The Geography of Consumption and Local Economic Shocks: The Case of the Great Recession

Authors Abe Dunn and Mahsa Gholizadeh[1]
U.S. Bureau of Economic Analysis

Date October 2020

Abstract Geographic analysis of consumption is often constrained by geographic borders such as counties, but economic agents often cross borders to consume. Using unique card transaction data, we estimate across-county spending flows between firms and consumers for every county in the United States and for 15 industries to provide a new consumption-link across counties that has not been previously studied. To demonstrate the importance of this consumption link, we reexamine the 2007–2009 Great Recession following the work of Mian and Sufi (2013) and Mian, Rao, and Sufi (2014), who demonstrate that counties with the greatest decline in housing net worth also had the largest declines in consumption and employment. We show that the effect of the housing wealth decline crosses borders to affect consumption and employment in a pattern consistent with our spending flows, even for the non-tradable sector. We find that not accounting for cross-border effects tends to understate the impact of local housing wealth shocks on employment and spending by 26 and 17 percent, respectively; it also misallocates where those effects are occurring, by about 11 percent for both spending and employment.

Keywords Regional spending, regional employment, Great Recession, housing

JEL Code R12, R20, E20, E24

1. The authors would like to thank Christian Awuku-Budu, Mary Bohman, Ben Bridgman, Eva De Francisco, Lasanthi Fernando, Dennis Fixler, Kyle Hood, Matt Knepper, Justine Mallatt, and Scott Wentland for comments. We would also like to thank Fiserv for the use of their data and the substantial work of the employees at Palantir who helped manage and work with the enormous Fiserv database, especially Albert Altarovici, Brady Fowler, and Daniel Williams. We would also like to thank Ledia Guci for some preliminary analysis of Fiserv data for purposes of measuring regional consumption.

1. Introduction

Consumption accounts for around 70 percent of gross domestic product (GDP), but there is large variability in per capita consumption across the United States that also affects local labor markets. The ability to study consumption and employment at granular geographic levels is important. Changes or economic shocks in one local market may be distinct from other markets and these differences can be used to understand and test economic theories. However, at more granular geographies, consumers are more likely to leave geographic areas to consume. This fact can reduce the usefulness of extremely rich data sources that provide detailed and nearly complete coverage of both consumption (e.g., the Economic Census (EC)) and employment (e.g., the Quarterly Census of Employment and Wages (QCEW)) for every county in the United States , but these data sources are based on the location of the firm and not the location of the consumer.

To fully utilize these rich data sources, we introduce a complementary new data source on spending flows between consumer and firm locations for all counties in the United States, providing a new consumption link between counties. We show how this information may be used to improve the precision of the estimated effects from local economic shocks, and it also improves understanding of how and why different geographies are differentially affected by local economic shocks.

We construct the spending flow estimates using card transaction data from Fiserv, one of the largest card transaction intermediaries in the country, with well over $2 trillion in card volume going through their system worldwide annually. Typically, when a firm uses Fiserv services, all associated debit and credit card transactions go through their systems. At a micro level, these data include information about both the location of consumers' residence as well as the physical location of firms, allowing the measurement of cross-county spending flows. The data are aggregated and anonymized across firms and consumers by county and by three- digit North American Industry Classification System (NAICS) industry codes. While there are around 4.5 million establishments underlying the data, they still represent a sample of the total establishments in the country. These data are combined with EC data along with other sources to build representative estimates of spending flows across all counties in the United States for 15 three-digit NAICS categories for the year 2015. The focus of the analysis is on brick-and-mortar stores and excludes the non-store retail category that includes e-commerce firms such as Amazon and eBay.[2]

The NAICS categories we study account for a total of about 79 percent of consumer spending nationally, excluding housing, health care, and financial services. On average, we find that around 62 percent of expenditures take place in the same county in which individuals reside and that

2. Our paper is related to Dolfen, Einav, Klenow, Klopack, Levin, Levin and Best (2019) that uses detailed VISA data on consumer location and spending habits across locations to assess the gains in e-commerce. They find large gains from the introduction and expansion of e-commerce. In contrast, our paper focuses more explicitly on brick-and-mortar stores for two reasons. First, the coverage of our data set is more complete and accurate for brick-and-mortar stores. With additional data, the basic approach laid out in our paper could be adapted to e-commerce sales. Second, the analysis in our paper focuses on the period during the Great Recession when e-commerce was a much smaller share of spending.

about 80 percent of spending occurs within a 100 mile radius of the home county. While these statistics show that spending typically occurs near where individuals reside, spending outside the home county still makes up a substantial share of total spending and may vary greatly depending on the local geography and industry. This turns out to be extremely important for some industries, such as accommodations, where only 12 percent of spending occurs in a person's home county, but less important for other industries, such as food and beverage stores, where over 75 percent of spending takes place in the home county.

We demonstrate the importance of these cross-border effects in two ways. First, we show how these spending flows form a basic part of regional accounting. The total consumption of individuals that reside in a county equals total final consumption sold in that county minus net exports of consumption (i.e., the total amount sold by firms to individuals outside of the county, minus total amount consumers purchase outside of the county in which they reside). We form a simple empirical test of this accounting relationship and find evidence that this relationship holds in the data and has significant explanatory power. Moreover, we show that the across-county spending flows estimated for 2015 are relevant throughout the period from 2002 to 2017. This spending flow information and analysis provides an important step toward creating more detailed regional and local economic accounting.

Next, we apply the across-county consumption flows to re-examine the effects of housing wealth declines from the 2007–2009 Great Recession. In particular, we follow the well-known work of Mian, Rao, and Sufi (2013) and Mian and Sufi (2014) to study how local changes in housing wealth affect consumption and local employment. Mian, Rao, and Sufi (2013) find that housing wealth declines at the county level have a significant negative effect on consumer spending. Rather than addressing the cross-border issue they turn to an alternative card transaction data source that contains information on spending based on the location of the consumer. In contrast, our paper starts with spending estimates from official sources that are centered around the location of a firm and considers the housing wealth of all consumers, including both local consumers and those traveling from other counties, in determining the effect of housing wealth declines on firm revenue. The across market flow estimates provide detailed information regarding the location of potential customers across areas. We find that firms are affected in proportion to the change in housing wealth of their customers, even if their customers reside in another county.

In a follow up paper, Mian and Sufi (2014) show significant negative effects on employment in those counties with the largest decline in net housing wealth. In their paper, there is no adjustment for the cross-border effects of spending on employment. In our paper, we show that the same spending flows that impact the amount of final consumption sold by firms also affects employment, even for the non-tradable sector. Overall, the main point of our paper is not to challenge the results of Mian, Rao, and Sufi (2013) and Mian and Sufi (2014), as our estimates confirm their main findings. Instead, the housing wealth shock is used in our paper to demonstrate the importance of cross-border spending flows, which link the effects of the housing wealth decline across counties.

Our elasticity of housing wealth change to spending is 0.19, which implies a marginal propensity to consume out of housing wealth (MPCH) of 7.6 cents on the dollar, although the estimate is around 6.4 cents on the dollar if spending flows are not accounted for.[3] These estimates are quite close to the estimates of Mian, Rao and Sufi (2013) who find a MPCH of 7.2 cents on the dollar, although we expand their analysis from 900 counties to over 3,000 counties in the United States and use alternative methods for constructing spending estimates from official sources. When employment is used as the dependent variable we find an elasticity of 0.15, which if used as a proxy for MPCH, would indicate an estimate of 5.9 cents on the dollar, which is comparable to Mian and Sufi (2014) who find an MPCH between 4.1 and 7.3 cents on the dollar based on estimates using employment.[4]

We find spending flows are important for obtaining the appropriate measure of the housing wealth change relevant to firms. We show that firms in high consumption export counties, those counties with higher levels of consumption from outside the county, are relatively unaffected by housing wealth changes within their own county, but are instead affected by housing wealth changes from the export counties where their customers' reside (e.g., Clark County, Nevada). Alternatively, those counties with low consumption exports are unaffected by housing wealth changes in other counties, and are only affected by housing wealth changes in their own county. Ignoring spending flows across counties tends to reduce the elasticity of housing wealth changes on spending and local employment.

Our results are robust to a number of alternative specifications. Other robustness checks include both panel regression models and instrumental variable estimates around the Great Recession following Guren et al. (2020). These robustness checks also show that ignoring across-county flows tends to understate the effects of the decline in housing wealth.

Similar to previous studies in this literature, we find that counties with the greatest drop in net housing wealth show the largest declines in consumer spending and local employment. However, we show that this effect is not isolated to county borders. We show that not accounting for these cross-border effects leads to an underestimate of the effects of the housing net wealth shock on both spending and employment by around 26 percent and 17 percent, respectively, and a misallocation of where these effects occur of around 11 percent for both spending and employment.

3. As discussed later, the MPCH is determined by dividing the spending elasticity by the ratio of housing wealth to personal consumption.

4. The calculation using employment as a proxy for consumption assumes a one-to-one relationship between employment and consumption.

2. Data

The card transaction data source used in this paper is from Fiserv, a card transaction intermediary, which processes transactions for establishments around the world, including credit, debit, and prepaid gift cards that includes all types of card transactions (e.g., Visa, MasterCard, Discover and others).[5] The unit of observation on the Fiserv system is a single transaction at a firm. Once a firm signs up for Fiserv services, typically all card transactions go through the Fiserv system. However, we do not see the data at this level of detail. Fiserv works with a company, Palantir, which is a software company that specializes in the management and analysis of big data. Fiserv and Palantir have aggregated and anonymized transaction data to the county level in a way that provides detailed and meaningful economic information, while still protecting the identity of both firms and individuals. The data contain millions of firms and transactions that span all states in the United States and the District of Columbia. The data includes transactions from e-commerce (primarily captured in NAICS category 454 for non-store retailers), but the coverage for this category is relatively poor, so we exclude e-commerce firms.

For counties within the United States, the home location of each cardholder is estimated based on the transaction history of the card using information on all transactions across all industries. The home location algorithm is optimized based on a subset of cards within the Fiserv database, where the home location of the cardholder is known.[6]

The Fiserv data we use is from 2015 and includes aggregate county-level information by three- digit NAICS industry. For every county-industry combination, the data contains an estimate of the share of revenues for establishments in that county coming from consumers residing in one of the more than 3,000 counties in the United States. For instance, this data includes information on the share of accommodation revenues (NAICS 721) in Clark County, Nevada (i.e., Las Vegas) coming from Orange County, California. The total shares across all areas add up to one. Our study focuses on 15 select three-digit industries that have good coverage in the Fiserv data. These select industries account for 64 percent of personal consumption spending excluding housing and financial services. It accounts for 79 percent of consumer spending if health care is also excluded.[7]

To protect the anonymity of firms and consumers in the Fiserv data, information on the transaction flows across geographies are suppressed in some cases. This is especially common in areas where

5. Other electronic card transactions are also included, such as Electronic Benefit Transfer.

6. As an additional check on the home-location algorithm, we also have a version of the data based solely on those consumers for whom the home location is known. This data are also similarly aggregated and anonymized to the county level. We find the two estimates of spending flows to be quite similar.

7. The 15 select industries account for 41 percent of total consumption, including all consumption categories.

revenues for the industry for a particular county are small.[8] Using information from the EC, we find that about 15 percent of spending is suppressed for these select industries.

For those county-industry pairs with suppressed flow information, we apply flexible models based on observable transactions in the database across areas to generate estimates of transaction flows across all areas in the country. To impute spending flows, information for those industries in which transactions are observed in a county (e.g., the category restaurants and bars (NAICS 722), where 98 percent of spending is unsuppressed), combined with information on distances traveled, revenues estimated based on the EC, and other covariates to impute the remaining spending flows. For instance, if we are missing accommodations flows in an area, but we observe flows of restaurant services, we can use information on the restaurant service flows between areas, combined with information on how far individuals typically travel to purchase accommodation services, as well as other information such as population and revenues, to impute the flows for accommodation services between two areas. We have explored a variety of flexible models to impute this missing information and selected our current specification using a holdout sample and cross-validation. We chose the method with the lowest mean squared error in our holdout sample. Additional details of this imputation method are described in the appendix.

We also form estimates of spending and employment. For the employment data we use the QCEW, which is a data source that includes quarterly employment and wage estimates for 95 percent of jobs at the county level and by detailed NAICS industry category. The source of the QCEW is administrative data from state unemployment insurance programs. While nearly all employment is included, it excludes select areas such as proprietors and the self-employed. QCEW is the same data source used by Guren et al. (2020). However, our version of the QCEW data includes complete coverage of all counties at the three-digit industry level from 2002 to 2017.[9]

For the spending estimates, we use the Geographic Area Series of the 2002, 2007, 2012, and 2017 ECs that contains information on revenues and establishment counts by NAICS industry and county-level geographies.[10] Next, to estimate spending for all of the intercensal years, we use the QCEW growth rates to interpolate county-level growth rates by NAICS. Specifically, the annual QCEW growth rates are rescaled by the ratio of the annualized 5-year EC growth rate and the annualized 5-year QCEW growth rate. This method essentially anchors the annual growth rates in

8. The specific rule is that there needs to be 10 or more firms in that three-digit NAICS, with no firm having more than a 20 percent market share. In addition to these criteria, some merchants have agreements with Fiserv to “opt out” of their data being used and their data are not included.

9. Mian and Sufi (2014) use County Business Patterns data from Census, which also provides information on employment and earnings. The CBP data are annual and QCEW data are quarterly, and there are also slight differences in coverage. Overall the two sources are similar for the industry categories studied here.

10. A subset of counties in the EC contain suppressions at the three-digit industry level, representing about 1 to 2 percent of spending. The estimates for suppressed counties are imputed using state-level EC data and QCEW data to create estimates for all counties in the United States for these benchmark years.

QCEW wages to match the average growth rate in the EC (see the appendix for additional details). A similar method is applied in the Bureau of Economic Analysis (BEA) regional economic accounts and private sector organizations such as Moody's and the Survey of Buying Power, as historically there is a high correlation between the growth rate in the EC and wages from the QCEW. In the appendix, we show that wage data performs quite well in predicting growth rates in revenues based on the EC years. Guren et al. (2020) also use employment data as a proxy for changes in spending, which we agree is a good proxy. However, we view growth rates from the QCEW as distinct from our spending estimates, as our spending estimates are anchored to the EC around the Great Recession for the years 2007 and 2012. This distinction appears to matter for our estimates, as we generally find higher elasticities for spending than for employment. Additional details of our county estimates of spending are outlined in the appendix.

Another important data set used in our analysis is from Zillow, which contains home pricing information from 1996 to January of 2020 on a monthly basis for more than 2,000 counties.11 The remaining counties are relatively small rural counties with relatively little economic activity. For our key analysis, we focus on the change in home prices at the end of 2006 to the beginning of 2009, which we calculate directly with the Zillow data, similar to Mian, Rao, and Sufi (2013) and Mian and Sufi (2014). For the missing counties, we assume the price decline is equal to the median price decline across counties in the same state. While this is a strong assumption, these are very rural counties and this has very little effect on the estimates and allows us to examine effects of housing wealth declines across all counties. The measure of housing wealth decline is calculated as:

$$\Delta HNW_i = \frac{P_{h,i}^{2009} - P_{h,i}^{2006}}{P_{h,i}^{2006}}$$

where ΔHNW_i is the change in housing wealth computed by the change in housing price from December 2006 to the end of 2009, where $P_{h,i}^{t}$ is the housing price for county i in year t. The Zillow data is also used later in the paper to help form an instrumental variable for the housing wealth change following Guren et al. (2020). Details of this strategy are discussed in the robustness section of this paper.

11. The data was downloaded from www.zillow.com/research/data.

3. Descriptive Statistics

Table 1 shows descriptive statistics for industry NAICS total estimated spend in 2015, in which the total is decomposed into the percent of spending that is observed, the percent of spending that was imputed, and the percent that could not be imputed. Across market spending flows are observed for about 86 percent of spending, therefore no additional imputation is required. About 14 percent of the flow shares are imputed using the method described previously. For less than 0.1 percent of spending, it was not possible to impute the flows across areas. The amount of imputation needed varies greatly by industry. For food service and drinking places (NAICS 722) we observe 98 percent of spending flows, but we observe just 63 percent for performing arts, spectator sports, and related industries.

Table 1. Spending by Industry

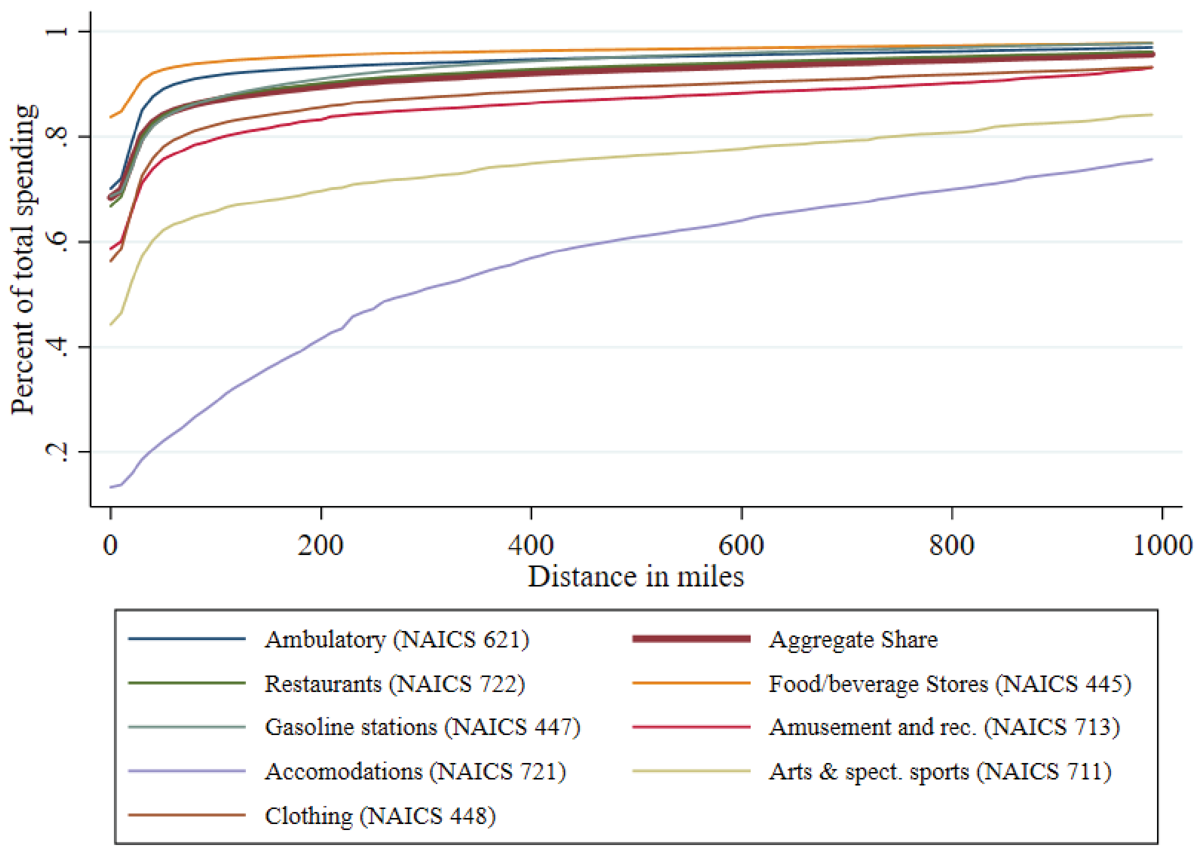
	Total $ millions	% Observed	% Imputed	% Unknown
Accommodation (NAICS 721)	225,765.3	85.59	14.34	0.07
Ambulatory health care services (NAICS 621)	960,110.3	95.74	4.21	0.04
Amusement, gambling, and recreation industries (NAICS 713)	119,829.8	86.72	13.21	0.07
Building material and garden equipment and supplies dealers (NAICS 444)	341,689.4	70.10	29.76	0.14
Clothing and clothing accessories stores (NAICS 448)	232,950.2	95.87	4.11	0.02
Food services and drinking places (NAICS 722)	660,300.4	98.31	1.68	0.01
Food and beverage stores (NAICS 445)	720,160.9	87.92	12.04	0.04
Furniture and home furnishings stores (NAICS 442)	126,712.8	82.53	17.34	0.13
Gasoline stations (NAICS 447)	523,039.2	84.00	15.94	0.06
General merchandise stores (NAICS 452)	749,349.3	66.71	33.16	0.13
Miscellaneous store retailers (NAICS 453)	138,279.0	95.36	4.61	0.03
Performing arts, spectator sports, and related industries (NAICS 711)	104,468.5	63.40	36.53	0.07
Personal and laundry services (NAICS 812)	110,372.0	94.68	5.20	0.12
Repair and maintenance (NAICS 811)	181,223.3	89.66	10.23	0.11
Sporting goods, hobby, book, and music stores (NAICS 451)	103,789.1	82.05	17.86	0.09
Total	**5,298,039.5**	**85.99**	**13.95**	**0.07**

Notes. The total spending estimates are based on our estimate of total spending at merchants in each county for 2015. The share imputed is based on total revenues in the counties in which spending flows are not observed for a three-digit NAICS category.
NAICS North American Industry Classification System.

3.1 Geography of Consumption

The amount individuals travel to consume varies greatly by industry. Figure 1 shows the cumulative distribution of spending by NAICS for the first 1,000 miles away from a firm's home location, where the location within each county is based on the population centroid.[12]

Figure 1. Cumulative Distribution of Spending by Distance, Truncated at 1,000 Miles



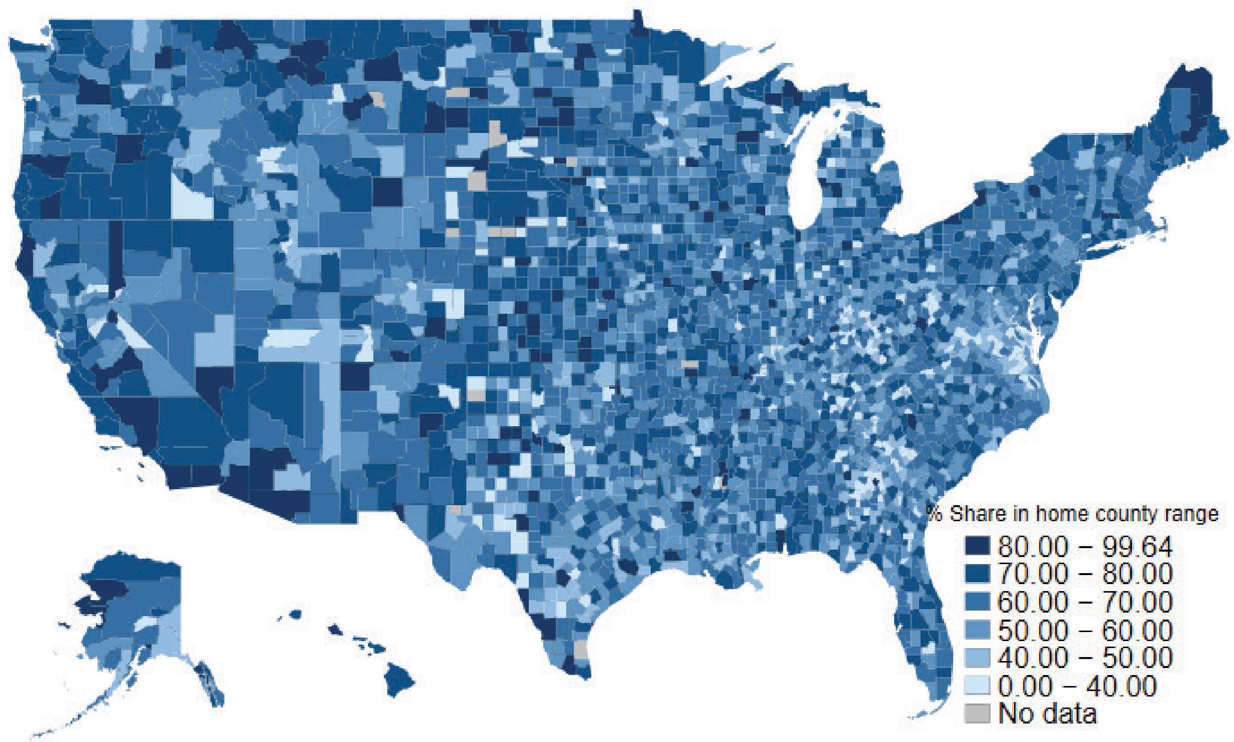
Notes: The cumulative spending is calculated for each NAICS category based on the total share of spending occurring within a distance radius of the merchants location where the location in each county is determined by the population centroid.
NAICS North American Industry Classification System.

Categories such as food and beverage stores, health care, and restaurants are among those in which most consumption occurs locally. The finding that preferences for food and beverage stores (i.e., grocery stores) is highly localized relates to the literature on food deserts and local availability on consumption (Allcott et al. 2019). In contrast, people tend to travel farther for arts and spectator sports, and accommodations.

12. We truncated the distribution at 1,000 miles to better highlight the differences across industries.

While there is considerable variation across industries, both the geography of different locations as well as the concentration of different industries and populations across the United States leads to large variation in how much consumers spend outside of the county in which they reside. Figure 2 shows the share of consumption that is consumed in a consumer's home county, with darker shades indicating that more consumption is occurring in the home county. Figure 2 shows that for most counties more than 50 percent of consumption occurs in the home county, and this is particularly true in large cities.[13] In contrast, in more rural areas consumers tend to travel to consume.

Figure 2. Share of Consumer Consumption in Home County



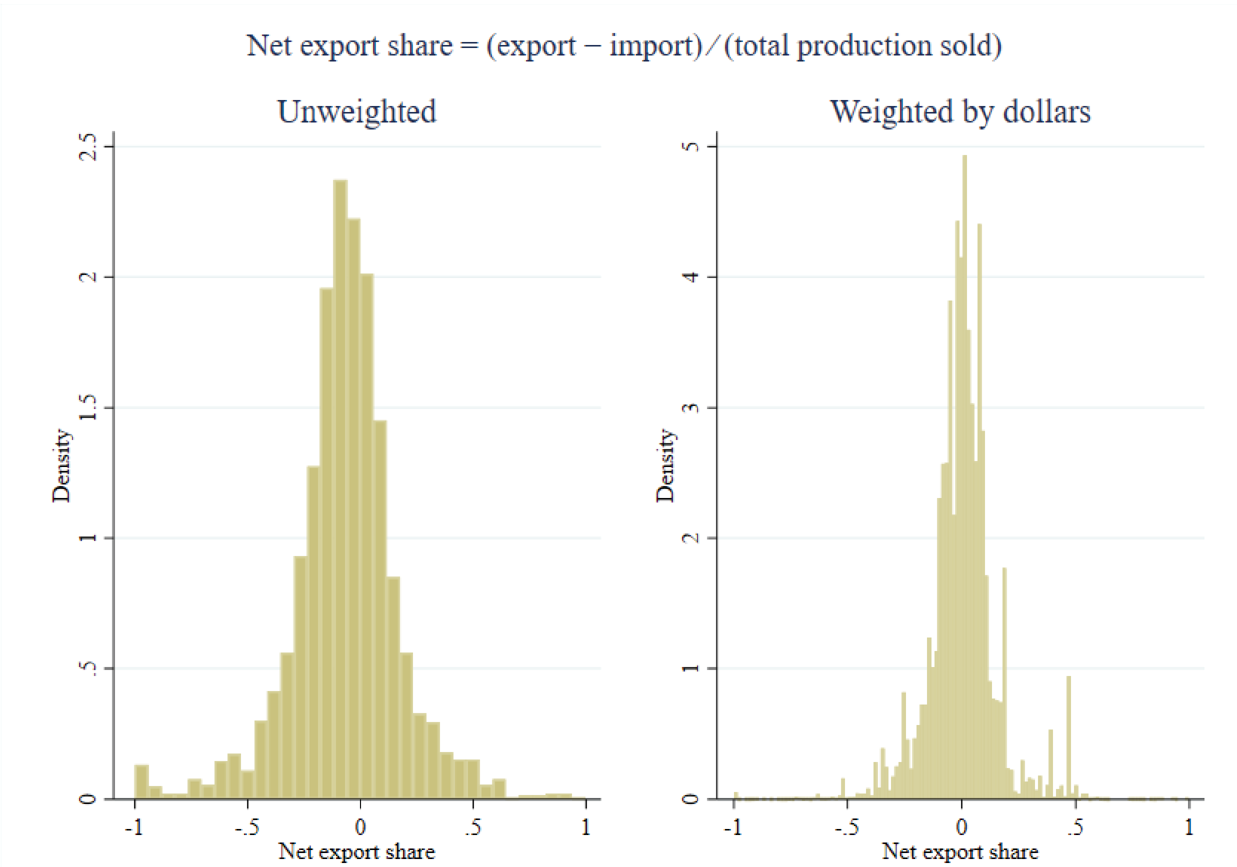
Notes: For each county we use all 15 NAICS categories and spending flow estimates for all counties to calculate the total spending by consumers in their home county and the total spending by consumers across all counties. We then take the ratio of home county consumption to total consumption.

Counties may differ greatly in how much spending flows into it from other locations and how much flows out as consumers purchase goods and services in counties outside of their home county. The net difference may not be symmetric. We summarize the share of net flows by calculating the total exports (i.e., firm revenues from consumers outside of the county), minus imports (i.e., the total amount of revenue from consumers leaving the county), divided by the total amount of final consumption sold in the county. Figure 3 shows the distribution of net exports across the United

13. We construct the same figure based on firm revenue share occurring in the home county, but it essentially shows the same pattern, in which most consumption occurs locally in more populated areas of the country.

States, both unweighted and weighted by the final consumption sold in the county. The distribution has been winsorized at –1 to avoid the long tail of rural counties that import most of their consumption. Figure 3 shows a large variation across the United States, especially for more rural counties, which are more represented in the unweighted distribution.

Figure 3. Distribution of Net Export Share

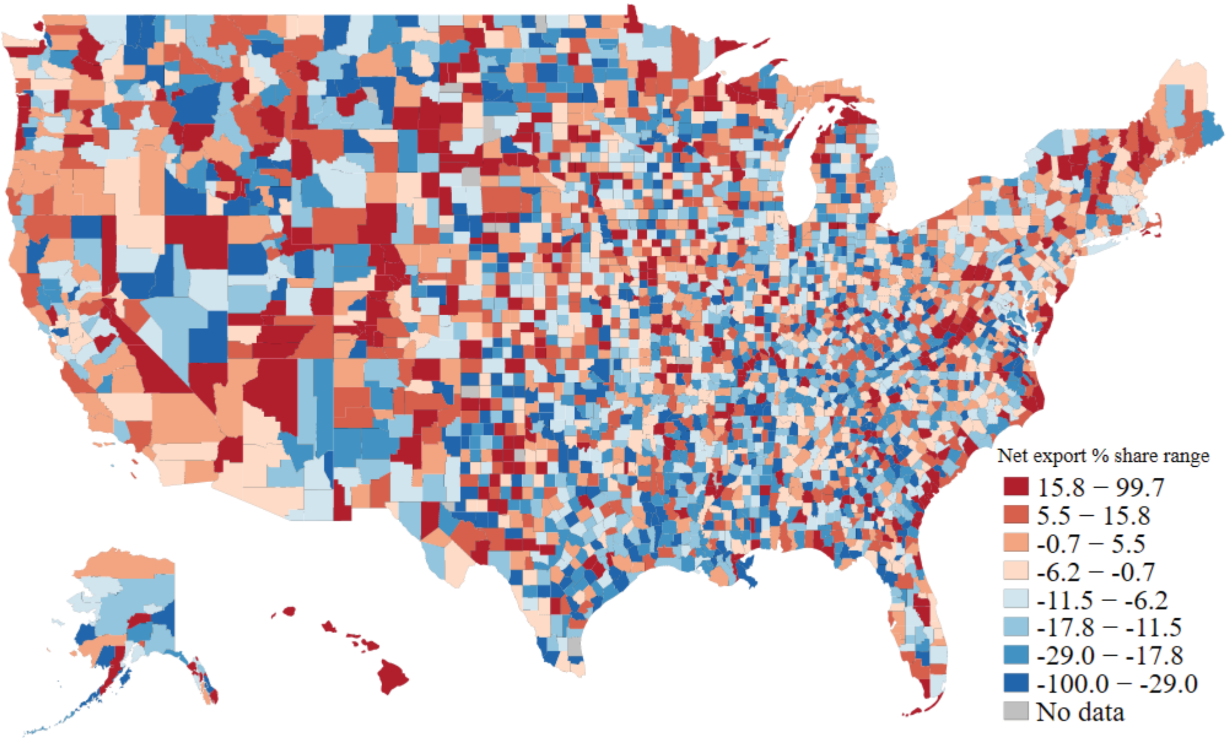


Notes: The net export share of each county is calculated as total exports (i.e., firm revenues from consumers outside of the county), minus imports (i.e., the total amount of revenue from consumers leaving the county), divided by the total amount of final consumption sold in the county. The distribution has been winsorized at –1 to avoid the long tail of rural counties that import most of their consumption.

Next we show this distribution in the form of a map, with figure 4 showing the distribution of net export shares across the United States with darker shades of red indicating a high net export share, while darker shades of blue indicate a higher import share. Here we see many expected patterns, including high export shares from places like Nevada and Hawaii, which are top tourist destinations. Overall, these patterns in figures 3 and 4 highlight the idea that counties are interconnected through consumption, indicating the potential importance of across-county consumption patterns.

Figure 4. Distribution of Net Export Share in the United States

Red—high export share **Blue—high import share**



Notes: The net export share of each county is calculated as total exports (i.e., firm revenues from consumers outside of the county), minus imports (i.e., the total amount of revenue from consumers leaving the county), divided by the total amount of final consumption sold in the county. Positive values indicate higher net export share and are shown in red and negative values indicate lower net export share and are shown in blue.

4. Consumption Flow Accounting: A Simple Test

The level of spending by consumers (i.e., consumption) that reside in a county must be equal to the amount of final consumption sold, minus the export of consumption to other areas by firms in the county, plus the imports of consumption by consumers traveling to other counties to consume, as shown in equation (1).

$$\begin{aligned}\text{Household Consumption} = \text{Final Product Sold} - \text{Export of Consumption} \\ + \text{Imports of Consumption}\end{aligned} \tag{1}$$

We use this basic accounting relationship to both test the validity of the data and also highlight the importance of these cross-market spending flows in understanding the consumption link across counties. To test this relationship, we first need empirical counterparts for each element.

Moving from left to right, the first estimate that is needed is an independent measure of household consumption. Household consumption at the county level is not an official statistic that currently exists. Indeed, one motivation for working with spending flow measures is to obtain a county-level measure of consumption from the right-hand side of the accounting relationship. However, we can empirically approximate an independent value assuming that consumer preferences are homothetic at the county level. This allows us to assume a constant share of income is devoted to the goods and services in our 15 select NAICS categories. We further assume that this budget share is constant across the entire United States for a given year. With this assumption, we then look at the national budget share of consumption going to our NAICS categories, which averages to be 38 percent of income. Next, we multiply the national budget share in each year times the income in each county from the BEA to obtain an estimate of consumption in county j, $\widehat{\text{Household Consumption}}_{j,t}$.

The next necessary element for equation (1) is an estimate of $\widehat{\text{Final Product Sold}}_j$ in county j. This estimate is taken directly from our spending estimates based on the EC data where the total spending over industries n is aggregated:

$$\widehat{\text{Final Product Sold}}_j = R_j = \Sigma_{\forall n \in I} R_{j,n}$$

where $R_{j,n}$ is the total sold by firms in the county j for industry n and set of industries I.

The estimate of the exports of consumption is the total amount sold by firms in the county to consumers that reside outside of the county. This is calculated as:

$$\widehat{\text{Exports of Consumption}}_j = \Sigma_{\forall n \in I} \Sigma_{\forall i \in C s.t. i \neq j} R_{j,n} S_{i,j,n}$$

where $S_{i,j,n}$ is the total share of revenues for firms in industry n located in county j selling to consumers that reside in county i. The estimated share, $S_{i,j,n}$, is based on 2015 estimates, so the implicit assumption is that these shares are constant across years in the sample.

Finally, to estimate the dollar amount of imports coming from a county, a similar exercise is conducted. The estimate of consumption import is the total amount consumed outside of a county by consumers that reside in county *j*. This amount may be estimated as:

$$\widehat{\text{Imports of Consumption}}_j = \Sigma_{\forall n \in I} \Sigma_{\forall k \in C, s.t. i=j, k \neq j} R_{k,n} S_{i,k,n}$$

After obtaining the empirical counterpart for each element of (1), we can estimate a simple regression model to test the accounting relationship:

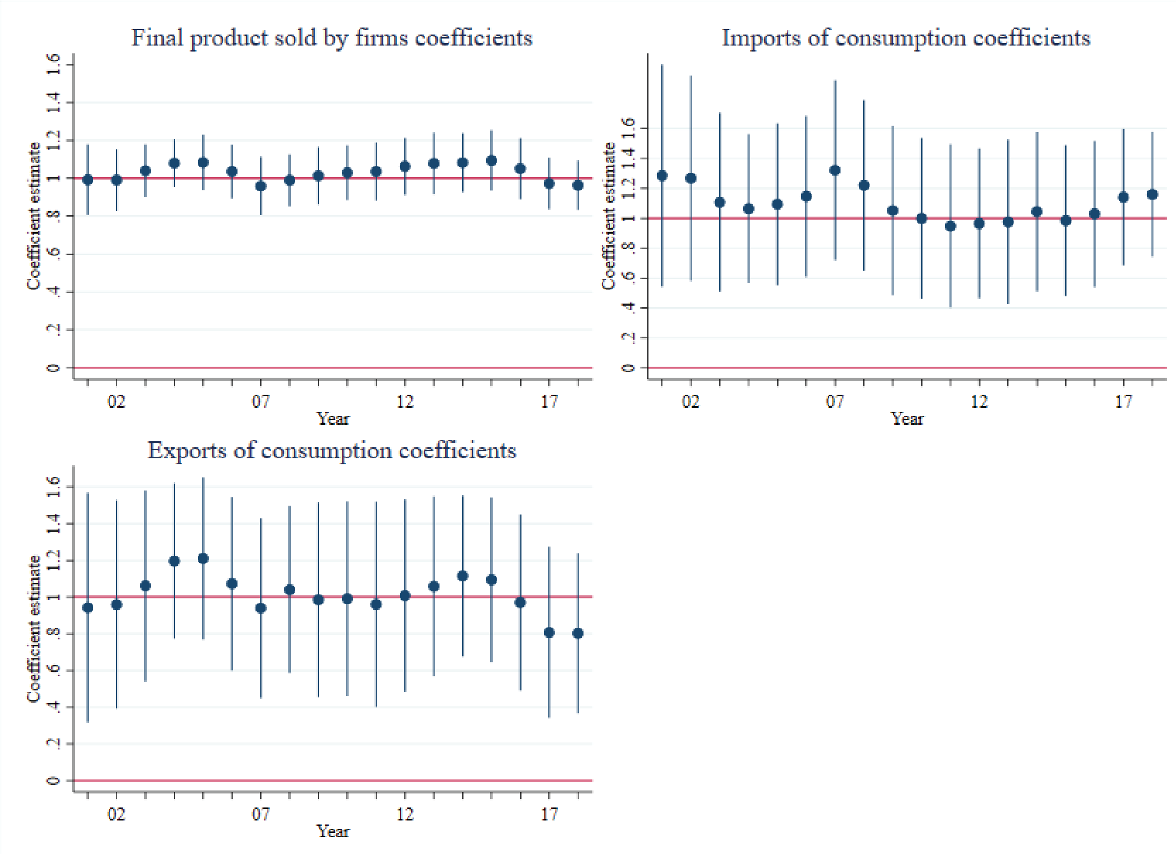
$$\begin{aligned}\widehat{\text{Household Consumption}}_{j,t} = \beta_1(\widehat{\text{Final Product Sold}}_{j,t}) - \beta_2(\widehat{\text{Exports of Consumption}}_{j,t}) \\ + \beta_3(\widehat{\text{Imports of Consumption}}_{j,t}) + \varepsilon_{j,t}\end{aligned} \tag{2}$$

If consumption flows are important, we should reject the hypothesis that they are equal to zero $\beta_2 = \beta_3 = 0$. In addition, if the accounting relationship holds, then we should not be able to reject the hypothesis $\beta_2 = \beta_3 = 1$. Before estimating this equation, it is important to highlight that measurement error enters the equation from multiple sources, increasing the likelihood of attenuating these estimates and reducing the statistical significance of the import and export variables. In particular, there may be measurement error from assuming shares $S_{i,k,n}$ are constant across years, from the Fiserv data measurement error, from assuming homothetic preferences across counties, and errors coming from the imputation of county spending in the intercensal years.

The empirical test is run in a joint regression for every year and county in our data from 2002 to 2017, but with different coefficients for each year. The coefficient for each year is shown in figure 5. Across all years we see that we can strongly reject the hypothesis that our consumption import and export measures are insignificant $\beta_2 = \beta_3 = 0$, as the estimates are significantly different from zero in each year. The import and export coefficients center around 1 across all years, and we cannot reject the hypothesis that estimates are equal to 1 in any year. In other words, we cannot reject the hypothesis that this accounting relationship holds in the data.

These estimates suggest that the right-hand side of the accounting relationship provides meaningful information about the components of consumption at the county level, which will be the focus of the analysis of the Great Recession.

Figure 5. Regression Coefficients from Accounting Tests Across Years



Notes: This figure shows the coefficient estimates from the regression equation 2. The regression is run on the full sample of counties and years with interactions of both counties and years using the income in the county in 2007 as a weight and clustering the standard errors at the state level. The upper left box shows the coefficient based on total sales by firms in the county. The upper right box shows the coefficient on imports of consumption. The lower left box shows the coefficient on exports of consumption. The blue dots represent the point estimates for the coefficient and the vertical lines represent the 95 percent confidence interval of the coefficients.

5. Empirical Application: The Case of the Great Recession

In this section we re-examine the Great Recession and the effect of housing wealth on spending and employment across areas. We specifically look at the effects of the recession on aggregate spending and employment for firms. We focus on firms, rather than consumers directly, as the spending flow information from Fiserv is based on firm-level data, and it also allows us to analyze the different components of wealth shocks affecting firms, such as wealth shocks to consumers.

Following the specification of Mian, Rao, and Sufi (2013) and Mian and Sufi (2014), growth rates are computed as percent changes between years t and $t - 2$:

$$\Delta Y_{j,t} = \frac{Y_{j,t} - Y_{j,t-2}}{Y_{j,t-2}}$$

where $t = 2009$ is our main specification. The paper by Mian and Sufi (2014) focuses on the 2007 to 2009 period, but the paper by Mian, Rao and Sufi (2013) focuses on the 2006 to 2009 period. Given that the Great Recession did not start until December 2007, we use 2007 as the base year for both our spending and employment analysis.

The variable $\Delta Y_{j,t}$ is either the growth rate in spending or employment. For our main specification, the 15 NAICS categories included in both the spending and employment estimates, $\Delta Y_{j,t}$, corresponds to the same NAICS categories used in the flow estimates.[14]

Assuming our spending flow estimates are more broadly representative of spending flows more generally, we can expand the number of industry categories included in the employment and spending effects. We obtain similar results when including additional non-tradable categories in our estimates.

5.1 Weighting Housing Wealth Change by Spending Flows

Our base measure of housing wealth change for consumers residing in county i is ΔHNW_i. Assuming that consumption does not cross county borders, then the wealth change relevant for firms in county j is then ΔHNW_i where $i = j$.

The hypothesis in this paper is that the effect of the change in housing wealth is not constrained to county borders. To distribute housing wealth shocks to firms more accurately, we use an aggregate measure of consumption flow across all industries in our data based on where consumers reside.

14. As a robustness check and for comparison, we have also estimated spending, employment and spending flow estimates focusing only on the non-tradable categories, as defined by Mian, Rao and Sufi (2013) and Mian and Sufi (2014). We obtain estimates very similar to those shown here.

The aggregate expenditure flows are measured as the share of revenues coming from each industry, weighted by the industry spending in the county:

$$S_{i,j}^{AGG} = \frac{\sum_{\forall n \in I} R_{j,n}.S_{i,j,n}}{\sum_{\forall n \in I} R_{j,n}}$$

This share, $S_{i,j}^{AGG}$, better captures the likely or potential customers from location i for firms located in county j. To better understand this flow variable, consider a hypothetical example, county A. In this example, if only about 50 percent of a firms revenue in county A comes from the home county A, $S_{i=A,A}^{AGG}$ = 50%, then we should expect changes in the wealth of those potential customers in A to account for around 50 percent of the total effect. The remaining 50 percent would be from exports (i.e., consumption from customers that reside outside of the county).

Taking these shares as fixed over time, the housing wealth change that is more relevant for firms in county j is then:

$$\Delta HNW_j^{FLOW} = \sum_{\forall i \in C} (\Delta HNW_i).S_{i,j}^{AGG} \quad (3)$$

Continuing with the example, suppose the local housing decline was 20 percent in the home county, A, that has 50 percent of the customers, but just a 2 percent decline for counties outside of the home county, then the associated decline for firms located in county A would be ΔHNW_A^{FLOW} = 20% • 50% + 2% • 50% = 11%.

This can be broken out into two components of the housing wealth change—one measure from consumers that reside in the same county as the firm, and another measure from consumers outside of the county: $\Delta HNW_j^{FLOW} = \Delta HNW_j^{Home} + \Delta HNW_j^{Export}$. More specifically these can be measured as:

$$\Delta HNW_j^{HOME} = (\Delta HNW_{i=j}).S_{i=j,j}^{AGG}$$

and also a separate measure from consumers that reside outside the county:

$$\Delta HNW_j^{EXPORT} = \sum_{\forall i \neq j \in C} (\Delta HNW_i).S_{i,j}^{AGG}$$

The regression we analyze then takes the form:

$$\Delta Y_{j,t} = \beta_1 f(\Delta HNW_j, S_{i,j}^{AGG}) + \beta_2 X_{j,t} + \varepsilon_{j,t} \quad (4)$$

where $f(\Delta HNW_j, S_{i,j}^{AGG})$ is a function of housing wealth changes and across market spending flows. We examine two types of housing wealth measures: 1) those that ignore across-county consumption flows ΔHNW_j, and 2) those that use the across-county consumption flows by including ΔHNW_j^{FLOW} or by including both ΔHNW_j^{Home} and ΔHNW_j^{Export}.

As mentioned previously, the dependent variable, $\Delta Y_{j,t}$, will be either changes in spending or employment. The first differencing in the estimation essentially makes this a difference-in- difference analysis, comparing spending and employment changes in areas that are more or less affected by housing wealth changes. The key controls included in the estimation are two-digit industry shares in each county that account for the general growth rate of different sectors over this time period. The inclusion of industry share mitigates the potential endogeneity concern that industry structure could be associated with changes in housing wealth. In addition, some of the identifying variation is plausibly exogenous as it depends on housing fluctuations of the counties of customers that are outside the county in which the firm is located.

This specification is a simple ordinary least squares (OLS) regression model. The advantage of the OLS model is that it more directly shows the correlation in the housing wealth shock on across-county spending and employment. The cross-county shares are fixed over time, so they are exogenous by construction. However, there is still the possibility that the estimates are affected by endogenous factors. For instance, employment declines could cause a downward shift in housing prices. We address these concerns in our robustness section, in which we apply both panel and instrumental variable (IV) specifications.

6. Results

The first set of regression results are shown in table 2. The first specification (1) uses the housing wealth shock that is in the same county as the firm is located, which ignores spending flows. The effect of the housing wealth shock on spending is positive and significant, as expected and consistent with previous work, with an elasticity of 0.16. If housing wealth declines by 10 percent, there is a 1.6 percent reduction in spending. In specification (2) we form our preferred specification that includes the weighted consumption flows, which also shows a positive and significant coefficient, but the magnitude is about 25 percent larger with an elasticity of around 0.19. To compare this estimate to other work in the literature, we convert the elasticity of spending to housing wealth to a marginal propensity to consume out of housing wealth by dividing these elasticities by the ratio of housing wealth to consumption, which we estimate to be 2.47.[15] The marginal propensity to consume based on the estimates without the flows is 6.4 cents on the dollar. Our preferred specification (2) indicates a value of 7.7 cents on the dollar. This estimate matches closely with the estimate from Mian, Rao, and Sufi (2013) that finds an estimate of around 7.2 cents on the dollar. Our estimates are surprisingly similar given that many aspects of our data and analysis are distinct. For instance, our estimates are based on over 3,000 counties, while they looked at about 900; we use different spending estimates based on the EC; and we adjust for consumer location using across-county spending flows.[16] These estimates are based on simple OLS regressions, but in our robustness section we show that these estimates correspond quite closely to our IV and panel estimates.

We include additional specifications to demonstrate the economic importance of including spending flows. The third specification (3) presents a test of the relative importance of these two alternative measures, which includes both the net housing wealth change variable with and without the spending flow weights. The measure of the net housing wealth change that ignores the flows, appears to be statistically insignificant, while the measure of the net wealth change with the flows is positive and statistically significant. This indicates that the estimates with the associated weighted spending flows is producing a more accurate measure of the associated housing wealth shock. In other words, all of the explanatory variation loads onto the explanatory variable that includes the flows, which suggests it is the better measure. Specification (4) includes the net housing wealth effect that ignores the flows, but also includes two flow weighted measures:

15. The elasticity captures the percent change in spending from a percent change in housing prices. To arrive at a dollar change in spending from a dollar change in housing wealth, we need to divide the elasticity by housing wealth and multiply it by the level of consumption. Following Guren et al. (2020) we estimate the value of the housing spending based on the market value of owner-occupied real estate from the Flow of Funds and we estimate the value of consumption based on total personal consumption expenditures net of housing and utilities. We calculate the average of the consumption and housing value components over the period 2000–2019 and then form the ratio.

16. Our results are also similar in range to Di Maggio et al. (2020) that examines the MPCH based on stock returns and find estimates of 5 cents on the dollar or more. Aladangady (2017) estimates a MPCH based on microdata of 4.7 cents for homeowners and finds no effects for renters. Based on a homeownership rate of 65 percent, this corresponds to an MPCH of 3.1 cents overall. Guren et al. (2020) find a MPCH of around 2.4 cents on the dollar.

the home net housing wealth effect and the export net housing wealth effect. The two flow-weighted estimates are again significant, but the estimates without weighted flows are insignificant. The last specification (5) is the same as specification (4), but excludes the unweighted housing wealth change. The results show positive and significant effects of changes in housing wealth on spending, whether it is from an export county or import county. The magnitude of the effect appears to be slightly higher from net housing wealth changes from exports relative to changes in net housing wealth from the home location, but we cannot reject the hypothesis that the coefficients are equal.

Table 2. Housing Wealth Change on Spending Growth

	(1) % Chg. Spend	(2) % Chg. Spend	(3) % Chg. Spend	(4) % Chg. Spend	(5) % Chg. Spend
Δ HNW (no flow)	0.158*** (0.0196)		−0.0841 (0.0861)	−0.0700 (0.0834)	
Δ HNW (total flow)		0.191*** (0.0232)	0.290*** (0.100)		
Δ HNW (home)				0.265*** (0.0969)	0.179*** (0.0254)
Δ HNW (export)				0.316** (0.119)	0.254*** (0.0788)
Observations	3063	3062	3062	3062	3062

Notes: Standard errors in parentheses and are clustered by state. Estimates are weighted by 2007 population levels.
* p<0.10, ** p<0.05, *** p<0.01

Table 3 is the same as table 2, but examines the effect on employment as a dependent variable rather than spending. The magnitude of the effect is smaller, with an elasticity of 0.11 in specification (1) without using spending flows. Similar to the spending estimates, the magnitude of the estimate increases with the incorporation of the spending flows in specification (2). When both measures of housing wealth change are included together in specification (3), the measure that excludes the spending flows is insignificant. Interestingly, employment is affected more by export housing wealth changes relative to changes in the home market (specifications (4) and (5)).

The result of our main specification (2) is similar to Mian and Sufi (2014) where the housing wealth effect on employment that they observe implies estimates of MPCH of between 4.1 and 7.3 cents on the dollar, while our main estimate implies an MPCH of 5.9 cents on the dollar.

Table 3. Housing Wealth Change on Employment Growth

	(1) % Chg. Emp.	(2) % Chg. Emp.	(3) % Chg. Emp.	(4) % Chg. Emp.	(5) % Chg. Emp.
Δ HNW (no flow)	0.120*** (0.0203)		−0.108 (0.0656)	−0.0743 (0.0545)	
Δ HNW (total flow)		0.147*** (0.0242)	0.273*** (0.0865)		
Δ HNW (home)				0.214*** (0.0686)	0.123*** (0.0214)
Δ HNW (export)				0.337*** (0.113)	0.271*** (0.0863)
Observations	3103	3102	3102	3102	3102

Notes. Standard errors in parentheses and are clustered by state. Estimates are weighted by 2007 population levels.
* p<0.10, ** p<0.05, *** p<0.01

6.1 Regression by Export Quartile

All of the analysis above relies on interactions of spending flows and housing wealth changes. In this section, we highlight the importance of the flows by discretely categorizing counties into export quartiles. If the consumption flows are meaningful, then we should expect the export housing wealth changes to have larger effects in the high-export quartile and to have less effect in the low-export quartile. Similarly, we should expect the home wealth change to have larger effects in those counties that export less spending. To perform this exercise, we construct a measure of average housing wealth change from consumers that reside outside of the county, and another measure for the average net wealth change from the home location.

The average net wealth shock from the home location is the average home wealth shock divided by the share of spending from the home location, which simplifies to the housing wealth change that excludes flows:

$$Average\Delta HNW_j^{HOME} = \frac{\Delta HNW_j^{HOME}}{S_{i=j,j}^{AGG}} = \frac{\Delta HNW_{i=j}.S_{i=j,j}^{AGG}}{S_{i=j,j}^{AGG}} = \Delta HNW_j$$

Returning to our previous example for county A, this would be the housing wealth decline in the home county, which was equal to 20 percent.

The average wealth change from outside the county is just the export housing wealth change divided by the export share:

$$Average\Delta HNW_j^{EXPORT} = \frac{\Delta HNW_j^{EXPORT}}{\Sigma_{\forall i \neq j \in C} S_{i,j}^{AGG}} = \frac{\Sigma_{\forall i \neq j \in C}(\Delta HNW_i).S_{i,j}^{AGG}}{\Sigma_{\forall i \neq j \in C} S_{i,j}^{AGG}}$$

In our previous example for county A, this average would be the average housing wealth decline outside of the home county, which was equal to 2 percent.

Both of these measures are simply average measures of net wealth changes across their expected customers, which ignore the share of consumption coming from outside or inside the county, which were both 50 percent in our example. In the county A example, if the potential customers are primarily from the home county, say with a 90 percent share of spending, then the home price decline of 20 percent should be more salientf. However, if potential customers are primarily from outside the home county, say a 10 percent share of spending come from the home county, then the home price decline of 2 percent should be more salient.

The estimates for spending by export quartile are shown in table 4. The estimates show that for higher export counties, the housing wealth changes from export counties are significantly more important. As expected, in the fourth quartile, the export coefficient is significant and also larger in magnitude, as would be expected, since a greater share of the change in housing wealth is coming from consumption exports. The magnitude of the net housing wealth effect from exports declines for those counties in which exports are lower, as we should expect. For the highest export quartile, the home net wealth shock is statistically insignificant, but becomes statistically significant for the lowest two quartiles, in which most of the consumption occurs locally. Table 5 shows a very similar pattern, but for employment.

These estimates show that for high export counties, focusing only on local shocks to consumers can be highly misleading.

6.2 Robustness Checks

The above specifications could potentially be affected by endogeneity problems, as the decline in employment could be a cause, and not a result, of the housing wealth decline. To address this issue we examine some alternative specifications.

Table 4. Housing Wealth Change from Home and Export Counties on Spending Growth: By Quartile of Export Share

	(1) Quartile 4	(2) Quartile 3	(3) Quartile 2	(4) Quartile 1
Δ HNW (no flow)	0.0262 (0.0446)	0.0194 (0.0499)	0.0834** (0.0371)	0.192*** (0.0438)
Average Δ HNW (export)	0.248*** (0.0892)	0.208*** (0.0745)	0.126* (0.0715)	−0.0157 (0.0982)
Observations	756	763	772	770

Notes. Standard errors in parentheses and are clustered by state. Estimates are weighted by 2007 population levels.
* p<0.10, ** p<0.05, *** p<0.01

Table 5. Housing Wealth Change from Home and Export Counties on Employment Growth: By Quartile of Export Share

	(1) Quartile 4	(2) Quartile 3	(3) Quartile 2	(4) Quartile 1
Δ HNW (no flow)	0.00842 (0.0324)	0.0393 (0.0351)	0.0825** (0.0353)	0.0802*** (0.0216)
Average Δ HNW (export)	0.210*** (0.0700)	0.184*** (0.0542)	0.0733 (0.0481)	0.105 (0.0653)
Observations	770	780	776	776

Notes. Standard errors in parentheses and are clustered by state. Estimates are weighted by 2007 population levels.
* p<0.10, ** p<0.05, *** p<0.01

As one alternative we estimate a panel model, which can reduce endogeneity by controlling for local factors affecting growth leading up to the Great Recession.

$$\Delta Y_{j,t} = \beta_1 f(\Delta HNW_j, S_{i,j}^{AGG}) \bullet (t{=}2009) + \beta_2 X_{j,t} + \gamma_j + \tau_t + \Delta\varepsilon_{j,t} \quad (5)$$

In addition to the 2007—2009 period, the panel model includes 2005–2007 and 2003–2005. The model includes the addition of a county-specific fixed effects γ_j that captures the unique growth factors associated with a particular county. The model also includes a year fixed effects, τ_t, capturing national trends in growth rates over each period. The estimates also includes controls for industry share, as in the OLS specification. The estimates from the panel specification on spending and employment are shown in tables 6 and 7. The results are qualitatively similar in many respects to the simple OLS estimates. The effect of net wealth on employment and spending is positive and the net wealth effects based on the flows are larger than those excluding the flows, by around 20 percent for both spending and employment. The effect on the housing wealth change from export s is positive and significant for both spending and employment.

Table 6. Panel Regression Model of Spending

	(1) % Chg. Spend	(2) % Chg. Spend	(3) % Chg. Spend
Δ HNW (no flow)	0.217*** (0.0375)		
Δ HNW (total flow)		0.253*** (0.0437)	
Δ HNW (home)			0.238*** (0.0636)
Δ HNW (export)			0.325* (0.167)
Observations	12201	12198	12198

Notes. Standard errors in parentheses and are clustered by state. Estimates are weighted by 2007 population levels.
* p<0.10, ** p<0.05, *** p<0.01

Table 7. Panel Regression Model of Employment

	(1) % Chg. Emp.	(2) % Chg. Emp.	(3) % Chg. Emp.
Δ HNW (no flow)	0.137*** (0.0312)		
Δ HNW (total flow)		0.161*** (0.0352)	
Δ HNW (home)			0.152*** (0.0525)
Δ HNW (export)			0.205* (0.113)
Observations	12347	12343	12343

Notes. Standard errors in parentheses and are clustered by state. Estimates are weighted by 2007 population levels.
* p<0.10, ** p<0.05, *** p<0.01

Even with a panel specification, there may be concerns of endogeneity that have been raised in previous research by Mian, Rao, and Sufi (2013), Mian and Sufi (2014), and Guren et al. (2020). For instance, the shock to income or employment could have initiated the decline in housing prices in the area. The instrument used in Mian, Rao, and Sufi (2013) and Mian and Sufi (2014) are based on estimates from Saiz (2010), capturing the housing supply elasticity for a subset of metropolitan statistical areas, but we are attempting to capture effects for all counties in the United States. Moreover, this instrumental variable strategy has been critiqued by Guren et al. (2020) and Davido (2016) as potentially being correlated with other city characteristics leading to potential biases. Therefore, as an alternative instrument, we follow Guren et al. (2020), which uses a history of housing pricedata that captures systematic differences in exposure to regional price fluctuations. The basic idea behind the instrument is to identify those regions in the country that have a particularly strong response to national or regional fluctuations in price. Therefore, the instrument is based on the general price sensitivity in the county, and not on other local factors that may be occurring directly around the Great Recession event date.

Following Guren et al. (2020), we use historical information on local area housing price responsiveness to regional price movements to estimate instruments for the level of sensitivity in local markets to regional shocks. Using Zillow data from January 1996 to January 2020, we estimate the responsiveness of county-level housing prices to regional changes in housing prices. The estimated county-specific responsiveness to regional price movements is the instrument that we apply in our estimates. The spending flow data are used to weight the instrument across different counties. Additional details of the formation of this instrument are included in the appendix.

Tables 8 and 9 show alternative models that include IV specifications. Specification (1) includes an IV model that excludes accounting for spending flows, along with IV models and IV panel models that account for the cross-market spending flows. The estimates are again qualitatively similar to those found using the simple regression models. We see the magnitude of the estimates show that accounting for spending flows, specification (2), exceed the estimates that do not include spending flows, specification (1). Specifications (4) and (5) apply the IV strategy to our panel estimates and we again obtain similar results.

Table 8. Instrumental Variable Regression Model for Spending

	(1) IV no flow	(2) IV flows	(3) IV flows	(4) Panel IV flows	(5) Panel IV flows
Δ HNW (no flow)	0.133*** (0.0245)				
Δ HNW (total flow)		0.167*** (0.0298)		0.211*** (0.0420)	
Δ HNW (home)			0.147*** (0.0339)		0.151*** (0.0558)
Δ HNW (export)			0.262*** (0.0917)		0.482*** (0.152)
Observations	3063	3062	3062	12194	12194

Notes. Standard errors in parentheses and are clustered by state. Estimates are weighted by 2007 population levels.
* p<0.10, ** p<0.05, *** p<0.01

Table 9. Instrumental Variable Regression Model for Employment

	(1) IV no flow	(2) IV flows	(3) IV flows	(4) Panel IV flows	(5) Panel IV flows
Δ HNW (no flow)	0.110*** (0.0209)				
Δ HNW (total flow)		0.135*** (0.0244)		0.124*** (0.0375)	
Δ HNW (home)			0.109*** (0.0284)		0.0796 (0.0524)
Δ HNW (export)			0.259*** (0.0983)		0.322*** (0.124)
Observations	3103	3102	3102	12342	12342

Notes. Standard errors in parentheses and are clustered by state. Estimates are weighted by 2007 population levels.
* p<0.10, ** p<0.05, *** p<0.01

As another robustness check on the estimates we re-estimate tables 8 and 9, but include additional industry categories for our spending and employment estimates, including all non- tradable categories. The basic idea is that the spending flows may provide reasonable proxies for all economic activity between areas. We again find results very similar to those presented here. We have also estimated the model using only those categories used in Mian and Sufi (2014) and obtain qualitatively similar results.

7. Economic Implications

We next evaluate the local economic effect of the change in housing wealth on spending and employment for the entire nation. We do this using estimates from specification (2) of tables 2 and 3, which appear to produce estimates in a similar range to those based on the panel, IV and IV panel estimates. We contrast our preferred specification (2) to the local wealth effects from using specification (1), which assumes no cross-county spending flows , also reported in tables 2 and 3.

A summary of our estimates are reported in table 10. The estimates using flows show a decline in employment of 2.2 percent due to the change in housing wealth and a decline of 2.9 percent in spending. For employment, this estimate is about 26 percent larger than the effects based on ignoring the spending flow estimates. For spending, this estimate is about 17 percent larger than when we ignore spending flows. For both spending and employment, ignoring these cross-county flows tends to understate the magnitude of the housing wealth effects.

In addition, the estimates have different predictions regarding which counties are affected by the associated housing wealth change. The magnitude of the effect on firms will be distinct, depending on whether spending flows are considered or not, but the allocation of which firms are affected will also change. For instance, suppose county A has a large change in housing prices, but all of county A's consumption takes place in county B. Without using the flows, the decline in spending and employment at firms will entirely be attributed to county A, when these changes should actually be attributed to county B. We capture the differential allocation of effects across counties by using coefficient estimates from specification (2), to normalize the magnitude of the effect, but calculate the decline in spending and employment based on the two different estimates of the housing wealth shock. One estimate uses spending flows, ΔHNW_j^{FLOW}, and the second estimate ignores the spending flows, ΔHNW_j. To measure this difference, we calculate the absolute value of the difference in spending and employment, based on those two alternative measures $\frac{\beta_1(|\Delta HNW^{FLOW} - \Delta HNW_j|)}{\beta_1 \Delta HNW_j^{FLOW}}$. We then add up the absolute differences across all counties in the country. To obtain a percentage effect, we divide this total by full magnitude of the decline. We find a percent difference in allocation of over 11 percent for both employment and spending.

Table 10. Measuring Local Economic Effects

Employment effects (# of persons)			Spending effects (in millions)		
		% Chg			%Chg
Total labor in sector	22,628,312		Total revenues in sectors	$3,999,575	
Labor declined with flows	487,474	2.2%	Spend decline with flows	$116,386	2.9%
Labor declined with no flows	412,799	1.8%	Spend decline with no flows	$99,308	2.5%
		%Diff			%Diff
Relative to no flows prediction			Relative to no flows prediction		
Additional decline with flows	74,673	25.5%	Additional decline with flows	$17,078	17.2%
Allocative difference with flows	57,135	11.7%	Allocative difference with flows	$13,294	11.2%

8. Conclusion

Local measures of spending and employment are of great interest to economists studying both microeconomic and macroeconomic questions. As key data sources are often available at the county level, many studies focus on policy effects or economic shocks at this level of disaggrega tion. In this paper, we introduce a consumption link across counties using cross-county spending flows based on card transaction data. We show net exports of consumption vary greatly across counties, especially in more rural areas, and this has implications for how each county is affected by local economic policies and shocks.

We provide evidence of the importance of across-county spending flows based on a simple accounting relationship and also based on evidence from the Great Recession. We show that not accounting for these cross-border spending flows leads to less precise estimates that understate the magnitude of housing wealth changes from 2007 to 2009 on employment and spending and also misallocates where those shocks have an effect. Our work shows that spending flows may be important for measuring local economic shocks, but it also has implications for policy design. In particular, the effect of local targeted policies on either firms or consumers may have broader effects outside of local markets, depending on the spending patterns of consumers.

More generally, the across-county consumption link is an important aspect of spatial economics that has received relatively little attention, likely due to data limitations (Redding and Rossi-Hansberg 2017). There are many potential applications to the data constructed in this study. These data may be used to help understand the effects of local tax policies, income shocks to consumers, or policies that affect the population heterogeneously, such as the ACA. In addition, these data can be used to help define local consumption markets, akin to how labor markets are defined using commuting data to construct commuting zones. Across county links in goods and factor markets have been shown to be empirically important, such as in the work by Monte, Redding, and Rossi-Hansberg (2018), in which they examine labor demand shocks on employment elasticities using a general equilibrium framework. The across-county consumption link may be an important addition to this literature.

In addition to applications of the spending flow estimates presented here, there are many potential avenues for improvements to our estimates. One area where additional work may be useful is e-commerce. This was not a limitation for our application over the 2007 to 2009 period, when e-commerce was a relatively small share of consumption, but this is an area of growing importance. Researchers may want to turn to alternative data sources to capture this aspect of spending. Also, in our work we excluded foreign spending to simplify the analysis, but it may be of particular interest in future work to better understand how foreign consumption spending can impact local markets. Finally, for our analysis we focus on a single cross-section in 2015, and assume that the share of spending across each location remained fixed. While we provide evidence that this assumption is reasonable, it may be of interest in future work to look at changes in spending flows and the determinants of across-county spending flows over time.

References

Aladangady, A. 2017. Housing Wealth and Consumption: Evidence from Geographically Linked Microdata. *American Economic Review* 107, no. 11: 3415–3446.

Allcott, H., R. Diamond, J. Dubé, J. Handbury, I. Rahkovsky, and M. Schnell. 2019. "Food Deserts and the Causes of Nutritional Inequality," *Quarterly Journal of Economics*, 134, no. 4: 1793–1844.

Davidoff, T. (2016). Supply constraints are not valid instrumental variables for home prices because they are correlated with many demand factors. *Critical Finance Review*, 5(2):177–206.

Di Maggio, M., A. Kermani and K. Majlesi. 2020. "Stock Market Returns and Consumption." *Journal of Finance*, forthcoming.

Dolfen, P., L. Einav, P. J. Klenow, B. Klopack, J. D. Levin, L.Levin. Year. "Assessing the Gains from E-commerce." Working Paper.

Guren, A., A. McKay, E. Nakamura, and J. Steinsson. 2020. "Housing Wealth Effects: The Long View." *Review of Economic Studies*, forthcoming,.

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

Mian, A. and A. Sufi, (2014) "What Explains the 2007-2009 Drop in Employment?". *Econometrica*, 82(6), 2197-2223.

Monte, F., S. J. Redding, and E. Rossi-Hansberg. 2018. "Commuting, Migration, and Local Employment Elasticities." *American Economic Review*, 10, no. 12:3855–3890.

Redding, Stephen, and Esteban Rossi-Hansberg. 2017. "Quantitative Spatial Economics." *Annual Review of Economics*, 9: 21–58.

Saiz, A. 2010. "The Geographic Determinants of Housing Supply Elasticity." *Quarterly Journal of Economics*, 125, no. 31253–1296.

Appendix

A.1 Economic Census Receipts

The Geographic Area Series of the EC is collected every 5 years at detailed geographic and NAICS industry levels. The EC contains information on industry-level revenues which are used in this study to create measures of consumer spending. Our study focuses on county level estimates for 15 industries that are important contributors to personal consumption expenditures, which also have good coverage in the Fiserv database. While EC provides detailed information for many industries at the county level, there are some geography and NAICS combinations that are suppressed. We have used county-level three-digit NAICS industries for 2002, 2007, 2012 and 2017 as our benchmark years.

Table A.1 shows list of industries included in our analysis with their associated share of suppressed revenues to total revenues for each census year.[17] The level of these suppressions vary across industry, but in general they are extremely low. While industries such as gasoline stations have high coverage, only 0.5 percent of total receipts are suppressed, others like performing arts and amusement and recreation have higher suppression rates of around 10 percent in 2002 and 2007, but the 2012 and 2017 suppression rate decreases to 6.5 percent and 3 percent respectively.

Table A.1. Share of Suppressed Revenues to Total in Selected NAICS Industries (percentages)

NAICS	NAICS description	2017	2012	2007	2002
442	Furniture and home furnishings stores	1.8	3.0	2.0	2.0
443	Electronics and appliance stores	2.5	1.4	1.7	3.0
444	Building material and garden equipment	1.1	1.6	0.7	0.7
445	Food and beverage stores	1.3	1.4	0.9	0.5
446	Health and personal care stores	1.2	1.5	1.6	0.5
447	Gasoline stations	0.7	0.5	0.5	0.5
448	Clothing and clothing accessories stores	0.7	0.5	0.9	0.8
541	Professional and scientific services	2.5	4.0	6.0	5.0
621	Ambulatory health care services	1.8	3.0	4.0	4.0
711	Performing arts and spectator sports	3.5	3.0	10	10
713	Amusement, gambling, and recreation	5.1	6.5	11	15
721	Accommodations	1.0	1.2	2.8	1.3
722	Food services and drinking places	1.1	1.1	1.8	1.4
811	Repair and maintenance	0.7	1.0	1.9	1.5
812	Personal and laundry services	0.6	0.7	2.2	2.3

Source: Authors' calculation.
NAICS North American Industry Classification System

17. The rate of suppression is determined by comparing to national estimates that are unsuppressed.

A.1.1 Imputing Revenue for Suppressed Values in Economic Census Benchmark Years

Overall suppression in EC years is quite low, but to obtain complete coverage across counties, we perform some imputations. To address the issue of suppression in the benchmark years, the annual series of QCEW is used to create full set of revenues for all county-NAICS combinations. Annual QCEW data for privately owned establishments provide information on payroll, employment, and wages, and the version of the QCEW used does not contain any suppression across counties. The method used for these imputations is relatively simple and uses wage data to allocate missing revenues across counties.[18]

To impute the revenues in benchmark years we take three steps. First, we use wages on QCEW to impute missing payroll data on EC. Second, we calculate the ratio of payroll to revenue for the non-suppressed receipts by industry. Third, we multiply the payroll data from the QCEW to the ratio of revenue to payroll by industry to impute the missing revenue for NAICS-county combinations.[19]

A.1.2 Imputing Revenues for Intercensal Years

For the two benchmark years t to t+5 the revenues are observed $Revenue_t$ and $Revenue_{t+5}$. For the years between ECs, we interpolate revenues using annual QCEW wage data.

The interpolation adjusts revenues based on the growth rate in wages, but there is an annual adjustment to account for the divergence in growth rates between revenues and wages over the 5 years of the EC. Let t represent a benchmark year, and let $t+n$ be an intercensal year where n is between 1 and 4. The revenue in year $t+n$ is calculated as:

$$Revenue_{t+n} = \left(\frac{Revenue_{t+n}/Wage_{t+n}}{Revenue_t/Wage_t}\right)^{(n/5)} \bullet \frac{Wage_{t+n}}{Wage_t} \bullet Revenue_t$$

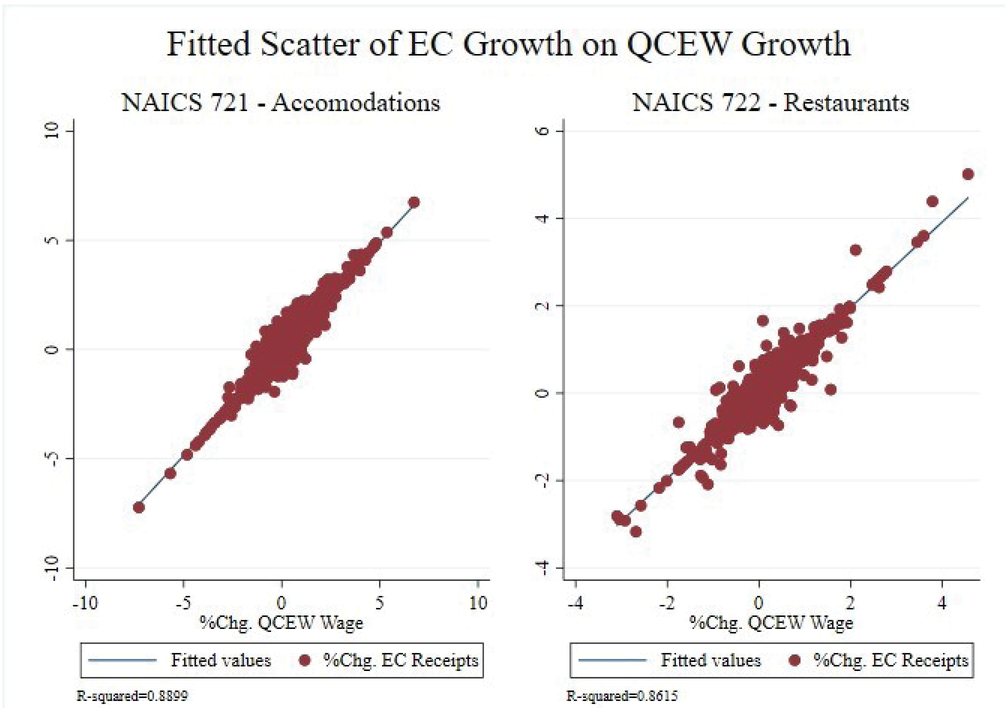
The second term $\frac{Wage_{t+n}}{Wage_t} \cdot Revenue_t$ is the estimated annual revenue based solely on the growth rate in wages. The first term, $\left(\frac{Revenue_{t+n}/Wage_{t+n}}{Revenue_t/Wage_t}\right)^{(n/5)}$, is the annual adjustment to better align changes in wages to predicted revenues. This first term suggests that our estimated changes in revenues may deviate from changes in wages.

18. The method used here is consistent with the method used by the BEA to create consumption estimates using EC revenues.

19. The assumption is that if there are wages being paid in that NAICS industry there should be revenue associated with the wage being paid. Only if both QCEW and census receipt are missing or are zero in a location for a specific industry, it is assumed that the revenue is zero.

While revenue growth is constrained to the growth rate in benchmark revenues, the year-to-year allocation of the 5-year revenue growth is determined by wages. To determine if applying wage data in this way is a reasonable, we examine how well wages do at predicting revenues in benchmark years. Figure A.1 is the graphical representation of regressing growth rates of EC revenues in the benchmark years on QCEW wage growth rates over the same periods for accommodations (NAICS 721) and restaurants (NAICS 722). The QCEW growth rates are closely correlated with EC growth rates. The R2 for both accommodations and restaurants is around 89 percent.

Figure A.1. Economic Census and QCEW Growth Rate Correlation



Source: Authors' calculation.
EC Economic Census
QCEW Quarterly Census of Employment and Wages

This method does quite well more generally. Table A.2 shows the R^2 estimate from that same regression for many NAICS industry categories. The three-digit NAICS categories used in our analysis are highlighted in red. The R^2 for our select industries are all above 0.70, except for NAICS categories 447 and 451 that have R^2 of around 0.5. The low R^2 for 447 is likely due to gas price fluctuation. Overall, the interpolation of revenue growth using wage data appears to do quite well at approximating revenues for many industries.

Table A.2. Regression Economic Census Growth Rates on the QCEW Growth Rate for Selected Industries for Census Years 2002, 2007, 2012 and 2017

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
NAICS	441	**442**	443	**444**	**445**	446	**447**	**448**	**451**	**452**
R^2	0.691	0.899	0.785	0.872	0.748	0.689	0.530	0.934	0.552	0.955
	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
NAICS	**453**	454	481	483	484	485	486	487	488	492
R^2	0.835	0.490	0.674	0.667	0.775	0.915	0.855	0.976	0.879	0.867
	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)	(29)	(30)
NAICS	493	511	512	515	517	518	519	521	522	523
R^2	0.661	0.656	0.930	0.674	0.902	0.485	0.850	0.589	0.891	0.918
	(31)	(32)	(33)	(34)	(35)	(36)	(37)	(38)	(39)	(40)
NAICS	524	531	532	533	541	551	561	562	611	**621**
R^2	0.955	0.688	0.856	0.556	0.584	0.937	0.178	0.874	0.608	0.800
	(41)	(42)	(43)	(44)	(45)	(46)	(47)	(48)	(49)	(50)
NAICS	622	623	624	**711**	712	**713**	**721**	**722**	**811**	**812**
R^2	0.923	0.902	0.613	0.707	0.659	0.810	0.868	0.905	0.786	0.711

Source: Authors' calculations.
NAICS North American Industry Classification System.
QCEW Quarterly Census of Employment and Wages

A.1.3 Fiserv Data, Spending Flows, and the Home Location Algorithm

The micro level data from Fiserv contains transaction level information for each firm in their data. Fiserv data contain well over one-third of all U.S. credit card transaction spending which includes more than 4.5 million U.S. firm locations and dollar amounts equal to 10 percent of the total GDP of the United States. To maintain the anonymity of cardholders and firms there are a number of suppression rules. The following suppression rules are applied: (1) no series has observation within a given NAICS and geography containing fewer than ten firms, and (2) across the series, no firm makes up more than 20 percent of the transaction volume. The card transactions flows include information on hashed card number, firm ID, transaction date and transaction amount. For each firm, the firm ID is mapped to the address and firm category code (MCC), which indicates the type of firm, which is mapped to its corresponding NAICS category.

The level of observations is a single transaction, although we do not see the data at this level of detail. As mentioned previously, the data has been aggregated and anonymized by Fiserv and Palantir. Prior to aggregation they apply an algorithm to predict the home location (HL) of each card holder, in order to construct the spending flow estimates used in our analysis. The HL algorithm uses transaction patterns to determine the most likely HL of a particular card based on all of that card's transactions across all firms. The raw data for modeling the location of the consumer consists of aggregated transaction counts for each card by three-digit NAICS categories and information on the firm zip codes. The estimated HL is formed based on estimates using a subset of cards for whom the HL of the card holder is known and a prediction is formed using a holdout sample of cards. The algorithm predicts the correct county around 75 percent of the time. Overall, the spending flow patterns from the known-cardholder data matches well with the patterns based on the full sample in which the HL algorithm is applied.

For our analysis we could have chosen either the known HL sample or the full predicted HL sample, as the two are quite similar. However, we chose the full predicted HL sample because it is based on more observations and can also help correct for the cases in which the zip code indicated by the card does not match where the individual actually resides.

A.2 Estimating Final Expenditure Flows

To obtain a complete system of consumption flows for the United States, we need to estimate the consumption flows in locations where the Fiserv estimates are suppressed. Overall, this accounts for about 15 percent of spending for our select categories. The goal of our imputation is to provide the best possible estimate for these missing expenditures. We examined a variety of flexible linear models to impute the missing spending flows, then we chose the method that performed the best based on cross-validation, a model validation technique, from a holdout sample.[20]

One factor that helps with imputation is that even when spending flows are suppressed, our data provides information regarding the set of counties where consumers are coming from, so we do not need to impute the set of potential counties. For instance, if NAICS category 448 is suppressed in Montgomery County, Maryland, we still observe the set of counties that customers came from to purchase in 448, but we do not observe the actual spending shares across locations. To impute the share of revenues for firms in industry n and county j going to location i, we estimate a flexible linear regression model with the log share of spending on the left-hand side $log(S_{i,j,n})$. Importantly, the right-hand side of the equation includes a county-pair fixed effect $\tau_{i,j}$ to capture economic activity occurring between two counties, using shares observed in other industries to help impute the industry share. For instance, suppose the share of a firm's revenues from a particular county for general merchandise stores is missing, but restaurants are observed. The county-pair fixed effect will capture the observed economic activity between locations in food services to help infer the amount of activity between areas for general merchandise stores. The right-hand side also includes a number of additional covariates, including revenues ($R_{j,n}$), distance ($distance_{i,j}$), population (pop_i), industry fixed-effects ($industry_n$). The function $f\,(\,)$ is specified as a flexible model that includes

20. The holdout method randomly divides the data into training and testing sets. To find the best model, each model is estimated using the training set only. Then the model is used to predict the output values for the data in testing set.

interactions of these variables and polynomials of distance. For instance, it includes polynomial of distance interacted with industry fixed effects and distance interacted with revenues and population. The model is specified as:

$$\log(S_{i,j,n}) = f(R_{j,n}, distance_{i,j}, pop_j, industry_n) + \tau_{i,j} + \varepsilon_{i,j,n}. \tag{6}$$

The term $\varepsilon_{i,j,n}$ is the error term. The imputed share is then calculated using the exponential of the expected value: $ImputedShare_{i,j,n} = \frac{exp(\widehat{\log(S_{i,j,n}))}}{\sum_i exp(\log(S_{i,j,n}))}$. For the relatively small number of areas where the county-pair fixed effects cannot be included, we use flexible linear regression models without fixed effects to impute these values.

Using cross-validation we test a variety of alternative models and examine the fit based on mean squared error and mean absolute deviation. We selected the methodology with the smallest mean squared error and mean absolute deviation based on a 5 percent holdout sample.

A.3 Instrumental Variable

In this section we outline the steps used to form the instrumental variables applied in the paper, following the work of Guren, McKay, Nakamura, and Steinsson (2020). Specifically, we estimate the following regression model:

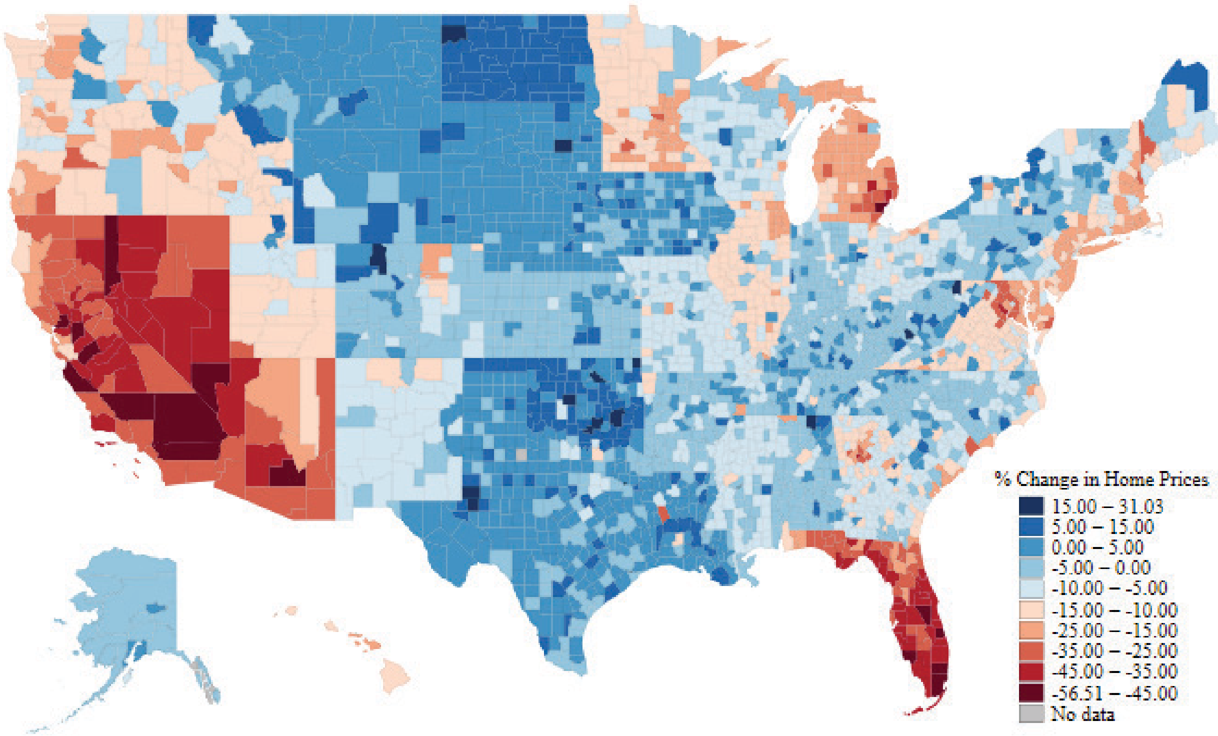
$$HousingPriceGrowth_{i,t}\, a_i + \beta_i \bullet RegionalHousingPrice_{R,t} + \beta \bullet X_{i,t} + \varepsilon_{i,t}$$

The model includes the average housing price growth of county *i* for two-years ending in year *t* on the left-hand side of the model. The right-hand side includes a county-level fixed-effect a_i and a county-level coefficient on the responsiveness of regional housing price movements β_i along with additional controls $X_{i,t}$, which are other factors that may influence local housing price changes. This includes two-digit industry share, growth in receipts at the county level and growth in receipts at the regional level.

A.4 Zillow Home Value Index

Zillow home value index (ZHVI) is seasonally adjusted measure of typical home value and marketchanges across a given region and housing type. Zillow publishes ZHVI for all single-family residences, for condo or coops, for all homes with 1, 2, 3, 4 and 5 and more bedrooms, and the ZHVI per square foot. The data is available at www.zillow.com/research/data. We focus on change in home prices using county-level data which covers approximately 2000 counties within the US for years 2006 and 2009. For the missing counties, mostly rural counties, we assume the price decline is equal to the median price decline across counties in the same state. Figure A.2 shows percent change in home prices across counties in the United States between December of 2006 and December of 2009 with darker shades of red indicating larger declines in home prices, while the darker shades of blue indicate a handful of counties that experienced an increase in home prices.

Figure A.2. Percent Change in Zillow Home Prices between 2006 and 2009



Source: Zillow website and authors' calculation