Lost Spending Heterogeneity in Zip Codes*

Sinem Hacıoğlu-Hoke *
Federal Reserve Board
CEPR

Leo Feler†
Numerator

Jack Chylak‡
Numerator

Preliminary and Incomplete
Please do not quote without the authors' permission
March 19, 2025

Abstract

Using zip code level income as a proxy for household income in analyzing spending masks the heterogeneity among households that emerged after 2021. We reach this conclusion by using a panel of 150,000 representative US households for whom we have both zip code and categorical income information alongside their total monthly retail spending from 2018 to 2024. We construct average retail spending measures by reported income groups and zip code level income. When we use reported income, spending patterns of low- versus high-income households diverge starting in 2021. The divergence in spending disappears when we use zip code level incomes. We find 30 to 70 percent discrepancy between groups classified based on reported versus zip code level income.

Keywords: Spending, Income, Heterogeneity, Zip code
JEL Classification:

*sinem.haciogluhoke@frb.gov Web: www.sinemhaciogluhoke.com
†leo.feler@numerator.com Web: www.leofeler.com
‡jack.chylak@numerator.com

* The authors would like to thank Tomaz Cajner, Ryan Decker, Aaron Flaaen. The views expressed in this paper are the responsibility of the author and should not be interpreted as reflecting the view of the Board of Governors of the Federal Reserve System or of any other person associated with the Federal Reserve System, or Numerator. All errors remain our own.

1 Introduction

Household income is a key metric in social sciences to understand changes in consumer behavior. It is common to categorize households as low-, middle- or high-income households to assess the heterogeneity in e.g. consumers' spending behavior, sentiment about the economy, housing situation etc. When household incomes are not observed, it has been the practice to use the median incomes in zip codes where households live in, generally obtained by population survey, as a proxy for income. The reasons behind such an approximation of household income using zip codes is two fold: first, where households afford to live is correlated with their income level; second, it is difficult to come by datasets which provide information on household income whereas it is more common to observe households' zip codes and infer their incomes based on zip code income. Due to these challenges, it is difficult to accurately gauge the heterogeneity in consumer behavior that emerges from differences in income, and assess if zip code income is a valid proxy for household income.

This paper has two objectives that lead to novel contributions to existing studies in documenting heterogeneity in spending growth. First, using a detailed micro dataset, we construct a measure of real average retail spending for low-, middle- and high-income households using households' reported income groups. The data we use is a panel dataset of 150,000 representative US households from Numerator which is a consumer data and survey company. Numerator obtains physical and online receipts from a rotating panel of 150,000 US households for whom we have detailed information on household attributes. We first compare our overall measure of monthly average retail spending between 2018 and 2024 against the Census Bureau's Advance Monthly Sales for Retail and Food Services report (MARTS) estimates and establish our measure's reliability to a great detail.

Having established that the aggregated Numerator retail spending series closely matches the Census Bureau's retail sales series, we examine spending patterns based on household income. We analyze the average monthly household retail spending for low-income households ($0-$60K in annual household income), for middle-income households ($60K-$100K), and for high-income households ($100K+) from 2018 to 2024. Published measures do not provide details on *which* consumers' spending has remained resilient and

hence fall short of identifying vulnerabilities in the economy originating from specific groups, so our analysis fills this gap. Our results suggest that retail spending growth evolves similarly for all households before the pandemic. However, a divergence emerges between the spending of high-income households vs low- and middle-income households: high-income households' spending fuel the economy while low- and middle-income households' spending lag behind after 2021. This finding is supported by many[1] and can be rationalized by low-income households' depleting their pandemic-era savings (Abdelrahman et al., 2024) and government support programs helping low- and middle-income households expiring after the pandemic, while high-income households' benefit from welfare gains.

The second objective and novelty of this paper is to use our micro-data to put the implication of using zip code level aggregation into a test. Specifically, we construct spending measures for different income groups that rely on households' proxied income using the zip codes they live in, instead of using households' reported income. In the onset of the pandemic, the Economic Tracker by Chetty et al. (2023) provided novel high-frequency estimates of the contribution of different income groups' spending from 2020 onward and illustrated the downside risks to the economy using credit card spending data. Their analysis proxies income by the median zip code level income using the American Community Survey (ACS). Our analysis of using Numerator data and replicating zip code level aggregation is unique in its nature of allowing a direct comparison of spending measures based on reported versus proxied income. We repeat the zip code level aggregation with our micro dataset by using median zip code level incomes reported in the ACS to classify households as low-, middle- and high-income households, instead of using households' reported income groups. We document similar pre-pandemic spending

[1]Such analysis of documenting the discrepancies in spending growth is sparse at best in academic platforms, mostly due to data availability to assess spending by income groups. One such analysis is Moody's Analytics', which show that top earners – top 10% of US households in terms of earnings – drive nearly half of all consumer spending. For the details of the analysis https://www.wsj.com/economy/consumers/us-economy-strength-rich-spending-2c34a571?st=2odEgM&reflink=article_copyURL_share. The Moody's Analytics analysis postdates our initial analysis which shows the divergence in spending between income groups in Hacioglu Hoke et al. (2024). Another related study is by Bank of America Institute that documents the slowing pace of lower-income households' spending: https://institute.bankofamerica.com/economic-insights/consumer-checkpoint-march-2024.html. Finally, Morning Consult analysis also documents slower spending by low- and middle-income households: https://pro.morningconsult.com/analysis/consumer-spending-september-2024.

dynamics however the divergence we observe in the post-2021 spending disappears. This results is at odds with the analysis showing that the wealthiest households are driving spending in the US in the post pandemic period.[2]

A potential concern related to the observed divergence between the spending of different income groups might be that it is driven by compositional changes in income groups. If more households become high-income over time and fewer households remain low-income, that affects the composition of households in each income group and their average monthly retail spending. To control for the higher mobility between income groups in 2022 and 2023 due to higher inflation adjustments on income, we use a bespoke survey, run by Numerator, to obtain a smaller subset of households' nominal incomes in 2019 and 2024. We calculate the share of respondents that move across income groups from 2019 to 2024 based on our definition of low-, middle- and high-income households. Consistent with the share of households that move across income groups based on the survey responses, we design a simulation exercise where we artificially move households from their income groups in 2019 to new income groups in 2024. We show that the change in the composition of income groups due to rise in incomes in the post-pandemic period does not drive the divergence between low- versus high-income households and divergence in spending remains as a robust finding.

The validity of using zip code income (or in general household characteristics) as a proxy for household income has faced doubts mostly in the health literature. Several studies explore the validity of using aggregate characteristics based on geography to proxy household characteristics, such as education or income. One such study is Geronimus et al. (1996), which documents the issues associated with using aggregate census-based variables as proxies for micro-level socioeconomic characteristics. Another is Hanley and Morgan (2008), which uses Canadian administrative data to document the large variability between household-level income and area-based income. More recently, Buajitti et al. (2020) which use Canadian data calls for caution when using area-level data as measure of individual-level socioeconomic status.

Our analysis mainly builds on studies that assess how consumer spending evolves during and after the pandemic using private sector data, see Vavra (2021) and Brodeur

[2]See footnote 1 for the references of these analyses.

et al. (2021) and there references therein. As we discussed in detail above, Chetty et al. (2023) provides regular updates on how spending by income group changes, however their income groups rely on zip code level income information. Another such study is Cox et al. (2020) which documents the heterogeneous effects of the pandemic on households using household-level bank account data. Their analysis of spending by income quartiles sheds light onto the differences in spending on the onset of the pandemic. However, as they acknowledge as a limitation, the micro-data on income in the bank account data are not available at the time of their writing, so they use publicly available data to simulate how incomes changed in the first few months of the pandemic. Our analysis contributes to this literature by i) using a micro dataset where we observe income groups reported by households and hence are readily available; ii) exploring changes to spending by income groups for a relatively long time period that includes pre- and post-pandemic periods; iii) comparing the implications of using reported income groups versus proxied income using zip code incomes in assessing changes in spending. Moreover, we have the ability to update our analysis frequently, as in Chetty et al. (2023), so that we can assess the changes to spending dynamics almost in real-time. This unique feature of our analysis enables us to easily detect the divergence in spending that emerges in the post-pandemic period.

The rest of the paper proceeds as follows. Section 2 provides the details of the Numerator data and establishes the reliability of its overall real average retail spending measure by comparing it officially published statistics. In Section 3, we show the evolution of seasonally adjusted real average retail spending measures for low-income, middle-income, and high-income households, constructed by reported income groups, between 2018 and 2024. Section 4 abstracts from using Numerator's reported income groups in constructing the measures of spending by income groups. Instead, a la Chetty et al. (2023)'s Economic Tracker, we use zip code level aggregation. We document that zip code level aggregation leads to similar results for the pre-pandemic period dynamics of retail spending but it masks the heterogeneity in spending that emerged in the post-pandemic period. Section 5 concludes with a summary of our results. Appendix provides additional analysis and results.

2 Data

The Numerator data contain 150,000 users who self-identify as the primary shopper in the household and whose transactions are continuous and complete for a period of at least 12 months. These 150,000 users are selected from over 1 million users in a way that is demographically and nationally representative. Numerator provides weights for households to ensure a match with Census demographic data and to ensure their detailed purchases, when summed by retailer or by manufacturer, align with quarterly earnings reports of major retailers and consumer packaged goods companies.

Numerator collects data from households in several ways using a mobile phone app. First, consumers can upload a picture of their paper receipts. Second, they can allow Numerator to scrape their emails for digital receipts. Third, users can link loyalty and membership accounts sot hat their transactions are automatically recorded.[3] Users' identifiable information is anonymized but Numerator collects survey information on household income, zip code, age, education, household size, race/ethnicity, and various other measures.

The Numerator data allow for 16 income groupings that panelist self-report.[4] Numerator re-surveys all panelists approximately every 12 months and often more frequently if panelists report life events. Panelists report any changes to their income when they are re-surveyed. Although there may be a lag in detecting changes in household income, this timing provides confidence that the margin of error in accounting for changes in household income is within reasonable limits.

To establish the reliability of the average retail spending measure we construct using the Numerator data, we compare it against the official statistics published for retail spending. Every month, the Census Bureau publishes the MARTS to provide an estimate of retail and food services sales. To prepare the MARTS report, the Census Bureau col-

[3]Numerator collects these data from households using a mobile phone app called Receipt Hog. Panelists are rewarded with coins redeemable for Amazon or Visa gift cards or for cash through PayPal. Panelists are rewarded directly with coins and with virtual spins of a slot machine that also rewards coins, which panelists can redeem for Amazon or Visa gift cards or for cash through PayPal. On average, Numerator rewards panelists approximately $10 per month for providing their purchase information and completing surveys.

[4]The income groups start with <$20K and go up to $100K with $10K increments. After $100K, they go up with $25K increments up to $250K with the last group being $250K+.

lects the dollar value of sales directly from retailers by fielding establishment level surveys and estimating total results and the results by the relevant 3-digit NAICS categories.

The Numerator panel data include information on households' purchases from paper receipts, online receipts, and transaction records. The purchase information contains the quantity and price of each item purchased, the overall total on the receipt, and where and when the purchase took place. These details allow us to classify households' spending into the relevant 3-digit NAICS categories for retail sales categories to construct a spending measure that compares with US Census Bureau's retail sales measure in the MARTS report.[5] Note that our measure excludes spending on motor vehicles and parts, a spending type Numerator data do not capture.

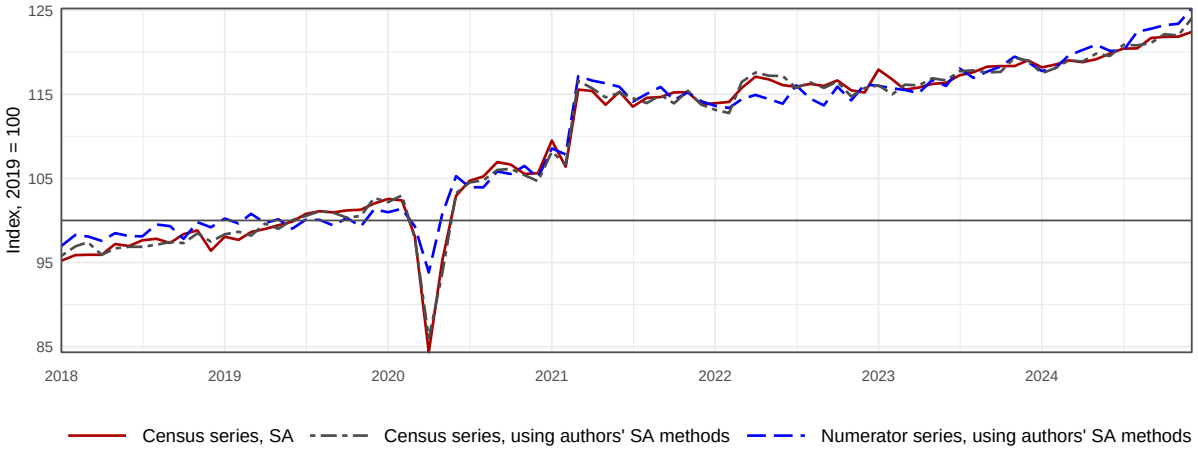
Figure 1 compares how the retail spending measure we construct from Numerator compares with Census Bureau's published series. The red line indicates the Census Bureau's published seasonally-adjusted total retail sales excluding motor vehicles and parts. The blue dashed line is the retail spending measure we construct from the Numerator data, which we seasonally adjust using the Census Bureau's X13-ARIMA-SEATS package.[6] The gray dash-dotted line is the non-seasonally adjusted Census series that we then seasonally adjust using the same X13-ARIMA-SEATS methodology we use for Numerator data. The purpose of the gray line is to show that our seasonal adjustment methodology aligns closely with the Census methodology, and therefore, the differences between the Numerator seasonally-adjusted series and the official Census seasonally-adjusted series is not due to differences in seasonal adjustment.[7]

[5]These eleven spending categories are Furniture & home furn. stores (NAICS 442), Electronics & appliance stores (NAICS 443), Building material & garden eq. & supplies dealers (NAICS 444), Food & beverage stores (NAICS 445), Health & personal care stores (NAICS 446), Gasoline stations (NAICS 447), Clothing & clothing accessories stores (NAICS 448), Sporting goods, hobby, musical instrument, & book stores (NAICS 451), General merchandise stores (NAICS 452), Miscellaneous store retailers (NAICS 453), Non-store retailers (NAICS 454), and Food services & drinking places (NAICS 722). According to the Census Bureau: MARTS covers firms classified in the Retail Trade and Food Services sectors as defined by the North American Industry Classification System (NAICS). Retail Trade, as defined by NAICS sectors 44-45, includes establishments engaged in selling merchandise in small quantities to the general public, without transformation, and rendering services incidental to the sale of merchandise. Food and drinking services to final consumers, subsector 722, is also included in the survey. More information on how the data are collected is available at `https://www.census.gov/retail/how_surveys_are_collected.html`.

[6]In addition, due to a change in Numerator methodology starting in April 2021, we apply a fixed level-correction to the Numerator series for April 2021 and onwards. This level-correction is equivalent to a fixed-effect that accounts for the change in Numerator's methodology.

[7]It is not entirely possible for us to replicate the Census Bureau's seasonal adjustment methodology.

FIGURE 1: Retail sales, total excluding motor vehicles and parts constructed using Census Bureau estimates and Numerator panel data



Note: The solid red line is the Census Bureau's seasonally-adjusted retail sales series excluding motor vehicles and parts. The dashed blue line is the retail spending measure we construct from the Numerator data, which we seasonally adjust. The dashed gray line is the non-seasonally adjusted Census series that we then seasonally adjust. The data are monthly from January 2018 through December 2024. Authors' seasonal adjustment method uses X13-ARIMA-SEATS. Inflation adjustment uses the chain indexed PCE deflator for goods and food services excluding motor vehicles. All three lines are indexed to 100 in 2019.

The correlation between the Numerator series and the Census Bureau's official seasonally-adjusted retail sales series is 0.94 in levels and 0.83 when looking at monthly percent changes. Figure 1 and the correlation of Numerator's retail spending measure with the officially published retail sales measure allow us to verify that the bottom-up construction of the Numerator data, captured from household purchases, closely matches the top-down construction of the Census Bureau's retail sales series, captured from surveys of retail and food service establishments. Appendix A provides additional figures and reports correlations for sub-spending categories.

Retail sales subcategories are seasonally adjusted separately, sometimes even at a more granular level, and then combined to sum to the total retail sales series published by the Census Bureau. This is the main reason why we perform our own seasonal adjustment in the series we work with. Due to the time span of the data and the reasons we explained above, we can get close to but cannot fully replicate the officially published seasonally-adjusted retail sales series. More information on the Census Bureau's seasonal adjustment can be found at https://www.census.gov/retail/marts/www/timeseries.html.

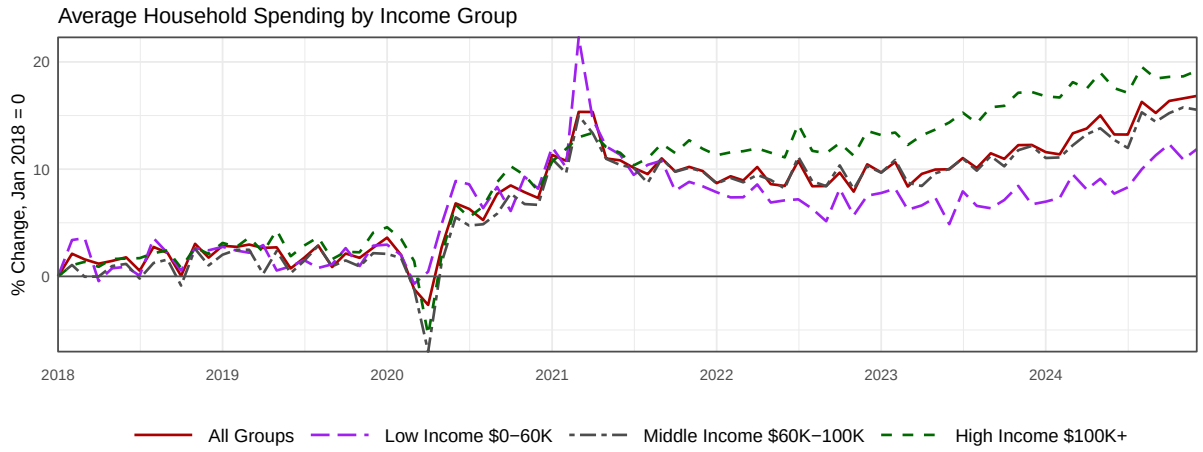
3 Spending by reported income

This section reports our findings on average retail spending by income. Using the categorical income groups recorded in the Numerator data, we can decompose retail spending, the blue line in Figure 1, into spending by difference income groups and investigate how spending differs between low, middle, and high-income households. We then provide additional evidence for the validity of the divergence in spending between high- vs low-income households. We first show that spending decomposed by education level (as a rough proxy for income) depicts the same divergence. Second, we present a simulation exercise for spending in 2024 taking into account movements across income groups based on the survey results we obtain from a subset of the panelists in the Numerator data.

3.1 Average household spending by income

Figure 2 reports growth since January 2018 in average monthly household spending, adjusted for inflation and seasonality, overall and by three income groups: low ($0-$60K in annual household income), middle ($60K-$100K), and high ($100K+).[8]

FIGURE 2: Growth of average retail spending decomposed by household income



Note: The data are monthly from January 2018 through August 2024. All series are adjusted for inflation using the chain indexed PCE deflator for goods and food services excluding motor vehicles and are shown as growth relative to January 2018. All series are seasonally adjusted using X13-ARIMA-SEATS.

[8]We select the cutoffs for low-, middle-, and high-income groups based on the following Census Bureau report that is based on 2022 incomes: https://www.census.gov/library/publications/2023/demo/p60-279.html. Minor modifications to these income groupings do not change our qualitative results.

The growth in average household retail spending is similar across all income groups in the pre-pandemic era. In the onset of the pandemic, all households' spending declines. Various studies, e.g. Cox et al. (2020), show that spending declined due to the direct effects of pandemic, rather than loss of income due to any disruptions in the labor market. Following the decline in the beginning of the pandemic, average household spending for all income groups bounces back, reaching levels higher than the pre-pandemic levels.

During the pandemic recovery period and shortly after, low-income households appear to benefit more from pandemic-era stimulus. In percentage terms, low-income households increase their spending by more than middle and high-income households in 2020 and early 2021. The large jump in the average spending of low-income households in early 2021 coincides with the last stimulus program which provided more stimulus, as a percentage of their income, to low-income individuals than others. The finding that high- and middle-income households did not spending too much of the last stimulus checks and low-income households' spending increased more after the last stimulus check in March 2021 is consistent with Chetty et al. (2023)'s findings.

Despite similar pre-pandemic and early-pandemic era trends, the spending behavior of low and high-income households starts to diverge in mid-2021. Between mid-2021 and mid-2023, middle and high-income households maintain or increase their real average spending; low-income households reduce their real average spending compared with mid-2021 levels. Since mid-2023, low, middle, and high-income households have all been increasing their real average spending. As of August 2024, real average spending by low-income households is up 7.9% relative to January 2018, real average spending by middle-income households is up 13.3%, and real average spending by high-income households is up 16.7%.

While our analysis does not offer evidence for the differences in spending we observe for low-, middle-, and high-income households, we offer a few hypotheses for why these differences have emerged. First, low-income households depleted pandemic-era excess savings earlier as documented by Abdelrahman et al. (2024). Second, government support programs such as SNAP emergency allotments, enhanced unemployment insurance, and enhanced child tax credits that disproportionately benefited low-income households expired and no longer provide a boost to these households' spending. Third, high-income

households might experience a wealth effect as their homes and investments increase in value, while they also receive more interest and investment income during periods of higher interest rates, all providing a stimulus for sustained levels of spending.

Note that our analysis and inflation adjustment likely under-counts the extent to which low-income households experienced inflation. Cavallo and Kryvtsov (2024) document that the prices of products low-income households tend to purchase increased at a faster pace. If there is heterogeneity in households' inflation experience, with low-income households experiencing higher inflation, low-income households' spending after the appropriate inflation adjustment would be lower than what our analysis has shown.

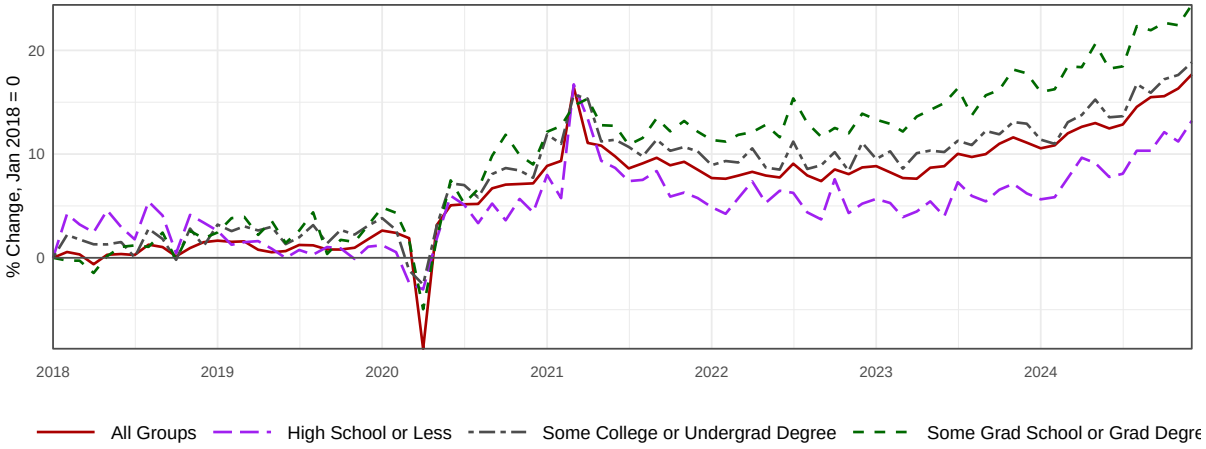
3.2 Divergence in Spending: Spending by Education

To support our results on the divergence of spending across income groups, we provide additional evidence in this subsection and the next. First, Numerator data contain several different household attributes we could use to decompose retail spending. Among a few attributes that were explored in Hacioglu Hoke et al. (2024), one of them is education. One benefit of decomposing retail sales by education rather than income is that education remains broadly stable between 2018 and 2024, whereas household incomes increase with inflation. The stability of the education groups over time allows us to check whether our results based on household income might have just been driven by compositional changes in income groups (e.g., if more households become high-income over time and fewer households remain low-income, that affects the composition of households in each income group and their average monthly retail spending).

In Figure 3, we show average household retail spending by education. Education is based on the self-reported level of education of the primary shopper of a household in Numerator panel data. Our results for education are broadly similar to our results for household income, with households with lower education levels (high school completed or less) experiencing a pullback in real average household spending on retail goods and food services between mid-2021 and mid-2023, similar to households with lower income levels, while households with middle and higher education levels (some college or undergraduate degree and some graduate school or graduate degree) showing resilience and sustaining

higher levels of retail spending on goods and food services since mid-2021.

FIGURE 3: Growth of average retail spending decomposed by education



Note:The data are monthly from January 2018 through December 2024. All series are adjusted for inflation using the chain indexed PCE deflator for goods and food services excluding motor vehicles and are shown as growth relative to January 2018. All series are seasonally adjusted using X13-ARIMA-SEATS.

3.3 Divergence in Spending: Survey Evidence and Simulation

To overcome the rigidity in income groups we use over time, we conducted a survey on a smaller yet a representative set of Numerator panelists to gauge the compositional changes in income groups. We asked panelists their annual income in 2019 and 2024 as well as a multiple choice question of if they and their household earn more, less, or about the same income in 2024 compared to 2019. After the data cleaning steps outlined in Appendix B, the final data set had 5644 respondents' reliable nominal income information. We classify households as high-, middle- and low-income consistent with the categorization we use for Figure 2, and calculate the percentage of respondents who stay in the same income category or those that move up or down in income groups.

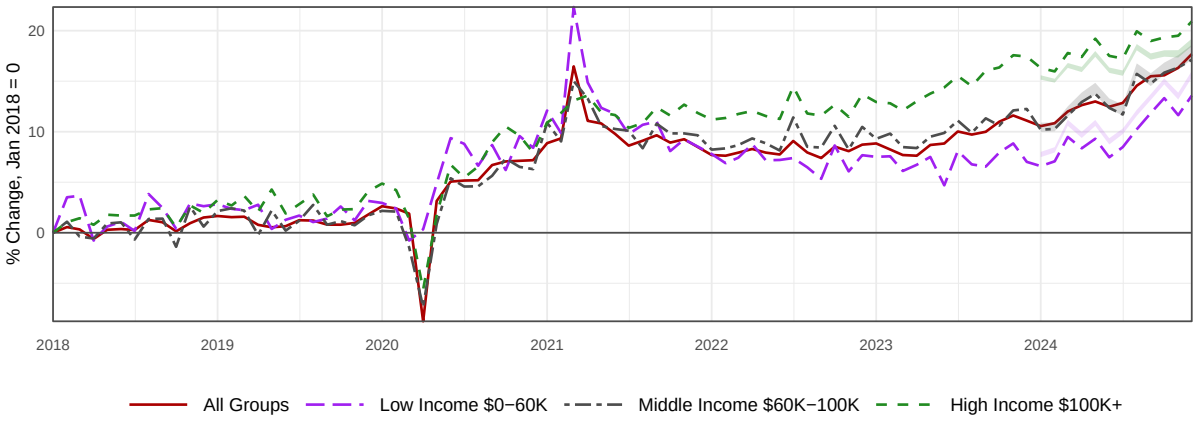
The share of respondents in each income group in 2019 and 2024 is reported in Table B.1. Most of the respondents stay in the bucket they are originally in but there is both upward and downward movement across buckets. If respondents move across buckets, the biggest moves are to the nearest bucket, not across multiple buckets. Maybe surprisingly, there is downward movement and we can infer to a couple of reasons using the commentary

they put in while filling out the survey. Some of these respondents report job losses since 2019. However, most of the downward movement is most often by older respondents, such as baby boomers, who retire during the pandemic period.

We proceed to design a simulation exercise to use the survey results reported in Table B.1 in combination with our results in Figure 2. The essence of this exercise is that we artificially move panelists in between income groups. Namely, we take each panelist's spending in 2024. For each month, we take e.g. a random 2.7% of high income panelists and change their income group to middle income based on the the percentage of panelists that moved from high to middle income groups, reported in Table B.1. We repeated this step for all other shares reported in the table from the leftover panelists. Once we move panelists to their 'new' income groups in 2024, we recalculate the average spending of low-, middle- and high-income households, which is seasonally and inflation adjusted. We repeat these steps 1,000 times which allows us to randomly sample and reshuffle panelists across income groups. We then calculate confidence bands by taking the range of the 10th and the 90th percentiles of spending across simulations.

In Figure 4, we overlay the confidence bands of the simulation exercise on our main results in Figure 2. Moving panelists across income group leads middle- and high-income households' average spending to decline, and low-income households' spending to increase which is expected. The noteworthy observation from the simulation results is the the shaded areas which show that the divergence between the low vs high income groups remains even if we were to move households across income buckets.

FIGURE 4: Growth of average retail spending decomposed by household income with simulated confidence bands based on the survey results



Note:The data are monthly from January 2018 through December 2024. All series are adjusted for inflation using the chain indexed PCE deflator for goods and food services excluding motor vehicles and are shown as growth relative to January 2018. All series are seasonally adjusted using X13-ARIMA-SEATS. The confidence bands show the range of the 10th and 90th percentile of spending across simulations that is designed based on the survey results reported in Table B.1.

4 Spending by zip code income

Numerator data gives us the advantage of exploring whether using the zip code-level income as a proxy for income leads to similar results. So in this section, we apply zip code level aggregation to Numerator's spending data to compare the results for spending by reported income in the previous section. The analysis of spending by zip code level income is a la Chetty et al. (2023).

Chetty et al. (2023) link zip codes of the card transactions to the median household income in zip codes reported in the 2014-2018 American Community Survey. They then split the sample between high- (household income greater than $78,000 per year), middle- (household income between $46,000 and $78,000 per year) and low-income households (household income less than $46,000 per year). The dollar values for annual income correspond to top, middle two and bottom quartiles of median household income for high-, middle- and low-income households, respectively. They discuss that although zip codes are a strong predictor of income in the US, they are not a perfect proxy for income, see also Chetty et al. (2020). Chetty et al. (2023) circumvent this issue by comparing their

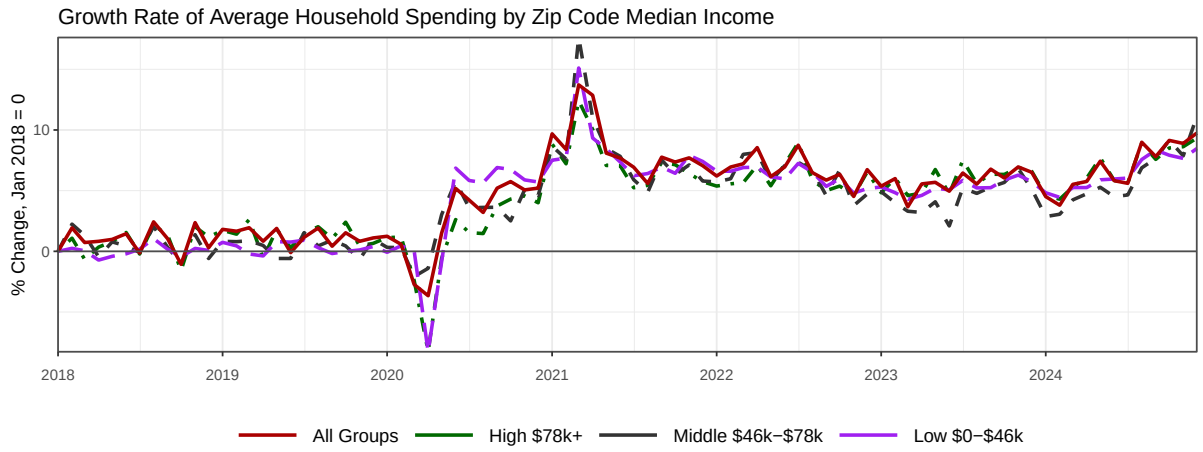
estimates with that of Cox et al. (2020) who analyze the initial impact of the pandemic using US household-level bank account data. However, Cox et al. (2020) report that their analysis is based on simulated income information from public use data due to the unavailability of micro income data at the time of writing.

To replicate a similar analysis using Numerator data, we aggregate households' spending by the income groups that are given by the zip code level income. We map Numerator panelists' zip codes to the real median household income over the preceding 12 months from 2014 to 2018 in zip codes reported in the 2014-2018 American Community Survey (ACS). The ACS publishes household income by 5-digit ZIP Code Tabulation Area (ZCTA). We use a crosswalk to match 5-digit zip codes to ZCTA.[9] For consistency with Chetty et al. (2023)'s analysis, we define the income groups in a similar way, i.e. high-(household income greater than $\$78,000$ per year), middle-(household income between $\$46,000$ and $\$78,000$ per year) and low-income households (household income less than $\$46,000$ per year). Appendix C also compares Numerator spending series based on zip code aggregation with Chetty et al. (2023)'s.

Figure 5 reports the average household spending when we zip code aggregation instead of using panelists' reported income. The spending patterns are similar to those reported in Figure 2 until roughly mid-2021, including the pre-pandemic period. The final stimulus check in March 2021 leads to a bigger jump for low-income households. However, the divergence starting in mid-2021 between low- and high-income households evident in the analysis that uses reported income information disappears at the zip code level. Note that the income group categorization we use in the previous section is only slightly different than Chetty et al. (2023)'s and reassuringly the findings in this section do not change when we use the same income groups as in the previous section, as Figure C.2 shows.

[9]The crosswalk from ZCTA to zip codes is from UDS Mapper. Source: `https://www.graham-center.org/maps-data-tools/uds-mapper.html`

FIGURE 5: Average household spending by zip code



Note: The data are monthly from January 2018 through December 2024. All series are adjusted for inflation using the chain indexed PCE deflator for goods and food services excluding motor vehicles and are in percent changes compared to January 2018. All series are seasonally adjusted using X13-ARIMA. Purple, grey and green lines plotting the spending by low-, middle- and high-income households.

Table D.1 shows the discrepancy between the zip code implied income versus Numerator income groups. The values in the third column are the share of households living in the low-, mid-, high-income zip codes (reported in the first column) who reported the income groups in the second column in Numerator data. In other words, the third column shows the error that zip code-based income grouping makes for each group when the first and second column do not match. The 'correct' shares are in green shades. Based on Table D.1, 70% of low income households live in low income zip codes whereas 22% that live in low income zip codes is actually is middle income households, and 8% is high-income households. The difference between zip code versus reported income becomes more stark for middle and high-income zip codes. 56% of households that live in middle income zip codes is low income households based on their reported income group and only 30% is actually middle income. Similarly, only 29% of households that live in high-income zip codes report high income in Numerator data, while 71% is low and middle-income.

The scenarios that lead to such a discrepancy are not difficult to come by. Households with low or middle income such as retirees or those that inherit houses from their parents can live in high-income zip codes. Also households with middle to high-income can live in low-income areas where house are more affordable with the prospect of these areas to

be gentrified in the near future. In summary, survey based zip code level income masks a fair bit of heterogeneity which is important to account for in understanding consumers' spending behavior. Moreover, the ACS we use, also used in Chetty et al. (2023), reports median zip code income for the 2014-2018 period. The characteristics of US households, and the landscape in rural and urban areas have changed significantly since then. All of these factors we discuss in this section collectively point out that zip code level income is far from being a good proxy for household income and hence for gauging heterogeneity in consumer behavior, in our case spending.

5 Conclusion

The analysis in this paper has two objectives. First, we construct novel spending measures for low-, middle- and high-income households using a micro-level household retail spending data from 150,000 representative US households. We validate our overall measure using official published statistics. We regularly update our measure to track the state of the US consumer. Second, and more importantly, our analysis attempts to document the discrepancy in using reported household groups versus zip code implied household income levels. The latter masks the heterogeneity in spending that arises after the peak spending in the post-pandemic period in 2021, after which low-income households' spending slows down while high-income households drive the strength in spending. This analysis calls for caution in using zip code level income as a proxy for household income and highlights the importance of using alternative datasets with micro-level information for macroeconomic results in economics.

References

Abdelrahman, H., L. E. Oliveira, and A. Shapiro (2024). Pandemic-era liquid wealth is running dry.

Brodeur, A., D. Gray, A. Islam, and S. Bhuiyan (2021). A literature review of the economics of covid-19. *Journal of Economic Surveys 35*(4), 1007–1044.

Buajitti, E., S. Chiodo, and L. C. Rosella (2020). Agreement between area- and individual-level income measures in a population-based cohort: Implications for population health research. *SSM - Population Health 10*, 100553.

Cascaldi-Garcia, D. (2022, August). Pandemic Priors. International Finance Discussion Papers 1352, Board of Governors of the Federal Reserve System (U.S.).

Cavallo, A. and O. Kryvtsov (2024). Price discounts and cheapflation during the post-pandemic inflation surge. *Journal of Monetary Economics 148*, 103644. Inflation in the COVID Era and Beyond.

Chetty, R., J. N. Friedman, E. Saez, N. Turner, and D. Yagan (2020, 02). Income segregation and intergenerational mobility across colleges in the united states*. *The Quarterly Journal of Economics 135*(3), 1567–1633.

Chetty, R., J. N. Friedman, M. Stepner, and O. I. Team (2023, 10). The Economic Impacts of COVID-19: Evidence from a New Public Database Built Using Private Sector Data*. *The Quarterly Journal of Economics 139*(2), 829–889.

Cox, N., P. Ganong, P. Noel, J. Vavra, A. Wong, D. Farrell, F. Greig, and E. Deadman (2020). Initial impacts of the pandemic on consumer behavior: Evidence from linked income, spending, and savings data. *Brookings Papers on Economic Activity*, 35–69.

Geronimus, A. T., J. Bound, and L. J. Neidert (1996). On the validity of using census geocode characteristics to proxy individual socioeconomic characteristics. *Journal of the American Statistical Association 91*(434), 529–537.

Hacioglu Hoke, S., L. Feler, and J. Chylak (2024, October). A better way of understanding the us consumer: Decomposing retail sales by household income. FEDS Notes. Washington: Board of Governors of the Federal Reserve System.

Hanley, G. and S. Morgan (2008). On the validity of area-based income measures to proxy household income. *BMC Health Services Research 8*(79).

Vavra, J. (2021, August). Tracking the pandemic in real time: Administrative micro data in business cycles enters the spotlight. *Journal of Economic Perspectives 35*(3), 47–66.

Appendix: Not For Publication

A Benchmarking Numerator to MARTS

Numerator retail spending data allow us to decompose aggregate retail spending to 12 categories reported in the US Census Bureau's Monthly Sales for Retail and Food Services (MARTS) report. These categories are spending in 3-digit North American Industry Classification System (NAICS) level for the sectors listed in Table A.1. Hacioglu Hoke et al. (2024) establishes the reliability of Numerator's retail spending measure.According to Census Bureau 'MARTS covers firms classified in the Retail Trade and Food Services sectors as defined by the North American Industry Classification System (NAICS). Retail Trade, as defined by NAICS sectors 44-45, includes establishments engaged in selling merchandise in small quantities to the general public, without transformation, and rendering services incidental to the sale of merchandise. Two principal types of establishments classified in retail trade can be distinguished: i) Store retailers operate fixed point-of-sale locations, located and designed to attract a high volume of walk-in customers. They have extensive displays of merchandise, use mass-media advertising to attract customers and typically sell merchandise to the general public for personal or household use. Some store retailers also provide after-sales services, such as repair and installation; for example, new automobile dealers; ii) Nonstore retailers also serve the general public, but their retailing methods differ. Such methods include paper and electronic catalogs, door-to-door solicitation, in-home demonstration, "infomercials,' selling from portable stalls or through vending machines. Food services, as defined by NAICS subsector 722, include establishments that prepare meals, snacks, and beverages to customer order for immediate on-premises and off-premises consumption.'. Figures A.1 through A.3 to shows the month-over-month changes of Numerator and MARTS series for the 12 spending categories reported in Table A.1 and reports their correlations.

TABLE A.1: Spending categories at 3 digit NAICS level

NAICS	Spending Type
442	Furniture & home furniture stores
443	Electronics & appliance stores
444	Building material & garden equipment & supplies dealers
445	Food and beverage stores
446	Health & personal care stores
447	Gasoline stations
448	Clothing & clothing accessories stores
451	Sporting goods, hobby, musical instrument & book stores
452	General merchandise stores
453	Miscellaneous store retailers
454	Non-store retailers
722	Food services & drinking places

FIGURE A.1: Comparing Numerator spending categories with MARTS

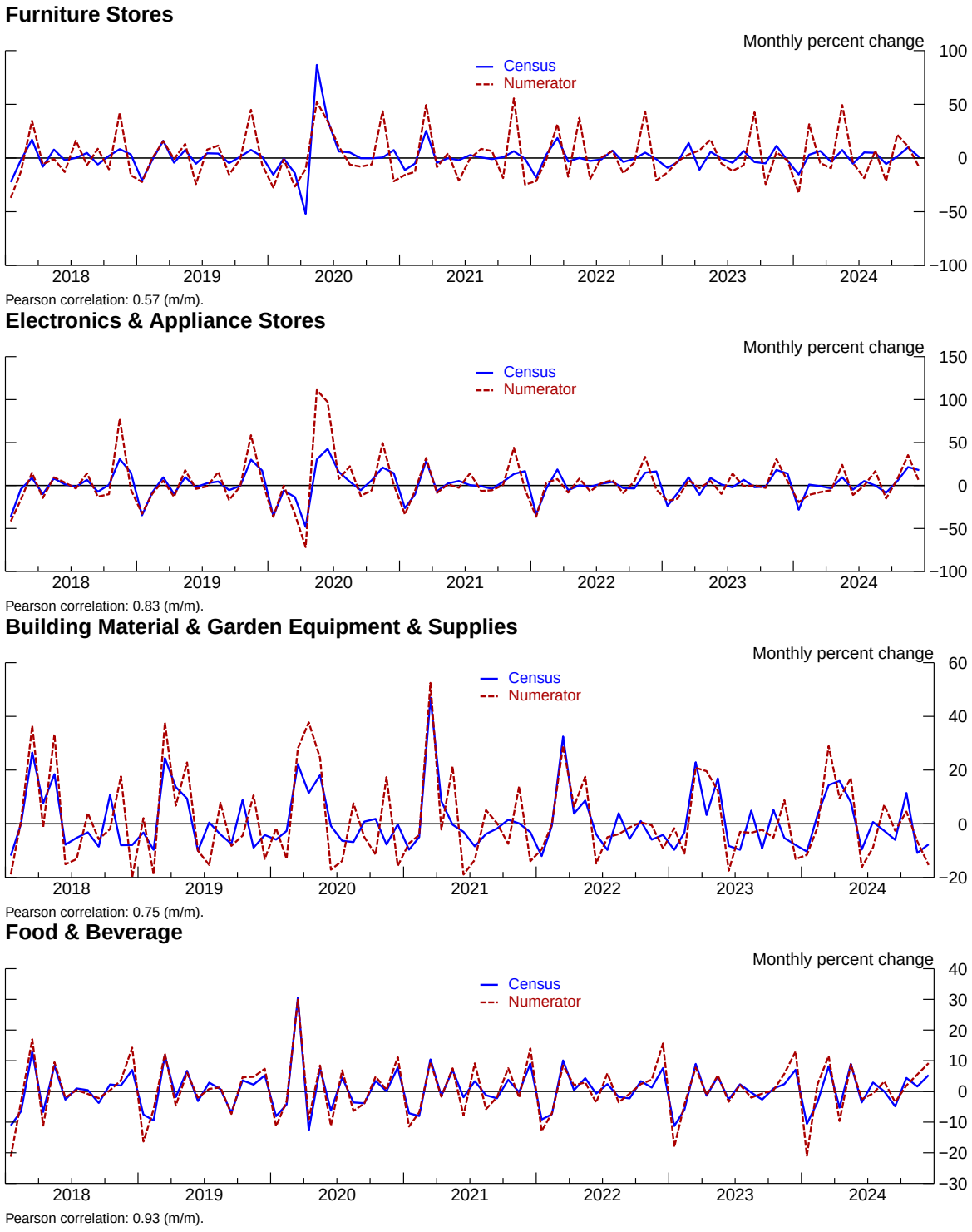


FIGURE A.2: Comparing Numerator spending categories with MARTS

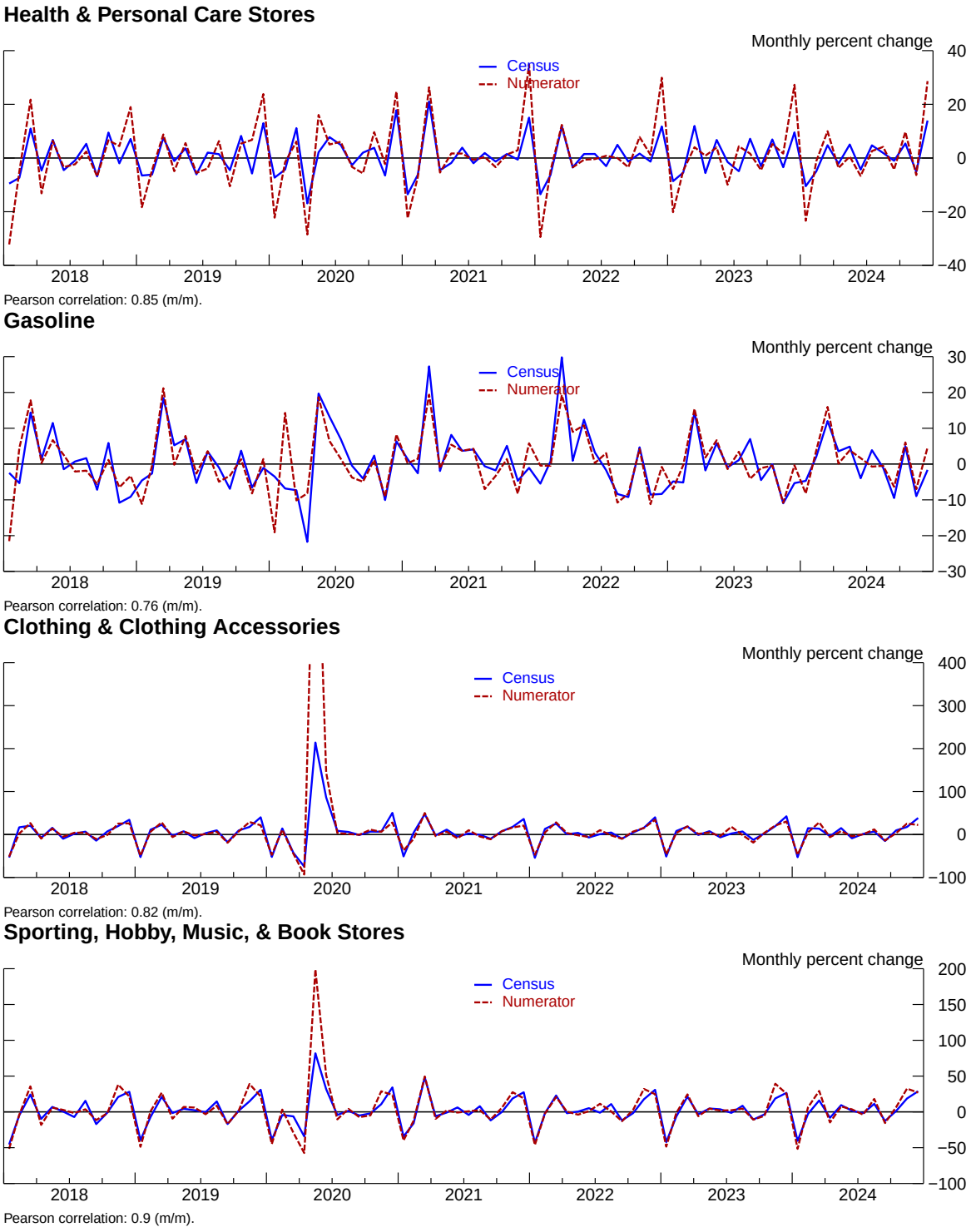
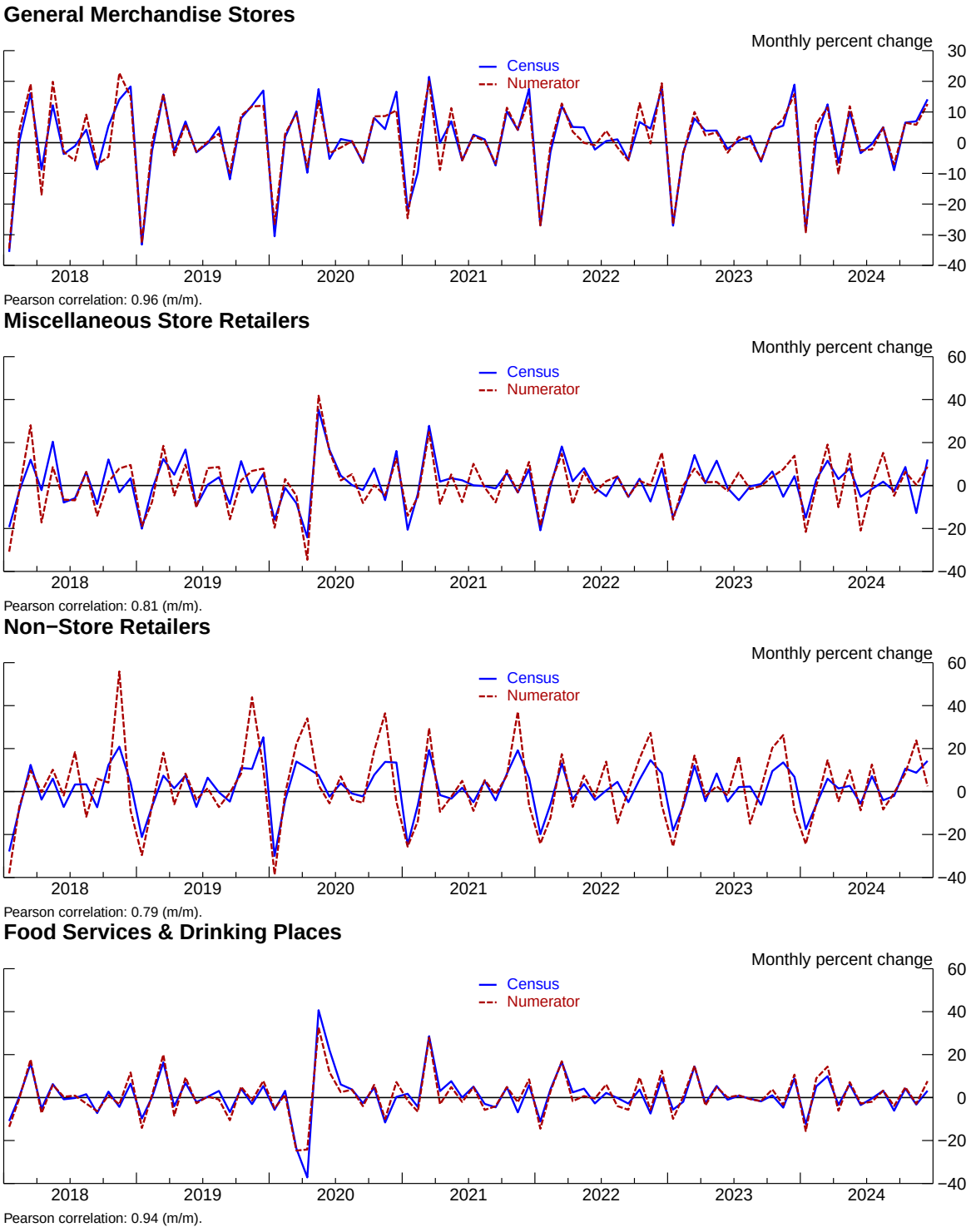


FIGURE A.3: Comparing Numerator spending categories with MARTS



B Survey Data and Results

We conducted a survey among Numerator panelists that would help us address any concerns related to the construction of the income buckets in Section 3. We ran a test survey on 100 people between October 25 and 28 to gauge the appropriateness of the questions and the validity of respondents' answers. The first wave of the survey was rolled out to random 10,000 Numerator panelists on October 28, and 4600 panelists responded by November 3. The second wave started on 10,000 different Numerator panelists on November 12 and concluded on November 19 with 4618 responses. Overall, we have 9218 respondents. The survey consisted of 13 questions and a commentary box. In the survey design, we incorporated questions to understand the change in income from 2019 to 2024. Three questions from the survey that help us in our analysis are as follows.

1. *Do you and your household earn more, less, or about the same income now compared with 2019?* Respondents choice from three options: 'We earn more now'; 'We earn about the same now'; 'We earn less now'.

2. *About how much do you and your household earn now per year?*

3. *About how much did you and your household earn per year in 2019?*

Among the 9218 responses we received, there were inconsistent entries of income which are outlined with the examples below. Once we clean the data for these issues, the final dataset has 5644 respondents' reliable income information alongside all the other household characteristics Numerator panel data records on these households.

- **Missing zeros:** Reporting \$150,000 for 2024 income but \$147 for 2019 income. Similarly if 2024 and 2019 income are e.g. \$50 and \$48. As long as the 2024 income provided by the panelists is consistent with the categorical bin information Numerator records for panelists, and their qualitative answer is also consistent with their entry, we corrected for the missing zeros.

- **Missing zero:** Reporting \$28,500 for 2024 income but \$2,400 for 2019. As long as the 2024 income provided by the panelists is consistent with the categorical bin

information Numerator records for panelists, and their qualitative answer is also consistent with their entry, we corrected for the missing zero.

- **Reporting an additional zero:** Reporting $1,000,000 for 2024 income while reporting $76,000 for 2019 income. As long as the 2024 income provided by the panelists is consistent with the categorical bin information Numerator records for panelists, and their qualitative answer is also consistent with their entry, we corrected for the additional zero.

- **Inconsistent incomes:** If either 2019 or 2024 income is reported, and the other income is missing but the percent change in income is provided with the info on if they earn more, less or the same now, I applied that percent change to their reported income to fill in the missing income. I did this in very rare occasions when all the other responses made sense.

- **Erroneous or misleading entries:** If 2024 and 2019 incomes are inconsistent or they do not seem to reflect respondents' approximate income levels (e.g. reporting $123,456 or reporting $3,000 for 2019 then $85,000 for 2024), and information on how much they earn now, and income bucket do not give us a way to identify their incomes, we eliminated these users.

- **Outliers:** After all these cleaning steps, we look for outliers in percent change in income between 2019 and 2024 and eliminate survey respondents who are below the 5% and above the 95% percentiles. We eliminate survey respondents that report earning more in 2024 with negative percent change in nominal income, and similarly eliminate respondents that report earning less in 2024 with positive percent change in nominal income. For respondents that report earning the same, we eliminate those whose income changes are anything but in 0-5% range to account for the small changes that might have affected these households' income which they neglected to report in the qualitative question of earning more, less or the same.

Table B.1 presents the share of respondents in each income group in 2019 vs 2024. For example, 90.5% of the respondents who were high income households in 2019 remained high income in 2024 while 19.5% of middle-income households moved to high income in

2024. Expectedly, most of the respondents stay in the income group they are originally in but there is both upward and downward movement across buckets. If respondents move across buckets, the biggest moves are to the nearest bucket, not across multiple buckets. Maybe surprisingly, there is downward movement, and we can infer to a couple of reasons using the commentary they put in while filling out the survey. Some of these respondents report job losses since 2019. However, most of the downward movement is most often by older respondents, such as baby boomers, who retired during the pandemic period.

TABLE B.1: Shares of respondents in each income group in 2019 versus 2024

2019 \ 2024	High Income	Middle Income	Low Income
High Income	90.5	9.5	0
Middle Income	19.5	70.0	10.5
Low Income	<0.1	9.4	90.5

Notes: Share of respondents in each income group in 2019 vs 2024 for three income groups: low ($0-$60K in annual household income), middle ($60K-$100K), and high ($100K+). Information on annual income in 2019 and 2024 collected by the custom survey run by Numerator in October and November 2024. Final data set has 5644 respondents' reliable nominal income information.All the VARs include the term premium, industrial production, consumer price index, unemployment rate and average household spending by low- or middle- of high-income households. Industrial production, consumer price index and average household spending enter the VAR in log levels and rates enter in levels. The chart shows the responses of different income groups overlapped from the results of separate VARs. The VAR includes pandemic dummies from March through August 2020, in the spirit of Cascaldi-Garcia (2022).

Note that younger respondents, such as those in early or mid 20s in 2024 record big jumps in their incomes over the five year horizon during which we explore the change in income using the survey. This is because these respondents probably got their first jobs upon graduating high school or university between 2019 and 2024. To address the big shift in their income bins, we check if our results look similar when we exclude Gen Z. In this case the dataset comprises 5510 respondents. The differences in the results are negligible.

C Replicating Chetty et al. (2023)

Replication documentation of the Economic Tracker as introduced by Chetty et al. (2023) can be found in `https://github.com/OpportunityInsights/EconomicTracker/tree/`

main/docs. In a nutshell, Economic Tracker uses aggregated and anonymized purchase data from consumer credit and debit card spending from Affinity Solutions which is available from 13 January 2020. Spending by consumer zip code income relies on matching the zip codes they live with the measured median household income provided by surveys. To replicate Chetty et al. (2023)'s analysis, we use Table S1903 from the 2018 American Community Survey, which contains the real median household income over the preceding 12 months from 2014 to 2018.[10]

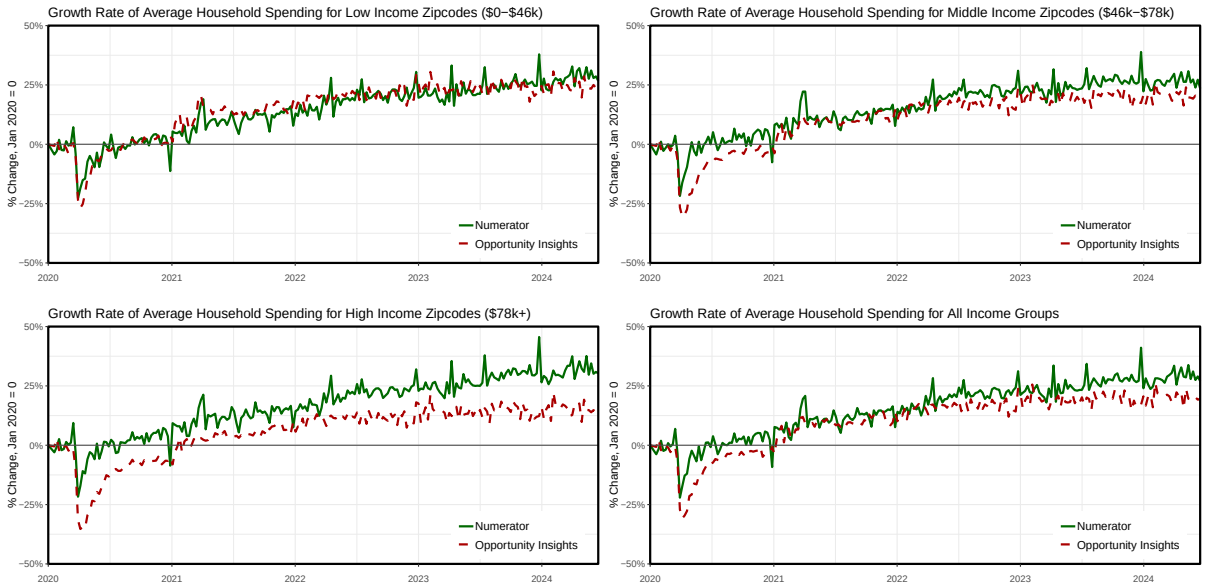
The American Community Survey publishes household income by 5-digit ZCTA instead of 5-digit zip code. We use a ZCTA to zip code crosswalk to identify the income levels that correspond to panelists' zip codes. The income groups in Chetty et al. (2023)'s analysis are slightly different from our original analysis. However, to replicate their results and compare how Numerator's zip code level income measures perform in comparison, we inherit their income group classification. That is,

- High Income (top quartile of median household income; approximately greater than $78,000 per year)
- Middle Income (middle two quartiles of median household income; approximately between $46,000 per year and $78,000 per year)
- Low Income (bottom quartile of median household income; approximately less than $46,000 per year)

The Economic Tracker is in weekly frequency until May 2024, so we construct weekly average spending measures using Numerator data through May 2024 to compare both spending spending measures. Figure C.1 shows weekly Numerator spending series and the Economic Tracker spending series for different income groups and for all income groups combined when households are grouped based on zip code income. Numerator and Economic Tracker series are very close for low-income households. There are differences for middle- and high-income households and overall spending which generally emerge in the post 2022 period. Numerator only captures retail sales whereas card spending should capture services in theory. The differences in spending, especially for high income

[10]The 5-year American Community Survey data cover all geographic units whereas the 1-year estimates only account for areas with at least 65,000 people.

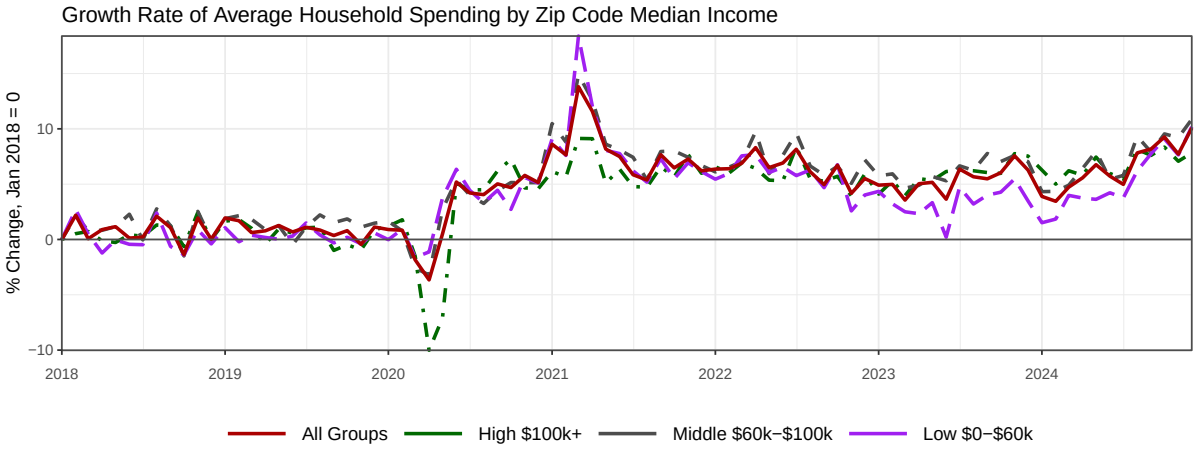
FIGURE C.1: Average household spending by zip code income



Note: The data are non-seasonally adjusted, weekly from January 2020 through December 2024 and presented in percent changes compared to January 2020 as in Chetty et al. (2023). Start from January 2020 due to the data availability of Opportunities Insight series. Red line indicates the percent change in spending implied by Opportunity Insights. Green line indicates the percent change in spending implied by Numerator. Top left: low-income zip codes ($0-$46K); top right: middle-income zip codes ($46K-$78K); bottom left: high-income zip codes ($78K+); bottom right: all income groups.

households might be due to the fact that services spending was more suppressed during Covid in Numerator data.

FIGURE C.2: Average household spending by zip code using the income classification in Section 3



Note: The data are monthly from January 2018 through December 2024. All series are adjusted for inflation using the chain indexed PCE deflator for goods and food services excluding motor vehicles and are in percent changes compared to January 2018. All series are seasonally adjusted using X13-ARIMA. Purple, grey and green lines plotting the spending by low-, middle- and high-income households.

D Reported versus zip code income discrepancy

Numerator data provides us income groups self-reported by households. Section 3 reports our analysis of average monthly household spending for three income groups: low ($0-$60K in annual household income), middle ($60K-$100K), and high ($100K+). We also have information on which zip codes they live in, which we use in our analysis in Section 4.

We classify households low-, middle- and high-income based on their zip codes using the median zip code income information in ACS as in Section 4 and also based on the reported income group. We calculate the shares of households that report low-, middle- and high-incomes which live in low-, middle- and high-income zip codes. Table D.1 reports these results. The first column specifies zip code income group and the second column shows the Numerator income group. Share reports the percent share of Numerator panelist the live in respective zip code income areas and report the respective incomes.

TABLE D.1: The discrepancy between the zip code implied income vs Numerator income groups

Zip code Income Group	Numerator Income Group	Share
Low	Low	0.70
	Middle	0.22
	High	0.08
Middle	Low	0.56
	Middle	0.30
	High	0.14
High	Low	0.39
	Middle	0.32
	High	0.29

Notes: The table calculates the percent share of Numerator panelists that live in low-, middle- and high-income zip codes in column 1 and reports the incomes in column 2. Three income groups: low ($0-$60K in annual household income), middle ($60K-$100K), and high ($100K+). Numerator income information updates periodically. The calculation in this table is as of end-2024. The green shades indicate when zip code income group matches Numerator income group, i.e. unshaded rows indicate the discrepancy between zip code level and reported income.