From Transactions Data to Economic Statistics: Constructing Real-time, High-frequency, Geographic Measures of Consumer Spending

Aditya Aladangady, Shifrah Aron-Dine, Wendy Dunn, Laura Feiveson,
Paul Lengermann and Claudia Sahm[1]

Abstract

Data on consumer spending are important for tracking economic activity and informing economic policymakers in real time. This paper describes our construction of a new data set on consumer spending. We transform anonymized card transactions from a large payment technology company into daily, geographic estimates of spending that are available only a few days after the spending occurred. The Census Bureau's monthly survey of retail sales is a primary source for monitoring the cyclical position of the economy, but it is a national statistic which is not well suited to study localized or short-lived shocks. Moreover, lags in the release of the survey and subsequent—sometimes large—revisions can diminish its usefulness for policymakers. Expanding the official survey to include more detail and faster publication would be expensive and add substantially to respondent burden. Our approach helps fill these information gaps by using data on consumer spending with credit and debit cards and other electronic payments from a private company. Our daily series are available from 2010 to the present and can be aggregated to generate national, monthly growth rates similar to official Census statistics. As an application of the new, higher-frequency, geographic information in our data set, we quantify in real-time the effects on spending of Hurricanes Harvey and Irma.

[1] We thank especially Dan Moulton, Aaron Jaffe, Felix Galbis-Reig, and Kelsey O'Flaherty for extensive work and conversations in constructing these new spending indexes. We also thank Zak Kirstein, Tommy Peeples, Gal Wachtel, Chris Pozzi, Dan Williams, and their colleagues at Palantir who have been integrally involved in implementation. The views expressed here are those of the authors and not necessarily those of other members of the Federal Reserve System.

1 Introduction

Data on consumer spending are crucial for monitoring macroeconomic conditions. In the United States, personal consumption expenditures comprise nearly 70 percent of gross domestic product and are therefore a key determinant of the cyclical position of the economy. However, consumption can be hard to measure in practice, particularly in a timely and detailed manner. Existing official statistics on consumer spending are extremely useful but limited. For instance, retail sales from the Census Bureau's surveys are published for the nation as whole at a monthly frequency. The monthly figures are available after about two weeks and often revise considerably. The Census statistics also do not include any subnational detail, so for analysis of regional shocks, researchers and policymakers must rely on other data sources, such as the BEA's quarterly regional accounts, or household expenditure surveys like the Consumer Expenditure Survey. Most of these data sources have limited sample sizes at smaller geographies and are only available after a lag of a year or two. Real-time access to geographic spending data would allow for better monitoring of shocks at the regional level and could alert policymakers before they turn into a national shock. Indeed, research on the Great Recession, such as Mian, Sufi, and Rao (2013), has shown that consumption declines were larger and appeared sooner in areas with declining house prices.

This paper describes our efforts to use transactions data from the card payments system to fill these information gaps—in particular, to provide more timely, higher-frequency measures of consumer spending for detailed geographies. The massive amount of raw data created by consumers using credit and debit cards and other electronic payments are a promising new source of information on consumer spending. Cards are the prevailing method of payment for most retail purchases in the United States. Survey data from financial institutions indicate that total card payments were $6.5 trillion in 2017 (Federal Reserve Board, 2018). The companies that process the payments associated with card-based transactions generate and store large amounts of data as part of their normal business operations. Industry analysts and market researchers have long tapped into such transactions data to observe retail shopping behavior and market trends. More recently, however, economic researchers have also begun to utilize these

and other nontraditional data sources in empirical studies of consumption.[2] These new data are not only timely, but often contain extremely detailed information on the buyers, sellers, and items purchased. And because they encompass such a broad swath of overall spending activity, card-based transactions data are also especially promising as a new source for tracking macroeconomic conditions. Yet, these new data sources also pose a myriad of new challenges, including protecting the privacy of individuals and businesses, confirming the quality of the data, and adjusting for non-representative samples.

In this project, we develop a comprehensive research dataset of spending activity using transactions data from First Data, a global payment technology company that processes $2 trillion dollars in annual card transaction volumes. We filter, transform, and aggregate their card transactions into economic statistics. To maintain confidentiality of the merchants and customers, First Data does not give us access to the transaction-level micro data. Instead, we work with Palantir—First Data's technology business partner—to build the new series to our specifications, which are fully anonymized before we receive them.[3] We currently have created estimates of daily retail spending from 2010 to the present for a number of industry categories, at the national, state, and metropolitan statistical area (MSA) level.

Our merchant-centric data on spending is, in some ways, conceptually similar to the Census Bureau's Monthly Retail Trade Survey (MRTS). As with the Census survey, our transaction data are organized by the type of the merchant making the sale. We adopt the same industry categories as Census, so we can access how well the national, monthly growth rates from our new data relate to the Census estimates. However, an important difference is the sample selection. The Census Bureau uses a statistical sampling and survey design to select the MRTS sample of about 13,000 employer firms that own or control one or more retail establishments. The survey is used to produce estimates that are representative of all retail activity.[4] First Data's

[2] Some recent examples are Mian, Rao, and Sufi (2013) using credit card company data, Farrell and Grieg (2015) using accounts from a large bank, as well as Baker (2018) and Gelman, Kariv, Shapiro, Silverman, and Tadelis (2014) using data from apps used by households.

[3] Specifically, Palantir suppresses any spending estimate based on fewer than ten merchants or where a single merchant comprises over 20 percent of the total transaction volume. In addition, some merchants also have "opt out" agreements with First Data, and their transaction data are not used in any of the analysis.

[4] For more details on the survey construction, see the Census Bureau's, "Monthly Retail Trade Survey Methodology" https://www.census.gov/retail/mrts/how_surveys_are_collected.html. Note also that a merchant in First Data is similar conceptually to an establishment in Census.

client merchants are not necessarily representative of all retailers and some First Data client merchants do not permit use of their data for research.

We are encouraged to find that, despite being constructed from very different underlying raw data sources, the Census retail sales figures and our own time series are strongly, positively correlated. Validation exercises have given us some confidence in the soundness of our methodology and the quantitative reliability of our estimates. This was an important hurdle to clear before using the high-frequency dimension of our research dataset to track economic conditions in real time for policymakers.

Our first application has been to use weather events to study high-frequency fluctuations in consumer spending in relatively small geographic areas. Specifically, we examine the impact of Hurricanes Irma and Harvey on daily spending. Our results reveal that the hurricanes significantly reduced—not just delayed—consumer spending in the affected states in the third quarter of 2017. Although the level of spending after the storms quickly returned to normal, very little of the shortfalls during the storm were made up in the subsequent weeks. Thus, on net over the span of several weeks, the hurricanes reduced spending. The declines in activity are most apparent in the discretionary components of spending, such as restaurants, as opposed to necessities, such as grocery stores. While these findings perhaps are not so surprising, we find it interesting and rewarding to see the expected patterns so clearly confirmed in the data. This also marks the first time we were able to monitor and report on the spending effects of a major weather event within days of its occurrence.

Initial applications of our new spending indexes have been promising, but our ongoing work underscores the many significant challenges with using raw transactions data to construct economic statistics. Most importantly, because these data are a by-product of ongoing business operations, changes in those operations—if not detected and addressed—can diminish the quality of the resulting statistics. In addition, the sheer quantity of transactions makes the choice of methodologies used to filter, classify, aggregate, and seasonally adjust the data essential to its ultimate usefulness. Developing these methodologies involved an extended learning process that spanned many months and involved repeated trial and error. We began these efforts in the summer of 2016 in a pilot collaboration between the First Data Merchant Services, LLC (First Data), Palantir Technologies, and the Federal Reserve Board. First Data has supported the use of its data for statistical and research purposes, while maintaining strict controls on the privacy of

merchants and consumers. Palantir's knowledge of First Data's business has been key to identifying and handling the intricacies of the data. We believe this collaboration may be a useful model for future attempts to adapt other sources of proprietary "big data" into usable statistics.

The remainder of this paper is organized as follows. In section two, we describe the transaction data from First Data. In section three, we detail the methodology we use to construct our spending series from the raw transactions data. In section four, we compare our new series with official estimates from the Census Bureau as a data validation exercise. Finally, we show how we used the transactions data to track in real time the effect on spending of Hurricanes Harvey and Irma.

2 Description of the Transactions Data

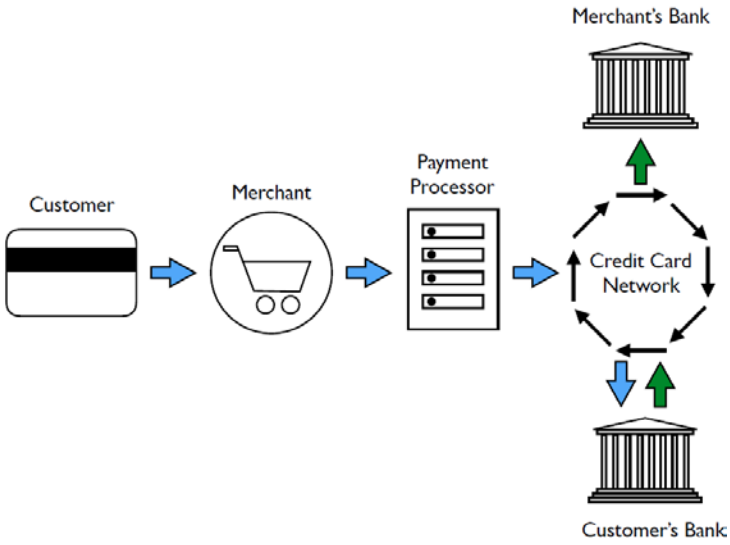
The unit of analysis in our underlying data from First Data is the swipe of a card by a customer paying at a merchant or the entering of card information online for an electronic purchase. The total dollar amount of the purchase, as well as when and where it occurred are recorded.[5] Only card or electronic transactions at merchants that work with First Data (or one of their subsidiaries) are captured in our transactions database. All cash payments as well as card payments at merchants of First Data clients with op-out clauses for further use of their data will be missed in our database. Geography is determined by the location of the merchant, which may differ from where the consumer lives.

First Data is a global payment technology company and the largest credit card transaction processor in the United States. As of 2016, First Data processes approximately $2 trillion of card payments a year. First Data serves two roles in the credit card payments market. They act as both a merchant acquirer and a payment processor. As a merchant acquirer First Data sells credit card terminals to merchants and signs them onto First Data's transaction processing network. As a payments processor First Data provides the 'plumbing' to help credit card terminals process payment authorization requests and settlements (irrespective of whether or not they are First Data terminals). They provide similar services for some types of electronic payments.

[5] The name and zip code of the merchant are in the raw data. Bank Identification Numbers (BIN) can be mapped to the card numbers and in some cases we have a flag as to whether the card was present for the transaction (in store) or not (online). While these data are initially recorded by First Data, they are only available to us in an aggregated and anonymized form.

Figure 1 illustrates the role of payment processors in a credit card transaction. When a consumer makes a purchase at a First Data merchant, First Data serves as the intermediary between the merchant and the various credit card networks. When a consumer swipes a card at a merchant's point of sale system, the processor sends the transaction information through the credit card network to the consumer's bank which then decides whether or not to authorize the transaction. That information is then relayed back to the point of sale system and the transaction either goes through or is denied. When the transaction is settled, the final transaction amount (for example including tip) is transferred from the customer's account to the merchant's account. There may be a lag of several days between the authorization and the settlement due to individual bank procedures, but both of these dates are available to us in the data.[6] Banks batch payments to other banks at various frequencies (daily, biweekly, weekly).

Figure 1. The Role of Payment Processors in Credit Card Transactions



[6] For January 2012 to the present, First Data reports both authorization and settlement dates and amounts. The authorization date should be the same as the purchase date. Thus the most accurate representation of a purchase is the authorization time stamp and the settlement amount. The settlement amount is more accurate because it would include tips which are typically not included in the authorization amount. When available, we combine data on both authorizations and settlements to characterize each transaction. The date of the transaction is the time stamp of the authorization request (when the credit card was swiped) and value of the transaction is the settlement amount (so as to include tip, or any revision in the original authorization amount). When a valid authorization time stamp is not available, we use both the time stamp and value of the settlement. From January 2010 until January 2012, First Data only reports transaction settlement dates and amounts. Due to batch processing by consumers' banks, the settlement date can be days after the actual purchase data. We used the older database to extend our time series back to 2010 by adjusting the timing of transactions with only settlement data according to the average difference in timing between settlement and authorization.

The First Data microdata includes details about every card transaction including the authorization and settlement amount and date, the merchant address, the merchant name, and the merchant category code (MCC).[7] Even though First Data only covers the portion of sales paid for by card, the possible universe of consumer spending transactions we could hope to observe with this transaction dataset is quite large. According to the 2017 Diary of Consumer Payment Choice, consumers use credit and debit cards for 30.3 percent of their payments, in dollar value, while they use cash for just 8.5 percent of dollars paid (Greene and Stavins, 2018). For the categories that we focus on—retail goods and restaurant meals—the card share is even higher.

Our analysis of First Data transactions focuses on a subset of retail stores and restaurants, which we refer to in this paper as the "retail sales group" or RSG. To compare our estimates with official Census statistics, we convert the MCC codes available from First Data to 4-digit NAICS codes. We use the unique mapping that was developed by staff at the Census Bureau and the Bureau of Economic Analysis, shown in Appendix A. The retail sales group includes total retail sales (including food services) excluding sales at building materials stores and gasoline stations and covers about one-third of total consumer spending.[8]

Finally, it is worth emphasizing that, because First Data has business relationships with merchants, not consumers, our data provide a merchant-centric view of spending. While the raw microdata also technically includes an anonymized identifier for each credit and debit card, we can only draw on the subset of that card's transactions that occur at First Data merchants. Furthermore, we do not have any information about the cardholders.

[7] First Data client merchants decide their own MCC identification. MCC is an industry standard, but the accuracy of MCC assignments is not integral to the payment processing. Palantir staff have found cases when the assigned MCC is inconsistent with the type of business that the merchant does (based on the name of the merchant). A client merchant can also have multiple MCCs, for example a grocery store with an affiliated gas stations could have one MCC for terminals in the grocery and one for terminals at the gas pumps.

[8] The retail sales group is the subset of retail and food service industries in the Census retail sales survey that are also used to estimate approximately one-third of aggregate personal consumption expenditures in the National Income and Product Accounts. It includes the following NAICS categories: 4413 - Auto Parts, Accessories, and Tire Stores, 442 - Furniture and Home Furnishings Stores, 443 - Electronics and Appliance Stores, 445 - Food and Beverage Stores, 446 - Health and Personal Care Stores, 448 - Clothing and Clothing Accessories Stores, 451 - Sporting Goods, Hobby, Book, and Music Stores, 452 - General Merchandise Stores, 453 - Miscellaneous Store Retailers, 454 - Non-store Retailers, 722 - Food Services and Drinking Places. It is worth noting that First Data also has ample coverage of several other NAICS categories not included in RSG: 444 - Building Material and Garden Equipment and Supplies Dealers, 447 - Gasoline Stations, 721 – Accommodation, 713 - Amusement, gambling, and recreation industries.

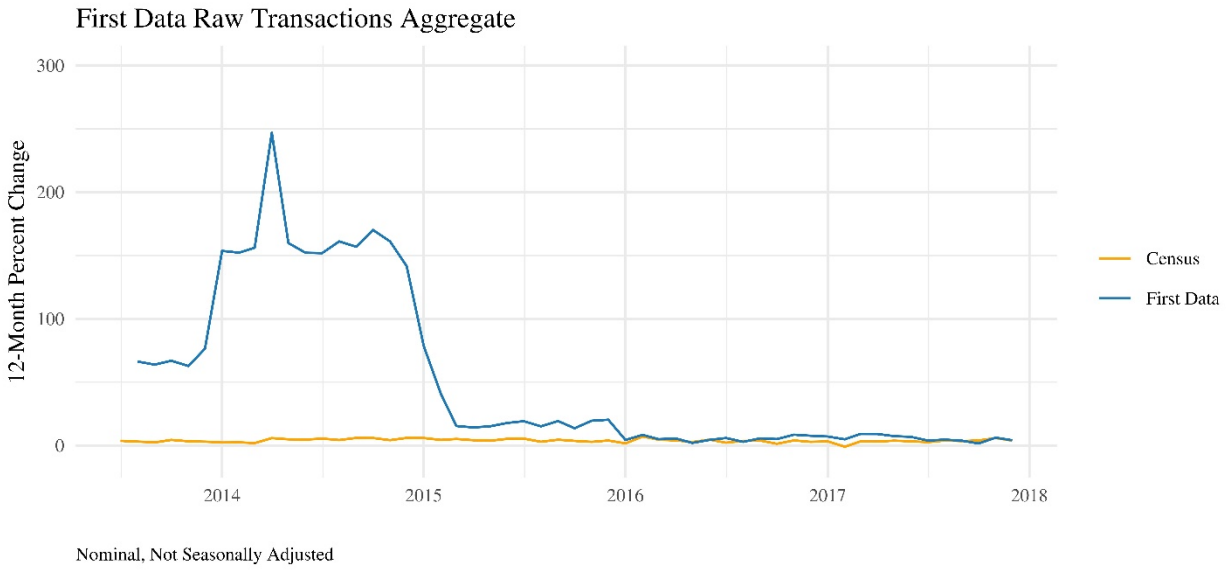
3 Methodology

In this section, we describe the methodology used to filter and transform the raw First Data transactions data into daily spending indexes for different industries and geographies. We developed the filtering strategies to isolate, as much as possible, the economic changes in spending in the data from the changes driven by business decisions at First Data. In the absence of a statistical sampling frame, the filtering of transactions is an important step in the analysis of firm-sourced "big data."

3.1 Filtering with 14-month constant merchant samples

First Data's raw payment transactions are not suitable, on their own, as statistics of overall consumer spending. Figure 2 compares the 12-month percent changes in the raw sum of transactions at RSG First Data merchants and the equivalent RSG series from the Census retail trade survey. It is immediately clear that the two series are not measuring the same underlying phenomenon. The huge swings in the First Data growth rate in early- and mid-2014 reflect their acquisition of other payment processing firms. The raw index also reflects choices by individual merchants to start, end, or continue their contract with First Data as their payment processor. One of the major challenges with these transactions data is that unlike in the Census Retail Trade Surveys we do not have a sample frame and we thus cannot distinguish actual merchant births and deaths from merchants that have decided to start or stop using First Data as their payment processor, a phenomenon we refer to as "client merchant churn."

Figure 2. Raw Sum of Transactions Retail Sales Group 12-Month Percent Change



To address client merchant churn, we use a "constant merchant approach" to filter transactions to derive an analytical sample. The basic principle is to base our *growth* estimates between two time periods on the subset of First Data merchants that regularly transact between the two periods. This filtering concept excludes *both* client merchant churn and economic births and deaths. Given the rapid expansion in First Data's business and the economic growth in the retail sector, it would be too restrictive to select merchants that transact in the full data set from 2010 onward. At the other extreme, using very short windows for the constant merchant approach, such as comparing transactions one day to the next or even one month to the next, could also be problematic because of the strong seasonal and day-of-the-week patterns in retail spending.

We chose 12-month growth rates as the focus to filter to constant merchants. For each month from January 2011 to present, we define a 14-month constant-merchant sample. Each sample includes "well-behaved" merchants (criteria described below) in the reference month and the preceding 13 months. The 14-month window includes the 12-month change ending in the reference month (which require 13 months of transactions) in our merchant sampling criteria. The extra month at the start of the sample window insures that merchants who become First Data clients in the middle of the month do not affect the 12-month change. To give a concrete example—shown in the first row Figure 3—the constant merchant sample in January 2017 is a subset of the client merchants that transacted in each month from December 2015 to January

2017. The sample for December 2016—in the second row—is based on transactions from November 2015 to December 2016. The same merchant may appear in multiple monthly samples, but it will depend on the merchant's transactions behavior within the specific 14-month window.

Figure 3. Illustration of Overlapping of 14-Month Constant Merchant Samples

	Dec-15	Jan-16	Feb-16	Mar-16	Apr-16	May-16	Jun-16	Jul-16	Aug-16	Sep-16	Oct-16	Nov-16	Dec-16	Jan-17
	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0
…	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	
…	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0		
…	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0			
…	-9	-8	-7	-6	-5	-4	-3	-2	-1	0				
…	-8	-7	-6	-5	-4	-3	-2	-1	0					
…	-7	-6	-5	-4	-3	-2	-1	0						
…	-6	-5	-4	-3	-2	-1	0							
…	-5	-4	-3	-2	-1	0								
…	-4	-3	-2	-1	0									
…	-3	-2	-1	0										
…	-2	-1	0											
…	-1	0												
…	0													

One implication of the 14-month trailing window for each constant merchant sample is that for any calendar month, we will have multiple estimates of spending from different samples. For instance, the shaded area in Figure 3 shows the 14 different merchant samples that we use to estimate spending in December 2015. The reference months for the samples range from December 2015 to January 2017. We discuss below how we combine the estimates across the reference samples into a single time series. This overlapping sample methodology is applied independently to each 3-digit NAICS code and geography.

3.2 Additional criteria for selecting "well-behaved" merchants

We applied several other filtering criteria for selection into each 14-month constant merchant sample: [9]

[9] The underlying raw sample (before filtering) excludes merchants that have opted out of having their data shared. We also control for the introduction of new payment processing platforms by imposing a three month lag before

1. Misclassified MCCs to NAICS mapping: Some merchants were determined by Palantir to be paired with inaccurate MCCs and were subsequently dropped from our analysis. For example, MCC code 5962 (Merchandising Machine Operators) was found to contain many merchants that should be classified as Travel Vendors.

2. Batch processors: Merchants cannot have more than 40% of their transaction volume concentrated in one day in a month. This cutoff is well above the typical transaction distribution for extreme days such as Black Fridays and the days before Christmas. The goal of this filter is to remove merchants who batch their transactions over several days for processing.

3. Minimum monthly spending/transaction days: Merchants must transact more than 4 days and clear at least 20 dollars in every month of the sampling window. This filter removes merchants who effectively leave the First Data platform but still send in occasional transactions to avoid inactivity/early termination fees. It also removes any merchants that may be batching transactions at a lower frequency that were not captured above.

4. Growth outliers: The 12-month percent change in each merchant's sales must be within the inner 99.99% of the distribution of growth rates of merchants at that NAICS code and geography combination.

Table 1 shows the fraction of spending removed from our sample in each filtering step for the 14-month window ending in January 2017. The denominator throughout is the set of merchants in MCC codes associated with the retail sales group that do not opt out of the sample or are not transacting on platforms that are dropped from the sample. Our final, filtered sample, shown in the last row of the table, accounts for a little over half of the dollar transaction volume in the raw

merchants on the new platform can appear in the sample since merchants often exhibit volatile behavior in the data when a new platform comes online. One small platform with several data quality issues is dropped from our sample.

sample, but reflects a set of merchants with a stable attachment to First Data and for whom sales growth appears well-measured by the data.

Table 1: Filtering Steps – 14-month window ending Jan 2017

	Cumulative % Dollar volume not filtered	Cumulative % Merchants not filtered
Opt-out Merchants and Platform Introduction	100.0%	100.0%
Misclassified MCCs to NAICS Mapping	86.7%	89.5%
Batch Processors	85.2%	81.5%
Minimum Monthly Spending/Transaction Days	85.2%	80.2%
14-Month Constant Merchant Sample	52.7%	29.1%
Growth Outliers	51.4%	29.1%

Note: Table shows fraction of merchants and associated transaction volumes that meet each successive filtering criteria in the 14-month window from December 2015 to January 2017.

3.3 Adjustments to the first and last month of the constant-merchant sample

Before we combine information from the overlapping 14-month merchant samples, we need to correct for a bias at the beginning and end of the samples. For each month in the dataset (excepting the first 13 months and the most recent 13 months), there are exactly fourteen 14-month samples that have a sales estimate for that month, and thirteen 14-month samples that have a monthly sales *growth* estimate for that month (which requires that months *t* and *t-1* be in the sample). Although the monthly level of sales in each sample is highly dependent on the merchant births, deaths, and business acquisitions between overlapping 14-month merchant samples, we find that the estimates of monthly growth in different samples are, on average, similar, with two notable exceptions: The first monthly growth estimate from a 14-month merchant sample is biased upwards, and the last monthly growth estimate is biased downwards. To make things more explicit, call g_t^{t+j} the estimate of monthly growth in time *t* that comes from the 14-month sample ending in month *t+j*. For each month *t*, we construct the average growth rate, $\bar{g}_t$ using all 14-month samples that include an estimate of the growth rate in *t*:

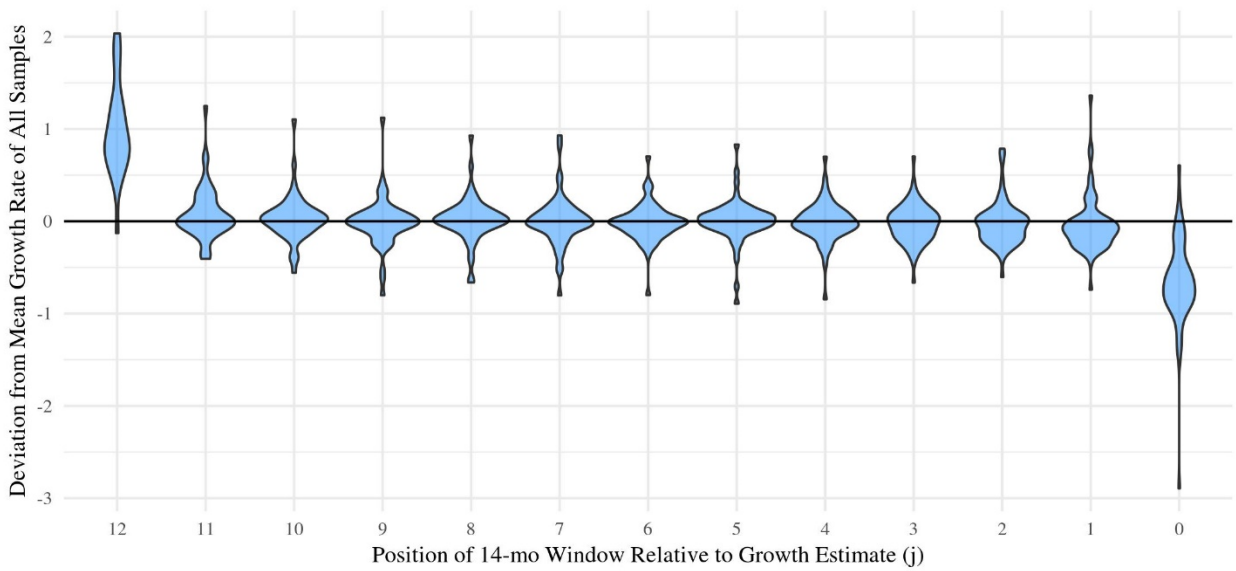
$$\bar{g}_t = \frac{1}{13}\sum_{j=0}^{12} g_t^{t+j}$$

Next, we calculate the deviation of the growth estimate *t* from a merchant sample t+*j* relative to the average across all samples:

$$deviation\ from\ mean\ (j,t) = g_t^{t+j} - \bar{g}_t$$

In Figure 4, we plot the distribution of deviations in all calendar months in the dataset, based on where growth estimate the falls in the merchant sample window (the index *j*).[10] The upward bias at the beginning of the 14-month sample—i.e. the growth rate at time *t* for the sample which runs from *t-1* through *t+12*—comes from a "birthing" bias due to firms that were just born and who are therefore ramping up sales. Equivalently, the downward bias at the end of a sample—the growth rate in which runs from *t-13* through *t*—are from the fact that firms that are about to die (say in time *t*+1, just after the sample ends) tend to have falling sales.

Figure 4. Deviation from Mean Growth in Each Month of the 14-Month Sample



To address this issue, we apply a simple correction model to fix the first and last month's estimate based on the mean growth rates from other sample estimation windows. Assuming that the size of the bias varies by month of the year (*m*), we estimate a separate correction factor β_m^j

[10] Figure 4 shows the results for the national RSG, although the picture is similar for other NAICS codes and geographies.

for each month of the year, for both the 14-month sample ending in t+12 (j=12), and the sample ending in *t (j=0)*, as:

$$\bar{g}_{t,m} = \beta_m^j g_{t,m}^{t+j} + \varepsilon_t$$

The β_m^j applies a correction that results in adjusting up the growth estimates from the end of a 14-month sample, and adjusting down the growth estimate from the beginning of a 14-month sample. We run these regressions separately for every NAICS code and geography.

To apply this fix to the daily values within the first and last month, we assume that the magnitude of the last-month bias increases and the first-month bias decreases over the course of the month. If Δ is defined as the dollar value of the adjustment for a particular month's estimate, the daily dollar adjustment amount for day *d* in a month of length D is:

$$\frac{2\Delta \mathrm{d}}{D^2 + D}$$

This correction is particularly important to achieve unbiased readings of spending for the most recent months of the data output. The index that covers recent months will necessarily only depend on the 14-month samples that end with those months (since the subsequent 14-months samples do not yet exist), their growth rates would be severely biased downward without this correction.

3.4 Combining constant merchant samples to create a daily spending index

We next combine our adjusted 14-month constant merchant samples to construct a daily index of spending for each 3-digit NAICS code and in each geography. Since the data are daily, we cannot simply back out an index by cumulating the average monthly growth rates from our 14-month samples, as might have been the most natural approach if we were using monthly data. Instead, we choose to take a weighted average of the *level* of the 14-month samples for each day.

There are three basic steps to combine the 14-month samples. In the first, we "level adjust" the 14-month samples by multiplying them by factors, as described below, to smooth

through shifts in the level due to merchant churn or platform acquisition. In the second step, we average together the level-adjusted overlapping 14-month samples for each day. This creates a continuous daily index. In the final step, which we discuss in the next subsection, we renormalize this index such that the 2012 level of spending matches the Economic Census value for that NAICs code and geography.

Denote the estimate of daily sales on day *i* of month *t* from the 14-month sample series ending in month *t*+*j* as a_{it}^{t+j}. In the level-adjustment step, we scale each successive 14-month sample by a factor, f_t, such that the average of spending over the first 13 months of the series is equal to the average spending of those same 13 months in the preceding, and already scaled, 14-month sample. These factors are multiplicative, $f_t = \prod_{s=0}^{t-1} q_{t-s}$ where $q_t = \frac{\sum_{k=1}^{13} \sum_{i \in t-k} a_{it-k}^{t}}{\sum_{k=1}^{13} \sum_{i \in t-k} a_{it-k}^{t-1}}$..

In the second step, we average together the 14 indexes that cover each day's spending to get our daily spending series:[11]

$$x_{it} = \frac{1}{14} \sum_{j=0}^{13} f_{t+j} a_{it}^{t+j}$$

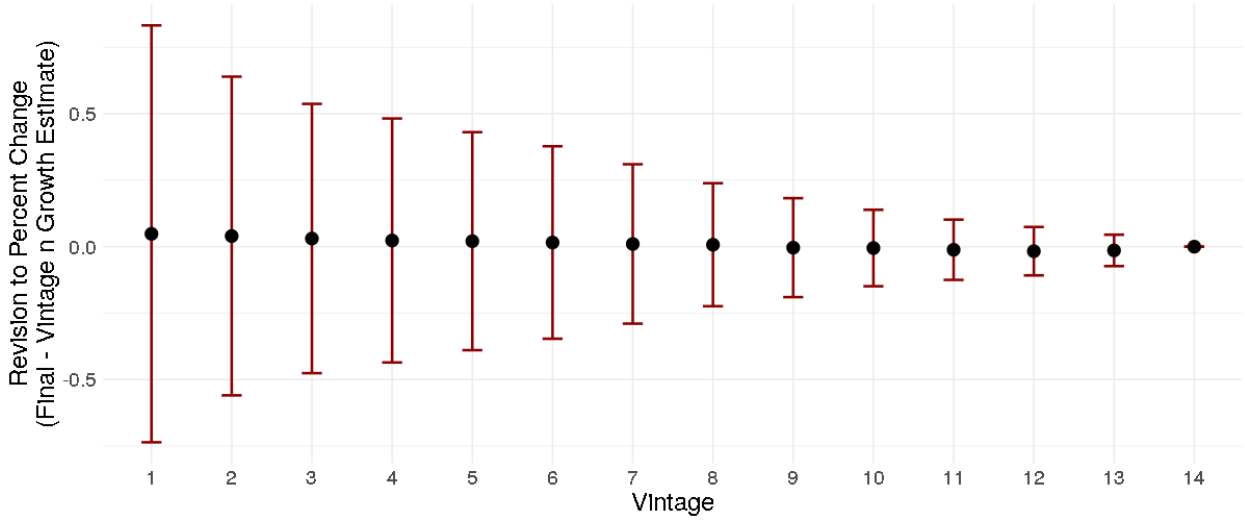
The result is a continuous daily series that can be used to derive growth rates of spending in different geographic regions, for different industries, and over different time frames. As is desirable, the monthly growth rates that derive from this index are weighted averages of the monthly growth rates in each contributing 14-month sample.[12]

One consequence of using this methodology is that the estimates for the most recent thirteen months of the series revise when we receive a new complete month of data (and thus a new 14-month sample). Figure 5 shows the magnitudes of the revisions between a given real-time estimate and the final estimate of growth in a given month. The dots and bars reflect the means and 90% confidence intervals, respectively, for differences between the final growth rate estimate and each vintage growth rate estimate, based on monthly observations from April 2011 to December 2017.[13] The vintage describes the number of months passed at the time of the

[11] For days in the months at the start or end of the existing data span, we average together whatever indexes are available for that period, which will be less than 14.
[12] See Appendix B.
[13] At the time of this writing, this date range includes all of the months that have the full 14 vintages associated with them and are no longer subject to revision.

estimate: for instance, the first vintage of February 2017 will be the first monthly growth rate estimate for February in hand at the beginning of March. The second vintage will be the monthly growth rate estimate for February in hand at the beginning of April, and so on. The final growth rate estimate for a month is the growth rate of that month once it no longer revises, which occurs after 14 months. The revisions, particularly for the first few vintages, can be as high as around plus or minus 0.8 percentage point at the outer range of the 90% confidence interval, although they appear to be essentially unbiased. The preliminary estimates of monthly RSG growth from Census have roughly comparable standard errors.[14] As we make further refinements to our data estimation methods, including adjustments like those discussed in section 3.3, we anticipate that the revision standard errors will shrink.

Figure 5. Revision Properties of First Data Retail Sales Group Monthly Growth Rates



Note: Black dots show the mean revision to monthly seasonally adjusted growth rates, and red bars show the 90% confidence interval (i.e. 1.65 times the standard deviation.)

[14] The standard deviation of the revisions to the preliminary Census monthly growth rate is 0.4 percentage point, as compared to 0.5 percentage point in the First Data.

3.5 Using the Economic Census as benchmarks

In the final step of our index creation, we reweight each 3-digit NAICS daily spending series to better represent the composition of spending across industries in national and regional aggregates.

The only source of retail sales volume with comparably detailed industry and geographic detail is the Economic Census. The most recently available census was conducted in 2012. For each of our 3-digit NAICS level series, we normalize the level in 2012 to equal the Economic Census level for that region and industry.[15,16] Then, to construct spending indexes for the Retail Sales Group, or any other grouping of 3-digit NAICs codes, we sum over the 3-digit NAICS codes that compose those series. This provides a set of series that better reflects the distribution of spending across industry categories. We use this same weighting scheme separately to generate estimates for national, state, and MSA level spending series.

The Economic Census also allows us to check how well the First Data indexes cover the universe of sales in the country. For each year, the "coverage ratio" of each index is computed by dividing the total First Data sales that are used in the creation of the index by the total estimated sales in the region.[17] Figure 6 shows that the coverage ratio of the national RSG has increased from roughly 5.5% in 2010 to 7.5% in 2018. However, the coverage is not uniform across the country. Figure 7 plots the coverage ratio of the RSG in each state in 2018. Some states, such as North Dakota and Iowa, have low coverage at 3.4 and 3.3 percent, respectively, while others have excellent coverage such as Nevada with 13.8 percent and Alaska with 10.8 percent.

[15] These weights correct for biases in industry composition, but do not necessarily correct for other ways in which First data firms may differ from those in the economy. For example, if First Data firms are smaller than those in the economy as a whole and if spending growth differs systematically by firm size, the series may be biased towards smaller firms.

[16] For those region-NAICS code pairs for which the 3-digit NAICS code is suppressed in the Economic Census, we impute them using the number of firms in that industry and region. When the First Data index is suppressed for 2012, we instead normalize the first full year of the First Data index to the Economic Census level for that region-industry that is grown out using the national growth rates for the 3-digit NAICS.

[17] Sales in the region in a given year are estimated using 2012 Economic Census sales in a given region grown out by national growth in 3-digit NAICs categories from the Census MRTS.

Figure 6. First Data Coverage of National Retail Sales Group Sales

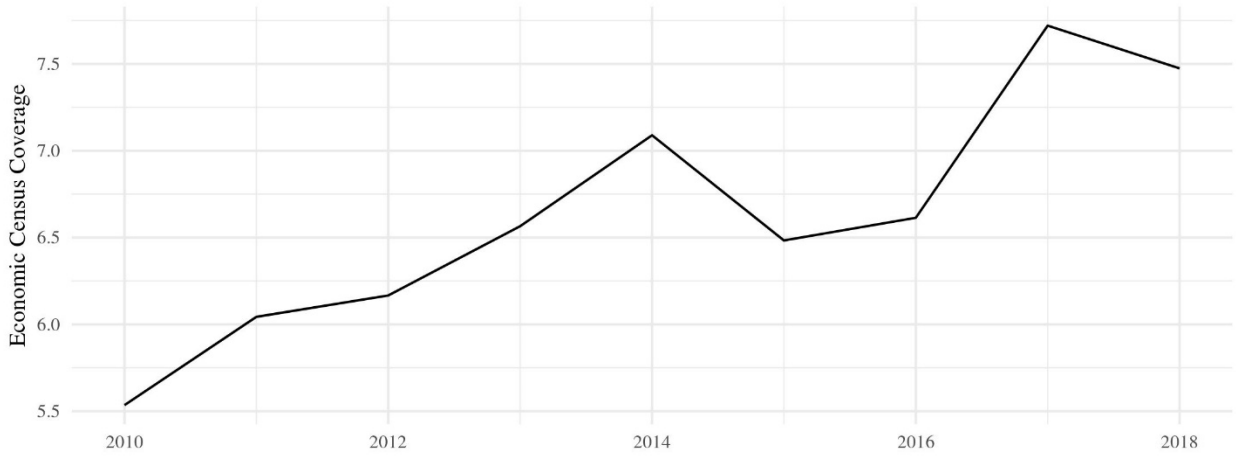
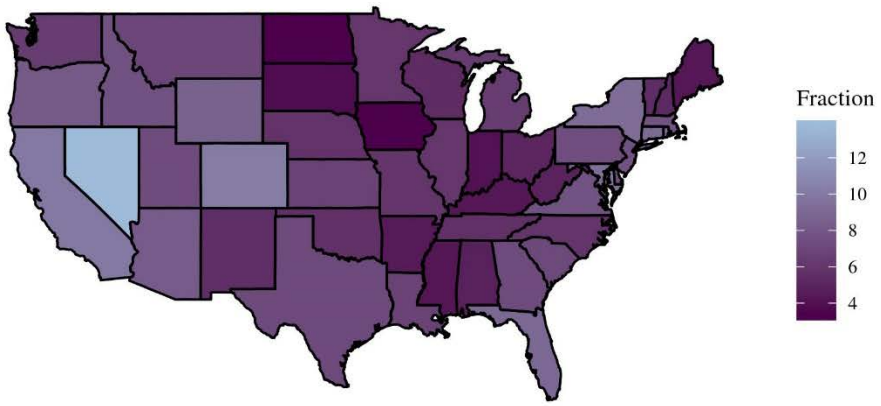


Figure 7. First Data Coverage of Economic Census Retail Sales Group Sales by State, 2018



3.6 Seasonal Adjustment

In order to use our spending indexes for time-series analysis, we also need to filter the indexes to remove regular variation related to week days, holidays, and other calendar effects. After exploring several alternative strategies, we have taken a parsimonious approach: We seasonally adjust the data by summing the daily transactions by calendar month and running the monthly series through the X-12 ARIMA program maintained by the Census Bureau. An advantage of this method is that it is also used to seasonally adjust the Census retail sales data, which we use for comparison with our own monthly estimates.

Seasonal adjustment of the daily data is more challenging, partly because the methods for estimating daily adjustment factors are not as well established. That said, working with daily data offers some potential advantages in this regard. As pointed out by Leamer (2014), with daily data we can directly observe the distribution of spending across days of the week, and this allows for a relatively precise estimation of weekday adjustment factors. Indeed, we find that retail transaction volumes vary markedly by the day of the week—the highest spending days appear to be Thursday, Friday, and Saturday, and the lowest spending day by far is Sunday. Interestingly, there also appears to be a slow shift in the composition of spending by day of week, toward Fridays and Saturdays and away from Mondays and Tuesdays. This pattern is likely capturing trends in the timing of shopping activity, though it may also be partly due to an unobserved change in the composition of merchants represented in our sample.

Working with the daily data series also allows us to uncover the recurring fluctuations in activity around major holidays, which play a critical role in seasonal adjustment. Moreover, the daily data reveal that spending patterns in the days surrounding major holidays can vary substantially depending on the day of the week that the holiday occurs. In future work, we intend to document these and other seasonal patterns revealed in the daily data and to discuss their implications for the construction of seasonally adjusted indexes at the daily or weekly frequency.

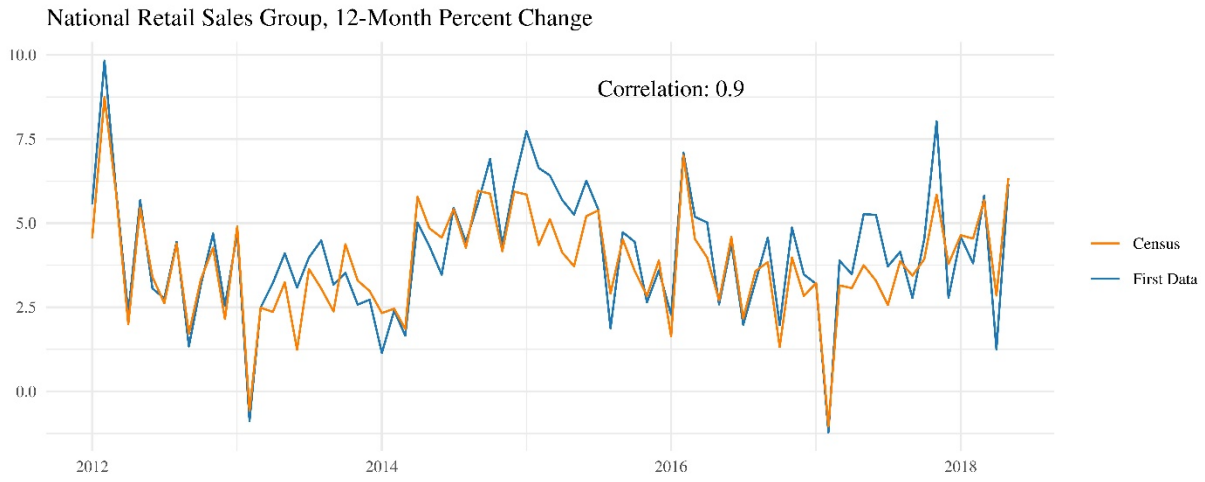
4 Comparisons of Spending Growth in New Series with Official Statistics

An important step in the development of our new spending indexes has been making comparisons to official estimates, specifically the Census Monthly Retail Trade Survey (MRTS). The MRTS is a comprehensive survey of retail sales at a monthly frequency and at the national

level. Since it is administered to firms with at least one retail establishment, it is a useful benchmark against which to compare the indexes that we derive from aggregating the First Data merchant-level data. The Census surveys roughly 13,000 firms monthly, with the full sample being re-selected every 5 years.[18] Firm births and deaths are incorporated quarterly.

Even if we have isolated the true signal for economic activity from First Data transactions, we would not expect a perfect correlation with the MRTS. In reality, the First Data transactions offer an independent, albeit noisy, signal of economic activity. Moreover, the Census estimates are also subject to measurement error, such as sampling error. Figure 8 shows the 12-month percent change in the national retail sales group from the First Data indexes and Census retail sales. Our spending indexes capture both a similar seasonal variation as well as the broad contour of growth rates over time.

Figure 8: National Retail Sales Group, 12-Month Percent Change



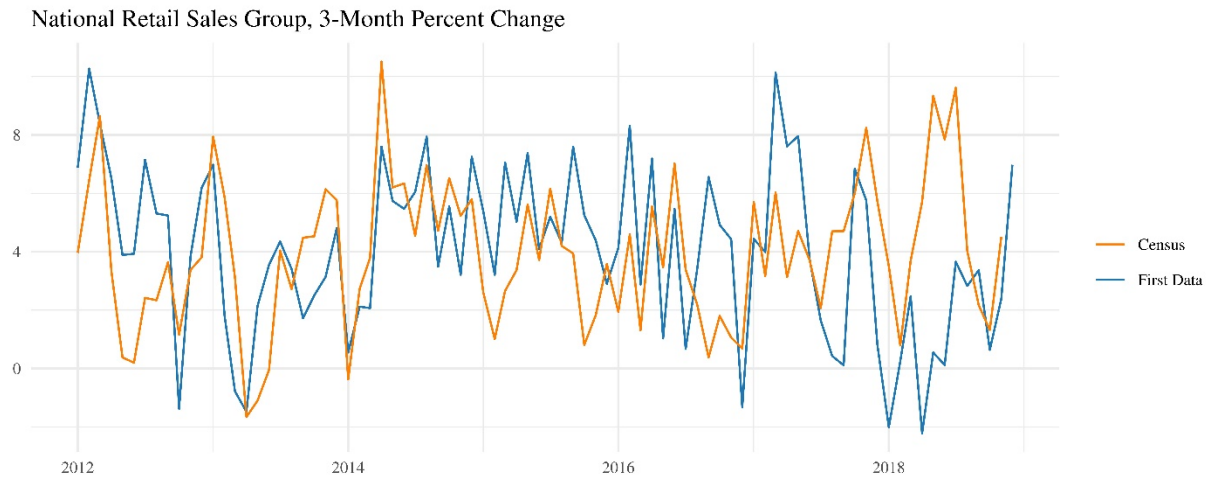
Source: Census and First Data, not seasonally adjusted.

Figure 9 shows 3-month percent changes in seasonally adjusted versions of both Census and First Data series. While the co-movement between the series is certainly weaker than the 12-

[18] The Census Bureau's initial estimate of retail sales for a month comes from the "Advance" Monthly Retail Trade Survey, which has smaller sample of firms, roughly 5,000. The results from the Advance survey are released for a specific month about two weeks after the month end. The MRTS for that same month is released one month later. Because firms are often delayed in their responses, the MRTS can undergo major revisions as additional firms report sales in subsequent months or in the annual retail sales survey, released each March.

month NSA changes in Figure 8, the broader contour of growth in the two series remain quite correlated even at a higher-frequency.

Figure 9: National Retail Sales Group, 3-Month Percent Change, SAAR



Source: Census and First Data, seasonally adjusted, annualized 3 month growth)

The results in this section have made us confident that we are, in fact, measuring monthly growth in consumer spending well. Furthermore, the signal derived from the First Data series provides a read on spending that is more timely than the official statistics. For any particular month, the initial reading on retail spending from First Data comes only three days after the completion of the month, while the Census's initial read lags by two weeks. Moreover, while the First Data series provides an independent read on retail spending, it also enhances our ability to forecast the final growth estimates published by Census, even when controlling for the preliminary estimates from Census.[19] This timeliness and incremental signal content allows policymakers—particularly the members of the Federal Open Market Committee deciding monetary policy—to base their decisions on a more accurate assessment of the current cyclical state of the economy.

[19] A regression of the final 3-month Census RSG growth rate on the preliminary 3-month Census RSG growth rate has an adjusted R-squared of 0.48, while the addition of the preliminary First Data series pushes up the adjusted R-squared to 0.55. While the incremental improvement in forecasting revisions in general is small, the First Data estimates are particularly helpful as an independent signal when Census preliminary estimates show an unusually large change in sales.

5 Application: Spending Effects of Hurricanes Harvey and Irma

Weather events, such as hurricanes, are a useful application of our new high-frequency, geographic spending indexes. The disruptions to spending during the storm are often severe but localized and short-lived, so that the lost spending is hard to quantify with monthly, national statistics. Moreover, the sampling frame used for a national survey may not be adequate to capture such geographic shocks to spending. In this section, we discuss our analysis of Hurricanes Harvey and Irma using our spending indexes.

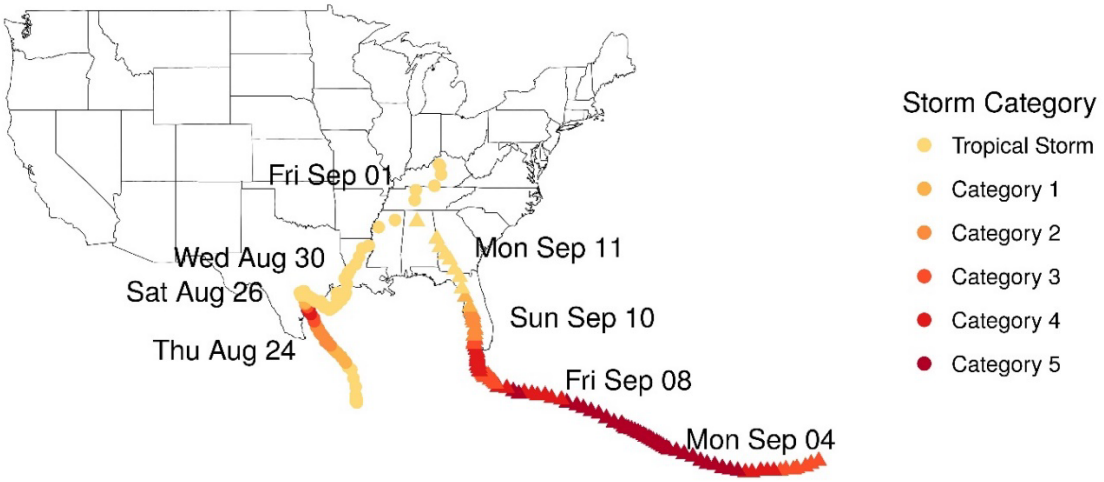
Our results reveal that the hurricanes significantly reduced consumer spending in the affected states in the third quarter of 2017. Although the level of spending after the storms quickly returned to normal, very little of the preceding shortfalls appear to have been made up in the subsequent weeks, suggesting that, on net over the span of several weeks, both hurricanes had negative effects on spending. The declines in activity were most apparent in the discretionary components of spending, such as restaurants, as opposed to necessities, such as grocery stores. We interpret the net negative effect on spending as implying that inter-temporal smoothing through temporary weather shocks is either incomplete or slow to occur.

5.1 Hurricanes Harvey and Irma

The 2017 Atlantic hurricane season was unusually active, with 17 named storms over a three-month period. Two of these hurricanes—Harvey and Irma—were especially large and severe. On August 28, Hurricane Harvey made landfall in Texas. Historic rainfall and widespread flooding severely disrupted life in Houston, the fifth largest metropolitan area in the United States. Less than two weeks later, Hurricane Irma made landfall in South Florida after causing mass destruction in Puerto Rico, and then proceeded to track up the western coast of the state, bringing heavy rain, storm surge, and flooding to a large swath of Florida and some areas of Georgia and South Carolina. By Monday September 11, 2017, more than 7 million U.S. residents of Puerto Rico, Florida, Georgia and South Carolina were without power. In Figure 10, panel A depicts the path of the two hurricanes and panel B the Google search intensity during the two storms.

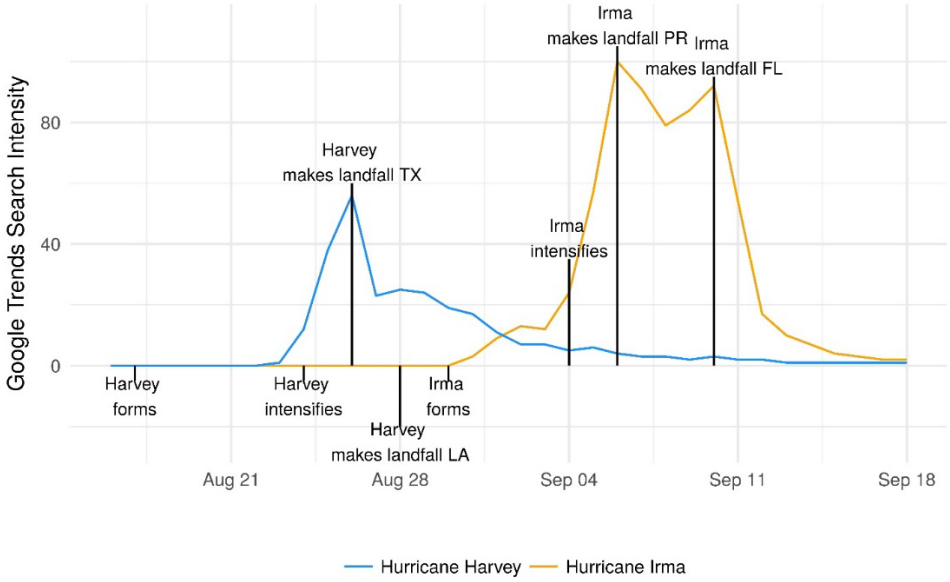
Figure 10. Path and Timing of Hurricanes Harvey and Irma

Panel A: Paths of Hurricanes Harvey and Irma



Source. NOAA

Panel B: Hurricane Timelines and Google Search Intensity



Source. Google Trends search intensity for the terms “Hurricane Harvey” and “Hurricane Irma.”

5.1 The Effects of Hurricanes Harvey and Irma on National and Local Spending

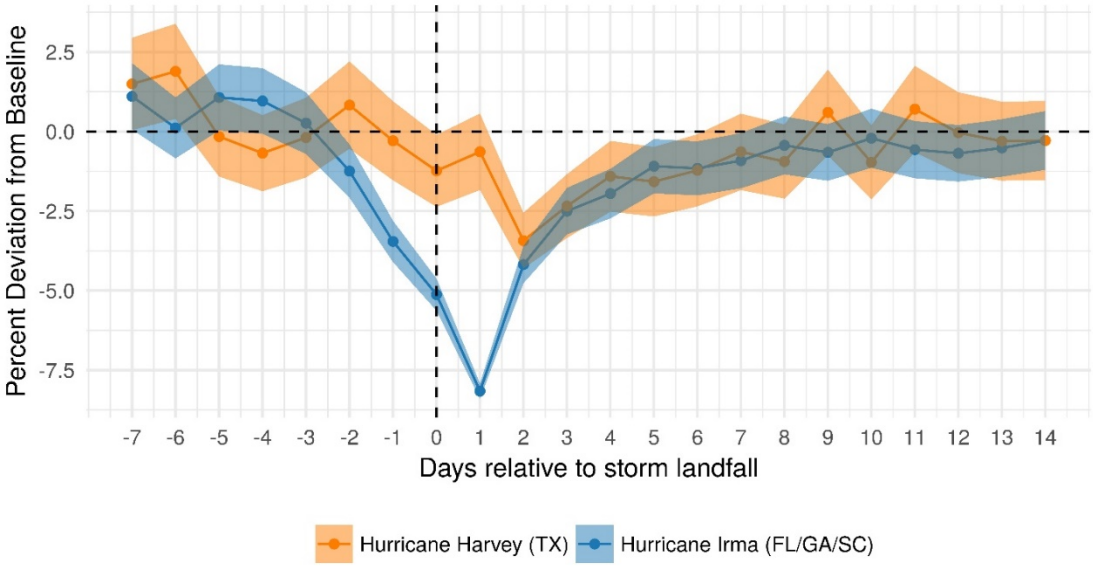
Using daily, state and MSA-level indexes, we examined the pattern of activity in the days surrounding the landfalls of Hurricanes Harvey and Irma. To quantify the size of the hurricane's effect, we estimated the following regression specification for each affected state:

$$\ln(Spending_t) = \sum_{i=-7}^{i=14} \beta_i * H_{t-i} + \sum_{w=Mon}^{w=Sun} \delta_w * I(Day_t = w) + \sum_{m=July}^{m=Nov} \delta_m * I(Month_t = m) + T_t + \varepsilon_t$$

The state-specific hurricane effects are captured by the coefficients on the indicator variables, H_{t-i}, which equal one if the hurricane occurred on day *t-i*, and zero otherwise. The regression also controls for variation in spending due to the day of week, the month of year, and a linear time trend (T_t). The coefficient β_0 is thus the estimated effect on (log) spending in that state on the day the hurricane struck.

Figure 11 illustrates the results of the regression for Hurricanes Harvey and Irma effects on national daily retail sales group spending. For this broad category of retail spending, there is little evidence of spending in advance of the storm. In the days following the landfall of Hurricane Harvey, daily retail sales group was about 3 percent lower than what normally would have occurred without a hurricane. In the case of Hurricane Irma, the disruption in spending was larger, reducing national retail sales group spending by more than 7 percent in the day after landfall. However, the level of spending rebounded quickly after the both hurricanes and within a week of landfall was back to normal levels. On balance, these data suggest that little of the reduced spending associated with Hurricanes Harvey and Irma was offset by higher spending in the days before or just after the storms.

Figure 11. Effects of Hurricanes on National Retail Sales Group Spending



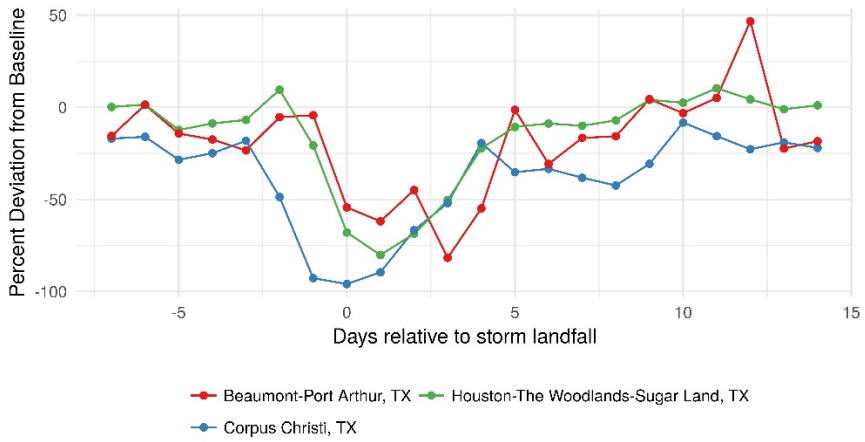
It is useful exercise to translate the daily effects on national spending to quarterly GDP growth. To roughly gauge the direct reduction in GDP, we first sum the percentage deviation from baseline in daily retail group spending from both hurricanes (in figure 11). We then divide this total by the 92 days in the quarter and scale the effects by the retail sales group's share of GDP (about 0.25). By this measure, we find that together both hurricanes reduced GDP growth by almost ½ percentage point (annual rate) in the third quarter of 2017. The gradual makeup, unlike the sharp drop on impact, is difficult to distinguish from the usual variability in daily spending, so our direct estimate may overstate the negative effect of the hurricanes. In addition, this estimate is derived only from behavior in retail sales group spending and therefore excludes other consumption, like recreation services, or unplanned inventory accumulation or other production disruptions, see also Bayard, Decker, and Gilbert (2017). Our spending indexes, albeit incomplete, may still be able to capture the GDP effects better than official statistics on retail sales. The national sampling frame of such survey measures may not measure localized shocks well.

In addition to tracking the effects of hurricanes on national spending, our new data set allows us to study local effects (figure 12). Both in Texas (panel a) and in Florida (panel b), the hurricanes brought spending in their direct path to a near halt. Daily, geographic data can trace out the economic effects of the hurricanes, and specific circumstances, such as evacuation orders, power outages, or flooding, with greater clarity than the national monthly statistics. With these

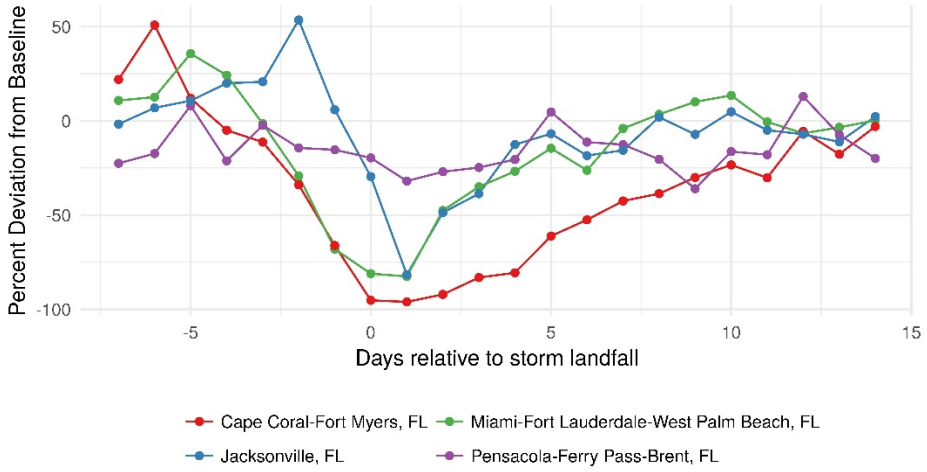
data it would also be possible to explore, possible shifts in spending to nearby areas and other spending categories, such as sales at gasoline stations or hotel accommodations, that are not included in the retail sales group.

Figure 12. Effects on Local Retail Sales Group Spending

Panel A: Houston and Texas Metros



Panel B: Miami and Other Florida Metros

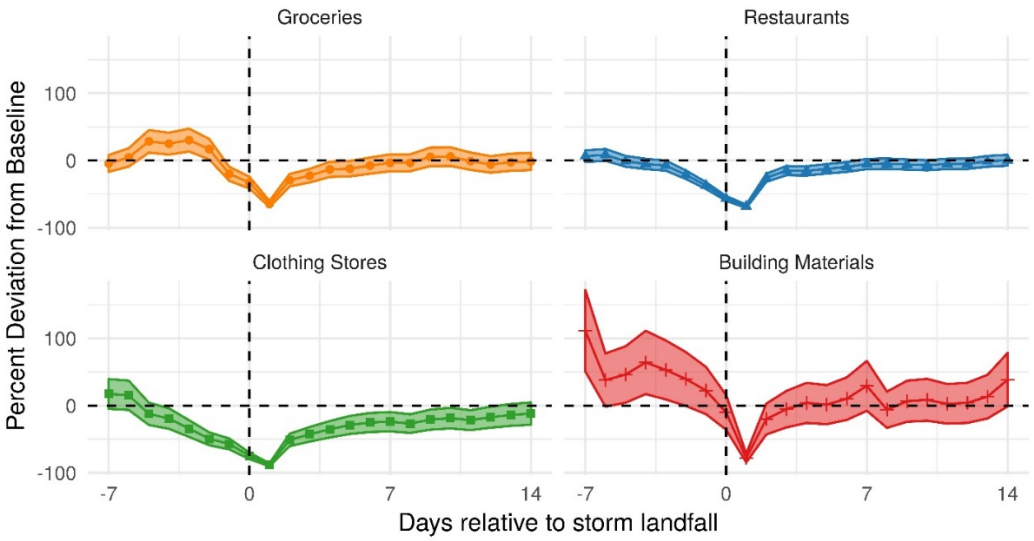


5.2 Using Hurricanes to Study Response of Consumers to Unanticipated Events

To further unpack this result, we also estimated the same regression using more detailed categories of spending in Hurricane Irma in Florida (figure 13). Interestingly, responses around the day of the Hurricane Irma varied noticeably among these categories. Spending at building materials stores actually ramped up before the hurricane and rebounded afterwards, such that the

net effect for this category is positive (12 percent for the month). Spending at grocery stores also ramped up before the hurricane but did not rebound afterwards so that the net effect was negative (-3.5 percent for the month). By adjusting the timing of purchases, consumers smoothed out the temporary disruption of the hurricane, with little effect on their overall groceries spending.

Figure 13. Effect of Hurricane Irma on Selected Components of Spending in Florida



However, other retail categories look quite different, showing no evidence of a ramp up in spending prior to the storm or a quick make-up in spending afterwards. In these cases, the spending lost during the storm appears to be largely foregone, at least in the near term. For example, our estimates indicate net reductions in spending in October due to the hurricane at restaurants (-9.5 percent) and clothing stores (-21 percent).

One possible explanation for the lack of a quick reversal in spending is that some purchases are tied together with time use. For example, going out to eat requires time spent at a restaurant. If the storm makes it more difficult to spend time on such activities, then individuals are likely to cut back on restaurant spending, and some may substitute to alternatives such as buying groceries to eat at home. In addition, purchases that are directly tied to an experience, such as an afternoon out with friends, may be foregone or postponed for some time. See also our related discussion of Hurricane Matthew in Aladangady et al (2016).

Another potential explanation for the apparent lack of make-up spending is that some portion of spending are "impulse purchases" that arise from a mood or temptation in the moment.[20] If bad weather disrupts a shopping trip or damps the mood of consumers, then these impulse purchases may never happen. Such psychological factors seem like a plausible explanation for the lack of make-up spending in several types of purchases, like clothing. Of course, we cannot rule out that the make-up in spending was gradual enough that the estimated effects in the days following the storm cannot be statistically distinguished from zero.[21] Furthermore, we cannot observe whether consumers make up spending in online sales rather than brick-and-mortar establishments. Even so, the transactions aggregates provide suggestive evidence that temporary disruptions like hurricanes can have persistent effects on some types of spending.

6 Conclusion

In this paper, we have presented our methodology for transforming transactions data from a large payment processing company into new statistics for consumer spending. Raw payment transaction volumes are clearly not suitable as statistics of overall consumer spending. This is because the true underlying trends in spending that policymakers and researchers care about are obscured by "merchant churn," namely the choices of individual merchants to start, end, or continue their contract with their payment processor.

Transforming the payments data into sensible measures of consumer spending required us to address a host of thorny measurement challenges. Most notably, unlike the Census Bureau's monthly retail trade survey, we do not have a sample frame and thus cannot distinguish actual merchant births and deaths from merchant churn. Thus, the filtering of transactions is an important step in the analysis of firm-sourced "big data." We are quite certain that the steps we took to address these challenges can be improved upon in the future, as researchers and statistical agencies continue to deepen their expertise in working with big data. At the same time, we are

[20] As some examples of related research, Busse, Pope, Pope, and Silva-Risso (2014) find that weather has a psychological effect on car purchases and Spies, Hesse, and Loesch (1997) argue that mood can influence purchases

[21] We also tested specifications that allowed for hurricane effects more than 7 days after the storm. The longer window did not materially change the results, and estimated coefficients for 7 to 21 days after the storm were not statistically different from zero.

confident that the spending series we have developed—in the span of just two years—have valuable economic content and will prove useful going forward.

At the monthly frequency, our series provides a timely and independent signal about the cyclical position of the economy. The series we produce are released a few weeks before their Census counterparts, and they are available when the Census estimates are delayed—as was the case during the recent government shut down. The First Data series are highly correlated with those published by Census, even though they are derived from a completely independent data source. Furthermore, despite the fact that the First Data series revise, we have shown that their preliminary estimates of monthly growth can be used to predict revisions to the Census monthly retail sales estimates. Finally, the First Data spending series provide a signal about spending that is independent from the Census series, which themselves are subject to sampling errors. In future work, once a longer time series for our data is available, we plan to develop spending estimates that extract and combine the signal content from both our own spending index and the Census survey, along the lines that Cajner et al (2018) have implemented for employment by blending the BLS's Current Establishment Survey data with a monthly payroll estimate that they derive from ADP microdata.

Beyond the monthly estimates, we see the timely higher-frequency, geographic detail in our spending series as the most unique contribution of our work. Census monthly statistics do not include any subnational detail, so for analysis of regional shocks, researchers and policymakers must rely on other data sources, such as the BEA's quarterly regional accounts, or household expenditure surveys like the Consumer Expenditure Survey. But most of these data sources have limited sample sizes at smaller geographies and are only available after a considerable lag.

Our spending estimates at the daily frequency and at detailed geographies can be used to examine a host of economic questions. As one example, in this paper we considered the effects of Hurricanes Harvey and Irma, and found that they significantly reduced consumer spending in the affected states in the third quarter of 2017. The declines in activity were most apparent in the discretionary components of spending, such as restaurants, as opposed to necessities, such as grocery stores. And although spending quickly returned to normal after the storms, very little of the preceding shortfalls were made up in the subsequent weeks. In separate work, our daily,

geographic series have also been used to study the spending effects of sales tax holidays and of delays in Earned Income Tax Credit refund payments.[22]

Looking ahead, there are a number of methodological next steps: Our efforts to refine the geographic detail of our estimates beyond the state level and for selected MSAs is ongoing, and we also hope to take up the challenge of seasonally adjusting our daily spending series. Another significant improvement to our current methodology would be to account for the effect of establishment births and deaths, possibly through the use of alternative data sources. The main disadvantage of the constant merchant methodology underpinning our analysis is that we cannot account for the effect of merchant births and deaths on spending since true births and deaths cannot be distinguished from merchant churn.[23]

Stepping back, and to conclude with some perspective on developing new economic statistics in the century ahead, we believe that our collaborative workflow setup, with researchers focusing on the economic statistics, software engineers handling computations on the raw data, and a private firm allowing the controlled access to its data could be a useful model for other big data projects going forward.

[22] See Aladangady et al (2016) and Aladangady et al (2018).

[23] In Appendix C, we demonstrate that "true" spending growth g_t is approximately equal to the growth rate derived from our constant merchant First Data series, $\hat{g}_t^{FD}$, plus the fraction of sales accounted for by true births less the fraction at true deaths: $g_t \approx \hat{g}_t^{FD} + b_t - d_{t-1}$. Thus, as long as true births and deaths are small and/or roughly offsetting, the constant merchant growth rate should do well at approximating total growth. However, a major concern is that shifts in *b-d* may occur at turning points.

References

Aladangady, Aditya, Shifrah Aron-Dine, Wendy Dunn, Laura Feiveson, Paul Lengermann, and Claudia Sahm (2016). "The Effect of Hurricane Matthew on Consumer Spending," FEDS Notes. Washington: Board of Governors of the Federal Reserve System, December 2, 2016, https://doi.org/10.17016/2380-7172.1888.

Aladangady, Aditya, Shifrah Aron-Dine, Wendy Dunn, Laura Feiveson, Paul Lengermann, and Claudia Sahm (2017). "The Effect of Sales-Tax Holidays on Consumer Spending ," FEDS Notes. Washington: Board of Governors of the Federal Reserve System, March 24, 2017, https://doi.org/10.170162380-7172.1941.

Aladangady, Aditya, Shifrah Aron-Dine, David Cashin, Wendy Dunn, Laura Feiveson, Paul Lengermann, Katherine Richard, and Claudia Sahm (2018). "High-frequency Spending Responses to the Earned Income Tax Credit," FEDS Notes. Washington: Board of Governors of the Federal Reserve System, June 21, 2018, https://doi.org/10.17016/2380-7172.2199.

Baker, Scott (2018). "Debt and the Response to Household Income Shocks: Validation and Application of Linked Financial Account Data," *Journal of Political Economy*. 126(4): 1504-1557.

Bayard, Kimberly, Ryan Decker, and Charles Gilbert (2017). "Natural Disasters and the Measurement of Industrial Production: Hurricane Harvey, a Case Study," FEDS Notes. Washington: Board of Governors of the Federal Reserve System, October 11, 2017, https://doi.org/10.17016/2380-7172.2086.

Busse, Meghan R., Devin G. Pope, Jaren C. Pope, and Jorge Silva-Risso (2015). "The Psychological Effect of Weather on Car Purchases." *Quarterly Journal of Economics.* 130(1): 371-414.

Cajner, Tomaz, Leland Crane, Ryan Decker, Adrian Hamins-Puertolas, Christopher Kurz, and Tyler Radler (2018). "Using Payroll Processor Microdata to Measure Aggregate Labor Market Activity," Finance and Economics Discussion Paper Series 2018-05. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2018.005.

Farrell, Diana and Fiona Grieg (2015). "Weathering Volatility: Big Data on the Financial Ups and Downs of U.S. Individuals," *JPMorgan Chase & Co. Institute*, May.

Federal Reserve Board (2018). "The Federal Reserve Payments Study: 2018 Annual Supplement," Washington: Board of Governors of the Federal Reserve System, https://www.federalreserve.gov/paymentsystems/fr-payments-study.htm.

First Data, *First Data Retail volume aggregates*, https://www.firstdata.com/en_us/home.html.

Galbraith, John W. and Greg Tkacz (2018). "Nowcasting with payments system data," International Journal of Forecasting, Elsevier, vol. 34(2), pages 366-376.

Gelman, Michael, Shachar Kariv, Matthew D. Shapiro, Daniel Silverman, and Steven Tadelis (2014). "Harnessing naturally occurring data to measure the response of spending to income." *Science*, 345(6193): 212-215.

Greene, Claire, and Joanne Stavins (2018). "The 2017 Diary of Consumer Payment Choice," *Federal Bank of Atlanta Research Data Reports* No. 18-5.

Leamer, Edward, (2014), "Workday, holiday and calendar adjustment: Monthly aggregates from daily diesel fuel purchases." *Journal of Economic and Social Measurement*, issue 1-2, p. 1-29, https://EconPapers.repec.org/RePEc:ris:iosjes:0005.

Matheny, Wendy, Shaun O'Brien, and Claire Wang (2016). "The State of Cash: Preliminary Findings from the 2015 *Diary of Consumer Payment Choice*." Cash Product Office, Federal Reserve System.

Mian, Atif, Kamalesh Rao, and Amir Sufi (2013). "Household Balance Sheets, Consumption, and the Economic Slump," *Quarterly Journal of Economics* 128(4): 1687-1726.

Roth Tran, Brigitte (2016). "Blame it on the Rain: Weather Shocks and Retail Sales." Working Paper.

Spies, Kordelia, Friedrich Hesse, and Kerstin Loesch (1997). "Store Atmosphere, Mood and Purchasing Behavior." *International Journal of Research in Marketing.* 14(1): 1-17.

Starr-McCluer, Martha (2000). "The Effects of Weather on Retail Sales." Finance and Economics Discussion Series, 2000-08. Board of Governors of the Federal Reserve System.

Appendix A: Mapping of MCC to NAICS for Retail Stores and Restaurants

MCC	MCC Description	NAICS2	NAICS3	NAICS	NAICS Description
5533	Automotive Parts, Accessories Stores	44	441	441310	Automotive Parts and Accessories Stores
5531	Automobile Supply Stores	44	441	441310	Automotive Parts and Accessories Stores
5996	Swimming Pools: Sales, Service, and Supplies	45	45	45	#N/A
5997	Electric Razor Stores: Sales and Service	45	45	45	#N/A
5998	Tent and Awning Shops	45	45	45	#N/A
5940	Bicycle Shops: Sales and Service	45	451	451110	Sporting Goods Stores
5941	Sporting Goods Stores	45	451	451110	Sporting Goods Stores
5970	Artist's Supply and Craft Shops	45	451	451120	Hobby, Toy, and Game Stores
5945	Hobby, Toy, and Game Shops	45	451	451120	Hobby, Toy, and Game Stores
5949	Sewing, Needle, Fabric, and Price Goods Stores	45	451	451130	Sewing, Needlework, and Piece Goods Stores
5733	Music Stores, Musical Instruments, Piano Sheet Music	45	451	451140	Musical Instrument and Supplies Stores
5942	Book Stores	45	451	451211	Book Stores
5994	News Dealers and Newsstands	45	451	451212	News Dealers and Newsstands
5735	Record Shops	45	451	451220	#N/A
10	#N/A	45	452	452111	Department Stores (except Discount Stores)
5311	Department Stores	45	452	452111	Department Stores (except Discount Stores)
5310	Discount Stores	45	452	452112	Discount Department Stores
5300	Wholesale Clubs	45	452	452910	Warehouse Clubs and Supercenters
5331	Variety Stores	45	452	452990	All Other General Merchandise Stores
5399	Misc. General Merchandise	45	452	452990	All Other General Merchandise Stores
5992	Florists	45	453	453110	Florists
5978	Typewriter Stores: Sales, Rental, Service	45	453	453210	Office Supplies and Stationery Stores
5943	Stationery Stores, Office and School Supply Stores	45	453	453210	Office Supplies and Stationery Stores
5947	Card Shops, Gift, Novelty, and Souvenir Shops	45	453	453220	Gift, Novelty, and Souvenir Stores
5932	Antique Shops	45	453	453310	Used Merchandise Stores
5931	Used Merchandise and Secondhand Stores	45	453	453310	Used Merchandise Stores
5937	Antique Reproductions	45	453	453310	Used Merchandise Stores
5995	Pet Shops, Pet Foods, and Supplies Stores	45	453	453910	Pet and Pet Supplies Stores
5971	Art Dealers and Galleries	45	453	453920	Art Dealers
9	#N/A	45	453	453930	Manufactured (Mobile) Home Dealers
5271	Mobile Home Dealers	45	453	453930	Manufactured (Mobile) Home Dealers
5993	Cigar Stores and Stands	45	453	453991	Tobacco Stores
5972	Stamp and Coin Stores: Philatelic and Numismatic Supplies	45	453	453998	All Other Miscellaneous Store Retailers
5974	#N/A	45	453	453998	All Other Miscellaneous Store Retailers
5973	Religious Goods Stores	45	453	453998	All Other Miscellaneous Store Retailers
5999	Miscellaneous and Specialty Retail Stores	45	453	453998	All Other Miscellaneous Store Retailers

(continued)

Source: Staff at the Census Bureau and the Bureau of Economic Analysis developed this mapping from MCC to NAICs. Other MCC/NAICs outside of retail stores and restaurants not shown here.

Appendix A - continued: Mapping of MCC to NAICS for Retail Stores and Restaurants

(continued)

MCC	MCC Description	NAICS2	NAICS3	NAICS	NAICS Description
5961	Mail Order Houses Including Catalog Order Stores, Book/Record Clubs	45	454	454113	Mail-Order Houses
5983	Fuel: Fuel Oil, Wood, Coal, Liquefied Petroleum	45	454	454311	#N/A
5960	Direct Marketing- Insurance Service	45	454	454390	Other Direct Selling Establishments
5962	Direct Marketing: Travel Related Arrangements Services	45	454	454390	Other Direct Selling Establishments
5967	Direct Marketing: Inbound Teleservices Merchant	45	454	454390	Other Direct Selling Establishments
5969	Direct Marketing: Not Elsewhere Classified	45	454	454390	Other Direct Selling Establishments
5422	Meat Provisioners: Freezer and Locker	45	454	454390	Other Direct Selling Establishments
5963	Door-to-Door Sales	45	454	454390	Other Direct Selling Establishments
4815	VisaPhone	45	454	454390	Other Direct Selling Establishments
5966	Direct Marketing- Outbound Telemarketing Merchant	45	454	454390	Other Direct Selling Establishments
5964	Direct Marketing: Catalog Merchant	45	454	454390	Other Direct Selling Establishments
5965	Direct Marketing: Catalog and Catalog and Retail Merchant	45	454	454390	Other Direct Selling Establishments
5968	Direct Marketing: Continuity/Subscription Merchant	45	454	454390	Other Direct Selling Establishments
5812	Eating places and Restaurants	72	722	722110	#N/A
5814	Fast Food Restaurants	72	722	722211	#N/A
5811	Caterers	72	722	722320	Caterers
5813	Drinking Places (Alcoholic Beverages), Bars, Taverns, Cocktail lounges, Nightclubs and Discotheques	72	722	722410	Drinking Places (Alcoholic Beverages)

Source: Staff at the Census Bureau and the Bureau of Economic Analysis developed this mapping from MCC to NAICs. Other MCC/NAICs outside of retail stores and restaurants not shown here.

Appendix B: Decomposing monthly growth rates of the series into a weighted average of the monthly growth rates from the contributing 14-month samples

Given the daily series, x_{it}, the monthly growth rates for the months in the middle of our sample can be derived as shown in the equation below:

$$1 + g_t = \frac{\sum_{i \in t} x_{it}}{\sum_{i \in t-1} x_{it-1}} = \frac{\sum_{j=0}^{13} f_{t+j} \sum_{i \in t} a_{it}^{t+j}}{\sum_{j=0}^{13} f_{t-1+j} \sum_{i \in t-1} a_{it-1}^{t-1+j}}$$

Define a_t^j to be the total sales in a 14-month sample j in month t, such that $a_t^j = \sum_{i \in t} a_{it}^j$. Furthermore, as in the previous section, define g_t^{t+k} to be the average monthly growth in time *t* within the 14-month series ending in *t+k* for *k≥0*, such that $g_t^{t+k} = \frac{f_{t+k} a_t^{t+k}}{f_{t+k} a_{t-1}^{t+k}} - 1$. For *k=-1*, we define $g_t^{t-1} = \frac{f_{t+13} a_t^{t+13}}{f_{t-1} a_{t-1}^{t-1}} - 1$, which is the monthly growth rate achieved from using the normalized monthly value for month *t* from the 14-month sample ending in time *t+13* and the normalized monthly value for month *t-1* from the 14-month sample ending in time *t*-1. We can then rearrange the above equation to show the monthly growth rate of our series is a weighted average of these monthly growth rates:[24]

$$g_t = \sum_{k=0}^{13} g_t^{t+k-1} * \frac{f_{t+k} a_{t-1}^{t+k-1}}{\sum_{j=0}^{13} f_{t+j-1} a_{t-1}^{t+j-1}}$$

The equation above is instructive as it shows us that the monthly growth rates derived from our daily index can be naturally interpreted as a weighted average of monthly growth rates for each

[24] For the 13 months at the beginning of our index and the 13 months at the end of our index, this equation will be slightly modified to account for the fact that there are fewer than 14 14-month samples that cover those months. The modified growth equations for these months can still be written as a weighted average of the growth estimates from the available 14-month estimates. Most notably, the growth rate for the last month of the series can be written as the average of two growth rates: $g_t = g_t^t * \frac{f_t a_{t-1}^t}{f_t a_{t-1}^t + f_{t-1} a_{t-1}^{t-1}} + g_t^{t,t-1} * \frac{f_{t-1} a_{t-1}^{t-1}}{f_t a_{t-1}^t + f_{t-1} a_{t-1}^{t-1}}$, where g_t^t is the growth rate between the last two months of the 14-month sample ending in t, and $g_t^{t,t-1}$ is the growth rate between the normalized sales in month t of the 14-month sample ending in t, and the sales in month t-1 from the 14-month sample ending in t-1.

constant merchant sample that contains those months (in addition to one final “faux” monthly growth rate using the first and last 14-month samples that contain those months).

Appendix C. Mathematical Derivation of Birth and Death Bias

The main disadvantage of the constant merchant methodology described above is that we cannot capture true births and deaths. To show the bias that may result, we introduce some notation. In a given month t let x_t be the total consumer spending in that month so that the true monthly growth rate of consumer spending is simply:

$$g_t = \frac{x_t}{x_{t-1}} - 1$$

Some set of firms transact in both period t and $t-1$ and we can call the spending at these firms in time *t*, s_t^- (where the minus denotes that these are the firms that existed in both that period and the previous one, so t and t-1) and, in time *t-1*, s_{t-1}^+ (where the plus denotes the firms that existed in both that period and the following one, so t-1 and t). The growth rate of spending for merchants who transact in both periods, what we will refer to as "constant merchant" growth, is simply:

$$\hat{g}_t = \frac{s_t^-}{s_{t-1}^+} - 1$$

However we know that in every period new establishments are born and we assume that they make up some fraction b_t of the sales in the previous period so that their total sales in the current period t are $b_t x_{t-1}$. Similarly some fraction, d_t, of total sales are by firms that die at the end of the period such that total sales in period $t-1$ can be expressed as:

$$x_{t-1} = \frac{s_{t-1}^+}{(1 - d_{t-1})}$$

And sales in period t can be written as:

$$x_t = s_t^- + b_t \frac{s_{t-1}^+}{(1 - d_{t-1})}$$

Assuming that births and deaths are a small fraction of the total spending in our sample we derive an approximate expression for total growth:

$$g_t = \left(s_t^- + b_t \frac{s_{t-1}^+}{(1-d_{t-1})}\right) / \left(\frac{s_{t-1}^+}{(1-d_{t-1})}\right) - 1$$

In simplifying this equation, we see that growth is approximately equal to "constant merchant" growth plus the rate of births minus the rate of deaths.

$$g_t = \left(\frac{s_t^-}{s_{t-1}^+}(1-d_{t-1}) + b_t\right) - 1$$

$$g_t \approx \hat{g}_t + b_t - d_{t-1}$$

The constant merchant methodology described in the previous sections yields an estimate of $\hat{g}_t$, using the constant merchants within the First Data platform. Thus, if we assume that the First Data merchant sample is close to representative, we see that "true" growth is approximately equal to the growth rate derived from the First Data, $\hat{g}_t^{FD}$, plus the true birth rate minus the true death rate.

$$g_t \approx \hat{g}_t^{FD} + b_t - d_{t-1}$$

Thus, the cost of the constant merchant methodology is that we are necessarily missing true births and deaths, but as long as they are small and/or roughly offsetting, the constant merchant growth rate would do well at approximating total growth. One particular concern is that shifts in *b-d* may occur at turning points.