This PDF is a selection from a published volume from the National Bureau of Economic Research

Volume Title: Big Data for Twenty-First-Century Economic Statistics

Volume Authors/Editors: Katharine G. Abraham, Ron S. Jarmin, Brian Moyer, and Matthew D. Shapiro, editors

Volume Publisher: University of Chicago Press

Volume ISBNs: 978-0-226-80125-4 (cloth), 978-0-226-80139-1 (electronic)

Volume URL: https://www.nber.org/books-and-chapters/big-data-twenty-first-century-economic-statistics

Conference Date: March 15-16, 2019

Publication Date: Februrary 2022

Chapter Title: From Transaction Data to Economic Statistics: Constructing Real-Time, High-Frequency, Geographic Measures of Consumer Spending

Chapter Author(s): Aditya Aladangady, Shifrah Aron-Dine, Wendy Dunn, Laura Feiveson, Paul Lengermann, Claudia Sahm

Chapter URL: https://www.nber.org/books-and-chapters/big-data-twenty-first-century-economic-statistics/transaction-data-economic-statistics-constructing-real-time-high-frequency-geographic-measures

Chapter pages in book: p. 115 – 145

4

From Transaction 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

4.1 Introduction

Access to timely, high-quality data is crucial for the ability of policymakers to monitor macroeconomic developments and assess the health of the economy. Consumer spending—70 percent of overall GDP—is key in policy deliberations about the economy. Existing official statistics on consumer spending are extremely useful, but they have limitations. For instance, the official retail sales data from the Census Bureau's surveys are only published

Aditya Aladangady is a principal economist in the Microeconomic Surveys Section of the Board of Governors of the Federal Reserve System.

Shifrah Aron-Dine is a graduate student in the Department of Economics at Stanford University.

Wendy Dunn is an adviser of the Program Direction Section of the Board of Governors of the Federal Reserve System.

Laura Feiveson is chief of the Household and Business Spending Section of the Board of Governors of the Federal Reserve System.

Paul Lengermann is assistant director of the Program Direction Section of the Board of Governors of the Federal Reserve System.

Claudia Sahm is a senior fellow at the Jain Family Institute. She is currently a contributing writer for *New York Times Opinion* and *Bloomberg Opinion*, and founder of Stay-At-Home-Macro Consulting.

We especially thank 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. For acknowledgments, sources of research support, and disclosure of the authors' material financial relationships, if any, please see https://www.nber.org/books-and-chapters/big-data-21st-century-economic-statistics/transactions-data-economic-statistics-constructing-real-time-high-frequency-geographic-measures.

for the nation as a whole and only at a monthly frequency.[1] The monthly figures are available two weeks after the end of the month and are subject to substantial revisions. Until recently, for analysis of regional shocks, researchers and policymakers had to rely on other data sources, such as the quarterly regional accounts from the Bureau of Economic Analysis (BEA), or household expenditure surveys like the Consumer Expenditure Survey. These more detailed data sources have limited sample sizes at smaller geographies and are only available a year or two after the fact. Our new real-time geographic data on spending data allow for better monitoring of shocks at the regional level and have the potential to serve as an early warning system to policymakers. 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 subsequent collapses in house prices. Our prior research shows other examples of how real-time geographic data are useful for studying economic events, such as Hurricane Matthew, sales tax holidays, and legislative hold on disbursement of Earned Income Tax Credit (EITC) in Aladangady et al. (2016, 2017, 2019, respectively).

The question motivating our research is whether alternative data sources can provide a timelier and more granular—but still reliable—picture of consumer spending. A promising new source of information on retail spending is the massive volume of data generated by consumers using credit and debit cards and other electronic payments.[2] Industry analysts and market researchers have long tapped into such transaction data to observe retail shopping behavior and market trends. Recently, economic researchers have also begun to use these and other nontraditional data, such as scanner data or online financial websites, in empirical studies of consumption.[3] These new data can offer timely and extremely detailed information on the buyers, sellers, and items purchased, yet they also pose myriad challenges, including protecting the privacy of individuals and businesses, ensuring the quality of the data, and adjusting for nonrepresentative samples.

In this project, we develop a comprehensive research dataset of spending activity using transaction data from First Data Merchant Services LLC (First Data, now Fiserv), a global payment technology company that processes $2 trillion dollars in annual card transaction volumes. We filter, aggre-

1. In September 2020, the Census Bureau began publishing 12-month percent changes (not seasonally adjusted) in state-level retail sales estimates. They used existing Census surveys as well as private Big Data sources. See for more details: https://www.census.gov/retail/state_retail_sales.html. The Bureau of Economic Analysis, the Bureau of Labor Statistics, and other statistical agencies have also begun using private data sources. Many of those efforts are detailed in this volume.

2. Moreover, cards—as we use in our new series—are now 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).

3. 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 et al. (2014) using data from apps used by households.

gate, and transform the card transactions into economic statistics. To protect the anonymity of all merchants and customers, we are restricted from accessing the transaction-level data. Instead, we worked with Palantir Technologies from 2016 to 2019—First Data's technology business partner—to build the new, fully-anonymized series to our specifications.[4] We currently have created estimates of daily retail spending from 2010 to the present for several 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 classification of the merchant making the sale. We adopt the same industry categories as the MRTS, which allows us to compare the national estimates from our new dataset to the corresponding Census estimates. However, an important difference in our approach is how we construct our sample. The Census Bureau uses a statistical sampling and survey design of tax records to select its 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 in the United States.[5] In contrast, First Data's client merchants that we use are not necessarily representative of all retailers, and some First Data client merchants do not permit us access to their data. In this paper, we describe the multi-stage process we developed to obtain high-quality, representative estimates of spending that are used for economic analysis at the Federal Reserve.

Despite being constructed from very different underlying raw data sources and methods, our new spending series and the Census retail sales data exhibit remarkably similar time-series patterns. The strong correlation of our new national series with the official statistics validates the soundness of our methodology and the reliability of our estimates. It showed that our new series was of high enough quality to use in policy analysis.

In this paper, we present two examples of how our new series could have been used to inform policy. First, we show how our series provided valuable insights on economic activity during the 2019 government shutdown, when the publication of official statistics was delayed. During a time of heightened uncertainty and financial market turbulence, it was crucial for policymakers to fill this information gap. Months before the Census data became available, we were able to see that spending slowed sharply early in the shutdown

4. Specifically, Palantir suppresses any spending estimate based on fewer than 10 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 analyses.

5. 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. wNote also that a merchant in First Data is similar conceptually to an establishment in Census.

but rebounded soon after, implying that the imprint of the shutdown on economic activity was largely transitory.

Second, we describe how we used the geographic detail in our daily data to track the effects of Hurricanes Irma and Harvey on spending. We showed 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 quickly returned to normal after the storms, very little of the lost activity during the storm was made up in the subsequent weeks. Thus, on net over the span of several weeks, the hurricanes reduced spending. This episode was an example of how it is possible to create reliable estimates of the effects of a natural disaster in real time.

The remainder of this paper is organized as follows. In section 4.2, we describe the transaction data from First Data. Section 4.3 details the methodology we use to construct our spending series from the raw transaction data. In section 4.4, we compare our new series with official estimates from the Census Bureau as a data validation exercise. Finally, in section 4.5 we show how we used the transaction data to track consumer spending during the government shutdown in early 2019 and in the weeks surrounding Hurricanes Harvey and Irma in 2017. Section 4.6 concludes.

4.2 Description of the Transaction Data

Our daily estimates are built up from millions of card swipes or electronic payments by customers at merchants that work with First Data. The total dollar amount of the purchase and when and where it occurred are recorded.[6] Only card or electronic transactions at merchants that work with First Data (or one of their subsidiaries) are included in our data. Cash payments as well as card payments at First Data merchant clients that do not allow further use of their data are also omitted. Geography of spending is determined by the location of the merchant, which may differ from the location of the purchaser.

First Data (now Fiserv) is a global payment technology company and one of the largest electronic payment processors in the United States. As of 2016, First Data processed approximately $2 trillion in card payments a year. First Data serves multiple roles in the electronic payments market. As a merchant acquirer, First Data sells 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 they are on First Data card terminals). Transactions at both types of merchant-clients are included in our data.

6. The name and zip code of the merchant are in the raw data. Bank Identification Numbers (BINs) 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.

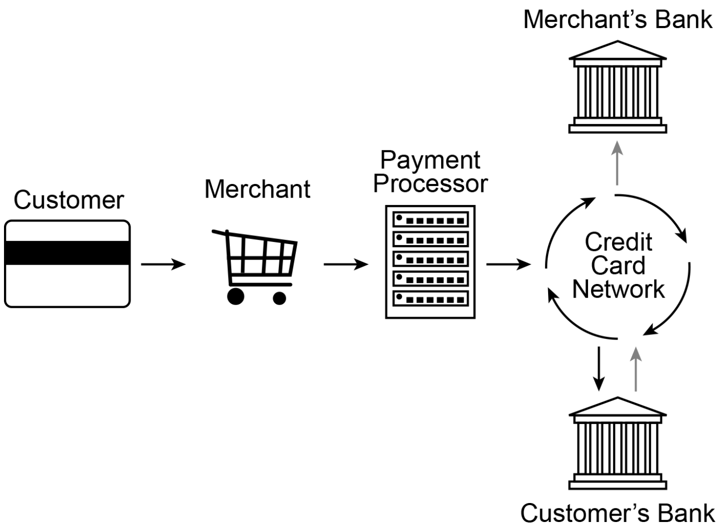

Fig. 4.1 The role of payment processors in credit card transactions

Figure 4.1 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 to authorize the transaction. That information is then relayed back to the point-of-sale system and the transaction is either approved or 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 authorization and settlement due to individual bank procedures. These two dates and the transaction amounts at authorization and settlement are in our data.[7]

7. 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 timestamp and the settlement amount. The settlement amount is more accurate than the authorization amount because it would include tips, which are typically not in the authorization amount. When available, we combine data from both authorizations and settlements to characterize each transaction. The date of the transaction is the timestamp 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.

First Data has details about every card transaction including the authorization and settlement amount and date, merchant address, merchant name, and merchant category code (MCC).[8] Even though First Data only covers a portion of purchases made with cards, the number of consumer spending transactions we observe with these data 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 of transactions is even higher. For example, it is nearly twice as high among groceries. (Cohen and Rysman 2013).

In this paper, we focus on a subset of First Data transactions at retailers and restaurants, which we refer to as the "retail sales group." The retail sales group is a key aggregate from the Census Bureau that the Federal Reserve and other macroeconomic forecasters track closely, because these data inform the estimates for about one third of personal consumption expenditures.[9] To create a comparable subset in our data, we map the available MCCs to 3-digit North American Industry Classification System (NAICS) categories in the Census data. We use a mapping tool developed by staff at the Census Bureau and the Bureau of Economic Analysis, shown in appendix A.

Because First Data has business relationships with merchants, not consumers, our data provide a merchant-centric view of spending. While technically a customer initiates a transaction and the data have an anonymized identifier for each credit and debit card, we do not observe the purchases that individuals make at merchants who are not in the First Data network. Moreover, we have information on merchants, not customers. Our merchant-centric orientation is the same as Census Retail Sales, which surveys firms. In contrast, other data sources on spending like the Consumer Expenditure Survey are household-centric. Both have advantages and disadvantages.

8. 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 station could have one MCC for terminals in the grocery and one for terminals at the gas pumps.

9. 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 the retail sales group: 444—Building Material and Garden Equipment and Supplies Dealers, 447—Gasoline Stations, 721—Accommodation, 713—Amusement, gambling, and recreation industries.

4.3 Methodology

In this section, we describe the methodology we instructed Palantir to use to filter, aggregate, and transform the raw transaction data into daily spending indexes for different industries and geographies. One of the major challenges with using nontraditional data like these for economic analysis is that we do not have a statistical sample frame. Our set of merchants is not representative of all US merchants, and it does not come with a well-established method to statistically reweight the sample, as in the Census survey. We had to develop new procedures that would yield usable statistics.

4.3.1 Filtering with 14-Month Constant-Merchant Samples

First Data's unfiltered universe of merchant clients and their associated payment transactions are not suitable, on their own, as economic statistics of retail spending. In the absence of a statistical sampling frame, the filtering of transactions is an important first step in the analysis of these nontraditional data. The filtering strategy is necessary to remove movements in the data resulting from changes in the First Data client portfolio, rather than those driven by changes in economic activity.

As shown in figure 4.2, there are vast divergences in year-over-year changes in the unfiltered sum of retail sales group transactions and in the equivalent Census series. The huge swings in the First Data series in 2014 and 2015 reflect their business acquisitions of other payment processing platforms. The unfiltered index of all merchants and all transactions includes the true birth and death of merchants; however, it also reflects choices by individual merchants to start, end, or continue their contract with First Data as their payment processor.

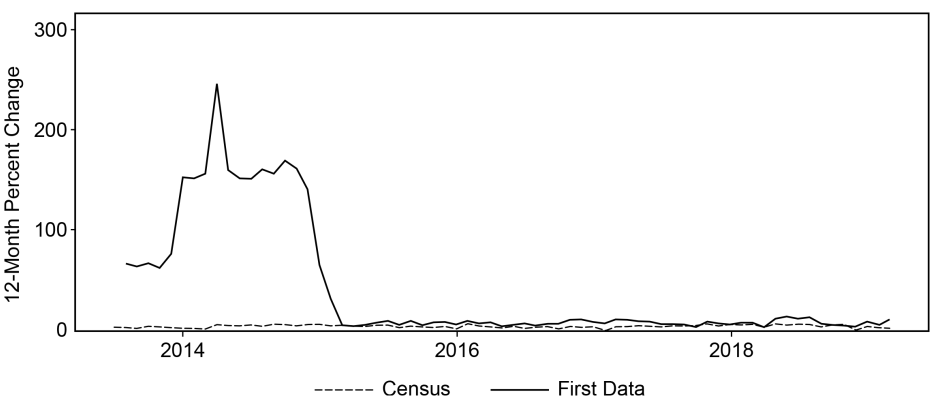

Fig. 4.2 Unfiltered sum of retail sales group transactions
Source: First Data and Census, authors' calculations.
Note: Not seasonally adjusted.

The first challenge for our filter is the considerable entry and exit of merchants in the transaction data. Some instances of this so-called merchant churn are to be expected and reflect economic conditions. For example, the decision to open a new business or to close an existing one is a normal occurrence that should be reflected in our statistics. In fact, the Census Bureau has adopted formal statistical procedures to capture these "economic births and deaths" in its monthly estimates of retail sales. Our unfiltered data include merchant churn based on those economic decisions; however, the data also include a large amount of merchant churn related to First-Data-specific business decisions, which should be excluded from our spending measures. Specifically, the decision of a merchant to contract with First Data as their payment processor should not be included in economic statistics. Given the rapid expansion of First Data over the past decade, client merchant churn is a big problem in the unfiltered data and must be effectively filtered from our spending series. To address this phenomenon, we developed a "constant-merchant" sample that restricts the sample to a subset of First Data merchants that exhibit a steady flow of transactions over a specific time period. Our method is aggressive in that it filters out economic births and deaths over that period, along with the First Data client churn. A future extension of our work is to create a statistical adjustment for economic births and deaths, but even without it, our current filter delivers sensible economic dynamics. Given the rapid expansion in First Data's business, and the economic growth in the retail sector overall, it would be far 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, would also be problematic because of strong seasonal and day-of-week patterns in retail spending.

To balance these tradeoffs, we combine a set of 14-month windows of constant-merchant samples. Each sample is restricted to include only those merchants that were "well-attached" to First Data (criteria described below) over the 14 months ending in the reference month of a given spending estimate. We need only 13 months to calculate a 12-month percent change but including an additional month at the start of the filtering window ensures that merchants who begin to register First Data transactions in the middle of a month do not enter the 12-month percent change calculations. We do not include a 15th month at the end of each window because it would delay our spending estimates for the most recent month and defeat a key purpose of making timely economic statistics available.

To give a concrete example—shown in the first row of figure 4.3—the constant-merchant sample of January 2017 is the subset of well-attached 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

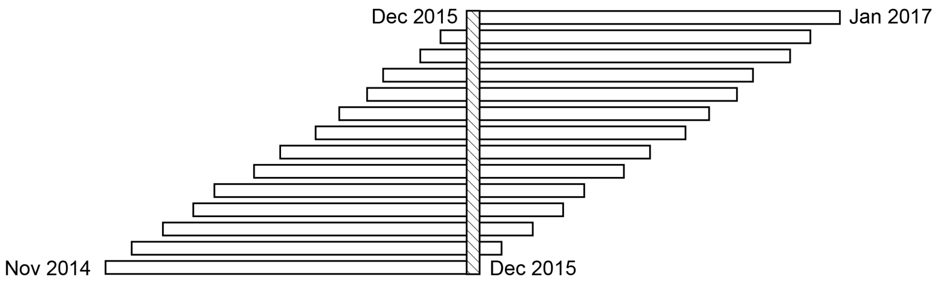

Fig. 4.3 Illustration of overlapping of 14-month constant-merchant samples

may appear in multiple overlapping monthly samples, but it will depend on the merchant's transaction behavior within each 14-month window.

An implication of this method of constructing 14-month constant-merchant samples is that, for any calendar month, we have multiple samples from which to estimate spending in a given reference month. For instance, the shaded area in figure 4.3 shows the 14 different merchant samples that we use to estimate spending in December 2015. The reference months for the constant-merchant samples shown in figure 4.3 range from December 2015 to January 2017. We discuss below how we combine the estimates across the separate merchant samples into a single time series. This overlapping sample methodology is applied independently to each 3-digit NAICS category and geography.

4.3.2 Additional Criteria for Selecting "Well-Attached" Merchants

We applied several other filtering criteria for selection into each 14-month constant-merchant sample:[10]

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 percent 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.

10. 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 merchants on the new platform can appear in the sample because merchants often exhibit volatile behavior in the data when a new platform comes online. Three small platforms with several data quality issues are dropped from our sample.

Table 4.1 **Filtering steps—14-month window ending January 2017**

Filter criteria applied in the step	Cumulative dollar volumes remaining (percent of raw sample)	Cumulative merchants remaining (percent of raw sample)
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 criterion in the 14-month window from December 2015 to January 2017.

3. *Minimum monthly spending/transaction days:* Merchants must transact more than four 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 percent of the distribution of growth rates of merchants at that NAICS 3-digit industry and geography combination.

Table 4.1 shows how our filtering techniques affect the number of First Data merchants and transactions in our series. Specifically, we report the fraction of spending removed from our sample in each filtering step for the 14-month window for January 2017. The denominator throughout is the unfiltered set of merchants in the retail sales group that do not have opt-out agreements with First Data. 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 unfiltered data, but it reflects a set of merchants with a stable attachment to First Data, and for whom sales growth appears well-measured by the data.

4.3.3 Combining Constant-Merchant Samples

After applying the filtering methods described above, we combine our adjusted 14-month constant-merchant samples to produce a daily index of spending growth and then monthly estimates of growth for each NAICS 3-digit industry and geography. The technical details here will be of interest to researchers who are applying our techniques to other data. For others, much of this section can be skipped. Since the transaction data at a specific merchant in our 14-month constant-merchant sample are daily, we cannot simply back out an index by cumulating the average monthly growth rates from our 14-month samples. That approach would have been the most natural if we were using monthly transaction data. Instead, for a given day we take a weighted average of the *level* across the 14-month samples that include that day. The weights remove level differences across the samples due to client-merchant churn. The result is a single, continuous daily index for each NAICS 3-digit industry and geography.

More precisely, we scale each successive 14-month sample by a factor, f_t, such that the average of spending over the first thirteen months of the series is equal to the average spending of those same thirteen months in the preceding, and already scaled, 14-month sample.[11] These factors are multiplicative; $f_t = \prod_{s=0}^{t-1} q_{t-s}$ where $q_t = (\sum_{k=1}^{13} \sum_{i \in t-k} a_{it-k}^{t}) / (\sum_{k=1}^{13} \sum_{i \in t-k} a_{it-k}^{t-1})$ and a_{it}^{t+j} denotes the estimate of daily sales on day i of month t from the 14-month sample series ending in month $t + j$. Then, we average together the 14 indexes that cover each day's spending to get our daily spending series:[12]

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

We obtain estimates of monthly growth from our daily indexes. See also appendix C.

In our method, each month's estimate relies on multiple constant-merchant samples, so the most recent month's estimate will revise as additional samples are added over time. Figure 4.4 shows the magnitudes of the revisions between the first growth estimate for a month (vintage 0) and its final estimate (vintage 13) when all the merchant samples are available. The dots and bars reflect the average revision at each vintage and its 90 percent confidence intervals. The revision is the final estimate of a month's growth rate (at vintage 14) minus the growth estimate at a specific vintage (from 1 to 13). The figure covers the period from April 2011 to December 2017. The range of revisions, particularly for the first few vintages, is high, with a 90 percent confidence interval of around plus or minus 0.8 percentage point. The average revision is near zero, so early estimates are not biased. It is worth noting that the preliminary estimates of monthly retail sales growth from Census have roughly comparable standard errors to our estimates.[13] As we make further refinements to our data estimation methods, we anticipate that the revision standard errors will shrink (for further details, see appendix).

In the final step, we create dollar-value estimates. Benchmarking is an important step when using a nonrepresentative sample and incomplete data. If some industries are over- or underrepresented among First Data merchants relative to all US merchants, or if use of noncard payments for spend-

11. Prior to this step, and as described in appendix B, we make a statistical adjustment to the first and final month of each 14-month sample. The adjustment attempts to correct bias due to our inability to perfectly filter new and dying merchants at the beginning and end of the sample. The notation for variable *a* in the equation above reflects the series after the correction has been applied.

12. 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.

13. 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.

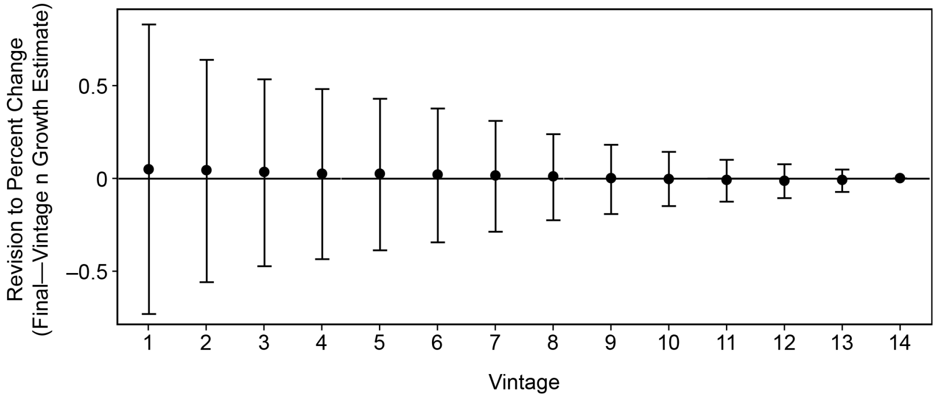

Fig. 4.4 Revision properties of First Data retail sales group monthly growth rates

Source: First Data, authors' calculations.

Note: Black dots show the mean revision to monthly seasonally adjusted growth rates, and bars show the 90% confidence interval; that is, 1.65 times the standard deviation.

ing differs across industries, a simple aggregation of our industry indexes would not accurately reflect overall growth.

The Economic Census—conducted every five years—is the only source of retail sales data with sufficient industry and geographic detail to serve as our benchmark. The most recent census available is from 2012. With each of our industry indexes for a specific geography, we set the average level in 2012 equal to the level in the Economic Census for that industry and geography.[14] We then use our daily indexes from First Data transactions to extrapolate spending from the Census level in 2012. Our final spending series in nominal dollars reflects the Census levels, on average, in 2012 and the First Data growth rates at all other times. This approach provides spending indexes in which the nominal shares of each industry are comparable to those across all US merchants, not just First Data clients. Then, to construct total spending indexes for the Retail Sales Group, or any other grouping of retail industries, we simply sum over the benchmarked industry indexes that compose the desired aggregate. We use this benchmarking procedure to create levels indexes for national-, state-, and MSA-level spending.

Prior to benchmarking, 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

14. For those geography-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.

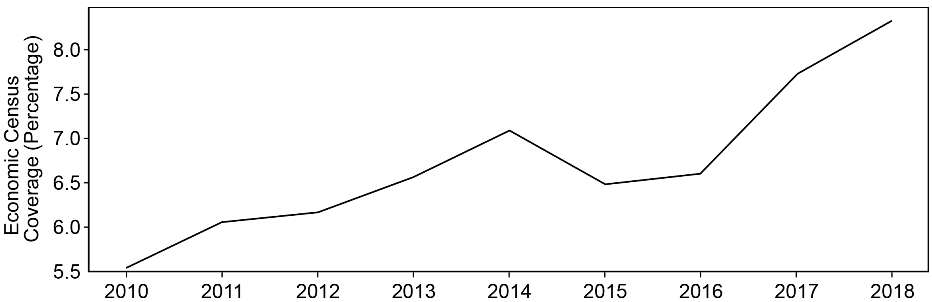

Fig. 4.5 First Data coverage of national retail sales group sales
Source: First Data and Census, authors' calculations.

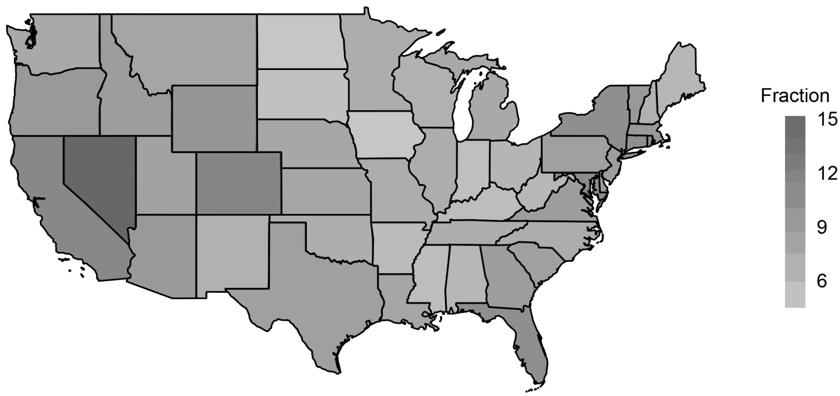

Fig. 4.6 First Data coverage of Economic Census retail sales group sales by state, 2018
Source: First Data and Census, authors' calculations.

the total First Data sales that are used in the creation of the index by the total estimated sales in the region.[15] Figure 4.5 shows that the coverage ratio of the national retail sales group has increased from roughly 5.5 percent in 2010 to 8.3 percent in 2018. However, the coverage is not uniform across the country. Figure 4.6 plots the coverage ratio of the retail sales group in each state in 2018. Some states, such as North Dakota and Iowa, both have low coverage at 3.7 percent, while others have higher coverage such as Nevada with 15.1 percent and Alaska (not shown) with 11.6 percent.

15. For years other than 2012, estimates from Economic Census for a specific industry and geography are grown out using national growth estimates for that industry from the Census Monthly Retail Trade Survey.

4.3.4 Seasonal Adjustment

In order to use our monthly spending indexes for time-series analysis, we also need to filter the indexes to remove regular variation related to weekdays, 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. We do not seasonally adjust our daily estimates; instead, we include day of the week and holiday controls when using them in analysis.[16]

4.4 Comparing Our Spending Measures with Official Statistics

An important step in the development of our new spending indexes has been making comparisons to official Census estimates of retail sales. Because the Census survey 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 reselected every five years.[17] 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 Census series. In reality, the First Transaction data offer an independent, albeit noisy, signal of economic activity. Moreover, the Census estimates are also subject to measurement error, such as sampling error. Figure 4.7 shows the 12-month percent change in the national retail sales group from the First Data indexes and Census retail sales. Our spending indexes and the Census

16. 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.

17. The Census Bureau's initial estimate of retail sales for a month comes from the "Advance" Monthly Retail Trade Survey, which has a 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.

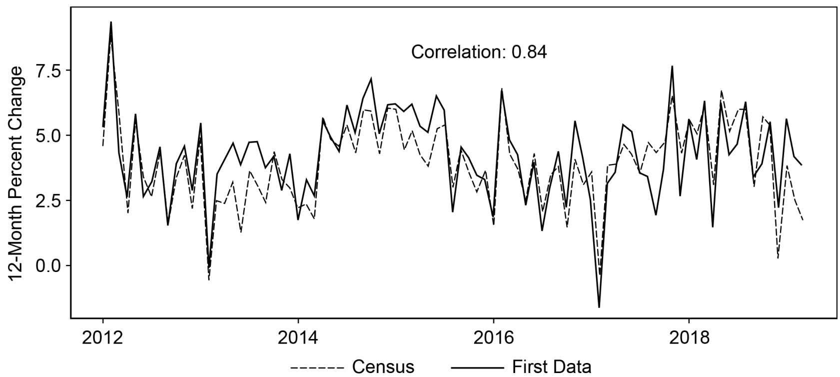

Fig. 4.7 National retail sales group (12-month percent change)
Source: First Data and Census, authors' calculations.
Note: Not seasonally adjusted.

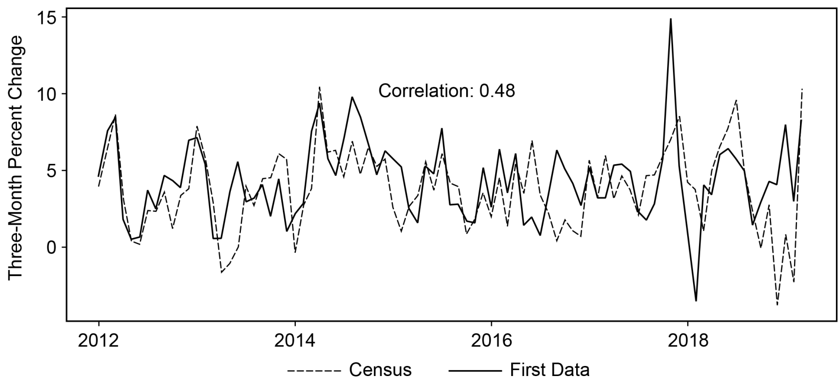

Fig. 4.8 National retail sales group (3-month percent change)
Source: First Data and Census, authors' calculations.
Note: Seasonally adjusted, annualized growth rate.

estimates clearly share the same broad contours, as one would expect from two noisy estimates of the same underlying phenomenon.

Figure 4.8 shows three-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-month NSA changes in figure 4.7, the broad contour of growth in the two series remains quite correlated even at a higher frequency. The standard deviation of the growth rates is also similar.

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 timelier 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. A regression of the final three-month Census retail sales group growth rate on the preliminary three-month Census growth rate has an adjusted R^2 of 0.48, while the addition of the preliminary First Data series raises the adjusted R^2 to 0.55. While the incremental improvement in forecasting revisions is small, the First Data estimates are particularly helpful as an independent signal when Census preliminary estimates show an unusually large change in sales. This timeliness and incremental signal content allow policymakers, such as 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.

4.5 Applications: Real-Time Tracking of Consumer Spending

The First Data indexes developed in this paper can improve the information set of policymakers, including at the Federal Reserve. In this section, we discuss how our First Data indexes helped policymakers during the partial government shutdown in 2019 and in the wake of Hurricanes Harvey and Irma in 2017.

4.5.1 The Partial Government Shutdown in 2019

In December 2018 and January 2019, heightened turmoil in global financial markets raised concerns about the pace of economic activity; as a result, policymakers were acutely focused on the incoming economic data to inform their decisions. Unfortunately, a government shutdown delayed the publication of many official statistics, including December retail sales—ordinarily one of the timeliest indicators of consumer spending—leaving policymakers with less information to assess current economic conditions.

The First Data spending index remained available during the shutdown. In contrast to the worrying signs in financial markets, the December reading from First Data indicated only a modest decline in retail spending, as shown in figure 4.9.

When the shutdown ended and Census published its first estimate of December retail sales (on February 14, a month later than usual), it showed an exceptionally large decline. At that point, however, the January First Data reading was also available, and it pointed to a solid rebound in spending. Indeed, the first Census reading for January also popped back up when it was eventually published on March 11.

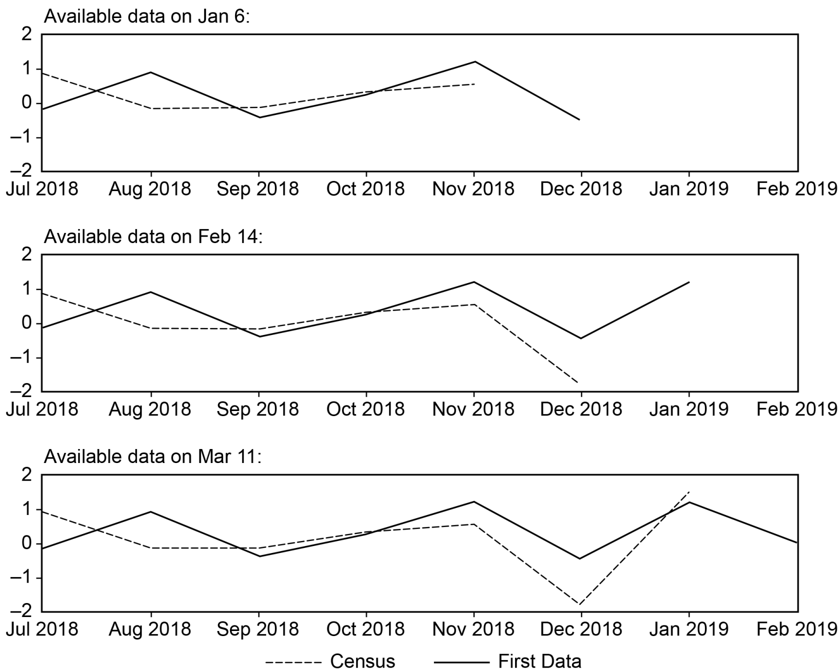

Fig. 4.9 Retail sales data releases during 2019 government shutdown
Source: First Data and Census, authors' calculations.
Note: Monthly growth rates of latest vintage available.

4.5.2 Hurricanes Harvey and Irma in 2017

Another useful application of our data is for assessing the impact of severe weather events, like hurricanes. The disruptions to spending during a storm are often severe but localized and short-lived, so that the lost spending is hard to quantify with monthly national statistics where the sampling frame may be inadequate to capture geographic shocks. Moreover, policymakers ultimately care about the extent to which swings in aggregate spending reflect the effect of a large, short-run disruption like a hurricane versus a change in the underlying trend in spending.

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 US residents of Puerto Rico, Florida, Georgia, and South Carolina were without power.[18] In figure 4.10, panel A depicts the path of the two hurricanes and panel B the Google search intensity during the two storms.

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 4.11 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 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.

It is a 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, shown in figure 4.11. 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

18. Because our data do not cover Puerto Rico, we could not conduct a comparable analysis of Hurricane Maria, which devastated Puerto Rico several weeks later.

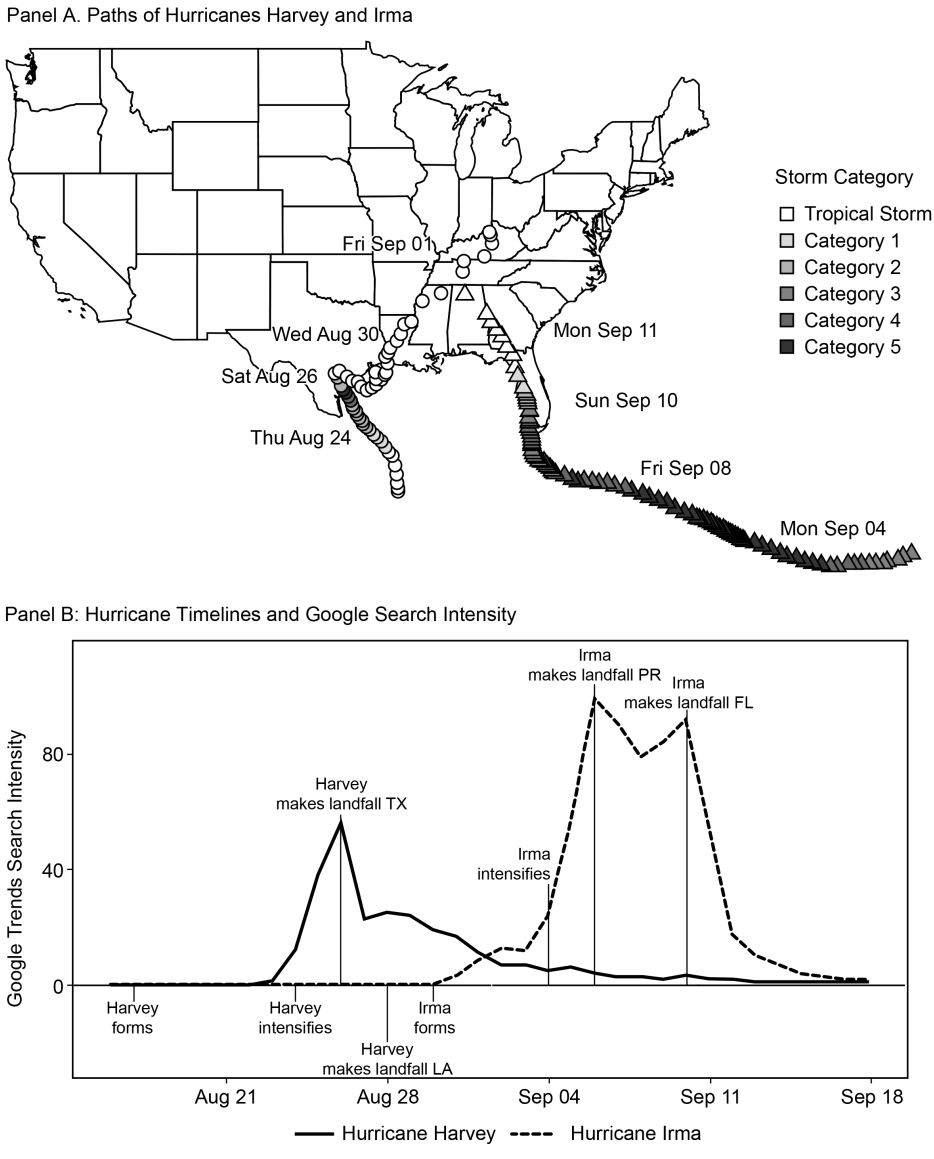

Fig. 4.10 Path and timing of Hurricanes Harvey and Irma

Panel A. Paths of Hurricanes Harvey and Irma

Source: National Oceanic and Atmospheric Administration.

Panel B. Hurricane timelines and Google search intensity

Source: Google Trends search intensity for the terms "Hurricane Harvey" and "Hurricane Irma."

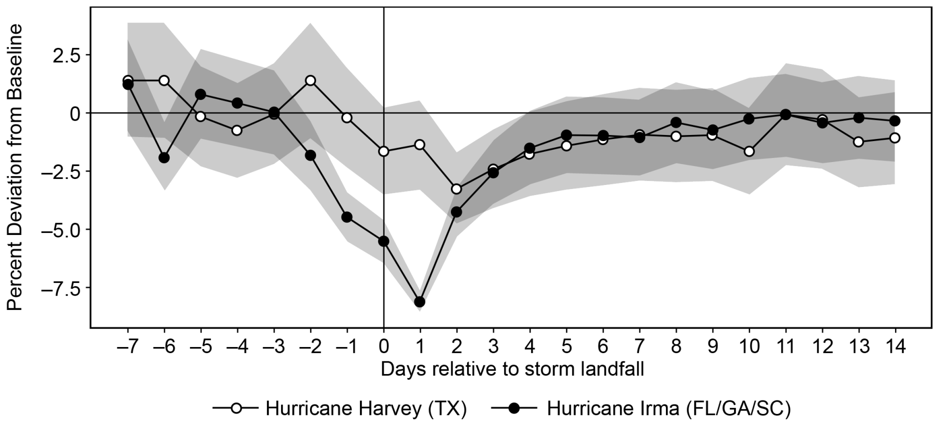

Fig. 4.11 Effects of hurricanes on national retail sales group spending
Source: First Data, authors' calculations.

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 dataset allows us to study local effects. As seen in figure 4.12, in both Texas (panel A) and 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, which are not included in the retail sales group.

To further unpack our results, we also estimated the same regression using more detailed categories of spending in Hurricane Irma in Florida (figure 4.13). Interestingly, responses around the day of 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

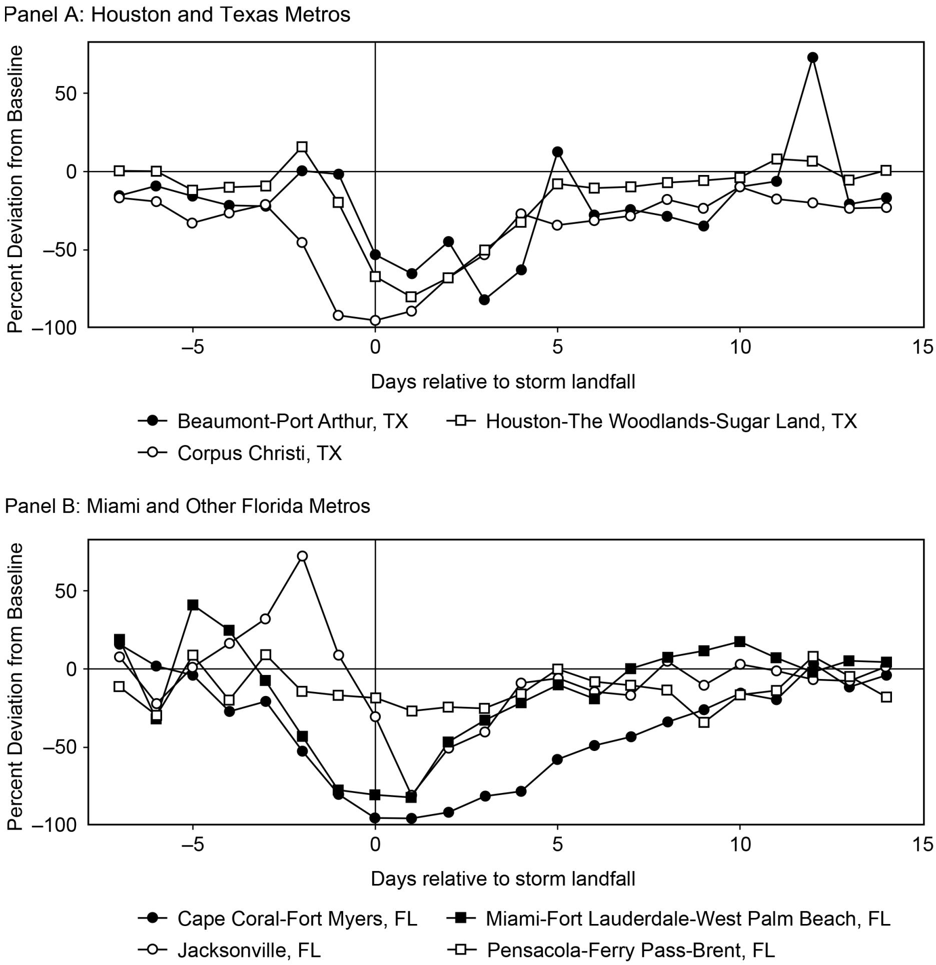

Fig. 4.12 Effects on local retail sales group spending
Source: First Data, authors' calculations.

the temporary disruption of the hurricane, with little effect on their overall grocery spending.

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 forgone, 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

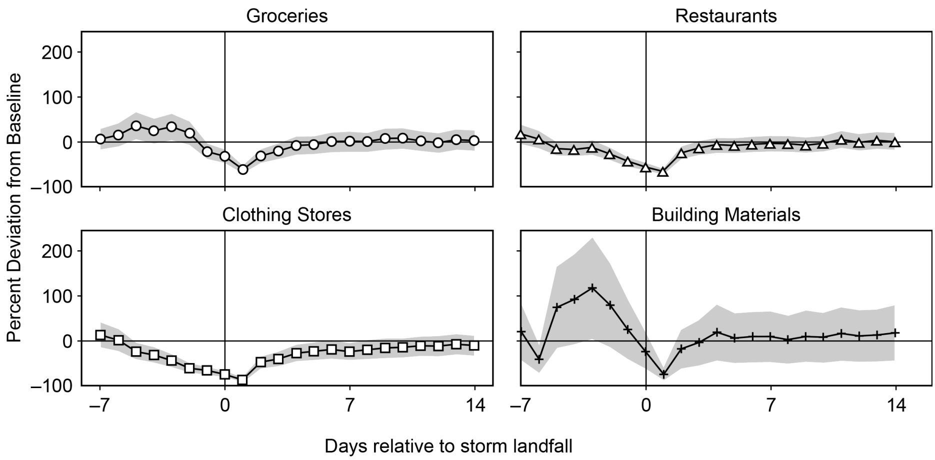

Fig. 4.13 Effect of Hurricane Irma on selected components of spending in Florida
Source: First Data, authors' calculations.

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 forgone 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 is "impulse purchases" that arise from a mood or temptation in the moment.[19] If bad weather disrupts a shopping trip or dampens 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.[20] Furthermore, we cannot observe whether consumers make up spending in online sales rather than brick-and-mortar establishments. Even so, the transaction aggregates provide suggestive evidence that temporary disruptions like hurricanes can have persistent effects on some types of spending.

19. As some examples of related research, Busse et al. (2015) find that weather has a psychological effect on car purchases and Spies, Hesse, and Loesch (1997) argue that mood can influence purchases.

20. We also tested specifications that allowed for hurricane effects more than seven 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.

4.6 Conclusion

In this paper, we present our methodology for transforming transaction data from a large payment processing company into new statistics of consumer spending. Raw payment transaction volumes are clearly not suitable and transforming payments data into sensible measures required us to address a host of thorny measurement issues. The steps we took to address these challenges can be improved upon; nevertheless, the spending series we developed have already proven to be a timely and independent signal about the cyclical position of the economy.

Our spending estimates at the daily frequency and at detailed geographies can be used to examine several economic questions. In this paper, we considered the high-frequency spending responses to Hurricanes Harvey and Irma. In other work, we used our series to study sales-tax holidays and delays in EITC refund payments.[21]

Looking ahead, we plan to refine our methodology. We would like to produce estimates for more detailed geographies, such as counties. With a longer time series, we will also be able to improve the seasonal adjustment of our spending series. Another significant improvement to our current methodology would be to account for establishment births and deaths (see appendix D).

To conclude with a broader perspective, we believe that nontraditional data can be used successfully to produce new economic statistics. In fact, several statistical agencies, including Census Bureau, the Bureau of Economic Analysis, and the Bureau of Labor Statistics are now using private Big Data to improve existing data series and expand their data offering. The collaborative efforts in our project—and by many other agencies detailed in this volume—with researchers focusing on the economic statistics, software engineers handling the computations with the raw data, and a private firm allowing controlled access to its data could be a useful model for other Big Data projects going forward.

Finally, we would note that the project discussed in this paper represents our third attempt over several years to obtain promising new data sources and use them to create spending statistics. Through earlier false starts, we learned valuable lessons about the many challenges that must be overcome to convert proprietary Big Data into functional economic statistics. This paper details the ingredients for our eventual success, including a private company supportive of our statistical efforts, skilled staff from a technology company to process the raw data, and rich data structured in a way that we could map to Census retail sales.

21. See Aladangady et al. (2016) and Aladangady et al. (2018).

Appendix A

Table 4A.1 **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 piece 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
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.
Note: Other MCC/NAICs outside of retail stores and restaurants not shown here.

Appendix B

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, $\underline{g}_t$ using all 14-month samples that include an estimate of the growth rate in t:

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

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

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

In figure 4B.1, we plot the distribution of deviations in all calendar months in the dataset, based on where the growth estimate falls in the merchant sample window (the index j).[22] The upward bias at the beginning of the 14-month sample—that is, the growth rate at time t for the sample that 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 that runs from $t - 13$ through t—is 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.

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

22. Figure 4B.1 shows the results for the national retail sales group, although the picture is similar for other NAICS codes and geographies.

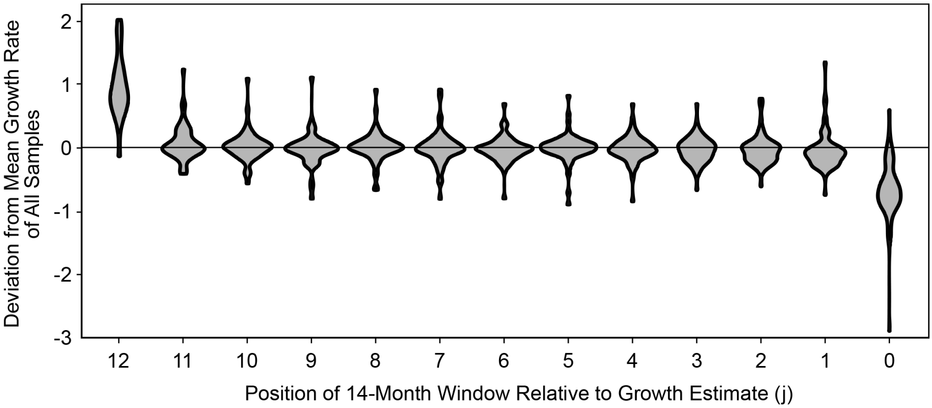

Fig. 4B.1 Deviation from mean growth in each month of the 14-month sample
Source: First Data.

of the year (m), we estimate a separate correction factor β^j_m 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:

$$\underline{g}_{t,m} = \beta^j_m g^{t+j}_{t,m} + \varepsilon_t.$$

The β^j_m applies a correction that results in adjusting up the growth estimates from the end of a 14-month sample and adjusting down the growth estimates 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:

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

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.

Appendix C

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 appendix B, define g_t^{t+k} to be the average monthly growth in time t within the 14-month series ending in $t + k$ for $k \geq 0$, such that $g_t^{t+k} = [(f_{t+k} a_t^{t+k}) / (f_{t+k} a_{t-1}^{t+k})] - 1$. For $k = -1$, we define $g_t^{t-1} = [(f_{t+13} a_t^{t+13}) / (f_{t-1} a_{t-1}^{t-1})] - 1$, 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:[23]

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

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 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 D

Mathematical Derivation of Birth and Death Bias

The main disadvantage of the constant-merchant methodology described above is that we cannot capture true economic births and deaths. To show

23. 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.

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 $b_2 x_{t-1}$. 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.

References

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, June 21. Washington, DC: Board of Governors of the Federal Reserve System. https://doi.org/10.17016/2380-7172.2199.

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, December 2. Washington, DC: Board of Governors of the Federal Reserve System. https://doi.org/10.17016/2380-7172.1888.

———. 2017. "The Effect of Sales-Tax Holidays on Consumer Spending." FEDS Notes, March 24. Washington, DC: Board of Governors of the Federal Reserve System. https://doi.org/10.170162380-7172.1941.

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): 1504–57.

Bayard, Kimberly, Ryan Decker, and Charles Gilbert. 2017. "Natural Disasters and the Measurement of Industrial Production: Hurricane Harvey, a Case Study." FEDS Notes, October 11. Washington, DC: Board of Governors of the Federal Reserve System. 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.

Cohen, Michael, and Marc Rysman. 2013. "Payment Choice with Consumer Panel Data." Research Department Working Paper Series No. 13–6, Federal Reserve Bank of Boston.

Farrell, Diana, and Fiona Grieg. 2015. *Weathering Volatility: Big Data on the Financial Ups and Downs of U.S. Individuals*. Washington, DC: JPMorgan Chase & Co. Institute. https://www.jpmorganchase.com/content/dam/jpmc/jpmorgan-chase-and-co/institute/pdf/54918-jpmc-institute-report-2015-aw5.pdf.

First Data. n.d. *First Data Retail Volume Aggregates*. https://www.firstdata.com/en_us/home.html.

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): 212–15.

Greene, Claire, and Joanne Stavins. 2018. "The 2017 Diary of Consumer Payment Choice." *Federal Bank of Atlanta Research Data Reports*, No. 18-5. https://www.atlantafed.org/banking-and-payments/consumer-payments/research-data-reports/2018/the-2017-diary-of-consumer-payment-choice.

Leamer, Edward. 2014. "Workday, Holiday and Calendar Adjustment: Monthly Aggregates from Daily Diesel Fuel Purchases." *Journal of Economic and Social Measurement* (1–2): 1–29. https://EconPapers.repec.org/RePEc:ris:iosjes:0005.

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

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