BEA Working Paper Series, WP2026-21

Containing Consumption Zones

Authors Andrea Batch, Abe Dunn, Daniel Fernandez, Mahsa Gholizadeh, Samuel Stern, and Steven Weinberg[1]

Contact abe.dunn@bea.gov

Date November 2026

Abstract Researchers at the U.S. Bureau of Economic Analysis recently introduced a new economic-geographic unit, the consumption zone (ConZ), to characterize the geography of consumption, much as commuting zones (CZs) are used to delineate labor markets. In this paper, we revisit ConZs using more recent Mastercard transaction data in the USA and revise the hierarchical clustering algorithm applied in prior work to more closely align with the methodology originally developed for CZs. We also introduce a new containment-based stopping rule that determines the number of clusters based on the share of economic activity occurring within each zone, which provides a stopping criteria that is both easily understood and economically meaningful. Setting a baseline containment level of 80%, we obtain consumption zones that differ substantially from earlier estimates and that vary markedly across industries, underscoring how the geography of consumption depends on the sector studied. For nonlocal services where consumer travel is common, such as accommodations, performing arts, and other recreational activities, we report alternative nonlocal consumption zones constructed at a 50% containment threshold, which yields more useful clusters for these industries. Overall, we view these zones as preferable to prior estimates because they follow the original algorithm of CZs more precisely and apply more transparent and meaningful stopping rule.

Keywords Consumption zones (ConZs), consumption, geographic markets, concentration, card data.

JEL codes R12, P23

[1]Batch, Dunn, and Gholizadeh: U.S. Bureau of Economic Analysis (BEA); Fernandez, Stern, and Weinberg: Mastercard Economics Institute. We thank Benjamin Bridgman, Nick Bunker, Aditya Jyoti, Swapnil Kulkarni, and Yashaswi D for discussions and contributing to data analysis. We thank Alberto Cavallo for helping to encourage the BEA and Mastercard collaboration.

The views expressed in this paper are those of the authors and do not necessarily represent the U.S. Bureau of Economic Analysis or the U.S. Department of Commerce.

Suggested citation:
Batch, Andrea, Abe Dunn, Daniel Fernandez, Mahsa Gholizadeh, Samuel Stern, and Steven Weinberg. *Containing Consumption Zones.* Working paper no. WP2026–21. Washington, DC: U.S. Bureau of Economic Analysis, November 2026, https://doi.org/10.66137/CBIL1862.

1. Introduction

To study economic differences across geographic markets, it is important to have a systematic methodology for defining those markets, ideally one grounded in an underlying economic pattern in the data. Tolbert and Sizer (1987) applied a hierarchical clustering algorithm to county-to-county commuting flows to delineate the geographic extent of a labor market, an area now referred to as a commuting zone (CZ). More recently, Foote, Kutzbach and Vilhuber (2021) and Fowler (2024) revisited this approach using more recent data. Building on Foote, Kutzbach and Vilhuber (2021), Batch et al. (2025) applied the same algorithm to consumer spending across geographic markets to construct what they call consumption zones (ConZs), which are geographic areas that define a market for consumption. The methodology rests on the intuition that areas with large bilateral flows are economically similar and should be grouped into the same cluster.[2]

The data underlying the work of Batch et al. (2025) are now more than a decade old. In this paper, we revisit this topic using more recent consumption flow data using information from Mastercard transaction data in the United States at the county level for the years 2019 and 2024. In addition, we review the prior algorithm and find that the precise code applied in Foote, Kutzbach and Vilhuber (2021) and followed by Batch et al. (2025) made some implicit assumptions that do not explicitly match with the algorithm of Tolbert and Sizer (1987). In particular, the implementation of the clustering algorithm effectively normalized the distance metric across industries.[3] This feature has the practical advantage of allowing a common clustering height to be applied across industries. At the same time, because industries differ substantially in the geographic dispersion of their consumption flows, a common height need not imply a common degree of economic containment. We therefore explore an alternative stopping rule that makes this economic property explicit. In this paper, we apply the algorithm as originally formulated and also describe an approach that generates more meaningful and intuitive zones across industries.

The first step of the hierarchical clustering algorithm is to define a distance metric. Following Tolbert and Sizer (1987), distance is determined by the bilateral flows between areas, so that larger flows imply a smaller distance. The algorithm iteratively groups areas until all remaining distances exceed a chosen "height," at which point it stops. Determining this stopping height is a central step. The implementation of Foote, Kutzbach and Vilhuber (2021) applied by Batch et al. (2025) implicitly normalized the heights across industries, allowing a single height to apply to all of them. This normalization, however, differs from Tolbert and Sizer (1987) formulation and is less straightforward.

[2] Batch et al. (2025) argue that, under certain assumptions, the algorithm relates to cross-price elasticities across areas.

[3] This is important, as it meant that a single "height", which determines the stopping rule for the clustering algorithm, could be the same across industries. Batch et al. (2025) used this feature to define ConZs. The rationale for this normalization was not made explicit in Foote, Kutzbach and Vilhuber (2021) or Batch et al. (2025).

We find that the original Tolbert and Sizer (1987) metric, based on bilateral flows, implies very different distances across industries. Flows for grocery stores, which are highly local, differ sharply from those for food services and restaurants, which involve more travel, and the two generate distinct distance metrics. The key challenge is to construct a stopping rule that yields economically meaningful zones across such different industries. Although we explored several alternatives, our preferred stopping rule is based on containment, the share of aggregate consumption that remains within a zone's borders, indicating that economic activity and the determination of prices are largely determined within these borders. We view this as a natural metric in guiding the construction of ConZs as Fowler (2024) states, "…containment is the measure most closely associated with the intent of CZ's…." Containment can also be understood as the amount of economic activity that does not leak outside the zones. The concept is related to the macroeconomic notion of leakage used in studying cross-sectional fiscal multipliers (Chodorow-Reich, 2019) and regional input–output multipliers (Bess, Ambargis et al., 2011), both of which seek to account for economic activity that remains within a local geographic area.

As a baseline, we set containment to 80%. That is, we continue the algorithm's iterative merges until 80% of all spending occurs within the same cluster. An advantage of this approach is that every industry's zones share the same property that, on average, 80% of economic activity occurs within the zone. The ConZs can therefore be applied consistently across industries, and zones in different industries carry a common economic meaning. This is useful for researchers studying how economic shocks or policies affect particular geographic areas. When consumption zones are used, one can be more confident that most consumption activity takes place within the defined zones.

For some industries, containment is more limited. For accommodations, performing arts, and other recreational industries, consumption naturally occurs over a wider region. For these industries, we construct alternative nonlocal or "traded" zones using a 50% containment threshold. While this fundamentally changes the meaning of the zones, it provides more useful geographies for understanding the markets for these industries.

Improving ConZs is important for researchers studying local economic shocks to consumption or understanding local consumption markets, as argued by Batch et al. (2025). In addition, accurate information about local geographic markets may lead to more efficient investment, as suggested by Chi (2025), who finds that retail and restaurant establishments fail at higher rates when they rely on outdated Census demographics, with the effect concentrated precisely among industries that serve small trade areas, such as grocery stores and restaurants. That finding assumes a local market over which the relevant demographic information is measured; yet the geography of these markets differs sharply across industries and locations, as we show in this paper. Defining consumption markets from spending directly, as we do with ConZs, offers a more accurate method for measuring local conditions, which may in turn support more productive investment across

geographic markets.

The updated estimates yield four main findings. An explicit containment target gives the zones a common, interpretable definition and, as we show below, redistributes zones markedly across industries even when the aggregate level of containment is held fixed. The containment level then determines the overall number of zones: at 80% containment the algorithm produces 819 aggregate consumption zones from the 2015 Fiserv data, compared with the 1,235 in Batch et al. (2025). Based on this threshold, the resulting zones are closer in scale to commuting zones in size, unlike the earlier estimates. They also carry a transparent interpretation when comparing across industries, as they are all set to attain the same containment level.

Second, consumption markets differ sharply across industries. Highly local industries, such as food and beverage stores, building materials, and ambulatory care, require many small zones to reach the containment target, implying that the relevant market is often close to, or perhaps even smaller than, a county. Industries that involve travel or tourism—accommodations, performing arts, and recreation—result in a few large regional markets when containment is set at 80%. This motivates the lower-threshold of 50% for "traded" zones we report for these sectors.

Third, the Mastercard-based zones are generally stable between 2019 and 2024. The aggregate number of zones changes only modestly for most categories, and most counties remain grouped with the same neighbors, despite the well-documented shifts in consumption associated with the pandemic and the rise of remote work.

Fourth, we revisit the prior application of Batch et al. (2025) and look at industry concentration applying the updated zones. We find the updated zones deliver concentration measures that are qualitatively consistent with prior work, with high levels of concentration for general merchandise stores, gas stations, food and beverage stores, and building material. As before, we find that for these industries, ConZs typically imply lower measured concentration than counties while still flagging the same industries as relatively concentrated as in Batch et al. (2025). Together, these findings suggest that the updated zones provide a practical and economically meaningful geography for studying local consumption markets.

We also explore a number of alternative methodologies that could improve on features of the original algorithm by Tolbert and Sizer (1987). We view this as a valuable contribution since the clustering algorithm, originally developed to construct CZs, is more than 30 years old. While we identify many possible limitations to the original CZ algorithm, we found no obvious alternative that clearly outperformed the original algorithm. The alternatives tend to add complexity without substantially more benefit. That said, these alternatives may prove more useful with additional research or when moving to finer levels of geographic detail (e.g., zip code level data).

2. Review of the Algorithm

We construct a dissimilarity matrix D that captures the economic distance between counties:

$$D_{i,j} = 1 - \frac{f_{i,j} + f_{j,i}}{\min(C_i, C_j)}. \tag{1}$$

Here, $f_{i,j}$ and $f_{j,i}$ denote bilateral spending flows, and C_i and C_j are county measures of consumption by consumers in these counties. By construction, the distance metric is a value between 0 and 1.[4] The distance metric links directly to the concept of containment, since by definition it is the dollar of revenue contained per dollar of revenue of the smaller area's consumption. Scaling by the minimum consumption emphasizes links between smaller areas and larger economic centers.[5] A parallel formulation is used for commuting zones, replacing spending flows and revenues with commuter counts and resident populations.[6]

The matrix D is an input to an agglomerative hierarchical clustering procedure. Starting with each county as its own cluster, distances between clusters C_K and C_L are defined as the average pairwise dissimilarity:

$$D_{K,L} = \frac{1}{N_K N_L} \sum_{i \in C_K} \sum_{j \in C_L} D_{i,j}. \tag{2}$$

Clusters are merged sequentially based on the smallest $D_{K,L}$, with distances recomputed after each merge. And the process continues until all counties are grouped or a cutoff threshold H is reached, so that clusters with $D_{K,L} > H$ are not combined.

The choice of H reflects a tradeoff: low values produce a high number of clusters, while high values can force together economically unrelated areas, generating "greedy" clusters.

Although not stated in the paper, Foote, Kutzbach and Vilhuber (2021) applied a distance metric that was based on the Euclidian distance between areas, not directly applying equation 1. This effectively led to some normalization of the distance metric across industries, so that a single H could be applied to different industries, which is the approach taken by Batch et al. (2025), who apply the code of Foote, Kutzbach and Vilhuber (2021) directly. Upon reviewing the code, we believe this distance metric was not intentional and is not justified by either empirical arguments or theory.[7] In this paper, we propose a distinct approach for selecting a stopping rule across

[4] By construction, $D_{i,j} = D_{j,i}$. We additionally impose the condition $D_{ii} = 0$, and cap the dissimilarity values at 0 or 1 when they fall outside the interval $[0, 1]$.

[5] In the prior work, the firms' revenue was placed in the denominator in the formula, but in practice, using consumption is preferred, as there are fewer zeros, so it works well across areas where there is little production occurring.

[6] For commuting, $D_{i,j} = 1 - \frac{e_{i,j}+e_{j,i}}{\min(ResPop_i, ResPop_j)}$.

[7] Additional details on the code applied in Foote, Kutzbach and Vilhuber (2021) and Batch et al. (2025) are available in Section C.

industries, which we view as both intuitive and an improvement relative to Batch et al. (2025).

3. Selecting Stopping Rule for H

The selection of the stopping rule for each industry follows the idea that the goal of the zones is containment of economic activity. This goal fits well with the distance metric that is the area where the greatest amount of flows occur per dollar of economic activity in the smaller area. In this paper we select a cutoff for containment of 80%, meaning that 80% of the dollars spent are from consumers within the same zone.

For an industry with county-to-county spending flows f_{ij} (spending by residents of county j at merchants in county i) and a partition of counties into zones $\{Z_1, \ldots, Z_M\}$, define *containment* as the share of total spending whose consumer and merchant counties fall in the same zone:[8]

$$\Gamma = \frac{\sum_m \sum_{i \in Z_m} \sum_{j \in Z_m} f_{ij}}{\sum_i \sum_j f_{ij}}. \tag{3}$$

Containment rises weakly with each merge, from a minimum at the all-singleton partition (where only within-county spending $\sum f_{ii}$ is contained) to 1 when all counties form a single zone. Our stopping rule selects the partition along the agglomerative path at which Γ first reaches the target level (80% in the baseline, 50% for traded zones).

This threshold requires that a large majority of spending remains within the constructed area, while allowing for some degree of leakage across neighboring areas. For instance, if one is studying a local economic shock to income or prices that might affect consumption and one applies ConZs as the geographic unit, one would expect that most of the shock would be contained within these zones.

The algorithm is simple and essentially chooses the H that produces a selected containment value of 80% for the corresponding industry, which which provides a common and directly interpretable characterization of the zones across industries. As mentioned in the introduction, a single contain-

[8] An alternative method for calibrating H would be to control for minimum containment (or even a quantile of the containment distribution) at the county level:

$$\Gamma_{\min} \equiv \min_k \Gamma_k \equiv \min_k \frac{\sum_{j \in Z_k} f_{kj}}{\sum f_{kj}}.$$

We ultimately decided to use definition (3) instead, as the focus of this paper is on average outcomes, as we viewed this as easier to communicate to users.

ment value does not work for all industries or applications. In particular, we suggest an alternative traded services consumption zones where the containment value is just 50%.

Our goal is to construct economically meaningful geographies that capture broad patterns of consumption activity. Our choice of an 80% containment threshold is not meant to identify a precise theoretical cutoff, but is instead motivated by related work. As mentioned previously, the concept behind CZs is the idea of containing economic activity for labor around an economic center. Although a specific containment threshold is not applied in the analysis, it is thought to be a fundamental motivation for the algorithm and is an important descriptive property of CZs (Fowler (2024)). About 93% of commuters live in their CZs, although it is important to note that commuting is distinct from consumption, so the exact same threshold may not apply. In another area of research for health care, Dartmouth researchers for the Evaluative Clinical Sciences (1996) define hospital referral regions (HRRs) using hospital claims data, with zip-code level information. They similarly use flows across areas in their process of defining regional markets and specify a minimum of at least 65% of patients obtaining care within the region as one of their criteria, which suggests the average share of patients coming within the HRR is typically over 65%. While they do not use containment as an explicit stopping rule, they do use it as a diagnostic. The median share that is local is 87.5%, although HRRs are intended to be very broadly defined geographic areas. Dartmouth also develops smaller markets called hospital service areas, which typically have a much lower containment share, with the median share being around 70%. Also in the context of hospital services consumption applied in antitrust work, the Elzinga–Hogarty approach used flow-based criteria to define geographic markets, which distinguished between a "weak" market at approximately 75% internal flows and a "strong" market at approximately 90% internal flows (Frech, Langenfeld and McCluer (2004)).[9] Our 80% threshold therefore falls around the range from this past literature to characterize markets with substantial internal activity, while avoiding the very large regions that would result from imposing a 90% threshold. Selecting 80% containment requires that most spending occur within the zone while still allowing for some cross-zone leakage. Importantly, the threshold holds in the aggregate rather than for every individual zone, so some zones may have more or less than 80% containment, allowing for flexibility in how zones are defined locally.

3.0.1. Alternative Approaches

While our main estimates use the baseline hierarchical clustering algorithm from Tolbert and Sizer (1987), we also considered several alternatives. First, we examined alternative containment thresholds. Lower thresholds generate smaller and more numerous zones, while higher thresholds

[9] We do not interpret this threshold as meaningful for rigorously defined antitrust markets, as issues with applying containment for defining antitrust markets have been highlighted in past work. Important critiques of the Elzinga–Hogarty approach have emphasized limitations of using flow-based criteria to define individual geographic markets for antitrust purposes (Capps et al., 2001, 2019)

generate larger zones that contain a greater share of spending. We have selected 80% as we view this as a reasonable baseline, but there is no definitively correct threshold amount for all purposes. We make alternative thresholds available so that users can select a geography that better matches their empirical setting.

We also considered alternative distance metrics. As discussed in Appendix B, the original bilateral-flow metric has several attractive features, but it can generate unintuitive rankings in certain stylized cases. For example, the metric may be sensitive to differences in county size, may not fully capture cases where a small county is economically dependent on a larger hub, and may understate relationships between counties when most spending occurs locally. The appendix therefore considers a sequence of modified distance metrics that use alternative normalizations based on inflows, outflows, local spending, and bilateral capacity. These alternatives are designed to address particular weaknesses of the baseline metric and are evaluated using a set of stylized examples.

The baseline algorithm relies on bilateral flows between counties, so that two counties are treated as closer when there is greater spending activity between them. This approach has the advantage of being transparent and closely related to the commuting-zone methodology. However, it also has limitations. For example, the bilateral-flow measure does not fully account for substitution patterns involving third counties. If residents of county C regularly choose between merchants in counties A and B, that relationship may be relevant for defining the market even if the direct bilateral flow between A and B is limited. We considered distance metrics that incorporate these broader substitution patterns, but at the county level these alternatives produced results that were broadly similar to the baseline estimates. This likely reflects the fact that most observed spending flows are concentrated across a relatively small number of nearby counties.

While the discussion in the previous two paragraphs highlights potential issues with the distance metric, there are also potential issues with the hierarchical algorithm weighting. Under the baseline approach, once two counties are combined into a cluster, each county receives equal weight when calculating distances to other clusters. This can be problematic when a very small county is grouped with a much larger county, since the two counties may then contribute equally to subsequent merging decisions despite representing very different amounts of population or spending. We explored procedures that reweight counties after they are merged, so that small and large counties do not contribute equally to subsequent clustering decisions, but these approaches were more computationally intensive and did not produce qualitatively superior zones.

Overall, we find that these alternatives are useful for clarifying the tradeoffs embedded in the clustering procedure, but they do not clearly dominate the baseline approach in our county-level application. The more complex metrics often address one concern while introducing additional complexity or creating new potentially problematic cases. We therefore retain the baseline

algorithm for the main estimates, while viewing the alternative metrics as a useful direction for future work, especially at finer geographic levels such as zip codes.

4. Data

The primary transaction data used in this analysis are provided by Mastercard, one of the world's largest payment networks. According to the Nilson Report (HSN Consultants, Inc. (2025)), Mastercard accounted for approximately 26 percent of card spending for goods and services in the United States in 2024, behind Visa (61 percent). While Mastercard does not capture the majority of card transactions, its coverage is substantial in absolute terms—representing nearly $3 trillion in annual U.S. purchase volume (HSN Consultants, Inc. (2026))—across millions of cardholders and merchant locations. Importantly, Mastercard cards are issued by thousands of financial institutions across the United States, from the largest national banks to small community banks and credit unions, ensuring broad geographic and demographic coverage. Because our identification strategy relies on the relative distribution of spending across county pairs rather than on total spending levels—what is required for our analysis is that the geographic pattern of consumer-to-merchant flows observed through Mastercard be representative of the broader population of card transactions. Since both Mastercard and Visa are accepted at virtually all U.S. merchant locations, and cards from both networks are issued across the full spectrum of financial institutions and consumer segments, there is no obvious mechanism that would cause Mastercard cardholders in a given county to systematically differ from other cardholders where they shop.

For this study, we use anonymized and aggregated spending data corresponding to card-present transactions conducted at merchant locations across the United States. Card-present transactions—those in which the physical card (or its digital equivalent) is presented at a point of sale—are particularly well suited to the study of geographic consumption patterns, as they require the cardholder to be physically present at the merchant's location. This ensures that the spending flows we observe reflect actual spatial movements of consumers between their county of residence and the county in which the merchant operates. By excluding card-not-present (i.e., online or remote) transactions, we avoid conflating e-commerce activity—where the consumer's physical location is irrelevant—with in-person consumption that defines the geographic reach of local product markets. At the transaction level, the data record the total dollar amount and the merchant's category code but do not contain information on specific items purchased, their prices, or quantities. For our purposes, however, the relevant unit of observation is the aggregate dollar flow between county pairs within an industry, so item-level granularity is not required. The data encompass both credit and debit card transactions (including consumer and corporate cards), thereby capturing a broad cross-section of consumer payment behavior. The inclusion of debit cards is particularly important for representativeness. Although credit cards have recently overtaken debit in total payment share, debit cards remain the preferred in-person payment instrument for U.S.

consumers—with 40 percent of consumers preferring debit for in-person transactions (Federal Reserve Financial Services (2026))—and are especially prevalent in everyday spending categories such as groceries and gasoline (PULSE Network (2025)). Moreover, lower- and middle-income households rely disproportionately on debit, so excluding debit transactions would systematically underweight these consumers in the geographic flow matrices.

The information underpinning our estimates is sourced from Mastercard's Data Warehouse, which records the universe of credit and debit card transactions cleared through the Mastercard network.[10] Additional detail regarding the Mastercard data are available in Appendix Section A.1.

4.1. Fiserv Data

The historical consumption zone estimates are based on aggregated and anonymized payment card transactions from Fiserv, a major payment processor. These data were used to construct the original consumption zones and are described in detail in Dunn and Gholizadeh (2025) and Batch et al. (2025). In brief, the transactions are aggregated into county-to-county spending flow matrices that link consumers' counties of residence to the merchant counties where they spend, by industry. We use these flows here only to construct the zones to compare to prior work as we can hold the data fixed to that from prior work while varying the algorithm. The results are discussed in Section 6. The updated zones rely on the Mastercard data described above.

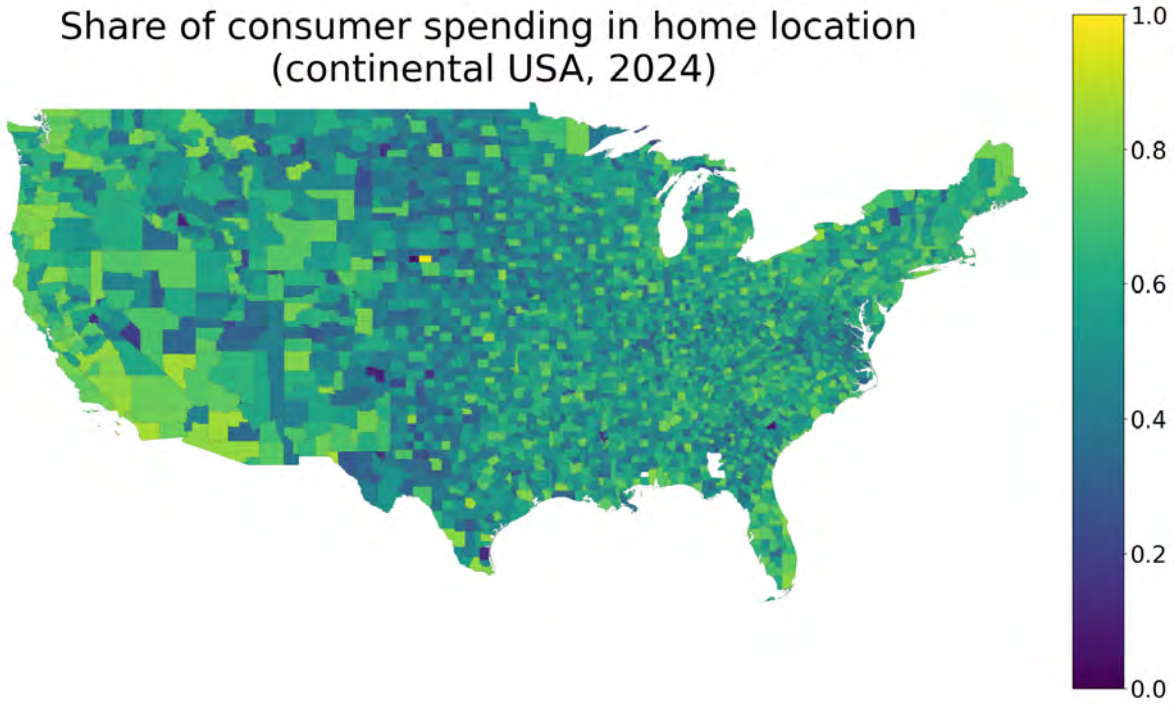
5. Descriptive Statistics

In this section we provide some basic descriptive statistics that highlight the importance of constructing ConZs and also the importance of measuring these zones by industry. Figure 1 shows the share of spending by consumers occurring in the home location relative to other areas in 2024. The more yellow the shade, the greater share occurring in the home location, while a darker shade of green and then blue indicates that more spending is leaving the local area. The wide variation in share of spending outside the local area motivates the need for ConZs because county political boundaries do not consistently contain consumption across areas, so alternative borders are needed to generate a more consistent consumption geography. A similar pattern is observed for estimates in 2019 and reported in the appendix in Figure A2.

Consumption patterns, however, vary substantially across industries. To illustrate this, we construct for each industry the cumulative distribution of spending by distance from the home county,

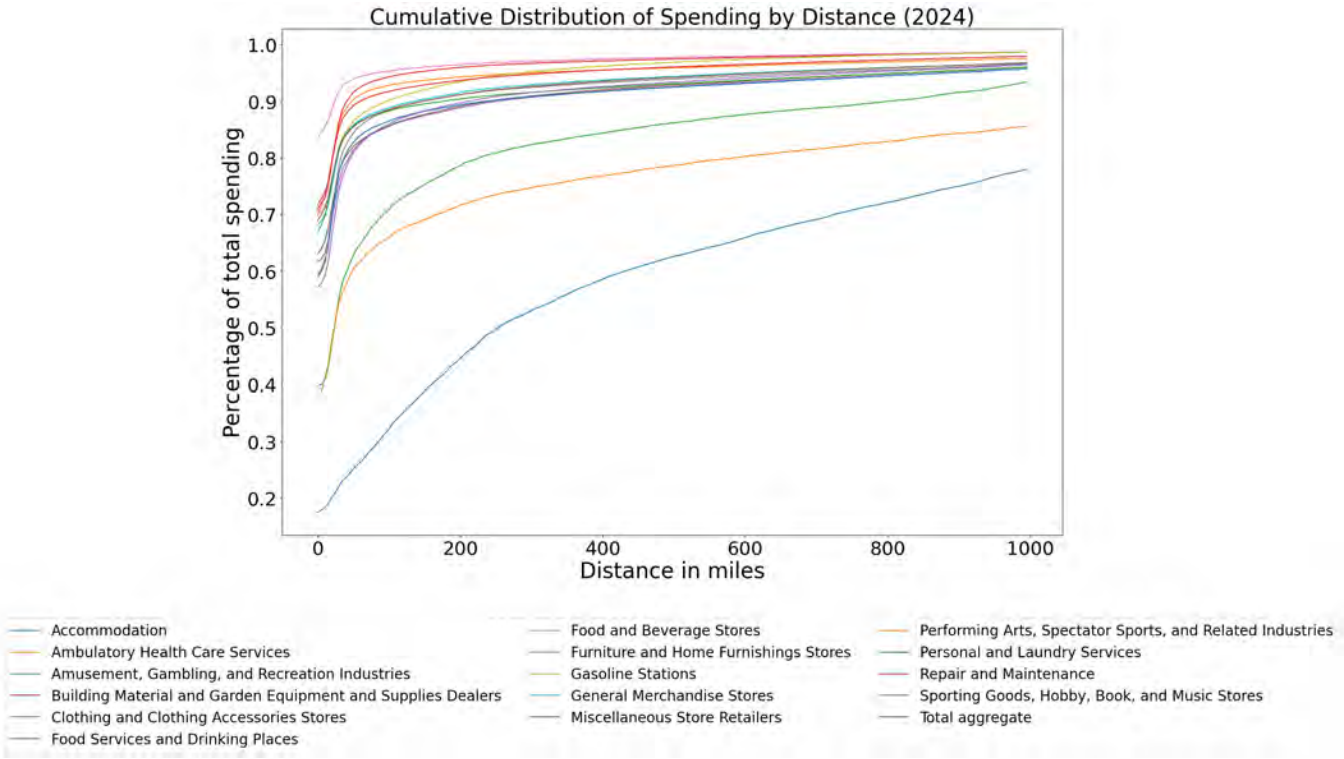
[10] All stages of data extraction, aggregation, and processing for this project are conducted entirely within Mastercard's secure internal infrastructure. At no point do individual transaction records, cardholder-level information, or the intermediate flow matrices leave Mastercard's systems; only the final output of the clustering algorithm—the consumption zone assignments—is shared externally with the Bureau of Economic Analysis.

Figure 1. Share of Aggregate Spending Occurring in Home County in 2024



Note. This figure displays the share of spending occurring in the home county in 2024. The denominator is total consumer spending in that county, and the numerator is total dollar spending by consumers in the county. The figure is constructed from aggregated and anonymized Mastercard transaction data.

Figure 2. Cumulative Distribution of Spending by Industry in 2024



Note. This figure displays the cumulative distribution of spending by industry. All industries show a greater share of spending captured as one moves farther away from the home county. The figure is constructed from aggregated and anonymized Mastercard transaction data.

shown in Figure 2. The contrast across industries is stark. For food and beverage stores, about 83% of spending occurs in the home county itself (distance zero), reflecting a highly local market. Accommodations lie at the other extreme: only about 20% of spending occurs in the home county, rising to roughly 60% within 500 miles, as consumers travel much farther for these purchases. In addition, for accommodations there is a large share of spending coming from consumers over 500 miles away. A very similar pattern emerges from the 2019 data, reported in Appendix Figure A3.

These large differences in spending patterns across industries provide clear motivation for constructing ConZs separately by industry. They also demonstrate why we adopt two containment cutoffs, one for our main local ConZs and a lower one for trade ConZs. Three industries, including performing arts, accommodations, and amusement, are distinct from the rest, drawing a much larger share of their spending from consumers far from the home county. For these, the 80% containment threshold used elsewhere would group most of the country into a handful of zones, so we treat them separately by applying a 50% threshold.

6. Results

In this section we present the updated consumption zones and compare them to those in Batch et al. (2025). We proceed in steps designed to separate the role of the algorithm from that of the data. We first hold the data fixed, using the same 2015 Fiserv flows as Batch et al. (2025), and apply our updated algorithm with the 80% containment stopping rule, so that any differences relative to prior work are attributable to the method alone. We then turn to the updated Mastercard data for 2019 and 2024, which form the basis for the remainder of the paper, and assess how the zones differ across industries and how stable they are over time. Throughout, the baseline zones are defined by 80% containment; for a small set of nonlocal industries we also report zones constructed at a 50% threshold.

6.1. Analysis and Basic Comparison with Prior Clusters Using Data from Batch et al. (2025)

The results from the updated clustering algorithm are shown in Table 1. The first column repeats the approach taken in Batch et al. (2025), and the second applies our updated algorithm to the same 2015 Fiserv data, with the stopping rule set to 80% containment. Because the data are held fixed, the differences reflect the algorithm alone. The aggregate number of zones is shown in the bottom row where we obtain 819 zones, compared to 1,235 in Batch et al. (2025). The new zones are fewer but larger, because the containment threshold needed to reach the number of zones in Batch et al. (2025) is 78%.[11] Accordingly, the average aggregate zone covers 4,300 square miles with a population of 392,377, up from 2,851 square miles and 260,208 in Batch et al. (2025). These values are closer to those of commuting zones, which Batch et al. (2025) report as 810 zones averaging 4,348 square miles and with a population of 396,737.

A key advantage of the containment-based zones is that they share an explicit economic meaning, on average, 80% of industry spending is contained within the resulting zones. The earlier approach instead applied a common clustering height across industries. This consistency also produces much wider variation in the number of zones across industries, which we view as intuitive and as informative about the unique geography of each industry. For industries with similar local consumption structures, these approaches can yield broadly comparable economic geographies, while the containment criterion becomes more consequential as spatial consumption patterns differ across industries.

Local industries require many small zones to reach 80% containment. The clearest case is food

[11] Table A3 shows the resulting zones if 78% containment is applied rather than 80%. Note that when we set containment to 78%, it matches the top line number of zones from Batch et al. (2025), but when the 78% containment is applied to the other industry, a vastly different number of zones is obtained.

Table 1. Comparing Old Zones and New Zones With 80% Containment Both Using 2015 Fiserv Data

Industry	Old Zones	New Zones	New Land Area Mean	New Population Mean	Share New Crossed by Old	Share Old Crossed by New
442 Furniture and Home Furnishings	570	386	9,123	832,531	0.671	0.102
444 Building Materials	1,194	2,398	1,469	134,010	0.013	0.619
445 Food and Beverage	1,862	3,128	1,126	102,736	0.000	0.499
447 Gas Stations	972	555	6,345	579,022	0.519	0.140
448 Clothing Stores	567	142	24,800	2,263,078	0.958	0.134
451 Sporting Goods Books	554	224	15,721	1,434,629	0.848	0.125
452 General Merchandise	1,198	2,388	1,475	134,572	0.009	0.596
453 Misc. Retail	900	143	24,626	2,247,252	0.944	0.079
621 Ambulatory Care	800	1,509	2,334	212,960	0.076	0.541
711 Performing Arts	322	10	352,156	32,135,700	0.600	0.059
713 Recreation	530	72	48,910	4,463,292	0.903	0.081
721 Accommodation	383	6	586,926	53,559,500	0.333	0.021
722 Food Services	997	586	6,009	548,391	0.529	0.141
811 Repair Maintenance	1,009	1,676	2,101	191,740	0.042	0.502
812 Personal Services	955	1,519	2,318	211,558	0.079	0.462
Aggregate	1,235	819	4,300	392,377	0.376	0.060

Note. This table compares the consumption zones of Batch et al. (2025) to zones constructed with our updated algorithm, holding the underlying data fixed at the 2015 Fiserv flows so that differences reflect the algorithm alone. The first column reports the original zone counts from Batch et al. (2025); the remaining columns apply the updated algorithm with a stopping rule set to 80% containment, meaning that on average 80% of spending in each industry occurs within the zone. Land area is reported in square miles, and population is the resident population, each averaged across zones within the industry. The last two columns compare the two geographies: "Share New Crossed by Old" is the share of new zones that contain more than one old zone, and "Share Old Crossed by New" is the share of old zones that contain more than one new zone. The bottom row reports zones for aggregate consumption, constructed from all industry flows combined.

and beverage stores (NAICS 445), where every county is its own zone (3,128 in total), suggesting that the relevant market is often smaller than a county and that a finer geography such as zip codes may be a better starting point. The number of zones also rises sharply for building materials (NAICS 444, from 1,194 to 2,398), general merchandise (NAICS 452, from 1,198 to 2,388), and ambulatory care (NAICS 621, from 800 to 1,509)— roughly doubling in each case—and increases substantially for repair and maintenance (NAICS 811) and personal services (NAICS 812).

Other industries move in the opposite direction. Where consumption involves tourism, or a mix of tourism and local spending, far fewer zones are needed: accommodations (NAICS 721) falls from 383 to just 6, and performing arts and spectator sports (NAICS 711) from 322 to 10. At this scale, the areas may be better described as regions than as zones. Dividing these markets more finely requires a lower containment threshold, which we pursue below as a set of "traded" or nonlocal geographic markets.

The remaining columns describe these contrasts in more detail. Columns 3 and 4 report average land area and population, which move with the number of zones. Columns 5 and 6 compare the two geographies directly: only 6% of old zones are crossed by a new zone (i.e., contain more than one new zone), so old zones typically sit within a single new zone, while the larger new zones are frequently crossed by old ones (about 38%).

This exercise shows that the algorithm plays an important role in defining ConZs. The underlying

data, however, are now relatively old, so for the remainder of the paper we focus on the updated Mastercard zones, which are more relevant for users analyzing recent data.

6.2. New Consumption Zones Based on Mastercard Data 2019 and 2024

We start by reporting some basic descriptive statistics on these new zones, first for 2019 and 2024, along with some simple comparisons of the prior zones used in Batch et al. (2025). The updated zones are reported in Table 2. There are a few points worth noting. First, the algorithm applied generally matters substantially more than the data or time period applied. While there are some notable differences between the zones obtained when using the 2015 Fiserv data and the Mastercard data, there are much larger differences in the zones that comes from applying the alternative algorithm. [12] Another notable feature of the zones, is that they are relatively persistent over time for most categories, despite external evidence of major shifts in consumption around the COVID pandemic (Chetty et al. (2024)).

The Table 2 reports, for each industry and year, the number of zones and the mean land area and population. The aggregate counts illustrate the first point. We obtain 776 zones in 2019 and 727 in 2024, both close to the 819 zones produced when the same algorithm is applied to the 2015 Fiserv data (Table 1), and all far from the 1,235 zones of Batch et al. (2025). Moving across data sources and years shifts the aggregate count only modestly, within a range of roughly 730 to 820, whereas changing the algorithm moved it by several hundred. The same ordering holds for most individual industries.

The second point is the persistence of the zones over time. For most industries the number of zones is similar in 2019 and 2024, despite external evidence of large shifts in consumption around the pandemic. Gas stations (726 to 691), general merchandise (729 to 708), and personal services (540 to 561) are nearly unchanged; food and beverage stores remain close to one zone per county in both years; and the nonlocal industries stay very small, with accommodations at 4 and then 5 zones and performing arts at 30 and then 20. The clearest exception is miscellaneous retailers, which falls from 318 zones to 21, with clothing and sporting goods also declining by roughly a third. We examine these changes more formally in the next section.

As in the previous table, mean land area and population move inversely with the number of zones. The local industries form many small zones, food and beverage stores average around 1,000 square miles and 100,000 people, while the nonlocal industries form a handful of very large areas. Accommodations is the extreme case, with only 4 to 5 zones that each span on the order of 700,000 square miles and tens of millions of residents. This supports the earlier point that for

[12] Unfortunately, due to data restrictions, we are unable to make a direct comparison between the Fiserv flows and Mastercard flows, to further evaluate the source of potential differences.

these industries the relevant geography is closer to a region than to a local market.

Table 2. Consumption Zones Using Mastercard Data 2019 and 2024 (80% containment)

	2019			2024		
Industry	Num. Zones	Land Area	Population	Num. Zones	Land Area	Population
442 Furniture and Home Furnishings	411	7,061	753,983	337	10,301	942,479
444 Building Materials	1,341	2,164	231,086	1,136	3,056	279,591
445 Food and Beverage	2,994	969	103,503	3,114	1,115	101,996
447 Gas Stations	726	3,997	426,842	691	5,024	459,646
448 Clothing Stores	221	13,131	1,402,203	142	24,447	2,236,728
451 Sporting Goods Books	235	12,349	1,318,668	160	21,696	1,985,096
452 General Merchandise	729	3,981	425,085	708	4,903	448,609
453 Misc. Retail	318	9,126	974,487	21	165,305	15,124,542
621 Ambulatory Care	1,467	1,978	211,239	1,615	2,149	196,666
711 Performing Arts	30	96,731	10,329,565	20	173,570	15,880,769
713 Recreation	29	100,067	10,685,757	32	108,481	9,925,480
721 Accommodation	4	725,486	77,471,736	5	694,281	63,523,072
722 Food Services	291	9,972	1,064,904	214	16,222	1,484,184
811 Repair Maintenance	1,254	2,314	247,119	1,322	2,626	240,254
812 Personal Services	540	5,374	573,865	561	6,188	566,159
Aggregate	776	3,740	399,339	727	4,775	436,885

Note. This table reports consumption zones constructed from Mastercard county-to-county spending flows (aggregated and anonymized) for 2019 and 2024, using our updated algorithm with a stopping rule set to 80% containment (on average, 80% of an industry's spending occurs within the zone). For each year, the columns report the number of zones and the mean land area (in square miles) and mean resident population across zones within the industry. A larger number of zones corresponds to smaller average land area and population. The bottom row reports zones for aggregate consumption, constructed from all industry flows combined.

6.3. The stability of zones over time

When introducing the Mastercard zones we noted that the number of zones is broadly similar in 2019 and 2024 for most industries. Equal counts, however, don't imply that the same counties are grouped together; a stable count could mask substantial reshuffling. We therefore assess stability directly by treating each year's assignment as a partition of the counties and comparing the two partitions industry by industry. Because zone labels are assigned separately in each year and carry no cross-year meaning, every measure we use depends only on which counties are grouped together, not on the labels used to identify the zones.

Our primary measure is the share of counties that keep their grouping. For each 2019 zone we identify the 2024 zone into which the largest number of its counties move, and count those counties as retained:

$$R = \frac{1}{N} \sum_{k} \max_{\ell} n_{k\ell} \in [0, 1], \tag{4}$$

where $n_{k\ell}$ is the number of counties shared by 2019 zone k and 2024 zone ℓ and N is the total

number of counties; $R = 1$ when every zone maps cleanly onto a single successor. We also report a population-weighted version R_w, which for each 2019 zone selects the 2024 zone capturing the largest share of its population. Comparing R_w with R indicates whether the counties that change grouping are predominantly small ones, in which case $R_w > R$.

As a more conservative complement, we report the Jaccard index of co-membership. Among all pairs of counties, let a be the number assigned to the same zone in both years, and let b and c be the number grouped together in only 2019 or only 2024. Then

$$J = \frac{a}{a+b+c} \in [0,1] \tag{5}$$

is the share of ever-linked county pairs that remain linked. Because it counts pairs rather than counties, J responds sharply to the splitting or merging of large zones, and so serves as a stricter lower benchmark against which to read R. We likewise report a population-weighted Jaccard J_w and, to ease comparison across industries with differing numbers of zones, the chance-corrected adjusted Rand index (ARI). Full definitions are given in Appendix D.

The results, reported in Table 3, confirm that the zones are largely stable over time. For the aggregate, 88% of counties keep their grouping ($R = 0.875$), rising to 90% on a population-weighted basis, and the pairwise and chance-corrected measures tell the same story ($J = 0.544$, ARI $= 0.704$). Stability is high across most individual industries, and R_w generally exceeds R, indicating that the counties that do change grouping tend to be smaller, lower-population ones. The clearest exception is miscellaneous retailers, where the collapse from 318 zones to 21 leaves the plurality measure high ($R = 0.917$) but drives the pairwise measures sharply down ($J = 0.031$, ARI $= 0.054$), which is exactly the pattern the Jaccard index is designed to identify, since merging many small zones into a few large ones reshuffles a large number of county pairs while leaving most individual counties with their former neighbors. Overall, the consumption zones appear to be a stable construct over this period, despite the well-documented shifts in consumption around the pandemic.

6.4. Mapping New Clusters

Figures 3 and 4 display the updated zones on the map, using the Mastercard 2024 data at 80% containment. The aggregate clusters in Figure 3 show small zones concentrated in densely populated areas, where a single county or a few counties contain most spending, and larger zones spanning the more sparsely populated interior are combined. The result is a partition of the country into local consumption markets that respects population density without imposing political boundaries.

Figure 4 makes clear that there is no single consumption geography. The same containment rule

Table 3. Similarity Measures Across Years

Industry	Zones 2019	Zones 2024	Same-grp	Same-grp (pop-wt.)	Jaccard	Jaccard (pop-wt.)	ARI
442 Furniture and Home Furnishings	411	327	0.812	0.918	0.432	0.511	0.602
444 Building Materials	1,341	1,101	0.895	0.934	0.501	0.381	0.668
445 Food and Beverage	2,993	2,993	1.000	1.000	1.000	1.000	1.000
447 Gas Stations	726	666	0.807	0.879	0.463	0.449	0.633
448 Clothing Stores	221	139	0.852	0.906	0.421	0.556	0.589
451 Sporting Goods Books	235	157	0.865	0.868	0.448	0.418	0.616
452 General Merchandise	729	687	0.785	0.874	0.417	0.536	0.588
453 Misc. Retail	318	21	0.917	0.957	0.031	0.048	0.054
621 Ambulatory Care	1,467	1,565	0.838	0.928	0.515	0.616	0.680
711 Performing Arts	30	20	0.729	0.813	0.347	0.581	0.390
713 Recreation	29	32	0.701	0.862	0.456	0.792	0.604
721 Accommodation	4	5	0.855	0.860	0.606	0.610	0.503
722 Food Services	291	208	0.823	0.910	0.412	0.698	0.582
811 Repair Maintenance	1,254	1,284	0.809	0.895	0.437	0.290	0.608
812 Personal Services	540	548	0.756	0.841	0.428	0.382	0.598
Aggregate	776	700	0.875	0.904	0.544	0.533	0.704

Note. This table measures how much the consumption zones change between 2019 and 2024, treating each year's assignment as a partition of counties and comparing the two partitions industry by industry. All measures depend only on which counties are grouped together, not on the zone labels. "Zones 2019" and "Zones 2024" report the number of zones in each year, but are limited to those counties where we are able to match with population measures for those counties. "Same-grp" is the share of counties that remain with the plurality of their 2019 zone (R in Equation 4), and "Same-grp (pop-wt.)" is its population-weighted analogue (R_w). "Jaccard" is the share of ever-co-grouped county pairs that remain co-grouped in both years (J in Equation 5), with "Jaccard (pop-wt.)" its population-weighted version (J_w); because they count pairs rather than counties, the Jaccard measures fall sharply when large zones are split or merged. "ARI" is the chance-corrected adjusted Rand index, which equals 1 for identical partitions and approximately 0 for unrelated ones. Higher values indicate greater stability for all measures. The bottom row reports the aggregate consumption zones. Full definitions are given in Appendix D. This table is constructed using aggregated and anonymized Mastercard transaction data.

produces strikingly different maps across industries. Panels (a)–(d) show the local industries—building materials, food and beverage stores, food services, and ambulatory care, which fragment the country into many small zones, most densely in the eastern half of the country, which is more densely populated. Food and beverage stores (panel b) are the extreme case, with nearly every county forming its own zone. Panels (e) and (f) show the opposite pattern for two nonlocal industries. Accommodations (panel e) becomes just a handful of very large regions, and performing arts (panel f) similarly forms a small number of expansive zones. These maps put the variation documented in the preceding tables in visual form: the geography of consumption is highly local for some industries and effectively regional for others, and a single fixed threshold of 80% containment is enough to reveal that contrast.

Figures 5 and 6 provide two pictures of containment within ConZs for aggregate spending. Figure 5 shows containment at the county level, which is the share of spending at the county level that

stays within its corresponding ConZ,[13] and Figure 6 shows average containment at the ConZ level, so all counties within a cluster receive the same shading.[14] The figures correspond to setting average containment (according to Eq. (3)) of 80%.

Note that the local share of spending shown in Figure 1 in the descriptive statistics section is equivalent to a measure of county containment when the county itself forms a ConZ. This implies we can directly compare Figures 1 and 5. Comparing Figure 1 to Figure 5 , we observe that the color distribution shifts notably more yellow, indicating a much greater level of containment in the defined ConZs.

In Figure 6, counties are colored according to the containment of the ConZ to which they belong. Note that, despite fixing an average containment of 80% (according to Eq. (3)), individual ConZs may deviate from this level as indicated by some light shades of green in the figure. In particular, many of the darker counties are singletons (i.e., they form their own ConZ).

6.5. Nonlocal Consumption Zones

For the nonlocal industries, the 80% containment rule produces only a handful of very large zones, which are better understood as regions than as local markets. To recover meaningful within-industry geography for these cases, we lower the containment threshold to 50%, defining what we term traded consumption zones. Table 4 reports the result for the three most nonlocal industries. The lower threshold yields far more zones: performing arts (NAICS 711) rises from 30 zones at 80% containment to 1,919 in 2019, and recreation (NAICS 713) from 29 to 2,264, while accommodations (NAICS 721) remains comparatively coarse at 47 zones, reflecting how widely accommodation spending is dispersed relative to the consumer's home county. As with the baseline zones, mean land area and population move inversely with the number of zones. The 2024 counts are broadly similar for recreation but fall markedly for performing arts and accommodations, a pattern we would read alongside the stability measures rather than as evidence of a sharp change in underlying geography.

[13] The containment shown in Figure 5 is defined, for a county k, as

$$\Gamma_k \equiv \frac{\sum_{j \in Z_k} f_{kj}}{\sum_{j} f_{kj}},$$

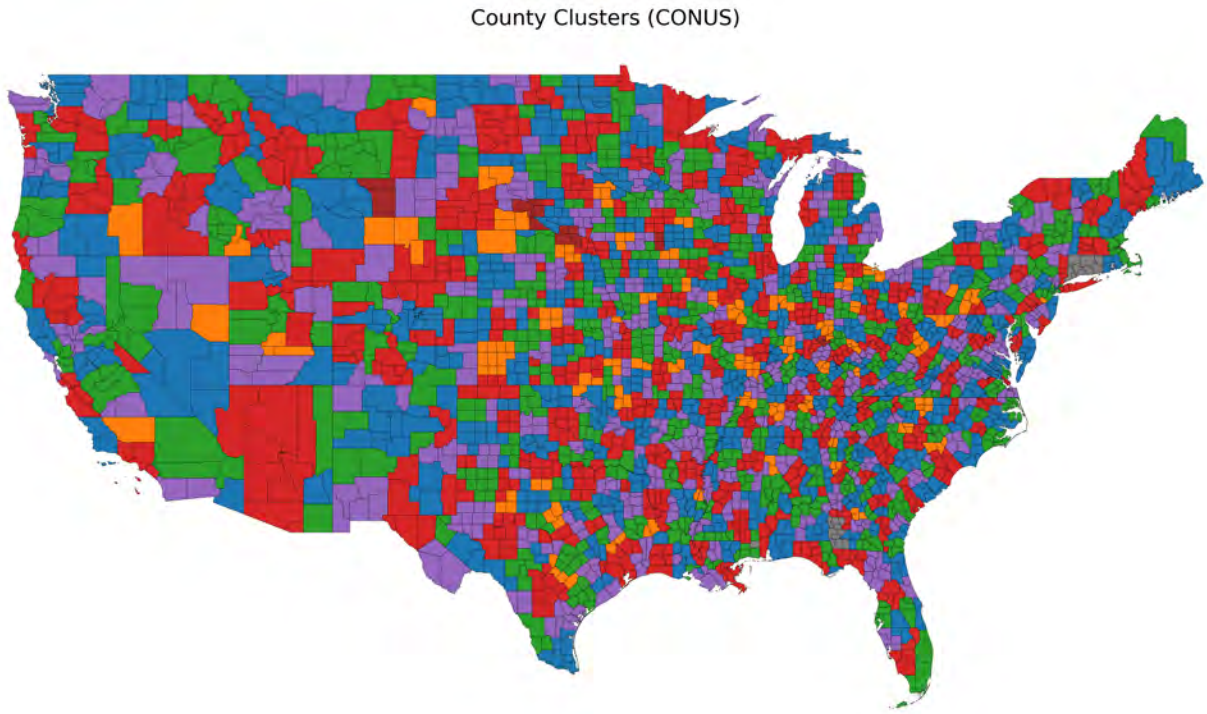
where Z_k denotes the ConZ to which county k belongs. Figure 5 is directly comparable to Figure 1.

[14] The containment of a given ConZ Z is defined as

$$\Gamma_Z \equiv \frac{f_{ZZ}}{\sum_{K \in \Omega} f_{ZK}},$$

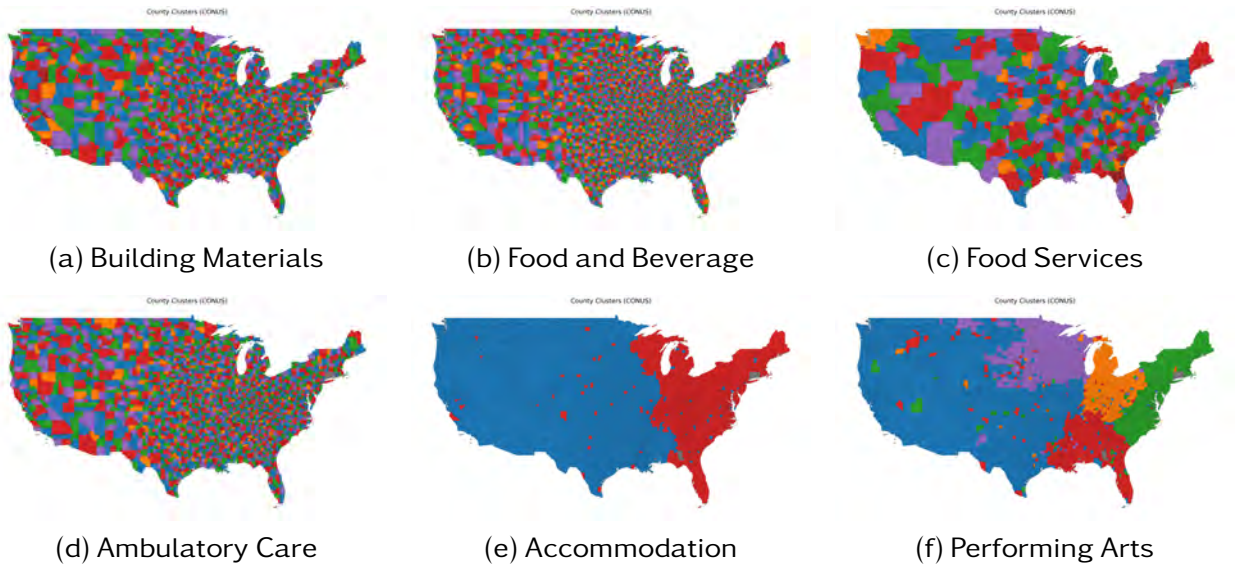
where Ω denotes the set of all ConZs.

Figure 3. Aggregate County Clusters Using Mastercard 2024 and 80% Containment



Note. This figure displays aggregate consumption zones for the contiguous United States, constructed from Mastercard county-to-county spending flows (aggregated and anonymized) for 2024 using a stopping rule set to 80% containment. Each color represents a distinct zone that may contain one or more counties; the same color may be reused for non-contiguous zones. Aggregate zones are constructed from all industry flows combined.

Figure 4. County Clusters Mastercard 2024 Data Under Alternative Industry Definitions

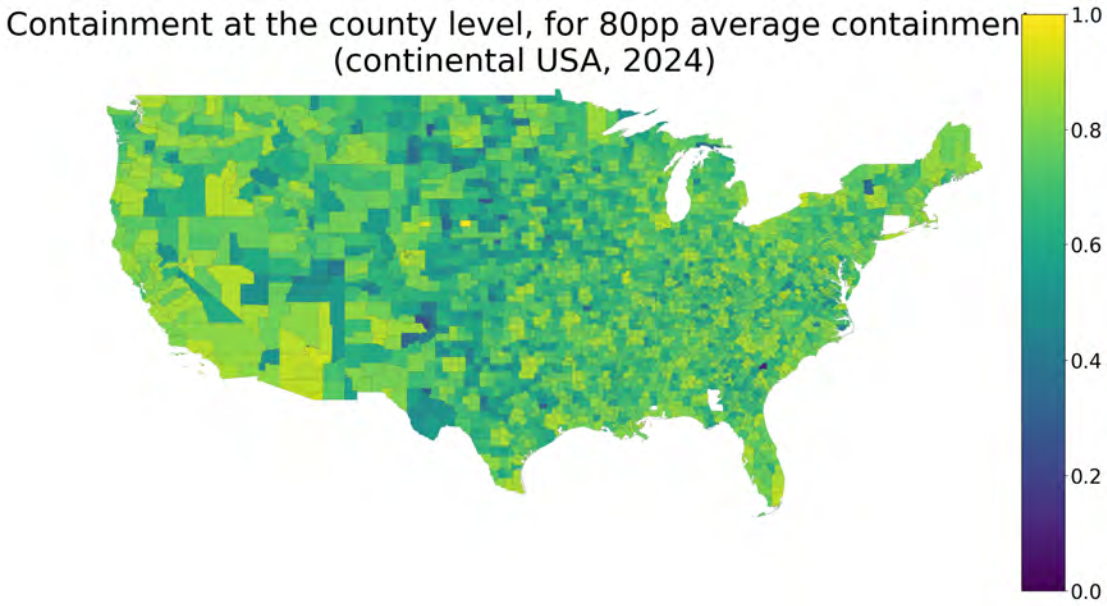


(a) Building Materials (b) Food and Beverage (c) Food Services

(d) Ambulatory Care (e) Accommodation (f) Performing Arts

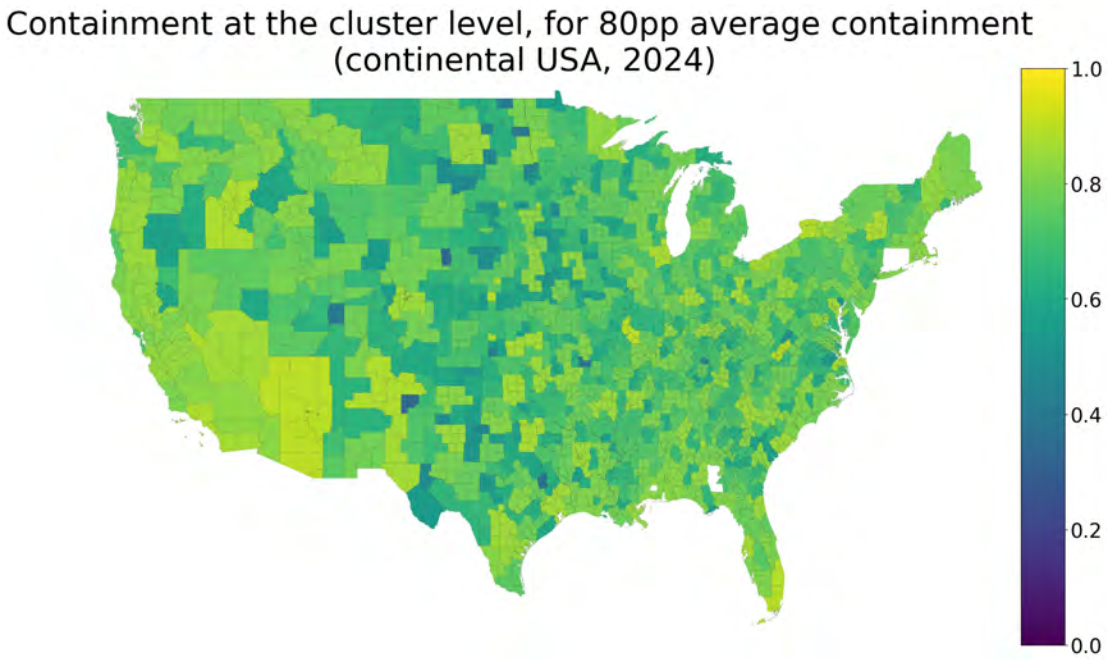
Note. Each panel displays consumption zones for the contiguous United States under a different industry definition, constructed from Mastercard county-to-county spending flows (aggregated and anonymized) for 2024 using a stopping rule set to 80% containment: (a) building materials (NAICS 444), (b) food and beverage stores (NAICS 445), (c) food services (NAICS 722), (d) ambulatory care (NAICS 621), (e) accommodation (NAICS 721), and (f) performing arts (NAICS 711). Each color represents a distinct zone that may contain one or more counties; the same color may be reused for non-contiguous zones. Panels (e) and (f) cover nonlocal industries and therefore resolve into a small number of very large zones at the 80% threshold.

Figure 5



Note. This figure is constructed using Mastercard aggregated and anonymized transaction data.

Figure 6



Note. This figure is constructed using Mastercard aggregated and anonymized transaction data.

Table 4. Nonlocal Consumption Zones Using Mastercard Data 2019 and 2024 Using 50% Containment

	2019			2024		
Industry	Num. Zones	Land Area	Population	Num. Zones	Land Area	Population
711 Performing Arts	1,919	1,512	161,484	493	7,041	644,250
713 Recreation	2,264	1,282	136,876	2,300	1,509	138,094
721 Accommodation	47	61,743	6,593,340	32	108,481	9,925,481

Note. This table reports consumption zones for the three most nonlocal industries, constructed from Mastercard county-to-county spending flows (aggregated and anonymized) for 2019 and 2024 using a stopping rule set to 50% containment (on average, 50% of an industry's spending occurs within the zone). We term these traded consumption zones. The lower threshold is used because the 80% rule applied elsewhere yields only a handful of very large zones for these industries; see Table 3 for the corresponding counts at 80% containment. For each year, the columns report the number of zones and the mean land area (in square miles) and mean resident population across zones within the industry. A larger number of zones corresponds to smaller average land area and population.

Figure 7 maps the traded consumption zones for the three nonlocal industries, using the Mastercard 2024 data at 50% containment. At this lower threshold, the broad regions of the 80% maps break apart into much finer geographies. Accommodations (panel a) and performing arts (panel b) resolve into many distinct zones rather than the handful produced at 80% containment, and amusement and recreation (panel c) similarly fragments across the country. The maps confirm that the coarseness of these industries at the baseline threshold reflects how widely their spending is dispersed, not an absence of underlying structure: once the containment requirement is relaxed, meaningful within-industry zones emerge.

Figure 7. Nonlocal Consumption Clusters Mastercard 2024 Data Using 50% Containment



(a) Accommodations

(b) Performing Arts, Spectator Sports, and Related Subsectors

(c) Amusement, Gambling, and Recreation Industries

Note. Each panel displays traded consumption zones for the contiguous United States, constructed from Mastercard county-to-county spending flows (aggregated and anonymized) for 2024 using a stopping rule set to 50% containment: (a) accommodations (NAICS 721), (b) performing arts, spectator sports, and related industries (NAICS 711), and (c) amusement, gambling, and recreation industries (NAICS 713). Each color represents a distinct zone that may contain one or more counties; the same color may be reused for non-contiguous zones. The 50% threshold is used in place of the 80% baseline applied elsewhere because these nonlocal industries yield only a handful of very large zones at 80% containment.

7. Reexamining Industry Concentration Measures Using Updated Zones

In Batch et al. (2025), ConZs were used to examine how these alternative boundaries affect measured concentration relative to county (FIPS) and state geographies. Because the zones have changed considerably under the updated clustering algorithm and Mastercard data, we re-examine that analysis here. Using the identical NETS data described in Batch et al. (2025) for 2019, we compute the Herfindahl–Hirschman index (HHI) by industry for each geography; the HHI ranges from 0 (no concentration) to 10,000 (a monopoly). Following the 2010 DoJ Horizontal Merger Guidelines, we flag markets into three zones: a "Red Zone" (HHI above 2,500), presumed highly concentrated; a "Yellow Zone" (1,500–2,500); and a "Green Zone" (below 1,500), presumed unconcentrated. Because only the ConZ definitions depend on the spending data, the county and state HHIs are unchanged from Batch et al. (2025), so any differences in the ConZ column reflect the new zones alone.

The results, reported in Table 5, are broadly similar to those in Batch et al. (2025) despite the change in algorithm and the different number of zones. As before, ConZ concentration is lower than at the county level, and the same industries stand out: gasoline stations and general merchandise stores fall in the Red Zone, while building materials and food and beverage stores fall in the Yellow Zone. The remaining industries stay in the Green Zone under both versions. The few notable differences appear among the tourist- and infrequent-purchase industries, such as performing arts, amusement, and accommodations, whose ConZ HHIs drop substantially, reflecting the broader markets these zones now span; these industries remain in the Green Zone in either case. Overall, the updated zones produce concentration metrics qualitatively the same as Batch et al. (2025).

Table 5. Concentration Metrics for 2019 Using Updated Mastercard Data with 80% Containment

	FIPS	ConZ	State
Furniture and Home Furnishings (NAICS 442)	1,003	448	226
Building Material and Garden Equip. (NAICS 444)	2,471	2,070	1,463
Food and Beverage Stores (NAICS 445)	2,095	2,047	1,129
Gasoline Stations (NAICS 447)	3,783	3,076	2,059
Clothing and Clothing Accessories (NAICS 448)	668	333	243
Sporting Goods, Hobby, Books, and Music (NAICS 451)	1,072	493	360
General Merchandise Stores (NAICS 452)	3,372	2,958	2,637
Miscellaneous Store Retailers (NAICS 453)	970	504	294
Ambulatory (NAICS 621)	629	463	138
Performing Arts, Spec Sports, and Related Ind. (NAICS 711)	1,683	187	559
Amusement, Gambling, and Recreation (NAICS 713)	1,563	202	493
Accommodations (NAICS 721)	1,333	98	412
Food Service and Drinking Places (NAICS 722)	198	75	42
Repair and Maintenance (NAICS 811)	571	411	88
Personal and Laundry Services (NAICS 812)	582	331	127

Note. This table reports sales-weighted Herfindahl–Hirschman indexes (HHIs) for three geographies—county (FIPS), consumption zone (ConZ), and state—for 2019. The HHI ranges from 0 (no concentration) to 10,000 (a monopoly). The ConZ columns use the updated zones constructed from Mastercard county-to-county spending flows (aggregated and anonymized) with a stopping rule set to 80% containment; the establishment-level sales used to compute the HHIs are from the NETS data described in Batch et al. (2025), so the county and state values do not depend on the zone definition. Following the 2010 Department of Justice Horizontal Merger Guidelines, we classify markets into three zones: a "Red Zone" for an HHI above 2,500, marked red; a "Yellow Zone" for an HHI between 1,500 and 2,500, marked orange; and a "Green Zone" for an HHI below 1,500, which is unmarked.

8. Conclusion

This paper revisits the construction of consumption zones using updated transaction data and a revised implementation of the original hierarchical clustering algorithm. Rather than relying on a normalized stopping rule with limited economic interpretation, we propose selecting the clustering threshold based on containment, which is the share of spending that remains clusters in the aggregate. This approach provides a transparent and economically meaningful criterion that can be applied consistently across industries while naturally allowing the size and number of markets to differ according to observed consumer behavior. Using updated Mastercard transaction data from 2019 and 2024, we show that the resulting consumption zones differ substantially across industries, are generally stable over time despite the disruptions associated with the COVID–19 pandemic, and continue to produce intuitive measures of local market concentration.

More broadly, our results suggest that there is no single geographic market appropriate for

studying consumption. Local industries, such as grocery stores and ambulatory care, require highly localized markets, while industries such as accommodations and performing arts are better characterized by much larger geographic regions. We therefore provide both baseline consumption zones based on 80% containment and alternative traded consumption zones for industries where travel is an important feature of demand. We hope these updated geographies will provide a useful resource for researchers and statistical agencies studying local economic activity, competition, regional policy, and the geographic effects of economic shocks, while also serving as a foundation for future work using increasingly granular transaction data to measure consumer markets.

References

Batch, Andrea, Benjamin Bridgman, Abe Dunn, and Mahsa Gholizadeh. 2025. "Consumption zones." *Journal of Economic Geography*, 25(2): 191–213.

Bess, Rebecca, Zoë O Ambargis, et al. 2011. "Input-output models for impact analysis: Suggestions for practitioners using RIMS II multipliers." 23–27, Southern Regional Science Association Morgantown, WV.

Capps, Cory, David Dranove, Shane Greenstein, and Mark A Satterthwaite. 2001. "The silent majority fallacy of the Elzinga-Hogarty criteria: a critique and new approach to analyzing hospital mergers."

Capps, Cory, Laura Kmitch, Zenon Zabinski, and Slava Zayats. 2019. "The continuing saga of hospital merger enforcement." *Antitrust Law Journal*, 82(2): 441–496.

Chetty, Raj, John N Friedman, Michael Stepner, and Opportunity Insights Team. 2024. "The Economic Impacts of COVID-19: Evidence from a New Public Database Built Using Private Sector Data." *Quarterly Journal of Economics*, 139(2): 829–889.

Chi, Feng. 2025. "Information waves and firm investment." *Available at SSRN 4269750*.

Chodorow-Reich, Gabriel. 2019. "Geographic cross-sectional fiscal spending multipliers: What have we learned?" *American Economic Journal: Economic Policy*, 11(2): 1–34.

Dunn, Abe, and Mahsa Gholizadeh. 2025. "The geography of consumption and local economic shocks." *American Economic Journal: Macroeconomics*.

Federal Reserve Financial Services. 2026. "2026 Findings from the Diary of Consumer Payment Choice." Federal Reserve Bank of Atlanta, Atlanta, GA.

Foote, Andrew, Mark J. Kutzbach, and Lars Vilhuber. 2021. "Recalculating–How uncertainty in local labor market definitions affects empirical findings." *Applied Economics*, 53(14): 1598–1612.

for the Evaluative Clinical Sciences, Dartmouth Medical School. Center. 1996. *The Dartmouth atlas of health care*. AHA, published.

Fowler, Christopher S. 2024. "New Commuting Zone delineation for the US based on 2020 data." *Scientific Data*, 11(1): 975.

Frech, HE, James Langenfeld, and R Forrest McCluer. 2004. "Elzinga-Hogarty tests and alternative approaches for market share calculations in hospital markets." *Antitrust law journal*, 71(3): 921–947.

HSN Consultants, Inc. 2025. "General Purpose Cards–U.S. 2024." *The Nilson Report*.

HSN Consultants, Inc. 2026. "Purchase Transactions on Global Cards in 2025." *The Nilson Report*.

PULSE Network. 2025. "2025 PULSE Debit Issuer Study." Discover Financial Services, Houston, TX.

Tolbert, Charles M., and Molly Sizer. 1987. "Labor Market Areas for the United States." United States Department of Agriculture, Economic Research Service Staff Report AGES870721.

A. Data

A.1. Detailed description of Mastercard transaction data

A.1.1. Geographic and industry dimensions

The data are aggregated at the county level, identified by Federal Information Processing Standards (FIPS) codes. Each observation in the dataset represents the total dollar amount of spending at merchants located in county i by cardholders residing in county j, for a given industry and year. This yields a flow matrix $f_{ij}^{(k)}$ for each industry k, where the rows index the merchant (destination) county and the columns index the consumer (origin) county. The flow matrices are aggregated annually and cover the years 2019 and 2024. The merchant county is determined using Mastercard's Member Merchant Hierarchy (MMH), which assigns geographic coordinates and zip code to merchant locations. The MMH is enriched with third-party data sources (e.g., chain store directories, geocoding services) to validate and correct merchant location information, including the Merchant Category Code (MCC) assigned to each merchant. The consumer zip code is inferred using internal models at Mastercard based on spending patterns, and then it is mapped from zip code to county FIPS using standard crosswalk files. The industry dimension is defined using the North American Industry Classification System (NAICS). Because every transaction processed by the Mastercard network is assigned a four-digit Merchant Category Code (MCC) that denotes the type of business providing the good or service, we aggregate MCCs to NAICS codes using Mastercard's internal mapping tables. This mapping allows us to construct spending flows for industries that are directly comparable to those used in official economic statistics and in the original consumption zones framework of Batch et al. (2025).[15] Specifically, we construct flow matrices for 15 retail and service industries that require a local consumer presence, plus an aggregate of all industries combined. These industries are:

1. Food and Beverage Stores (NAICS 445)
2. Food Services and Drinking Places (NAICS 722)
3. Gasoline Stations (NAICS 447)
4. General Merchandise Stores (NAICS 452)
5. Clothing and Clothing Accessories Stores (NAICS 448)
6. Sporting Goods, Hobby, Book, and Music Stores (NAICS 451)

[15] Note, however, that unlike Batch et al. (2025), and for simplicity, we assume a reasonable degree of representativeness in the Mastercard data and do not rescale our flows to match total sales estimates constructed by combining Economic Census and QCEW data.

7. Building Material and Garden Equipment and Supplies Dealers (NAICS 444)
8. Furniture and Home Furnishings Stores (NAICS 442)
9. Accommodation (NAICS 721)
10. Amusement, Gambling, and Recreation Industries (NAICS 713)
11. Performing Arts, Spectator Sports, and Related Industries (NAICS 711)
12. Personal and Laundry Services (NAICS 812)
13. Repair and Maintenance (NAICS 811)
14. Ambulatory Health Care Services (NAICS 621)
15. Miscellaneous Store Retailers (NAICS 453)

These 15 industries correspond to the same sectors analyzed in Batch et al. (2025), which facilitates direct comparisons between the consumption zones derived in this paper and those estimated in the original study.

A.1.2. Data construction and preprocessing

The raw transaction-level data are processed entirely within Mastercard's secure analytics environment, and the data shared with the Bureau of Economic Analysis consist exclusively of the final output of the clustering algorithm—namely, the consumption zone assignments of counties to clusters. At no point are individual transaction records, cardholder-level information, or the intermediate county-to-county flow matrices shared externally. Several preprocessing steps are applied during the construction of the flow matrices:

- MCC-to-NAICS aggregation. Individual transactions, each tagged with a four-digit MCC, are mapped to their corresponding NAICS industry code. For aggregated merchants (i.e., identified chain retailers), the MCC assignment is validated using third-party data, resulting in higher mapping accuracy. For non-aggregated merchants, the MCC recorded at the time of transaction clearing is used directly.
- County assignment. Merchant locations are geocoded and assigned to FIPS county codes using Mastercard's MMH and supplementary geocoding data. Consumer zip codes are inferred using internal models at Mastercard based on cardholders' spending patterns, and then mapped to FIPS county codes using standard crosswalk files. A county-FIPS reconciliation is performed using the 2022 Census Bureau Gazetteer file to ensure consistency with current county definitions.

- Privacy filtering. The data are subject to Mastercard's privacy and data protection requirements. Appropriate filtering rules have been applied throughout the data construction process to ensure that the resulting consumption zones do not reveal the identity of any individual merchant or cardholder.

B. Alternative dissimilarity metrics

The original dissimilarity metric proposed by Tolbert and Sizer (1987), which we denote in this section as $D^{(1)}$, has some weaknesses that we highlight below. In the process of fixing/improving these weaknesses, we construct a sequence of different distance metrics: $D^{(2)}$, $D^{(3)}$, $D^{(4)}$, $D^{(5)}$, and $D^{(6)}$. Each of these metrics is constructed in a way that it fixes a new identified issue and still behaves well in all the previously considered cases.

The different metrics that we consider and the issues that each of them solves are detailed below (they are applicable to both the CZ or ConZ contexts). We use the following notation:

$$D_{ij} \equiv 1 - P_{ij}\,, \tag{6}$$

where P_{ij} represents the interaction strength between the two counties. And we consider different types of fluxes:

$$f_{ij} \quad \text{(flow from county } i \text{ to county } j)\,, \tag{7}$$

$$I_i \equiv \sum_{j \neq i} f_{ji} \quad \text{(total inflow to county } i)\,, \tag{8}$$

$$O_i \equiv \sum_{j \neq i} f_{ij} \quad \text{(total outflow from county } i)\,, \tag{9}$$

$$L_i \equiv f_{ii} \quad \text{(local flow inside county } i)\,. \tag{10}$$

B.1. $L_j + O_j$ much larger than $L_i + O_i$

Example:

$L_i + O_i = 100$, $L_j + O_j = 10{,}000$, $L_k + O_k = 120$

$f_{ij} = 5$, $f_{ji} = 4{,}000 \Rightarrow P_{ij} = 0.05$, $P_{ji} = 0.40$

$f_{ik} = 30$, $f_{ki} = 30 \Rightarrow P_{ik} = 0.30$, $P_{ki} = 0.25$

The county i sends 30% of its outflow to k but only 5% to j. **The natural cluster is** (i,k).

We propose the following modified metric:

$$P_{ij}^{(2)} \equiv \frac{1}{2}\left(\frac{f_{ij}}{L_i + O_i} + \frac{f_{ji}}{L_j + O_j}\right) \tag{11}$$

pair	$D^{(1)}$	$D^{(2)}$
(i,j)	0.603	0.775
(i,k)	0.727	0.725
clusters i with	j	k ✓

B.2. f_{ji} large compared to $L_i + O_i$

Example:

$L_i + O_i = 1,000$, $L_j + O_j = 100,000$, $L_k + O_k = 1,200$

$L_i + I_i = 900$, $L_j + I_j = 120,000$, $L_k + I_k = 1,100$

$f_{ij} = 20$, $f_{ji} = 600$, $\Rightarrow f_{ij}/(L_i + O_i) = 0.02$, $f_{ji}/(L_j + O_j) = 0.006$, $f_{ji}/(L_i + I_i) = 0.667$, $f_{ij}/(L_j + I_j) = 0.0002$

$f_{ik} = 100$, $f_{ki} = 80$, $\Rightarrow f_{ik}/(L_i + O_i) = 0.10$, $f_{ki}/(L_k + O_k) = 0.067$, $f_{ki}/(L_i + I_i) = 0.089$, $f_{ik}/(L_k + I_k) = 0.091$

County i sends only 2% of its outflow to j, and j sends only 0.6% to i – but 67% of all inflow arriving at i comes from j. **County i's economy is dominated by j.**

We propose the following modified metric:

$$P_{ij}^{(3)} \equiv \frac{1}{4}\left(\frac{f_{ij}}{L_i + O_i} + \frac{f_{ji}}{L_j + O_j} + \frac{f_{ji}}{L_i + I_i} + \frac{f_{ij}}{L_j + I_j}\right) \tag{12}$$

pair	$D^{(2)}$	$D^{(3)}$
(i,j)	0.987	0.827
(i,k)	0.917	0.913
clusters i with	k	j ✓

B.3. I_i much smaller than O_i

Example:

$L_i + O_i = 10,000$, $L_j + O_j = 500,000$, $L_k + O_k = 13,000$

$L_i + I_i = 28$, $L_j + I_j = 531,505$, $L_k + I_k = 13,000$

$f_{ij} = 5$, $f_{ji} = 18$, $f_{ik} = 2,000$, $f_{ki} = 3$.

We propose the following modified metric:

$$P_{ij}^{(4)} \equiv \frac{f_{ij} + f_{ji}}{2} \left(\frac{1}{2L_i + I_i + O_i} + \frac{1}{2L_j + I_j + O_j} \right). \tag{13}$$

pair	$D^{(3)}$	$D^{(4)}$
(i,j)	0.8391	0.9988
(i,k)	0.8847	0.8601
clusters i with	j	k ✓

B.4. Ceiling problem for satellite–core pairs.

Example:

Consider the case $f_{ij} = L_i + O_i$, $f_{ji} = L_i + I_i$, $L_j + O_j + I_j \gg L_i + O_i + I_i$, $\Rightarrow P^{(4)} \simeq 0.5$. Even a county that is entirely a bedroom community of a large city can never achieve a dissimilarity below ~ 0.5.

We propose the following modified metric:

$$P_{ij}^{(5)} \equiv \frac{f_{ij} + f_{ji}}{\min(L_i + O_i, L_j + I_j) + \min(L_j + O_j, L_i + I_i)} \equiv \frac{f_{ij} + f_{ji}}{\mathrm{Cap}_{ij}}, \tag{14}$$

where Cap_{ij} is the *bilateral capacity*.

B.5. Probability density compressed near $D^{(5)} = 1$

In $D^{(5)}$, we are dividing purely external flows by a contribution which is dominated by internal flows, giving too much weight to the peak of the distribution at the right.

Instead, we propose the following modified metric (i.e., we now divide by the *external bilateral capacity*):

$$P_{ij}^{(6)} \equiv \frac{f_{ij} + f_{ji}}{\min(O_i, I_j) + \min(O_j, I_i)} \equiv \frac{f_{ij} + f_{ji}}{\mathrm{Cap}_{ij}^{\mathrm{ext}}} \,. \tag{15}$$

B.6. Cross-sectional comparison

We run 15 tests for the 6 distance metrics (the results are shown in Figure A1), structured in the following way:

Category	Tests
Symmetry & baseline	1, 2
Directional dependence	3, 4, 5, 6
Scale / asymmetry	7, 8, 10
Network structure	9a, 9b
Mathematical validity	11, 12
Self-flow (consumption)	13, 14
Clustering behavior	15

More explicitly, these are the 15 tests considered:

1. **Balanced symmetric (with external alternative).** Two counties A and B of similar total activity each spend a large fraction internally (self-flow of 800), and each direct their remaining external activity partly to the other and partly to a larger third county C that acts as the ″rest of the world″ alternative. Specifically, A and B exchange 200 units in each direction (one-third of their external flow) and send 400 units in each direction to C (the remaining two-thirds). This is the canonical baseline case of a genuine but partial bilateral relationship: A and B are connected, but neither county's external life is exclusively tied to the other. A good dissimilarity measure should return a moderate value — clearly below 1 because the link is real, but well above 0 because the bilateral exchange is only a fraction of

each county's external activity. The test checks that the measure behaves sensibly in the most ordinary scenario, where two counties are partially integrated but retain meaningful economic alternatives.

2. **Size asymmetry (moderate)**. One small county and one much larger county exchange flows. From the small county's perspective, the link may represent a meaningful share of its activity, while from the large county's perspective the same flow is negligible. The test checks whether the measure is unduly dominated by the larger county's marginals. A measure that simply averages over both counties' total activity will tend to underestimate the connection, because the large county's contribution dilutes the small county's perspective.

3. **Bedroom community ($L_i + I_i = 0$)**. County i sends 100% of its outflow to county j and receives no flow from anywhere. This is the textbook case of perfect functional dependency: i has no economic life outside its connection to j. A correct dissimilarity measure should give a value very close to 0, signaling that the two counties must cluster together. Measures that ignore inflow or use total marginals in the denominator tend to overestimate the distance and fail this test.

4. **Employment hub**. County i receives a massive inflow from county j but sends very little back. This captures the situation of a small employment or commercial hub serving residents of a larger neighboring county. The test checks whether the measure correctly recognizes inflow as a signal of functional integration. Measures that normalize only by outflow (such as $D^{(1)}$ or $D^{(2)}$) are blind to this kind of asymmetric dependency.

5. **One-directional** ($i \to j$ only). A moderate share — for example 40% — of i's outflow goes to j, with no reverse flow. The link is strong but not exclusive, since iii still distributes the majority of its activity elsewhere. A good measure should return a moderate-to-low value: clearly lower than the baseline but not as low as the bedroom community case. This test checks how the measure scales with the strength of a one-directional connection.

6. **Near-perfect bilateral**. Two counties exchange nearly all of their flow with each other in both directions. This represents the case of two fully integrated economic units. The dissimilarity should be very close to 0, since virtually all of each county's external activity is captured by the bilateral relationship. This is the canonical "should obviously cluster together" case.

7. **Small I_i vs large O_i**. A county has a very small total inflow but a large total outflow. The test stresses measures that divide by the small marginal, such as $D^{(3)}$'s term f_{ji}/I_i, which can overreact to tiny absolute flows and produce unstable values. The test checks robustness when one of the four marginals (O_i, I_i, O_j, I_j) is several orders of magnitude smaller than the others.

8. **Large f_{ji} relative to $L_i + O_i$**. County i receives an inflow from j that is much larger than its own outflow — for example, 50 times $L_i + O_i$. This captures a county whose economic life

is dominated by being a destination, not an origin. The test checks whether the measure recognizes a strong functional link even when it is driven entirely by inflow on one side. Measures that normalize only by outflow miss this entirely.

9a **Chain: A–B (adjacent)**. In a three-county chain A–B–C, counties A and B exchange substantial flow directly. The test checks whether the measure correctly identifies the adjacency by producing a low dissimilarity for this directly connected pair. It also serves as the reference point for the non-adjacent test that follows.

9b **Chain: A–C (non-adjacent)**. In the same A–B–C chain, A and C have no direct flow — all interaction must pass through B. The measure should return a dissimilarity of 1, reflecting the absence of any direct bilateral link. This test checks that the measure does not invent spurious indirect similarity through shared neighbors. A large gap between the values of tests 9a and 9b indicates good resistance to chaining artifacts.

10 **Extreme size asymmetry (90% outflow to large county)**. A very small county sends 90% of its outflow to a mega-county that is orders of magnitude larger. This is the most demanding asymmetric case: the link is essential from the small county's perspective but negligible from the large county's perspective. The measure should give a low value, recognizing that the small county has nowhere else to go. Measures like $D^{(4)}$, which average across both counties' total activity, run into a structural ceiling (around 0.50) and fail to express the strength of the connection.

11 **Boundedness**. The bilateral flow $f_{ij} + f_{ji}$ exceeds the denominator used by $D^{(1)}$, which is $\min(L_i + O_i, L_j + O_j)$. The test checks whether the measure stays within the desirable range [0,1]. $D^{(1)}$ shows negative values, whereas all other measures should remain bounded by construction.

12 No bilateral flow. Two counties have no flow between them in either direction. The dissimilarity should be exactly 1, signaling maximal dissimilarity. This is the baseline "completely disconnected" case and serves as a sanity check that the measure correctly identifies the absence of any link.

13 **High self-spending neighbors**. Two neighboring counties each spend a large fraction of their activity locally (for example, 90% and 80% self-spending), but they also exchange a meaningful amount of flow with each other. This is the canonical consumption-data scenario. The test checks whether the measure can detect the genuine inter-county link without being washed out by self-flow. $D^{(5)}$ fails here because total marginals are inflated by self-spending, compressing the dissimilarity near 1. $D^{(6)}$ correctly identifies the proximity by normalizing only by external capacity.

14 **High self-spending distant**. Same setup as test 13 (two counties with high self-spending), but now the pair has a much weaker bilateral exchange — they are spatially or functionally

distant. The dissimilarity should be higher than in test 13. This test checks that the measure can still distinguish near from far counties in the high self-spending regime, which is essential for producing meaningful zones in consumption data.

15 $D^{(4)}$ **merge-order stress (100% one-directional to large county)**. A small county sends 100% of its outflow to a large hub, with no reverse flow. This is a classic "must-merge" situation — the small county has no functional alternative. A correct measure should give a value close to 0, so that this pair is the first to be merged in a hierarchical clustering algorithm. $D^{(4)}$ fails decisively, returning approximately 0.50 due to its structural ceiling, which causes weaker but bilaterally balanced pairs to be merged before this essential link. The test directly exposes how a flawed dissimilarity produces an incorrect merge order in the clustering pipeline.

However, although $D^{(6)}$ has the best theoretical properties, in practice we observe only small differences in the ConZ configurations obtained using $D^{(1)}$ and $D^{(6)}$. For this reason, we conduct all the analysis in this paper using only $D^{(1)}$.

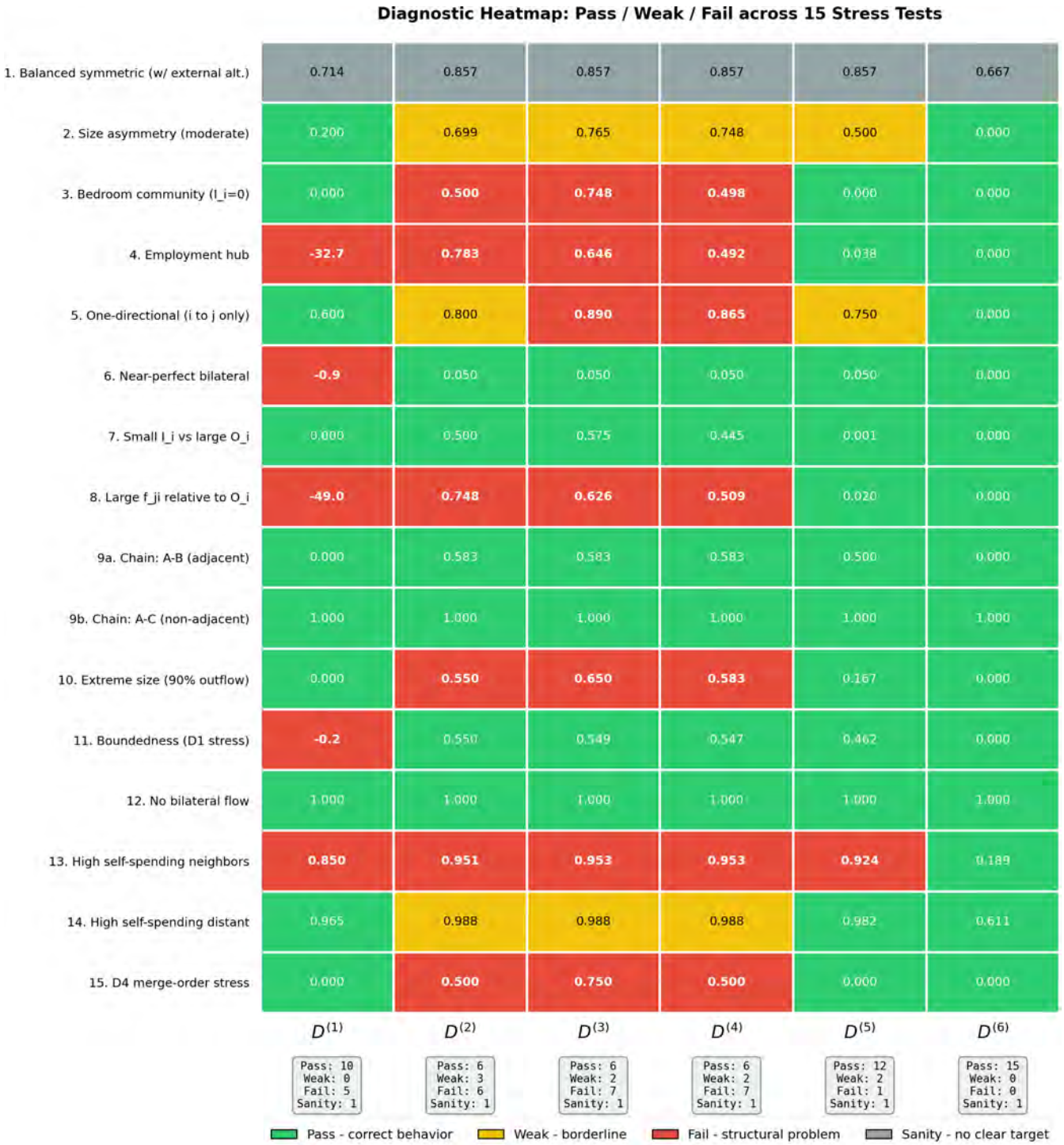
C. Prior Algorithm Applied by Foote, Kutzbach and Vilhuber (2021)

The algorithm applied in Foote, Kutzbach and Vilhuber (2021) effectively used a Euclidean geometry to adjust the distance, as this is a default option in the SAS hierarchical clustering algorithm, basically calculating the Euclidean distances between the distances metric provided. This approach led to suboptimal clusters and less transparency as it is not using the distances directly provided by the Tolbert and Sizer (1987) algorithm, but is, instead, computing the Euclidian distances of those distances. We believe this generated some normalization in the distances across industries in Batch et al. (2025), allowing for a single height to reasonably be applied across industries, although we believe this approach led to suboptimal clusters and less transparency in the algorithm.

D. Measuring the stability of zones over time

Set-up and notation. Fix an industry and let $\mathcal{C}$ be the set of counties observed in both 2019 and 2024, with $N = |\mathcal{C}|$. The zone assignment in each year is a partition of $\mathcal{C}$: write $\mathcal{P} = \{P_1, \ldots, P_K\}$ for the 2019 zones and $\mathcal{Q} = \{Q_1, \ldots, Q_L\}$ for the 2024 zones. Let $n_{k\ell} = |P_k \cap Q_\ell|$ be the number of counties in 2019 zone k that lie in 2024 zone ℓ, with row and column totals $n_{k\cdot} = |P_k|$ and $n_{\cdot\ell} = |Q_\ell|$. Each county i carries a population weight w_i (we use 2015 population). For any subset $G \subseteq \mathcal{C}$ we

Figure A1. Comparison of performance across the different dissimilarity metrics based on 15 tests.



write $w(G) = \sum_{i \in G} w_i$ and $W = w(\mathcal{C})$. Every measure below is invariant to how the zones are labeled, since each depends only on the co-membership counts $\{n_{k\ell}\}$ and not on the identities of the zones; this is essential because the labels are produced independently in each year.

Fraction of counties retaining their grouping (primary measure). Our main measure is the share of counties that stay with the bulk of their 2019 zone, which we view as the clearest summary of stability. For each 2019 zone k we identify the 2024 zone receiving the largest number of its counties and treat those counties as retained:

$$R = \frac{1}{N} \sum_{k} \max_{\ell} n_{k\ell}. \tag{16}$$

The population-weighted version selects, for each 2019 zone, the 2024 zone capturing the largest share of its population:

$$R_w = \frac{1}{W} \sum_{k} \max_{\ell} w(P_k \cap Q_\ell). \tag{17}$$

R is the accuracy of the best plurality map from 2019 zones to 2024 zones. Because two 2019 zones may send their pluralities into the same 2024 zone, R is an upper bound on the accuracy of the optimal one-to-one zone matching; in practice the two coincide unless zones overlap heavily. Both versions condition on the 2019 partition; the quantities that condition on 2024 are defined symmetrically. A county that is its own zone in both years counts as retained, since its 2019 zone of size one maps entirely into whichever 2024 zone that single county joins.

Pairwise (Jaccard) similarity (conservative complement). The retention rate asks only whether each county lands with the *plurality* of its former neighbours. As a stricter complement we compare the two partitions pair by pair, which additionally penalizes any reshaping of a zone short of a wholesale move. Consider the $\binom{N}{2}$ unordered pairs of counties; a pair is *co-grouped* in a year if both counties lie in the same zone that year. Define

$$a = \sum_{k} \sum_{\ell} \binom{n_{k\ell}}{2}, \qquad a + b = \sum_{k} \binom{n_{k\cdot}}{2}, \qquad a + c = \sum_{\ell} \binom{n_{\cdot\ell}}{2}, \tag{18}$$

so a counts pairs co-grouped in both years, while b and c count pairs co-grouped in only 2019 or only 2024. The Jaccard index is

$$J = \frac{a}{a+b+c} = \frac{\sum_{k,\ell} \binom{n_{k\ell}}{2}}{\sum_{k} \binom{n_{k\cdot}}{2} + \sum_{\ell} \binom{n_{\cdot\ell}}{2} - \sum_{k,\ell} \binom{n_{k\ell}}{2}}, \tag{19}$$

which equals 1 when the partitions coincide and falls toward 0 as counties are regrouped; we report $1 - J$ as the share of ever-linked pairs that are reshuffled. Because it is computed over pairs,

the splitting or merging of a single large zone moves many pairs at once, so J lies systematically below R and serves as a demanding lower benchmark.

Population-weighted Jaccard. To ask whether instability is driven by small counties, we weight each pair $\{i, j\}$ by the product $w_i w_j$. For any subset G the total pair weight is

$$\pi(G) = \sum_{\{i,j\}\subseteq G} w_i w_j = \tfrac{1}{2}\left[w(G)^2 - \sum_{i\in G} w_i^2 \right]. \tag{20}$$

Replacing each pair count by its weighted analogue gives

$$a_w = \sum_{k,\ell} \pi(P_k \cap Q_\ell), \quad (a+b)_w = \sum_k \pi(P_k), \quad (a+c)_w = \sum_\ell \pi(Q_\ell), \tag{21}$$

$$J_w = \frac{a_w}{(a+b)_w + (a+c)_w - a_w}. \tag{22}$$

The unweighted index (19) is the special case $w_i \equiv 1$. When $J_w \gg J$, high-population counties remain co-grouped while the reshuffling is concentrated among small counties.

Adjusted Rand Index. As a chance-corrected check we report the Adjusted Rand Index,

$$\mathrm{ARI} = \frac{a - \mathbb{E}}{\frac{1}{2}[(a+b)+(a+c)] - \mathbb{E}}, \qquad \mathbb{E} = \frac{(a+b)(a+c)}{\binom{N}{2}}, \tag{23}$$

where $\mathbb{E}$ is the number of co-grouped-in-both pairs expected if the two partitions were independent given their zone sizes; $\mathrm{ARI} = 1$ for identical partitions and $\mathrm{ARI} \approx 0$ for unrelated ones. Because chance correction has no standard population-weighted analogue, we report the ARI in unweighted form only.

Treatment of singletons. A county that forms a zone of size one, which we refer to as a singleton, generates no within-zone pairs. A county that is a singleton in *both* years is therefore invisible to J and J_w (it contributes nothing to a, b, or c) yet still enlarges the pair universe $\binom{N}{2}$ entering the ARI. Let $\mathcal{S} = \{\, i : n_{k(i)\cdot} = 1 \text{ and } n_{\cdot\ell(i)} = 1 \,\}$ collect these counties and let $s = |\mathcal{S}|$. We consider three conventions. *(i) Ignore*: retain them but leave them uncredited, as in the formulas above. *(ii) Drop*: remove $\mathcal{S}$ before computing; this leaves J and J_w unchanged but removes the singletons' distortion of the ARI. *(iii) Match*: credit each both-year singleton as one unit of agreement, on the ground that a county that remains its own zone has not been regrouped. Concretely we add s to a and to both year totals, so that

$$J^{\mathrm{match}} = \frac{a+s}{(a+s)+b+c}, \tag{24}$$

add $\sigma = \sum_{i\in\mathcal{S}} w_i^2$ to a_w and both weighted totals for J_w, and extend the pair universe in (23) to $\binom{N}{2} + s$ so that the ARI remains internally consistent. The retention measure R is unaffected by this

choice, as it already credits both-year singletons. Our baseline tables use the *match* convention; results under the other two are similar, and by construction the *ignore* and *drop* conventions leave J and J_w unchanged.

E. Sensitivity Analysis, Validation, and Limitations

E.1. Data and calibration robustness

We tested whether the consumption zone methodology is robust to data source and calibration strategy by comparing the number of consumption zones produced using different data sources and containment thresholds. Table A1 presents these results, showing the number of consumption zones generated when using three different data source—Mastercard 2024 transactions, Mastercard 2019 transactions, and Fiserv 2015 flows—and four alternative containment levels—50, 75, 80, and 85%. The zone counts are also compared to the number of zones in Batch et al. (2025).

This table shows encouraging consistency across data source. By and large, the number of consumption zones across Fiserv and Mastercard data are substantially higher in local industries such as *Food Services and Drinking Places* and fewer zones for comparatively nonlocal sectors such as *Gasoline Stations*.

Relative patterns in consumption zones are also consistent across different containment thresholds. Table A1 compares the consumption zones count for our baseline results (using 80% containment) to two other candidate thresholds consistent with the related literature. As expected, the number of consumption zones decreases as the required level of containment increases, and we see similar relative patterns in the number of zones across industries. This robustness check suggests that the comparative results outlined above do not depend on the choice of data or containment threshold.

E.2. Geographically disperse consumption

Consumption zones best capture consumer behavior for local goods and services. Spending categories with weak local geographic ties do not have a clear interpretation when grouped into consumption zones.

Online spending provides the best example of this limitation in the consumption zones framework. Online purchases are not limited by the physical location of retailers and do not provide information about the geographic area in which consumers travel to make purchases. Consumption zones

Table A1. Cluster Counts by Industry and Calibration Method

Source	Batch et al			Fiserv 2015				Mastercard 2019				Mastercard 2024	
Calibration	H=0.9	85%	80%	75%	50%	85%	80%	75%	50%	85%	80%	75%	50%
Industry													
Aggregated consumption	1235	199	819	1823	3128	213	776	1843	2994	208	727	1866	3114
Accommodation	383	6	6	7	33	4	4	5	47	5	5	6	32
Ambulatory Services	800	675	1509	2442	3128	611	1467	2482	2994	638	1615	2593	3114
Amusement, Gambling...	530	17	72	268	3128	17	29	59	2264	21	32	69	2300
Building Material and Garden...	1194	1306	2398	3128	3128	603	1341	2357	2994	540	1136	2146	3114
Clothing	567	39	142	426	3128	72	221	560	2994	47	142	383	3114
Food Services and Drinking Places	997	182	586	1514	3128	80	291	794	2994	62	214	559	3114
Food and Beverage Stores	1862	2612	3128	3128	3128	2953	2994	2994	2994	2658	3114	3114	3114
Furniture Stores	570	79	386	866	3128	139	411	884	2994	108	337	700	3114
Gasoline Stations	972	128	555	1640	3128	231	726	1942	2994	223	691	1697	3114
General Merchandise Stores	1198	1034	2388	3128	3128	221	729	1792	2994	230	708	1647	3114
Miscellaneous Store Retailers	900	28	143	483	3128	97	318	851	2994	19	21	46	3114
Performing Arts, Sports...	322	7	10	20	1398	29	30	36	1919	18	20	22	493
Personal Services	955	602	1519	2845	3128	149	540	1262	2994	201	561	1325	3114
Repair and Maintenance	1009	781	1676	2686	3128	437	1254	2514	2994	497	1322	2583	3114
Sporting Goods, Hobby, Book...	554	57	224	728	3128	73	235	570	2994	57	160	385	3114

Note. This table is constructed using both Fiserv and Mastercard aggregated and anonymized transaction data.

constructed with online spend data would over-agglomerate counties, highlighting the disperse nature of e-commerce sales and limitations of the clustering methodology.

Consumption zones may be relevant for nonlocal industries after excluding major national destinations. For instance, *Accommodations* spending is disproportionately concentrated in a handful of major tourist destinations and commercial hubs. These counties attract a larger share of national spending due to long-distance travel and destination-specific demand, which limits their relevance when connecting local consumer behavior. *Performing Arts* has a similar pattern of national destinations: Mastercard spending data shows that New York City and Los Angeles dominate national spending, but local theaters and performing arts centers are still important spending destinations.

We conduct additional sensitivity analysis to test whether local consumption zones of geographically disperse sectors such as *Accommodations* and *Performing Arts* emerge after excluding destination counties. For both sectors, we compare the number of clusters when excluding the top 10 and 100 counties with the most spending. Table A2 shows the share of spending that occurs in the top 10 and 100 counties for each industry, with the top counties comprising a disproportionately higher share of spending for both *Accommodations* and *Performing Arts*. When excluding these counties from the consumption zones algorithm, we see little increase in the number of zones: *Performing Arts* consumption zones increased by 11 and 15 when excluding the top 10 and 100 counties, respectively; *Accommodations* adds only one zone when excluding either the top 10 or 100 counties. These results suggest that national destinations cannot explain the low level of consumption zones on their own, but regional or even local consumption zones are best revealed by lowering the containment threshold.

E.3. Validation exercises

Sector consumption zones were validated against other characteristics of spending data. We first identified potentially concerning sectors using the ratio of the number of consumption zones for each industry relative to *Gasoline Stations* spending as the benchmark sector. Sectors with extreme ratios were further analyzed to uncover possible data or methodology issues. For instance, *Building Material and Garden Equipment and Supplies Dealers* have many more consumption zones than *Gasoline Stations* despite only one third of unique merchant locations.

Validated sectors were compared to *Gasoline Stations* across several dimensions including the modal distance traveled by consumers and the cumulative distribution of spending by distance. When comparing the relative number of consumption zones and distance traveled for a fixed containment threshold, we found a consistent ratio of consumption zones and the square of the

Table A2. Top N Counties Share of Spending by Industry (%)

Top N Counties Industry	10	100
Performing Arts, Spectator Sports, and Related Industries	43	88
Accommodation	29	66
Clothing and Clothing Accessories Stores	26	63
Amusement, Gambling, and Recreation Industries	25	60
Miscellaneous Store Retailers	21	54
Personal and Laundry Services	21	60
Food Services and Drinking Places	18	53
Furniture and Home Furnishings Stores	17	58
Aggregated consumption	16	49
Ambulatory Health Care Services	15	53
Food and Beverage Stores	15	49
Repair and Maintenance	14	47
Sporting Goods, Hobby, Book, and Music Stores	14	49
General Merchandise Stores	12	42
Building Material and Garden Equipment and Supplies Dealers	11	40
Gasoline Stations	10	35

Note. This table is constructed using Mastercard aggregated and anonymized transaction data.

distance traveled by consumers.[16] This consistency suggests that the size of consumption zones comoves with the distance traveled by consumers.

F. Additional Tables and Figures

F.1. Comparison With Prior Zones

[16] This argument assumes that the number of ConZs for a given spending category is approximately

$$\#(ConZ) \sim \frac{\text{area(USA)}}{\pi r^2},$$

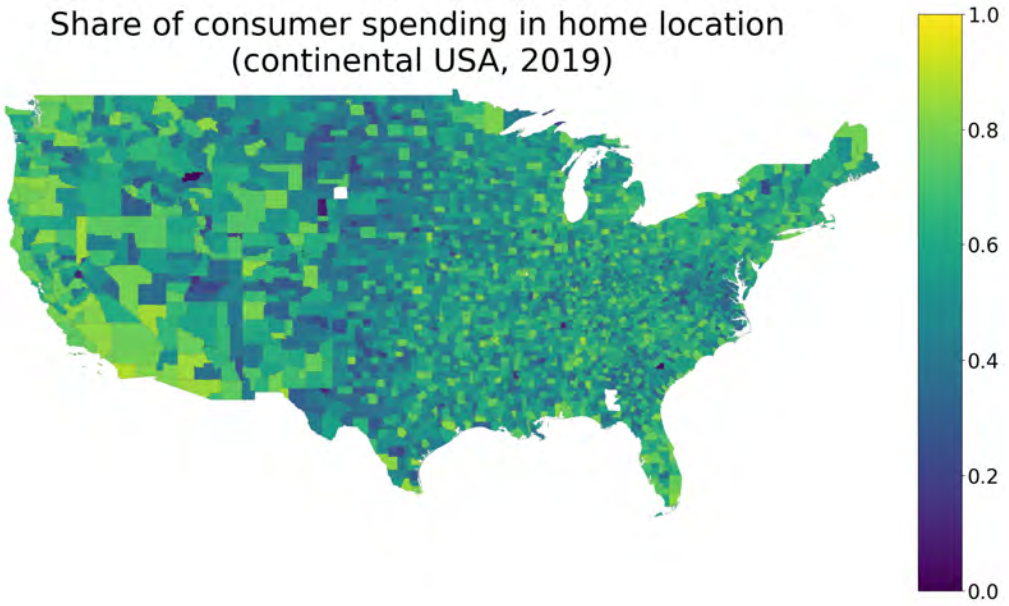
where r is the distance inferred from Figure 2 or Figure A3 for a given level of containment (approximately the share of total consumption corresponding to that travel distance).

Table A3. Number of Consumption Zones by Industry Across Data Sources 78% containment

Industry	Batch et al.	Updated ConZs	Mastercard 2024
Aggregated consumption	1235	1235	1330
Accommodation	383	7	17
Ambulatory Health Care Services	800	1926	2554
Amusement, Gambling, and Recreation Industries	530	163	188
Building Material and Garden Equipment	1194	2757	2086
Clothing and Clothing Accessories Stores	567	234	461
Food Services and Drinking Places	997	866	448
Food and Beverage Stores	1862	3128	3028
Furniture and Home Furnishings Stores	570	559	1346
Gasoline Stations	972	958	1506
General Merchandise Stores	1198	2859	1745
Miscellaneous Store Retailers	900	269	74
Performing Arts, Spectator Sports, and Related Ind.	322	12	52
Personal and Laundry Services	955	2220	1493
Repair and Maintenance	1009	2127	2985
Sporting Goods, Hobby, Book, and Music Stores	554	437	610

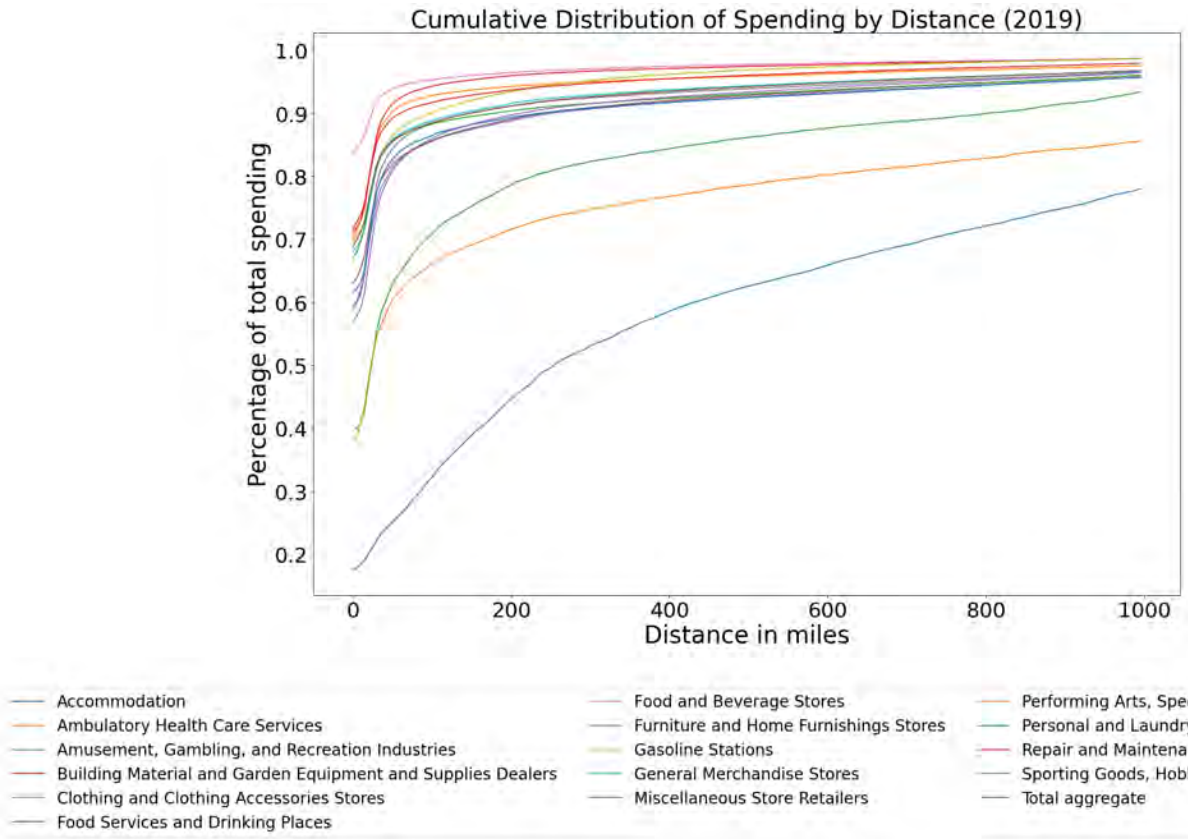
Note. We report 78% containment here because it is the level at which the original zones of Batch et al. (2025) contain spending, allowing a like-for-like comparison of zone counts." This ties the 78% calibration fact directly to the 80% baseline. This table is constructed from both Fiserv and Mastercard aggregated and anonymized transaction data.

Figure A2. Share of Aggregate Spending Occurring in Home County in 2019



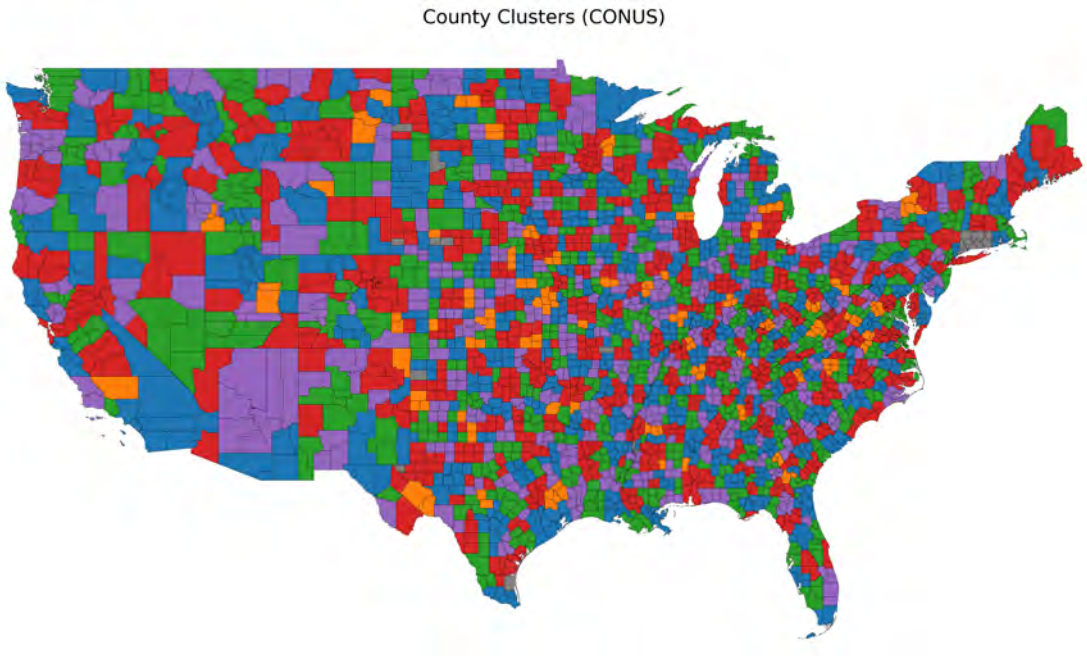
Note. This figure displays aggregate share of a consumers spending occurring in the home county, relative to other locations for the year 2019 based on Mastercard aggregated and anonymized transaction data.

Figure A3. Cumulative Distribution of Spending by Industry in 2019



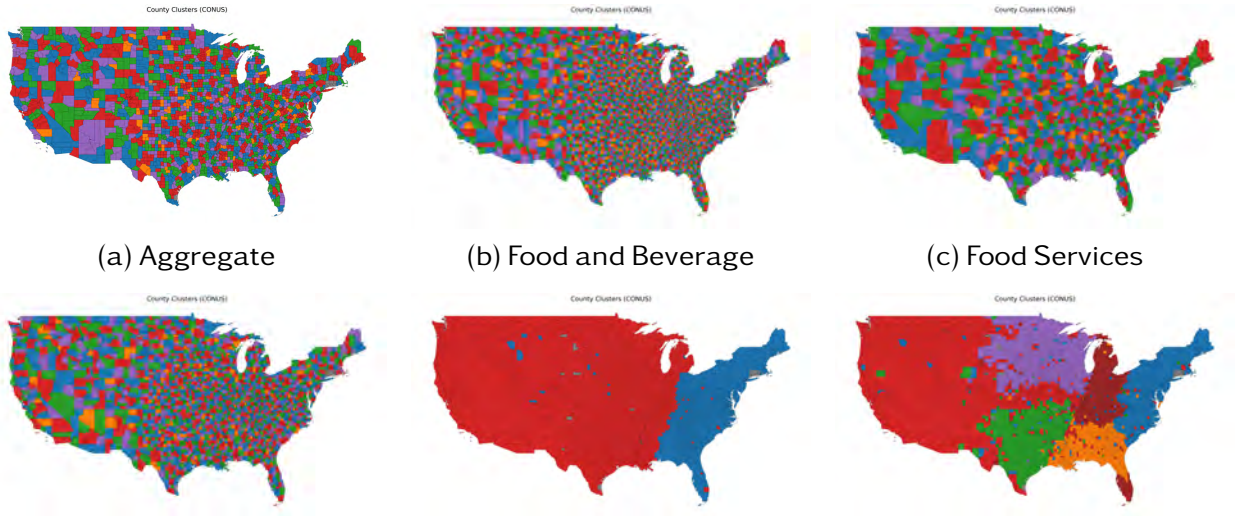
Note. This figure displays the cumulative distribution of spending by industry using Mastercard aggregated and anonymized transaction data. All industries show a greater share of spending captured as one moves farther away from the home county.

Figure A4. Aggregate County Clusters (Fiserv 2015 Containment 80%)



Note. This figure shows that corresponding ConZ maps using Fiserv data from 2015, rather than Mastercard data.

Figure A5. County Clusters Fiserv 2015 data containment 80%



(a) Aggregate (b) Food and Beverage (c) Food Services

(d) Ambulatory Care (e) Accommodation (f) Performing Arts

Note. Each panel displays county clusters constructed from Fiserv transaction flows.

Table A4. HHI Sales Weighted for 2019 Using New Fiserv Zones

	FIPS	ConZ	CZ	State
Furniture and Home Furnishings (NAICS 442)	1,003	442	617	226
Building Material and Garden Equip. (NAICS 444)	2,471	2,300	1,949	1,463
Food and Beverage Stores (NAICS 445)	2,095	2,095	1,762	1,129
Gasoline Stations (NAICS 447)	3,783	3,115	3,281	2,059
Clothing and Clothing Accessories (NAICS 448)	668	302	468	243
Sporting Goods, Hobby, Books, and Music (NAICS 451)	1,072	497	750	360
General Merchandise Stores(NAICS 452)	3,372	3,232	2,998	2,637
Miscellaneous Store Retailers (NAICS 453)	970	390	656	294
Ambulatory (NAICS 621)	629	495	422	138
Performing Arts, Spec Sports, and Related Ind. (NAICS 711)	1,683	184	1,222	559
Amusement, Gambling, and Recreation (NAICS 713)	1,563	391	1,184	493
Accomodations (NAICS 721)	1,333	97	926	412
Food Service and Drinking Places (NAICS 722)	198	92	114	42
Repair and Maintenance (NAICS 811)	571	435	336	88
Personal and Laundry Services (NAICS 812)	582	440	386	127

Note. This table is constructed using Mastercard aggregated and anonymized transaction data.