.00	1.29	1.51	1.49	1.69	1.95
6	2011	1.00	1.27	1.56	1.51	1.69	2.17
6	2012	1.00	1.20	1.41	1.35	1.47	1.68
8	2009	1.00	1.20	1.30	1.29	1.37	1.60
8	2010	1.00	1.20	1.34	1.33	1.44	1.66
8	2011	1.00	1.20	1.38	1.35	1.45	1.80
8	2012	1.00	1.12	1.27	1.23	1.31	1.46
10	2009	1.00	1.16	1.23	1.22	1.28	1.48
10	2010	1.00	1.16	1.26	1.25	1.33	1.52
10	2011	1.00	1.16	1.29	1.26	1.35	1.62
10	2012	1.00	1.10	1.20	1.17	1.24	1.35
12	2009	1.00	1.13	1.18	1.18	1.23	1.40
12	2010	0.98	1.14	1.21	1.20	1.27	1.44
12	2011	1.00	1.13	1.23	1.21	1.28	1.52
12	2012	0.99	1.08	1.15	1.14	1.19	1.29
Inf	2009	0.76	1.00	1.02	1.02	1.04	1.11
Inf	2010	0.73	1.00	1.02	1.01	1.04	1.12
Inf	2011	0.81	1.00	1.02	1.02	1.05	1.17
Inf	2012	0.81	0.97	1.00	0.99	1.02	1.05

Reference County: Washtenaw (FIPS 26161)

Our SUPI estimates for these product modules in 2009, 2010, 2011, and 2012 are mapped for $\sigma \in \{6, 8, 10, 12, \infty\}$ in Figures 5.1, 5.2, 5.3 and 5.4 respectively. Though we also estimate SUPIs for $\sigma \in \{2, 4\}$, we omit these from many of our visualizations because their values are significantly more extreme than the other indices, which makes the color scale difficult to interpret for the other comparisons on the map. We note that in each of these years and for each of the sampled product modules, our SUPI estimates indicate that the cost of living is significantly higher outside of a small cluster around the Detroit Metropolitan Area to the southeast of the state. As discussed in Chapter III, the UPI component indices are economically meaningful, and can help us to discover the drivers of this clustering effect. In Section 5.1.1 we assess the drivers of this result by examining the relationship of the SUPI to the UPI. In Section 5.1.2 we examine the relationship of the UPI to its component indices.

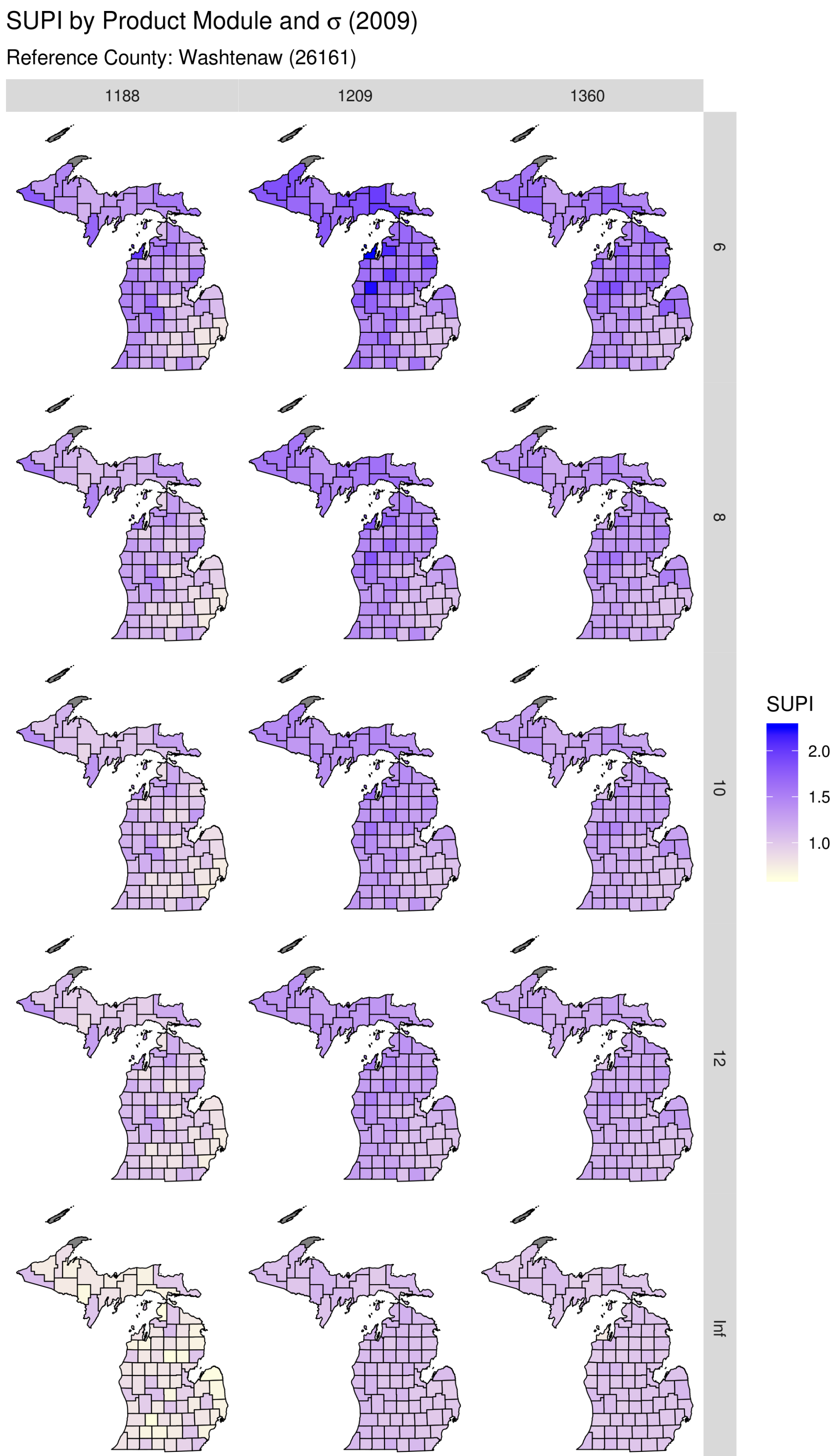

Figure 5.1: Product Module SUPI Choropleth (2009)

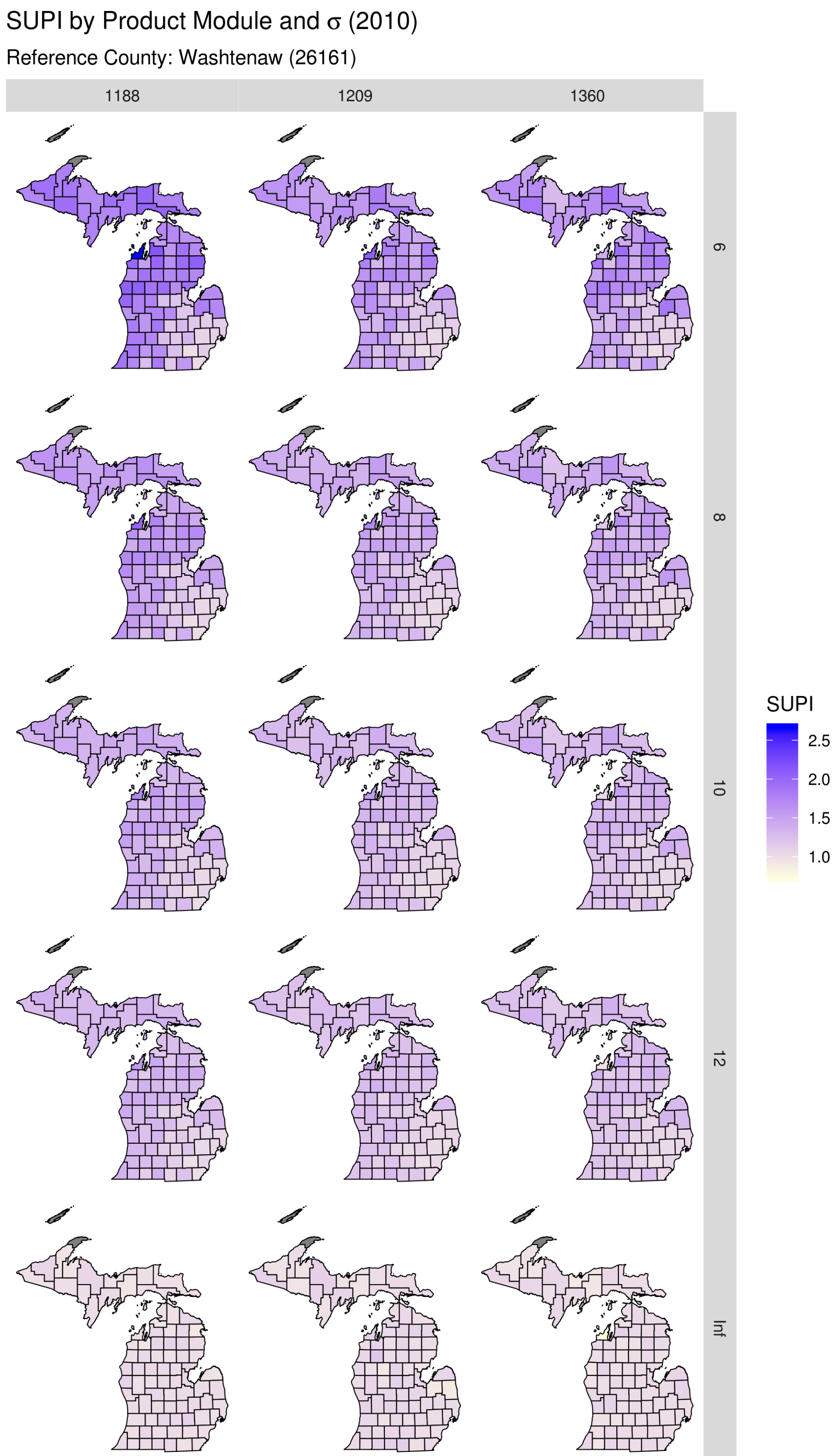

Figure 5.2: Product Module SUPI Choropleth (2010)

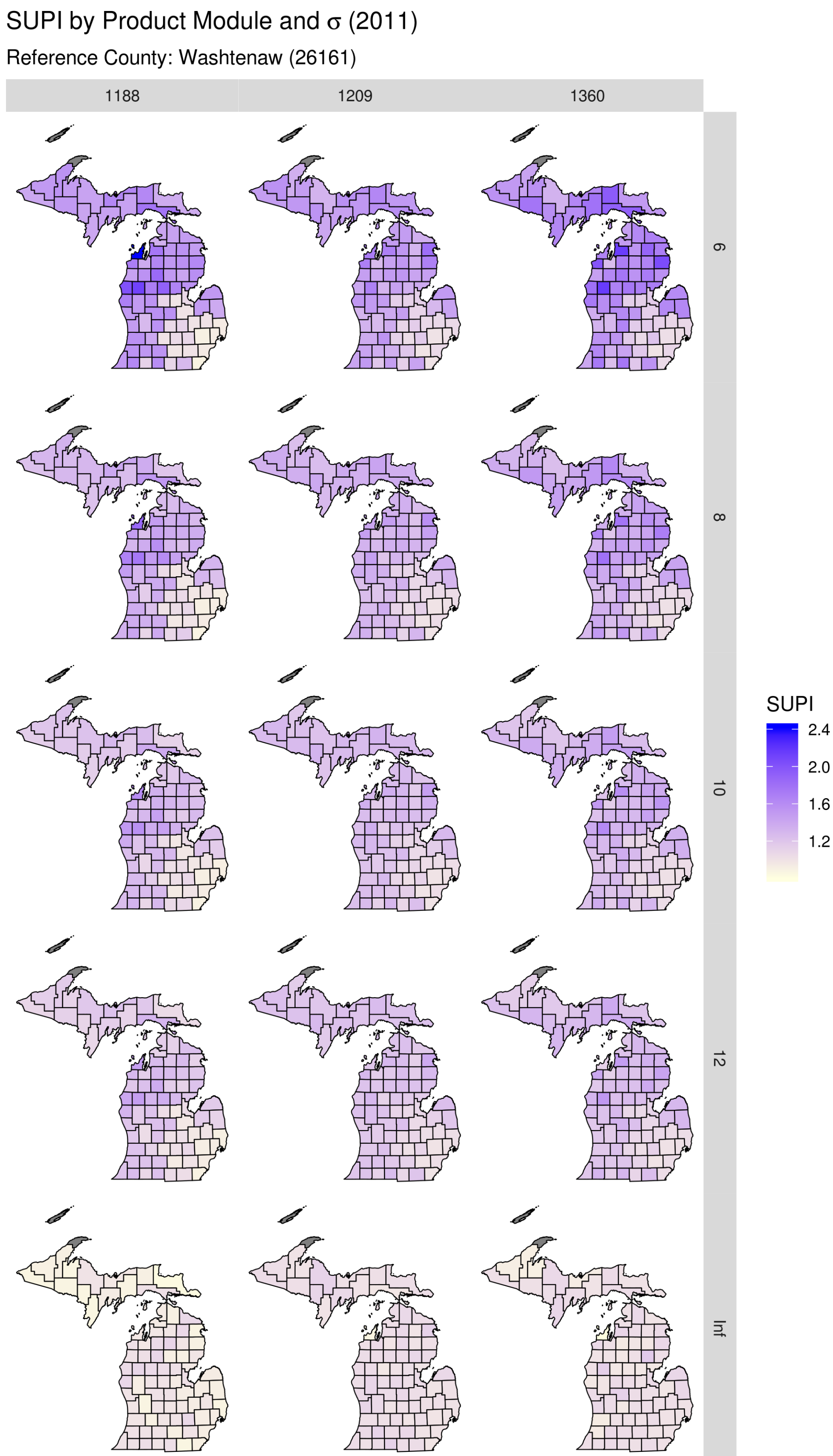

Figure 5.3: Product Module SUPI Choropleth (2011)

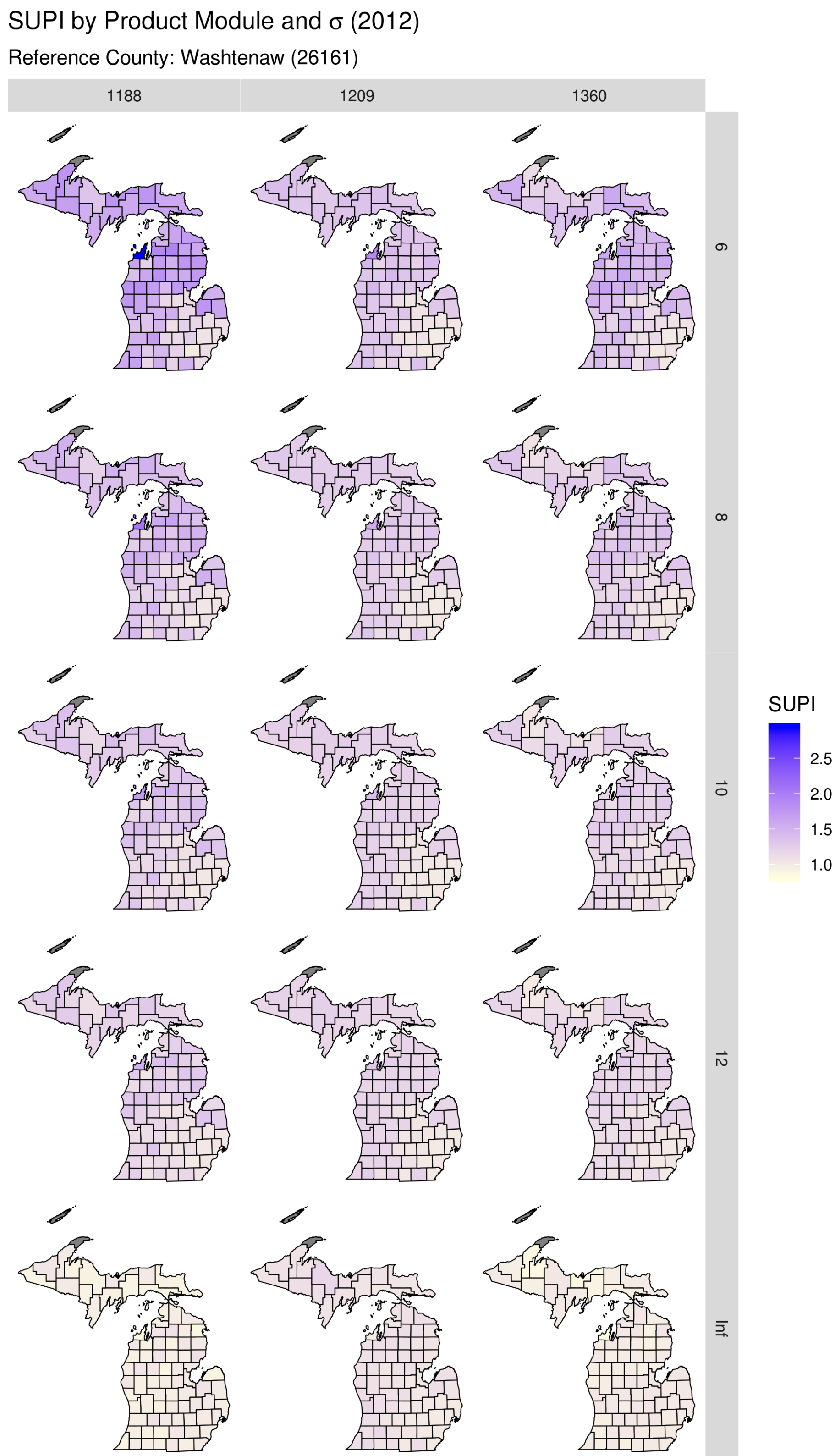

Figure 5.4: Product Module SUPI Choropleth (2012)

Versions of these maps based in a geometric average of all in-sample Michigan counties, rather than Washtenaw county as above, are included in Appendix A.

5.1.1 Product Module SUPIs vs. UPIs

Because SUPIs are indices derived from imposing transitivity on the UPIs according to a least squares loss function, as discussed in Chapter III, we might expect that our estimated SUPI values will appear similar to the UPIs that they are estimated from. From the scatterplots in Figures 5.5, 5.6, 5.7 and 5.8, we can see that this expectation is borne out for many of the indices. However, some scatterplots exhibit a clear "bent" pattern, where the SUPI values track the UPI closely at low values, but are systematically lower than the corresponding UPI at high values.

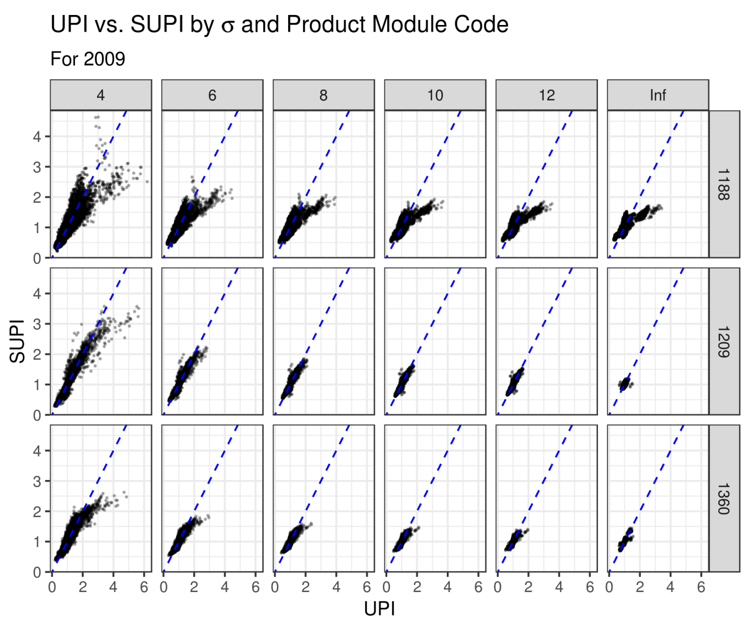

Figure 5.5: Scatterplot of UPI vs SUPI for 3 Product Modules (2009)

We can think about this feature of the SUPI positively or negatively. Because the SUPI needs to reconcile some large index values with other smaller index values involving the same area, many of the most extreme values wind up being "smoothed out." One might consider this a "loss of area characteristicity," as discussed in Section

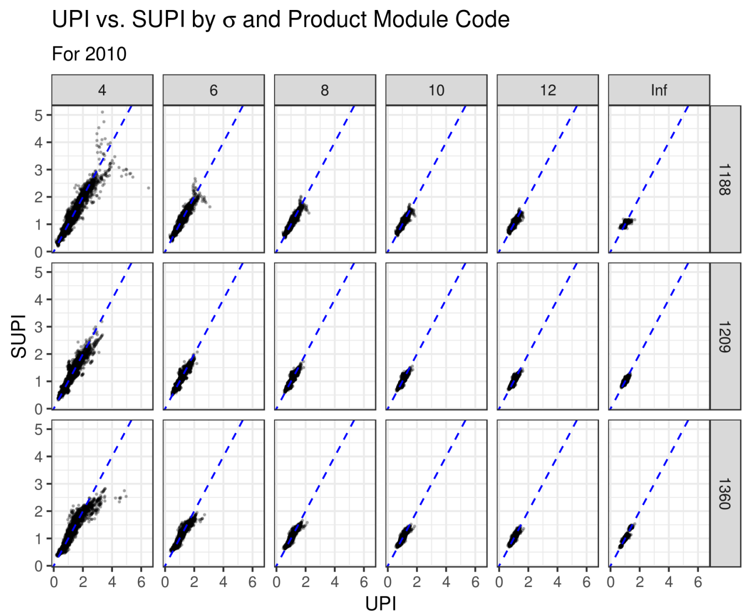

Figure 5.6: Scatterplot of UPI vs SUPI for 3 Product Modules (2010)

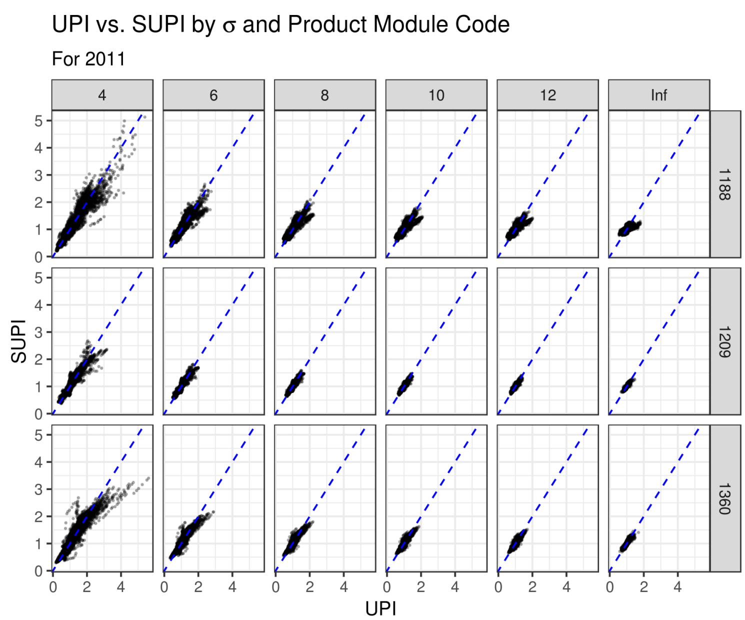

Figure 5.7: Scatterplot of UPI vs SUPI for 3 Product Modules (2011)

3.2. However, this same smoothing might be considered a positive factor in a context in which our UPI estimates are subject to significant sampling error. In such a circumstance, we might expect that the extremity of these estimates may reflect the influence of noise, rather than intransitive characteristics of the comparison areas. To

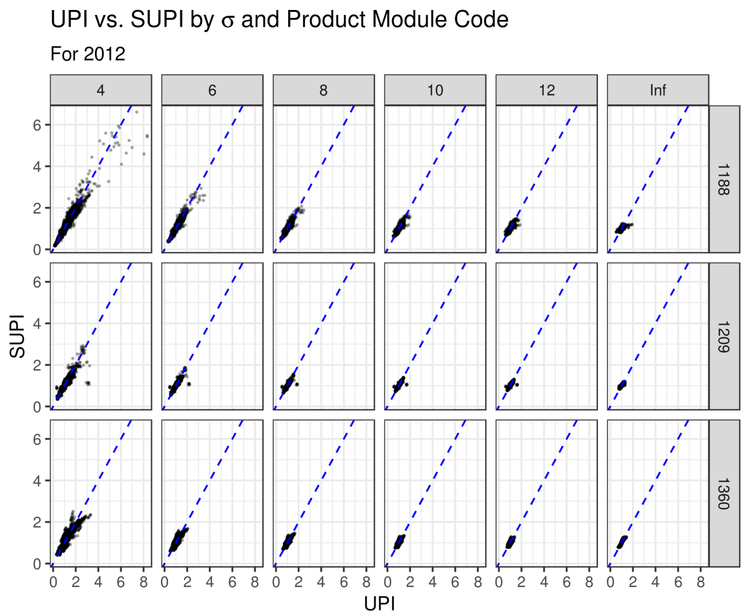

Figure 5.8: Scatterplot of UPI vs SUPI for 3 Product Modules (2012)

the extent that this is true, then the additional smoothing imposed by the SUPI may be a benefit of the procedure rather than a drawback.

5.1.2 Interpreting UPI Components

Because the UPIs form the basis from which the SUPIs are estimated, SUPIs are generally highly correlated with UPIs, as we saw in Section 5.1.1. We can thus understand our SUPI estimates by understanding the drivers of our UPI values. Note that on the log scale, the UPI is the sum of its logged subindices:

$$log(UPI_{a_1a_2}) = log(Jevons_{a_1a_2}) + log(SADJ_{a_1a_2}) + log(VADJ_{a_1a_2}) \tag{5.1}$$

Figures 5.9, 5.10 and 5.11 are stacked bar charts showing the log(UPI) for the products in Table 5.1 as the sum of these components for each county. Counties in these plots are identified by their Federal Information Processing Standard (FIPS) codes. We form these numbers by concatenating the FIPS code for the state of Michigan (26) with a three-digit county FIPS code corresponding to the individual county. These codes are matched to their corresponding county names in Table B.2.

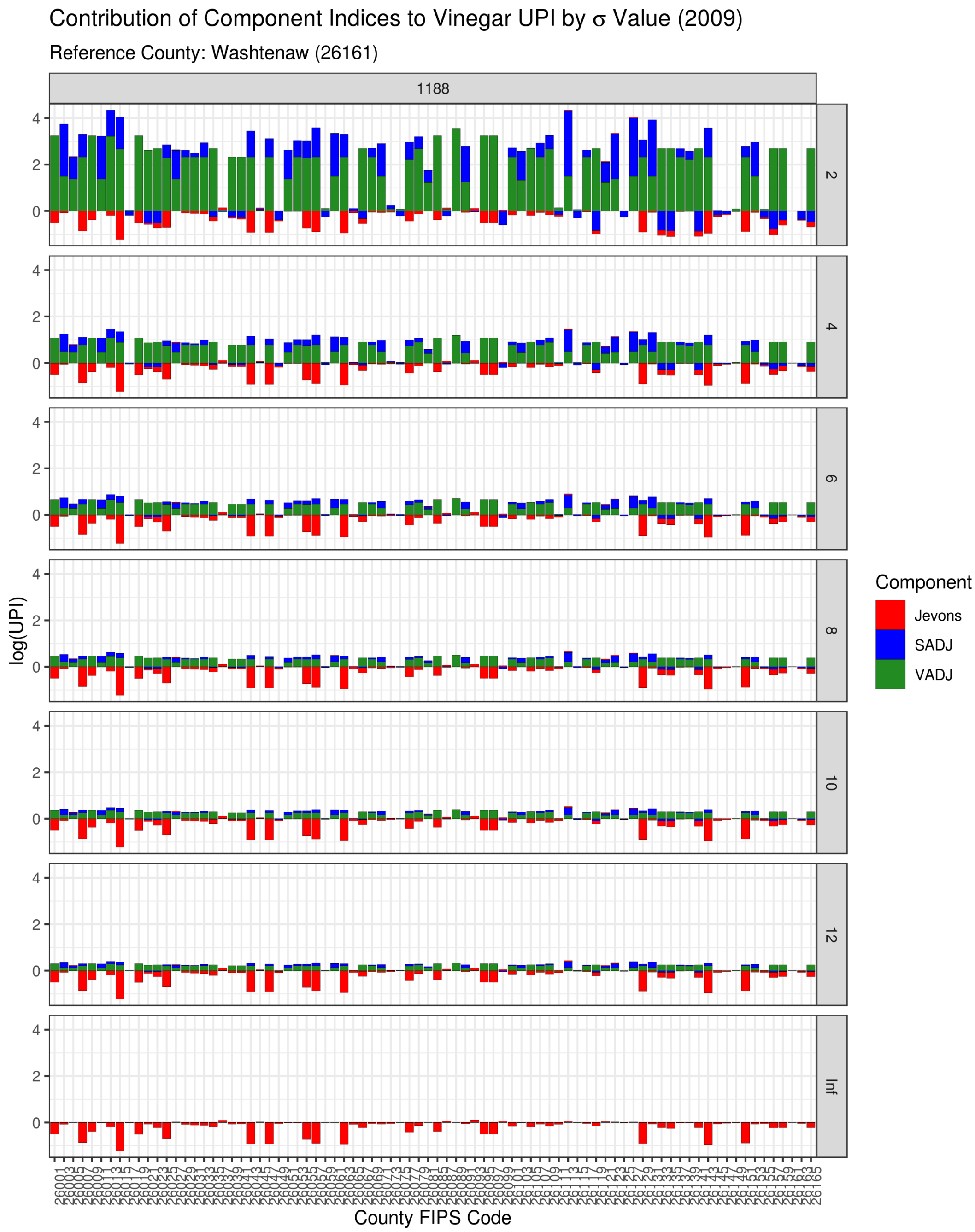

Figure 5.9: Contribution of Jevons, SADJ, and VADJ to Vinegar UPI (2009)

The most striking result we see in these bar charts is the strong influence of the SADJ and VADJ terms on the overall UPI result for low values of σ. The Jevons index, which reflects the differences in cost of living due to differences in the prices of

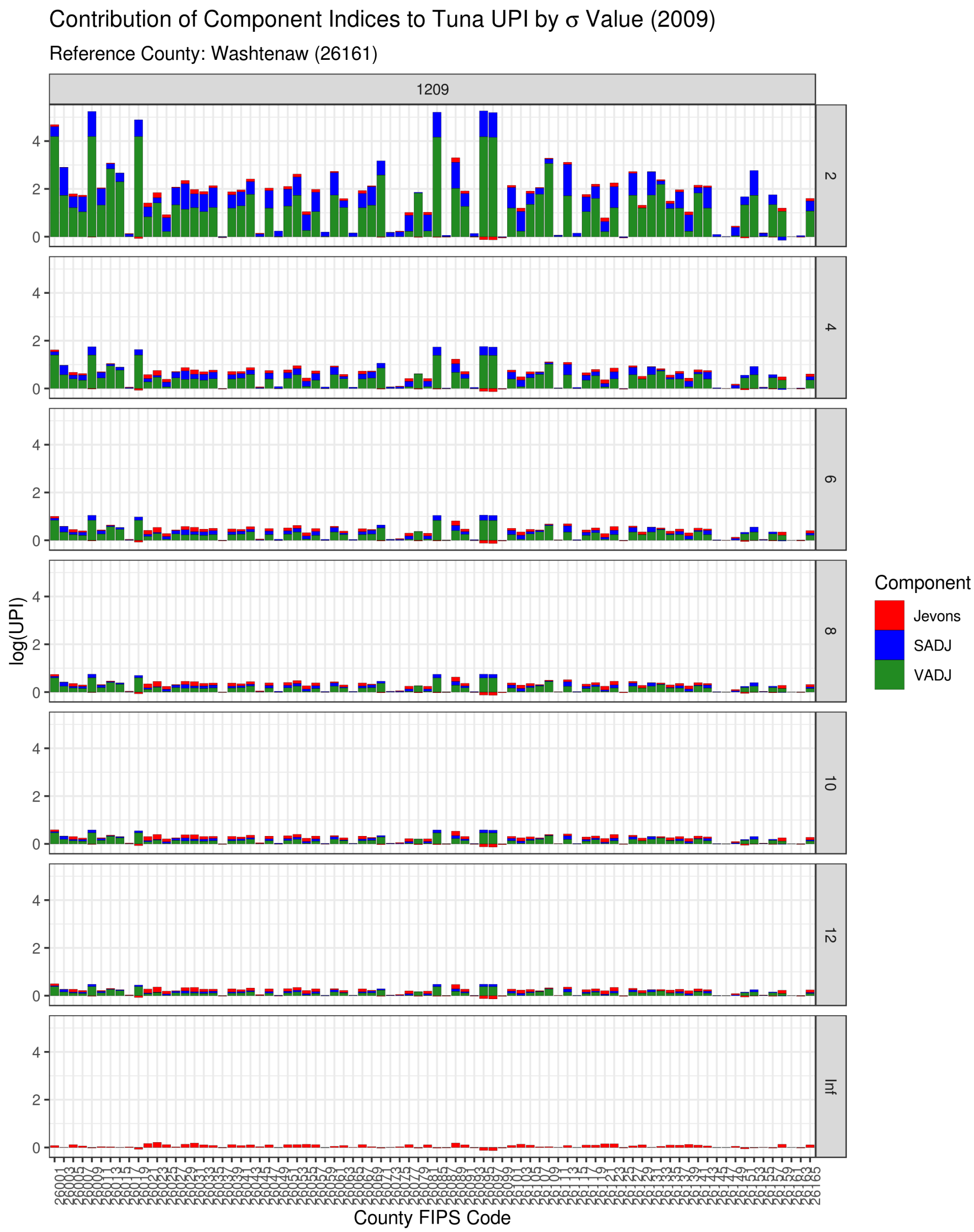

Figure 5.10: Contribution of Jevons, SADJ, and VADJ to Tuna UPI (2009)

product varieties sold in both of the comparison areas, generally accounts for a small proportion of the overall cost of living difference between counties when $\sigma < \infty$. In particular, the differences in product variety availability between counties captured by the VADJ term have dramatic effects on the cost of living when the elasticity of

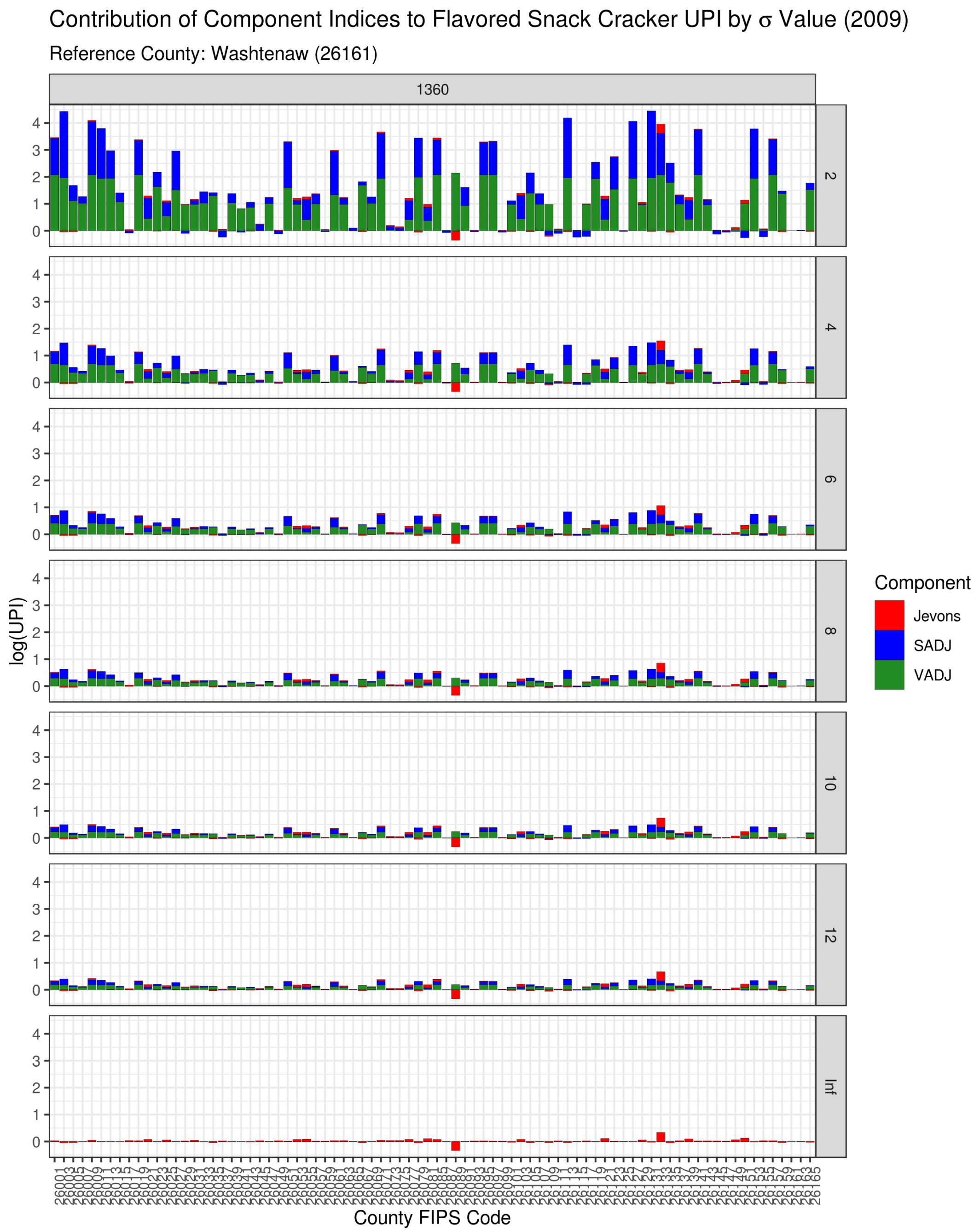

Figure 5.11: Contribution of Jevons, SADJ, and VADJ to Flavored Snack Cracker UPI (2009)

substitution is assumed to be very low.

We can also see this in Figures 5.12, 5.13 and 5.14, which display the relationship between the estimated SUPI, UPI, Jevons, SADJ and VADJ terms for each of the product modules in Table 5.1 in 2009. These graphs show that most of the higher estimated index values correspond to UPIs with Jevons and SADJ terms that are relatively close to 1.0, but high VADJ terms.

Despite this, many of the SUPIs associated with outlying UPI values are smoothed closer to the rest of the indices. As discussed in Section 5.1.1, this is a consequence of the imposition of transitivity via the GEKS procedure. Indices whose values are high enough that they are incommensurate with the other index values are "bent" closer towards the mean, as observed in Figures 5.5, 5.6, 5.7 and 5.8.

The magnitude of the SADJ and VADJ terms lessens as the assumed value of σ increases, as we would expect based on the theory in Chapter III. We can see the impact of this on our SUPI estimates in the clusters of vinegar SUPI estimates with Jevons and SADJ / VADJ terms that pull in opposite directions. For the lower values of σ in our plot, the SUPIs associated with these counties trend towards somewhere in the middle of the points representing each of the UPI component indices. As σ increases though, we see the value of these SUPIs fall closer to that of the Jevons as the influence of the SADJ and VADJ terms attenuates. The same pattern of behavior is visible in the flavored snack cracker SUPI for Leelanau county (FIPS 26089) in Figure 5.14.

From the maps in Section 5.1, we can observe that the between county variation appears to decrease over time in some product modules. Figure 5.15 shows the between-county standard deviation of the SUPI and its associated UPI component indices across time.

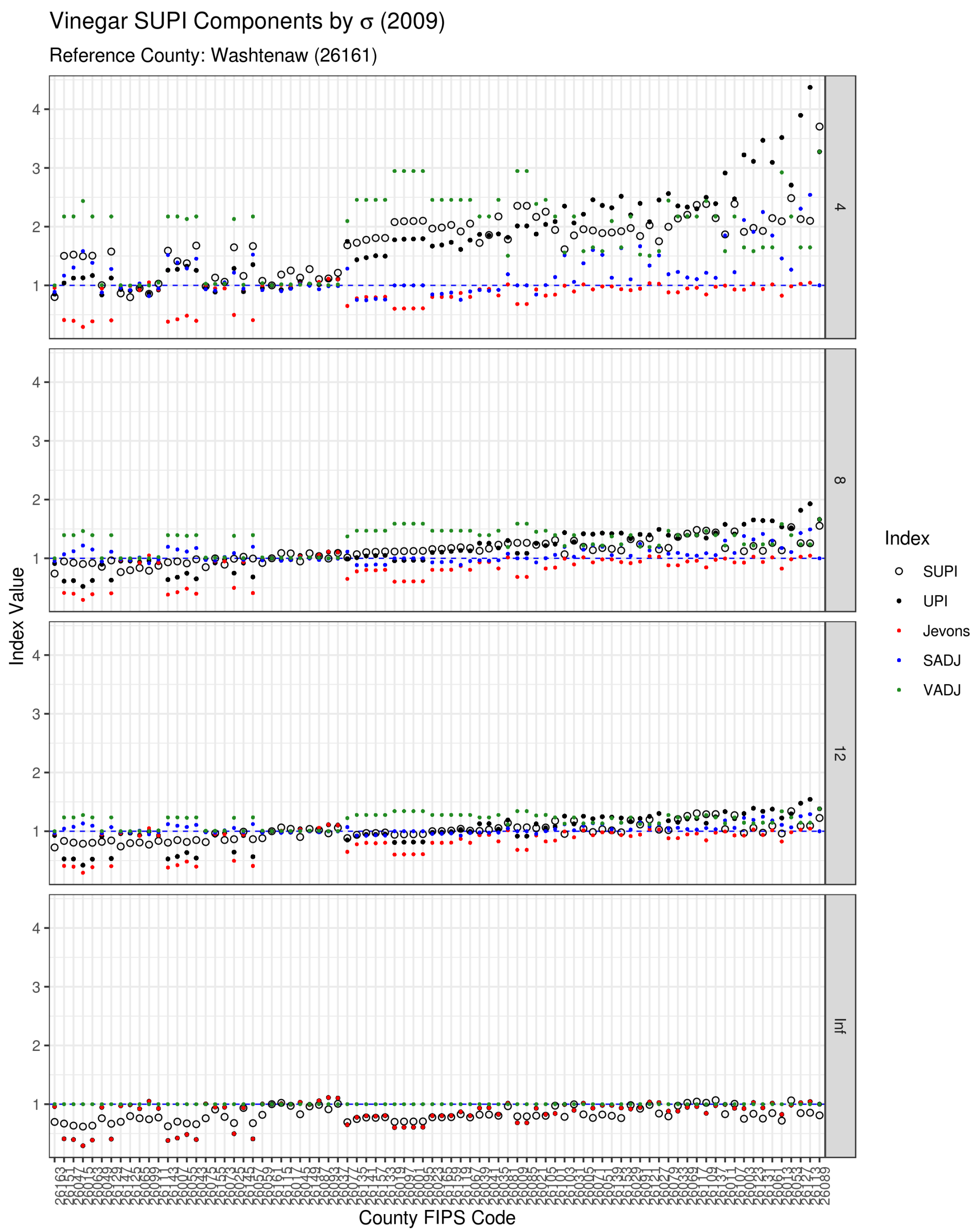

Figure 5.12: Vinegar SUPI, UPI, Jevons, SADJ, and VADJ by County and σ (2009)

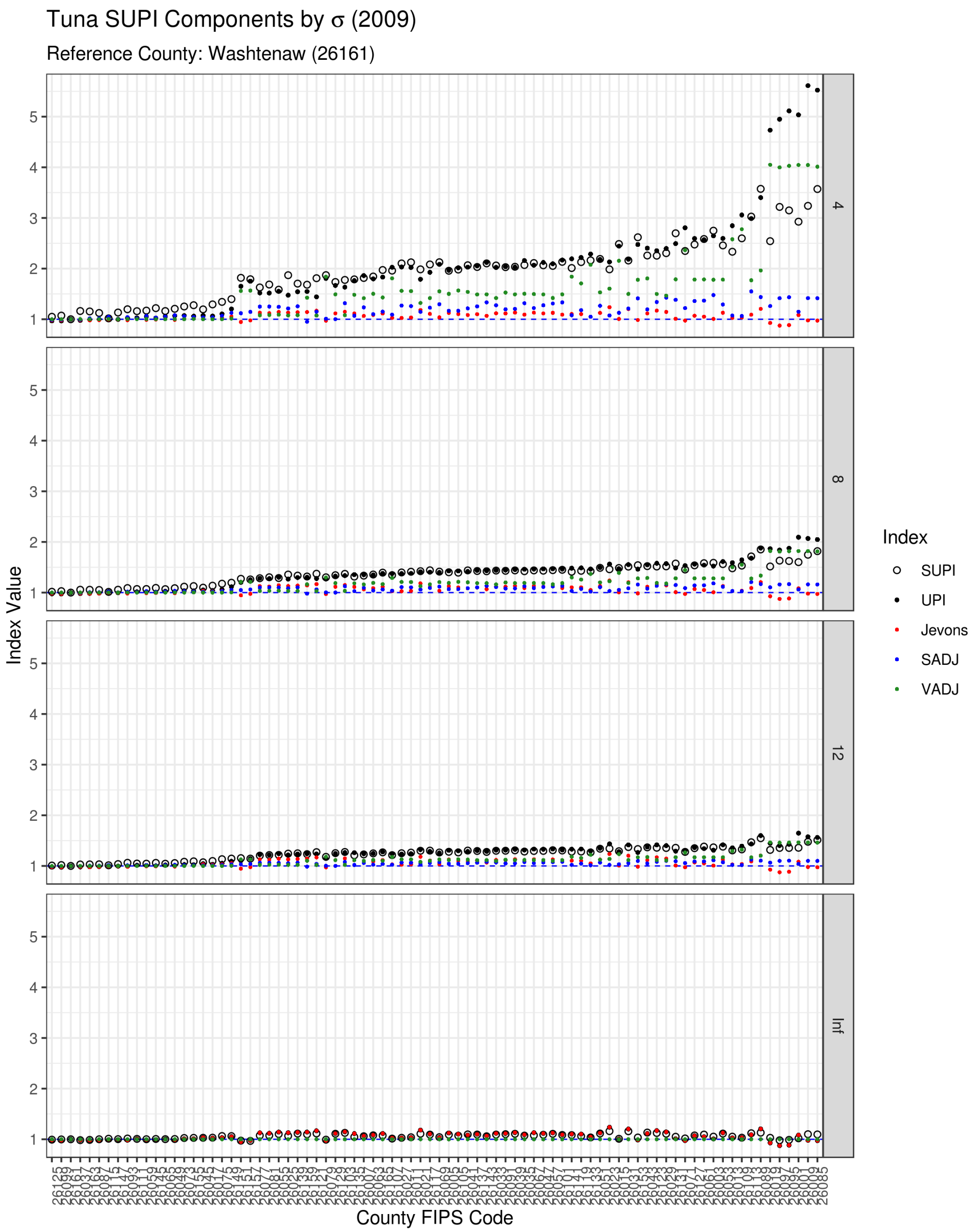

Figure 5.13: Tuna SUPI, UPI, Jevons, SADJ, and VADJ by County and σ (2009)

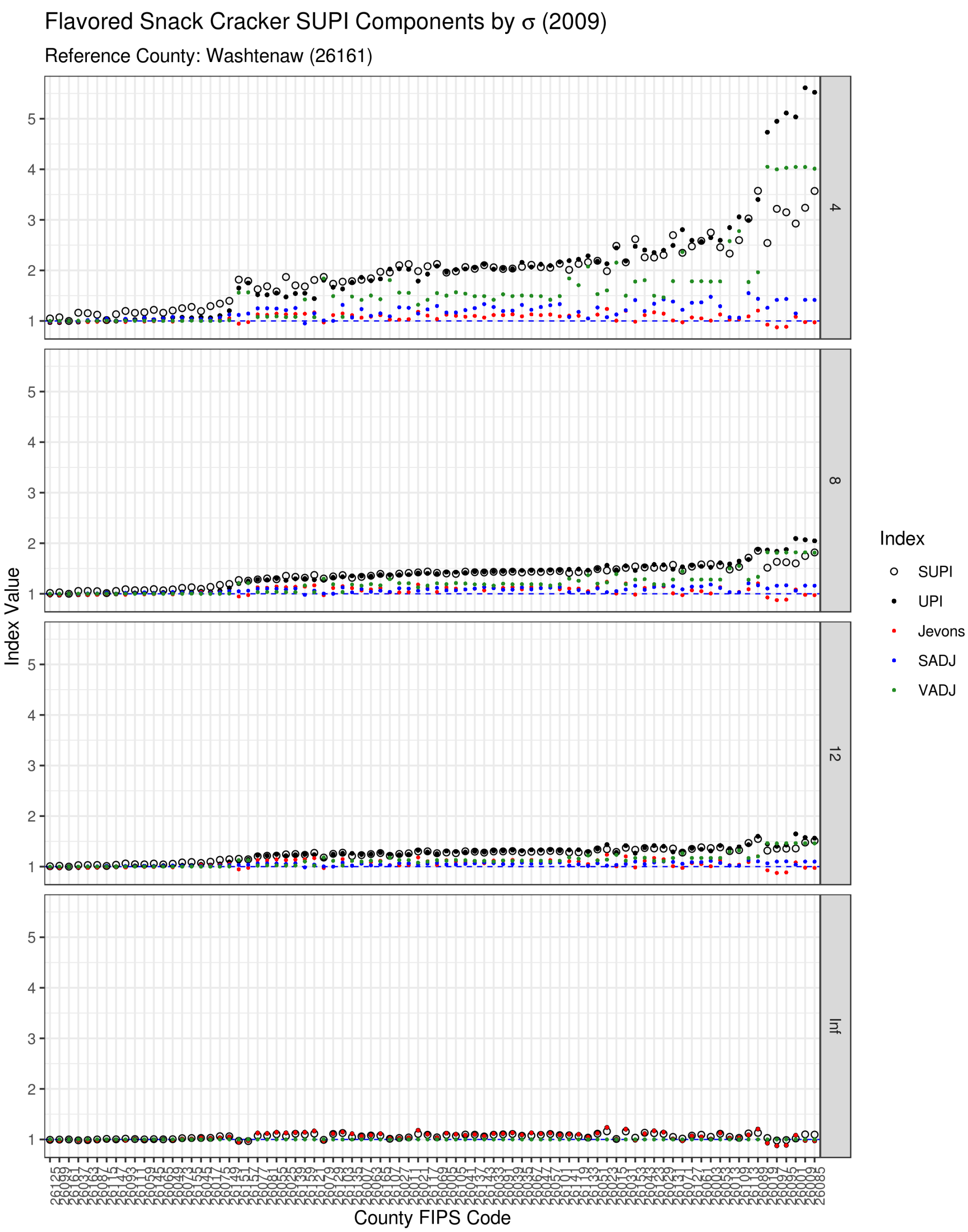

Figure 5.14: Flavored Snack Cracker SUPI, UPI, Jevons, SADJ, and VADJ by County and σ (2009)

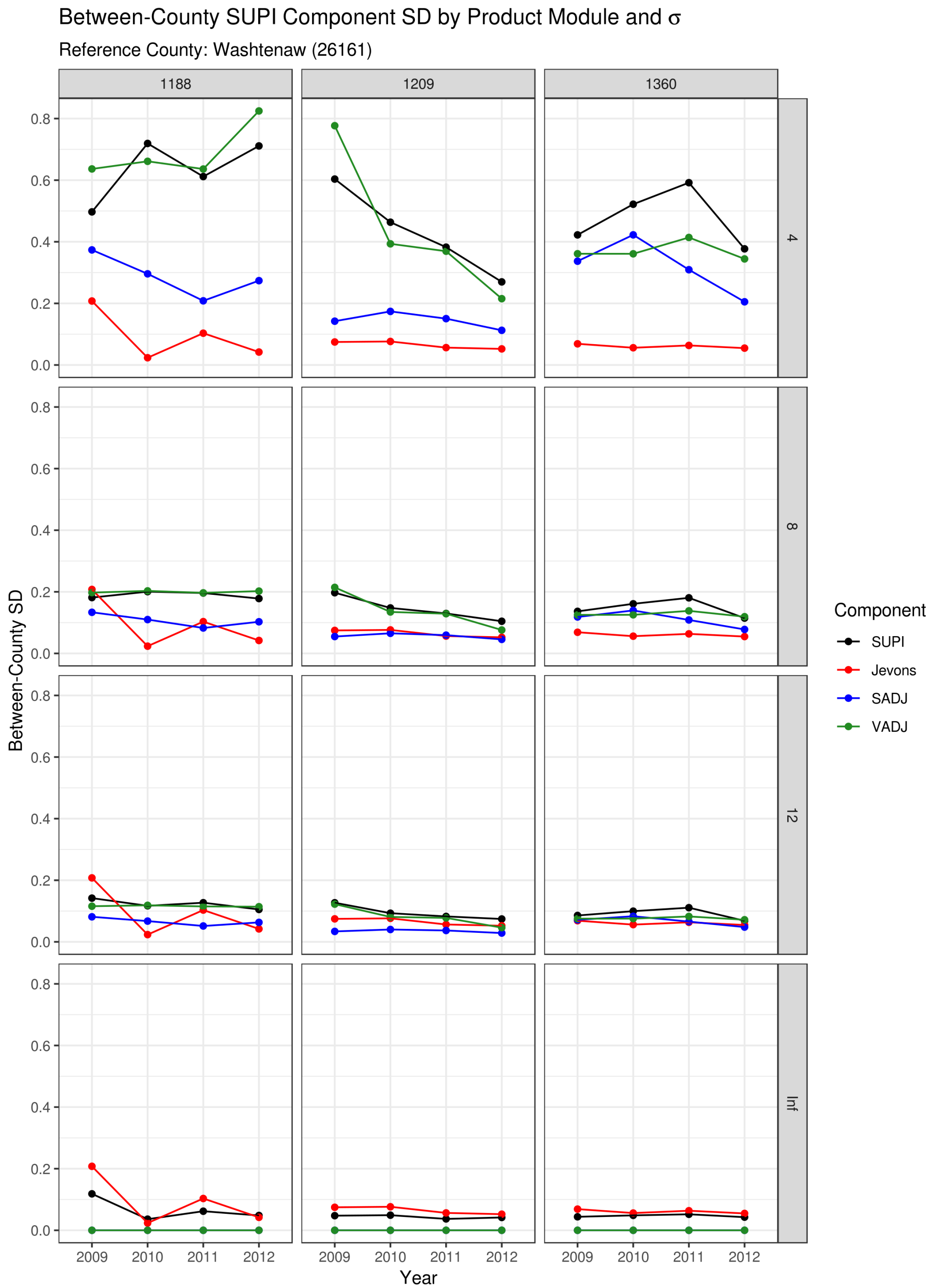

Figure 5.15: Between-County SUPI Standard Deviations by Product Module and σ

Table 5.5: Cross-County Vinegar SUPI Standard Deviations by Year and σ

	2	4	6	8	10	12	Inf
2009	9.26	0.50	0.25	0.18	0.15	0.14	0.12
2010	15.93	0.72	0.32	0.20	0.15	0.12	0.04
2011	11.34	0.61	0.29	0.20	0.15	0.13	0.06
2012	41.55	0.71	0.29	0.18	0.13	0.11	0.05

Reference County = Washtenaw (26161)

Table 5.6: Cross-County Tuna SUPI Standard Deviations by Year and σ

	2	4	6	8	10	12	Inf
2009	8.14	0.60	0.29	0.20	0.15	0.13	0.05
2010	5.12	0.46	0.22	0.15	0.11	0.09	0.05
2011	3.68	0.38	0.19	0.13	0.10	0.08	0.04
2012	2.82	0.27	0.14	0.10	0.09	0.07	0.04

Reference County = Washtenaw (26161)

Table 5.7: Cross-County Flavored Snack Cracker SUPI Standard Deviations by Year and σ

	2	4	6	8	10	12	Inf
2009	4.17	0.42	0.21	0.14	0.10	0.09	0.04
2010	6.20	0.52	0.25	0.16	0.12	0.10	0.05
2011	7.69	0.59	0.28	0.18	0.14	0.11	0.05
2012	3.93	0.38	0.18	0.11	0.08	0.07	0.04

Reference County = Washtenaw (26161)

This plot shows that for values of the elasticity of substitution $\sigma < \infty$, the between-county tuna SUPI variance declines monotonically across time. The magnitude of this decline depends on the assumed value of σ. If we assume that $\sigma = 4$, the between-county tuna SUPI standard deviation more than halves, from about 0.60 in 2009 to around 0.27 in 2012. If we instead assume that $\sigma = \infty$, there is effectively no change in this standard deviation at all. We can see that the decline in SUPI standard deviation tracks the decline in the VADJ standard deviation almost perfectly. This suggests that the apparent decline is driven primarily by the convergence of each county's tuna variety adjustment terms across time. The pattern is less clear for flavored snack crackers and vinegar. For a given value of σ, the estimated between-county standard deviations for these product modules fluctuate near their initial 2009 values without a consistent pattern across time, and without an obvious relationship to any single UPI subcomponent. Nevertheless, we can see that the estimated between-county variation is consistently higher when lower values of σ are assumed.

While we do not propose any method for estimating σ in this paper, we think that σ should be relatively large due to the similarity of the product varieties within each product module. Under this assumption, the extreme UPI values generated by the cases in which $\sigma = 2$ or $\sigma = 4$ are more useful for illustrative purposes than as true reflections of the cost of living. Regardless of what one assumes however, these results show the potential for significant biases when estimating the cost of living using traditional price indices such as the Jevons, which do not account for differences in consumer preferences between the comparison areas, or differences in the stock of product varieties that are available in each county.

5.1.3 Bootstrap Confidence Intervals

Up to this point, we've been examining the properties of fixed point estimates for the Vinegar, Tuna, and Flavored Snack Cracker SUPIs based on Nielsen's retail scanner data. However many of these estimates are based on only small numbers of observed product varieties within each area, and are highly sensitive to which product varieties are selected in the sample. In order to account for the effect of variation in the estimated area prices and particular product varieties selected on our SUPI estimates, we generated 100 bootstrap replicates of the data within each of the product modules, as described in Section 4.3.1. The results of this process for the product modules in Table 5.1 in 2009 are displayed in Figures 5.16, 5.17 and 5.18 for values of $\sigma \in \{6, 8, 10, 12, \infty\}$.

In these graphs, the black dot represents the geometric mean of the bootstrap

SUPI replications. The red dot is the SUPI estimate from the observed sample, and the blue line shows unity on the y-axis, i.e. the point where the comparsion county has the same cost of living as the reference county. The black error bars represent a confidence interval around the bootstrap means. Concretely, let m and s denote the bootstrap mean and standard deviation of $log(SUPI_{a_1a_2})$. The confidence interval for $log(SUPI_{a_1a_2})$ is then

$$(m - 2s, m + 2s) \tag{5.2}$$

Hence the confidence interval for the geometric mean SUPI value is

$$(e^{m-2s}, e^{m+2s}) \tag{5.3}$$

Note that although the choice to add and subtract two geometric standard deviations makes these bounds analagous to normal confidence intervals, they do not guarantee any particular coverage level unless SUPIs are assumed to be log-normally distributed. We do not assume that our SUPIs have any particular distribution, and producing reliable nonparametric confidence intervals would require significantly more than 100 bootstrap replications. Because producing this number of replications is computationally intensive, we restrict ourselves to estimates of the raw variability of our indices. Hence these error bounds reflect the variability of our estimates, but should not necessarily be interpreted as 95% confidence intervals.

A cursory glance at Figures 5.16, 5.17 and 5.18 shows that the observed SUPIs (red dots) are generally close to the geometric mean of the bootstrap replications (black dots). Bootstrap estimates of bias are generally close to zero, indicating that any bias in our SUPI estimator is negligible. We also note that several of our index estimates at the level of the individual product module are quite variable. At lower values of σ, the geometric standard deviation intervals for the most variable tuna SUPIs include values as high as about 3.75, and as low as about 1.25. In each of our

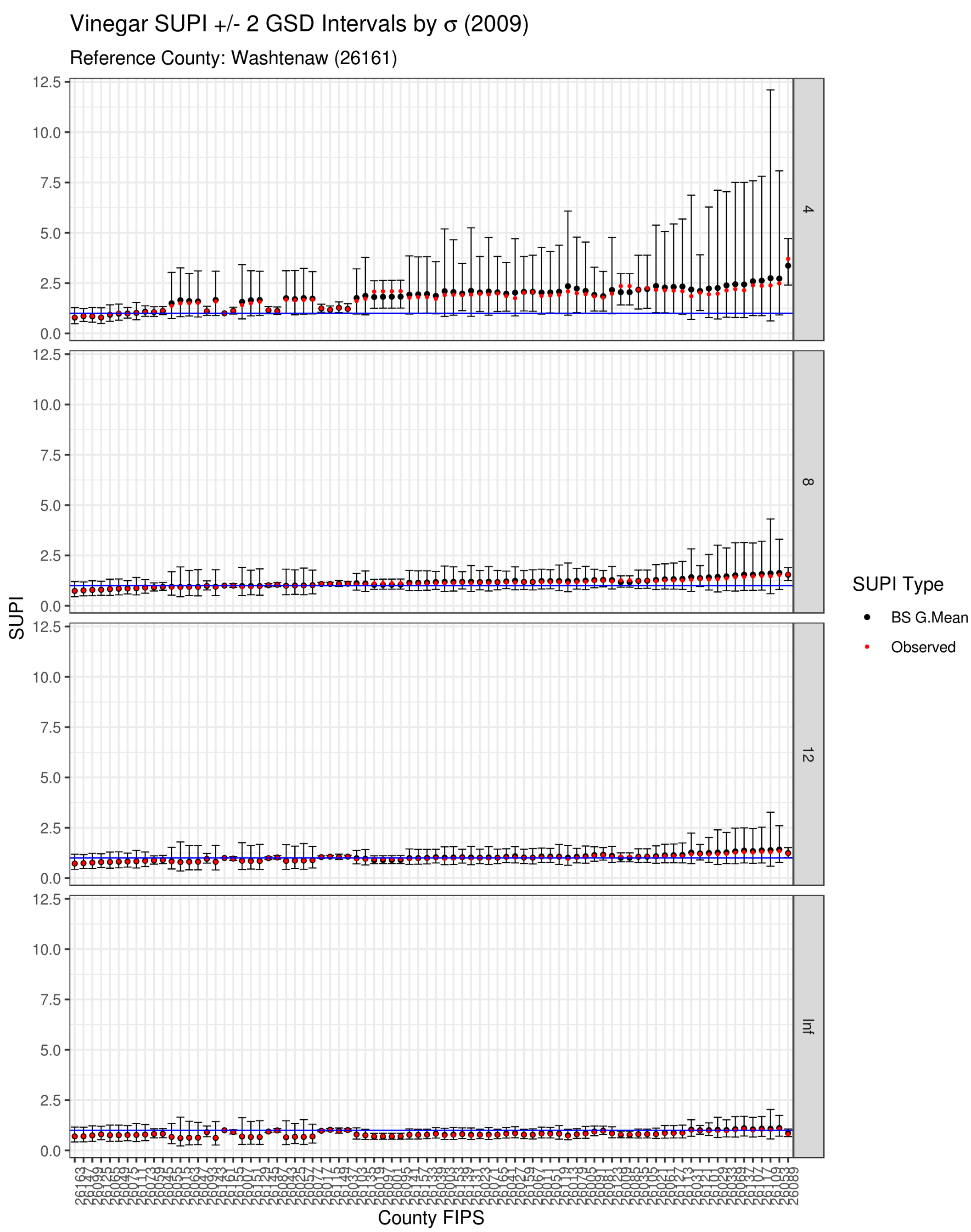

Figure 5.16: Vinegar SUPI Confidence Intervals (2009)

sampled product modules, estimates associated with smaller values of σ are generally more variable than estimates with larger values of σ. This makes intuitive sense, as variability in which product varieties are included in our sample will impact the value of our SUPIs less when products are assumed to be perfect substitutes than when the

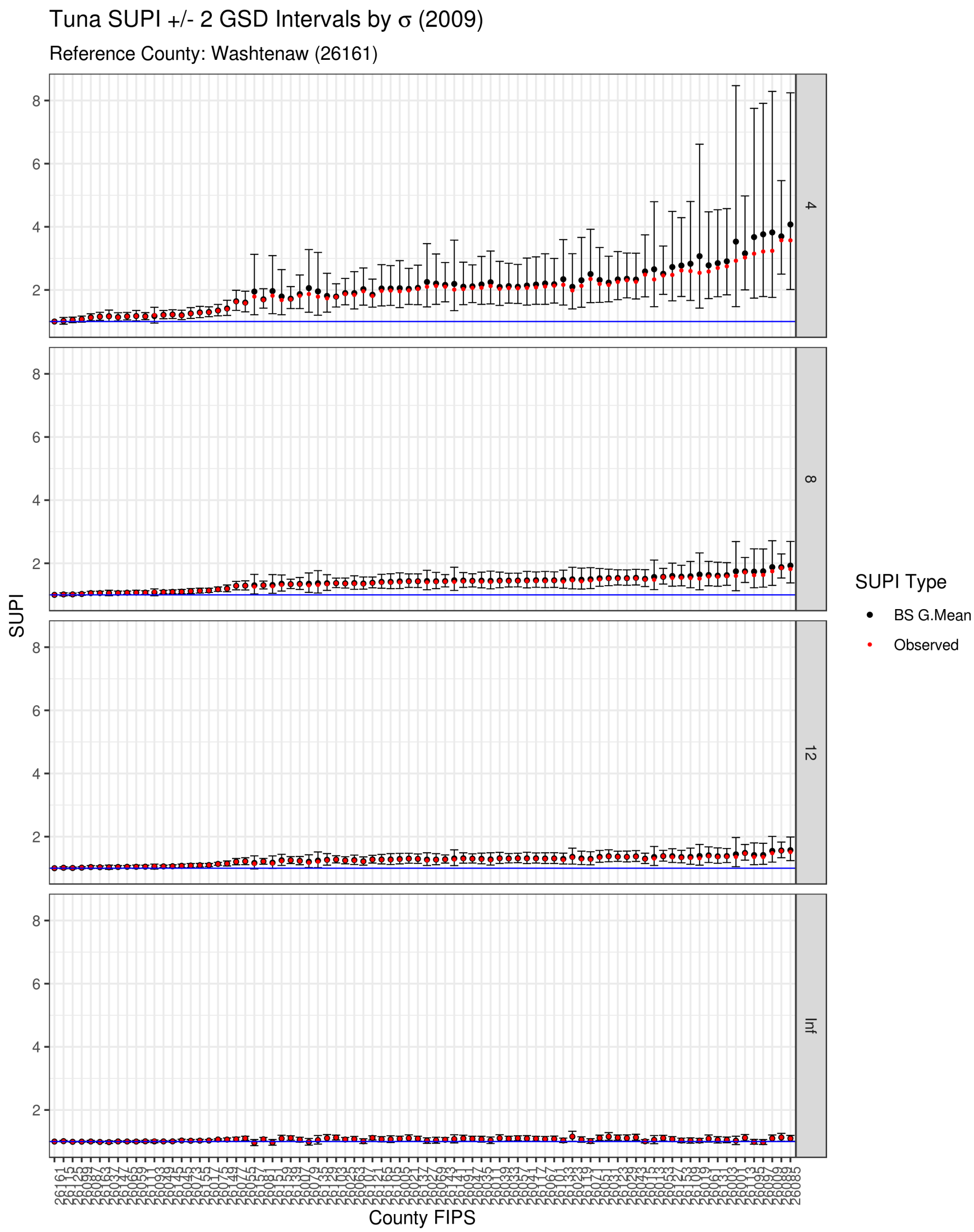

Figure 5.17: Tuna SUPI Confidence Intervals (2009)

varieties that are available for purchase matters to consumers.

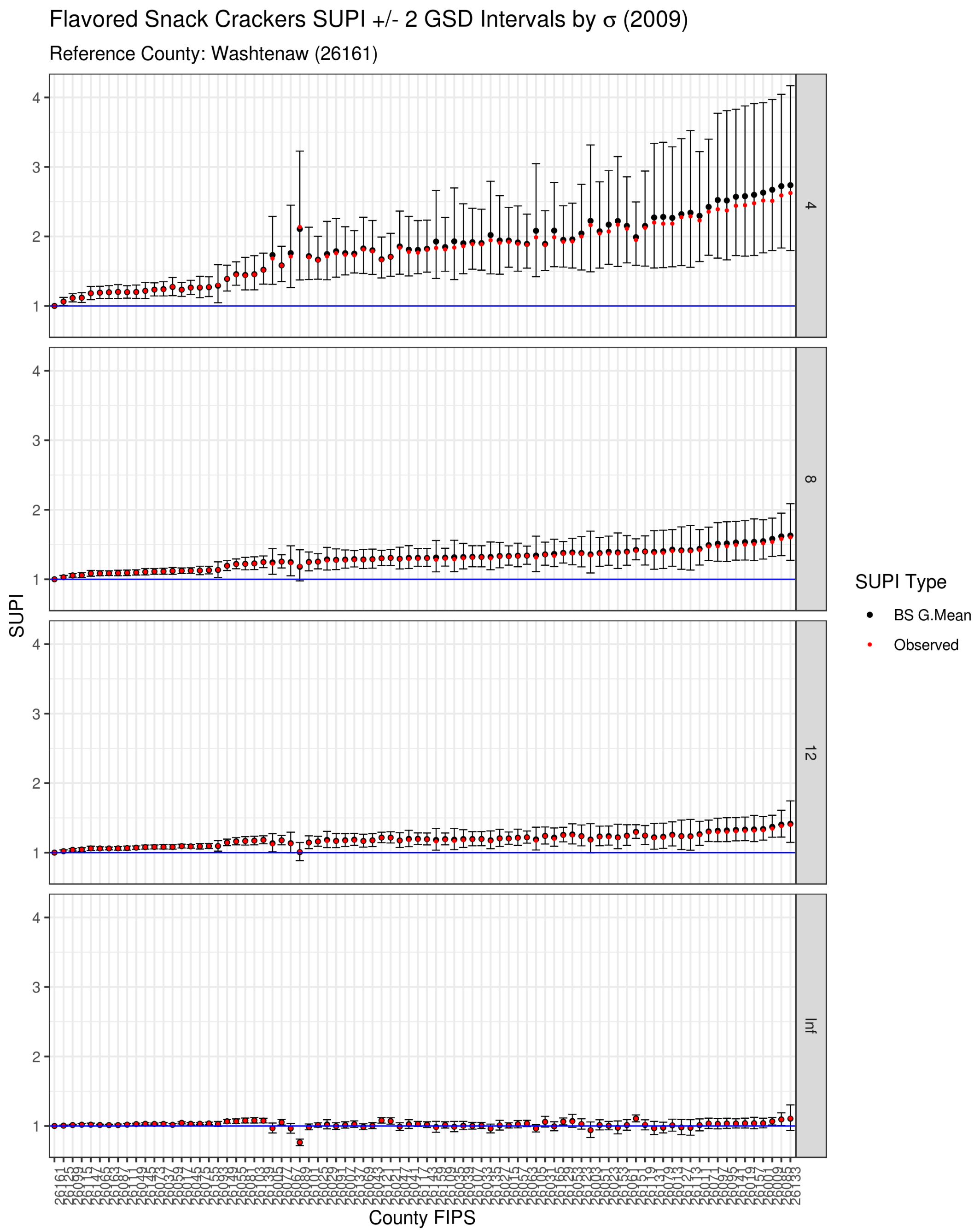

Figure 5.18: Flavored Snack Cracker Confidence Intervals (2009)

5.2 Food CSUPIs

In this section, we examine the relationship of the category-level SUPI estimates for all food to the average values of the product module SUPIs, UPIs, and component subindices. The food CSUPIs for 2009 - 2012 are summarized in Table 5.8 below.

The unabridged values of the food CSUPIs for 2009 - 2012 are included in Tables B.4,

Table 5.8: Food CSUPI Five Number Summaries by σ

σ	Year	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
2	2009	1.00	2.47	4.09	4.57	5.37	15.43
2	2010	1.00	2.73	4.61	5.00	5.98	15.61
2	2011	1.00	2.64	4.34	4.56	5.57	14.24
2	2012	1.00	2.27	3.84	3.95	4.63	12.95
4	2009	1.00	1.39	1.64	1.62	1.80	2.61
4	2010	1.00	1.43	1.69	1.66	1.85	2.60
4	2011	1.00	1.41	1.66	1.62	1.81	2.41
4	2012	1.00	1.34	1.59	1.55	1.69	2.27
6	2009	1.00	1.24	1.36	1.35	1.44	1.83
6	2010	1.00	1.26	1.38	1.37	1.46	1.82
6	2011	1.00	1.24	1.36	1.35	1.45	1.73
6	2012	1.00	1.20	1.33	1.31	1.38	1.65
8	2009	1.00	1.18	1.26	1.25	1.31	1.57
8	2010	1.00	1.19	1.27	1.26	1.32	1.56
8	2011	1.00	1.18	1.26	1.24	1.31	1.50
8	2012	1.00	1.15	1.23	1.22	1.27	1.45
10	2009	1.00	1.15	1.21	1.20	1.25	1.45
10	2010	1.00	1.15	1.22	1.20	1.25	1.43
10	2011	1.00	1.14	1.20	1.19	1.24	1.39
10	2012	1.00	1.12	1.18	1.17	1.21	1.36
12	2009	1.00	1.13	1.18	1.17	1.21	1.37
12	2010	1.00	1.13	1.18	1.17	1.21	1.35
12	2011	1.00	1.12	1.17	1.16	1.20	1.32
12	2012	1.00	1.10	1.15	1.14	1.18	1.30
Inf	2009	0.98	1.03	1.04	1.04	1.05	1.08
Inf	2010	0.96	1.02	1.03	1.03	1.04	1.06
Inf	2011	0.95	1.02	1.03	1.03	1.04	1.06
Inf	2012	0.95	1.01	1.02	1.02	1.03	1.06

Reference County: Washtenaw (FIPS 26161)

B.5, B.6, and B.7 respectively. These values are mapped in Figure 5.19. A version of this map based in a geometric average of all in-sample Michigan counties, rather than Washtenaw county, is included in Appendix A.

From these maps, we can see that for each given value of σ, the estimated differences in cost of living remain relatively stable across the years we study. The range of these indices is slightly smaller than the range of the SUPIs for individual product

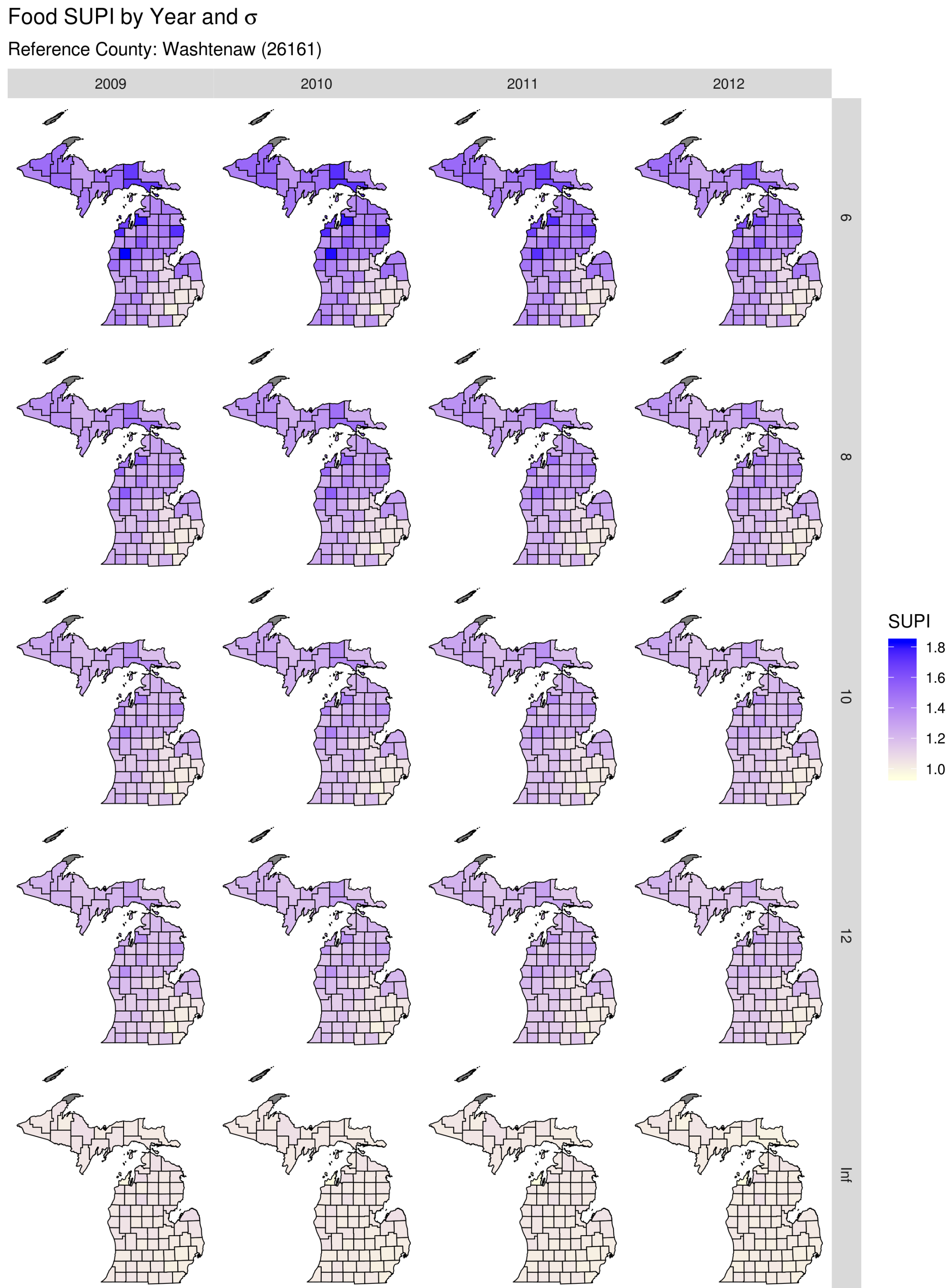

Figure 5.19: Food CSUPI Choropleth

modules, with the highest estimates in the sampled product modules being on the order of 2.5, whereas the highest food CSUPI estimates are around 1.8. Despite this, the cost of living still appears to be lower within a cluster surrounding the Detroit Metropolitan Area, particularly for lower values of σ. This pattern appears to be quite durable. We see it each year for the product modules listed in Table 5.1 in Figures 5.1, 5.2, 5.3 and 5.4, as well as in the aggregate measures in Figure 5.19.

5.2.1 Interpreting Food CSUPI Components

Because the food CSUPI is constructed as the geometric mean of a large number of product module SUPIs, there is no single set of UPIs or UPI components we can look at to analyze the resulting values. Instead, because each of the component SUPIs is estimated from a set of product module UPIs, we take the weighted geometric average of the resulting UPI indices and their component indices, and analyze the relationship between these values and the food CSUPIs. The food CSUPIs, along with the weighted geometric means of the product level UPIs, Jevons indices, SADJ terms and VADJ terms are plotted in Figures 5.20, 5.21, 5.22, and 5.23.

These plots are analogous to the plots in Section 5.1.2, and they tell a similar story to the plots in that section. On the aggregate level, the VADJ component of the UPI is primarily responsible for our SUPI estimates being larger than more traditional indices such as the Jevons, although the SADJ component also contributes to this outcome.

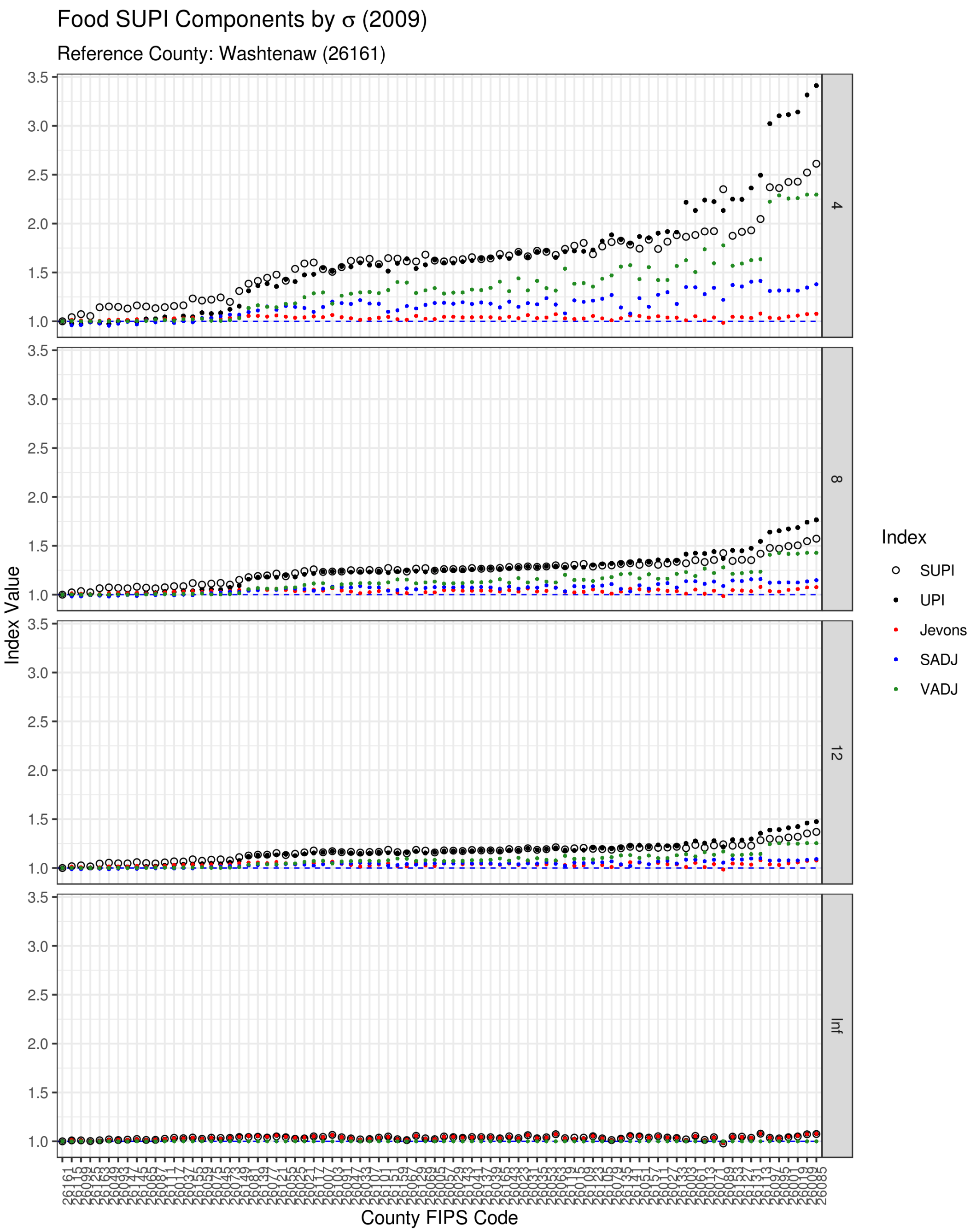

Figure 5.20: Food CSUPI Components (2009)

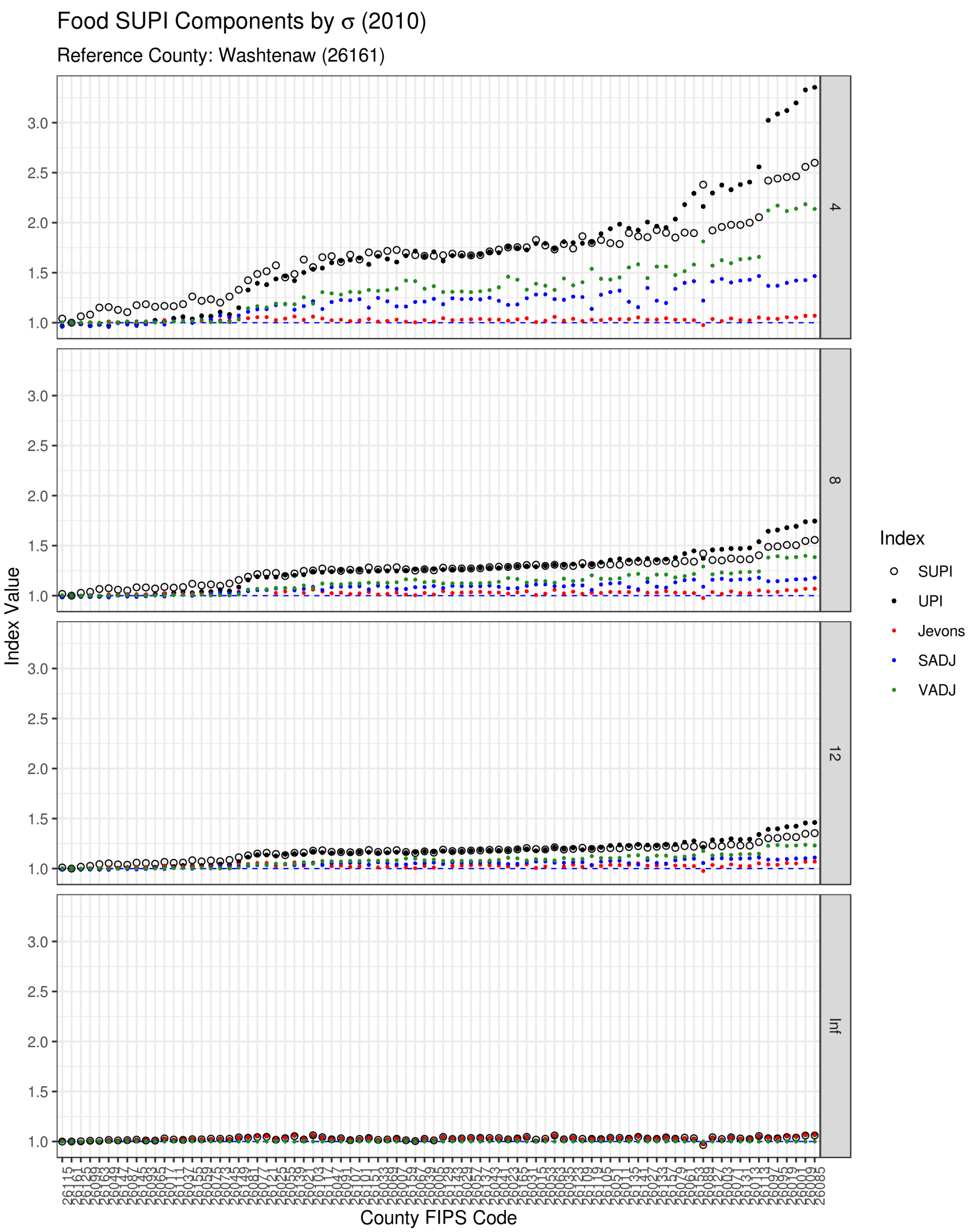

Figure 5.21: Food CSUPI Components (2010)

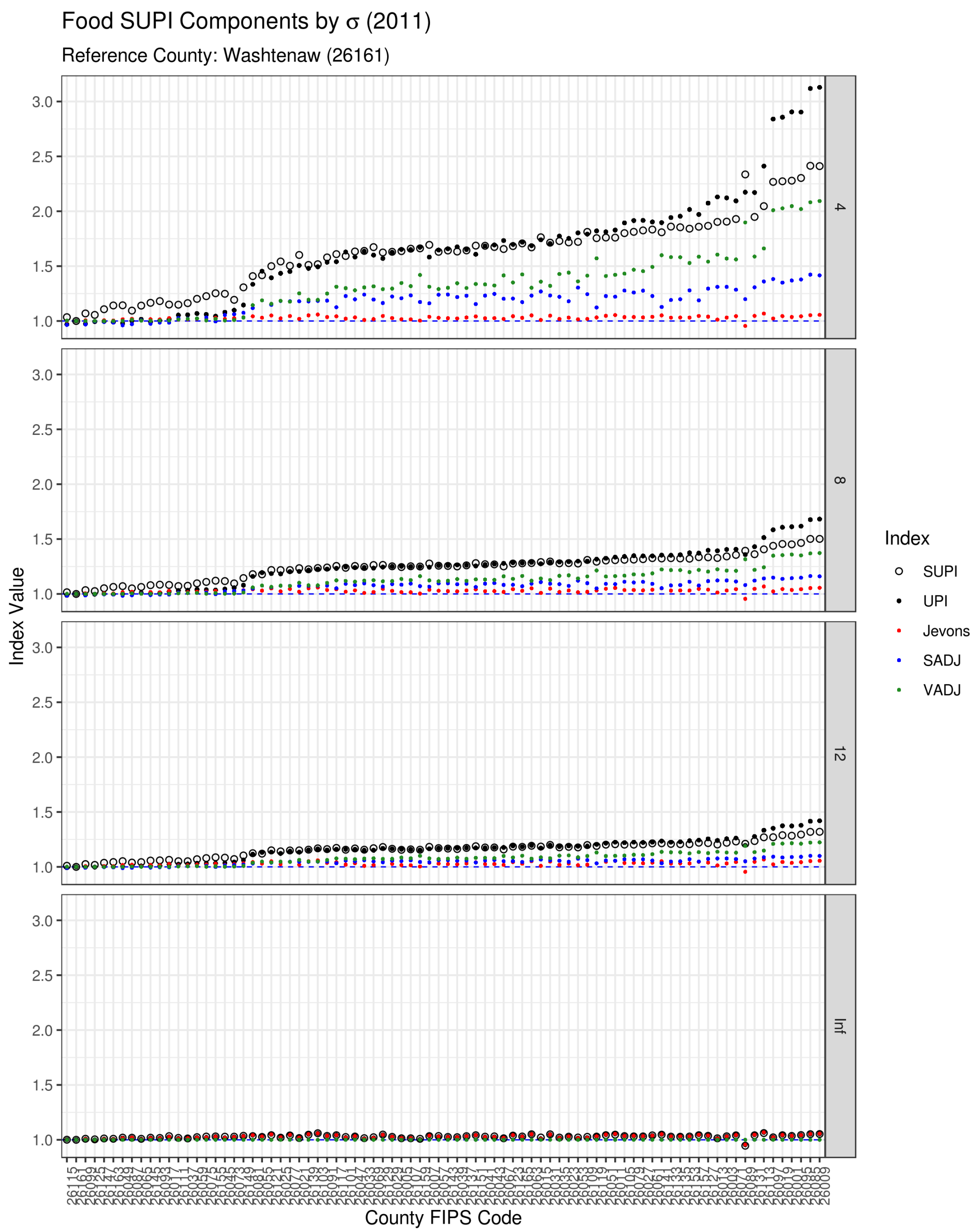

Figure 5.22: Food CSUPI Components (2011)

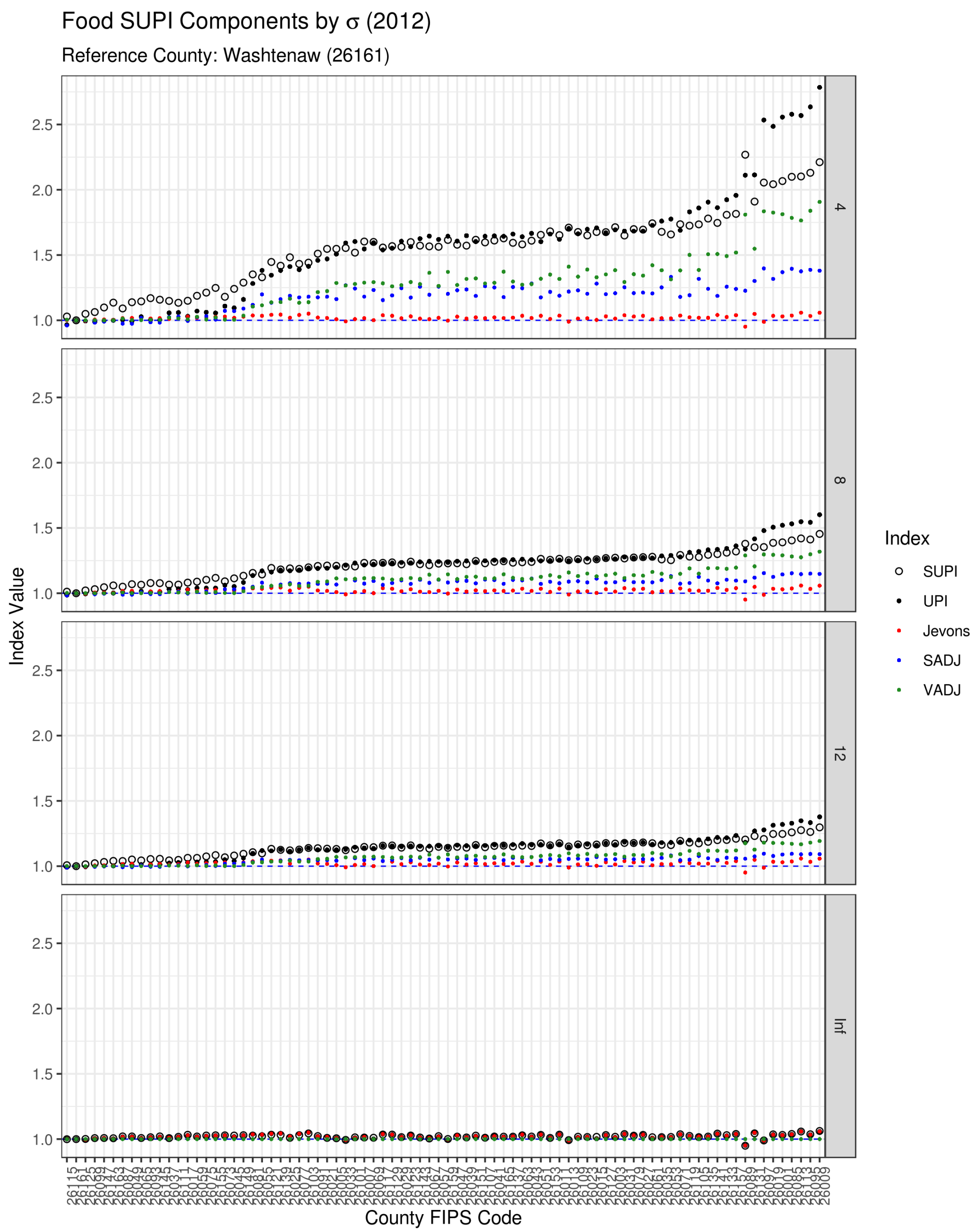

Figure 5.23: Food CSUPI Components (2012)

We also note that the pattern of estimated SUPI values remains similar from year to year. Figure 5.24 shows that our food CSUPI estimates don't change much over time, particularly at high values of σ. At lower values of σ, there is a bit more flux

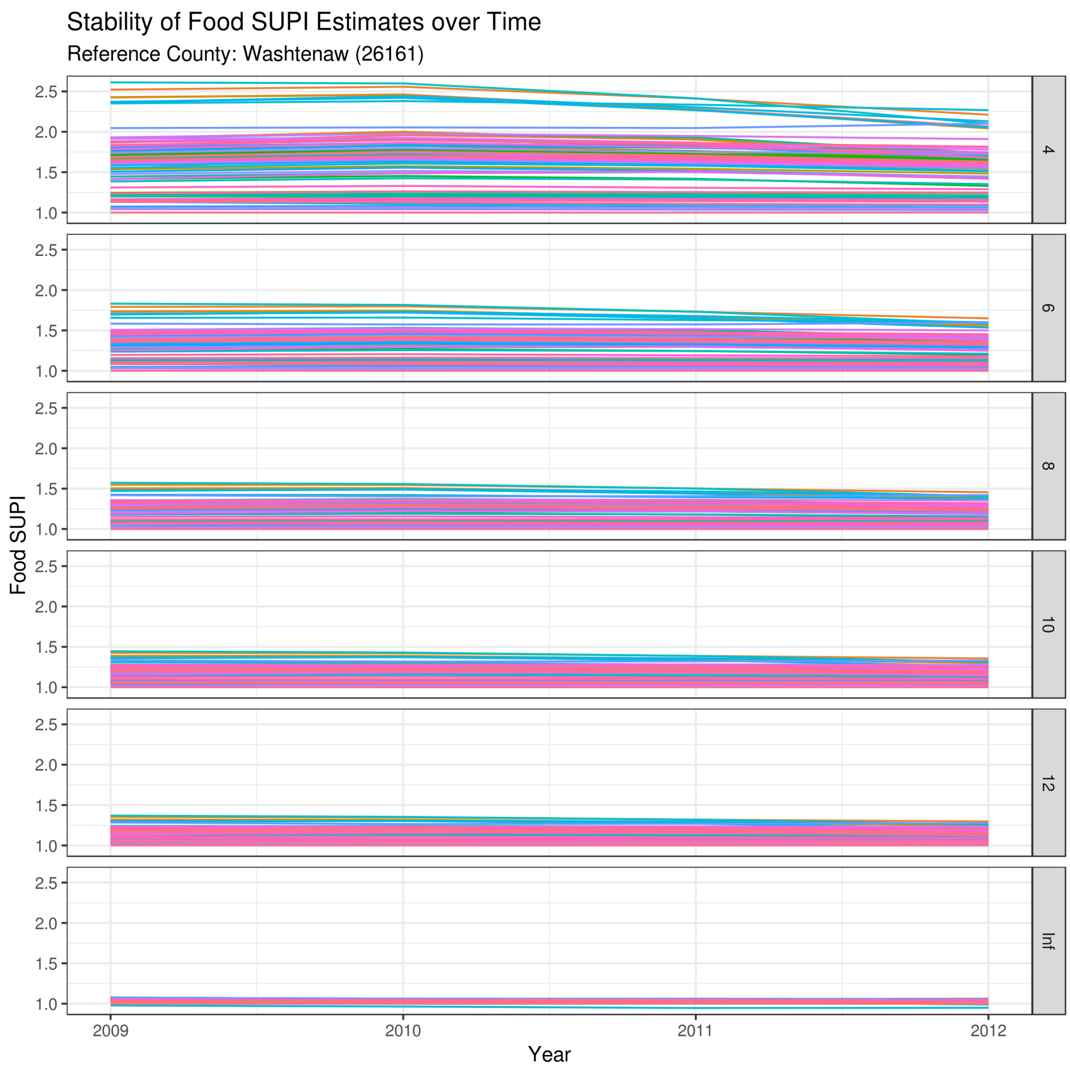

Figure 5.24: Stability of Food CSUPI Estimates over Time by σ

in our estimated values over time. Several counties with high SUPI estimates seem to fall closer to the mean as time goes on when $\sigma < \infty$. Figure 5.25 shows that the cross county variability in Food CSUPIs declines between 2009 and 2012, presumably due to this convergence. This decline appears to be correlated with the variance in the average cross county in the VADJ term, as the standard deviation of the average

SADJ and Jevons remain mostly flat throughout this period.

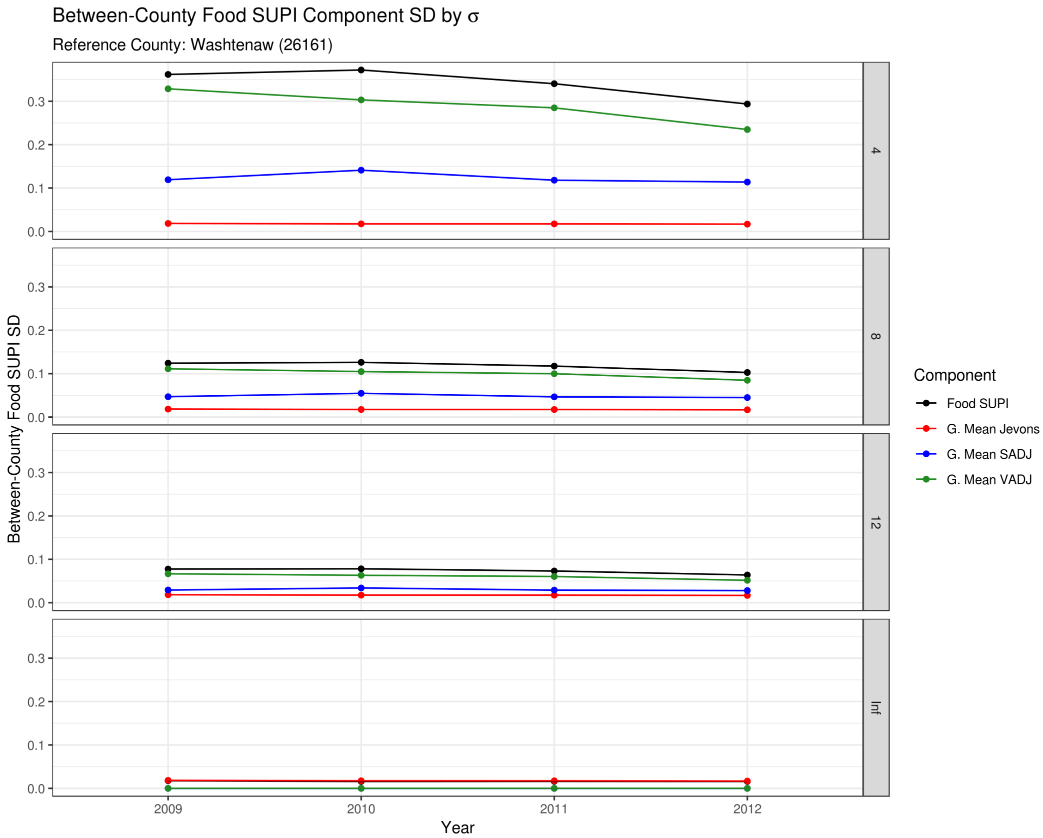

Figure 5.25: Between-County Food CSUPI Standard Deviations by σ

Table 5.9: Cross-County Food CSUPI Standard Deviations by Year and σ

	2	4	6	8	10	12	Inf
2009	3.18	0.36	0.18	0.12	0.09	0.08	0.02
2010	3.37	0.37	0.19	0.13	0.10	0.08	0.02
2011	2.84	0.34	0.17	0.12	0.09	0.07	0.02
2012	2.22	0.29	0.15	0.10	0.08	0.06	0.02

Reference County = Washtenaw (26161)

This may be due to changes in the availability of product varieties across time, or due to improvements in sample coverage rates in those counties. Alternatively, this pattern might be caused by highly variable SUPI estimates for a few outlying counties.

5.2.2 Bootstrap Confidence Intervals

In this section, we examine the uncertainty associated with our food CSUPIs. This uncertainty is visualized in Figures 5.26, 5.27, 5.28 and 5.29. These figures show the geometric mean of our food CSUPI replicates as black dots, the food CSUPIs estimated from the observed samples as red dots, and two geometric standard deviation intervals around the geometric mean of the replicates as black error bars. These plots are produced based on the data in Tables B.4, B.5, B.6, and B.7, and interpreted analogously to the confidence intervals that we discussed in Section 5.1.3.

As was the case for our estimates in that section, the bootstrap mean of the food CSUPI replicates concides almost perfectly with the food CSUPI estimates calculated from the sample. However, the bootstrap geometric standard deviation intervals for the food CSUPIs are much smaller than the analogous intervals were for the individual product module SUPIs.

Figure 5.30 shows these same values as Figures 5.26, 5.27, 5.28 and 5.29 on a common scale for three values of σ, so that the uncertainty and ordering of these estimates can be more easily compared across time. From this figure, we can see that both the estimated index values and the estimated uncertainty are roughly consistent with across the years that we consider.

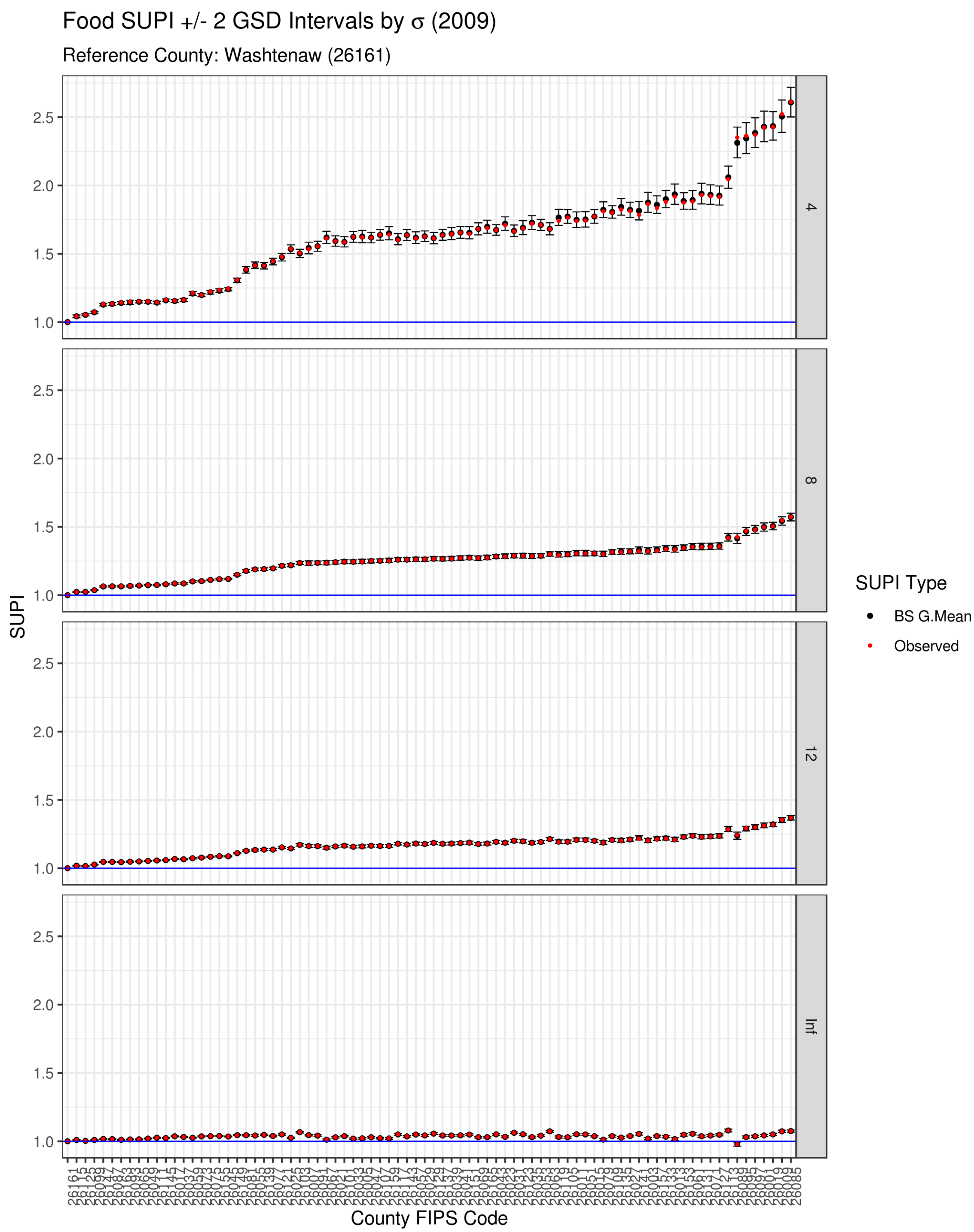

Figure 5.26: Food CSUPI Uncertainty Intervals (2009)

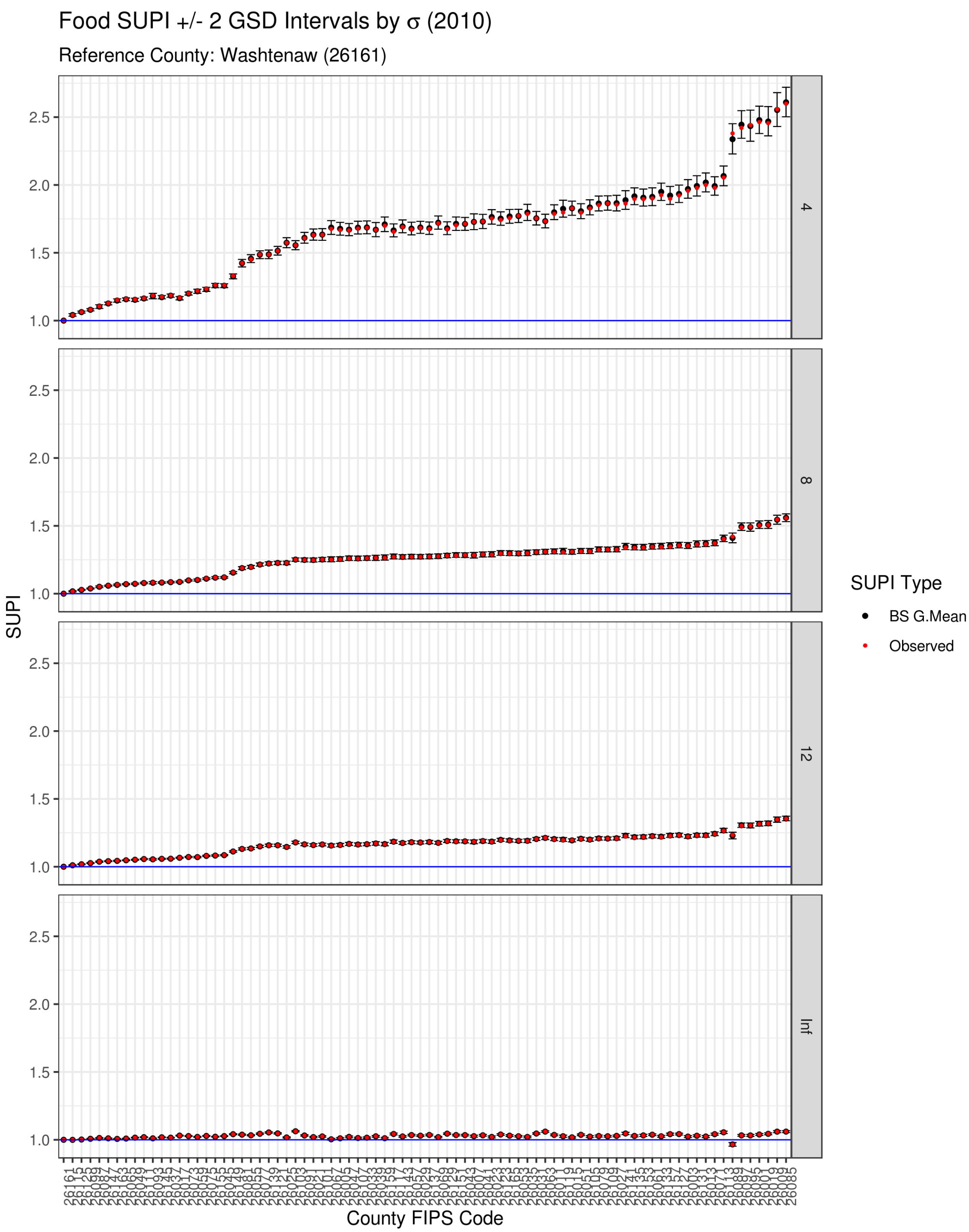

Figure 5.27: Food CSUPI Uncertainty Intervals (2010)

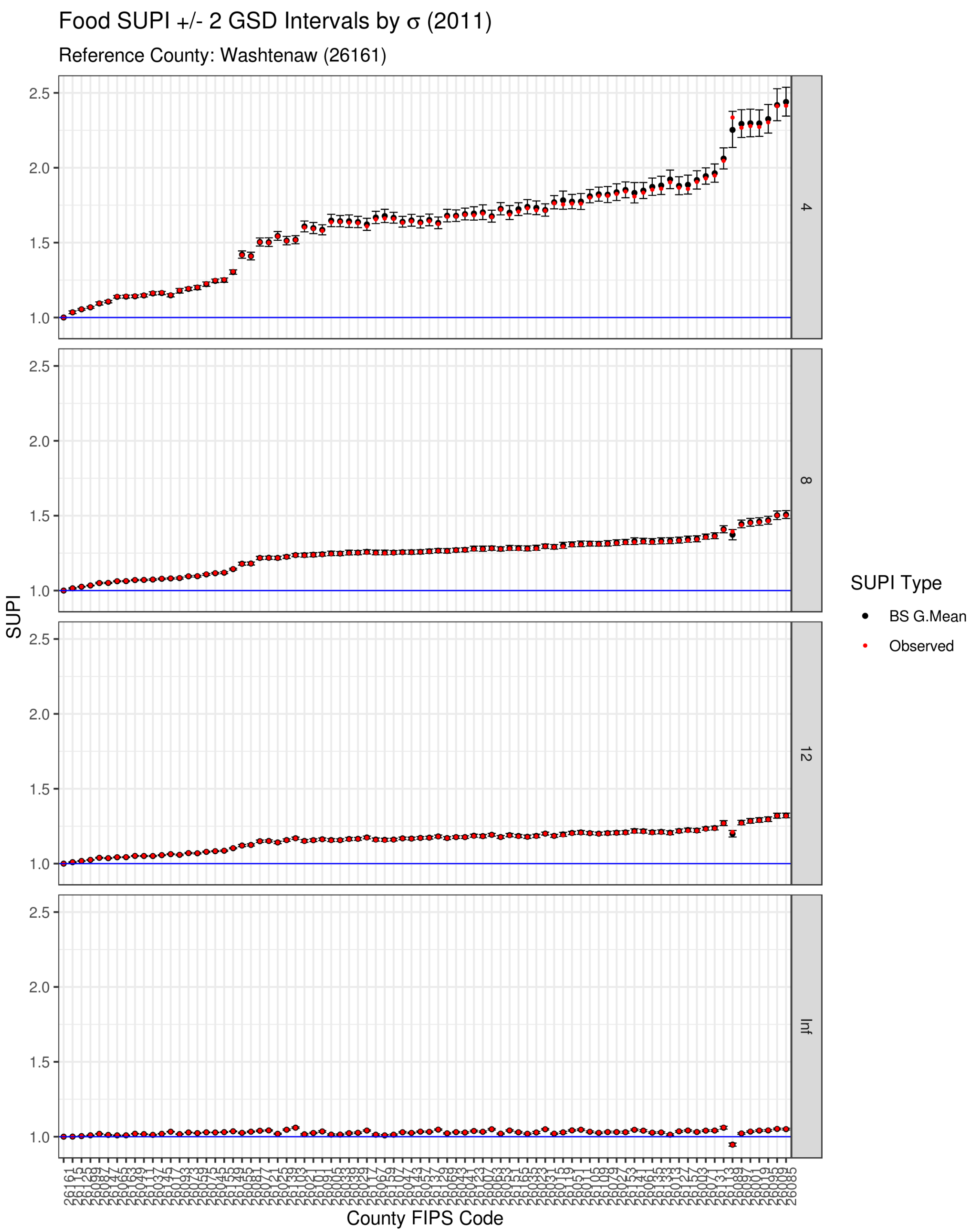

Figure 5.28: Food CSUPI Uncertainty Intervals (2011)

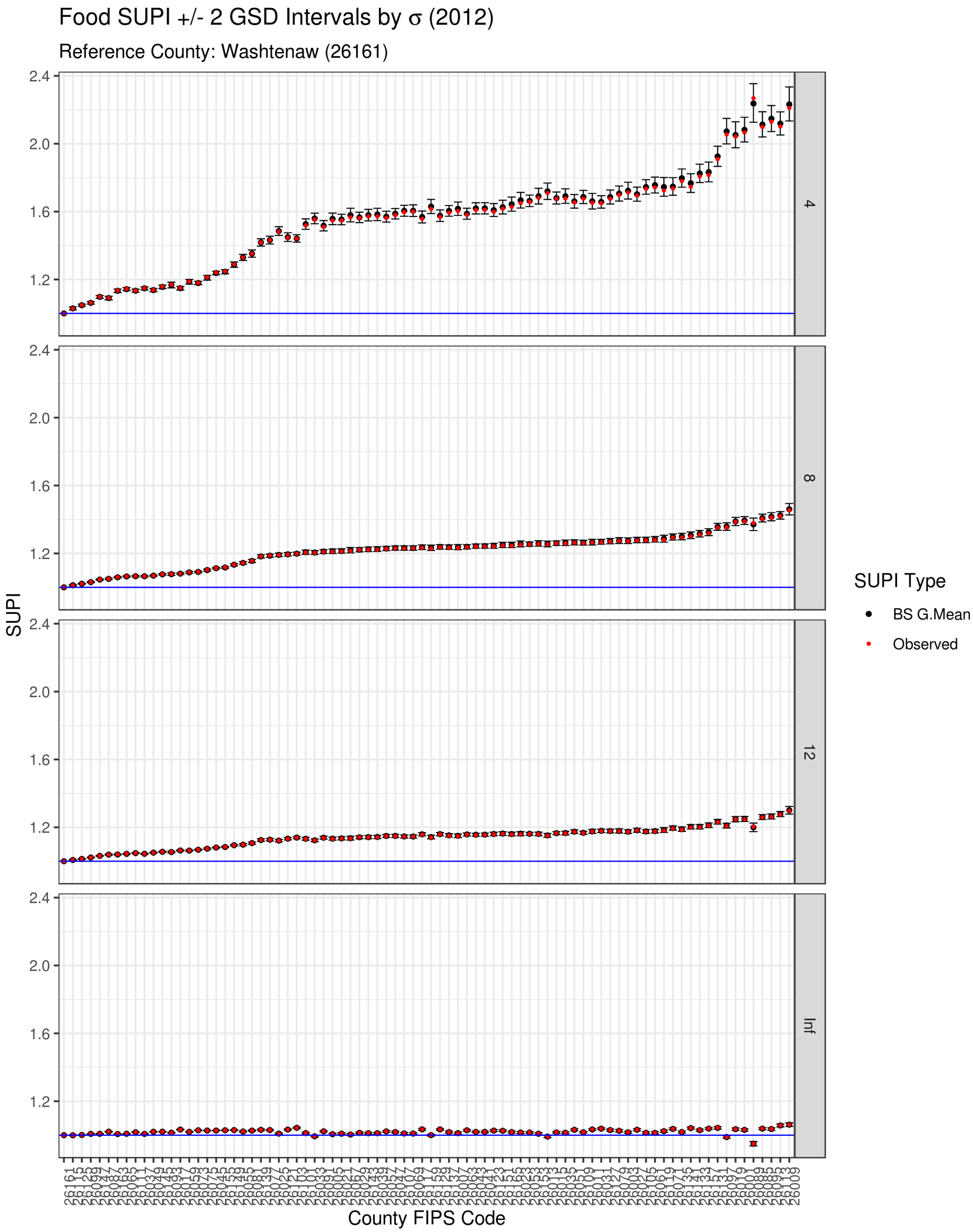

Figure 5.29: Food CSUPI Uncertainty Intervals (2012)

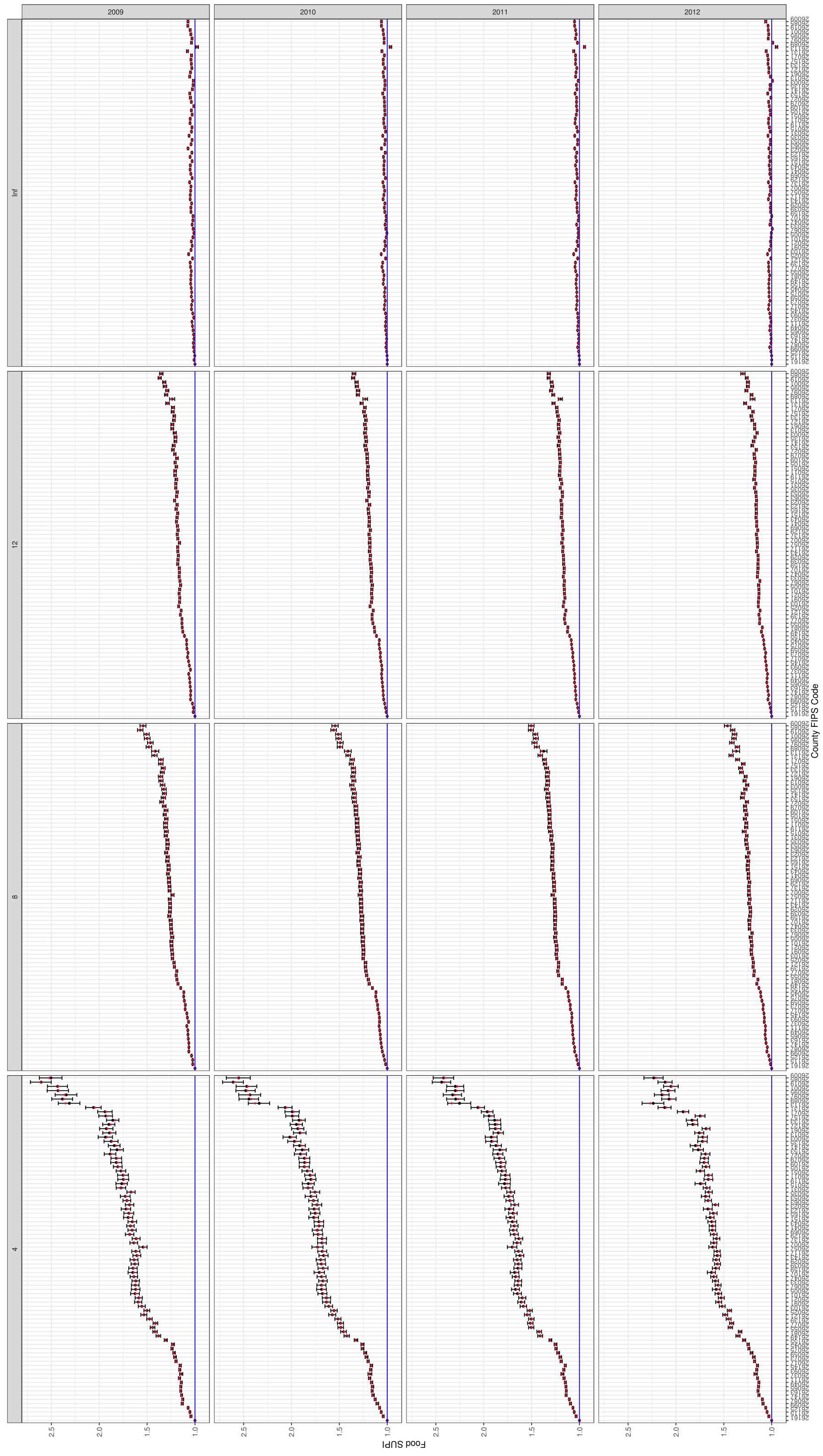

Figure 5.30: Food CSUPI Uncertainty Intervals by σ and Year

5.3 Analyzing Bootstrap Replication Failures

We can examine the number of replication failures to gain context for the comparisons in Figures 5.16, 5.17 and 5.18. For example, we can see from Figure 5.16 that the SUPI comparing vinegar in Leelanau (FIPS 26089) and Washtenaw counties has a lower estimated variance than counties with similar point estimates. 43% of the attempted bootstrap replications for this SUPI failed. This hints that our ability to estimate this index depends heavily on a small number of product varieties being included in the resample. The scatterplot in Figure 5.31 confirms this intuition, showing that in product module 1188 (vinegar), areas with low numbers of UPCs tend to have high numbers of replication failures.

This relationship looks cleaner and less variable than other potential drivers of replication failures. For example, Figure 5.32 shows the relationship in the sampled product modules between the average number of stores within an area and the average number of replication failures, and Figure 5.33 shows the relationship in the sampled product modules between the average number of in-sample price quotes across all product varieties and the average number of replication failures within that county.

Figures 5.34, 5.35 and 5.36 show that these relationships are more general. The average number of replication failures and the average number of product varieties within the associated areas are consistently related across time. To an even greater degree than we see in Figures 5.31, 5.32 and 5.33, the relationship between bootstrap replication failures and the number of observed UPCs seems clearer and more consistent than the relationships between replication failures and the other two variables.

There are two possible interpretations for why an area would have a high replication failure rate. The first is that the replication failures are caused by poor data quality in the affected areas. Under this interpretation, the high rate of replication failures indicates that our sample size is likely too small to accurately estimate the variability associated with our indices for this area, and that therefore the variabil-

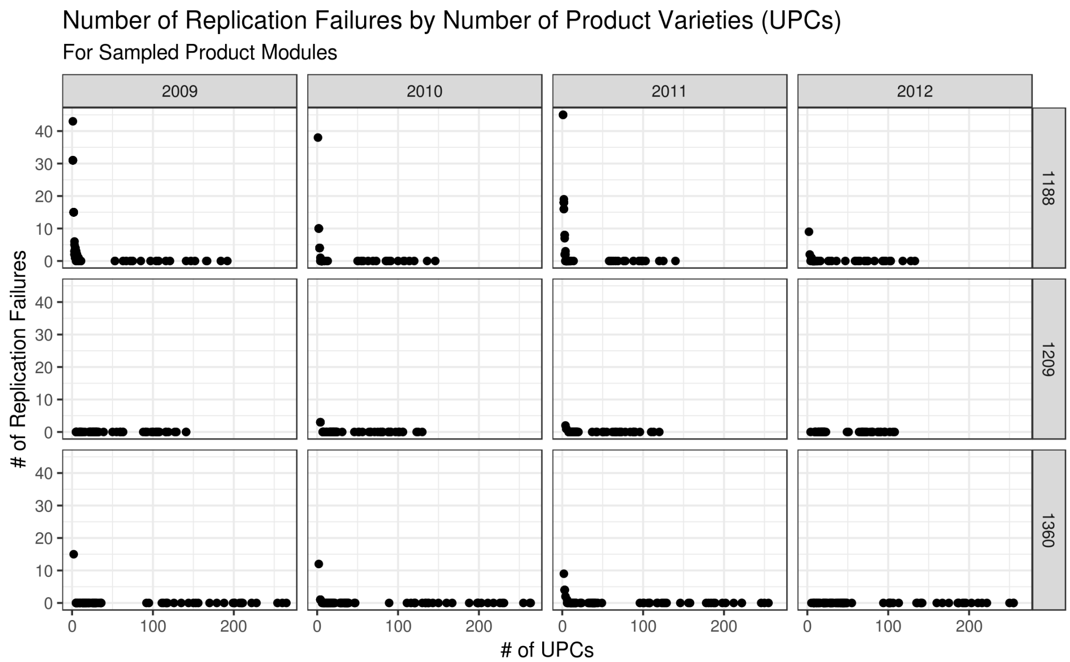

Figure 5.31: SUPI Replication Failures by UPCs and Product Module

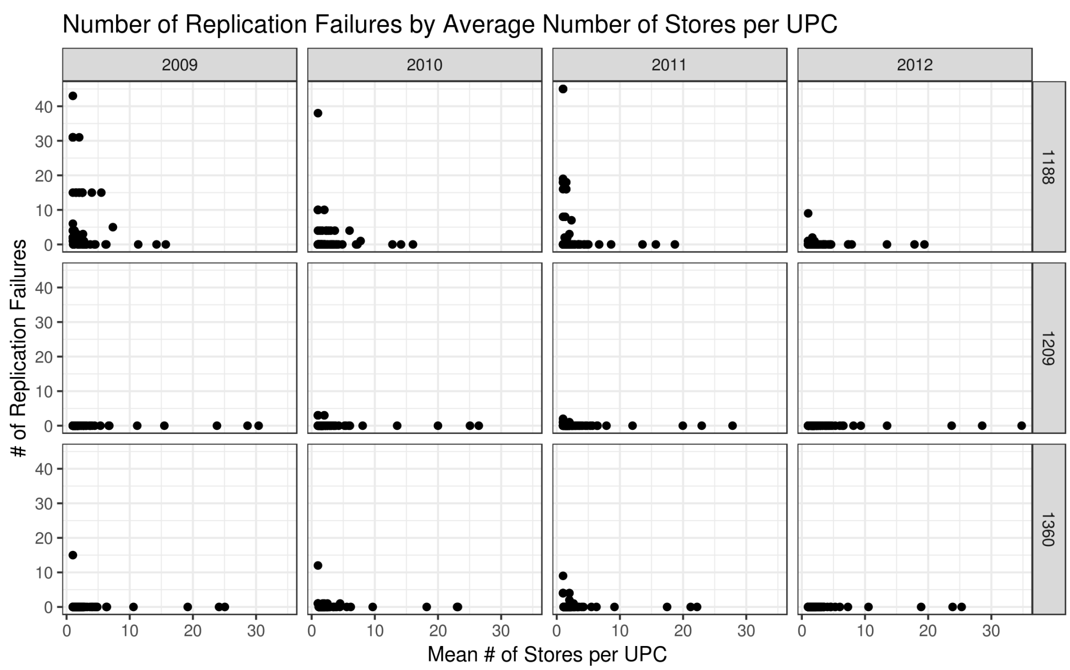

Figure 5.32: SUPI Replication Failures by Number of Stores and Product Module

ity of this index is potentially greater than these estimates would suggest. In such a circumstance, there might be a case for combining some of the smaller counties

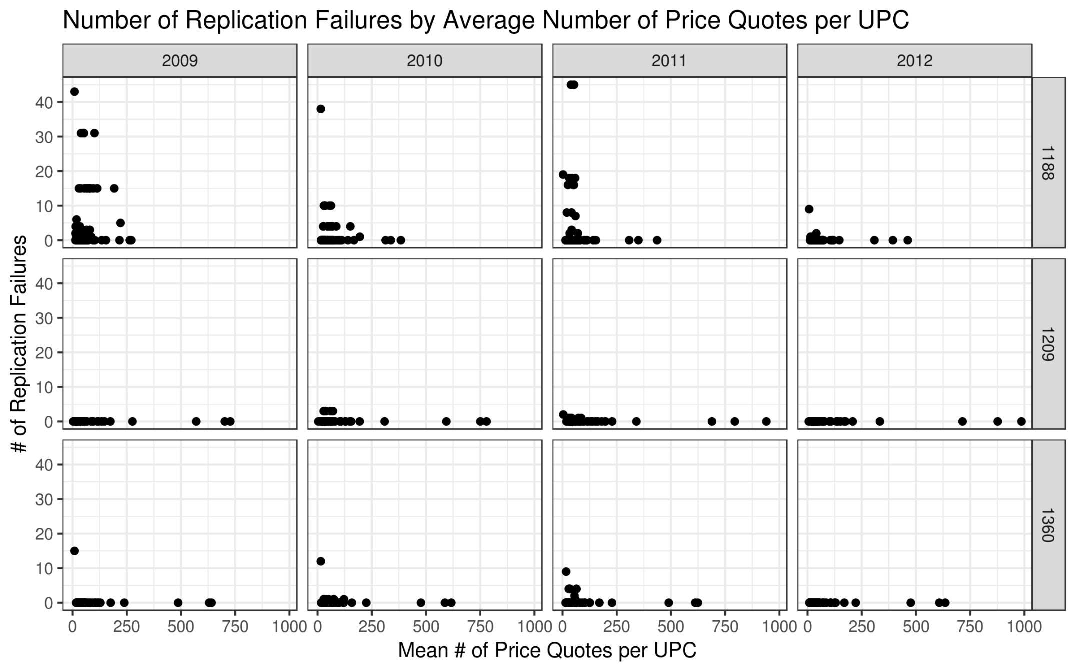

Figure 5.33: SUPI Replication Failures by Number of Price Quotes and Product Module

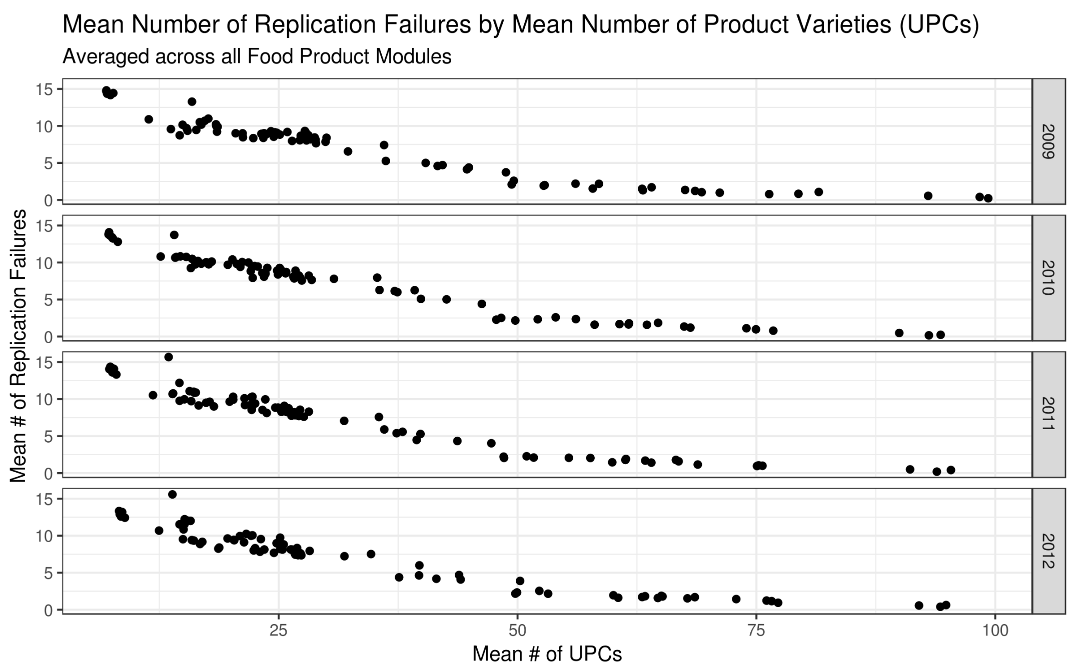

Figure 5.34: SUPI Replication Failures by UPCs per Product Module

together with adjacent areas into larger composite areas, in order to compensate for the sparsity of coverage. The second interpretation is that there really are only a

Figure 5.35: SUPI Replication Failures by Number of Stores and Product Module

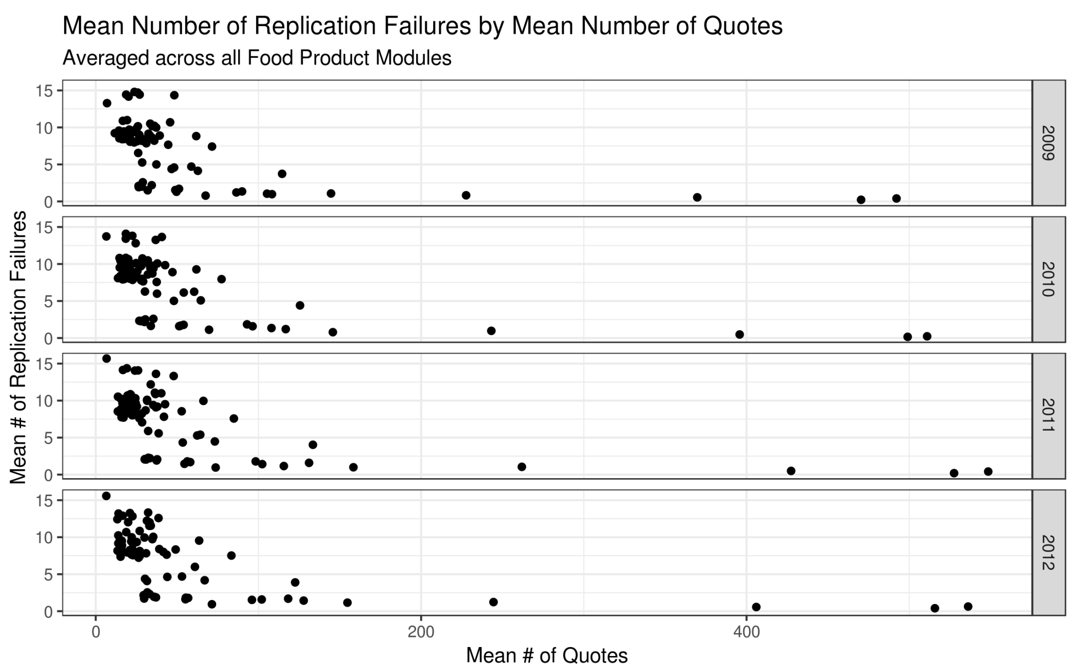

Figure 5.36: SUPI Replication Failures by Number of Quotes per Product Module

small number of product varieties available in these areas. Under this interpretation, a smaller variability estimate might be justified. Any attempt to collapse the areas would therefore cause us to lose potentially important information about the individual counties in question. Because we have no practical means to evaluate the degree of coverage in the Nielsen data within every product module, county and year it includes, we must remain agnostic about the causes of replication failures in each individual case. This complicates the interpretation of our variability estimates.

Though cases like these are worrisome, they are also relatively rare. Figure 5.37 shows a histogram of the number of times that any area's π value (defined as in Section 3.2) could not be replicated. We can see that for each of our sampled product modules, the bars in this histogram are generally clustered close to zero, with a long thin tail to the right. The frequency of failed replications also seems to drop off over time, possibly reflecting an increase in data availability over the period we study.

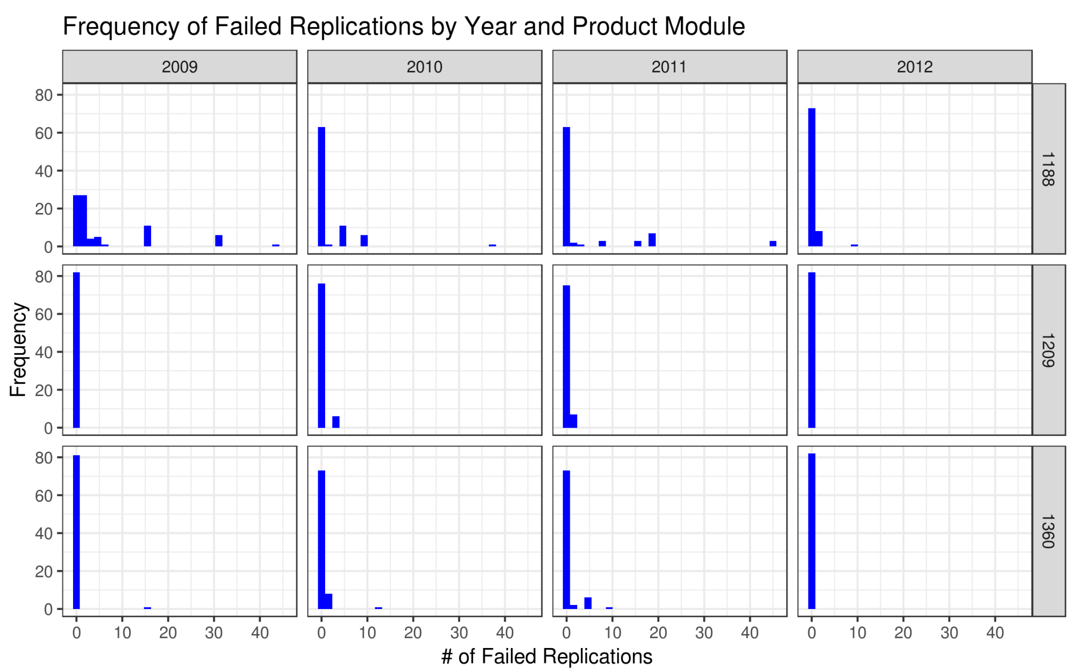

Figure 5.37: Histogram of Failed Replications by Year and Sampled Product Module

We can get a broader idea of how replication failures affect our bootstrap estimates by creating a similar histogram for each year pooling across all of the food product modules in our sample. This histogram, shown in Figure 5.38, relates the number of times that any area's π value could not be replicated across all 554 food product modules in which there are at least 3 total observed product varieties. Figure 5.39 is similar histogram in which the frequency of failed replications is pooled across only the "common product modules," which are observed in every county. These are the modules that are ultimately used to estimate our food CSUPIs. Although both of these histograms indicate that the majority of areas in the majority of product modules have no failed replications, there is a thicker tail and an observable multimodality in Figure 5.38 that does not exist to the same extent in Figure 5.39. This suggests that in the aggregate common products have better representation within each area, leading to smaller proportions of "brittle" SUPI estimates that depend heavily on only one or two product varieties.

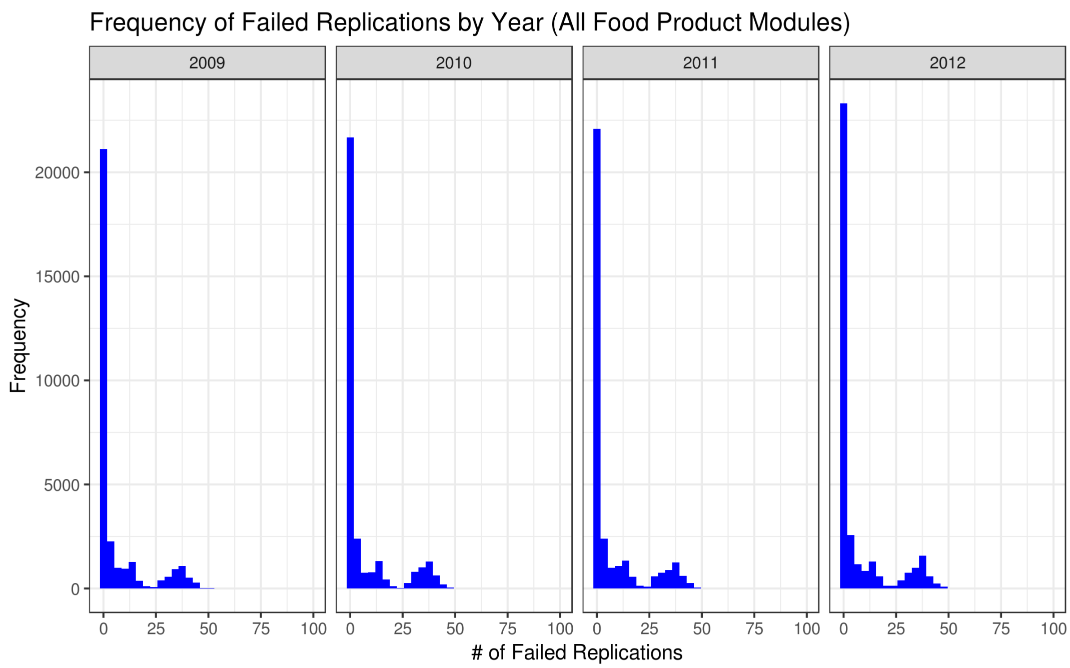

Figure 5.38: Histogram of Failed Replications by Year across All Product Modules

Within the common product modules, we can directly compare the proportion of replication failures within each area. Table B.3 contains the numerical table of these proportions across county and year, and Figure 5.40 is a barchart of the proportions in this table.

These results show that Leelanau county has the highest proportion of replication

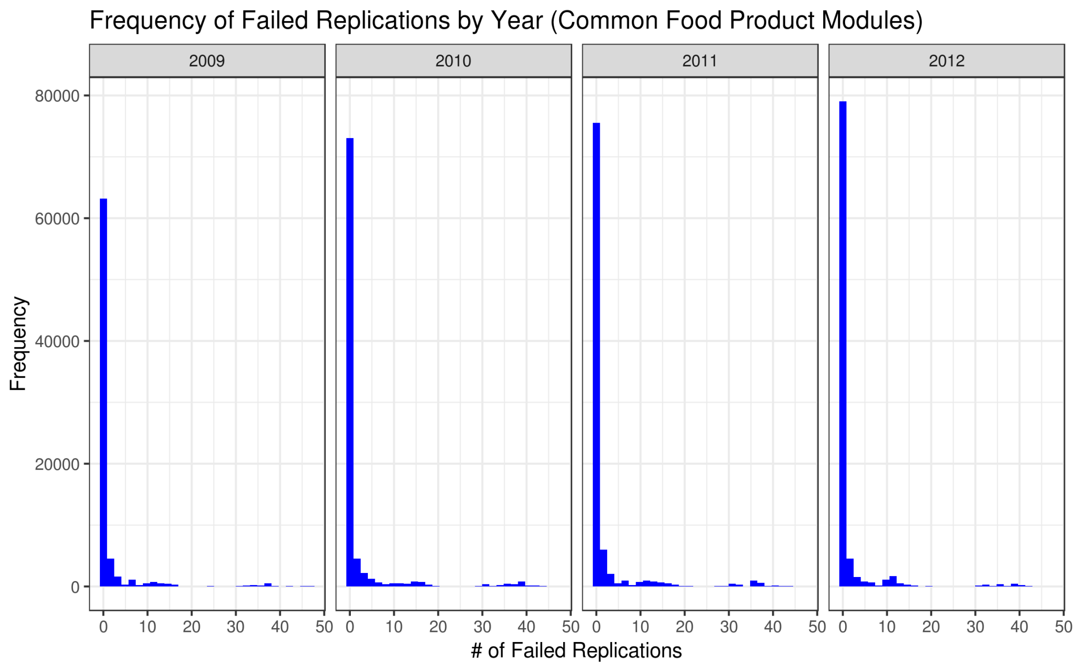

Figure 5.39: Histogram of Failed Replications by Year across Common Product Modules

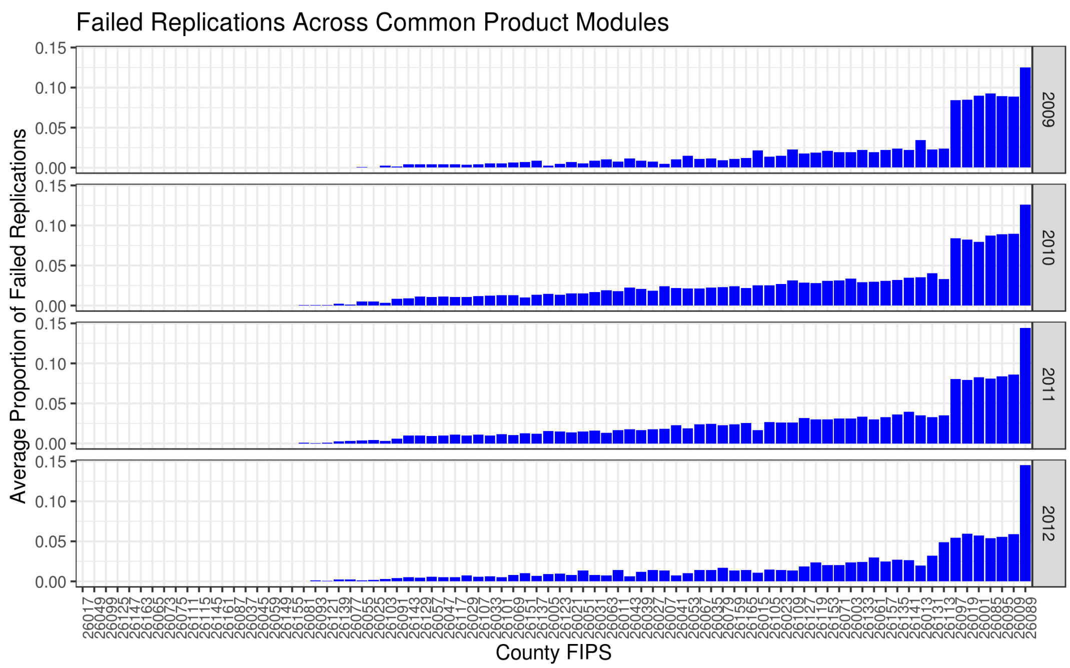

Figure 5.40: Proportion of Failed Replications by County across Common Product Modules

failures in our sample, with about 15% of bootstrap replications across all common product modules failing. The vast majority of the counties have average bootstrap failure rates of around 2-3%, with a small cluster of areas having rates closer to 10%. These aggregated failure rates are not as concerning as some of the bootstrap failure rates we see at the level of individual products, such as the 43% bootstrap failure rate for vinegar SUPIs in Leelanau county.

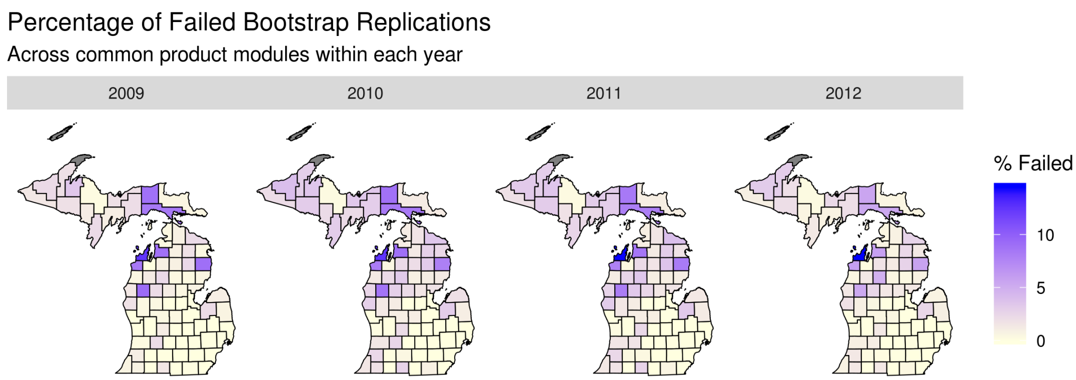

Figure 5.41: Choropleth of Failed Replications Percentages

As we can see from Figure 5.41, there is no obvious spatial pattern to the counties that experience higher rates of bootstrap failures, though the counties in which they do occur are consistent across time. Bootstrap variance estimates from these counties for the food CSUPIs we estimate in Section 5.2 should be viewed with some additional caution.

With that said, we expect the impact of product modules with high numbers of replication failures on the overall food CSUPIs to be smaller than the impact of other product modules. The reason for this is that product modules that experience high rates of replication failures tend to account for a smaller percentage of total food expenditure than other product modules, as we can see in Figure 5.42. This relationship is even stronger when we restrict consideration to the product modules common to all Michigan counties, as we can see in Figure 5.43.

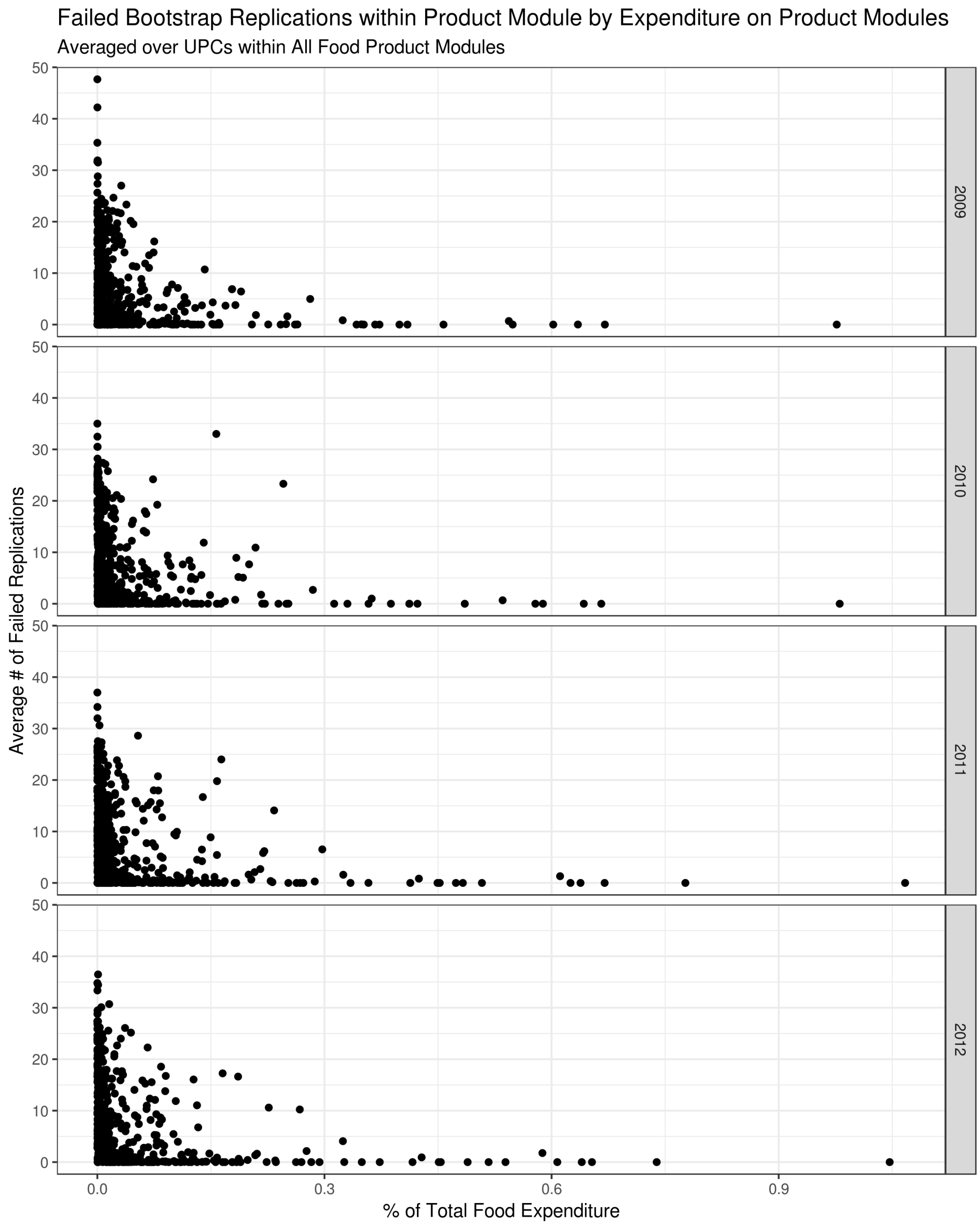

Figure 5.42: Average Failed Replications per Product Module by Percentage of Total Food Expenditure

Because product modules that have high numbers of replication failures also tend to account for low proportions of total food expenditure, they receive low expenditure

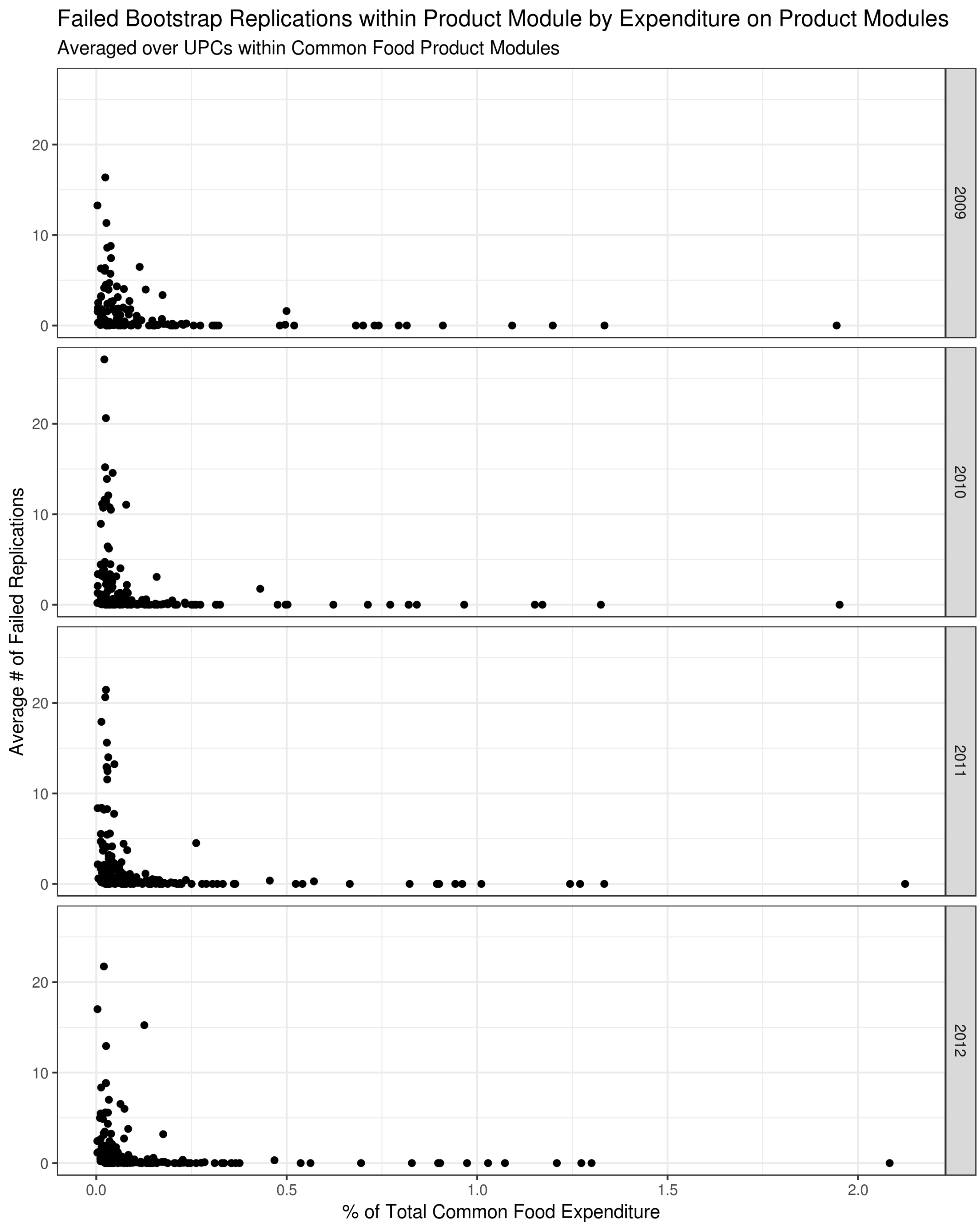

Figure 5.43: Average Failed Replications per Product Module by Percentage of Common Food Expenditure

weights ($\hat{\omega}_g$ as defined in Section 4.2.2), and thus make smaller contributions to the overall index values.

5.4 Face Validity

In this section, we briefly discuss the validity of our indices. As discussed in Chapter III, the SUPI and CSUPI estimates above are based on an economic model that attempts to describe how consumers value the consumption options available in each county. For our measurements to be valid, this model should ideally approximate the concerns of consumers in each county as closely as possible. Unfortunately, it is difficult to assess the extent to which it does. We are not aware of any available price indices that are published for low levels of spatial aggregation such as counties and include adjustments for biases related to consumer valuation and product turnover. Additionally, because we do not estimate the elasticity of substitution parameter σ, we do not know the correct SUPI values to use for such a comparison.

For these reasons, we address this question by informally evaluating the "face validity" of our results. In particular, three patterns from our results seem both intuitively plausible and largely consistent with our expectations. The first such pattern is that the cross-county variation of the SUPI is relatively low when $\sigma = \infty$. From Tables 5.5, 5.6 and 5.7, we can see that the between-county standard deviation for individual products is generally on the order of about 4-5%. Table 5.9 shows that the between-county standard deviation for food CSUPIs with $\sigma = \infty$ is around 2%. These results are in line with our expectation that average grocery prices will exhibit modest, but not extreme, variation between counties.

The second pattern is that the cost of living in rural areas increases as products are assumed to be less substitutable (i.e. as $\sigma \to 1$). This makes sense because less-populated areas have fewer potential customers, and as a result they can support fewer stores than more populated areas. If different stores carry different product varieties, then we might expect less populated areas to have less extensive selections of products on average. If consumers in less-populated counties would strongly prefer to consume particular product varieties that are unavailable where they live, then it would make sense that our CSUPIs show those places as having a higher cost of living.

Finally, Figure 5.24 shows that our food CSUPI estimates for each county are remarkably stable across time. If differences in the cost of living between counties are related to factors such as population and physical capital investment, we wouldn't expect these relationships to change very much on a year-to-year basis. Thus the fact that our CSUPIs do not change dramatically from year-to-year is an encouraging sign.

CHAPTER VI

Conclusions

6.1 Conclusions

Based on the results in Chapter V, we draw several broad conclusions. First, the potential impact of scanner data on estimating spatial price differences is as substantial as the impact of scanner data on estimating inflation indices, and for many of the same reasons. The increase in the quantity of available information made possible by scanner data makes comparisons of the cost of living possible at higher frequencies and finer levels of spatial resolution than were previously feasible. We demonstrated this by estimating spatial price indices comparing the cost of living between individual counties within the state of Michigan, without the use of imputation or regression estimates based on census data, or data from higher levels of aggregation. Similar procedures could be used to estimate price indices comparing any set of spatial units, given sufficient data.

Second, using the Nielsen retail scanner data, we are able to produce indices that account for aspects of consumer welfare that are difficult or impossible with more traditional methods. We identify spatial analogues to the biases due to differing product availability that Feenstra (1994)[17] identifies, and the consumer valuation bias identified by Redding and Weinstein (2016)[34][35]. Assuming that a CES utility function of the form assumed by Feenstra (1994)[17] and Redding and Weinstein (2016)[34][35] adequately describes the preferences of consumers, we are able to estimate the effects of these biases on the cost of living.

Third, we can use the above framework to draw some substantive conclusions about differences in the cost of living between Michigan counties. When variety and consumer valuation effects are taken into account, for most reasonable values of the elasticity of substitution parameter σ, a pattern emerges in which the cost of living associated with food is highest outside of a cluster centered around the Detroit metropolitan area. While consumer valuation bias contributes to the observed outcome, this pattern is driven primarily by differences in which product varieties are available in each county. In plain terms, we can interpret this result to mean that even in cases where the raw price of food is similar, consumers in less urbanized areas can be worse off when there is a smaller selection of product varieties available for them to choose between. How much worse off they are depends on how much they value consuming the particular varieties of food that are unavailable to them as compared to locally available substitutes.

We can see from Table and 5.9 that the cross-county variability in the cost of living reacts strongly to changes in the elasticity of substitution. In particular, increasing σ from 2 to 4 results in a factor of 8 reduction in the between-county food CSUPI standard deviation. The reduction can be even more dramatic in individual proudct modules, as we can see in Tables 5.5, 5.6, and 5.7. However, the magnitude of this decline decreases quickly as $\sigma \to \infty$. The cross-county food CSUPI standard deviation halves as σ increases from 4 to 6, but falls by only 0.1 or 0.2 as σ increases from 10 to 12. At the limit where consumers find all product varieties infinitely substitutable, the cross-county food CSUPI standard deviation is small, suggesting that differences in the prices of goods across counties are generally modest.

This implies that most of the food cost of living differences between urban and rural counties are due to differences in product availability. One might liken counties with few available food varieties to urban "food deserts," impoverished areas of cities that suffer due to the absence of nearby grocery stores. Residents of these food deserts may consume different diets than they would have if a broader selection of food items were available inside of their communities. While neighborhood-level food deserts are too granular to be visible in a county-level index, the gap between measured prices and cost of living differences in our results point to the possibility of a regional analogue to this phenomenon.

Because the SUPI includes the impact of product selection on consumer well-being, it is not appropriate for applications in which only differences in price matter. For example, a study whose goal is to compare food prices between Detroit and Michigan's upper peninsula would be better served by alternative indices that focus exclusively on price differences, such as the Jevons. Similarly, researchers seeking to establish the minimum wage required to purchase a set number of calories may not wish to use the SUPI, since it includes information that is irrelevant to this task. The SUPI is a better option than indices such as the Jevons for evaluating spatial differences in consumer welfare, or for studies that seek to compare quality of life more broadly. For instance, a homebuyer comparing the cost maintaining a set standard of living in different counties might prefer the SUPI. For the same reason, a researcher concerned with regional inequality might prefer to use the SUPI to adjust nominal wages for cost of living differences.

The SUPI is not the only approach to making spatial comparisons based on scanner data, nor is it the definitive solution to this problem. The work discussed above is highly exploratory, and these conclusions depend on a number of modeling decisions and assumptions. We discuss some of the challenges that our approach poses for statistical agencies in Section 6.2, and several ways that our indices could be improved upon in Section 6.3. Despite the contingent nature of these results, we hope that the approach we have advanced can become a serious option for statistical agencies or other interested parties to improve economic measurement.

6.2 Challenges to Implementation

Several practical challenges must be considered before indices like the SUPIs can be published by government statistical agencies. The most fundamental of these have to do with the raw data such indices require. For several reasons, it is unlikely that data gathered by organizations such as Nielsen would be sufficient for such a task. Most glaringly, the Nielsen retail scanner data do not contain information about regions outside the continental United States that statistical agencies would want to include. This would mean that SUPIs could not be produced for states like Alaska and Hawaii, or US territories such as Puerto Rico. Further, the retail scanner data are gathered from an opt-in, non-random sample of stores. This means that government statistical agencies relying on Nielsen would not be able to guarantee that indices are based on a representative sample of the population. Coverage rates could vary significantly across space and time depending on which stores choose to participate in each area. For example, if a large retail chain were excluded from the sample, this exclusion could have a disproportionate impact on coverage rates in less populated areas with fewer competing outlets. Additionally, Nielsen prohibits government organizations from using the retail scanner data without their explicit permission, and it is unclear what terms Nielsen would require for statistical agencies to obtain this permission.

For these reasons, statistical agencies attempting to incorporate indices based on scanner data into the national accounts will require additional sources of data. Implementation would involve combining information from multiple data sources, potentially including measurement firms such as Nielsen, retail chains, and/or individual stores at the point of purchase. Perhaps the least disruptive way to implement this would be to request scanner data from outlets that are selected by the Bureau of

Labor Statistics (BLS) using sampling frames derived from point of purchase survey (POPS). Unless the CPI index areas were also changed, this would not provide the necessary spatial resolution to estimate county level indices. An alternative approach might be to solicit information from the leadership of chain retailers directly, and thereby obtain scanner data from associated stores or franchises in the United States wherever they are located. In the long term, both of these approaches have the potential to reduce costs for participating businesses and statistical agencies by reducing the need for manual data collection. However, these data are produced by private organizations with different motivations and interests to protect. As a result, a major challenge for statistical agencies employing either approach will be incentivizing participation. If data are obtained from aggregators such as Nielsen, agencies might have to pay substantial sums for access. Another challenge will be combining data gathered from such heterogeneous sources into a standardized format from which indices can be estimated, despite differences in which information is recorded and how the data are organized.

Part of this process will include deciding upon appropriate rules to identify and remove outliers and mistaken observations. Though we trim outliers based on their distance in IQRs from the 25th and 75th percentiles, statistical agencies might also wish to consider different methods from the literature. For instance, the Bureau of Economic Analysis (BEA) screens for outliers using a procedure called Quaranta analysis, which they adapt from its initial use in the International Comparisons Program (ICP) organized by the United Nations and the World Bank.[4] Redding and Weinstein (2019) rely on a combination of trimming "purchases by households that reported paying more than three times or less than one third the median price for a good in a quarter or who reported buying twenty-five or more times the median quantity purchased by households buying at least one unit of the good," and winsorizing their data by "dropping observations whose percentage change in price or value were in the top or bottom one percent."[35] Statistical agencies will need to consider the impact of different methods of outlier screening on their estimates when choosing their approach to this problem.

When organizing these data, statistical agencies will also need to make choices about which product varieties should be modeled as substitutes. For convenience, this thesis uses the product module classifications chosen by Nielsen to identify categories of substitutable goods. Depending on how their data are collected, statistical agencies will have more options in this regard. They might choose to employ the elementary level item (ELI) classifications currently used by the Bureau of Labor Statistics, item classifications used by particluar businesses or groups of businesses, or something entirely new.

This choice will have a significant impact on the resulting indices, because it defines the goods for which SUPIs will be estimated. This definition can also impact the validity of other choices we make in this thesis. For example, if products are defined narrowly, dropping all products that are missing observations in even one of the 3,142 counties within the United States is likely to result in dropping every product. In such a circumstance, aggregating product-level SUPIs to CSUPIs using a weighted geometric mean as we have done may be infeasible. To address this, statistical agencies will need to be able to distinguish between product categories that are structurally missing, and those that are missing because the stores carrying these product categories are not included in the sample. Agencies might choose to impute index values that are missing due to sampling error, while approaching structural missingness differently depending on whether products are thought to be substitutable. Substitutable products with structural missingness can be aggregated by applying the UPI and the GEKS to the product-level SUPIs, rather than using a weighted geometric mean. Products with structural missingness that are not substitutable will be more complicated to aggregate. To cope with this issue, statistical agencies might need to combine several of the smaller counties into broader areas, in order to limit the number of missing products.

6.3 Suggestions for Future Research

There are several directions that future researchers might pursue to improve upon this work. Most conspicuously, we posit, rather than estimate, values for the crucial elasticity of substitution parameter σ. There are several competing ideas in the literature about how this parameter should be estimated in the context of inflation indices. Much of this discussion focuses on various ways to disentangle changes in demand from changes in price, so that σ can be inferred by observing changes in consumer expenditure across time.[34][35] The extent to which these ideas can or should be generalized to a spatial context remains unclear, and thus a promising area for future work. In the absence of a consensus method for identifying σ from observational data, future research might approach this problem from a Bayesian perspective. This could mean marginalizing SUPI estimates over an appropriate prior distribution for σ, which may be elicited from the literature or other relevant sources, rather than positing arbitrary values as we have done here.

We have also made fairly restrictive assumptions about the constancy of σ that future researchers might attempt to relax. For example, we have assumed that σ is constant within each individual product module. This implies that, for example, all flavored snack crackers are assumed to be equally substitutable for other flavored snack crackers. As Ehrlich et. al (2019)[14] point out, this assumption is questionable, and there is significant room to relax it by experimenting with different product groupings than the ones proposed by Nielsen. We have also assumed that the elasticity of substitution is constant *between* all food product modules, so that consumers are exactly as willing to substitute different kinds of snack crackers for each other as they would be to substitute different kinds of tuna for each other. This is also questionable, and could be improved upon with a proper method for estimating the values of σ in each product grouping.

Once the elasticity of substitution can be reliably estimated, the SUPI would benefit from future research assessing its validity. The most straightforward approach would be to compare SUPI and/or CSUPI estimates for some class of products to indices that use a different approach to account for the welfare impact of differences in consumer preferences and product varieties. This would enable researchers to assess the concurrent validity of both measures. At present, we are not aware of any available county-level price indices that would be appropriate for this comparison. One possible approach for producing such indices would be to estimate large-scale hedonic regression models to account for differences in the available product varieties and consumer tastes. This would require more detailed item characteristics than are available from the Nielsen retail scanner data, but may be feasible using other data sources. Alternatively, researchers could assess validity by comparing SUPI values to relevant measures of consumer satisfaction or behavior. In this approach, researchers would operationalize the concept of consumer welfare used by the CES utility function, and assess whether this construct is correlated with SUPI estimates.

Before researchers can rely on indices estimated from the Nielsen retail scanner data, it is necessary to assess the degree of population coverage these data have. There are several important questions that our research cannot address because we don't know whether the Nielsen data contains a representative sample of stores from the counties we estimate indices for. For example, we are uncertain about whether the product modules that are not observed in some counties are missing because they are unavailable in those places, or because the retailers that sell them have chosen not to participate. Because our food CSUPIs are based exclusively on product modules that are observed in all of the areas we wish to compare, we omit data on many goods that would otherwise be available. If we could distinguish between products that are structurally missing and those that are merely excluded from the sample, it would help us decide whether imputation is an appropriate response to this limitation. If index values for missing products could be validly and reliably imputed, then our food CSUPIs could include the impact of hundreds of products that are currently ignored. For the same reasons, it is difficult to know whether areas with small numbers of in-sample product varieties also have small numbers of population product varieties. This is relevant because we cannot say for certain whether an area has a large variety adjustment term due to its economic characteristics, or merely because our sample has sparse coverage there.

This problem also complicates the interpretation of our bootstrapped variance estimates. As we discuss in Chapter V, replication failures due to structurally missing product varieties should be interpreted differently than replication failures due to data quality issues. More generally, additional information about the methods by which the data were sampled, or how the sample in each area relates to the population, could make it possible to improve the quality of our variance estimates. Our cluster bootstrap estimates depend on the implausible assumption that observations are missing from our sample at random, and thus could benefit from further insight into sampling methodology or coverage rates.

Future work might also seek to assess the impact of other assumptions we have made on our uncertainty measures. In particular, our cluster bootstrap method assumes away any dependence between the prices of different product varieties in the same store. Researchers might quantify the impact of this simplification on the resulting variance estimates by measuring the dependence between product varieties. Alternatively, they could attempt to relax these problematic assumptions, and thereby avoid some of the compromises that our approach requires.

APPENDICES

APPENDIX A

Figures

This appendix contains versions of the choropleth maps in Chapter V that are normalized relative to the geometric mean price level of all counties, rather than Washtenaw couny. Formally, let the set of all counties be $\mathbb{A}$, and Washtenaw county be area w. Within each year, we denote the renormalized SUPIs and CSUPIs for a given county $c \in \mathbb{A}$ as $SUPI_{g\mu c}$ and $CSUPI_{\mu c}$ respectively.

We calculate the renormalized SUPIs for each product module $g \in \mathbb{G}$ as

$$SUPI_{g\mu c} = \frac{SUPI_{gwc}}{\left(\prod_{a \in \mathbb{A}} SUPI_{gwa}\right)^{\frac{1}{|\mathbb{A}|}}} \tag{A.1}$$

Similarly, the renormalized food CSUPIs are calculated as

$$CSUPI_{\mu c} = \frac{CSUPI_{wc}}{\left(\prod_{a \in \mathbb{A}} CSUPI_{wa}\right)^{\frac{1}{|\mathbb{A}|}}} \tag{A.2}$$

Because each of these indices are normalized within year, the product module level maps can have appreciably different color scales in different years.

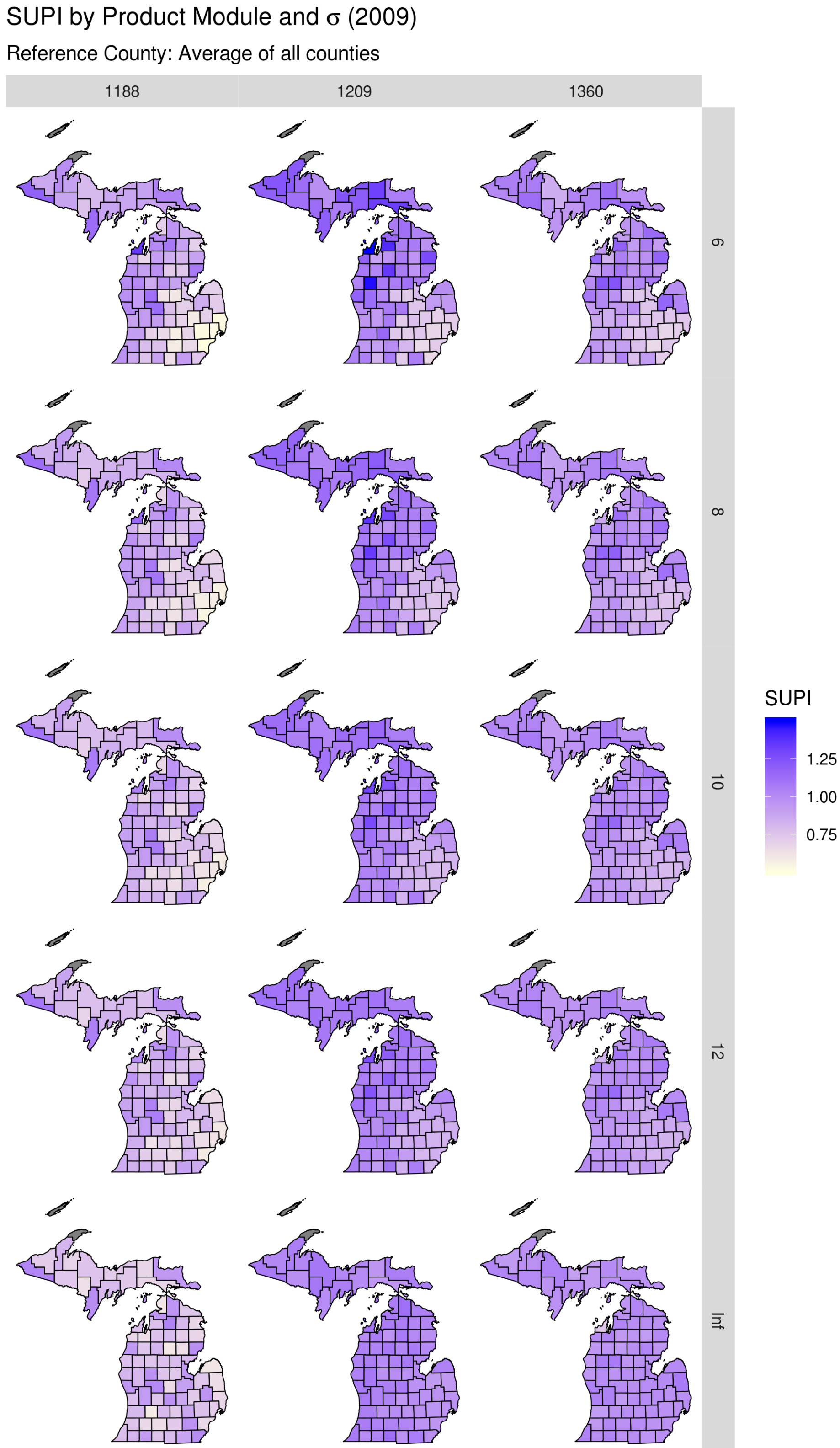

Figure A.1: Product Module SUPI Choropleth (2009)

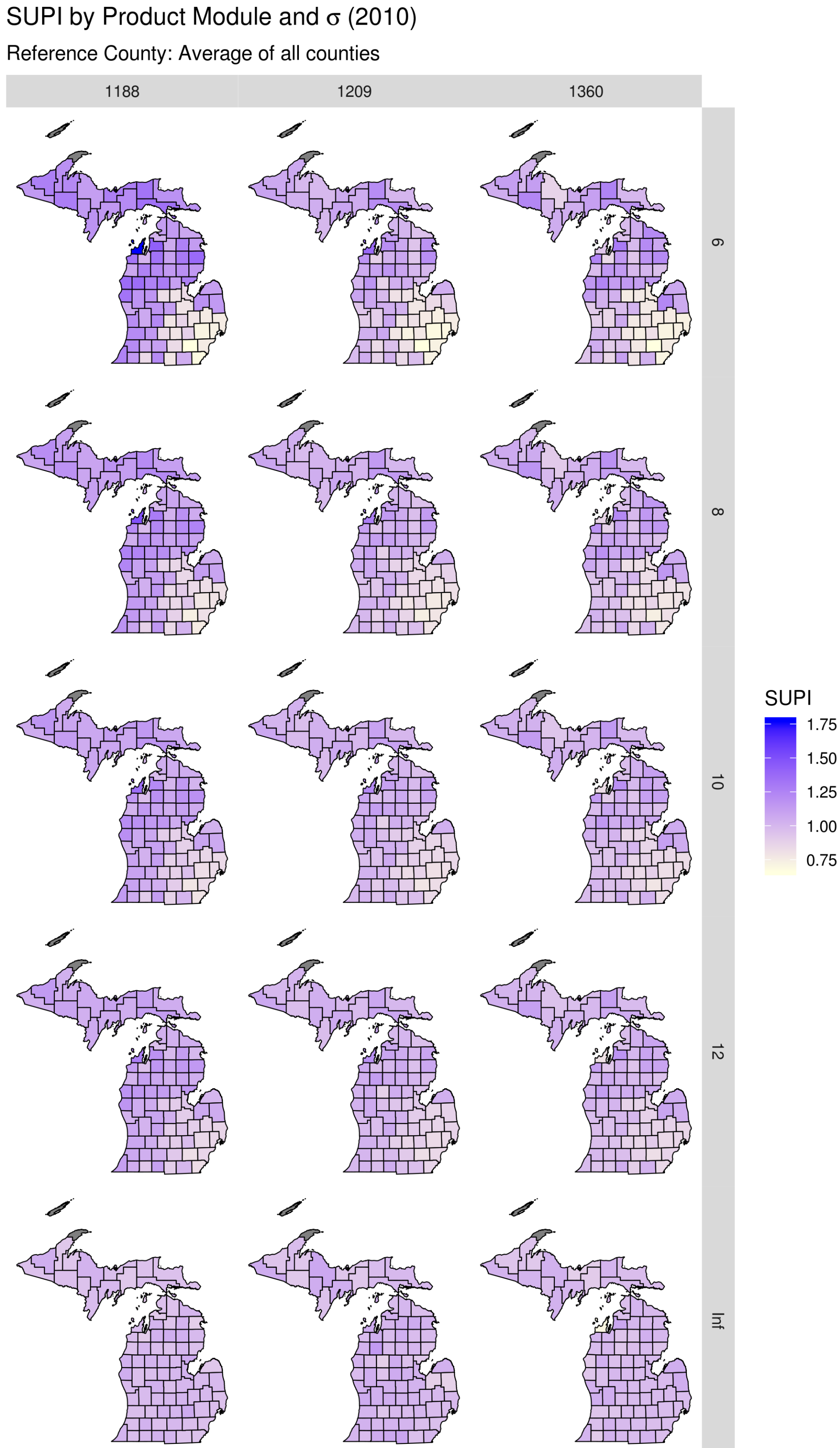

Figure A.2: Product Module SUPI Choropleth (2010)

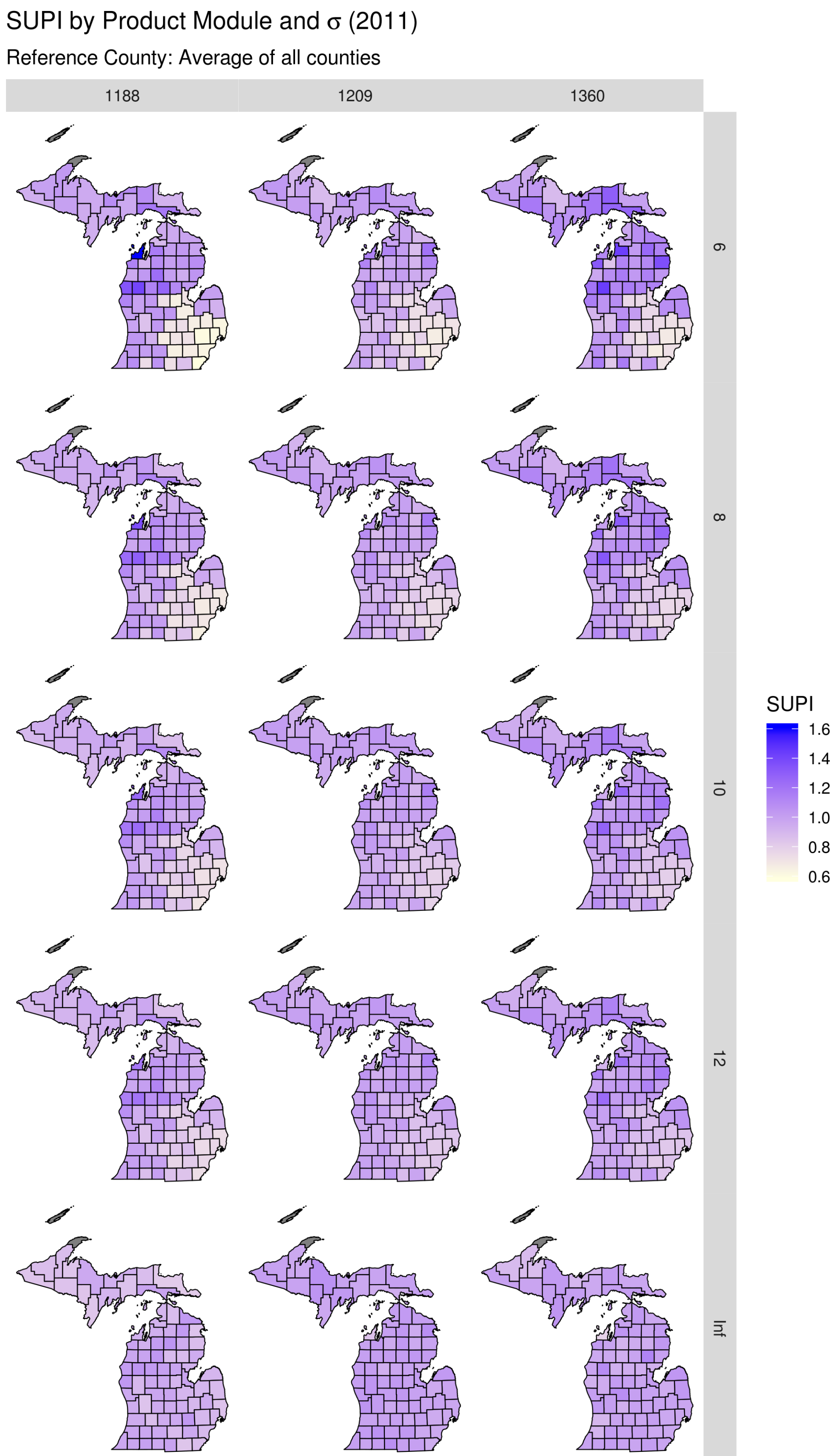

Figure A.3: Product Module SUPI Choropleth (2011)

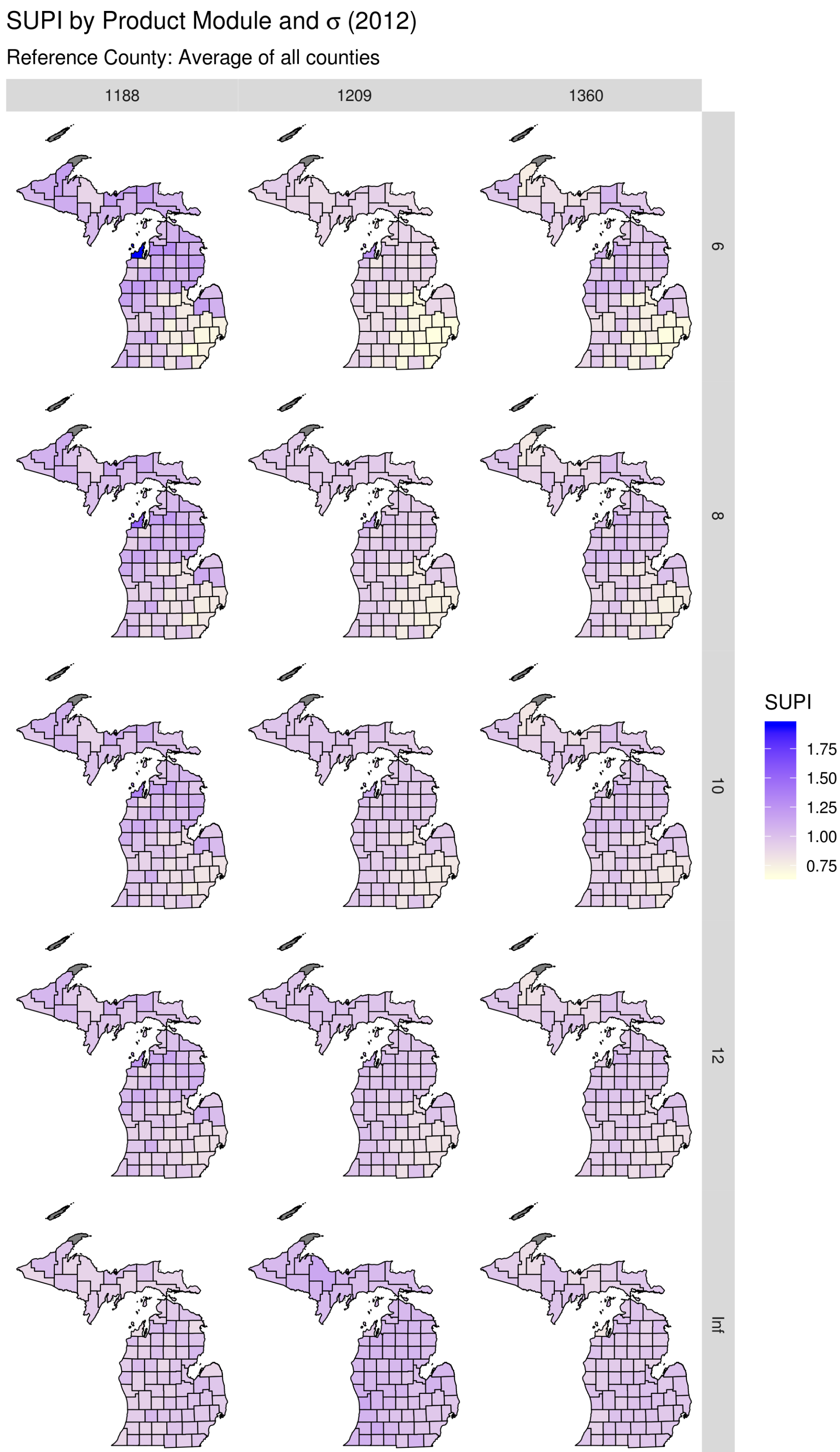

Figure A.4: Product Module SUPI Choropleth (2012)

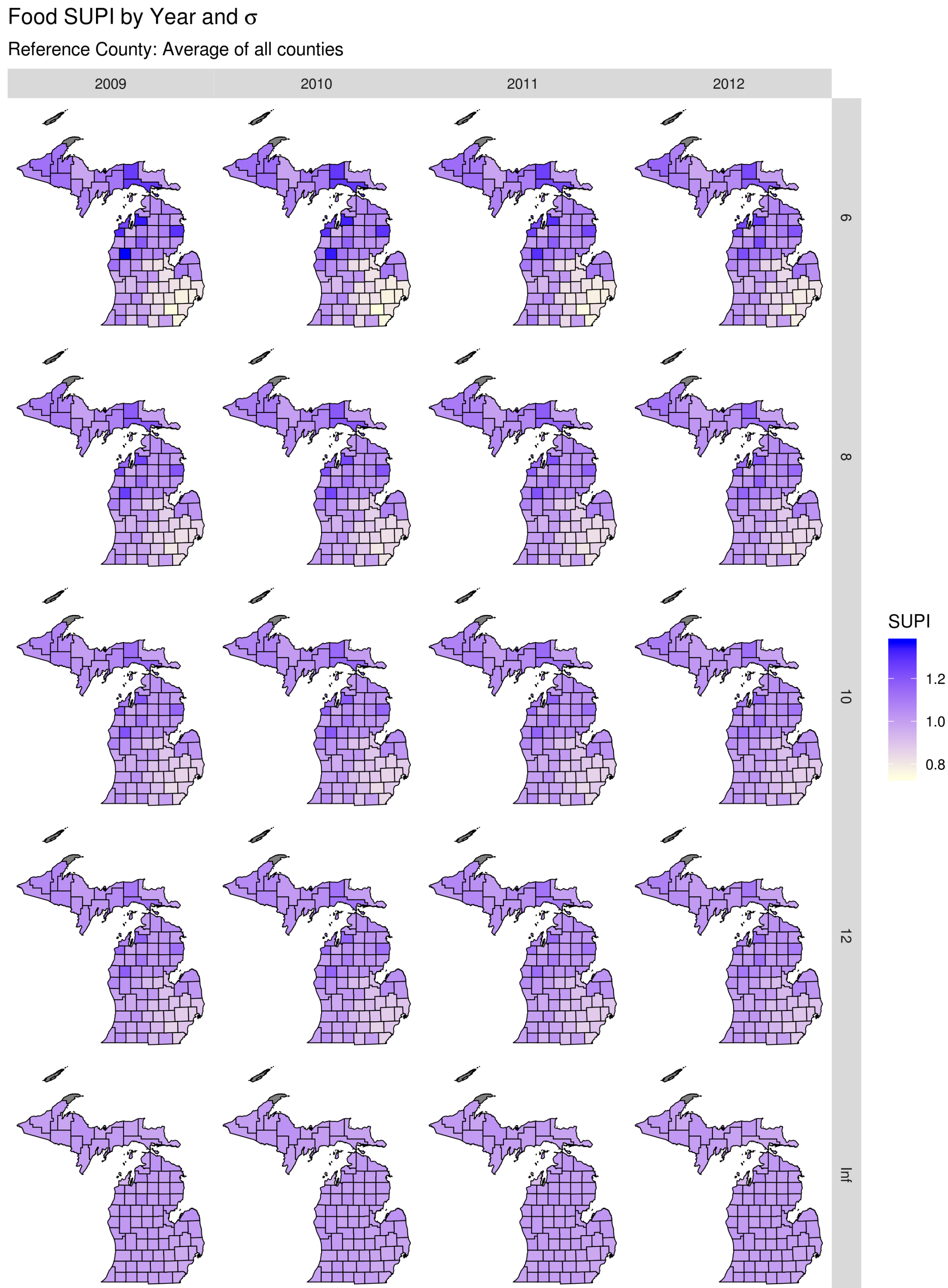

Figure A.5: Food CSUPI Choropleth

APPENDIX B

Tables

Table B.1: Food Product Groups

Product Group Code	Product Group Description
3001	DRESSINGS/SALADS/PREP FOODS-DELI
0503	CANDY
1007	CONDIMENTS, GRAVIES, AND SAUCES
0505	GUM
0501	BABY FOOD
4001	FRESH PRODUCE
1017	SPICES, SEASONING, EXTRACTS
2008	PREPARED FOODS-FROZEN
2010	VEGETABLES-FROZEN
1505	COOKIES
2503	COT CHEESE, SOUR CREAM, TOPPINGS
2506	MILK
1501	BREAD AND BAKED GOODS
0513	SOUP
2510	YOGURT
0514	VEGETABLES - CANNED
2501	BUTTER AND MARGARINE
2502	CHEESE
2007	PIZZA/SNACKS/HORS DOEURVES-FRZN
1020	TEA
2508	SNACKS, SPREADS, DIPS-DAIRY
1014	PICKLES, OLIVES, AND RELISH
1016	SHORTENING, OIL
1506	CRACKERS
1002	BAKING SUPPLIES
0510	PREPARED FOOD-READY-TO-SERVE

1019	TABLE SYRUPS, MOLASSES
0507	JUICE, DRINKS - CANNED, BOTTLED
2505	EGGS
1011	NUTS
1503	CARBONATED BEVERAGES
0511	PREPARED FOOD-DRY MIXES
1001	BAKING MIXES
0512	SEAFOOD - CANNED
1021	VEGETABLES AND GRAINS - DRIED
1507	SNACKS
1508	SOFT DRINKS-NON-CARBONATED
2002	BREAKFAST FOODS-FROZEN
2003	DESSERTS/FRUITS/TOPPINGS-FROZEN
1012	PACKAGED MILK AND MODIFIERS
1006	COFFEE
0506	JAMS, JELLIES, SPREADS
2001	BAKED GOODS-FROZEN
3002	PACKAGED MEATS-DELI
2005	ICE CREAM, NOVELTIES
1013	PASTA
1015	SALAD DRESSINGS, MAYO, TOPPINGS
1005	CEREAL
2009	UNPREP MEAT/POULTRY/SEAFOOD-FRZN
1010	FRUIT - DRIED
0504	FRUIT - CANNED
1008	DESSERTS, GELATINS, SYRUP
1018	SUGAR, SWEETENERS
2006	JUICES, DRINKS-FROZEN
1004	BREAKFAST FOOD
2504	DOUGH PRODUCTS
3501	FRESH MEAT
1009	FLOUR
2507	PUDDING, DESSERTS-DAIRY
2509	YEAST

Table B.2: In-Sample Michigan Counties (2009 - 2012)

FIPS	County Name	FIPS	County Name
26001	ALCONA	26085	LAKE
26003	ALGER	26087	LAPEER
26005	ALLEGAN	26089	LEELANAU
26007	ALPENA	26091	LENAWEE
26009	ANTRIM	26093	LIVINGSTON

26011	ARENAC	26095	LUCE
26013	BARAGA	26097	MACKINAC
26015	BARRY	26099	MACOMB
26017	BAY	26101	MANISTEE
26019	BENZIE	26103	MARQUETTE
26021	BERRIEN	26105	MASON
26023	BRANCH	26107	MECOSTA
26025	CALHOUN	26109	MENOMINEE
26027	CASS	26111	MIDLAND
26029	CHARLEVOIX	26113	MISSAUKEE
26031	CHEBOYGAN	26115	MONROE
26033	CHIPPEWA	26117	MONTCALM
26035	CLARE	26119	MONTMORENCY
26037	CLINTON	26121	MUSKEGON
26039	CRAWFORD	26123	NEWAYGO
26041	DELTA	26125	OAKLAND
26043	DICKINSON	26127	OCEANA
26045	EATON	26129	OGEMAW
26047	EMMET	26131	ONTONAGON
26049	GENESEE	26133	OSCEOLA
26051	GLADWIN	26135	OSCODA
26053	GOGEBIC	26137	OTSEGO
26057	GRATIOT	26139	OTTAWA
26055	GRD TRAVERSE	26141	PRESQUE ISLE
26059	HILLSDALE	26143	ROSCOMMON
26061	HOUGHTON	26145	SAGINAW
26063	HURON	26147	SAINT CLAIR
26065	INGHAM	26151	SANILAC
26067	IONIA	26153	SCHOOLCRAFT
26069	IOSCO	26155	SHIAWASSEE
26071	IRON	26149	ST JOSEPH
26073	ISABELLA	26157	TUSCOLA
26075	JACKSON	26159	VAN BUREN
26077	KALAMAZOO	26161	WASHTENAW
26079	KALKASKA	26163	WAYNE
26081	KENT	26165	WEXFORD

Table B.3: Proportion of Failed Bootstrap Replications across Common Product Modules

Year	County FIPS	Proportion of Failed Replications
2009	26001	0.09
2010	26001	0.08
2011	26001	0.08

2012	26001	0.06
2009	26003	0.02
2010	26003	0.03
2011	26003	0.03
2012	26003	0.02
2009	26005	0.00
2010	26005	0.01
2011	26005	0.02
2012	26005	0.01
2009	26007	0.00
2010	26007	0.02
2011	26007	0.02
2012	26007	0.01
2009	26009	0.09
2010	26009	0.09
2011	26009	0.09
2012	26009	0.06
2009	26011	0.01
2010	26011	0.02
2011	26011	0.02
2012	26011	0.01
2009	26013	0.03
2010	26013	0.04
2011	26013	0.04
2012	26013	0.02
2009	26015	0.02
2010	26015	0.03
2011	26015	0.02
2012	26015	0.01
2009	26017	0.00
2010	26017	0.00
2011	26017	0.00
2012	26017	0.00
2009	26019	0.08
2010	26019	0.08
2011	26019	0.08
2012	26019	0.06
2009	26021	0.01
2010	26021	0.02
2011	26021	0.01
2012	26021	0.01
2009	26023	0.01
2010	26023	0.03
2011	26023	0.03
2012	26023	0.01

2009	26025	0.00
2010	26025	0.01
2011	26025	0.00
2012	26025	0.00
2009	26027	0.01
2010	26027	0.02
2011	26027	0.02
2012	26027	0.01
2009	26029	0.00
2010	26029	0.01
2011	26029	0.01
2012	26029	0.01
2009	26031	0.01
2010	26031	0.02
2011	26031	0.02
2012	26031	0.01
2009	26033	0.01
2010	26033	0.01
2011	26033	0.01
2012	26033	0.01
2009	26035	0.01
2010	26035	0.02
2011	26035	0.02
2012	26035	0.01
2009	26037	0.00
2010	26037	0.00
2011	26037	0.00
2012	26037	0.00
2009	26039	0.01
2010	26039	0.02
2011	26039	0.02
2012	26039	0.01
2009	26041	0.01
2010	26041	0.02
2011	26041	0.02
2012	26041	0.01
2009	26043	0.01
2010	26043	0.02
2011	26043	0.02
2012	26043	0.01
2009	26045	0.00
2010	26045	0.00
2011	26045	0.00
2012	26045	0.00
2009	26047	0.00

2010	26047	0.01
2011	26047	0.01
2012	26047	0.01
2009	26049	0.00
2010	26049	0.00
2011	26049	0.00
2012	26049	0.00
2009	26051	0.01
2010	26051	0.02
2011	26051	0.01
2012	26051	0.01
2009	26053	0.01
2010	26053	0.02
2011	26053	0.02
2012	26053	0.01
2009	26055	0.00
2010	26055	0.01
2011	26055	0.00
2012	26055	0.00
2009	26057	0.00
2010	26057	0.01
2011	26057	0.01
2012	26057	0.01
2009	26059	0.00
2010	26059	0.00
2011	26059	0.00
2012	26059	0.00
2009	26061	0.02
2010	26061	0.03
2011	26061	0.03
2012	26061	0.03
2009	26063	0.01
2010	26063	0.02
2011	26063	0.01
2012	26063	0.01
2009	26065	0.00
2010	26065	0.00
2011	26065	0.00
2012	26065	0.00
2009	26067	0.01
2010	26067	0.02
2011	26067	0.02
2012	26067	0.01
2009	26069	0.01
2010	26069	0.01

2011	26069	0.01
2012	26069	0.01
2009	26071	0.02
2010	26071	0.03
2011	26071	0.03
2012	26071	0.02
2009	26073	0.00
2010	26073	0.00
2011	26073	0.00
2012	26073	0.00
2009	26075	0.00
2010	26075	0.00
2011	26075	0.00
2012	26075	0.00
2009	26077	0.00
2010	26077	0.00
2011	26077	0.00
2012	26077	0.00
2009	26079	0.01
2010	26079	0.02
2011	26079	0.02
2012	26079	0.02
2009	26081	0.00
2010	26081	0.00
2011	26081	0.00
2012	26081	0.00
2009	26085	0.09
2010	26085	0.09
2011	26085	0.08
2012	26085	0.05
2009	26087	0.00
2010	26087	0.00
2011	26087	0.00
2012	26087	0.00
2009	26089	0.13
2010	26089	0.13
2011	26089	0.14
2012	26089	0.14
2009	26091	0.00
2010	26091	0.01
2011	26091	0.01
2012	26091	0.00
2009	26093	0.00
2010	26093	0.00
2011	26093	0.00

2012	26093	0.00
2009	26095	0.09
2010	26095	0.09
2011	26095	0.08
2012	26095	0.06
2009	26097	0.08
2010	26097	0.08
2011	26097	0.08
2012	26097	0.05
2009	26099	0.00
2010	26099	0.00
2011	26099	0.00
2012	26099	0.00
2009	26101	0.01
2010	26101	0.01
2011	26101	0.01
2012	26101	0.01
2009	26103	0.00
2010	26103	0.00
2011	26103	0.00
2012	26103	0.00
2009	26105	0.01
2010	26105	0.03
2011	26105	0.03
2012	26105	0.01
2009	26107	0.00
2010	26107	0.01
2011	26107	0.01
2012	26107	0.01
2009	26109	0.02
2010	26109	0.03
2011	26109	0.03
2012	26109	0.01
2009	26111	0.00
2010	26111	0.00
2011	26111	0.00
2012	26111	0.00
2009	26113	0.02
2010	26113	0.03
2011	26113	0.03
2012	26113	0.05
2009	26115	0.00
2010	26115	0.00
2011	26115	0.00
2012	26115	0.00

2009	26117	0.00
2010	26117	0.01
2011	26117	0.01
2012	26117	0.01
2009	26119	0.02
2010	26119	0.03
2011	26119	0.03
2012	26119	0.02
2009	26121	0.00
2010	26121	0.00
2011	26121	0.00
2012	26121	0.00
2009	26123	0.00
2010	26123	0.01
2011	26123	0.01
2012	26123	0.01
2009	26125	0.00
2010	26125	0.00
2011	26125	0.00
2012	26125	0.00
2009	26127	0.02
2010	26127	0.03
2011	26127	0.03
2012	26127	0.02
2009	26129	0.00
2010	26129	0.01
2011	26129	0.01
2012	26129	0.00
2009	26131	0.02
2010	26131	0.04
2011	26131	0.03
2012	26131	0.03
2009	26133	0.02
2010	26133	0.03
2011	26133	0.03
2012	26133	0.02
2009	26135	0.02
2010	26135	0.03
2011	26135	0.04
2012	26135	0.03
2009	26137	0.01
2010	26137	0.01
2011	26137	0.01
2012	26137	0.01
2009	26139	0.00

2010	26139	0.00
2011	26139	0.00
2012	26139	0.00
2009	26141	0.02
2010	26141	0.03
2011	26141	0.04
2012	26141	0.03
2009	26143	0.00
2010	26143	0.01
2011	26143	0.01
2012	26143	0.01
2009	26145	0.00
2010	26145	0.00
2011	26145	0.00
2012	26145	0.00
2009	26147	0.00
2010	26147	0.00
2011	26147	0.00
2012	26147	0.00
2009	26149	0.00
2010	26149	0.00
2011	26149	0.00
2012	26149	0.00
2009	26151	0.01
2010	26151	0.01
2011	26151	0.01
2012	26151	0.01
2009	26153	0.02
2010	26153	0.03
2011	26153	0.03
2012	26153	0.02
2009	26155	0.00
2010	26155	0.00
2011	26155	0.00
2012	26155	0.00
2009	26157	0.02
2010	26157	0.03
2011	26157	0.03
2012	26157	0.02
2009	26159	0.01
2010	26159	0.02
2011	26159	0.02
2012	26159	0.01
2009	26161	0.00
2010	26161	0.00

2011	26161	0.00		
2012	26161	0.00		
2009	26163	0.00		
2010	26163	0.00		
2011	26163	0.00		
2012	26163	0.00		
2009	26165	0.01		
2010	26165	0.02		
2011	26165	0.03		
2012	26165	0.01		

Table B.4: 2009 Food SUPIs based in Washtenaw (FIPS 26161)

σ	Comparison FIPS	Food SUPI	BS SUPI GMean	BS SUPI GSD
2.00	26001	13.13	13.19	1.07
2.00	26003	6.23	6.34	1.06
2.00	26005	4.08	4.12	1.04
2.00	26007	3.30	3.36	1.04
2.00	26009	13.95	13.66	1.08
2.00	26011	4.77	4.84	1.05
2.00	26013	6.87	7.03	1.06
2.00	26015	5.18	5.18	1.04
2.00	26017	1.44	1.43	1.02
2.00	26019	12.99	13.08	1.07
2.00	26021	3.80	3.82	1.04
2.00	26023	4.69	4.78	1.04
2.00	26025	3.45	3.43	1.03
2.00	26027	5.54	5.61	1.05
2.00	26029	3.95	3.97	1.04
2.00	26031	4.09	4.11	1.04
2.00	26033	4.12	4.12	1.04
2.00	26035	4.80	4.86	1.04
2.00	26037	1.48	1.47	1.01
2.00	26039	4.09	4.12	1.04
2.00	26041	4.16	4.17	1.04
2.00	26043	4.24	4.25	1.03
2.00	26045	1.80	1.78	1.01
2.00	26047	3.99	4.00	1.04
2.00	26049	1.47	1.46	1.01
2.00	26051	4.80	4.87	1.05
2.00	26053	4.62	4.63	1.04
2.00	26055	2.61	2.62	1.02

2.00	26057	3.81	3.84	1.04
2.00	26059	1.69	1.68	1.02
2.00	26061	5.99	6.10	1.06
2.00	26063	4.12	4.14	1.04
2.00	26065	1.48	1.47	1.01
2.00	26067	4.09	4.15	1.04
2.00	26069	4.48	4.49	1.04
2.00	26071	6.54	6.64	1.06
2.00	26073	1.61	1.60	1.01
2.00	26075	1.70	1.68	1.02
2.00	26077	2.80	2.79	1.03
2.00	26079	5.79	5.90	1.05
2.00	26081	2.44	2.42	1.03
2.00	26085	15.43	15.36	1.07
2.00	26087	1.41	1.41	1.01
2.00	26089	13.59	12.89	1.07
2.00	26091	3.46	3.47	1.04
2.00	26093	1.47	1.46	1.02
2.00	26095	12.45	12.13	1.08
2.00	26097	12.41	12.62	1.07
2.00	26099	1.21	1.21	1.01
2.00	26101	3.69	3.71	1.03
2.00	26103	3.01	2.98	1.03
2.00	26105	5.19	5.25	1.04
2.00	26107	4.21	4.21	1.04
2.00	26109	5.43	5.47	1.04
2.00	26111	1.42	1.42	1.01
2.00	26113	7.37	7.51	1.06
2.00	26115	1.11	1.11	1.01
2.00	26117	3.72	3.76	1.04
2.00	26119	4.95	5.17	1.05
2.00	26121	2.92	2.91	1.03
2.00	26123	4.33	4.37	1.05
2.00	26125	1.16	1.16	1.01
2.00	26127	6.40	6.51	1.05
2.00	26129	3.76	3.77	1.04
2.00	26131	6.70	6.79	1.06
2.00	26133	6.22	6.42	1.05
2.00	26135	5.75	5.93	1.05
2.00	26137	4.02	4.05	1.04
2.00	26139	2.57	2.57	1.03
2.00	26141	5.11	5.38	1.06
2.00	26143	4.08	4.09	1.04
2.00	26145	1.50	1.49	1.01
2.00	26147	1.40	1.39	1.01

2.00	26149	2.06	2.04	1.02
2.00	26151	4.05	4.12	1.04
2.00	26153	6.00	6.11	1.05
2.00	26155	1.74	1.72	1.02
2.00	26157	5.73	5.96	1.05
2.00	26159	4.24	4.31	1.05
2.00	26161	1.00	1.00	1.00
2.00	26163	1.46	1.45	1.01
2.00	26165	4.53	4.61	1.04
4.00	26001	2.42	2.43	1.02
4.00	26003	1.86	1.88	1.02
4.00	26005	1.62	1.63	1.01
4.00	26007	1.53	1.54	1.01
4.00	26009	2.52	2.50	1.02
4.00	26011	1.74	1.75	1.02
4.00	26013	1.92	1.94	1.02
4.00	26015	1.77	1.77	1.01
4.00	26017	1.16	1.15	1.00
4.00	26019	2.43	2.43	1.02
4.00	26021	1.59	1.59	1.01
4.00	26023	1.71	1.72	1.01
4.00	26025	1.54	1.53	1.01
4.00	26027	1.81	1.82	1.02
4.00	26029	1.63	1.63	1.01
4.00	26031	1.67	1.67	1.01
4.00	26033	1.62	1.62	1.01
4.00	26035	1.72	1.73	1.01
4.00	26037	1.16	1.16	1.00
4.00	26039	1.64	1.65	1.01
4.00	26041	1.66	1.66	1.01
4.00	26043	1.67	1.67	1.01
4.00	26045	1.24	1.24	1.00
4.00	26047	1.62	1.62	1.01
4.00	26049	1.15	1.15	1.00
4.00	26051	1.74	1.75	1.02
4.00	26053	1.71	1.71	1.01
4.00	26055	1.42	1.42	1.01
4.00	26057	1.61	1.62	1.01
4.00	26059	1.21	1.21	1.01
4.00	26061	1.88	1.89	1.02
4.00	26063	1.68	1.68	1.01
4.00	26065	1.15	1.15	1.00
4.00	26067	1.61	1.62	1.01
4.00	26069	1.68	1.68	1.01
4.00	26071	1.92	1.93	1.02

4.00	26073	1.20	1.20	1.00
4.00	26075	1.22	1.22	1.01
4.00	26077	1.45	1.44	1.01
4.00	26079	1.81	1.82	1.02
4.00	26081	1.38	1.38	1.01
4.00	26085	2.61	2.61	1.02
4.00	26087	1.13	1.13	1.00
4.00	26089	2.35	2.31	1.02
4.00	26091	1.55	1.56	1.01
4.00	26093	1.15	1.14	1.01
4.00	26095	2.36	2.34	1.02
4.00	26097	2.37	2.38	1.02
4.00	26099	1.07	1.07	1.00
4.00	26101	1.58	1.59	1.01
4.00	26103	1.51	1.50	1.01
4.00	26105	1.77	1.77	1.01
4.00	26107	1.64	1.64	1.01
4.00	26109	1.80	1.81	1.01
4.00	26111	1.14	1.14	1.00
4.00	26113	2.05	2.06	1.02
4.00	26115	1.04	1.04	1.00
4.00	26117	1.60	1.61	1.01
4.00	26119	1.74	1.77	1.02
4.00	26121	1.48	1.48	1.01
4.00	26123	1.69	1.69	1.02
4.00	26125	1.05	1.05	1.00
4.00	26127	1.92	1.93	1.02
4.00	26129	1.61	1.61	1.01
4.00	26131	1.93	1.94	1.02
4.00	26133	1.88	1.90	1.02
4.00	26135	1.82	1.84	1.02
4.00	26137	1.64	1.64	1.01
4.00	26139	1.41	1.41	1.01
4.00	26141	1.78	1.81	1.02
4.00	26143	1.64	1.64	1.01
4.00	26145	1.16	1.16	1.00
4.00	26147	1.13	1.13	1.00
4.00	26149	1.31	1.31	1.01
4.00	26151	1.65	1.65	1.01
4.00	26153	1.87	1.89	1.02
4.00	26155	1.23	1.23	1.01
4.00	26157	1.84	1.86	1.02
4.00	26159	1.64	1.65	1.01
4.00	26161	1.00	1.00	1.00
4.00	26163	1.14	1.14	1.00

4.00	26165	1.69	1.70	1.01
6.00	26001	1.73	1.73	1.01
6.00	26003	1.46	1.47	1.01
6.00	26005	1.35	1.35	1.01
6.00	26007	1.32	1.32	1.01
6.00	26009	1.79	1.78	1.01
6.00	26011	1.42	1.43	1.01
6.00	26013	1.49	1.50	1.01
6.00	26015	1.43	1.43	1.01
6.00	26017	1.11	1.11	1.00
6.00	26019	1.74	1.74	1.01
6.00	26021	1.34	1.34	1.01
6.00	26023	1.40	1.40	1.01
6.00	26025	1.31	1.31	1.01
6.00	26027	1.45	1.45	1.01
6.00	26029	1.36	1.36	1.01
6.00	26031	1.39	1.39	1.01
6.00	26033	1.35	1.35	1.01
6.00	26035	1.40	1.41	1.01
6.00	26037	1.11	1.11	1.00
6.00	26039	1.37	1.37	1.01
6.00	26041	1.38	1.38	1.01
6.00	26043	1.39	1.39	1.01
6.00	26045	1.16	1.15	1.00
6.00	26047	1.35	1.35	1.01
6.00	26049	1.10	1.10	1.00
6.00	26051	1.42	1.43	1.01
6.00	26053	1.40	1.40	1.01
6.00	26055	1.25	1.25	1.01
6.00	26057	1.36	1.36	1.01
6.00	26059	1.13	1.13	1.00
6.00	26061	1.49	1.50	1.01
6.00	26063	1.40	1.41	1.01
6.00	26065	1.09	1.09	1.00
6.00	26067	1.34	1.34	1.01
6.00	26069	1.38	1.38	1.01
6.00	26071	1.50	1.51	1.01
6.00	26073	1.13	1.13	1.00
6.00	26075	1.14	1.14	1.00
6.00	26077	1.27	1.27	1.01
6.00	26079	1.44	1.44	1.01
6.00	26081	1.24	1.24	1.01
6.00	26085	1.83	1.83	1.01
6.00	26087	1.09	1.08	1.00
6.00	26089	1.66	1.64	1.02

6.00	26091	1.32	1.32	1.01
6.00	26093	1.09	1.09	1.00
6.00	26095	1.70	1.69	1.01
6.00	26097	1.70	1.71	1.01
6.00	26099	1.05	1.05	1.00
6.00	26101	1.34	1.34	1.01
6.00	26103	1.31	1.31	1.01
6.00	26105	1.42	1.43	1.01
6.00	26107	1.36	1.36	1.01
6.00	26109	1.44	1.45	1.01
6.00	26111	1.10	1.09	1.00
6.00	26113	1.58	1.59	1.01
6.00	26115	1.03	1.03	1.00
6.00	26117	1.35	1.36	1.01
6.00	26119	1.41	1.42	1.01
6.00	26121	1.29	1.29	1.01
6.00	26123	1.40	1.40	1.01
6.00	26125	1.03	1.03	1.00
6.00	26127	1.50	1.51	1.01
6.00	26129	1.36	1.36	1.01
6.00	26131	1.51	1.51	1.01
6.00	26133	1.48	1.49	1.01
6.00	26135	1.45	1.46	1.01
6.00	26137	1.37	1.37	1.01
6.00	26139	1.25	1.25	1.01
6.00	26141	1.45	1.46	1.01
6.00	26143	1.36	1.36	1.01
6.00	26145	1.10	1.10	1.00
6.00	26147	1.08	1.08	1.00
6.00	26149	1.20	1.19	1.00
6.00	26151	1.37	1.38	1.01
6.00	26153	1.49	1.49	1.01
6.00	26155	1.15	1.15	1.00
6.00	26157	1.46	1.47	1.01
6.00	26159	1.36	1.36	1.01
6.00	26161	1.00	1.00	1.00
6.00	26163	1.09	1.09	1.00
6.00	26165	1.39	1.39	1.01
8.00	26001	1.50	1.50	1.01
8.00	26003	1.32	1.32	1.01
8.00	26005	1.25	1.25	1.01
8.00	26007	1.23	1.23	1.01
8.00	26009	1.55	1.54	1.01
8.00	26011	1.31	1.31	1.01
8.00	26013	1.33	1.34	1.01

8.00	26015	1.31	1.30	1.01
8.00	26017	1.09	1.09	1.00
8.00	26019	1.50	1.51	1.01
8.00	26021	1.24	1.24	1.01
8.00	26023	1.28	1.29	1.01
8.00	26025	1.22	1.22	1.00
8.00	26027	1.32	1.32	1.01
8.00	26029	1.26	1.26	1.01
8.00	26031	1.29	1.29	1.01
8.00	26033	1.24	1.24	1.01
8.00	26035	1.28	1.29	1.01
8.00	26037	1.09	1.09	1.00
8.00	26039	1.27	1.27	1.01
8.00	26041	1.27	1.27	1.01
8.00	26043	1.28	1.28	1.01
8.00	26045	1.12	1.12	1.00
8.00	26047	1.25	1.25	1.01
8.00	26049	1.07	1.07	1.00
8.00	26051	1.31	1.31	1.01
8.00	26053	1.29	1.29	1.01
8.00	26055	1.19	1.19	1.00
8.00	26057	1.26	1.26	1.01
8.00	26059	1.10	1.10	1.00
8.00	26061	1.35	1.36	1.01
8.00	26063	1.30	1.30	1.01
8.00	26065	1.07	1.07	1.00
8.00	26067	1.24	1.24	1.01
8.00	26069	1.27	1.27	1.01
8.00	26071	1.35	1.36	1.01
8.00	26073	1.10	1.10	1.00
8.00	26075	1.11	1.11	1.00
8.00	26077	1.20	1.20	1.00
8.00	26079	1.30	1.30	1.01
8.00	26081	1.18	1.18	1.00
8.00	26085	1.57	1.57	1.01
8.00	26087	1.06	1.06	1.00
8.00	26089	1.42	1.42	1.01
8.00	26091	1.24	1.24	1.01
8.00	26093	1.07	1.07	1.00
8.00	26095	1.47	1.47	1.01
8.00	26097	1.48	1.48	1.01
8.00	26099	1.04	1.04	1.00
8.00	26101	1.24	1.24	1.01
8.00	26103	1.24	1.24	1.00
8.00	26105	1.30	1.30	1.01

8.00	26107	1.25	1.25	1.01
8.00	26109	1.31	1.32	1.01
8.00	26111	1.08	1.07	1.00
8.00	26113	1.42	1.42	1.01
8.00	26115	1.02	1.02	1.00
8.00	26117	1.26	1.26	1.01
8.00	26119	1.29	1.30	1.01
8.00	26121	1.22	1.21	1.00
8.00	26123	1.29	1.29	1.01
8.00	26125	1.02	1.02	1.00
8.00	26127	1.36	1.36	1.01
8.00	26129	1.27	1.27	1.01
8.00	26131	1.35	1.36	1.01
8.00	26133	1.34	1.34	1.01
8.00	26135	1.31	1.32	1.01
8.00	26137	1.27	1.27	1.01
8.00	26139	1.19	1.19	1.00
8.00	26141	1.32	1.33	1.01
8.00	26143	1.26	1.26	1.01
8.00	26145	1.08	1.08	1.00
8.00	26147	1.06	1.06	1.00
8.00	26149	1.15	1.15	1.00
8.00	26151	1.27	1.27	1.01
8.00	26153	1.34	1.35	1.01
8.00	26155	1.12	1.12	1.00
8.00	26157	1.33	1.33	1.01
8.00	26159	1.25	1.25	1.01
8.00	26161	1.00	1.00	1.00
8.00	26163	1.07	1.06	1.00
8.00	26165	1.27	1.28	1.01
10.00	26001	1.38	1.38	1.01
10.00	26003	1.25	1.25	1.01
10.00	26005	1.19	1.19	1.01
10.00	26007	1.19	1.19	1.01
10.00	26009	1.43	1.42	1.01
10.00	26011	1.24	1.25	1.01
10.00	26013	1.26	1.26	1.01
10.00	26015	1.24	1.24	1.01
10.00	26017	1.07	1.07	1.00
10.00	26019	1.39	1.39	1.01
10.00	26021	1.19	1.19	1.00
10.00	26023	1.22	1.22	1.01
10.00	26025	1.17	1.17	1.00
10.00	26027	1.25	1.25	1.01
10.00	26029	1.21	1.21	1.00

10.00	26031	1.23	1.23	1.01
10.00	26033	1.19	1.19	1.00
10.00	26035	1.22	1.22	1.01
10.00	26037	1.07	1.07	1.00
10.00	26039	1.21	1.21	1.01
10.00	26041	1.22	1.22	1.01
10.00	26043	1.23	1.23	1.00
10.00	26045	1.10	1.10	1.00
10.00	26047	1.20	1.20	1.00
10.00	26049	1.06	1.06	1.00
10.00	26051	1.24	1.25	1.01
10.00	26053	1.23	1.23	1.00
10.00	26055	1.15	1.15	1.00
10.00	26057	1.21	1.21	1.01
10.00	26059	1.08	1.08	1.00
10.00	26061	1.28	1.28	1.01
10.00	26063	1.25	1.25	1.00
10.00	26065	1.06	1.06	1.00
10.00	26067	1.18	1.18	1.01
10.00	26069	1.21	1.21	1.01
10.00	26071	1.28	1.28	1.01
10.00	26073	1.09	1.09	1.00
10.00	26075	1.10	1.09	1.00
10.00	26077	1.16	1.16	1.00
10.00	26079	1.23	1.23	1.01
10.00	26081	1.15	1.15	1.00
10.00	26085	1.45	1.44	1.01
10.00	26087	1.05	1.05	1.00
10.00	26089	1.31	1.30	1.01
10.00	26091	1.19	1.19	1.00
10.00	26093	1.06	1.06	1.00
10.00	26095	1.36	1.36	1.01
10.00	26097	1.37	1.37	1.01
10.00	26099	1.03	1.03	1.00
10.00	26101	1.20	1.20	1.00
10.00	26103	1.20	1.20	1.00
10.00	26105	1.23	1.23	1.01
10.00	26107	1.20	1.20	1.00
10.00	26109	1.25	1.25	1.01
10.00	26111	1.06	1.06	1.00
10.00	26113	1.34	1.34	1.01
10.00	26115	1.02	1.02	1.00
10.00	26117	1.21	1.21	1.00
10.00	26119	1.23	1.23	1.01
10.00	26121	1.18	1.18	1.00

10.00	26123	1.23	1.23	1.01
10.00	26125	1.02	1.02	1.00
10.00	26127	1.28	1.28	1.01
10.00	26129	1.22	1.22	1.00
10.00	26131	1.28	1.28	1.01
10.00	26133	1.26	1.27	1.01
10.00	26135	1.24	1.25	1.01
10.00	26137	1.21	1.21	1.00
10.00	26139	1.16	1.16	1.00
10.00	26141	1.26	1.26	1.01
10.00	26143	1.21	1.21	1.00
10.00	26145	1.07	1.07	1.00
10.00	26147	1.05	1.05	1.00
10.00	26149	1.13	1.13	1.00
10.00	26151	1.22	1.22	1.01
10.00	26153	1.27	1.27	1.01
10.00	26155	1.10	1.10	1.00
10.00	26157	1.26	1.26	1.01
10.00	26159	1.20	1.20	1.01
10.00	26161	1.00	1.00	1.00
10.00	26163	1.05	1.05	1.00
10.00	26165	1.22	1.22	1.01
12.00	26001	1.31	1.31	1.01
12.00	26003	1.20	1.20	1.01
12.00	26005	1.16	1.16	1.00
12.00	26007	1.16	1.16	1.00
12.00	26009	1.35	1.35	1.01
12.00	26011	1.21	1.21	1.01
12.00	26013	1.21	1.21	1.01
12.00	26015	1.20	1.20	1.00
12.00	26017	1.07	1.07	1.00
12.00	26019	1.32	1.32	1.01
12.00	26021	1.16	1.16	1.00
12.00	26023	1.19	1.19	1.00
12.00	26025	1.15	1.14	1.00
12.00	26027	1.21	1.21	1.01
12.00	26029	1.18	1.18	1.00
12.00	26031	1.20	1.20	1.00
12.00	26033	1.16	1.16	1.00
12.00	26035	1.19	1.19	1.00
12.00	26037	1.07	1.07	1.00
12.00	26039	1.18	1.18	1.00
12.00	26041	1.18	1.18	1.00
12.00	26043	1.19	1.19	1.00
12.00	26045	1.09	1.09	1.00

12.00	26047	1.17	1.16	1.00
12.00	26049	1.05	1.05	1.00
12.00	26051	1.21	1.21	1.01
12.00	26053	1.19	1.19	1.00
12.00	26055	1.13	1.13	1.00
12.00	26057	1.18	1.18	1.00
12.00	26059	1.07	1.07	1.00
12.00	26061	1.24	1.24	1.01
12.00	26063	1.21	1.21	1.00
12.00	26065	1.05	1.05	1.00
12.00	26067	1.15	1.15	1.00
12.00	26069	1.18	1.18	1.00
12.00	26071	1.23	1.23	1.01
12.00	26073	1.08	1.08	1.00
12.00	26075	1.08	1.08	1.00
12.00	26077	1.14	1.14	1.00
12.00	26079	1.19	1.19	1.01
12.00	26081	1.13	1.13	1.00
12.00	26085	1.37	1.37	1.01
12.00	26087	1.05	1.05	1.00
12.00	26089	1.24	1.24	1.01
12.00	26091	1.16	1.16	1.00
12.00	26093	1.05	1.05	1.00
12.00	26095	1.29	1.29	1.01
12.00	26097	1.30	1.30	1.01
12.00	26099	1.03	1.03	1.00
12.00	26101	1.17	1.17	1.00
12.00	26103	1.17	1.17	1.00
12.00	26105	1.19	1.19	1.01
12.00	26107	1.16	1.16	1.00
12.00	26109	1.21	1.21	1.00
12.00	26111	1.06	1.06	1.00
12.00	26113	1.28	1.29	1.01
12.00	26115	1.02	1.02	1.00
12.00	26117	1.18	1.18	1.00
12.00	26119	1.19	1.19	1.01
12.00	26121	1.15	1.15	1.00
12.00	26123	1.20	1.20	1.00
12.00	26125	1.02	1.02	1.00
12.00	26127	1.24	1.24	1.01
12.00	26129	1.19	1.19	1.00
12.00	26131	1.23	1.23	1.01
12.00	26133	1.22	1.22	1.01
12.00	26135	1.20	1.20	1.01
12.00	26137	1.18	1.18	1.00

12.00	26139	1.14	1.14	1.00
12.00	26141	1.22	1.22	1.01
12.00	26143	1.17	1.17	1.00
12.00	26145	1.06	1.06	1.00
12.00	26147	1.05	1.05	1.00
12.00	26149	1.11	1.11	1.00
12.00	26151	1.19	1.19	1.00
12.00	26153	1.23	1.23	1.01
12.00	26155	1.09	1.09	1.00
12.00	26157	1.21	1.22	1.01
12.00	26159	1.16	1.16	1.00
12.00	26161	1.00	1.00	1.00
12.00	26163	1.05	1.04	1.00
12.00	26165	1.18	1.18	1.00
Inf	26001	1.04	1.04	1.00
Inf	26003	1.02	1.02	1.00
Inf	26005	1.02	1.02	1.00
Inf	26007	1.05	1.05	1.00
Inf	26009	1.07	1.07	1.00
Inf	26011	1.05	1.05	1.00
Inf	26013	1.01	1.02	1.00
Inf	26015	1.04	1.04	1.00
Inf	26017	1.04	1.04	1.00
Inf	26019	1.05	1.05	1.00
Inf	26021	1.03	1.03	1.00
Inf	26023	1.03	1.03	1.00
Inf	26025	1.03	1.03	1.00
Inf	26027	1.04	1.04	1.00
Inf	26029	1.04	1.04	1.00
Inf	26031	1.06	1.06	1.00
Inf	26033	1.02	1.02	1.00
Inf	26035	1.03	1.03	1.00
Inf	26037	1.03	1.03	1.00
Inf	26039	1.04	1.04	1.00
Inf	26041	1.04	1.04	1.00
Inf	26043	1.05	1.05	1.00
Inf	26045	1.03	1.03	1.00
Inf	26047	1.03	1.03	1.00
Inf	26049	1.02	1.02	1.00
Inf	26051	1.05	1.05	1.00
Inf	26053	1.04	1.04	1.00
Inf	26055	1.04	1.04	1.00
Inf	26057	1.05	1.05	1.00
Inf	26059	1.03	1.03	1.00
Inf	26061	1.06	1.06	1.00

Inf	26063	1.07	1.07	1.00
Inf	26065	1.01	1.01	1.00
Inf	26067	1.01	1.01	1.00
Inf	26069	1.03	1.03	1.00
Inf	26071	1.04	1.04	1.00
Inf	26073	1.04	1.04	1.00
Inf	26075	1.04	1.04	1.00
Inf	26077	1.04	1.04	1.00
Inf	26079	1.01	1.01	1.00
Inf	26081	1.04	1.04	1.00
Inf	26085	1.07	1.07	1.00
Inf	26087	1.02	1.02	1.00
Inf	26089	0.98	0.98	1.01
Inf	26091	1.04	1.04	1.00
Inf	26093	1.01	1.01	1.00
Inf	26095	1.03	1.03	1.00
Inf	26097	1.04	1.04	1.00
Inf	26099	1.01	1.01	1.00
Inf	26101	1.04	1.04	1.00
Inf	26103	1.07	1.07	1.00
Inf	26105	1.03	1.03	1.00
Inf	26107	1.02	1.02	1.00
Inf	26109	1.04	1.04	1.00
Inf	26111	1.03	1.03	1.00
Inf	26113	1.08	1.08	1.00
Inf	26115	1.01	1.01	1.00
Inf	26117	1.05	1.05	1.00
Inf	26119	1.03	1.03	1.00
Inf	26121	1.05	1.05	1.00
Inf	26123	1.05	1.05	1.00
Inf	26125	1.00	1.00	1.00
Inf	26127	1.05	1.05	1.00
Inf	26129	1.06	1.06	1.00
Inf	26131	1.04	1.04	1.00
Inf	26133	1.03	1.03	1.00
Inf	26135	1.03	1.03	1.00
Inf	26137	1.04	1.04	1.00
Inf	26139	1.05	1.05	1.00
Inf	26141	1.05	1.05	1.00
Inf	26143	1.04	1.04	1.00
Inf	26145	1.02	1.02	1.00
Inf	26147	1.02	1.02	1.00
Inf	26149	1.04	1.04	1.00
Inf	26151	1.05	1.05	1.00
Inf	26153	1.05	1.05	1.00

Inf	26155	1.04	1.04	1.00
Inf	26157	1.04	1.04	1.00
Inf	26159	1.02	1.02	1.00
Inf	26161	1.00	1.00	1.00
Inf	26163	1.01	1.01	1.00
Inf	26165	1.03	1.03	1.00

Table B.5: 2010 Food SUPIs based in Washtenaw (FIPS 26161)

σ	Comparison FIPS	Food SUPI	BS SUPI GMean	BS SUPI GSD
2.00	26001	13.84	14.09	1.06
2.00	26003	7.14	7.29	1.05
2.00	26005	4.53	4.61	1.04
2.00	26007	4.88	4.89	1.05
2.00	26009	14.90	14.80	1.08
2.00	26011	5.31	5.42	1.05
2.00	26013	7.63	7.83	1.05
2.00	26015	5.88	5.89	1.04
2.00	26017	1.50	1.49	1.02
2.00	26019	13.57	13.79	1.07
2.00	26021	4.17	4.19	1.04
2.00	26023	5.16	5.26	1.04
2.00	26025	3.76	3.77	1.04
2.00	26027	6.03	6.13	1.05
2.00	26029	4.48	4.52	1.04
2.00	26031	4.93	4.93	1.04
2.00	26033	4.63	4.65	1.04
2.00	26035	5.46	5.56	1.05
2.00	26037	1.62	1.61	1.02
2.00	26039	4.40	4.43	1.05
2.00	26041	4.85	4.85	1.04
2.00	26043	4.69	4.69	1.04
2.00	26045	1.90	1.88	1.02
2.00	26047	4.42	4.47	1.04
2.00	26049	1.50	1.49	1.02
2.00	26051	5.38	5.48	1.05
2.00	26053	5.28	5.28	1.04
2.00	26055	2.87	2.89	1.03
2.00	26057	4.37	4.41	1.04
2.00	26059	1.74	1.73	1.02
2.00	26061	6.37	6.49	1.05
2.00	26063	4.62	4.63	1.04

2.00	26065	1.53	1.52	1.01
2.00	26067	4.67	4.75	1.04
2.00	26069	4.87	4.91	1.04
2.00	26071	7.13	7.28	1.05
2.00	26073	1.64	1.63	1.01
2.00	26075	1.78	1.76	1.02
2.00	26077	3.01	3.01	1.03
2.00	26079	6.01	6.12	1.04
2.00	26081	2.68	2.68	1.03
2.00	26085	15.61	15.80	1.06
2.00	26087	1.32	1.31	1.02
2.00	26089	14.48	13.68	1.07
2.00	26091	3.90	3.92	1.04
2.00	26093	1.63	1.61	1.03
2.00	26095	13.65	13.53	1.07
2.00	26097	13.30	13.69	1.06
2.00	26099	1.24	1.24	1.01
2.00	26101	4.14	4.15	1.04
2.00	26103	3.33	3.33	1.03
2.00	26105	5.81	5.89	1.04
2.00	26107	4.61	4.66	1.04
2.00	26109	6.15	6.17	1.04
2.00	26111	1.52	1.51	1.01
2.00	26113	7.79	7.91	1.05
2.00	26115	1.13	1.13	1.01
2.00	26117	4.16	4.24	1.04
2.00	26119	5.51	5.76	1.05
2.00	26121	3.17	3.17	1.03
2.00	26123	4.89	4.98	1.04
2.00	26125	1.20	1.20	1.01
2.00	26127	6.54	6.66	1.05
2.00	26129	4.30	4.34	1.04
2.00	26131	7.28	7.46	1.06
2.00	26133	6.78	7.03	1.05
2.00	26135	6.46	6.65	1.05
2.00	26137	4.38	4.43	1.04
2.00	26139	2.96	2.97	1.03
2.00	26141	5.91	6.14	1.06
2.00	26143	4.59	4.63	1.04
2.00	26145	1.56	1.55	1.02
2.00	26147	1.41	1.40	1.01
2.00	26149	2.17	2.15	1.02
2.00	26151	4.61	4.70	1.04
2.00	26153	6.39	6.50	1.05
2.00	26155	1.92	1.91	1.02

2.00	26157	6.31	6.54	1.05
2.00	26159	4.78	4.88	1.05
2.00	26161	1.00	1.00	1.00
2.00	26163	1.50	1.49	1.02
2.00	26165	5.09	5.20	1.04
4.00	26001	2.46	2.48	1.02
4.00	26003	1.96	1.97	1.02
4.00	26005	1.67	1.68	1.01
4.00	26007	1.73	1.73	1.02
4.00	26009	2.56	2.55	1.02
4.00	26011	1.79	1.80	1.02
4.00	26013	2.00	2.02	1.02
4.00	26015	1.83	1.83	1.01
4.00	26017	1.17	1.17	1.01
4.00	26019	2.46	2.47	1.02
4.00	26021	1.63	1.63	1.01
4.00	26023	1.75	1.76	1.02
4.00	26025	1.57	1.57	1.01
4.00	26027	1.86	1.87	1.02
4.00	26029	1.68	1.69	1.01
4.00	26031	1.75	1.75	1.01
4.00	26033	1.68	1.69	1.01
4.00	26035	1.79	1.80	1.02
4.00	26037	1.19	1.18	1.01
4.00	26039	1.67	1.67	1.02
4.00	26041	1.73	1.73	1.02
4.00	26043	1.71	1.71	1.01
4.00	26045	1.26	1.26	1.01
4.00	26047	1.66	1.67	1.01
4.00	26049	1.16	1.15	1.01
4.00	26051	1.80	1.81	1.02
4.00	26053	1.77	1.77	1.01
4.00	26055	1.45	1.46	1.01
4.00	26057	1.67	1.68	1.01
4.00	26059	1.22	1.22	1.01
4.00	26061	1.90	1.91	1.02
4.00	26063	1.73	1.73	1.01
4.00	26065	1.16	1.16	1.00
4.00	26067	1.68	1.69	1.02
4.00	26069	1.72	1.72	1.01
4.00	26071	1.98	1.99	1.02
4.00	26073	1.20	1.20	1.00
4.00	26075	1.23	1.23	1.01
4.00	26077	1.49	1.49	1.01
4.00	26079	1.85	1.86	1.01

4.00	26081	1.42	1.42	1.01
4.00	26085	2.60	2.61	1.02
4.00	26087	1.11	1.10	1.01
4.00	26089	2.38	2.34	1.02
4.00	26091	1.61	1.61	1.01
4.00	26093	1.18	1.18	1.01
4.00	26095	2.44	2.43	1.02
4.00	26097	2.42	2.44	1.02
4.00	26099	1.08	1.08	1.00
4.00	26101	1.63	1.63	1.01
4.00	26103	1.56	1.56	1.01
4.00	26105	1.83	1.83	1.02
4.00	26107	1.68	1.69	1.01
4.00	26109	1.86	1.87	1.01
4.00	26111	1.17	1.16	1.00
4.00	26113	2.05	2.07	1.02
4.00	26115	1.04	1.04	1.00
4.00	26117	1.65	1.66	1.01
4.00	26119	1.80	1.82	1.02
4.00	26121	1.51	1.51	1.01
4.00	26123	1.74	1.75	1.01
4.00	26125	1.06	1.06	1.00
4.00	26127	1.92	1.93	1.02
4.00	26129	1.67	1.68	1.01
4.00	26131	1.98	1.99	1.02
4.00	26133	1.93	1.95	1.02
4.00	26135	1.90	1.92	1.02
4.00	26137	1.68	1.68	1.01
4.00	26139	1.49	1.49	1.01
4.00	26141	1.86	1.89	1.02
4.00	26143	1.69	1.69	1.01
4.00	26145	1.18	1.17	1.01
4.00	26147	1.13	1.13	1.00
4.00	26149	1.33	1.33	1.01
4.00	26151	1.70	1.71	1.01
4.00	26153	1.90	1.91	1.02
4.00	26155	1.26	1.26	1.01
4.00	26157	1.90	1.92	1.02
4.00	26159	1.70	1.71	1.02
4.00	26161	1.00	1.00	1.00
4.00	26163	1.15	1.15	1.01
4.00	26165	1.76	1.77	1.01
6.00	26001	1.74	1.75	1.01
6.00	26003	1.51	1.52	1.01
6.00	26005	1.37	1.37	1.01

6.00	26007	1.40	1.40	1.01
6.00	26009	1.80	1.80	1.01
6.00	26011	1.44	1.44	1.01
6.00	26013	1.53	1.54	1.01
6.00	26015	1.45	1.45	1.01
6.00	26017	1.11	1.11	1.00
6.00	26019	1.74	1.75	1.01
6.00	26021	1.35	1.35	1.01
6.00	26023	1.41	1.42	1.01
6.00	26025	1.32	1.32	1.01
6.00	26027	1.47	1.47	1.01
6.00	26029	1.38	1.38	1.01
6.00	26031	1.43	1.43	1.01
6.00	26033	1.38	1.38	1.01
6.00	26035	1.43	1.43	1.01
6.00	26037	1.11	1.11	1.00
6.00	26039	1.37	1.37	1.01
6.00	26041	1.41	1.41	1.01
6.00	26043	1.40	1.40	1.01
6.00	26045	1.16	1.16	1.00
6.00	26047	1.37	1.37	1.01
6.00	26049	1.10	1.10	1.00
6.00	26051	1.44	1.45	1.01
6.00	26053	1.42	1.42	1.01
6.00	26055	1.27	1.27	1.01
6.00	26057	1.38	1.38	1.01
6.00	26059	1.14	1.13	1.00
6.00	26061	1.49	1.50	1.01
6.00	26063	1.42	1.42	1.01
6.00	26065	1.10	1.10	1.00
6.00	26067	1.37	1.37	1.01
6.00	26069	1.39	1.40	1.01
6.00	26071	1.53	1.54	1.01
6.00	26073	1.13	1.13	1.00
6.00	26075	1.15	1.14	1.00
6.00	26077	1.29	1.29	1.01
6.00	26079	1.46	1.47	1.01
6.00	26081	1.25	1.25	1.01
6.00	26085	1.82	1.82	1.01
6.00	26087	1.07	1.07	1.00
6.00	26089	1.66	1.64	1.02
6.00	26091	1.35	1.35	1.01
6.00	26093	1.11	1.11	1.01
6.00	26095	1.73	1.73	1.01
6.00	26097	1.72	1.73	1.01

6.00	26099	1.05	1.05	1.00
6.00	26101	1.36	1.36	1.01
6.00	26103	1.34	1.34	1.01
6.00	26105	1.45	1.45	1.01
6.00	26107	1.37	1.37	1.01
6.00	26109	1.47	1.47	1.01
6.00	26111	1.10	1.10	1.00
6.00	26113	1.57	1.58	1.01
6.00	26115	1.02	1.02	1.00
6.00	26117	1.38	1.38	1.01
6.00	26119	1.44	1.45	1.01
6.00	26121	1.31	1.31	1.01
6.00	26123	1.42	1.42	1.01
6.00	26125	1.04	1.04	1.00
6.00	26127	1.51	1.51	1.01
6.00	26129	1.39	1.39	1.01
6.00	26131	1.52	1.53	1.01
6.00	26133	1.50	1.51	1.01
6.00	26135	1.49	1.49	1.01
6.00	26137	1.38	1.38	1.01
6.00	26139	1.30	1.30	1.01
6.00	26141	1.48	1.49	1.01
6.00	26143	1.38	1.39	1.01
6.00	26145	1.11	1.11	1.00
6.00	26147	1.08	1.08	1.00
6.00	26149	1.21	1.20	1.00
6.00	26151	1.40	1.40	1.01
6.00	26153	1.49	1.49	1.01
6.00	26155	1.16	1.16	1.00
6.00	26157	1.49	1.50	1.01
6.00	26159	1.38	1.39	1.01
6.00	26161	1.00	1.00	1.00
6.00	26163	1.09	1.09	1.00
6.00	26165	1.42	1.42	1.01
8.00	26001	1.50	1.51	1.01
8.00	26003	1.35	1.36	1.01
8.00	26005	1.25	1.26	1.01
8.00	26007	1.28	1.28	1.01
8.00	26009	1.55	1.54	1.01
8.00	26011	1.31	1.31	1.01
8.00	26013	1.36	1.37	1.01
8.00	26015	1.31	1.31	1.01
8.00	26017	1.09	1.09	1.00
8.00	26019	1.51	1.51	1.01
8.00	26021	1.25	1.25	1.01

8.00	26023	1.29	1.29	1.01
8.00	26025	1.23	1.23	1.01
8.00	26027	1.32	1.33	1.01
8.00	26029	1.27	1.27	1.01
8.00	26031	1.31	1.31	1.01
8.00	26033	1.26	1.26	1.01
8.00	26035	1.30	1.30	1.01
8.00	26037	1.09	1.08	1.00
8.00	26039	1.26	1.26	1.01
8.00	26041	1.29	1.29	1.01
8.00	26043	1.28	1.28	1.01
8.00	26045	1.12	1.12	1.00
8.00	26047	1.26	1.26	1.01
8.00	26049	1.07	1.07	1.00
8.00	26051	1.31	1.32	1.01
8.00	26053	1.30	1.30	1.01
8.00	26055	1.20	1.20	1.00
8.00	26057	1.27	1.27	1.01
8.00	26059	1.10	1.10	1.00
8.00	26061	1.35	1.35	1.01
8.00	26063	1.31	1.31	1.01
8.00	26065	1.07	1.07	1.00
8.00	26067	1.25	1.25	1.01
8.00	26069	1.28	1.28	1.01
8.00	26071	1.37	1.37	1.01
8.00	26073	1.10	1.10	1.00
8.00	26075	1.11	1.11	1.00
8.00	26077	1.22	1.21	1.00
8.00	26079	1.32	1.33	1.01
8.00	26081	1.19	1.19	1.00
8.00	26085	1.56	1.56	1.01
8.00	26087	1.05	1.05	1.00
8.00	26089	1.42	1.41	1.01
8.00	26091	1.25	1.25	1.01
8.00	26093	1.08	1.08	1.00
8.00	26095	1.49	1.49	1.01
8.00	26097	1.49	1.49	1.01
8.00	26099	1.04	1.04	1.00
8.00	26101	1.25	1.25	1.01
8.00	26103	1.25	1.25	1.01
8.00	26105	1.31	1.31	1.01
8.00	26107	1.26	1.26	1.01
8.00	26109	1.33	1.33	1.01
8.00	26111	1.08	1.08	1.00
8.00	26113	1.40	1.41	1.01

8.00	26115	1.02	1.02	1.00
8.00	26117	1.27	1.27	1.01
8.00	26119	1.31	1.31	1.01
8.00	26121	1.23	1.23	1.00
8.00	26123	1.30	1.30	1.01
8.00	26125	1.03	1.03	1.00
8.00	26127	1.36	1.36	1.01
8.00	26129	1.28	1.28	1.01
8.00	26131	1.36	1.37	1.01
8.00	26133	1.34	1.35	1.01
8.00	26135	1.34	1.34	1.01
8.00	26137	1.27	1.27	1.01
8.00	26139	1.22	1.22	1.01
8.00	26141	1.34	1.35	1.01
8.00	26143	1.27	1.27	1.01
8.00	26145	1.08	1.08	1.00
8.00	26147	1.06	1.06	1.00
8.00	26149	1.16	1.15	1.00
8.00	26151	1.28	1.28	1.01
8.00	26153	1.34	1.34	1.01
8.00	26155	1.12	1.12	1.00
8.00	26157	1.35	1.35	1.01
8.00	26159	1.26	1.27	1.01
8.00	26161	1.00	1.00	1.00
8.00	26163	1.07	1.06	1.00
8.00	26165	1.30	1.30	1.01
10.00	26001	1.39	1.39	1.01
10.00	26003	1.27	1.27	1.01
10.00	26005	1.19	1.20	1.01
10.00	26007	1.22	1.22	1.01
10.00	26009	1.42	1.42	1.01
10.00	26011	1.24	1.24	1.01
10.00	26013	1.28	1.28	1.01
10.00	26015	1.24	1.24	1.01
10.00	26017	1.07	1.07	1.00
10.00	26019	1.39	1.39	1.01
10.00	26021	1.19	1.19	1.00
10.00	26023	1.22	1.23	1.01
10.00	26025	1.18	1.18	1.00
10.00	26027	1.25	1.25	1.01
10.00	26029	1.21	1.21	1.01
10.00	26031	1.24	1.24	1.01
10.00	26033	1.20	1.20	1.01
10.00	26035	1.23	1.23	1.01
10.00	26037	1.07	1.07	1.00

10.00	26039	1.21	1.21	1.01
10.00	26041	1.23	1.23	1.01
10.00	26043	1.22	1.22	1.01
10.00	26045	1.10	1.10	1.00
10.00	26047	1.20	1.20	1.01
10.00	26049	1.06	1.06	1.00
10.00	26051	1.25	1.25	1.01
10.00	26053	1.23	1.23	1.01
10.00	26055	1.16	1.16	1.00
10.00	26057	1.21	1.22	1.01
10.00	26059	1.08	1.08	1.00
10.00	26061	1.27	1.27	1.01
10.00	26063	1.25	1.25	1.01
10.00	26065	1.06	1.06	1.00
10.00	26067	1.19	1.19	1.01
10.00	26069	1.21	1.21	1.01
10.00	26071	1.29	1.29	1.01
10.00	26073	1.08	1.08	1.00
10.00	26075	1.09	1.09	1.00
10.00	26077	1.17	1.17	1.00
10.00	26079	1.25	1.25	1.01
10.00	26081	1.15	1.15	1.00
10.00	26085	1.43	1.43	1.01
10.00	26087	1.04	1.04	1.00
10.00	26089	1.30	1.30	1.01
10.00	26091	1.20	1.20	1.00
10.00	26093	1.07	1.06	1.00
10.00	26095	1.38	1.37	1.01
10.00	26097	1.37	1.38	1.01
10.00	26099	1.03	1.03	1.00
10.00	26101	1.20	1.20	1.01
10.00	26103	1.21	1.21	1.00
10.00	26105	1.24	1.24	1.01
10.00	26107	1.20	1.20	1.01
10.00	26109	1.25	1.25	1.01
10.00	26111	1.07	1.07	1.00
10.00	26113	1.32	1.32	1.01
10.00	26115	1.01	1.01	1.00
10.00	26117	1.22	1.22	1.01
10.00	26119	1.24	1.24	1.01
10.00	26121	1.18	1.18	1.00
10.00	26123	1.23	1.24	1.01
10.00	26125	1.02	1.02	1.00
10.00	26127	1.28	1.28	1.01
10.00	26129	1.22	1.22	1.01

10.00	26131	1.28	1.28	1.01
10.00	26133	1.27	1.27	1.01
10.00	26135	1.26	1.27	1.01
10.00	26137	1.22	1.22	1.01
10.00	26139	1.18	1.18	1.00
10.00	26141	1.27	1.27	1.01
10.00	26143	1.21	1.21	1.01
10.00	26145	1.07	1.07	1.00
10.00	26147	1.05	1.05	1.00
10.00	26149	1.13	1.13	1.00
10.00	26151	1.22	1.22	1.01
10.00	26153	1.26	1.27	1.01
10.00	26155	1.10	1.10	1.00
10.00	26157	1.27	1.28	1.01
10.00	26159	1.20	1.20	1.01
10.00	26161	1.00	1.00	1.00
10.00	26163	1.05	1.05	1.00
10.00	26165	1.23	1.23	1.01
12.00	26001	1.31	1.32	1.01
12.00	26003	1.22	1.22	1.01
12.00	26005	1.16	1.16	1.00
12.00	26007	1.18	1.18	1.01
12.00	26009	1.35	1.35	1.01
12.00	26011	1.20	1.20	1.00
12.00	26013	1.23	1.23	1.01
12.00	26015	1.19	1.19	1.00
12.00	26017	1.07	1.07	1.00
12.00	26019	1.32	1.32	1.01
12.00	26021	1.16	1.16	1.00
12.00	26023	1.18	1.19	1.01
12.00	26025	1.15	1.15	1.00
12.00	26027	1.21	1.21	1.01
12.00	26029	1.18	1.18	1.00
12.00	26031	1.20	1.20	1.00
12.00	26033	1.17	1.17	1.00
12.00	26035	1.19	1.19	1.01
12.00	26037	1.06	1.06	1.00
12.00	26039	1.17	1.17	1.01
12.00	26041	1.19	1.19	1.01
12.00	26043	1.19	1.19	1.00
12.00	26045	1.09	1.09	1.00
12.00	26047	1.17	1.17	1.00
12.00	26049	1.05	1.05	1.00
12.00	26051	1.20	1.21	1.00
12.00	26053	1.19	1.19	1.00

12.00	26055	1.13	1.13	1.00
12.00	26057	1.18	1.18	1.00
12.00	26059	1.07	1.07	1.00
12.00	26061	1.22	1.23	1.01
12.00	26063	1.21	1.21	1.00
12.00	26065	1.05	1.05	1.00
12.00	26067	1.15	1.16	1.01
12.00	26069	1.18	1.18	1.00
12.00	26071	1.24	1.24	1.01
12.00	26073	1.07	1.07	1.00
12.00	26075	1.08	1.08	1.00
12.00	26077	1.15	1.15	1.00
12.00	26079	1.21	1.21	1.01
12.00	26081	1.13	1.13	1.00
12.00	26085	1.35	1.36	1.01
12.00	26087	1.04	1.04	1.00
12.00	26089	1.23	1.23	1.01
12.00	26091	1.16	1.16	1.00
12.00	26093	1.06	1.05	1.00
12.00	26095	1.31	1.30	1.01
12.00	26097	1.30	1.31	1.01
12.00	26099	1.03	1.03	1.00
12.00	26101	1.16	1.16	1.00
12.00	26103	1.18	1.18	1.00
12.00	26105	1.20	1.20	1.01
12.00	26107	1.16	1.16	1.00
12.00	26109	1.21	1.21	1.00
12.00	26111	1.06	1.06	1.00
12.00	26113	1.27	1.27	1.01
12.00	26115	1.01	1.01	1.00
12.00	26117	1.18	1.18	1.01
12.00	26119	1.20	1.20	1.01
12.00	26121	1.16	1.16	1.00
12.00	26123	1.20	1.20	1.00
12.00	26125	1.02	1.02	1.00
12.00	26127	1.23	1.23	1.01
12.00	26129	1.19	1.19	1.00
12.00	26131	1.23	1.23	1.01
12.00	26133	1.22	1.22	1.01
12.00	26135	1.22	1.22	1.01
12.00	26137	1.18	1.18	1.00
12.00	26139	1.16	1.16	1.00
12.00	26141	1.22	1.23	1.01
12.00	26143	1.17	1.18	1.00
12.00	26145	1.06	1.06	1.00

12.00	26147	1.04	1.04	1.00
12.00	26149	1.11	1.11	1.00
12.00	26151	1.19	1.19	1.00
12.00	26153	1.22	1.22	1.01
12.00	26155	1.08	1.08	1.00
12.00	26157	1.23	1.23	1.01
12.00	26159	1.17	1.17	1.01
12.00	26161	1.00	1.00	1.00
12.00	26163	1.04	1.04	1.00
12.00	26165	1.19	1.19	1.00
Inf	26001	1.04	1.04	1.00
Inf	26003	1.02	1.02	1.00
Inf	26005	1.01	1.01	1.00
Inf	26007	1.03	1.03	1.00
Inf	26009	1.06	1.06	1.00
Inf	26011	1.04	1.04	1.00
Inf	26013	1.02	1.02	1.00
Inf	26015	1.02	1.02	1.00
Inf	26017	1.03	1.03	1.00
Inf	26019	1.04	1.04	1.00
Inf	26021	1.02	1.02	1.00
Inf	26023	1.02	1.02	1.00
Inf	26025	1.02	1.02	1.00
Inf	26027	1.03	1.03	1.00
Inf	26029	1.03	1.03	1.00
Inf	26031	1.05	1.05	1.00
Inf	26033	1.02	1.02	1.00
Inf	26035	1.02	1.02	1.00
Inf	26037	1.02	1.02	1.00
Inf	26039	1.03	1.03	1.00
Inf	26041	1.03	1.03	1.00
Inf	26043	1.04	1.03	1.00
Inf	26045	1.03	1.03	1.00
Inf	26047	1.02	1.02	1.00
Inf	26049	1.02	1.02	1.00
Inf	26051	1.04	1.04	1.00
Inf	26053	1.03	1.03	1.00
Inf	26055	1.03	1.03	1.00
Inf	26057	1.04	1.03	1.00
Inf	26059	1.02	1.02	1.00
Inf	26061	1.04	1.04	1.00
Inf	26063	1.06	1.06	1.00
Inf	26065	1.01	1.01	1.00
Inf	26067	1.00	1.00	1.00
Inf	26069	1.02	1.02	1.00

Inf	26071	1.04	1.04	1.00
Inf	26073	1.03	1.03	1.00
Inf	26075	1.03	1.03	1.00
Inf	26077	1.04	1.04	1.00
Inf	26079	1.03	1.03	1.00
Inf	26081	1.04	1.04	1.00
Inf	26085	1.06	1.06	1.00
Inf	26087	1.01	1.01	1.00
Inf	26089	0.96	0.97	1.01
Inf	26091	1.03	1.03	1.00
Inf	26093	1.01	1.01	1.00
Inf	26095	1.03	1.03	1.00
Inf	26097	1.03	1.03	1.00
Inf	26099	1.01	1.01	1.00
Inf	26101	1.03	1.02	1.00
Inf	26103	1.06	1.06	1.00
Inf	26105	1.02	1.02	1.00
Inf	26107	1.01	1.01	1.00
Inf	26109	1.03	1.03	1.00
Inf	26111	1.02	1.02	1.00
Inf	26113	1.05	1.06	1.00
Inf	26115	1.00	1.00	1.00
Inf	26117	1.04	1.04	1.00
Inf	26119	1.03	1.03	1.00
Inf	26121	1.05	1.05	1.00
Inf	26123	1.04	1.04	1.00
Inf	26125	1.00	1.00	1.00
Inf	26127	1.04	1.04	1.00
Inf	26129	1.05	1.05	1.00
Inf	26131	1.03	1.03	1.00
Inf	26133	1.03	1.03	1.00
Inf	26135	1.03	1.03	1.00
Inf	26137	1.04	1.04	1.00
Inf	26139	1.05	1.05	1.00
Inf	26141	1.05	1.05	1.00
Inf	26143	1.03	1.02	1.00
Inf	26145	1.02	1.02	1.00
Inf	26147	1.01	1.01	1.00
Inf	26149	1.04	1.04	1.00
Inf	26151	1.04	1.03	1.00
Inf	26153	1.03	1.03	1.00
Inf	26155	1.02	1.02	1.00
Inf	26157	1.04	1.04	1.00
Inf	26159	1.01	1.01	1.00
Inf	26161	1.00	1.00	1.00

Inf	26163	1.01	1.01	1.00
Inf	26165	1.03	1.03	1.00

Table B.6: 2011 Food SUPIs based in Washtenaw (FIPS 26161)

σ	Comparison FIPS	Food SUPI	BS SUPI GMean	BS SUPI GSD
2.00	26001	11.04	11.34	1.06
2.00	26003	6.50	6.63	1.05
2.00	26005	4.26	4.35	1.04
2.00	26007	4.55	4.62	1.04
2.00	26009	12.65	12.79	1.07
2.00	26011	4.97	5.10	1.04
2.00	26013	6.71	6.91	1.05
2.00	26015	5.25	5.32	1.04
2.00	26017	1.42	1.42	1.01
2.00	26019	10.82	11.18	1.06
2.00	26021	3.98	4.03	1.03
2.00	26023	4.78	4.92	1.04
2.00	26025	3.53	3.55	1.03
2.00	26027	5.72	5.83	1.04
2.00	26029	4.11	4.17	1.03
2.00	26031	4.56	4.60	1.04
2.00	26033	4.27	4.31	1.03
2.00	26035	4.96	5.06	1.04
2.00	26037	1.53	1.53	1.01
2.00	26039	4.15	4.23	1.04
2.00	26041	4.53	4.58	1.03
2.00	26043	4.42	4.46	1.03
2.00	26045	1.84	1.82	1.01
2.00	26047	4.11	4.15	1.03
2.00	26049	1.44	1.43	1.01
2.00	26051	5.02	5.14	1.04
2.00	26053	4.89	4.94	1.04
2.00	26055	2.70	2.72	1.02
2.00	26057	4.05	4.10	1.04
2.00	26059	1.65	1.65	1.02
2.00	26061	5.69	5.83	1.04
2.00	26063	4.23	4.27	1.04
2.00	26065	1.46	1.45	1.01
2.00	26067	4.42	4.52	1.04
2.00	26069	4.47	4.53	1.04
2.00	26071	6.63	6.79	1.04

2.00	26073	1.60	1.60	1.01
2.00	26075	1.74	1.73	1.02
2.00	26077	3.14	3.15	1.03
2.00	26079	5.60	5.68	1.04
2.00	26081	2.62	2.63	1.03
2.00	26085	12.73	13.16	1.06
2.00	26087	1.26	1.26	1.01
2.00	26089	14.24	12.75	1.08
2.00	26091	3.68	3.72	1.03
2.00	26093	1.59	1.58	1.02
2.00	26095	11.23	11.58	1.06
2.00	26097	11.15	11.55	1.06
2.00	26099	1.20	1.20	1.01
2.00	26101	3.84	3.88	1.03
2.00	26103	3.11	3.12	1.03
2.00	26105	5.46	5.56	1.04
2.00	26107	4.44	4.50	1.03
2.00	26109	5.64	5.76	1.04
2.00	26111	1.46	1.46	1.01
2.00	26113	7.61	7.79	1.05
2.00	26115	1.11	1.11	1.01
2.00	26117	3.84	3.94	1.04
2.00	26119	5.10	5.35	1.05
2.00	26121	3.11	3.13	1.03
2.00	26123	4.42	4.52	1.04
2.00	26125	1.17	1.16	1.01
2.00	26127	6.06	6.18	1.05
2.00	26129	3.91	3.96	1.04
2.00	26131	6.82	6.98	1.05
2.00	26133	6.08	6.30	1.05
2.00	26135	6.03	6.22	1.05
2.00	26137	4.16	4.22	1.03
2.00	26139	3.15	3.16	1.03
2.00	26141	5.40	5.62	1.06
2.00	26143	4.21	4.26	1.04
2.00	26145	1.52	1.52	1.01
2.00	26147	1.33	1.32	1.01
2.00	26149	2.08	2.06	1.02
2.00	26151	4.42	4.54	1.04
2.00	26153	5.89	5.99	1.04
2.00	26155	1.85	1.84	1.02
2.00	26157	5.92	6.19	1.05
2.00	26159	4.51	4.64	1.04
2.00	26161	1.00	1.00	1.00
2.00	26163	1.46	1.45	1.01

2.00	26165	4.69	4.83	1.04
4.00	26001	2.28	2.30	1.02
4.00	26003	1.91	1.92	1.02
4.00	26005	1.64	1.65	1.01
4.00	26007	1.69	1.70	1.01
4.00	26009	2.41	2.42	1.02
4.00	26011	1.76	1.77	1.02
4.00	26013	1.90	1.92	1.02
4.00	26015	1.76	1.77	1.01
4.00	26017	1.15	1.15	1.00
4.00	26019	2.27	2.30	1.02
4.00	26021	1.60	1.61	1.01
4.00	26023	1.72	1.73	1.01
4.00	26025	1.54	1.54	1.01
4.00	26027	1.83	1.84	1.02
4.00	26029	1.63	1.64	1.01
4.00	26031	1.71	1.72	1.01
4.00	26033	1.64	1.64	1.01
4.00	26035	1.73	1.74	1.01
4.00	26037	1.16	1.16	1.00
4.00	26039	1.63	1.64	1.01
4.00	26041	1.69	1.69	1.01
4.00	26043	1.67	1.68	1.01
4.00	26045	1.25	1.24	1.00
4.00	26047	1.63	1.64	1.01
4.00	26049	1.14	1.14	1.00
4.00	26051	1.76	1.77	1.01
4.00	26053	1.72	1.73	1.01
4.00	26055	1.42	1.42	1.01
4.00	26057	1.63	1.64	1.01
4.00	26059	1.20	1.20	1.01
4.00	26061	1.83	1.85	1.01
4.00	26063	1.67	1.68	1.01
4.00	26065	1.14	1.14	1.00
4.00	26067	1.66	1.67	1.01
4.00	26069	1.67	1.68	1.01
4.00	26071	1.93	1.94	1.01
4.00	26073	1.19	1.19	1.00
4.00	26075	1.23	1.22	1.01
4.00	26077	1.50	1.50	1.01
4.00	26079	1.81	1.82	1.01
4.00	26081	1.41	1.41	1.01
4.00	26085	2.41	2.44	1.02
4.00	26087	1.09	1.09	1.00
4.00	26089	2.34	2.25	1.03

4.00	26091	1.58	1.59	1.01
4.00	26093	1.18	1.18	1.01
4.00	26095	2.30	2.33	1.02
4.00	26097	2.27	2.29	1.02
4.00	26099	1.07	1.07	1.00
4.00	26101	1.59	1.60	1.01
4.00	26103	1.52	1.52	1.01
4.00	26105	1.80	1.81	1.01
4.00	26107	1.66	1.67	1.01
4.00	26109	1.81	1.82	1.01
4.00	26111	1.15	1.15	1.00
4.00	26113	2.05	2.06	1.02
4.00	26115	1.03	1.03	1.00
4.00	26117	1.61	1.62	1.01
4.00	26119	1.76	1.78	1.02
4.00	26121	1.50	1.50	1.01
4.00	26123	1.68	1.69	1.01
4.00	26125	1.06	1.05	1.00
4.00	26127	1.87	1.88	1.02
4.00	26129	1.62	1.63	1.01
4.00	26131	1.95	1.96	1.02
4.00	26133	1.86	1.88	1.02
4.00	26135	1.85	1.87	1.02
4.00	26137	1.64	1.65	1.01
4.00	26139	1.51	1.51	1.01
4.00	26141	1.81	1.83	1.02
4.00	26143	1.64	1.65	1.01
4.00	26145	1.17	1.16	1.00
4.00	26147	1.11	1.11	1.00
4.00	26149	1.31	1.30	1.01
4.00	26151	1.69	1.70	1.01
4.00	26153	1.84	1.85	1.01
4.00	26155	1.25	1.25	1.01
4.00	26157	1.86	1.89	1.02
4.00	26159	1.66	1.68	1.01
4.00	26161	1.00	1.00	1.00
4.00	26163	1.14	1.14	1.00
4.00	26165	1.71	1.72	1.01
6.00	26001	1.66	1.67	1.01
6.00	26003	1.49	1.50	1.01
6.00	26005	1.35	1.36	1.01
6.00	26007	1.39	1.39	1.01
6.00	26009	1.73	1.73	1.01
6.00	26011	1.43	1.44	1.01
6.00	26013	1.48	1.49	1.01

6.00	26015	1.42	1.42	1.01
6.00	26017	1.10	1.10	1.00
6.00	26019	1.66	1.67	1.01
6.00	26021	1.34	1.34	1.01
6.00	26023	1.40	1.41	1.01
6.00	26025	1.31	1.31	1.01
6.00	26027	1.45	1.46	1.01
6.00	26029	1.36	1.36	1.01
6.00	26031	1.41	1.41	1.01
6.00	26033	1.35	1.35	1.01
6.00	26035	1.40	1.40	1.01
6.00	26037	1.10	1.10	1.00
6.00	26039	1.35	1.36	1.01
6.00	26041	1.38	1.39	1.01
6.00	26043	1.38	1.38	1.01
6.00	26045	1.15	1.15	1.00
6.00	26047	1.36	1.36	1.01
6.00	26049	1.09	1.09	1.00
6.00	26051	1.43	1.43	1.01
6.00	26053	1.40	1.40	1.01
6.00	26055	1.25	1.25	1.01
6.00	26057	1.36	1.36	1.01
6.00	26059	1.13	1.13	1.00
6.00	26061	1.46	1.47	1.01
6.00	26063	1.39	1.39	1.01
6.00	26065	1.09	1.08	1.00
6.00	26067	1.36	1.37	1.01
6.00	26069	1.37	1.38	1.01
6.00	26071	1.51	1.51	1.01
6.00	26073	1.12	1.12	1.00
6.00	26075	1.14	1.14	1.00
6.00	26077	1.30	1.30	1.01
6.00	26079	1.45	1.45	1.01
6.00	26081	1.24	1.25	1.01
6.00	26085	1.73	1.74	1.01
6.00	26087	1.06	1.06	1.00
6.00	26089	1.63	1.59	1.02
6.00	26091	1.33	1.34	1.01
6.00	26093	1.11	1.11	1.00
6.00	26095	1.68	1.69	1.01
6.00	26097	1.65	1.66	1.01
6.00	26099	1.04	1.04	1.00
6.00	26101	1.34	1.34	1.01
6.00	26103	1.31	1.32	1.01
6.00	26105	1.44	1.45	1.01

6.00	26107	1.36	1.37	1.01
6.00	26109	1.44	1.45	1.01
6.00	26111	1.09	1.09	1.00
6.00	26113	1.57	1.58	1.01
6.00	26115	1.02	1.02	1.00
6.00	26117	1.35	1.36	1.01
6.00	26119	1.42	1.43	1.01
6.00	26121	1.30	1.30	1.01
6.00	26123	1.39	1.39	1.01
6.00	26125	1.03	1.03	1.00
6.00	26127	1.48	1.48	1.01
6.00	26129	1.36	1.37	1.01
6.00	26131	1.52	1.52	1.01
6.00	26133	1.47	1.48	1.01
6.00	26135	1.46	1.47	1.01
6.00	26137	1.36	1.37	1.01
6.00	26139	1.30	1.31	1.01
6.00	26141	1.45	1.46	1.01
6.00	26143	1.36	1.36	1.01
6.00	26145	1.10	1.10	1.00
6.00	26147	1.07	1.07	1.00
6.00	26149	1.19	1.19	1.00
6.00	26151	1.39	1.40	1.01
6.00	26153	1.46	1.46	1.01
6.00	26155	1.16	1.16	1.00
6.00	26157	1.47	1.49	1.01
6.00	26159	1.36	1.37	1.01
6.00	26161	1.00	1.00	1.00
6.00	26163	1.09	1.09	1.00
6.00	26165	1.39	1.40	1.01
8.00	26001	1.45	1.46	1.01
8.00	26003	1.34	1.35	1.01
8.00	26005	1.24	1.25	1.01
8.00	26007	1.28	1.28	1.01
8.00	26009	1.50	1.50	1.01
8.00	26011	1.31	1.31	1.01
8.00	26013	1.33	1.33	1.01
8.00	26015	1.29	1.29	1.01
8.00	26017	1.08	1.08	1.00
8.00	26019	1.46	1.46	1.01
8.00	26021	1.23	1.24	1.01
8.00	26023	1.28	1.29	1.01
8.00	26025	1.22	1.22	1.00
8.00	26027	1.32	1.32	1.01
8.00	26029	1.25	1.25	1.01

8.00	26031	1.30	1.30	1.01
8.00	26033	1.25	1.25	1.01
8.00	26035	1.28	1.28	1.01
8.00	26037	1.07	1.07	1.00
8.00	26039	1.25	1.25	1.01
8.00	26041	1.27	1.27	1.01
8.00	26043	1.27	1.27	1.01
8.00	26045	1.12	1.12	1.00
8.00	26047	1.25	1.26	1.01
8.00	26049	1.07	1.07	1.00
8.00	26051	1.30	1.31	1.01
8.00	26053	1.28	1.28	1.01
8.00	26055	1.18	1.18	1.00
8.00	26057	1.26	1.26	1.01
8.00	26059	1.10	1.10	1.00
8.00	26061	1.33	1.33	1.01
8.00	26063	1.28	1.28	1.01
8.00	26065	1.06	1.06	1.00
8.00	26067	1.25	1.25	1.01
8.00	26069	1.26	1.26	1.01
8.00	26071	1.36	1.36	1.01
8.00	26073	1.10	1.10	1.00
8.00	26075	1.11	1.11	1.00
8.00	26077	1.22	1.22	1.00
8.00	26079	1.31	1.32	1.01
8.00	26081	1.18	1.18	1.00
8.00	26085	1.50	1.51	1.01
8.00	26087	1.05	1.05	1.00
8.00	26089	1.39	1.37	1.01
8.00	26091	1.24	1.24	1.01
8.00	26093	1.08	1.08	1.00
8.00	26095	1.46	1.47	1.01
8.00	26097	1.44	1.44	1.01
8.00	26099	1.03	1.03	1.00
8.00	26101	1.24	1.24	1.01
8.00	26103	1.24	1.24	1.00
8.00	26105	1.31	1.31	1.01
8.00	26107	1.25	1.25	1.01
8.00	26109	1.31	1.31	1.01
8.00	26111	1.07	1.07	1.00
8.00	26113	1.41	1.41	1.01
8.00	26115	1.02	1.01	1.00
8.00	26117	1.26	1.26	1.01
8.00	26119	1.29	1.30	1.01
8.00	26121	1.22	1.22	1.00

8.00	26123	1.28	1.28	1.01
8.00	26125	1.03	1.03	1.00
8.00	26127	1.33	1.34	1.01
8.00	26129	1.26	1.27	1.01
8.00	26131	1.36	1.37	1.01
8.00	26133	1.33	1.33	1.01
8.00	26135	1.32	1.33	1.01
8.00	26137	1.26	1.26	1.01
8.00	26139	1.22	1.23	1.00
8.00	26141	1.32	1.33	1.01
8.00	26143	1.26	1.26	1.01
8.00	26145	1.08	1.08	1.00
8.00	26147	1.05	1.05	1.00
8.00	26149	1.14	1.14	1.00
8.00	26151	1.28	1.28	1.01
8.00	26153	1.32	1.32	1.01
8.00	26155	1.12	1.12	1.00
8.00	26157	1.34	1.34	1.01
8.00	26159	1.25	1.25	1.01
8.00	26161	1.00	1.00	1.00
8.00	26163	1.06	1.06	1.00
8.00	26165	1.28	1.28	1.01
10.00	26001	1.35	1.35	1.01
10.00	26003	1.27	1.27	1.01
10.00	26005	1.19	1.19	1.00
10.00	26007	1.22	1.22	1.01
10.00	26009	1.39	1.39	1.01
10.00	26011	1.24	1.25	1.01
10.00	26013	1.25	1.25	1.01
10.00	26015	1.22	1.23	1.00
10.00	26017	1.07	1.07	1.00
10.00	26019	1.35	1.35	1.01
10.00	26021	1.18	1.18	1.00
10.00	26023	1.22	1.22	1.01
10.00	26025	1.17	1.17	1.00
10.00	26027	1.25	1.25	1.01
10.00	26029	1.20	1.20	1.00
10.00	26031	1.24	1.24	1.00
10.00	26033	1.19	1.19	1.00
10.00	26035	1.22	1.22	1.01
10.00	26037	1.06	1.06	1.00
10.00	26039	1.20	1.20	1.01
10.00	26041	1.21	1.21	1.00
10.00	26043	1.21	1.21	1.00
10.00	26045	1.10	1.10	1.00

10.00	26047	1.20	1.20	1.00
10.00	26049	1.06	1.06	1.00
10.00	26051	1.24	1.24	1.01
10.00	26053	1.22	1.22	1.00
10.00	26055	1.14	1.14	1.00
10.00	26057	1.20	1.20	1.00
10.00	26059	1.08	1.08	1.00
10.00	26061	1.26	1.26	1.01
10.00	26063	1.23	1.23	1.00
10.00	26065	1.05	1.05	1.00
10.00	26067	1.19	1.20	1.01
10.00	26069	1.21	1.21	1.00
10.00	26071	1.28	1.28	1.01
10.00	26073	1.08	1.08	1.00
10.00	26075	1.09	1.09	1.00
10.00	26077	1.18	1.18	1.00
10.00	26079	1.24	1.25	1.01
10.00	26081	1.15	1.15	1.00
10.00	26085	1.39	1.39	1.01
10.00	26087	1.04	1.04	1.00
10.00	26089	1.28	1.26	1.01
10.00	26091	1.19	1.19	1.00
10.00	26093	1.07	1.07	1.00
10.00	26095	1.36	1.36	1.01
10.00	26097	1.33	1.34	1.01
10.00	26099	1.03	1.03	1.00
10.00	26101	1.19	1.19	1.00
10.00	26103	1.19	1.20	1.00
10.00	26105	1.24	1.25	1.01
10.00	26107	1.19	1.20	1.00
10.00	26109	1.24	1.24	1.00
10.00	26111	1.06	1.06	1.00
10.00	26113	1.32	1.32	1.01
10.00	26115	1.01	1.01	1.00
10.00	26117	1.20	1.21	1.00
10.00	26119	1.23	1.24	1.01
10.00	26121	1.18	1.18	1.00
10.00	26123	1.22	1.22	1.01
10.00	26125	1.02	1.02	1.00
10.00	26127	1.26	1.26	1.01
10.00	26129	1.21	1.21	1.00
10.00	26131	1.28	1.29	1.01
10.00	26133	1.25	1.26	1.01
10.00	26135	1.25	1.25	1.01
10.00	26137	1.21	1.21	1.00

10.00	26139	1.18	1.18	1.00
10.00	26141	1.26	1.26	1.01
10.00	26143	1.20	1.20	1.00
10.00	26145	1.07	1.07	1.00
10.00	26147	1.04	1.04	1.00
10.00	26149	1.12	1.12	1.00
10.00	26151	1.22	1.23	1.01
10.00	26153	1.25	1.25	1.01
10.00	26155	1.10	1.10	1.00
10.00	26157	1.26	1.27	1.01
10.00	26159	1.19	1.20	1.01
10.00	26161	1.00	1.00	1.00
10.00	26163	1.05	1.05	1.00
10.00	26165	1.22	1.22	1.00
12.00	26001	1.28	1.29	1.01
12.00	26003	1.22	1.22	1.01
12.00	26005	1.16	1.16	1.00
12.00	26007	1.18	1.18	1.00
12.00	26009	1.32	1.32	1.01
12.00	26011	1.21	1.21	1.01
12.00	26013	1.20	1.21	1.01
12.00	26015	1.19	1.19	1.00
12.00	26017	1.06	1.06	1.00
12.00	26019	1.29	1.29	1.01
12.00	26021	1.15	1.15	1.00
12.00	26023	1.18	1.19	1.00
12.00	26025	1.14	1.14	1.00
12.00	26027	1.20	1.21	1.01
12.00	26029	1.16	1.17	1.00
12.00	26031	1.20	1.20	1.00
12.00	26033	1.16	1.16	1.00
12.00	26035	1.18	1.18	1.00
12.00	26037	1.05	1.05	1.00
12.00	26039	1.16	1.16	1.00
12.00	26041	1.18	1.18	1.00
12.00	26043	1.18	1.18	1.00
12.00	26045	1.08	1.08	1.00
12.00	26047	1.17	1.17	1.00
12.00	26049	1.05	1.05	1.00
12.00	26051	1.20	1.20	1.00
12.00	26053	1.18	1.18	1.00
12.00	26055	1.12	1.12	1.00
12.00	26057	1.17	1.17	1.00
12.00	26059	1.07	1.07	1.00
12.00	26061	1.21	1.22	1.01

12.00	26063	1.19	1.19	1.00
12.00	26065	1.04	1.04	1.00
12.00	26067	1.16	1.16	1.00
12.00	26069	1.17	1.17	1.00
12.00	26071	1.23	1.23	1.00
12.00	26073	1.07	1.07	1.00
12.00	26075	1.08	1.08	1.00
12.00	26077	1.15	1.15	1.00
12.00	26079	1.20	1.20	1.00
12.00	26081	1.12	1.12	1.00
12.00	26085	1.32	1.32	1.01
12.00	26087	1.04	1.04	1.00
12.00	26089	1.21	1.20	1.01
12.00	26091	1.16	1.16	1.00
12.00	26093	1.06	1.06	1.00
12.00	26095	1.29	1.30	1.01
12.00	26097	1.27	1.27	1.01
12.00	26099	1.03	1.02	1.00
12.00	26101	1.16	1.16	1.00
12.00	26103	1.17	1.17	1.00
12.00	26105	1.20	1.20	1.00
12.00	26107	1.16	1.16	1.00
12.00	26109	1.20	1.20	1.00
12.00	26111	1.05	1.05	1.00
12.00	26113	1.27	1.27	1.01
12.00	26115	1.01	1.01	1.00
12.00	26117	1.17	1.17	1.00
12.00	26119	1.19	1.20	1.01
12.00	26121	1.15	1.15	1.00
12.00	26123	1.18	1.19	1.00
12.00	26125	1.02	1.02	1.00
12.00	26127	1.22	1.22	1.01
12.00	26129	1.18	1.18	1.00
12.00	26131	1.24	1.24	1.01
12.00	26133	1.21	1.21	1.01
12.00	26135	1.21	1.21	1.01
12.00	26137	1.17	1.17	1.00
12.00	26139	1.16	1.16	1.00
12.00	26141	1.22	1.22	1.01
12.00	26143	1.17	1.17	1.00
12.00	26145	1.06	1.06	1.00
12.00	26147	1.04	1.04	1.00
12.00	26149	1.10	1.10	1.00
12.00	26151	1.19	1.19	1.00
12.00	26153	1.21	1.21	1.00

12.00	26155	1.09	1.09	1.00
12.00	26157	1.22	1.22	1.01
12.00	26159	1.16	1.16	1.00
12.00	26161	1.00	1.00	1.00
12.00	26163	1.04	1.04	1.00
12.00	26165	1.18	1.18	1.00
Inf	26001	1.04	1.03	1.00
Inf	26003	1.03	1.03	1.00
Inf	26005	1.01	1.01	1.00
Inf	26007	1.03	1.03	1.00
Inf	26009	1.05	1.05	1.00
Inf	26011	1.05	1.05	1.00
Inf	26013	1.01	1.01	1.00
Inf	26015	1.02	1.02	1.00
Inf	26017	1.03	1.03	1.00
Inf	26019	1.04	1.04	1.00
Inf	26021	1.02	1.02	1.00
Inf	26023	1.03	1.03	1.00
Inf	26025	1.02	1.02	1.00
Inf	26027	1.03	1.03	1.00
Inf	26029	1.03	1.03	1.00
Inf	26031	1.05	1.05	1.00
Inf	26033	1.01	1.01	1.00
Inf	26035	1.02	1.02	1.00
Inf	26037	1.01	1.01	1.00
Inf	26039	1.02	1.02	1.00
Inf	26041	1.03	1.03	1.00
Inf	26043	1.03	1.03	1.00
Inf	26045	1.03	1.03	1.00
Inf	26047	1.03	1.03	1.00
Inf	26049	1.02	1.02	1.00
Inf	26051	1.04	1.04	1.00
Inf	26053	1.02	1.02	1.00
Inf	26055	1.03	1.03	1.00
Inf	26057	1.03	1.03	1.00
Inf	26059	1.02	1.02	1.00
Inf	26061	1.04	1.04	1.00
Inf	26063	1.05	1.05	1.00
Inf	26065	1.01	1.01	1.00
Inf	26067	1.01	1.01	1.00
Inf	26069	1.02	1.02	1.00
Inf	26071	1.04	1.04	1.00
Inf	26073	1.03	1.03	1.00
Inf	26075	1.03	1.03	1.00
Inf	26077	1.04	1.04	1.00

Inf	26079	1.03	1.03	1.00
Inf	26081	1.03	1.03	1.00
Inf	26085	1.05	1.05	1.00
Inf	26087	1.02	1.02	1.00
Inf	26089	0.95	0.95	1.01
Inf	26091	1.04	1.03	1.00
Inf	26093	1.02	1.02	1.00
Inf	26095	1.04	1.04	1.00
Inf	26097	1.02	1.02	1.00
Inf	26099	1.01	1.01	1.00
Inf	26101	1.03	1.02	1.00
Inf	26103	1.06	1.06	1.00
Inf	26105	1.03	1.03	1.00
Inf	26107	1.01	1.01	1.00
Inf	26109	1.03	1.03	1.00
Inf	26111	1.02	1.02	1.00
Inf	26113	1.06	1.06	1.00
Inf	26115	1.00	1.00	1.00
Inf	26117	1.04	1.04	1.00
Inf	26119	1.03	1.03	1.00
Inf	26121	1.04	1.04	1.00
Inf	26123	1.04	1.04	1.00
Inf	26125	1.00	1.00	1.00
Inf	26127	1.04	1.04	1.00
Inf	26129	1.05	1.05	1.00
Inf	26131	1.04	1.04	1.00
Inf	26133	1.03	1.03	1.00
Inf	26135	1.03	1.03	1.00
Inf	26137	1.03	1.03	1.00
Inf	26139	1.05	1.05	1.00
Inf	26141	1.05	1.05	1.00
Inf	26143	1.03	1.03	1.00
Inf	26145	1.02	1.02	1.00
Inf	26147	1.01	1.01	1.00
Inf	26149	1.04	1.04	1.00
Inf	26151	1.04	1.04	1.00
Inf	26153	1.03	1.03	1.00
Inf	26155	1.03	1.03	1.00
Inf	26157	1.04	1.04	1.00
Inf	26159	1.01	1.01	1.00
Inf	26161	1.00	1.00	1.00
Inf	26163	1.01	1.01	1.00
Inf	26165	1.03	1.03	1.00

Table B.7: 2012 Food SUPIs based in Washtenaw (FIPS 26161)

σ	Comparison FIPS	Food SUPI	BS SUPI GMean	BS SUPI GSD
2.00	26001	8.29	8.48	1.05
2.00	26003	4.86	4.94	1.04
2.00	26005	3.66	3.73	1.03
2.00	26007	4.02	4.12	1.03
2.00	26009	9.58	9.86	1.06
2.00	26011	4.22	4.29	1.04
2.00	26013	5.08	5.17	1.04
2.00	26015	4.55	4.59	1.03
2.00	26017	1.42	1.42	1.01
2.00	26019	7.94	8.05	1.06
2.00	26021	3.63	3.68	1.03
2.00	26023	4.35	4.49	1.04
2.00	26025	3.20	3.22	1.02
2.00	26027	4.57	4.63	1.04
2.00	26029	3.70	3.73	1.03
2.00	26031	4.16	4.22	1.03
2.00	26033	3.80	3.84	1.03
2.00	26035	4.58	4.70	1.04
2.00	26037	1.49	1.49	1.01
2.00	26039	3.79	3.87	1.03
2.00	26041	4.02	4.08	1.03
2.00	26043	4.01	4.08	1.03
2.00	26045	1.81	1.80	1.01
2.00	26047	3.79	3.86	1.03
2.00	26049	1.42	1.41	1.01
2.00	26051	4.25	4.30	1.04
2.00	26053	4.41	4.46	1.03
2.00	26055	2.25	2.25	1.02
2.00	26057	3.65	3.70	1.03
2.00	26059	1.61	1.61	1.02
2.00	26061	5.15	5.27	1.04
2.00	26063	3.74	3.78	1.03
2.00	26065	1.48	1.47	1.01
2.00	26067	3.81	3.91	1.04
2.00	26069	4.01	4.06	1.03
2.00	26071	4.86	4.96	1.04
2.00	26073	1.55	1.55	1.01
2.00	26075	1.69	1.68	1.02
2.00	26077	2.77	2.77	1.02
2.00	26079	4.65	4.72	1.04

2.00	26081	2.34	2.34	1.02
2.00	26085	8.56	8.75	1.05
2.00	26087	1.25	1.24	1.01
2.00	26089	12.95	12.42	1.07
2.00	26091	3.29	3.33	1.03
2.00	26093	1.56	1.55	1.02
2.00	26095	8.99	9.22	1.05
2.00	26097	8.86	9.10	1.06
2.00	26099	1.18	1.18	1.01
2.00	26101	3.41	3.47	1.03
2.00	26103	2.75	2.75	1.02
2.00	26105	5.07	5.16	1.03
2.00	26107	3.99	4.05	1.03
2.00	26109	4.55	4.63	1.03
2.00	26111	1.41	1.41	1.01
2.00	26113	8.29	8.50	1.05
2.00	26115	1.09	1.09	1.01
2.00	26117	3.53	3.60	1.03
2.00	26119	4.89	5.07	1.05
2.00	26121	2.84	2.85	1.03
2.00	26123	3.87	3.95	1.03
2.00	26125	1.15	1.15	1.01
2.00	26127	4.43	4.51	1.04
2.00	26129	3.60	3.66	1.03
2.00	26131	6.39	6.56	1.05
2.00	26133	5.57	5.73	1.04
2.00	26135	5.42	5.59	1.04
2.00	26137	3.90	3.98	1.03
2.00	26139	2.68	2.68	1.02
2.00	26141	4.90	5.09	1.05
2.00	26143	3.79	3.85	1.03
2.00	26145	1.49	1.48	1.01
2.00	26147	1.30	1.30	1.01
2.00	26149	2.02	2.01	1.02
2.00	26151	4.01	4.08	1.04
2.00	26153	4.67	4.75	1.04
2.00	26155	1.84	1.83	1.01
2.00	26157	5.54	5.71	1.05
2.00	26159	4.21	4.33	1.04
2.00	26161	1.00	1.00	1.00
2.00	26163	1.44	1.44	1.01
2.00	26165	4.16	4.28	1.04
4.00	26001	2.07	2.08	1.02
4.00	26003	1.71	1.72	1.01
4.00	26005	1.55	1.56	1.01

4.00	26007	1.60	1.62	1.01
4.00	26009	2.21	2.23	1.02
4.00	26011	1.65	1.66	1.01
4.00	26013	1.71	1.72	1.01
4.00	26015	1.68	1.68	1.01
4.00	26017	1.15	1.15	1.00
4.00	26019	2.04	2.05	1.02
4.00	26021	1.55	1.55	1.01
4.00	26023	1.65	1.67	1.01
4.00	26025	1.48	1.49	1.01
4.00	26027	1.69	1.70	1.01
4.00	26029	1.56	1.57	1.01
4.00	26031	1.65	1.66	1.01
4.00	26033	1.55	1.56	1.01
4.00	26035	1.68	1.69	1.01
4.00	26037	1.15	1.15	1.00
4.00	26039	1.57	1.58	1.01
4.00	26041	1.61	1.62	1.01
4.00	26043	1.61	1.62	1.01
4.00	26045	1.24	1.24	1.00
4.00	26047	1.58	1.59	1.01
4.00	26049	1.14	1.14	1.00
4.00	26051	1.65	1.66	1.01
4.00	26053	1.66	1.66	1.01
4.00	26055	1.33	1.33	1.01
4.00	26057	1.56	1.57	1.01
4.00	26059	1.19	1.19	1.01
4.00	26061	1.74	1.76	1.01
4.00	26063	1.58	1.59	1.01
4.00	26065	1.14	1.14	1.00
4.00	26067	1.57	1.58	1.01
4.00	26069	1.60	1.60	1.01
4.00	26071	1.74	1.75	1.01
4.00	26073	1.18	1.18	1.00
4.00	26075	1.21	1.21	1.01
4.00	26077	1.43	1.43	1.01
4.00	26079	1.70	1.71	1.01
4.00	26081	1.35	1.35	1.01
4.00	26085	2.10	2.11	1.02
4.00	26087	1.09	1.09	1.01
4.00	26089	2.27	2.24	1.03
4.00	26091	1.51	1.52	1.01
4.00	26093	1.17	1.17	1.01
4.00	26095	2.13	2.15	1.02
4.00	26097	2.06	2.07	1.02

4.00	26099	1.06	1.06	1.00
4.00	26101	1.52	1.53	1.01
4.00	26103	1.44	1.44	1.01
4.00	26105	1.73	1.74	1.01
4.00	26107	1.60	1.60	1.01
4.00	26109	1.68	1.69	1.01
4.00	26111	1.13	1.13	1.00
4.00	26113	2.10	2.12	1.02
4.00	26115	1.03	1.03	1.00
4.00	26117	1.56	1.57	1.01
4.00	26119	1.73	1.74	1.02
4.00	26121	1.45	1.45	1.01
4.00	26123	1.60	1.61	1.01
4.00	26125	1.05	1.05	1.00
4.00	26127	1.68	1.69	1.01
4.00	26129	1.57	1.58	1.01
4.00	26131	1.91	1.92	1.02
4.00	26133	1.81	1.82	1.01
4.00	26135	1.78	1.80	1.02
4.00	26137	1.59	1.60	1.01
4.00	26139	1.42	1.42	1.01
4.00	26141	1.75	1.77	1.02
4.00	26143	1.57	1.58	1.01
4.00	26145	1.16	1.16	1.00
4.00	26147	1.10	1.10	1.00
4.00	26149	1.29	1.29	1.01
4.00	26151	1.62	1.63	1.01
4.00	26153	1.68	1.69	1.01
4.00	26155	1.25	1.25	1.00
4.00	26157	1.82	1.83	1.02
4.00	26159	1.62	1.63	1.01
4.00	26161	1.00	1.00	1.00
4.00	26163	1.14	1.13	1.00
4.00	26165	1.63	1.64	1.01
6.00	26001	1.57	1.57	1.01
6.00	26003	1.39	1.40	1.01
6.00	26005	1.30	1.31	1.01
6.00	26007	1.33	1.34	1.01
6.00	26009	1.65	1.66	1.01
6.00	26011	1.37	1.37	1.01
6.00	26013	1.38	1.38	1.01
6.00	26015	1.37	1.37	1.01
6.00	26017	1.10	1.10	1.00
6.00	26019	1.56	1.56	1.01
6.00	26021	1.30	1.31	1.01

6.00	26023	1.36	1.37	1.01
6.00	26025	1.27	1.27	1.01
6.00	26027	1.39	1.39	1.01
6.00	26029	1.31	1.32	1.01
6.00	26031	1.37	1.37	1.01
6.00	26033	1.30	1.30	1.01
6.00	26035	1.37	1.38	1.01
6.00	26037	1.09	1.09	1.00
6.00	26039	1.32	1.32	1.01
6.00	26041	1.34	1.35	1.01
6.00	26043	1.34	1.35	1.01
6.00	26045	1.15	1.15	1.00
6.00	26047	1.33	1.33	1.01
6.00	26049	1.09	1.09	1.00
6.00	26051	1.37	1.37	1.01
6.00	26053	1.36	1.37	1.01
6.00	26055	1.20	1.20	1.00
6.00	26057	1.32	1.32	1.01
6.00	26059	1.12	1.12	1.00
6.00	26061	1.40	1.41	1.01
6.00	26063	1.33	1.33	1.01
6.00	26065	1.09	1.09	1.00
6.00	26067	1.31	1.32	1.01
6.00	26069	1.33	1.33	1.01
6.00	26071	1.41	1.42	1.01
6.00	26073	1.12	1.12	1.00
6.00	26075	1.13	1.13	1.00
6.00	26077	1.26	1.26	1.01
6.00	26079	1.39	1.39	1.01
6.00	26081	1.21	1.21	1.01
6.00	26085	1.58	1.59	1.01
6.00	26087	1.06	1.06	1.00
6.00	26089	1.60	1.59	1.02
6.00	26091	1.29	1.30	1.01
6.00	26093	1.10	1.10	1.00
6.00	26095	1.60	1.60	1.01
6.00	26097	1.53	1.54	1.01
6.00	26099	1.04	1.04	1.00
6.00	26101	1.29	1.30	1.01
6.00	26103	1.27	1.27	1.00
6.00	26105	1.40	1.40	1.01
6.00	26107	1.33	1.33	1.01
6.00	26109	1.37	1.38	1.01
6.00	26111	1.09	1.09	1.00
6.00	26113	1.60	1.60	1.01

6.00	26115	1.02	1.02	1.00
6.00	26117	1.32	1.33	1.01
6.00	26119	1.40	1.41	1.01
6.00	26121	1.26	1.27	1.01
6.00	26123	1.34	1.34	1.01
6.00	26125	1.03	1.03	1.00
6.00	26127	1.38	1.38	1.01
6.00	26129	1.33	1.33	1.01
6.00	26131	1.50	1.51	1.01
6.00	26133	1.44	1.45	1.01
6.00	26135	1.42	1.43	1.01
6.00	26137	1.33	1.34	1.01
6.00	26139	1.25	1.25	1.01
6.00	26141	1.42	1.43	1.01
6.00	26143	1.32	1.32	1.01
6.00	26145	1.10	1.10	1.00
6.00	26147	1.06	1.06	1.00
6.00	26149	1.18	1.18	1.00
6.00	26151	1.35	1.35	1.01
6.00	26153	1.37	1.37	1.01
6.00	26155	1.16	1.15	1.00
6.00	26157	1.45	1.46	1.01
6.00	26159	1.33	1.34	1.01
6.00	26161	1.00	1.00	1.00
6.00	26163	1.08	1.08	1.00
6.00	26165	1.35	1.36	1.01
8.00	26001	1.39	1.39	1.01
8.00	26003	1.27	1.27	1.01
8.00	26005	1.21	1.21	1.01
8.00	26007	1.23	1.24	1.01
8.00	26009	1.45	1.46	1.01
8.00	26011	1.26	1.27	1.01
8.00	26013	1.25	1.26	1.01
8.00	26015	1.26	1.26	1.01
8.00	26017	1.08	1.08	1.00
8.00	26019	1.39	1.39	1.01
8.00	26021	1.21	1.21	1.01
8.00	26023	1.25	1.26	1.01
8.00	26025	1.19	1.19	1.00
8.00	26027	1.28	1.28	1.01
8.00	26029	1.22	1.22	1.00
8.00	26031	1.27	1.27	1.01
8.00	26033	1.20	1.21	1.00
8.00	26035	1.26	1.26	1.01
8.00	26037	1.07	1.07	1.00

8.00	26039	1.22	1.23	1.01
8.00	26041	1.24	1.24	1.01
8.00	26043	1.24	1.24	1.00
8.00	26045	1.11	1.11	1.00
8.00	26047	1.23	1.23	1.01
8.00	26049	1.07	1.07	1.00
8.00	26051	1.26	1.26	1.01
8.00	26053	1.25	1.25	1.00
8.00	26055	1.14	1.14	1.00
8.00	26057	1.23	1.23	1.01
8.00	26059	1.09	1.09	1.00
8.00	26061	1.28	1.28	1.01
8.00	26063	1.24	1.24	1.01
8.00	26065	1.06	1.06	1.00
8.00	26067	1.22	1.22	1.01
8.00	26069	1.23	1.23	1.01
8.00	26071	1.29	1.30	1.01
8.00	26073	1.09	1.09	1.00
8.00	26075	1.10	1.10	1.00
8.00	26077	1.19	1.19	1.00
8.00	26079	1.27	1.28	1.01
8.00	26081	1.16	1.16	1.00
8.00	26085	1.40	1.41	1.01
8.00	26087	1.05	1.05	1.00
8.00	26089	1.38	1.37	1.01
8.00	26091	1.21	1.21	1.00
8.00	26093	1.08	1.08	1.00
8.00	26095	1.41	1.42	1.01
8.00	26097	1.35	1.36	1.01
8.00	26099	1.03	1.03	1.00
8.00	26101	1.21	1.21	1.01
8.00	26103	1.20	1.20	1.00
8.00	26105	1.28	1.28	1.01
8.00	26107	1.23	1.23	1.01
8.00	26109	1.26	1.26	1.01
8.00	26111	1.07	1.07	1.00
8.00	26113	1.42	1.42	1.01
8.00	26115	1.01	1.01	1.00
8.00	26117	1.23	1.24	1.01
8.00	26119	1.28	1.29	1.01
8.00	26121	1.19	1.19	1.00
8.00	26123	1.24	1.24	1.01
8.00	26125	1.02	1.02	1.00
8.00	26127	1.27	1.27	1.01
8.00	26129	1.24	1.24	1.01

8.00	26131	1.35	1.36	1.01
8.00	26133	1.31	1.32	1.01
8.00	26135	1.29	1.30	1.01
8.00	26137	1.23	1.24	1.01
8.00	26139	1.18	1.18	1.00
8.00	26141	1.30	1.31	1.01
8.00	26143	1.22	1.22	1.01
8.00	26145	1.08	1.08	1.00
8.00	26147	1.05	1.05	1.00
8.00	26149	1.13	1.13	1.00
8.00	26151	1.25	1.25	1.01
8.00	26153	1.26	1.26	1.01
8.00	26155	1.12	1.12	1.00
8.00	26157	1.32	1.32	1.01
8.00	26159	1.23	1.23	1.01
8.00	26161	1.00	1.00	1.00
8.00	26163	1.06	1.06	1.00
8.00	26165	1.25	1.25	1.01
10.00	26001	1.30	1.30	1.01
10.00	26003	1.21	1.21	1.01
10.00	26005	1.16	1.16	1.00
10.00	26007	1.18	1.18	1.01
10.00	26009	1.36	1.36	1.01
10.00	26011	1.21	1.21	1.01
10.00	26013	1.19	1.19	1.01
10.00	26015	1.20	1.20	1.00
10.00	26017	1.07	1.07	1.00
10.00	26019	1.30	1.30	1.01
10.00	26021	1.16	1.16	1.00
10.00	26023	1.19	1.20	1.01
10.00	26025	1.15	1.15	1.00
10.00	26027	1.22	1.22	1.01
10.00	26029	1.17	1.17	1.00
10.00	26031	1.21	1.21	1.00
10.00	26033	1.15	1.15	1.00
10.00	26035	1.20	1.20	1.01
10.00	26037	1.05	1.05	1.00
10.00	26039	1.17	1.17	1.00
10.00	26041	1.19	1.19	1.00
10.00	26043	1.19	1.19	1.00
10.00	26045	1.09	1.09	1.00
10.00	26047	1.18	1.18	1.00
10.00	26049	1.06	1.06	1.00
10.00	26051	1.21	1.21	1.01
10.00	26053	1.20	1.20	1.00

10.00	26055	1.12	1.12	1.00
10.00	26057	1.18	1.18	1.00
10.00	26059	1.07	1.07	1.00
10.00	26061	1.21	1.22	1.01
10.00	26063	1.19	1.19	1.00
10.00	26065	1.05	1.05	1.00
10.00	26067	1.16	1.17	1.01
10.00	26069	1.18	1.18	1.00
10.00	26071	1.23	1.23	1.01
10.00	26073	1.08	1.08	1.00
10.00	26075	1.09	1.08	1.00
10.00	26077	1.15	1.15	1.00
10.00	26079	1.21	1.22	1.01
10.00	26081	1.13	1.13	1.00
10.00	26085	1.31	1.32	1.01
10.00	26087	1.04	1.04	1.00
10.00	26089	1.27	1.26	1.01
10.00	26091	1.16	1.17	1.00
10.00	26093	1.06	1.06	1.00
10.00	26095	1.32	1.32	1.01
10.00	26097	1.26	1.27	1.01
10.00	26099	1.03	1.03	1.00
10.00	26101	1.16	1.16	1.00
10.00	26103	1.16	1.16	1.00
10.00	26105	1.21	1.22	1.01
10.00	26107	1.18	1.18	1.00
10.00	26109	1.20	1.20	1.00
10.00	26111	1.05	1.05	1.00
10.00	26113	1.33	1.33	1.01
10.00	26115	1.01	1.01	1.00
10.00	26117	1.19	1.19	1.00
10.00	26119	1.22	1.22	1.01
10.00	26121	1.16	1.16	1.00
10.00	26123	1.19	1.19	1.00
10.00	26125	1.02	1.02	1.00
10.00	26127	1.21	1.21	1.01
10.00	26129	1.19	1.19	1.00
10.00	26131	1.28	1.28	1.01
10.00	26133	1.24	1.25	1.01
10.00	26135	1.23	1.23	1.01
10.00	26137	1.18	1.18	1.00
10.00	26139	1.15	1.15	1.00
10.00	26141	1.24	1.24	1.01
10.00	26143	1.17	1.17	1.00
10.00	26145	1.06	1.06	1.00

10.00	26147	1.04	1.04	1.00
10.00	26149	1.11	1.11	1.00
10.00	26151	1.19	1.20	1.01
10.00	26153	1.20	1.20	1.01
10.00	26155	1.10	1.10	1.00
10.00	26157	1.25	1.26	1.01
10.00	26159	1.17	1.18	1.01
10.00	26161	1.00	1.00	1.00
10.00	26163	1.05	1.05	1.00
10.00	26165	1.19	1.19	1.01
12.00	26001	1.25	1.25	1.01
12.00	26003	1.17	1.17	1.00
12.00	26005	1.13	1.13	1.00
12.00	26007	1.15	1.15	1.00
12.00	26009	1.30	1.30	1.01
12.00	26011	1.18	1.18	1.00
12.00	26013	1.15	1.15	1.00
12.00	26015	1.17	1.17	1.00
12.00	26017	1.06	1.06	1.00
12.00	26019	1.25	1.25	1.01
12.00	26021	1.13	1.13	1.00
12.00	26023	1.16	1.16	1.01
12.00	26025	1.12	1.12	1.00
12.00	26027	1.18	1.18	1.00
12.00	26029	1.14	1.14	1.00
12.00	26031	1.18	1.18	1.00
12.00	26033	1.12	1.12	1.00
12.00	26035	1.16	1.17	1.00
12.00	26037	1.04	1.04	1.00
12.00	26039	1.14	1.14	1.00
12.00	26041	1.16	1.16	1.00
12.00	26043	1.15	1.16	1.00
12.00	26045	1.08	1.08	1.00
12.00	26047	1.15	1.15	1.00
12.00	26049	1.05	1.05	1.00
12.00	26051	1.17	1.17	1.00
12.00	26053	1.16	1.16	1.00
12.00	26055	1.10	1.10	1.00
12.00	26057	1.15	1.15	1.00
12.00	26059	1.06	1.06	1.00
12.00	26061	1.18	1.18	1.00
12.00	26063	1.16	1.16	1.00
12.00	26065	1.04	1.04	1.00
12.00	26067	1.13	1.14	1.00
12.00	26069	1.15	1.15	1.00

12.00	26071	1.19	1.20	1.01
12.00	26073	1.07	1.07	1.00
12.00	26075	1.07	1.07	1.00
12.00	26077	1.13	1.13	1.00
12.00	26079	1.18	1.18	1.01
12.00	26081	1.11	1.11	1.00
12.00	26085	1.26	1.26	1.01
12.00	26087	1.04	1.04	1.00
12.00	26089	1.20	1.20	1.01
12.00	26091	1.14	1.14	1.00
12.00	26093	1.05	1.05	1.00
12.00	26095	1.26	1.26	1.01
12.00	26097	1.21	1.21	1.01
12.00	26099	1.02	1.02	1.00
12.00	26101	1.13	1.13	1.00
12.00	26103	1.14	1.14	1.00
12.00	26105	1.17	1.18	1.00
12.00	26107	1.14	1.15	1.00
12.00	26109	1.17	1.17	1.00
12.00	26111	1.05	1.05	1.00
12.00	26113	1.28	1.28	1.01
12.00	26115	1.01	1.01	1.00
12.00	26117	1.16	1.16	1.00
12.00	26119	1.18	1.18	1.01
12.00	26121	1.13	1.13	1.00
12.00	26123	1.16	1.16	1.00
12.00	26125	1.01	1.01	1.00
12.00	26127	1.18	1.18	1.00
12.00	26129	1.16	1.16	1.00
12.00	26131	1.23	1.23	1.01
12.00	26133	1.20	1.20	1.01
12.00	26135	1.19	1.19	1.01
12.00	26137	1.15	1.15	1.00
12.00	26139	1.13	1.12	1.00
12.00	26141	1.20	1.20	1.01
12.00	26143	1.14	1.14	1.00
12.00	26145	1.06	1.06	1.00
12.00	26147	1.03	1.03	1.00
12.00	26149	1.09	1.09	1.00
12.00	26151	1.16	1.16	1.00
12.00	26153	1.16	1.16	1.00
12.00	26155	1.08	1.08	1.00
12.00	26157	1.21	1.21	1.01
12.00	26159	1.14	1.14	1.00
12.00	26161	1.00	1.00	1.00

12.00	26163	1.04	1.04	1.00
12.00	26165	1.16	1.16	1.00
Inf	26001	1.03	1.03	1.00
Inf	26003	1.02	1.02	1.00
Inf	26005	1.01	1.01	1.00
Inf	26007	1.01	1.01	1.00
Inf	26009	1.06	1.06	1.00
Inf	26011	1.03	1.03	1.00
Inf	26013	0.99	0.99	1.00
Inf	26015	1.02	1.02	1.00
Inf	26017	1.03	1.03	1.00
Inf	26019	1.04	1.04	1.00
Inf	26021	1.01	1.01	1.00
Inf	26023	1.02	1.02	1.00
Inf	26025	1.01	1.01	1.00
Inf	26027	1.03	1.03	1.00
Inf	26029	1.02	1.01	1.00
Inf	26031	1.04	1.04	1.00
Inf	26033	0.99	0.99	1.00
Inf	26035	1.02	1.01	1.00
Inf	26037	1.01	1.01	1.00
Inf	26039	1.01	1.01	1.00
Inf	26041	1.02	1.02	1.00
Inf	26043	1.02	1.02	1.00
Inf	26045	1.03	1.03	1.00
Inf	26047	1.02	1.02	1.00
Inf	26049	1.02	1.02	1.00
Inf	26051	1.03	1.03	1.00
Inf	26053	1.02	1.02	1.00
Inf	26055	1.02	1.02	1.00
Inf	26057	1.02	1.02	1.00
Inf	26059	1.02	1.02	1.00
Inf	26061	1.01	1.01	1.00
Inf	26063	1.03	1.03	1.00
Inf	26065	1.01	1.01	1.00
Inf	26067	1.00	1.00	1.00
Inf	26069	1.01	1.01	1.00
Inf	26071	1.04	1.04	1.00
Inf	26073	1.03	1.03	1.00
Inf	26075	1.03	1.03	1.00
Inf	26077	1.03	1.03	1.00
Inf	26079	1.03	1.03	1.00
Inf	26081	1.03	1.03	1.00
Inf	26085	1.04	1.04	1.00
Inf	26087	1.02	1.02	1.00

Inf	26089	0.95	0.95	1.01
Inf	26091	1.02	1.02	1.00
Inf	26093	1.01	1.01	1.00
Inf	26095	1.04	1.04	1.00
Inf	26097	0.99	0.99	1.00
Inf	26099	1.01	1.01	1.00
Inf	26101	1.01	1.01	1.00
Inf	26103	1.04	1.04	1.00
Inf	26105	1.02	1.01	1.00
Inf	26107	1.01	1.01	1.00
Inf	26109	1.02	1.02	1.00
Inf	26111	1.02	1.02	1.00
Inf	26113	1.06	1.06	1.00
Inf	26115	1.00	1.00	1.00
Inf	26117	1.03	1.03	1.00
Inf	26119	1.02	1.02	1.00
Inf	26121	1.03	1.03	1.00
Inf	26123	1.03	1.03	1.00
Inf	26125	1.00	1.00	1.00
Inf	26127	1.03	1.03	1.00
Inf	26129	1.03	1.03	1.00
Inf	26131	1.04	1.04	1.00
Inf	26133	1.03	1.03	1.00
Inf	26135	1.02	1.02	1.00
Inf	26137	1.02	1.02	1.00
Inf	26139	1.03	1.03	1.00
Inf	26141	1.04	1.04	1.00
Inf	26143	1.01	1.01	1.00
Inf	26145	1.02	1.02	1.00
Inf	26147	1.01	1.01	1.00
Inf	26149	1.03	1.03	1.00
Inf	26151	1.03	1.03	1.00
Inf	26153	1.01	1.01	1.00
Inf	26155	1.03	1.03	1.00
Inf	26157	1.04	1.04	1.00
Inf	26159	1.00	1.00	1.00
Inf	26161	1.00	1.00	1.00
Inf	26163	1.01	1.01	1.00
Inf	26165	1.02	1.02	1.00

BIBLIOGRAPHY

[1] Katharine G. Abraham et al. *Big Data for 21st Century Economic Statistics*. University of Chicago Press, 2019. URL: http://www.nber.org/books/abra-7.

[2] Bettina Aten. "Report on Interarea Price Levels". In: (2005).

[3] Bettina H Aten. "Regional price parities and real regional income for the United States". In: *Social Indicators Research* 131.1 (2017), pp. 123–143.

[4] Bettina Aten, Eric Figueroa, Troy Martin, et al. *Notes on Estimating the Multi-Year Regional Price Parities by 16 Expenditure Categories: 2005-2009*. BEA, 2011.

[5] C Lanier Benkard and Patrick Bajari. "Hedonic price indexes with unobserved product characteristics, and application to personal computers". In: *Journal of Business & Economic Statistics* 23.1 (2005), pp. 61–75.

[6] United States Census Bureau. *Census Quickfacts: Keeweenaw County, Michigan*. Accessed: 2019-06-08. URL: https://www.census.gov/quickfacts/fact/table/keweenawcountymichigan/PST045218.

[7] Council for Community & Economic Research. *2016 County Level Cost of Living Index*. Accessed: 2019-05-18. URL: http://coli.org/wp-content/uploads/2016/06/County-Samples.zip.

[8] Council for Community & Economic Research. *Cost of Living Index Manual*. 2016.

[9] Council for Community & Economic Research. *Frequently Asked Questions - C2ER Cost of Living Index*. Accessed: 2019-05-17. URL: http://coli.org/wp-content/uploads/2016/05/Frequently-Asked-Questions-%E2%80%93-C2ER-Cost-of-Living-Index.docx.

[10] J De Haan et al. "An overview of price index methods for scanner data". In: *room document for the meeting of the UNECE Group of Experts on Consumer Price Indices*. 2016.

[11] Erwin Diewert. "Weighted country product dummy variable regressions and index number formulae". In: *Review of Income and Wealth* 51.4 (2005), pp. 561–570.

[12] W Erwin Diewert. "Exact and superlative index numbers". In: *Journal of econometrics* 4.2 (1976), pp. 115–145.

[13] W Erwin Diewert. *The economic theory of index numbers: a survey*. Department of Economics, University of British Columbia, 1979.

[14] Gabriel Ehrlich et al. *Minding Your Ps and Qs: Going from Micro to Macro in Measuring Prices and Quantities*. Working Paper 25465. National Bureau of Economic Research, Jan. 2019. DOI: `10.3386/w25465`. URL: `http://www.nber.org/papers/w25465`.

[15] Gabriel Ehrlich et al. "Re-Engineering Key National Economic Indicators". In: Katharine G. Abraham et al. *Big Data for 21st Century Economic Statistics*. University of Chicago Press, Apr. 2019. URL: `http://www.nber.org/chapters/c14266`.

[16] Statistical Office of the European Communities. *Export and import price index manual: Theory and practice*. International Monetary Fund, 2009.

[17] Robert C Feenstra. "New product varieties and the measurement of international prices". In: *The American Economic Review* (1994), pp. 157–177.

[18] Robert C. Feenstra and Matthew D. Shapiro. "Introduction to 'Scanner Data and Price Indexes'". In: *Scanner Data and Price Indexes*. University of Chicago Press, Jan. 2003, pp. 1–14. URL: `http://www.nber.org/chapters/c9725`.

[19] Christopher A Field and Alan H Welsh. "Bootstrapping clustered data". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 69.3 (2007), pp. 369–390.

[20] Jenny Fitzgerald and Owen Shoemaker. "Evaluating the Consumer Price Index Using Nielsen's Scanner Data". In: (2013).

[21] Roy C Geary. "A note on the comparison of exchange rates and purchasing power between countries". In: *Journal of the Royal Statistical Society. Series A (General)* 121.1 (1958), pp. 97–99.

[22] IMF ILO. "Eurostat and World Bank (2004), Consumer Price Index Manual Theory and Practice (CPI-Manual), ed. by P". In: *Hill, Geneva* ().

[23] Lorraine Ivancic, W Erwin Diewert, and Kevin J Fox. "Scanner data, time aggregation and the construction of price indexes". In: *Journal of Econometrics* 161.1 (2011), pp. 24–35.

[24] Jahyeong Koo, Keith R Phillips, and Fiona D Sigalla. "Measuring regional cost of living". In: *Journal of Business & Economic Statistics* 18.1 (2000), pp. 127–136.

[25] Tomislav V Kovandzic and John J Sloan. "Police levels and crime rates revisited: A county-level analysis from Florida (1980–1998)". In: *Journal of Criminal Justice* 30.1 (2002), pp. 65–76.

[26] Tomislav V Kovandzic and Lynne M Vieraitis. "The effect of county-level prison population growth on crime rates". In: *Criminology & Public Policy* 5.2 (2006), pp. 213–244.

[27] Irving B Kravis, Alan Heston, and Robert Summers. *World product and income: International comparisons of real gross product*. The World Bank, 1982.

[28] Bureau of Labor Statistics. *Consumer Price Index Frequently Asked Questions*. Accessed: 2019-05-20. URL: https://www.bls.gov/cpi/questions-and-answers.htm.

[29] Bureau of Labor Statistics. *Consumer Price Index Handbook of Methods*. Accessed: 2019-05-20. URL: https://www.bls.gov/opub/hom/pdf/cpihom.pdf.

[30] *Main Features - Use of Transactions data in the Australian CPI*. Feb. 2019. URL: https://www.abs.gov.au/ausstats/abs@.nsf/7d12b0f6763c78caca257061001cc588/40fc971083782000ca25768e002c845b!OpenDocument.

[31] Naoko Muramatsu. "County-level income inequality and depression among older Americans". In: *Health services research* 38.6p2 (2003), pp. 1863–1884.

[32] Christopher JL Murray et al. "Eight Americas: investigating mortality disparities across races, counties, and race-counties in the United States". In: *PLoS medicine* 3.9 (2006), e260.

[33] DS Rao and Marcel P Timmer. "Purchasing Power Parities for Industry Comparisons Using Weighted Elteto–Koves–Szulc (EKS) Methods". In: *Review of Income and Wealth* 49.4 (2003), pp. 491–511.

[34] Stephen J Redding and David E Weinstein. *A unified approach to estimating demand and welfare*. Tech. rep. National Bureau of Economic Research, 2016.

[35] Stephen J Redding and David E Weinstein. *Measuring Aggregate Price Indexes with Taste Shocks: Theory and Evidence for CES Preferences*. Working Paper 22479. Revised in May 2019. National Bureau of Economic Research, Aug. 2016. DOI: 10.3386/w22479. URL: http://www.nber.org/papers/w22479.

[36] Marcel Van Kints. "Making Greater Use of Transactions Data to compile the Australian CPI". In: *15th Meeting of the Ottawa Group*. 2017.

[37] Hal R Varian. *Microeconomic analysis*. 338.5 V299m 1978. WW Norton, 1978.