# The Use of Real Time Spending Sources for Economic Measurement: An Analysis Using Multiple Data Sources

**Authors** Justine Mallatt, U.S. Bureau of Economic Analysis[1]

**Contact** justine.mallatt@bea.gov

**Date** April 2025

**Abstract** The rise of big data in the past 20 years has opened up the possibility of using new types of experimental data to measure consumer spending. Despite the promise of card transaction data, concerns remain about their representativeness and stability. This study makes a novel contribution by integrating transaction data from multiple sources to measure consumer spending, rather than relying on a single vendor. By blending several card datasets, the analysis mitigates the noise or biases specific to any one source. Three core applications help assess the robustness of these new data sources: (1) evaluating their correlation with established national benchmarks, (2) testing their forecasting ability for official measures, and (3) comparing spending trends at finer geographic levels. In all three applications, I find evidence of the potential usefulness of these sources to provide meaningful economic signals. Additionally, I find evidence that these data sources are complementary, finding a stronger signal when multiple data sources are combined, relative to when a single source is applied. However, considering any single source at the national level, I find several of the data sources provide a similar economic signal, demonstrating that the data sources are potential substitutes.

**Keywords** Real time spending; card transaction data; economic measurement

**JEL Codes** E01, E21, E27

1. I would like to thank Abe Dunn for his invaluable input, as well as Lasanthi Fernando, Calvin Ackley, Kyle Hood, Brian Quistorff, and Mahsa Gholizadeh for their comments and Ben Bridgeman, Jon Samuels, Marina Gindelsky, Gary Cornwall, and other researchers in attendance at BEA's internal seminar series for their helpful suggestions. I'd like to thank John Friedman, Michael Stepner, and fellow researchers at Opportunity Insights for their comments and suggestions.

Bureau of Economic Analysis
U.S. DEPARTMENT OF COMMERCE

# 1 Introduction

Since the Great Recession and economic downturn and recovery stemming from the COVID–19 crisis, the need for near real-time measures of consumer spending and economic well-being have been of interest to the public and policymakers. The rise of big data in the past 20 years has opened up the possibility of using new types of experimental data to measure consumer spending. The lag of official statistics during the COVID–19 crisis prompted economists, statisticians, and firms to call upon alternative data sources in order to glean real-time information on the state of economic activity. While some of these data sources were direct measures of economic transactions, like credit and debit card transactions, ATM withdrawals, and unemployment claims, other metrics were more indirect, measuring economic activity through mobility data, online search queries, restaurant reservations through online booking platforms, and fuel sales, among many others.

Throughout the COVID–19 crisis, researchers and agencies have released many experimental real-time economic measurement projects. Chetty and coauthors at Opportunity Insights launched an interactive platform where users can track trends in spending and unemployment from credit and debit card transactions (from Affinity Solutions), job postings, and unemployment claims over time (Chetty et al., 2023). The U.S. Bureau of Economic Analysis (BEA) provided weekly near-real-time spending estimates for several categories of retail spending based on transaction patterns from financial payment processing systems (from Fiserv) (Dunn et al., 2021). The New York Fed Staff Nowcast updates its forecasts each week as new sets of less-frequent data with staggered release timing enter its model (Baker et al., 2023).

Despite the potential of credit and debit card data in producing timely estimates, concerns remain about their representativeness and stability (Abraham, 2022; Dunn et al., 2024). In particular, most studies rely on only one source of card transactions, which is a convenience sample rather than a statistically designed sample of households or firms (Aladangady et al., 2019; Chetty et al., 2023). Different datasets may aggregate information at either the consumer or the firm level, and may vary in coverage of online versus in-person spending, or debit versus credit usage. Combining multiple data sources not only offers broader coverage but also reduces the risk of over-reliance on a single, potentially unrepresentative dataset. Moreover, as private-sector firms can exit the market or alter their data-sharing policies, understanding both the complementarity and substitutability of these data sources is crucial for the consistent use of transaction data in official statistics.

This study makes a novel contribution by integrating transaction data from multiple sources to measure consumer spending, rather than relying on a single vendor. By blending several card datasets, the analysis mitigates the noise or biases specific to any one source. While all datasets contain debit and credit card transactions, some also include checks, direct deposits, direct payer-to-payer transfers, or other financial flows. This diversity offers a more comprehensive snapshot of economic activity, which in turn enables investigation of dataset performance and potential improvements in data quality. In addition, three core applications help assess the robustness of these new data sources: (1) evaluating their correlation with established national benchmarks, (2) testing their forecasting ability for official measures, and (3) analyzing spending trends and correlations at finer geographic levels.

I first check whether aggregate movements in card transaction data align with official estimates from the U.S. Census Bureau's Monthly Retail Trade Survey (MRTS) and Quarterly Services Survey (QSS). The results show a high correlation between national-level card transactions and these survey-based metrics. In fact, multiple card datasets can match or exceed the performance of Fiserv data, which the BEA has already employed to anticipate QSS signals for selected service-spending categories in personal consumption expenditures (PCE).

Next, I explore whether the timeliness of card data can improve forecasts of monthly PCE, which is typically released with a 3- to 4-week lag. Because transaction data are available within about 8 days, they have the potential to offer advanced signals of shifts in consumption. Empirically, I find that these data outperform simple autoregressive (AR) models and deliver near-parity accuracy compared to the routine revisions made to PCE during the COVID–19 economic downturn.

Finally, I investigate card transaction data at the state level. National accounts are supported by carefully designed surveys, but subnational spending estimates are produced without any regional spending data, as no survey exists at that level of geographic granularity. Instead, to estimate state-level PCE, BEA uses regional wage information by industry combined with national spending trends to form the state-level spending estimates. Moreover, due to lags in state PCE source data, the state-level PCE estimates are produced with a 6-month lag, relative to state GDP estimates.[1] Therefore, the incorporation of card transaction data at the regional level may be important for two reasons: (1) it would bring actual regional spending information into the estimates, rather than imputing with wage data; (2) it could improve the timeliness of the estimate. While no official "gold standard" survey-based benchmarks currently exist for state-level spending, strong correlations among the various sources provide some reassurance that the sources are capturing the same underlying economic signal, and may be substitutes for one another. In addition, I compare card-based measures to the experimental Monthly State Retail Sales (MSRS) from the Cen-

[1]BEA will begin releasing quarterly GDP and PCE by state in 2025.

sus Bureau that uses a combination of survey data, wage data, and alternative spending information. The comparison with MSRS offers an important external check on the card transaction data sources. I find robust correlations for populous states, suggesting that card data can provide timely and valuable insights into regional economic trends. I also find that combining across multiple transaction data sources tends to increase correlations with state MSRS, highlighting the complementarity of gathering information from a variety of sources.

Collectively, these findings reveal strong promise in using multiple card transaction datasets to capture consumption patterns. By cross-verifying correlations with official benchmarks, testing forecasting capabilities, and extending the analysis to finer geographic scales, this research demonstrates that transaction data obtained through private-sector firms can serve as both a complementary and, in some contexts, substitutable source of near-real-time consumer spending information.

## 2 Background

Despite the promise of big data, researchers have raised concerns about their inherent differences from information derived through official surveys. Official surveys are based on random sampling and weighted to ensure representativeness of the overall U.S. population. In contrast, private-sector big data originates from non-designed samples, reflecting firms' business needs rather than statistical priorities. While surveys are tailored to extract information directly relevant to economic measures, big data often centers on accurate financial and operational records. Data quality is a cornerstone of federal surveys, which face challenges like question design and non-response (Groves and Lyberg, 2010; Groves et al., 2011). Big data introduces novel quality concerns, with significant variation across sources (Aladangady et al., 2019; Kurmann et al., 2021). Additionally, big data structures may not align with official statistical needs, access is contingent on data providers, and samples can shift over time, complicating long-term analysis.

Nonetheless, big data offers several notable advantages over traditional survey methods, including greater timeliness, higher-frequency updates, finer granularity in geographic and industry coverage, and immense scope—spanning billions of transactions annually. These benefits became particularly evident during the COVID–19 pandemic, which disrupted traditional survey methodologies and underscored the urgent need for near-real-time economic measures. In response, U.S. governmental organizations and academic groups leveraged alternative data sources, such as credit and debit card transactions, smartphone location data tracking retail visits (Couture et al., 2022), and payroll data (Cajner et al., 2020), to monitor the economic effects of shutdowns and changes in consumer behavior.

Internationally, statistical agencies and central banks also turned to alternative data during the pandemic to track economic conditions. Norway's Norges Bank used debit card data to forecast household consumption (Aastveit et al., 2020), Central Bank of Ireland analyzed card spending and ATM withdrawals (Cronin and McInerney, 2022), and Chile's Banco Central De Chile relied on electronic invoicing data (Carlomagno et al., 2023).[2]

Historically, official surveys like the Monthly Retail Trade Survey (MRTS), the Economic Census, and trends in Quarterly Survey of Employment and Wages' (QCEW's) wages have been the foundation of near-real-time economic estimates. However, declining response rates to surveys over recent decades have raised concerns about their representativeness (Czajka and Beyler, 2016). Simultaneously, data users increasingly demand more granular and timely statistics (Abraham, 2022). Experimental indexes, such as those developed by the Census Bureau's IDEA team and Federal Reserve Banks, have begun to combine card transaction data with other sources to create detailed and timely insights into economic activity. These methods rely an amalgamation of a single source of debit and credit card spending with other measures of economic activity to measure economic activity more abstractly or from a broader lens (U.S. Census Bureau, 2024a; Federal Reserve Bank of Chicago, 2024; Lewis and Stock, 2024; Brave and Butters, 2014; Stock and Watson, 1999). These approaches often use a factor model to combine disparate sources of data into a signal of consumer sentiment. While these experimental indexes are valuable, it is not clear how these broad measures of economic activity relate to consumer spending measures. In contrast, my paper focuses explicitly on consumer expenditure trends, providing a more precise and literal representation of consumer spending behavior.

While these card transaction data sources are relatively new, the use of private data sources for improving economic measurement is not. Agencies like BEA and the U.S. Bureau of Labor Statistics (BLS) have integrated private-sector data sources to improve their estimates. For instance, BEA uses Ward's Automotive Reports to measure spending in the auto industry, IQVIA data to track pharmaceutical expenditures, and airline industry data to monitor air travel. Similarly, BLS has investigated the use of Nielsen scanner data to enhance the Consumer Price Index (CPI) for select product categories (FitzGerald and Shoemaker, 2013). Unlike these established private industry sources, the card transaction data studied here are relatively new and have many of the potential issues associated with big data sources. In addition, transaction data analyzed in this paper are also different because they are collected from the payment infrastructure supporting consumer spending, offering a broader, cross-industry perspective rather than being limited to one industry.

[2]See Tissot and De Beer (2020) for a more extensive overview on how various central banks used experimental data during the COVID–19 economic downturn.

# 3 Data

## 3.1 Alternative Data Sources

Several data providers have furnished spending transaction data to BEA. Fiserv (formerly known as First Data Merchant Services LLC) is a payment processor that collects data on millions of credit and debit card payments to merchants each day. Fiserv acts as an intermediary between merchants and banks, allowing for credit and debit cards to pay for goods and services at the point of purchase. Fiserv aggregates data covering millions of transactions across the US across merchant category codes, then merges merchant category codes to three-digit NAICS codes. The BEA received data files each week that contain daily transaction payment amounts by three-digit NAICS code and state.

Earnest Analytics produces the Orion data product, which is a consumer-based spending dataset. The data cover a convenience sample of millions of constantly contributing individuals. Financial institutions link financial accounts to households and track transactions into and out of each account, tracking households' electronic spending (debit cards, credit cards, and person-to-person payments through apps) as well as bank withdrawals. The Orion Consistent Shopper panel covers 7.1 million active accounts. About 23 percent of covered spending is through credit cards, and 77 percent is through debit card spending and account withdrawals. Each transaction is assigned to a specific merchant, and data is aggregated to the merchant level to protect the anonymity of contributing individuals. The data includes both in-person and online purchases. Merchants are classified into categories (similar to merchant category code) which can be matched to three-digit NAICS codes. The data is provided on Earnest Analytics' proprietary Dash web portal and downloaded at the state and merchant category level. Data comes in as fine as weekly frequency, and is produced with an 8-day lag. The Orion data cover January 2018 and forward.

Similarly, Earnest Analytics also provides the Vela Gamma spending series, which are sourced from microdata at the card level. The Vela Gamma cover 12.2 million active accounts. Vela Gamma is more credit-heavy, with 78 percent of purchase amounts from credit cards, and 22 percent from debit spending. This dataset also includes both online and in-person purchases. Vela Gamma covers about twice as much expenditure as Orion. Time coverage of the data begins in January 2016 and are produced at a weekly frequency with an 8-day lag.

TransUnion produces Commerce Signals, which tracks credit and debit card spending from banks and payment processors. The data cover over 55 million US individuals, and are anonymized such that any measurement output is aggregated data of at least 30 buyers. The number of accounts and spending at a ZIP Code level are scaled upwards to match the number of bank accounts at the ZIP level to make the data more representative. The data is available at fine geographic levels, by merchant category, at a weekly frequency, with about a 3-week lag. While I demonstrate that this data is correlated with official series of spending, its 3-week lag prevents it from aiding in further advanced forecasts of national PCE, although it is still informative for regional spending estimates where data sources are limited.

Card and transaction data from Affinity Solutions covers 140 million credit and debit cards and more than 8 billion transactions annually comprising at least $500 billion in annual spending. Their Consumer Purchase series is aggregated to protect privacy of individual consumers and households. The data is cleaned and corrected for exit and entry of data contributors by researchers at Opportunity Insights. I have access to data at the merchant category and state level. Time coverage of the data begins in January 2019; the frequency is daily.

For each of the alternative datasets—Fiserv, Orion, Vela Gamma, TransUnion, and Affinity—I match merchant spending categories to retail industries and some service industries at the NAICS-3 classification level. I collapse the experimental data to match the time frequency of comparison data; comparison to MRTS occurs at the monthly level, comparison to QSS occurs at the quarterly level, and forecasting PCE utilizes weekly data.

## 3.2 Official Data Series

BEA's monthly personal consumption expenditures (PCE) account tracks all personal spending on goods and services within the economy. The consumer spending category of economic activity captured in PCE makes up 70% of GDP. Data sources behind PCE include U.S. Census Bureau statistical reports such as the Census Bureau's Economic Censuses, Annual Retail Trade Surveys, Service Annual Surveys, Quarterly Services Reports, and Monthly Retail Trade Surveys, as well as the Bureau of Labor Statistics' Consumer Price Indexes. BEA also relies on reports from other government agencies, on administrative and regulatory agency reports, and on reports from private organizations such as trade associations.

PCE is released monthly during the third or fourth week of the month, with estimates for the previous month's personal spending and updates to prior months' estimates of spending amounts. PCE is detailed by product type, rather than producing industry. PCE can be disentangled into expenditure on goods and services. In this study, I assess the forecasting potential of card transaction data in predicting PCE spending on goods, as merchant categories for services are less clearly matched to PCE service type.

While credit and debit card transaction data has the potential to capture many forms of consumer spending, it does not capture cash transactions or large monthly outlays. For example, transaction data based on electronic card swipes will not capture rent or mortgage payments, or spending on utilities which are commonly paid for with direct withdrawals. Furthermore, PCE and transaction data categorize spending in different schema. PCE classifies spending for groups of goods and services, whereas transaction data details spending by merchant category or industry.

Census' Monthly Retail Trade Survey (MRTS) measures retail spending activity nationwide by month. Census' state-level MSRS is an experimental product that measures retail spending (excluding NAICS 722 restaurants) for each state. I compare movements in retail categories within the card data to the MRTS at the national and to the experimental MSRS at the state level. I restrict our sample to retail categories only when comparing to the MRTS and MSRS (i.e., NAICS 441-453 in the MRTS and MSRS and 722 in the MRTS). I exclude NAICS 454 (non-store retailers) because it includes web retailers and it is unclear how to match merchant categories in card data with this category.

Census estimates quarterly revenue for service industries in the United States, including information services, health care and social assistance, professional, scientific and technical services, and many more. The survey is conducted quarterly and calls upon a random sample of businesses in the Census Bureau's Business Registrar. Census then weights survey responses to be representative of the U.S. service sector as a whole. Although the QSS provides essential data to calculate PCE, the source survey is released with a lag of 75 days following the end of the quarter, and BEA previously used Fiserv data to produce advanced estimates of quarterly PCE before the QSS became available.

BEA releases the annual state PCE in September of the following year. The Economic Census is conducted every 5 years and serves as the data spine for state PCE. In years between waves of the Economic Census survey, the BLS QCEW is used to approximate trends in expenditure receipts for industries that sell products and services to households. Remaining spending categories use data from the American Community Survey, the U.S. Bureau of Transportation Statistics, and the Centers for Medicare & Medicaid Services (U.S. Bureau of Economic Analysis, 2024). Because the state PCE is released with a months-long lag and at the annual frequency, I compare state-level measures from card transaction data to the MSRS, available at the monthly frequency.

Census calculates the MSRS as an experimental product to measure monthly retail spending at the state level. Depending on state and NAICS-3 industry, they calculate the change in spending by taking a variance-weighted average between a "top down" estimate and a "bottom up" estimate. The top-down estimate uses national-level industry spending and multiplies it by each state's share of annual wages for that industry that year. The bottom-up approach uses receipts data from Circana and Nielsen point-of-sale data in combination with MRTS establishments set in the state. Missing data for MRTS establishments and third-party data are imputed using Bayesian methods (U.S. Census Bureau, 2024b).[3] The MSRS is produced with about a 3-month lag.

# 4 Methods

This paper conducts three separate analysis, and there are three associated methodology sections. As described previously, the first set of analysis focuses on correlations of card series with official series, including both MRTS and QSS. The second section describes how the series provide expedited forecasts of PCE, as well as information evaluating the performance of those forecasts. The third section describes how these alternative data sources yield state-level estimates.

## 4.1 Correlations of Card Series and Official Series

I examine whether the card transaction data series vary similarly to spending from official sources, such as in the MRTS and QSS over time. I combine year-over-year measures of spending growth at the month and quarter level to compare to year-over-year growth in the MRTS and QSS. I calculate Pearson correlations that capture the co-variances between movements in spending measures of the same industries across data sources over time.

For this analysis, I developed crosswalks that link Earnest merchant categories to NAICS service categories based on two approaches: (1) the definition of each NAICS category, and (2) the Merchant Category Code (MCC code) definitions from a pre-existing MCC-to-NAICS crosswalk used by Fiserv. Depending on the approach, different Earnest merchant categories are mapped to each NAICS industry. The "Earnest Orion" estimate relies on NAICS definitions to select Earnest merchant categories, while the "Vela MCC" estimate applies Fiserv's MCC-to-NAICS crosswalk to match MCC categories to Earnest's merchant classification.

### 4.1.1 Assessing Timely Estimates of MRTS

For the MRTS correlations, the spending levels in data are first seasonally adjusted using Census' X-13ARIMA SEATS model, which removes seasonal components from monthly

[3]Since state, county, and MSA disaggregate measures of spending lack a "gold standard" outside of the Economic Census, the MSRS is calculated in a way that balances and minimizes variance and noise.

spending.[4] De-seasonalized monthly year-over-year series are then calculated.

For each monthly observation of growth from each of the five transaction datasets, I construct amalgamated measures of growth by taking the simple mean of the sample of growth rates across card transaction datasets but within an industry/month pair, and taking the median of the growth rates across card datasets within each industry/month pair. I then chain these average or median growth rates together across months to form a series of average monthly growth rates and median monthly growth rates.

Figure 1 graphs monthly year-over-year growth in national retail and food services spending for the average series, median series, and for spending in each of the five separate series: Fiserv, Earnest Orion, Earnest Vela Gamma, TransUnion and Affinity. This figure demonstrates the volatility of the separate, singular transaction data series in measuring MRTS-like retail categories, and shows the effects of combining the individual series into their amalgamated mean and median signals. While any one of the contributing transaction data series may be volatile, taking the median or average of the growth rate in each month creates a smoother series.

#### 4.1.2 Assessing Timely Estimates of QSS

Due to the production lag of QSS, the QSS data are not available at the time of the initial estimate of GDP. To predict certain services categories, BEA has previously used card transaction data from Fiserv as an input to produce the advanced estimates of spending in select service categories, as listed in Table 2. These initial estimates are subsequently updated in the second or third revision of the quarterly release, at which point the QSS replaced the advanced estimates informed by Fiserv estimates.

To construct alternative estimates of year-over-year spending growth at the quarterly frequency, I use data from Fiserv, Earnest Orion, and Earnest Vela. I use Census' X-13ARIMA SEATS model to remove seasonality from the quarterly data similarly to how I use it for the MRTS. The Earnest datasets classify transactions by merchant category, which is similar to the MCC code used in other card datasets, though not identical. The BEA provided Fiserv with a crosswalk to match MCCs to NAICS categories, and Fiserv delivers the data at the NAICS level. The matching of Earnest's merchant categories to NAICS service codes is designed to match NAICS definitions and to match the existing MCC-to-NAICS match used for Fiserv.

[4] While year-over-year series removes some seasonality from the data series, Census' seasonal adjustment model removes additional calendar effects and trading day effects.

## 4.2 Relative Performance of Estimating Timely National PCE

Official estimates of PCE are derived in part from the MRTS and the QSS. Because these surveys take time to collect and process, the first public release of monthly PCE usually appears several weeks after the close of each month. By contrast, transaction data from four of the five card providers used in this study (Fiserv, Earnest Orion, Earnest Vela Gamma, and Affinity) become available with an approximately eight-day lag. In practice, this means that a complete record of weekly or daily spending for a given month is typically in hand by the first week of the following month, enabling researchers to estimate month-over-month (MoM) PCE growth roughly 2 weeks before the official "advance" estimate.

BEA routinely revises PCE estimates as more comprehensive data become available. An "advanced" estimate of quarterly PCE is published about 1 month after the quarter ends, relying on partial and preliminary inputs. A second estimate follows a month later, incorporating additional or improved source data. About 3 months after the quarter, a third estimate is released, incorporating still more complete information—including data from the QSS, which is available quarterly. As an example, early estimates for "Motor vehicle maintenance and repair" initially depended on high-frequency payment card transactions (Fiserv) before being replaced by QSS data in the third estimate. Many other spending categories are initially measured using "judgmental trends" until official data arrive.

Transaction data's primary advantage over survey-based measures is its timeliness, but its accuracy must be gauged. I do this by comparing card-based forecasts of MoM PCE growth to two benchmarks: an autoregressive model (AR(1)); and BEA's routine revisions to PCE. A simple, backward-looking AR(1) model predicts next month's PCE growth using the previous month's official growth rate and historical patterns. This serves as a rough baseline that approximates what forecasters may expect for the next month. BEA's initial PCE estimates often differ from later revisions once more complete information arrives. The magnitude of these revisions indicates how much month-over-month growth figures can shift within the official estimation process itself. These two benchmarks demonstrate different aspects of forecast performance. While the AR(1) model shows how well a simple method based solely on past data might perform, the revision-based benchmark demonstrates the inherent initial variance in the official estimates themselves.

To implement the AR(1) forecast, I generate a rolling 1-month-ahead prediction of PCE growth. For example, an analyst forecasting March 2020 growth at the end of March 2020 would have: The *first* estimate of February 2020 PCE; the *second* estimate of January 2020 PCE; and the *third* estimate of December 2019 PCE. With these data points (and all months before them), the analyst estimates the AR(1) relationship, then applies it to the first estimate of February 2020 to forecast March 2020 growth. Each subsequent month, the same procedure is repeated, updating the model's parameters as new data come in. Because the AR(1) model relies exclusively on past PCE values, it can become less reliable during sudden economic disruptions.

Even within BEA's official process, PCE estimates for a given month undergo multiple rounds of revision—often three to five times before the quarter's third estimate is finalized. To measure the size of these revisions, I compare each monthly estimate $PCEEstimate_{t,r}$ (for revision $r \in \{1, 2, 3, 4\}$) with the fifth estimate, $PCEEstimate_{t,5}$, via:

$$PercentageRevision_{t,r} = \frac{PCEEstimate_{t,r}}{PCEEstimate_{t,5}} - 1 \tag{1}$$

Table 1 presents descriptive statistics on these revisions for national PCE spending on goods between January 2019 and December 2023. On average, the first estimate under- or overstates the eventual fifth estimate by about 0.3 percentage points, with a standard deviation of 0.9 percentage points. The most pronounced underestimation observed was 3.521 percentage points during the early 2020 COVID–19 shutdowns, while the largest overestimation reached 0.938 percentage points. Revisions tend to shrink with each new release, reflecting the integration of more complete source data.

To compare forecasts of month-over-month PCE growth, I convert the first and fifth estimates of PCE levels into growth rates:

$$MOM_{1st,t} = \frac{PCEEstimate_{1st,t}}{PCEEstimate_{2nd,t-1}} - 1 \tag{2}$$

Where $PCEEstimate_{1st,t}$ is the first estimate for month $t$ and $PCEEstimate_{2nd,t-1}$ is the second estimate for month $t-1$. The final or "true" growth rate is:

$$MOM_{5th,t} = \frac{PCEEstimate_{5th,t}}{PCEEstimate_{5th,t-1}} - 1 \tag{3}$$

I define the *revision benchmark* as the difference between these two:

$$Benchmark_t = MOM_{1st,t} - MOM_{5th,t} \tag{4}$$

which indicates how far the initial estimate of growth diverges from the final, fifth estimate.

To generate comparable MoM growth estimates from transaction data, I use weekly national spending series. Because conventional seasonal adjustment methods (such as the Census Bureau's ARIMA-13 SEATS) are not designed for weekly data, I first regress *log(weekly spending)* on dummies for major holidays (e.g., Christmas, Thanksgiving, Independence Day) and month fixed effects, then extract and exponentiate the residuals.

Next, I sum these holiday-adjusted and de-meaned weekly values for each calendar month. If a month contains five Saturdays, it will encompass five weekly observations, which I normalize to four weeks by multiplying by four-fifths. Dividing this month's adjusted spending sum by the previous month's sum yields an implied month-over-month growth rate. Finally, I compare this transaction-based growth rate to the *fifth* estimate of PCE growth, calculating absolute forecast errors to gauge the accuracy of the card-data-derived forecasts relative to official figures.

Each month, the BEA releases "Table 2.3.5U. Personal Consumption Expenditures by Major Type of Product and by Major Function," which details monthly PCE on goods and services. I investigate the transaction data's ability to forecast 28 different series of spending by product from this table. I implement a random forest model, which combines decision tree regressions and bootstrapping. Across a number of iterations set by the researcher, the random forest selects a random subsection of data and a random set of the researcher-provided covariates (set of card dataset growth rates) and runs decision tree regressions to predict the outcome variable (PCE growth in a product category). Each decision tree regression is calibrated to minimize the error in predicting the outcome variable. Coefficients associated with covariates across the iterations are then averaged.

Machine learning is useful in a situation where a researcher has many explanatory variables and many dependent variables to forecast, and the relationship between the predictors and variable of interest are not clear. For example, BEA produces PCE by product type, and my card transaction data classifies spending at the establishment industry level. Instead of hand-picking which of the card series (Affinity, Fiserv, Orion, Vela, the average or the median) will predict spending on each of 28 product categories and subcategories, I can have the random forest decide which of the card series are best at predicting the product series, and use those models to forecast out-of-sample PCE growth.

Similarly to forecasting PCE using a simple AR(1) model, I run the random forest on each of 28 product categories and subcategories using only the lagged PCE depending on the information set available at the time of forecast. I then run random forest models using the lagged value and the mean of all past values available at the time, and then on the lagged value and the median and average card series, and then on the lagged value and all separate card series.

## 4.3 State-Level Estimates

While there do not exist official series of spending at the state level for which to compare transaction data to, I correlate the transaction data series with movements in the experimental, state-level version of the MRTS, the MSRS. The state-level MSRS is based on a blended method combining regional information available in MRTS, alternative private-sector source data, and imputed values based on wage data. While the MSRS is also based on alternative data sources, they are constructed independently of the other data sources analyzed, offering a useful check on the other series.

I correlate the movements in the MSRS and card transaction data within each state over time. Resulting correlation coefficients are listed by state in Table 9, and the distribution of these correlations across fifty states are plotted in Figure 15 for visual comparison of state correlations across transaction datasets. I also provide the map-like Figure 18 that plots the difference over time between the MSRS series and the median card transaction series estimates for each state. In addition to comparing individual series to MSRS, I also examine how the average and the median across the different series correlate with MSRS.

# 5 Results

## 5.1 Correlation Results of Card Transaction Sources and Official Sources

### 5.1.1 MRTS and Retail Industries

Table 3 lists correlation coefficients between the MRTS monthly series and series of spending growth from transaction data sources. Columns (1)–(5) correlate spending growth in the MRTS with spending growth in the Fiserv, Earnest Orion, Earnest Vela Gamma, TransUnion, and Affinity series, respectively. Columns (6) and (7) correlate the MRTS growth in spending with chained growth rate series constructed by combining the five separate transaction datasets. Table 3's rows correspond to different retail industries; the first row lists correlation coefficients for total spending for retail and fast food industry codes, excluding non-store retailers. The second row and subsequent rows detail correlation coefficients for NAICS-3 retail industries (NAICS 441–453) and for in-person food and beverage service establishments (NAICS 722).

Overall, spending growth gleaned from transaction data is positively correlated with spending growth in the MRTS. The exception to this pattern appears to be NAICS 443 (Appliance and Electronic Stores). In many cases, using the median or average of growth rates across the contributing transaction series on spending leads to higher correlations with the MRTS.

Figures 2-14 plot similar correlation scatter plots and growth rates over time at the national level for each of the following NAICS industries; All retail and food service industries, 441–448 (retail), 451–453 (retail), and 722 (food and drinks). For all tested retail industries with the exception of 443 (appliances and electronic stores), growth rates in card data spending are highly correlated to the MRTS. These results suggest that card transaction data can be dis-aggregated by industry to provide timely information for spending growth for industry-specific measures.

### 5.1.2 QSS and Service Industries

BEA previously relied on Fiserv to provide early signals of growth in spending for service industries listed in Table 2. Table 4 shows the correlation coefficients between the Quarterly Services Survey (QSS) and quarterly changes in various card datasets, including Fiserv, Earnest Orion, Earnest Orion with an alternate MCC-to-NAICS mapping, and Earnest Vela.

Fiserv was an input to calculate advanced estimates of PCE for service industries listed in Table 2.[5] Fiserv is more correlated to QSS than the Earnest series and TransUnion for three out of the eight service industries. The Vela and VelaMCC series are more highly correlated with the QSS for the remaining five series. The combined average series is not reliably correlated with the QSS, as signals from the Orion and TransUnion series are impacting the average. However, the median series seems to produce correlations similar to Fiserv's or in some cases higher correlations to the QSS than Fiserv. Similar to correlations with the MRTS in the appliances and electronics stores retail category, the card data is not strongly correlated with movements in spending for industries 624 (social assistance) or for 813 (religious, grantmaking, civic and similar).

For several of these series, I find that multiple transaction data sources provide positive correlation with the associated service categories. This indicates that these series may be substitutes.

## 5.2 Results for Forecasting PCE

Table 6 lists the mean absolute error (MAE) associated with benchmark AR(1) one-step-ahead forecasting models and the revision magnitude between the 1st and 5th monthly PCE estimates. Columns (3)–(8) list the mean absolute error (MAE) in predicting PCE

[5]Due to budget considerations, the use of card transaction data for this purpose has been discontinued.

month-over-month growth with each card dataset separately—the Fiserv, Orion, Vela, and Affinity—and with the average growth series and median growth series. The second row of the table lists the mean absolute error scaled to the AR(1) model's mean absolute error. Predicting growth in PCE using card data leads to lower forecast errors than using the AR(1) model. These estimates would be available 2 to 3 weeks earlier than the official first PCE estimate. These estimates demonstrate the value of the more timely signal. For instance, the MAE for the benchmark AR(1) is 0.027, but the MAE of 0.018 from Fiserv is lower than this simple prediction model. The reduced MAE indicates the information gains from the availability of alternative data. Similar improvements are seen for the other card transaction series. As with the correlation results in Table 3, combining signals of separate card datasets by using means and medians has the potential to decrease forecasting errors as compared to forecasts from individual transaction series. Revision magnitudes are smaller in absolute magnitude than the errors of forecasting models, and revision amounts tend to be around 0.20 times the absolute error associated with the benchmark AR model.

Once the month's official first PCE estimate is available in weeks 3 or 4 of the following month, that first estimate is more accurate than the AR benchmark model and as well as simple forecasts using transaction data. That is, the simple forecasts based on transaction data perform worse and are not a replacement for the official estimates. The primary value-added of the transaction data is that it is available earlier than the first PCE estimate.

The lower panel of Table 6 lists mean absolute errors associated with forecasting March 2020 through May 2020 PCE movements, based on models estimated from prior months. The shocks to markets associated with COVID–19 and state lockdowns during the spring of 2020 present a situation where current data would be expected to perform far better than simple, backwards-looking models like the AR(1) benchmark. The mean absolute errors during the tumultuous time period are much lower for forecasts obtained from card transaction data than from the AR(1) model, as predicted. During the spring of 2020, the median card data series leads to a forecasting error that is about equivalent to the revisions to the first monthly PCE estimates for the months in this time period.

The relatively good performance of card data forecasts during the spring of 2020 merits further investigation. Figures 16 and 17 graph the error and absolute errors associated with forecasts using the median of the card series signal, in comparison with the forecast errors associated with the AR(1) forecast model and the standard PCE revision magnitudes. Figure 16 plots the errors of forecasts that could occur in the first week of following month. For example, researchers and policymakers investigating how COVID–19 impacted economic activity in April 2020 would have 4 weeks of April's card transaction data in the first week of May 2020. The analyst would have the first PCE estimate from March 2020 and the second PCE estimate from February 2020 to calculate lagged month-over-month growth. The analyst would use this lagged growth to predict what the month-over-month growth would have been between March 2020 and April 2020. The first estimates for PCE in April 2020 would only become available at the end of May 2020. The errors in the graph demonstrate that a backwards-looking model performs quite poorly in times of economic volatility; the card median forecast performs just about as well during the spring of 2020 as it does for the rest of the covered time periods. In more typical periods of economic activity like in late 2021 through mid 2023, the backwards-looking AR(1) forecast and card median forecast perform comparably well to the card data, with both series demonstrating absolute errors ranging between 0 and 0.03 (missing the true growth rate by 3 percentage points).

The graphs in Figure 17 plot the magnitude of the revision of PCE and the same forecast errors associated with the median card data forecasts in the previous graph. Note the figures are on a smaller scale. The card data forecasts are similarly noisy across the entire time period between 2019 and late 2023, whereas PCE estimates underwent relatively large revisions in April 2020, towards the end of 2020, and around the first quarter of 2021. The card transaction data series' regular forecast error volatility is about on par with PCE's larger revisions. The forecast errors are notably larger than the PCE revision magnitudes to month-over-month growth after the start of 2021. This may be because economic activity resorted back to more "normal" times after COVID–19 vaccines became available in the first half of 2021.

The transaction data appear to experience uniform variance in their forecast errors associated with predictions for PCE growth over time. This apparent uniform noisiness over time makes them perform similar to backwards-looking AR(1) models during ordinary times in the business cycle, and the errors in the card data's forecasts are quite a bit larger than the magnitude of routine revisions during stable times. However, during the COVID–19 shutdowns in the spring of 2020, backwards-looking models were rendered useless, and PCE's early predictions are more heavily reliant on judgmental trends before official data becomes available. In contrast, the timely transaction data that would be considered fairly noisy during normal times were suddenly relatively useful and available quickly. From these figures, I conclude that a blend of card transaction data can produce timely signals of spending during volatile times where PCE may require larger subsequent revisions, but forecasts from card data are relatively noisy during more typical economic times.

7 lists mean absolute errors (MAEs) associated with the random forest models. Columns (1)–(3) list the table line of the associated product spending category or subcategory corresponding to BEA's Table 2.3.5U, the category description, and the category code name. Each row of the table is the PCE category to be forecast (the dependent variable). Columns

(4)–(7) list the MAE for models with various sets of potential independent/predicting variables. Column (4) lists the MAE associated with models predicting PCE using only lagged PCE. Column (5) lists MAEs when using lagged PCE and the simple mean of past PCE growth. Column (6) lists MAEs using lagged PCE and the median and average growth rates among transaction series. Column (7) lists MAEs using lagged PCE and all card transaction series of growth in retail spending – Affinity, Fiserv, Orion, Vela, the median series, and the average series. For most PCE categories, forecast errors decrease when all card data series can be used to forecast PCE spending growth. Column (8) lists the amount that the MAE shrinks between a the model using only lagged PCE (Column (4)) and the model using all card series (Column (7)).

## 5.3 State-Level Correlations

Data users are demanding increasingly granular and timely statistics (Abraham, 2022; Abraham et al., 2021). BEA currently produces PCE at the state level, and Census produces the Monthly State Retail Sales (MSRS) series as a more geographically granular version of the MRTS. However, these official experimental series are produced with significant lags. In this section, I detail how the experimental MSRS correlates with transaction data series at the state and month level.

The Table 9 lists the mean of the standard errors accompanying the MSRS for each state and the coverage of the retail sector in the MSRS for each state, then lists Pearson correlation coefficients for correlations between each state's estimated growth in retail sales in the MSRS with growth in transaction spending series. Column (2) of Table 9 highlights that coverage of the retail sector by state varies. The column details the fraction of monthly observations within each state that have retail coverage below 10%; most state/month estimates are based on months with coverage of 10%–25% of the retail sector. States like Alaska, Michigan, and Oregon always have MSRS coverage of less than 10% of the retail sector for all observed months. The table demonstrates that transaction series are correlated with the MSRS. Larger states like California, New York, and Texas tend to have higher correlation coefficients, which is to be expected because larger states are likely to contribute more to transaction data and more likely to be more precisely measured in the MSRS. Consistent with correlations at the national level, combining state-level signals leads to more consistently positive correlations than using an individual transaction dataset. The median series in particular is always positively correlated with the MSRS in all states. These high correlations illustrate that transaction data has promise for estimating retail spending at the state level.

Figure 15 plots the distribution of Pearson correlation coefficients between the MSRS and each transaction spending series. Correlations are calculated within each of 50 states over time, and the resulting 50 correlation coefficients from each comparison are plotted as distributions in smoothed histograms. Correlation coefficients should cluster around +1.0 to signify similar patterns in series over time. The state coefficients from the median series are most clustered in the area between +0.5 and +1.0, with very little mass around 0.0. Figure 18 plots the difference between the MSRS monthly year-over-year growth estimate and the median transaction data signal measuring year-over-year growth for all retail within a state. This figure helps determine which states have more disagreement between experimental estimates; each state graph located in the figure plots the gap between the spending growth metrics over time. Some states have larger spikes around the spring of 2020; card data tends to have a growth rate higher than the MSRS. Subsequently in the spring of 2021, the year over year estimate from card data is lower than that of the MSRS because spring 2020 is in the denominator of the estimates. Some states like Texas, Louisiana, Mississippi have relatively little disagreement during the springs of 2020 and 2021. Broadly it appears that there is more agreement between estimates in Southern states and less agreement in the Mountain West states and in a few states in the Northeast.

# 6 Conclusion

In conclusion, the use of near-real-time credit and debit card spending data holds promise for measuring and predicting economic trends. This data can offer valuable insights into consumer spending and economic activity, particularly at more granular levels such as state and industry categories. The correlations observed between transaction data, the MRTS, and QSS suggest its potential as a tool for measuring fluctuations in consumer spending. This is particularly important for services where card transaction data may be used as an input for advanced estimates, as key input data from QSS arrives with a lag. The data also has important implications for regional PCE. In particular, this card transaction data may be used to help inform regional PCE estimates where currently no spending information is available.

Additionally, at the national level, credit and debit card data can serve as timely signals for nowcasting household consumption, prior to the availability of key PCE spending estimates. By using this data, policymakers and analysts can obtain more up-to-date assessments of consumer spending, enabling more informed decision-making and policy creation. Overall, the integration of credit and debit card data into economic analysis has the potential to enhance the accuracy and timeliness of economic indicators, contributing to improving economic measurement.

# References


Aastveit, K. A., T. M. Fastbø, E. Granziera, K. S. Paulsen, and K. N. Torstensen (2020). Nowcasting norwegian household consumption with debit card transaction data. *Journal of Applied Econometrics*.

Abraham, K. G. (2022, December). Big Data and Official Statistics. *Review of Income and Wealth 68*(4), 835–861.

Abraham, K. G., R. S. Jarmin, B. Moyer, and M. D. Shapiro (2021). *Big Data for Twenty-First-Century Economic Statistics*. University of Chicago Press.

Aladangady, A., S. Aron-Dine, W. Dunn, L. Feiveson, P. Lengermann, and C. Sahm (2019). From transactions data to economic statistics: Constructing real-time, high-frequency, geographic measures of consumer spending. Technical report, National Bureau of Economic Research.

Baker, K., M. Almuzara, H. O'Keeffe, and A. M. Sbordone (2023). Reintroducing the new york fed staff nowcast. Technical report, Federal Reserve Bank of New York.

Brave, S. A. and R. A. Butters (2014). Nowcasting using the chicago fed national activity index. *Economic perspectives 38*(1).

Cajner, T., L. D. Crane, R. Decker, A. Hamins-Puertolas, and C. J. Kurz (2020). Tracking labor market developments during the covid-19 pandemic: A preliminary assessment.

Carlomagno, G., N. Eterovic, and L. Hernández-Román (2023). *Disentangling Demand and Supply Inflation Shocks from Chilean Electronic Payment Data*. Banco Central de Chile.

Chetty, R., J. N. Friedman, M. Stepner, and the Opportunity Insights Team (2023, October). The Economic Impacts of COVID-19: Evidence from a New Public Database Built Using Private Sector Data. *The Quarterly Journal of Economics*.

Couture, V., J. I. Dingel, A. Green, J. Handbury, and K. R. Williams (2022). Jue insight: Measuring movement and social contact with smartphone data: a real-time application to covid-19. *Journal of Urban Economics 127*, 103328.

Cronin, D. and N. McInerney (2022). Atm cash withdrawals before, during and after the covid-19 pandemic. Technical report, Central Bank of Ireland.

Czajka, J. L. and A. Beyler (2016). Background paper declining response rates in federal surveys: Trends and implications. *Mathematica policy research 1*, 1–86.

Dunn, A., E. English, K. Hood, L. Mason, and B. Quistorff (2024). *Expanding the Frontier of Economic Statistics Using Big Data: A Case Study of Regional Employment.* US Department of Commerce, Bureau of Economic Analysis.

Dunn, A., K. Hood, A. Batch, and A. Driessen (2021). Measuring consumer spending using card transaction data: Lessons from the COVID-19 pandemic. In *AEA Papers and Proceedings*, Volume 111, pp. 321–25.

Federal Reserve Bank of Chicago (2024). Chicago fed national activity index (cfnai), current data. Updated November 25, 2024. Accessed on November 29, 2024.

FitzGerald, J. and O. Shoemaker (2013). Evaluating the consumer price index using nielsen's scanner data october 2013.

Groves, R. M., F. J. Fowler Jr, M. P. Couper, J. M. Lepkowski, E. Singer, and R. Tourangeau (2011). *Survey methodology.* John Wiley & Sons.

Groves, R. M. and L. Lyberg (2010). Total survey error: Past, present, and future. *Public opinion quarterly 74*(5), 849–879.

Kurmann, A., E. Lale, and L. Ta (2021). The impact of covid-19 on small business dynamics and employment: Real-time estimates with homebase data. *Available at SSRN 3954262*.

Lewis, Daniel J., M. K. and J. H. F. R. B. o. D. Stock (2024). Weekly economic index. Updated November 27, 2024. Accessed on November 29, 2024.

Stock, J. H. and M. W. Watson (1999). Forecasting inflation. *Journal of monetary economics 44*(2), 293–335.

Tissot, B. and B. De Beer (2020). *Implications of Covid-19 for official statistics: a central banking perspective.* Irving Fisher Committee on Central Bank Statistics, Bank for International ....

U.S. Bureau of Economic Analysis (2024). Personal consumption expenditures by state: Concepts, data and methods. Updated November, 2024. Accessed on December 17, 2024.

U.S. Census Bureau (2024a). Census bureau index of economic activity (idea). Updated November 27, 2024. Accessed on November 29, 2024.

U.S. Census Bureau (2024b). Monthly state retail sales technical documentation. Updated June 10, 2024. Accessed on December 17, 2024.

Table 1: Typical Revision Magnitudes Between Early and Fifth Estimate of Monthly PCE, Month-Over-Month Growth Rates at National/Month Level

| Comparison | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|
| Diff., Est 1 vs Est 5 | -.0033 | .00906 | -.03521 | .00938 |
| Diff., Est 2 vs Est 5 | -.0012 | .00519 | -.01511 | .0197 |
| Diff., Est 3 vs Est 5 | -.00075 | .00238 | -.01095 | .00159 |
| Diff., Est 4 vs Est 5 | -.00031 | .00132 | -.00791 | .00015 |

Notes. The table lists metrics associated with the relative difference between early monthly PCE estimates (estimates 1–4) and the fifth estimate obtained after official data sources are available.

Table 2: Description Of Industries Examined

| NAICS | Description |
|---|---|
| **MRTS Retail Categories** | |
| 441 | Motor Vehicle & Parts Dealers |
| 442 | Furniture & Home Furnishings Stores |
| 443 | Electronics & Appliance Stores |
| 444 | Building Material & Garden Equipment & Supplies Dealers |
| 445 | Food & 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 |
| 722 | Food Services |
| All Retail | 441-448, 451-453, & 722 |
| **QSS Services Categories** | |
| 541 | Professional, Scientific, & Technical Services |
| 561 | Administrative & Support Services |
| 624 | Social Assistance |
| 711 | Performing Arts, Spectator Sports, & Related Industries |
| 713 | Amusement, Gambling, & Recreation Industries |
| 811 | Repair & Maintenance |
| 813 | Religious, Grantmaking, Civic, Professional, & Similar Organizations |

Notes. The table lists the MRTS retail and QSS services industry categories of interest.

Table 3: Pearson Correlation Coefficients Between the Monthly Retail Trade Survey and Transaction Data

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| NAICS | Desc | Retail Share | Fiserv | Orion | Vela | TransU | Aff | Median | Average |
| 0 | All Retail Industries | 100 % | .792 | .704 | .877 | .631 | .942 | .966* | .964 |
| 441 | Motor Vehicle and Parts Dealers | 23 % | .847 | .84 | .836 | .879 | .918 | .935 | .954* |
| 442 | Furniture and Home Furnishings Stores | 2.1 % | .853 | .298 | .574 | .911* | .872 | .772 | .772 |
| 443 | Electronics and Appliance Stores | 1.5 % | .152 | -.445 | -.178 | .708* | -.219 | -.093 | -.095 |
| 444 | Building Material, Garden Equipment, Supplies | 7.5 % | .796* | .407 | .675 | .652 | .746 | .652 | .711 |
| 445 | Food and Beverage Stores | 14.7 % | .584 | .886 | .776 | .801 | .904* | .833 | .867 |
| 446 | Health and Personal Care Stores | 6.3 % | .654 | .86 | .905* | .761 | .481 | .849 | .869 |
| 447 | Gasoline Stations | 9.4 % | .974 | .93 | .954 | .977 | .986* | .975 | .972 |
| 448 | Clothing and Clothing Accessories Stores | 4.5 % | .932* | .776 | .828 | .842 | .887 | .88 | .882 |
| 451 | Sporting Goods, Hobby, Musical Instrument, Books | 1.5 % | .856 | .897 | .946* | .78 | .931 | .940 | .926 |
| 452 | General Merchandise Stores | 13.1 % | .117 | .584 | .821 | .823 | .841 | .918* | .837 |
| 453 | Miscellaneous Store Retailers | 2.5 % | .971* | .328 | .29 | .858 | .747 | .92 | .816 |
| 722 | Food Services | 13.9 % | .995* | .682 | .785 | .98 | .994* | .986 | .988 |

Notes. The table lists Pearson Correlation coefficients which capture the co-movement between series of spending data. A star * indicates the series with the highest correlation within an industry over time.

Table 4: Pearson Correlation Coefficients Between the Quarterly Services Survey and Transaction Data

| NAICS3 | NAICS Definition | (1) Fiserv | (2) Orion | (3) Vela | (4) VelaMCC | (5) TransUnion | (6) Average | (7) Median |
|---|---|---|---|---|---|---|---|---|
| 532 | Rental and Leasing Services | .748 | -.03 | .899* | .647 | .08 | .019 | .843 |
| 541 | Pro., Sci., and Tech. Services | .797* | .024 | -.02 | .561 | .02 | .02 | .358 |
| 561 | Administrative and Support Services | .408 | .008 | .805* | .522 | .049 | .015 | .742 |
| 624 | Social Assistance | -.23 | -.20 | .153* | -.23 | .06 | -.20 | -.24 |
| 711 | Performing Arts, Spectator Sports | .834* | .009 | .632 | .706 | -.01 | .063 | .746 |
| 713 | Amusement, Gambling, and Recreation | .826 | .807 | .72 | .847* | .167 | .503 | .819 |
| 811 | Repair and Maintenance | .702* | .217 | .632 | .648 | .18 | .217 | .692 |
| 813 | Religious, Grantmaking, Civic and Similar | -.3 | -.16 | .246* | .159 | -.18 | -.20 | -.13 |

Notes. The table lists Pearson Correlation coefficients which capture the co-movement between series of spending data. A star * indicates the series with the highest correlation within an industry over time.

Table 5: Summary Statistics For PCE Revisions, Month over Month Growth Rates

| | (1) Mean | (2) Std. Dev |
|---|---|---|
| Growth Fifth PCE Est | .0066 | (.035) |
| Growth First PCE Est | .0053 | (.0359) |
| Fiserv | .0054 | (.0431) |
| Orion | .0131 | (.031) |
| Vela | .0089 | (.0393) |
| Affinity | .0145 | (.0539) |
| Average | .0104 | (.0352) |
| Median | .011 | (.0378) |

Notes.

Table 6: Forecast Errors and Routine Revision Magnitudes

| | (1) Benchmark AR-1 | (2) Benchmark Fiserv | (3) Orion | (4) Vela | (5) Affinity | (6) Average | (7) Median | (8) Revision |
|---|---|---|---|---|---|---|---|---|
| Forecasting August 2019 through August 2023 | | | | | | | | |
| MAE | .027 | .018 | .019 | .017 | .023 | .015 | .015 | .006 |
| Relative MAE | 1 | .669 | .709 | .619 | .866 | .552 | .56600 | .209 |
| Forecasting March 2020 - May 2020 Only | | | | | | | | |
| MAE | .1 | .017 | .05 | .012 | .052 | .017 | .015 | .016 |
| Relative MAE | 1 | .175 | .499 | .122 | .52 | .168 | .149 | .159 |

Notes. Column (1) lists mean absolute error in predicting PCE with a one-step-ahead forecast using an AR-1 model. Column (8) lists the magnitude of four revisions between the 1st and 5th monthly PCE releases.

Table 7: Mean Absolute Errors of Machine Learning Model Forecasting PCE Categories

| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|
| Table Line | PCE Description | PCE Category | Lag 1 | Lag 1, Past Mean | Add Card Median, Avg | All Card | Perc Improve |
| 1 | Personal consumption expen | DPCERC | .0175 | .0194 | .014 | .0142 | -18.857% |
| 2 | Goods | DGDSRC | .0278 | .0283 | .0205 | .0203 | -26.978% |
| 3 | Durable goods | DDURRC | .0492 | .0468 | .0356 | .0351 | -28.659% |
| 4 | Motor vehicles and parts | DMOTRC | .0709 | .066 | .0583 | .058 | -18.195% |
| 5 | Furnishings and durable ho | DFDHRC | .0426 | .043 | .0322 | .0313 | -26.526% |
| 6 | Recreational goods and veh | DREQRC | .0352 | .0366 | .0303 | .0302 | -14.205% |
| 7 | Other durable goods | DODGRC | .0503 | .0566 | .0451 | .0454 | -9.7420% |
| 8 | Nondurable goods | DNDGRC | .0241 | .0251 | .0177 | .0171 | -29.046% |
| 9 | Food and beverages purchas | DFXARC | .0241 | .0259 | .0271 | .0253 | 4.979% |
| 10 | Clothing and footwear | DCLORC | .0572 | .0599 | .0478 | .0461 | -19.406% |
| 11 | Gasoline and other energy | DGOERC | .0811 | .0818 | .0746 | .0736 | -9.2479% |
| 12 | Other nondurable goods | DONGRC | .0162 | .0178 | .0137 | .0139 | -14.198% |
| 13 | Services | DSERRC | .014 | .0158 | .0125 | .0119 | -15% |
| 14 | Household consumption expend | DHCERC | .0164 | .0191 | .0144 | .0141 | -14.024% |
| 15 | Housing and utilities | DHUTRC | .0049 | .0048 | .004 | .0041 | -16.327% |
| 16 | Health care | DHLCRC | .0262 | .0302 | .0221 | .0212 | -19.084% |
| 17 | Transportation services | DTRSRC | .0505 | .0555 | .0471 | .0452 | -10.495% |
| 18 | Recreation services | DRCARC | .0587 | .0581 | .0505 | .0509 | -13.288% |
| 19 | Food services and accommod | DFSARC | .0515 | .0581 | .0454 | .0463 | -10.097% |
| 20 | Financial services and ins | DIFSRC | .0095 | .0082 | .0079 | .0083 | -12.632% |
| 21 | Other services | DOTSRC | .0206 | .0229 | .0194 | .0188 | -8.738% |
| 22 | Final consumption expenditur | DNPIRC | .0376 | .0435 | .037 | .036 | -4.255% |
| 23 | Gross output of nonprofit | DNPERC | .0073 | .0083 | .0069 | .0065 | -10.959% |
| 24 | Less: Receipts from sales | DNPSRC | .023 | .0291 | .0229 | .0222 | -3.478% |
| 25 | PCE excluding food and energ | DPCCRC | .0184 | .021 | .0151 | .015 | -18.478% |
| 26 | Energy goods and services | DNRGRC | .043 | .0433 | .0388 | .0389 | -9.535% |
| 27 | Market-based PCE | DPCMRC | .0198 | .022 | .0162 | .0164 | -17.172% |
| 28 | Market-based PCE excluding f | DPCXRC | .0212 | .0242 | .0181 | .0176 | -16.981% |

This table takes monthly PCE estimates from Table 2.3.5U. and runs a one-step-ahead forecast using transaction data. Mean Absolute Errors (MAE) are listed in columns (4)-(7). Percentage improvement of column (7) over column (4) is listed in column (8)

Table 8: Summary Statistics: State/Month MSRS and Card Data, 2019m1 - 2023m6

| Variable | (1) Mean | (2) Std. Dev | (3) Min | (4) Max | (5) 25th | (6) 50th | (7) 75th |
|---|---|---|---|---|---|---|---|
| Seasonally Adjusted Card Series, MSRS | | | | | | | |
| MSRS YoY | .061 | .11 | -.374 | .761 | .006 | .041 | .101 |
| Fiserv YoY | .047 | .092 | -.347 | .745 | -.003 | .035 | .084 |
| Orion YoY | .05 | .059 | -.129 | .292 | .011 | .045 | .081 |
| Vela YoY | .09 | .151 | -.383 | 1.155 | .012 | .062 | .124 |
| TransUnion YoY | .082 | .155 | -.59 | 1.664 | -.009 | .066 | .145 |
| Avg. Yoy | .068 | .078 | -.22 | .525 | .023 | .056 | .105 |
| Median Yoy | .056 | .066 | -.266 | .458 | .019 | .05 | .085 |
| Non Adjusted Card Series | | | | | | | |
| Fiserv YoY | .048 | .097 | -.362 | .795 | -.005 | .036 | .087 |
| Orion YoY | .051 | .061 | -.135 | .311 | .011 | .045 | .082 |
| Vela YoY | .091 | .153 | -.422 | 1.155 | .013 | .063 | .124 |
| TransUnion YoY | .066 | .146 | -.59 | 1.397 | -.007 | .062 | .128 |
| Avg. Yoy | .064 | .082 | -.226 | .54 | .021 | .053 | .095 |
| Median.Yoy | .057 | .075 | -.278 | .585 | .018 | .051 | .087 |
| Avg. Yoy | .064 | .082 | -.226 | .54 | .021 | .053 | .095 |
| Median.Yoy | .057 | .075 | -.278 | .585 | .018 | .051 | .087 |

Table 9: Correlation Coefficients: State Transaction Data and MSRS

| State | MSRS SE | Frac. Low Cov | Fiserv | Orion | Vela | TransUnion | Average | Median |
|---|---|---|---|---|---|---|---|---|
| AK | .072 | 1 | .583 | .213 | .738* | .234 | .61 | .715 |
| AL | .017 | .056 | .78 | .694 | .815 | .36 | .841 | .893* |
| AR | .022 | .037 | .643 | .729* | .036 | -.41 | .111 | .643 |
| AZ | .032 | .593 | .201 | .561 | .887* | .295 | .606 | .784 |
| CA | .009 | .167 | .808 | .754 | .891* | .38 | .794 | .888 |
| CO | .017 | .389 | .875 | .64 | .784 | .41 | .817 | .884* |
| CT | .021 | .019 | .747 | .644 | .815 | .431 | .849 | .854* |
| DE | .033 | .444 | .896 | .796 | .887 | .701 | .935 | .937* |
| FL | .011 | 0 | .869 | .689 | .908 | .52 | .874 | .922* |
| GA | .015 | .148 | .525 | .677 | .888* | .439 | .797 | .88 |
| HI | .177 | .741 | .787* | .687 | .426 | .504 | .694 | .784 |
| IA | .02 | 0 | .521 | .696 | .509 | .313 | .642 | .715* |
| ID | .032 | .87 | .807* | .47 | .687 | .512 | .742 | .8 |
| IL | .013 | .093 | .826 | .806 | .779 | .534 | .846 | .862* |
| IN | .016 | .796 | .807 | .655 | .412 | .391 | .756 | .839* |
| KS | .024 | .315 | .845* | .655 | .101 | .311 | .631 | .818 |
| KY | .018 | .481 | .665* | .651 | -.10 | .195 | .216 | .64 |
| LA | .017 | .019 | .68 | .741 | .541 | .518 | .805 | .866* |
| MA | .015 | 0 | .869* | .792 | .035 | .291 | .75 | .845 |
| MD | .018 | 0 | .842 | .802 | .886 | .79 | .918 | .922* |
| ME | .027 | 0 | .481 | .503 | .209 | -.01 | .356 | .509* |
| MI | .017 | 1 | .692* | .597 | -.20 | .586 | .473 | .648 |
| MN | .015 | .037 | -.25 | .556* | .5 | .147 | .253 | .355 |
| MO | .015 | .019 | .663 | .653 | .694 | .352 | .796 | .848* |
| MS | .022 | .074 | .662 | .767 | .77 | .321 | .812 | .889* |
| MT | .033 | .685 | .419 | .204 | .661* | .387 | .589 | .631 |
| NC | .013 | .037 | .806 | .662 | .792 | .2 | .773 | .858* |
| ND | .033 | 0 | .805 | .377 | .768 | .438 | .798 | .847* |
| NE | .03 | .074 | .714* | .593 | .062 | .257 | .54 | .691 |
| NH | .028 | 0 | .252 | .347 | -.32 | .615* | -.12 | .4 |
| NJ | .015 | .037 | .869 | .733 | .761 | .512 | .878 | .9* |
| NM | .026 | .074 | .367 | .493 | .736 | .455 | .755 | .803* |
| NV | .024 | .315 | .695 | .668 | .904* | .874 | .888 | .882 |
| NY | .013 | 0 | .826 | .815 | .838 | .476 | .827 | .843* |
| OH | .013 | .593 | .573 | .704* | .411 | .371 | .638 | .684 |
| OK | .019 | .037 | .728 | .698 | .262 | .434 | .774 | .826* |
| OR | .02 | 1 | .842 | .661 | .844 | .285 | .808 | .878* |
| PA | .011 | 0 | .78 | .766 | .852 | .432 | .858 | .888* |
| RI | .031 | .019 | -.05 | .518 | .6 | .548 | .644 | .688* |
| SC | .016 | .056 | .802 | .631 | .784 | .363 | .809 | .887* |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| SD | .037 | .185 | .815* | .647 | .484 | .444 | .779 | .792 |
| TN | .016 | .074 | .764 | .818 | .162 | .325 | .656 | .866* |
| TX | .01 | 0 | .722 | .721 | .847 | .407 | .783 | .886* |
| UT | .024 | .722 | .396 | .569 | .486 | .4 | .546 | .61* |
| VA | .014 | .056 | .802 | .707 | .183 | .796 | .872 | .899* |
| VT | .035 | .056 | .339 | .08 | -.33 | .385* | -.10 | .219 |
| WA | .017 | .741 | .867* | .493 | .6 | .717 | .785 | .832 |
| WI | .016 | .685 | .603 | .629 | .846* | .373 | .698 | .819 |
| WV | .022 | .278 | .495 | .497 | .522 | .482 | .679* | .627 |
| WY | .036 | .741 | .387 | .157 | .519 | .114 | .279 | .664* |

Notes: The table lists standard errors associated with the experimental MSRS estimates from each state in the first column. The second column contains the fraction of monthly MSRS observations within each state between January 2019 and June 2023 that report retail coverage of less than 10%; all other monthly observations have coverage between 10%-25% of retail spending in the state. Subsequent columns detail the Pearson correlation coefficients between the MSRS and each transaction data series at the state level. A star * indicates the column with the highest Pearson coefficient.

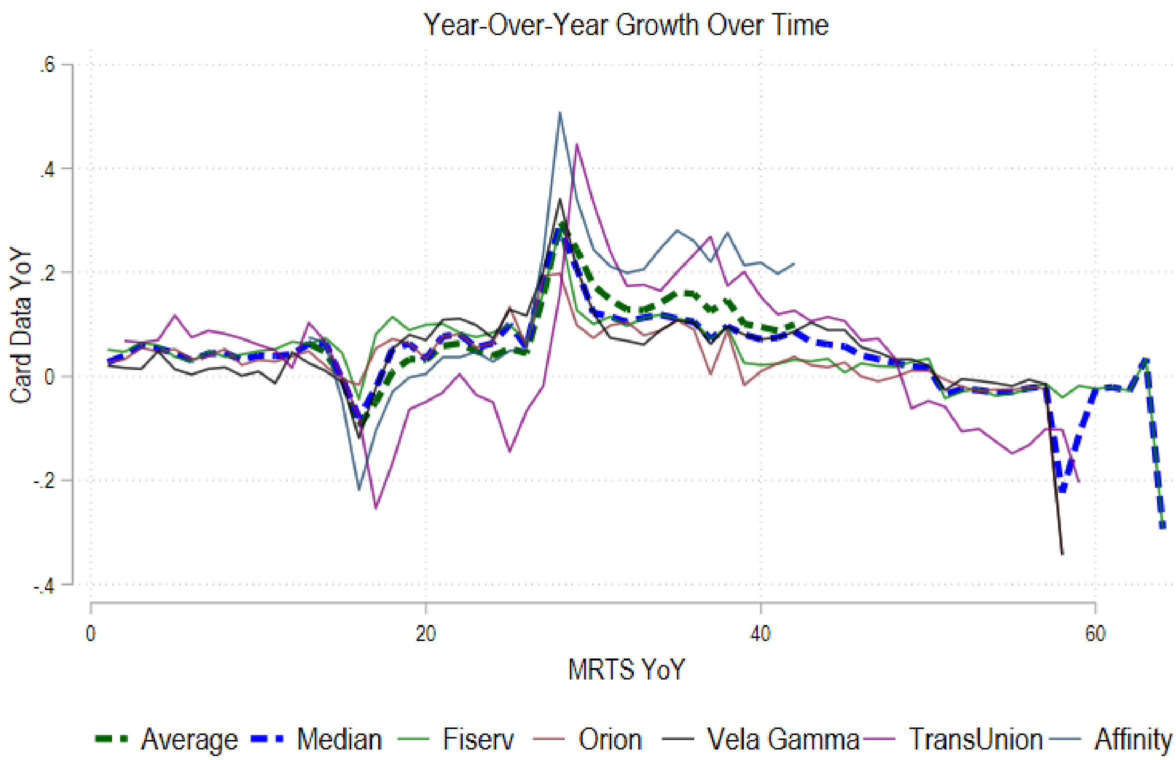


Figure 1: Growth Rates of Median, Average Growth Series and Five Transaction Series: Retail and Food Service Spending

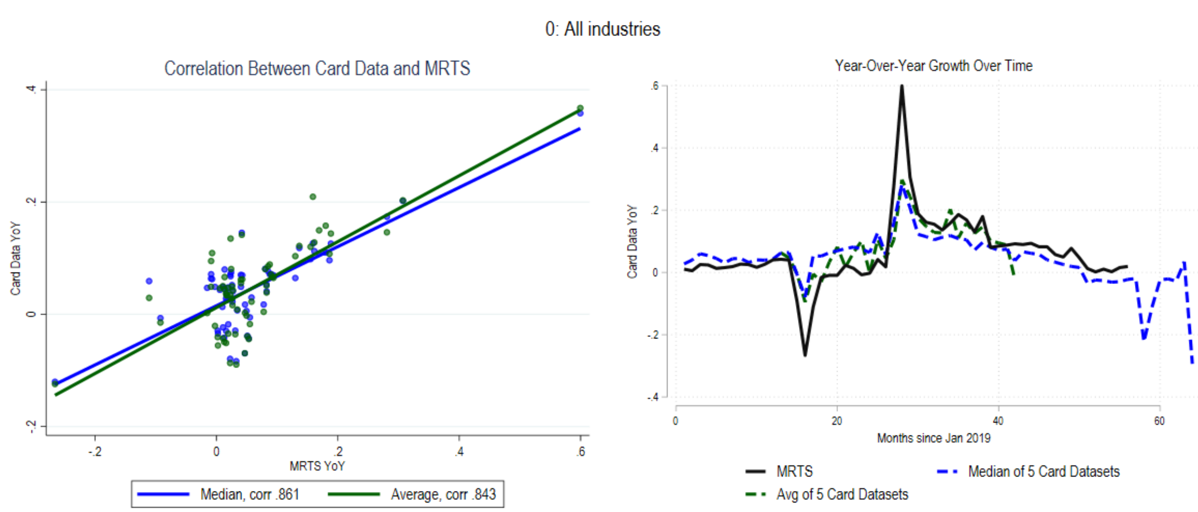


Figure 2: Correlation and Growth Rates of MRTS, Median and Average of Transaction Data, All Retail and Food Service

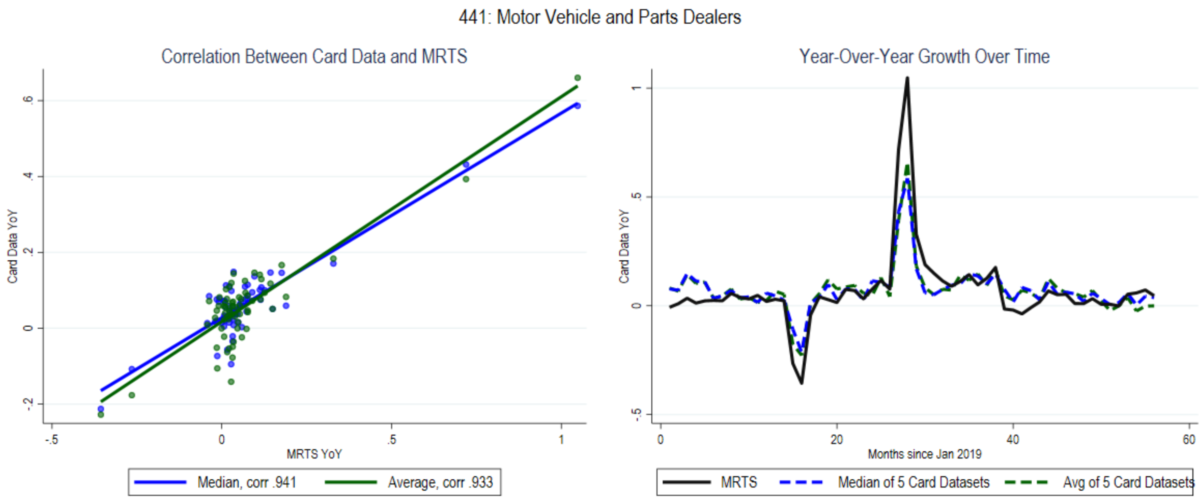


Figure 3: Correlation and Growth Rates of MRTS, Median and Average of Transaction Data

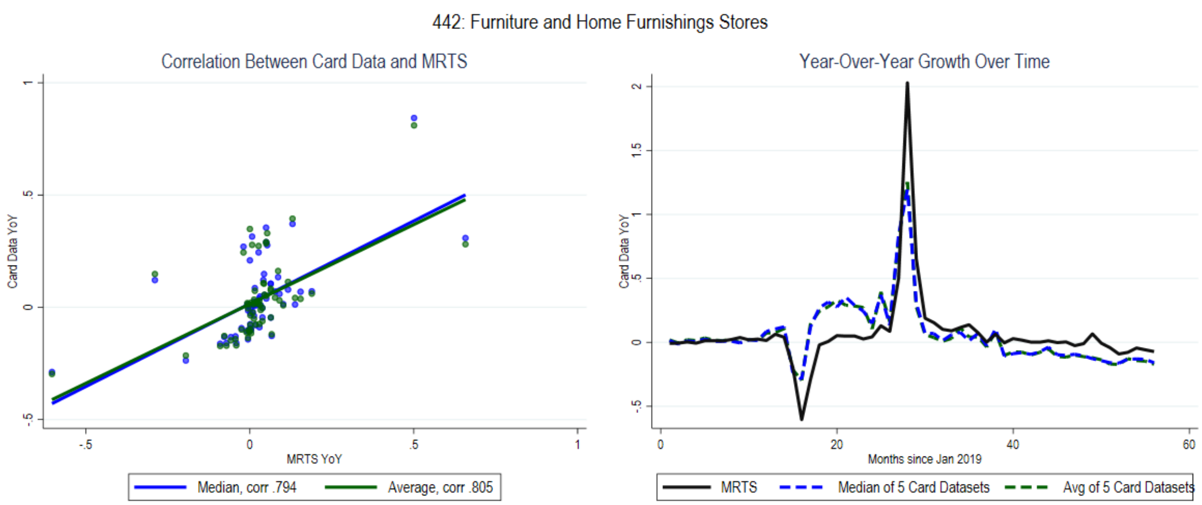


Figure 4: Correlation and Growth Rates of MRTS, Median and Average of Transaction Data

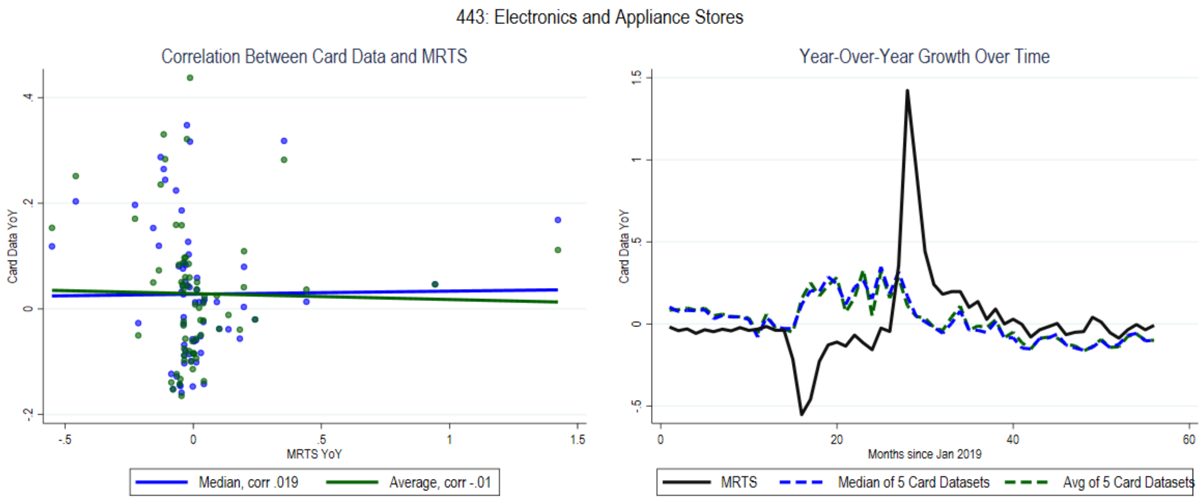


Figure 5: Correlation and Growth Rates of MRTS, Median and Average of Transaction Data

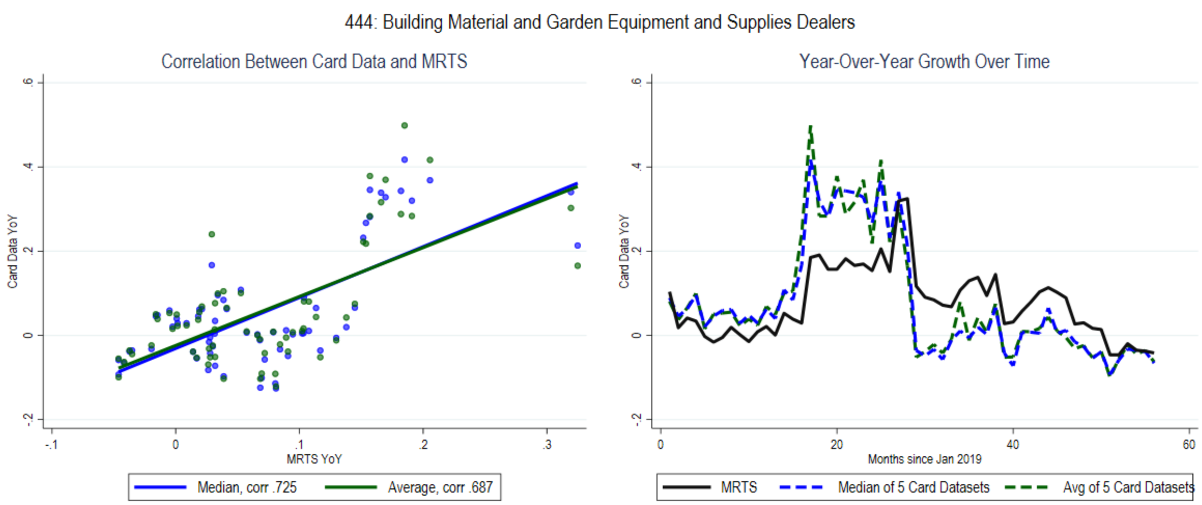


Figure 6: Correlation and Growth Rates of MRTS, Median and Average of Transaction Data

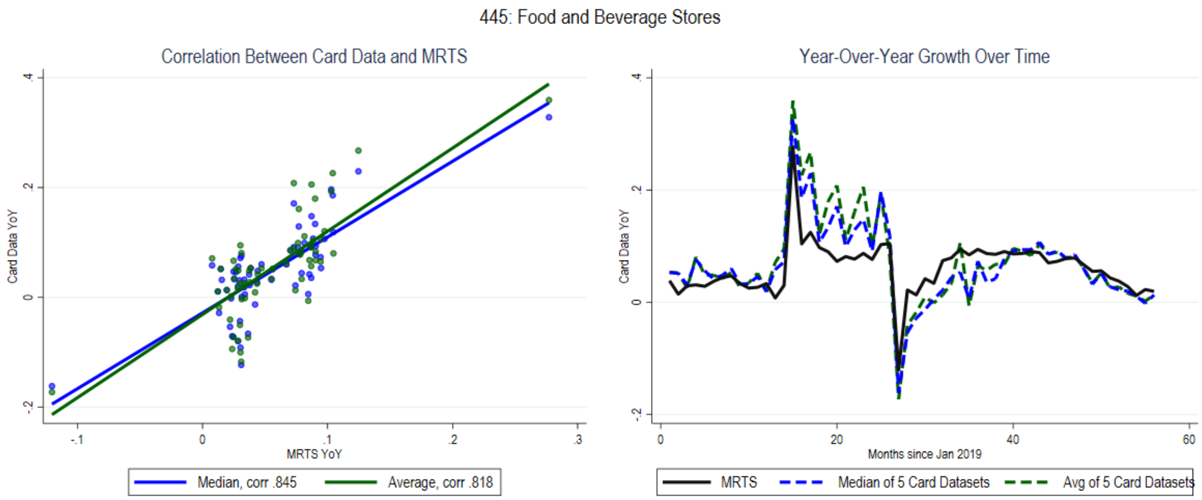


Figure 7: Correlation and Growth Rates of MRTS, Median and Average of Transaction Data

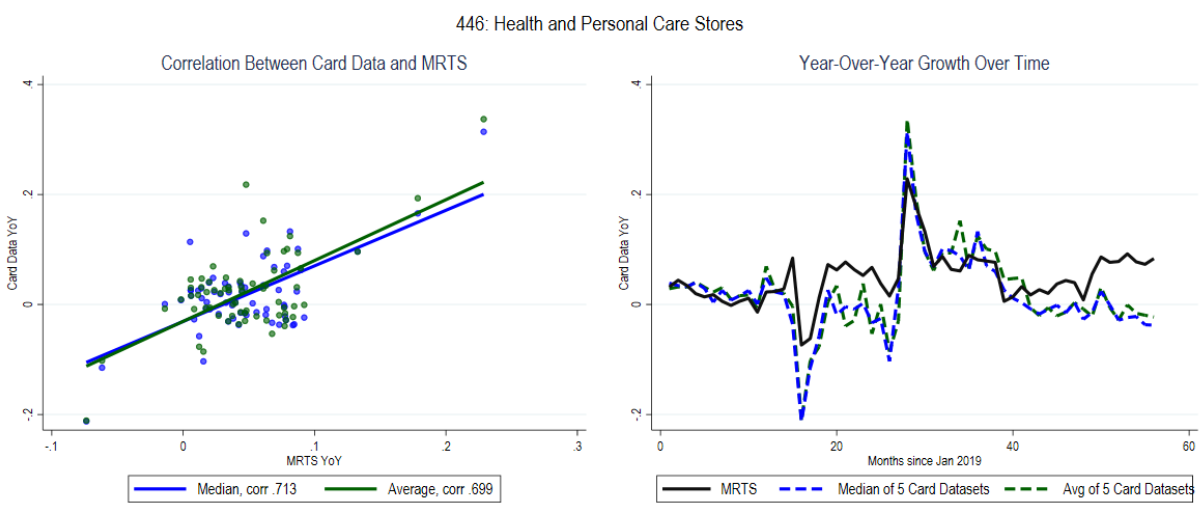


Figure 8: Correlation and Growth Rates of MRTS, Median and Average of Transaction Data

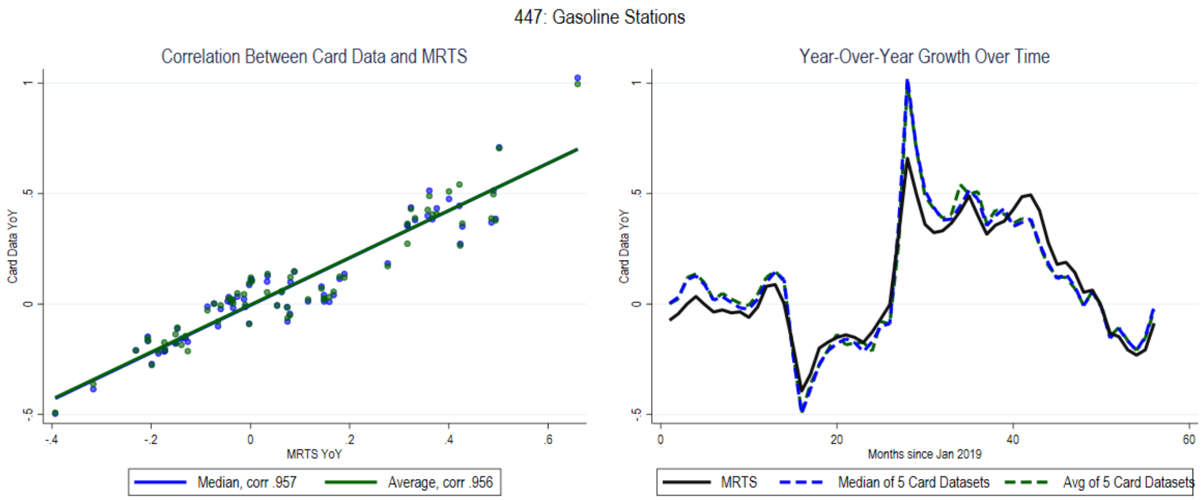


Figure 9: Correlation and Growth Rates of MRTS, Median and Average of Transaction Data

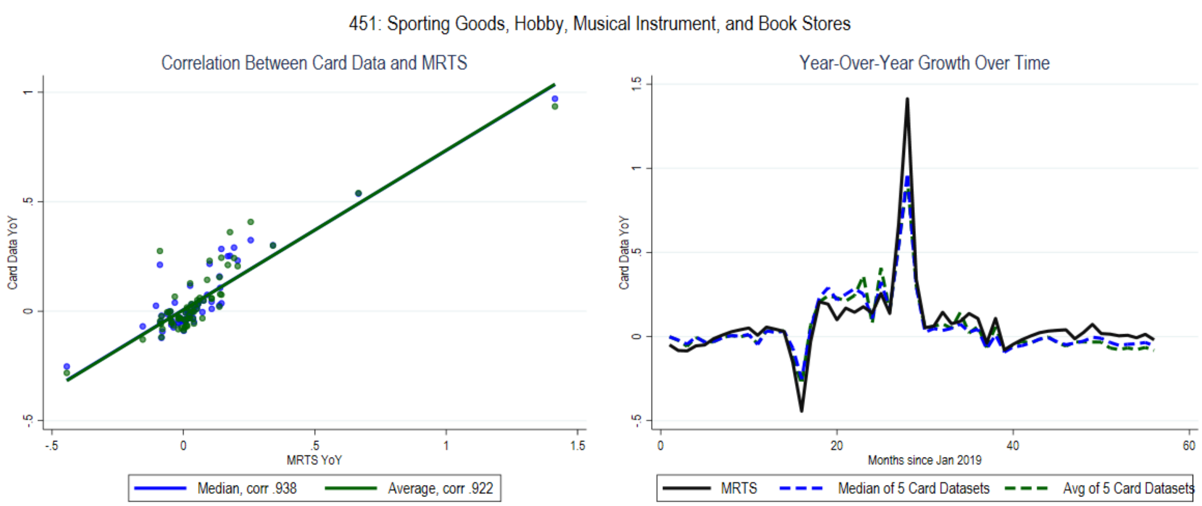


Figure 10: Correlation and Growth Rates of MRTS, Median and Average of Transaction Data

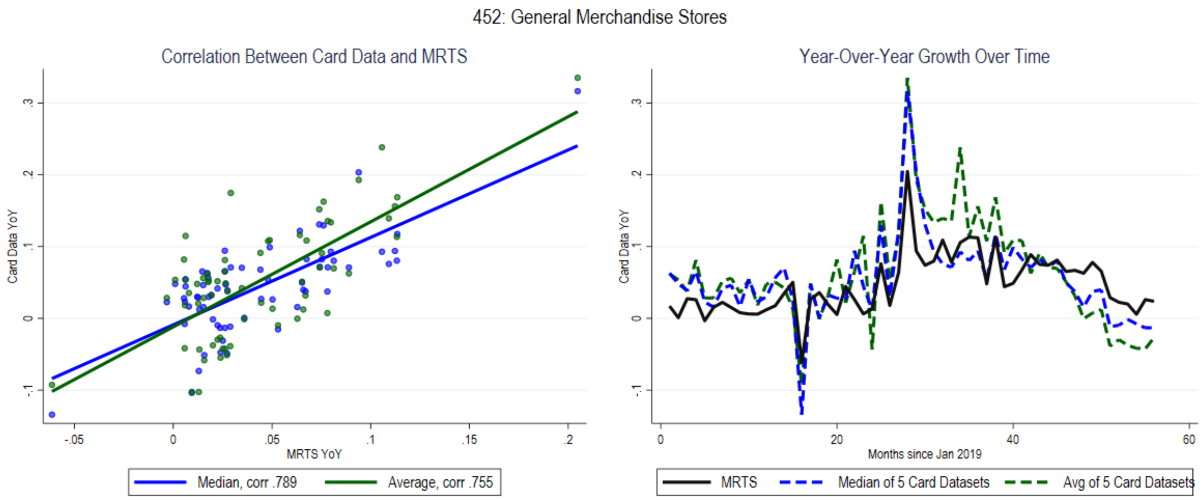


Figure 11: Correlation and Growth Rates of MRTS, Median and Average of Transaction Data

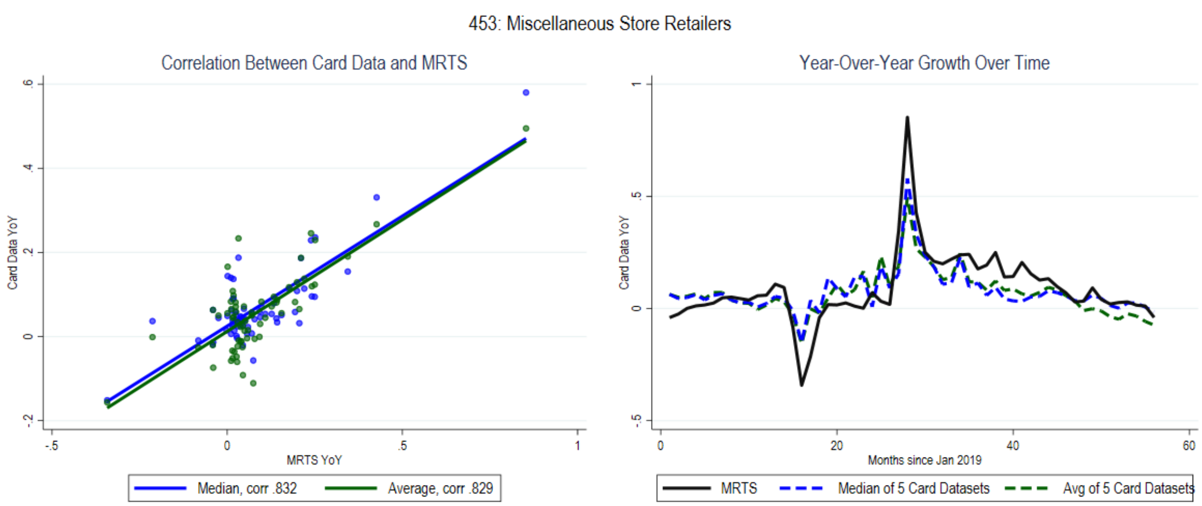


Figure 12: Correlation and Growth Rates of MRTS, Median and Average of Transaction Data

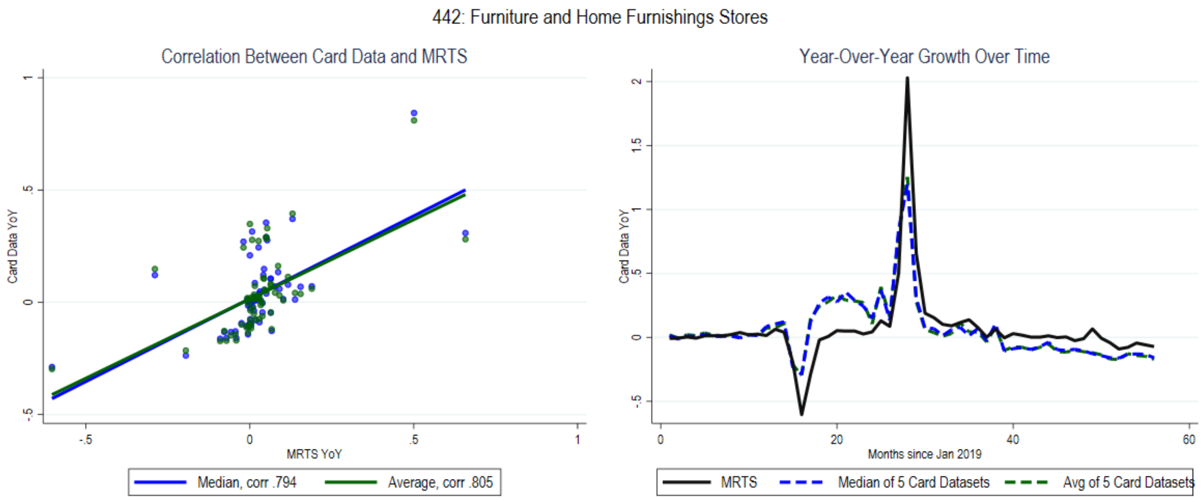


Figure 13: Correlation and Growth Rates of MRTS, Median and Average of Transaction Data

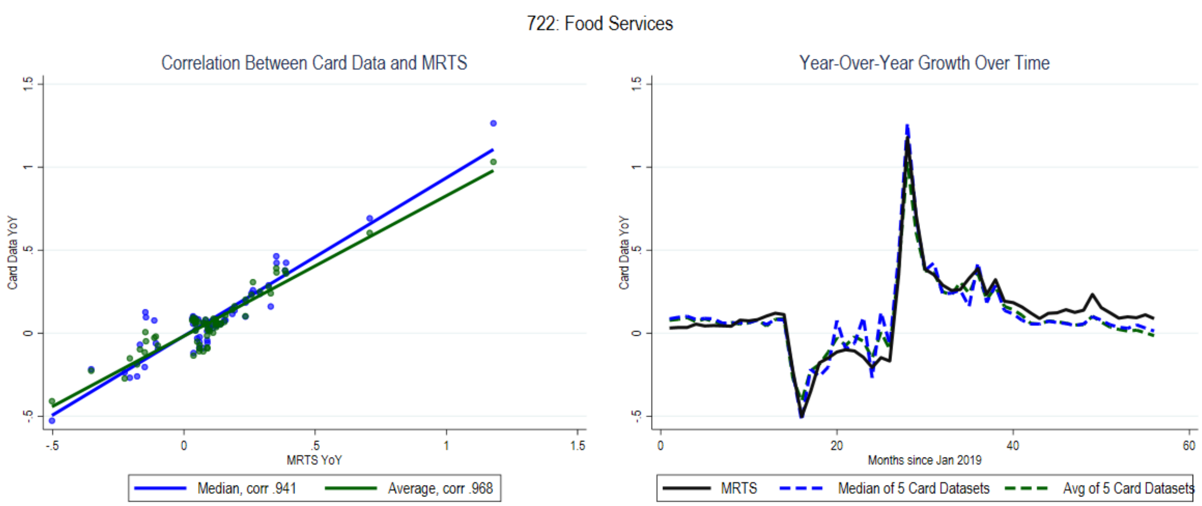


Figure 14: Correlation and Growth Rates of MRTS, Median and Average of Transaction Data

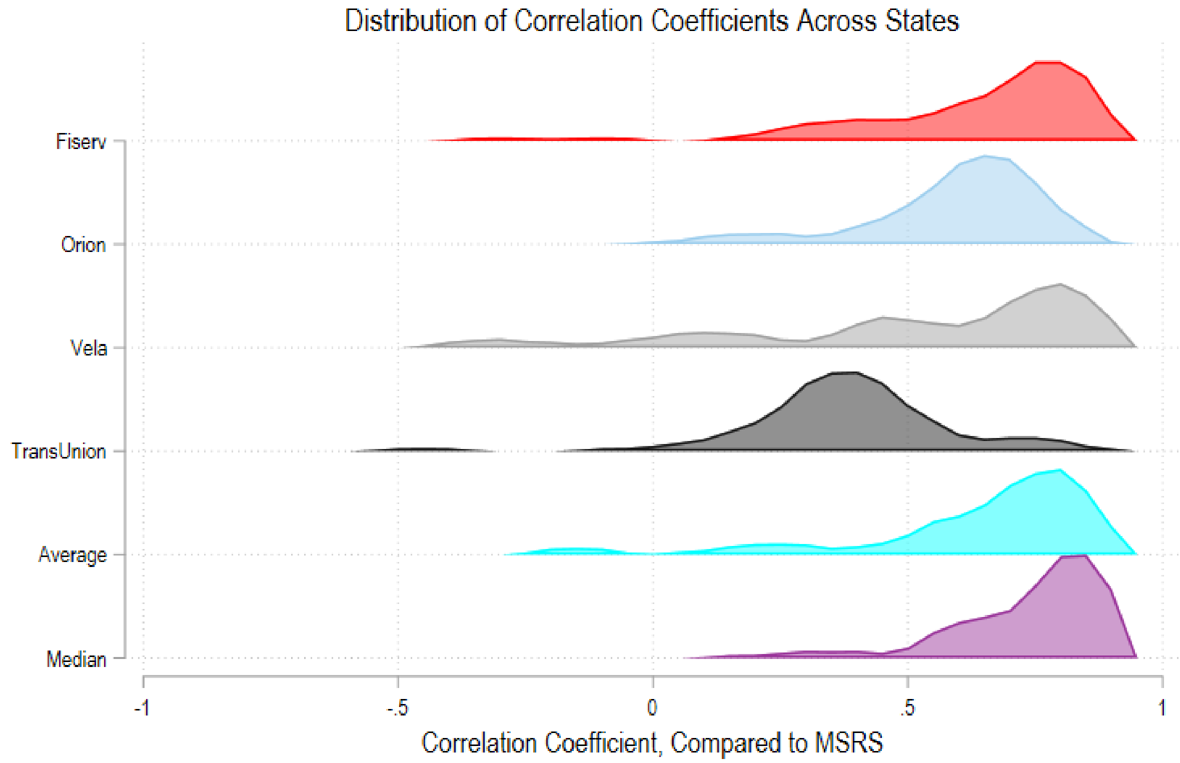


Figure 15: The Distributions of Pearson Correlation Coefficients between State MRTS and Transaction Spending Across States

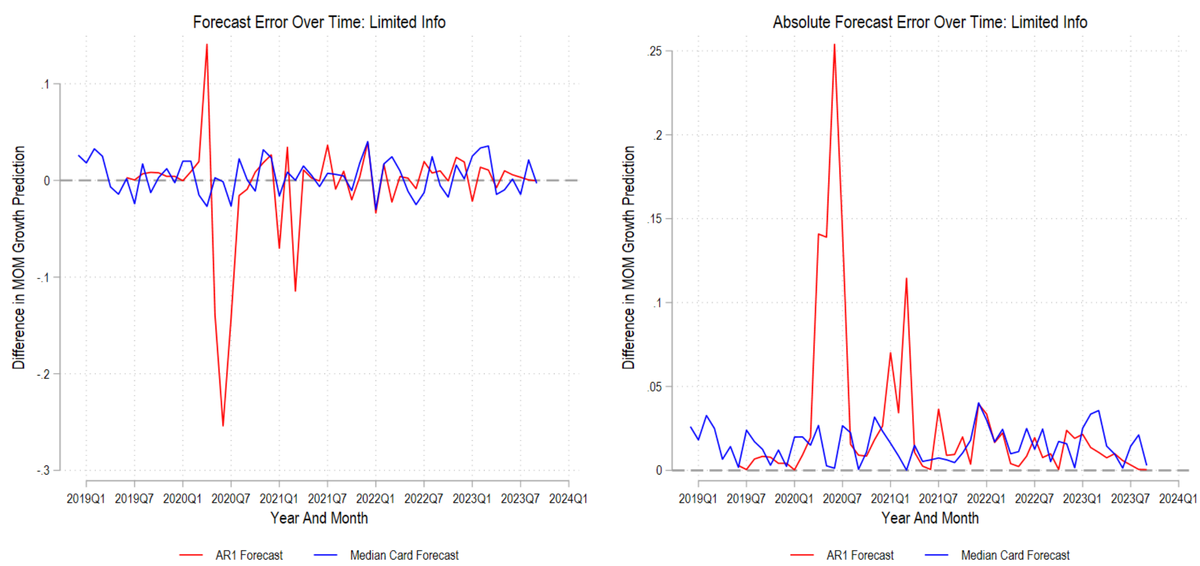


Figure 16: Forecast Errors Associated with AR(1) Forecast and Card Median Forecast

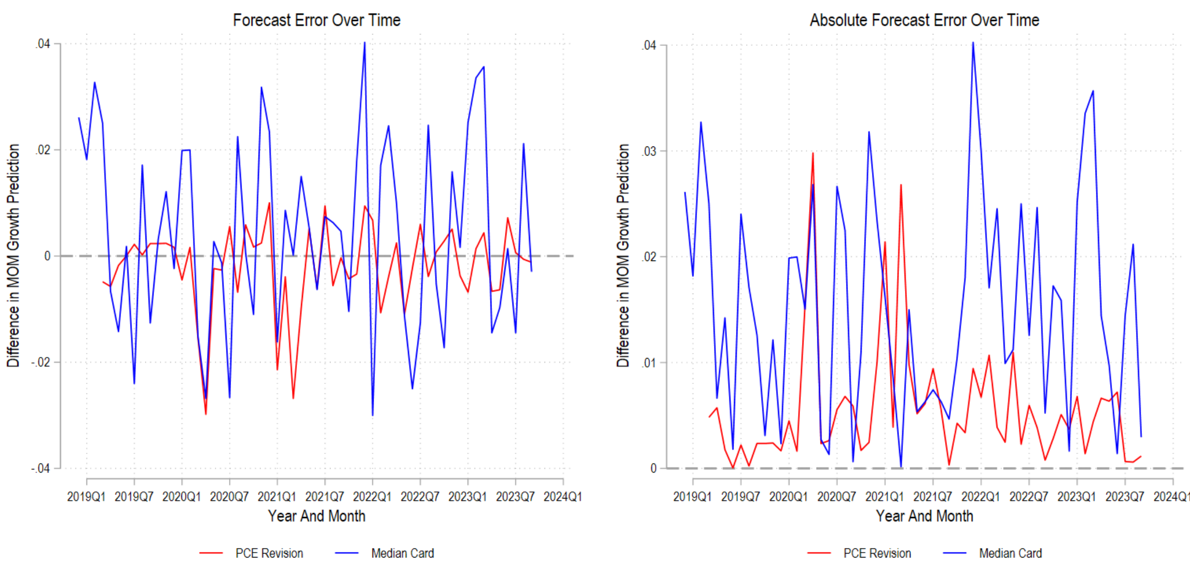


Figure 17: Revision Magnitude and Forecast Error with Card Median Forecast

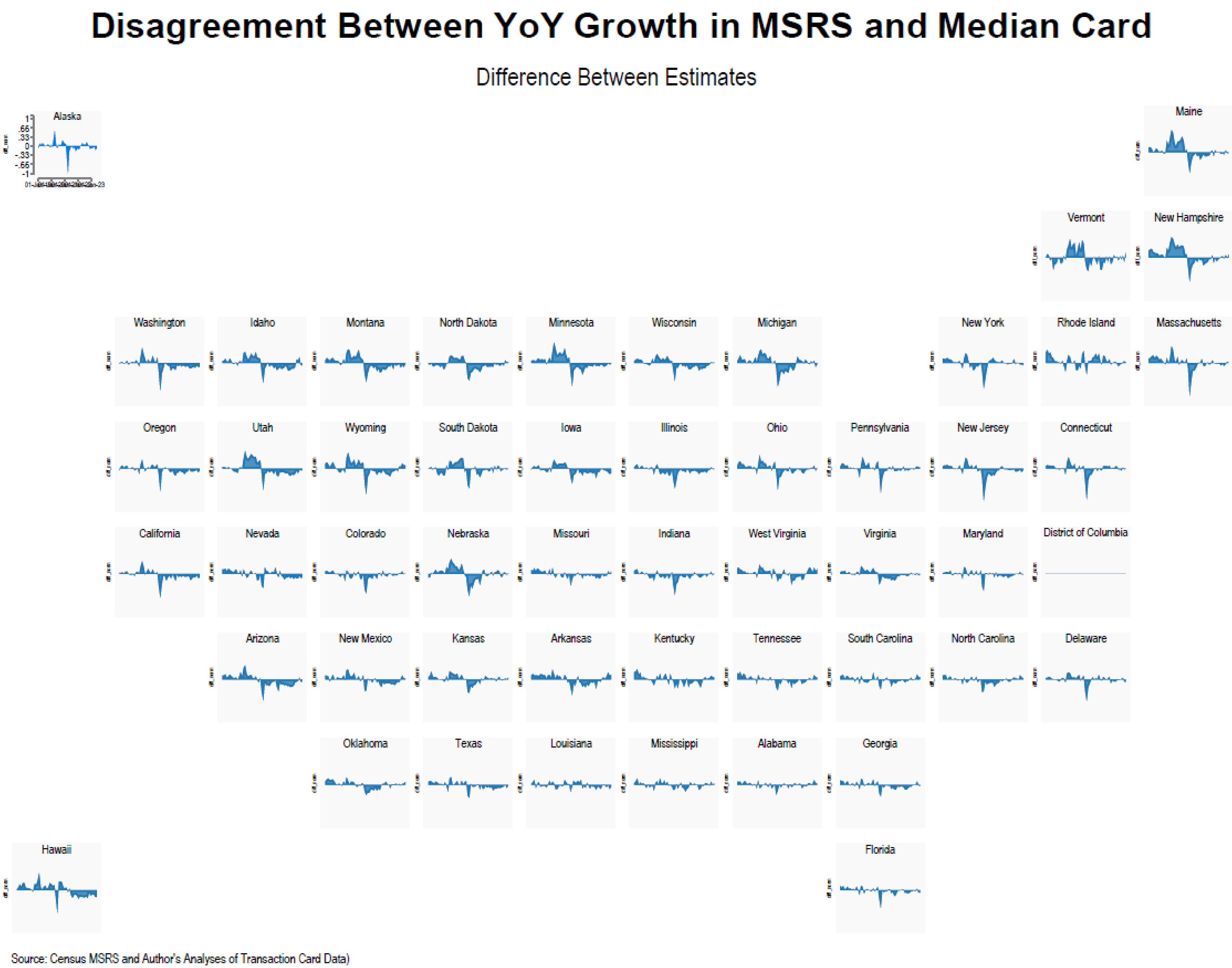




Figure 18: Visualization of Difference Between MSRS and Median Card Estimates

# A Additional Forecasting Models

Table 10 is an extension to the main Table 6, and includes mean absolute errors and relative mean absolute errors (scaled to AR(1)'s MAE) for one-month-ahead forecasts using Lasso models, the mean of past months, constrained linear regression, and an AR(13) model.

The Lasso model fits a model each month to: lagged PCE, four separate weeks each of Fiserv, Orion and Vela. The Lasso appears to not perform better than using the average or median card signals, and performs worse than the AR(1) during March through May 2020. The "Mean Past" model predicts the next month's PCE growth as the mean growth rate of past PCE growth, and performs better than the AR(1) in normal times. The constrained linear regression model fits a model using lagged PCE and the median card signal, forcing the two coefficients to add to one. The constrained model performs about similar to the AR(1) without the median card signal added in, suggesting the model weights the lagged PCE heavily compared to the median card signal. The AR(12) model fits a model based on 12 lags of PCE, and performs worse than AR(1).

Table 11 lists mean absolute errors by PCE category for random forest forecasting models and MIDAS forecasting models. Mixed data sampling (MIDAS) forecasting models are used to predict lower frequency economic indicators using higher frequency economic indicators. In this application I am using higher frequency card transaction data indicators to forecast monthly frequency PCE. The MIDAS model uses lagged weekly observations of card spending to predict monthly PCE, and the lagged observations are weighted based on recency. Column (5) applies a MIDAS model with equal weights for the past 4 weeks, column (6) applies a MIDAS model with an Almon (exponential) weighting scheme to aggregate the weekly data, and column (7) applies a MIDAS using both unweighted weekly data and the weighted aggregations across weeks. All models are fitted to the current observation of PCE growth using information available at the time, but errors are calculated using the fifth observed PCE spending.

# B State Scatterplots

Figure 20 depicts scatter plots for correlations between the state-level MSRS (on the x-axis) and the average and median card series on the y-axis. Graphs are at the state level, for CA, PA, TX, and FL – the states with the lowest standard error in the MSRS. Correlation coefficients are listed. Figure 19 plots the forecasts and actual PCE growth over time for overall PCE (PCE Table Line 1).

# C Comparing QCEW Wages and MSRS

BEA's state-level PCE currently sets spending levels based on the Economic Census and bridges across EC years using QCEW wages. Using wages to interpolate and extrapolate movements in spending may not be accurate during economic shocks that impact wages and spending differently. In Figure 21, I plot movements in QCEW wages with movements in MSRS spending for the states that have the lowest standard errors in the MSRS. Spending/Wages are for the retail categories 441–448, and 451–453 aggregated together. From first glance, it appears that wages lag spending changes by one month during COVID, and spending effects are larger in magnitude than wages changes. This suggests wage data may not be as reliable in measuring spending. However, the MSRS is an experimental product so might not capture true spending in states.

Figure 22 adds in the predicted growth in spending from state-level card spending growth. Surprisingly, the card data patterns align more closely with QCEW wages for these states than with the MSRS. The card data mirrors the QCEW wages in both timing and magnitude to a greater degree than it resembles the MSRS. It's unclear what is causing the discrepancies between the card data and the MSRS and the MSRS and QCEW. The card data, as always, raises concerns about representativeness. It's also unclear whether the MSRS or card data series better capture true state spending over time, since there are no good benchmarks for state spending.

Table 10: Forecast Errors and Routine Revision Magnitudes

| | (1) Benchmark AR-1 | (2) Benchmark Revision | (3) Fiserv | (4) Orion | (5) Vela | (6) Affinity | (7) Average | (8) Median | (9) Lasso | (10) Mean Past | (11) Constraints | (12) AR-12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Forecasting August 2019 through August 2023 | | | | | | | | | | | | |
| Relative MAE | 1 | .209 | .669 | .709 | .619 | .866 | .552 | .56600 | .838 | .722 | .911 | 40.112 |
| MAE | .027 | .006 | .018 | .019 | .017 | .023 | .015 | .015 | .023 | .02 | .025 | 1.088 |
| Forecasting March 2020 - May 2020 Only | | | | | | | | | | | | |
| | AR-1 | Revision | Fiserv | Orion | Vela | Affinity | Average | Median | Lasso | Mean Past | Constraints | AR-12 |
| Relative MAE | 1 | .159 | .175 | .499 | .122 | .52 | .168 | .149 | 1.023 | 1.017 | 1.088 | 1.446 |
| MAE | .1 | .016 | .017 | .05 | .012 | .052 | .017 | .015 | .102 | .101 | .109 | .144 |

Notes. Column (1) lists mean absolute error in predicting PCE with a one-step-ahead forecast using an AR-12 model. Column (2) lists the magnitude of four revisions between the 1st and 5th monthly PCE releases.

Table 11: Mean Absolute Errors of Machine Learning Model Forecasting PCE Categories

| Table Line No. | PCE Cat | (1) RF Lag 1 | (2) RF Past | (3) RF Avg & Med | (4) RF All Card | (5) MIDAS Indiv. | (6) MIDAS Almon | (7) MIDAS Combo |
|---|---|---|---|---|---|---|---|---|
| 1 | DPCERC | .0175 | .0194 | .014 | .0142 | .0153 | .0149 | .02 |
| 2 | DGDSRC | .0278 | .0283 | .0205 | .0203 | .0231 | .0207 | .0335 |
| 3 | DDURRC | .0492 | .0468 | .0356 | .0351 | .0373 | .0332 | .0514 |
| 4 | DMOTRC | .0709 | .066 | .0583 | .058 | .0548 | .0493 | .0751 |
| 5 | DFDHRC | .0426 | .043 | .0322 | .0313 | .0343 | .0323 | .0473 |
| 6 | DREQRC | .0352 | .0366 | .0303 | .0302 | .0306 | .0292 | .0462 |
| 7 | DODGRC | .0503 | .0566 | .0451 | .0454 | .0466 | .0441 | .0576 |
| 8 | DNDGRC | .0241 | .0251 | .0177 | .0171 | .0211 | .019 | .0276 |
| 9 | DFXARC | .0241 | .0259 | .0271 | .0253 | .0269 | .0225 | .0339 |
| 10 | DCLORC | .0572 | .0599 | .0478 | .0461 | .0476 | .0535 | .0727 |
| 11 | DGOERC | .0811 | .0818 | .0746 | .0736 | .0609 | .062 | .0735 |
| 12 | DONGRC | .0162 | .0178 | .0137 | .0139 | .0161 | .0149 | .0209 |
| 13 | DSERRC | .014 | .0158 | .0125 | .0119 | .0133 | .0131 | .0157 |
| 14 | DHCERC | .0164 | .0191 | .0144 | .0141 | .0155 | .0152 | .0195 |
| 15 | DHUTRC | .0049 | .0048 | .004 | .0041 | .0041 | .0043 | .0049 |
| 16 | DHLCRC | .0262 | .0302 | .0221 | .0212 | .0242 | .0241 | .0384 |
| 17 | DTRSRC | .0505 | .0555 | .0471 | .0452 | .0402 | .0434 | .0473 |
| 18 | DRCARC | .0587 | .0581 | .0505 | .0509 | .0495 | .052 | .0662 |
| 19 | DFSARC | .0515 | .0581 | .0454 | .0463 | .0487 | .0469 | .0527 |
| 20 | DIFSRC | .0095 | .0082 | .0079 | .0083 | .0076 | .0073 | .0082 |
| 21 | DOTSRC | .0206 | .0229 | .0194 | .0188 | .0188 | .018 | .0187 |
| 22 | DNPIRC | .0376 | .0435 | .037 | .036 | .0338 | .0329 | .0426 |
| 23 | DNPERC | .0073 | .0083 | .0069 | .0065 | .0058 | .0064 | .007 |
| 24 | DNPSRC | .023 | .0291 | .0229 | .0222 | .0213 | .022 | .032 |
| 25 | DPCCRC | .0184 | .021 | .0151 | .015 | .0167 | .0162 | .0207 |
| 26 | DNRGRC | .043 | .0433 | .0388 | .0389 | .0327 | .0326 | .0374 |
| 27 | DPCMRC | .0198 | .022 | .0162 | .0164 | .0178 | .0173 | .0239 |
| 28 | DPCXRC | .0212 | .0242 | .0181 | .0176 | .0198 | .019 | .0253 |

This table takes monthly PCE estimates from Table 2.3.5U. and runs a one-step-ahead forecast using transaction data. The first four columns detail the mean absolute errors in predicting PCE using Random Forest forecasting models, the following three columns list MAE from predicting PCE with MIDAS models.

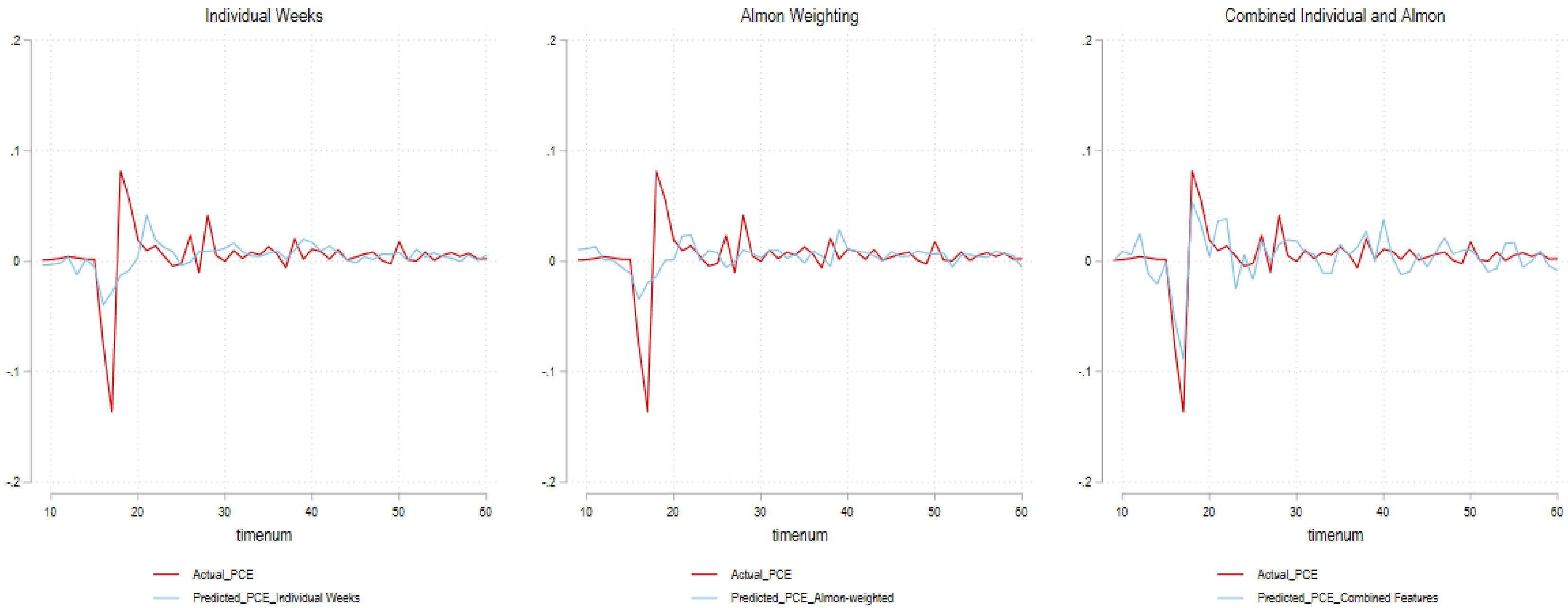


Figure 19: MIDAS Forecasts

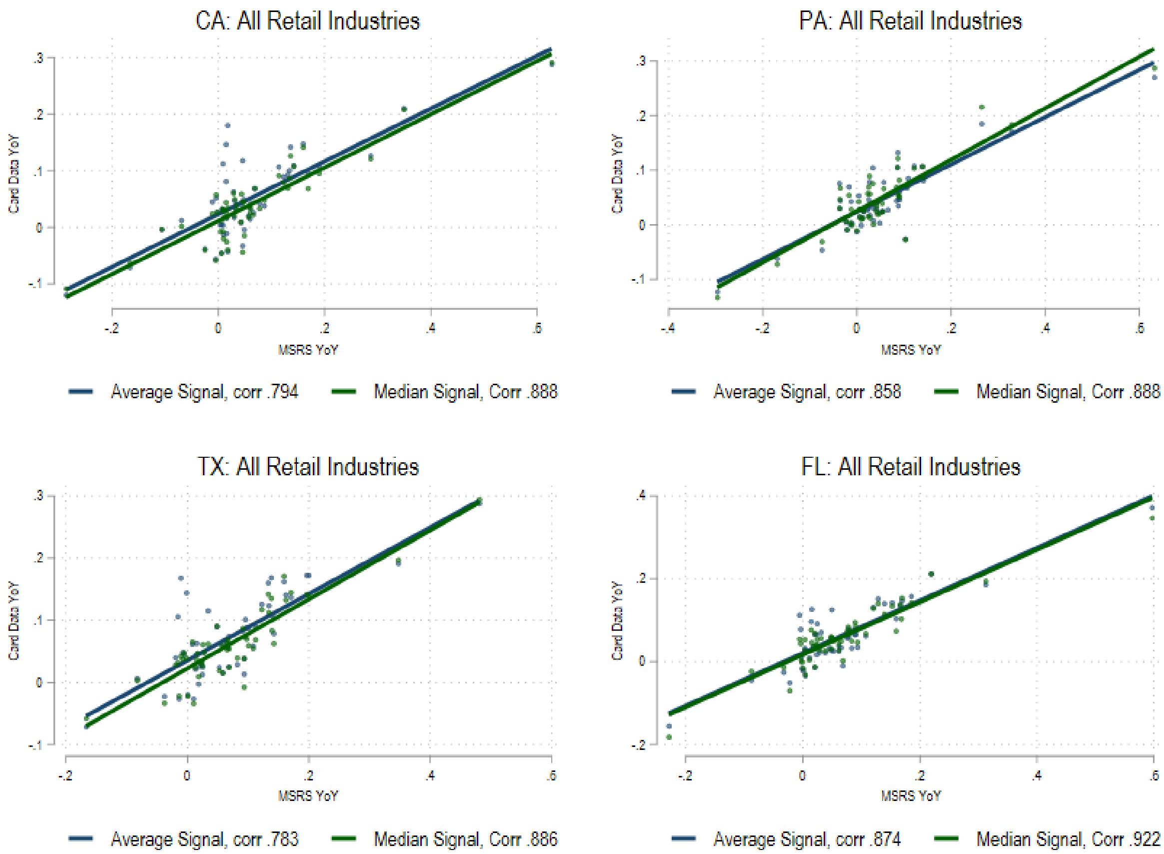


Figure 20: State Scatter Plots

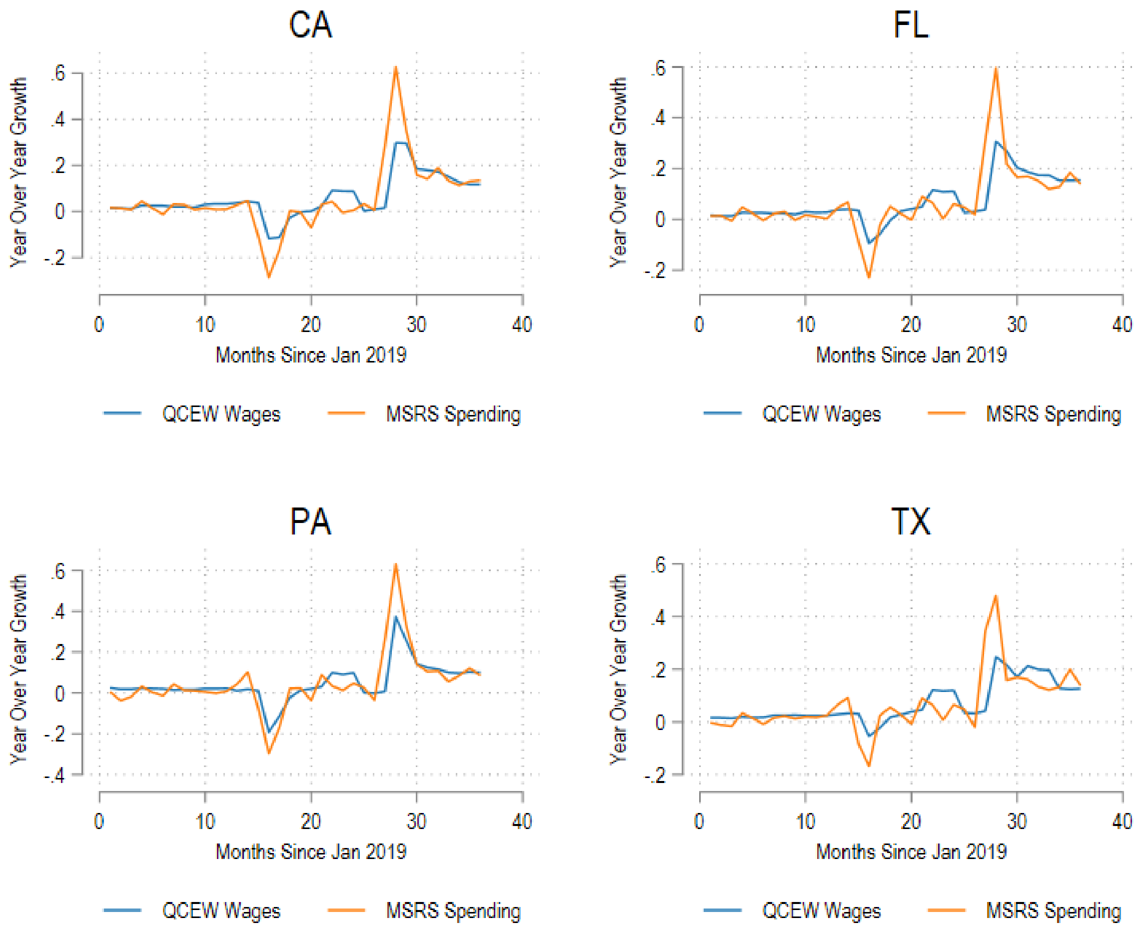


Figure 21: State-Level Patterns in QCEW and MSRS


Figure 22: State-Level Patterns in QCEW and MSRS, Adding Card Data