Impact of Cares Act Stimulus on Consumption: Evidence from Zip Code Level Transactions *

Zhen Yuan †

Abstract

We study the causal effect of the CARES Act on consumer expenditure during the Covid-19 pandemic with the Facteus data set. We use the transaction data at the Zip code level of over twelve million bank cards in eighteen central states to estimate the treatment effect of the stimulus. This question is identified because of the geographic and temporal variation in receiving the deposit across Zip codes. Our empirical findings suggest that in the first weekend receiving the stimulus, the overall spending increase by 53%. In particular, the expenditure on essential and financial categories increases by 62% and 76%. Unlike the tax rebates in 2001 and economic stimulus payments in 2008, people spend most of the payments on non-durable goods and services such as grocery, utility, money order, Etc. The estimated effect is substantial for districts with lower education, lower white ratio, and higher poverty, where poverty generates the most striking heterogeneity.

*We thank Vishal Singh for providing the Facetus data that made the analysis possible. We also thank Vishal Singh and Kotaro Yoshida for great advice.

†Corresponding Author: Computational Social Science program - Economics track, Social Science Division, The University of Chicago. Phone: (773)704-9287, email: zheny@uchicago.edu, Personal Website: yuan-zhen.com.

I Introduction

The outbreak of coronavirus disease 2019 disrupted people's lives immensely worldwide unprecedentedly. Intending to limit people's contact and slow the spread, the U.S. government announced stay-at-home orders to restrict people from leaving their homes unless necessary. At the same time, social distancing deteriorates the economy with significant losses in employment and a reduction in economic activities in many industries.

The U.S. government signed the CARES ACT into law that consists of various programs on March 27, 2020. One of the programs provides assistance for families and workers. Economic Impact Payments (henceforth called stimulus payment) were distributed to citizens and resident aliens with income of up to $ 75,000 for individuals, $112,500 for heads of household, or $ 150,000 for joint tax filers[14]. The first-round stimulus payments of up to $ 1,200 per adult and $ 500 per child (under 17) were deposited by IRS starting from April 10, 2020. By April 24, 2020, 89.5 million individuals had received the stimulus payments of around $ 160 billion from IRS for the first three weeks of this program [15].

Federal transfer policies naturally redistribute income across regions and households, with benefits tilted to low-income areas/households or areas/households receiving adverse shocks [20]. Some literature focuses on the marginal propensity to consume (MPC) of a household, i.e., the fraction of an extra dollar of aid that a household spends on consumption in response to the fiscal stimulus. Under the Economic Growth and Tax Relief Reconciliation Act of 2001, Johnson, Parker, and Souleles (2006) provide empirical evidence that the estimated MPC is largest for households with relatively low liquid wealth and low-income [16]. Parker et al. (2013) measure the response of household spending to economic stimulus payments (ESP) distributed in mid-2008. They observe a significant effect on the purchase of durable goods, where the response is more extensive for older and low-income households [19]. Hausman (2016) studies the veterans' bonus of 1936 and sheds light on that the spending was concentrated on durable goods such as cars and housing [12]. Romer and Romer (2016) show that the permanent increase in Social Security over the period 1952-1991 has a high impact around half a year, and vanish afterward [21]. Fuster, Kaplan, and Zafar (2018) use a hypothesis survey to identify MPCs in response to hypothetical cash windfall or loan and find heterogeneity and asymmetry in consumption behavior [11].

Many researchers have begun to study the impacts of the Covid-19 pandemic and the coordinating policies such as stay-at-home orders and the CARES Act on household consumption. Coibion, Gorodnichenko, and Weber (2020) estimate the causal effect of local lockdown policies identified by the difference in timing of lockdown using customized survey data. Their findings reveal that the lockdown accounts for effect on employment, consumer spending, and households' expectation on the economy [8]. Baker et al. (2020a) use transaction-level data from a personal finance website, SaverLife, and show that these events alter the users' spending in various categories. The responses depend on the local severity of Covid-19, and the demographic characteristics such as age and family structure [3]. Chetty et al. (2020) apply a novel platform of transaction data to analyze the economic impacts of Covid-19. Their findings reveal that high-income households reduce spending that requires physical interaction, which causes revenue losses in those sectors and layoffs of low-income [7]. Cox et al. (2020) analyze households' spending responses to the Cares Act stimulus with transaction-level data. Compared with economic stimulus programs in 2001 and 2008, they observe a faster effect, a minor effect on durable goods, and a more significant effect on food and financial payments. Chetty et al. (2020) apply a novel platform of transaction data to analyze the economic impacts of Covid-19. Their findings reveal that high-income households reduce spending that requires physical interaction, which causes revenue losses in those sectors and layoffs of low-income [7].

The literature mentioned above mainly focuses on heterogeneity in responses across households, but little is known about the potential geographical variation along with categorical variation under the effect of the fiscal stimulus. This ignored perspective is significant in the following aspects. First, geographical disparity systematically exists across the United States. For example, the Consumer Price Index (CPI) and population composition by age probably affect people's response to the stimulus. Districts with low price levels (e.g., Jackson, Mississippi) might spend less of the stimulus than districts with high price levels (e.g., NYC, New York). Besides, young people are more liquidity constrained than the elder, so areas with more millennial populations (e.g., Seattle, Washington) might need the rebates more urgently than those with more seniors (e.g., Augusta, Maine). Second, people living in different places might spend the payments disparately. People living in cities with higher rent-to-income ratios might spend more of the stimulus on the rents. Additionally, Blacks and Hispanics spend larger shares of their expenditure on visible goods such as clothing, jewelry, and cars than comparable Whites (Charles, Hurst, and Roussanov, 2009)[5]. Thus, areas with more black or Hispanic neighborhoods might spend a more significant proportion of the stimulus on unnecessary goods during the pandemic. Third, the impact of the pandemic varies much across the United States due to Covid-19 severity, stay-at-home orders, and the corresponding economic recession. The deposits are more needed to regions with manufacturing and tourism as the top industries. Our main contribution is to address the disparity in response to the stimulus across geographical locations and categories of spending under a natural experiment.

Our analysis is based on the Facteus data set. Two recent papers study the EIP payments using the same data. Misra, Singh, and Zhang (2020) use Zip code level daily transactions and estimate the MPC of the stimulus, and found Large cross-sectional variation [18]. Kager and Rajan (2020) analyze the transaction-level data and show MPC heterogeneity of recipients with different levels of income [17]. One paper has a similar setting to ours. Parker et al. (2013) [19] use consumer expenditure survey data to measure the response of household spending to economic stimulus payments distributed in mid-2008 and variation in randomized timing of the distribution of the payments. For the CARES Act, the payments were made by direct deposit from April 10 to April 15 when deposit information is available. Otherwise, paper checks would be mailed starting from April 24. The timing of deposit varied due to the schedule of IRS and different banks and financial institutes (Baker et al.,2020b) [4]. In our sample, some Zip codes received the payments earlier than others. The treatment Zip codes received the deposit on April 10, and the deposits in control Zip codes were made since April 13. We believe that a control group of Zip codes that do not receive stimulus on April 10 could be a natural basis for comparison with the stimulus experiences in the treatment Zip codes. Despite the significance of MPC, our paper estimate the immediate treatment effect of the stimulus on the Zip code level aggregated spending during this time interval with a two-way fixed-effect model. The identification will be discussed in the empirical methodology section.

Beyond the average treatment effect (ATE), we aim to study the cross-sectional heterogeneous response to the stimulus with machine learning methods. There has been a surge of literature in economics and marketing research applying statistical machine learning under the big data setting. Athey and Imbens (2016) provide a data-driven approach to partition data into subspace by constructing trees to estimating heterogeneity in causal effect in randomized experiments and observational studies when unconfoundedness assumption satisfies [2]. Wager and Athey (2017) propose a class of non-parametric causal forests for estimating heterogeneous treatment effect and show the causal forests are pointwise consistent for the true treatment effect under unconfoundedness assumption [23]. Histch and Misra (2018) introduce a direct uplift modeling approach, named causal KNN, to evaluate the profit of targeting policies. They find methods that directly predict the conditional average treatment effect (CATE) outperform methods that predict CATE indirectly via conditional expectation function [13]. We follow the approach presented in Histch and Misra (2018) in this paper. For the application, the propensity score means the probability of receiving the stimulus on April 10 for a Zip code. Then we select k treated and k untreated neighbors for each observation. The nearest neighbors are selected by the covariates, where we calculate the Euclidean distance of the standardized data. Then the treatment effect for a unit is defined as the difference in mean outcome between the k treated units and the k untreated units, which means we estimate CATE for each Zip code. Afterward, we group the observations by the selected characteristics to study the heterogeneity across Zip codes.

The rest of the paper proceeds as follows. Section 2 describes the data. Section 3 contains the identification and empirical strategies. Section 4 summarizes the results. Section 5 concludes with a discussion of the implications and limitations.

II Data

Our data comes from Facteus. Facteus is a data aggregation firm that provides daily transactions and spending at the Zip code level in the United States. Specifically, the data mainly comes from three sources: challenger banks, payroll cards, and government cards. Because of the characteristics of these cards, the majority of consumers in our study are young and low-income. We use Friday, Saturday, and Sunday transactions from 10921 zip codes belonging to 18 Midwestern and Southern states in the United States between March 7, 2020, and April 12, 2020. These states are Arkansas (AR), Iowa (IA), Kentucky (KY), Nebraska (NE), North Carolina (NC), North Dakota (ND), Oklahoma (OK), South Dakota (SD), Tennessee (TN), Illinois (IL), Indiana (IN), Kansas (KS), Minnesota (MN), Missouri (MO), Ohio (OH), Wisconsin (WI), and West Virginia (WV). Figures 1 panel (a) reports the geographic distribution of the treatment group and control group in these states, where some Zip codes in MO, MN, WI, KS, OH, OH, VA, IN, and WV receive the stimulus deposits on April 10. There are 2490 zip codes in the treatment group and 8431 zip codes in the control group. Panel (d) displays the distribution of the difference in weekend spending across Zip codes between the week of April 10 and the week of April 3. We calculate the difference of days in the same day-of-week to absorb the time fixed effect. The states with treatment Zip codes have an average increase in spending between 50% and 100%, while the spending in states without treatment Zip codes does not change on average. We plot the average number of transactions and spending in January and March to mid-April in panel (b) and panel (c). The transactions and spending are stable during January because it is not affected by the Covid19 pandemic and stay-at-home orders, while we notice the transactions and spending declines gradually and revive after the stimulus.

The data reports daily aggregated transactions and expenditures of 24 kinds of goods and services (See Appendix A.1). We further define them into four categories: essential, non-essential, financial, and others. Essential category includes utility, discount stores, grocery stores, health care, insurance and government services. Non-essential category includes taxis, variety stores, other stores, foreign currency and travel cheques, restaurants, and automobile dealers. Financial category includes money orders and wire transfer, manual cash disbursements, and merchandise and services. Other category includes miscellaneous goods and services that we are less interested in, and many of them are uncertain in categories across different districts and households. Note that essential goods in normal time such as travel- and commute-related can become non-essential anymore because they are restricted by the stay-at-home orders.

Factus provides this type of data from December 2019, and we only cover transactions in March 2020 and April 2020 to avoid other tax rebates. Figure 2 shows the change in spending overall and the four categories with the week of March 6 as the base week. In Panel (a), we observe that the overall spending of both the treatment group and control group follows the same pattern before the occurrence of stimulus payments: it declines gradually to around 25% of the original spending in the week of March 6. After receiving the treatment, the spending of the treatment group spikes with the highest increase on the second day of the payments, about 90%, while the control group slightly recovers but does not surpasses the original level. Panel (b) and Panel (c) respectively report the change in spending for the control and treatment groups. Likewise, each type of expenditure follows the same pattern in the pre-treatment period. The essential category is the least affected, and the financial category is the most affected. However, after the receiving rebates, essential and financial spending has nearly the same rise, achieving around 140% on the second of the treatment. Non-essential spending climbs mildly and surpasses the original point. The consumption of the control group is relatively unaffected, except that there is a moderate increase in essential spending. Furthermore, we also compare the spending composition of the treatment group and control group in pre-and post-treatment periods. Before April 10, these two groups have about the same proportion of spending on these four categories of goods and services: 33% on essential, 28% on financial, 26% on non-essential, and 21% on others. From April 10, the control group's spending remains stable, but that of the treatment group flows from non-essential and others to essential and financial. Although absolute spending does not drop, people spend most of the deposit on the essential and financial categories. Besides, these plots indicate that the treatment group and the control group are identical in many aspects, such as the pre-treatment trend and spending composition, which further imply that the treatment is randomly assigned.

There are several supplementary sources of data in research. First, the Zip code level characteristics, including demographics, income, education, commute, and rent are from U.S. Census and ACS. This information is significant in comparing the similarity of Zip codes and studying the heterogeneity in spending. Second, the monthly county-level employment data are from the U.S. Bureau of Labor Statistics (BLS). Third, the daily county-level Covid-19 cases and deaths are from New York Times. The employment and Covid-19 data are used in the robustness check.

III Empirical Methodology

III.I Identifying Strategies

Panel Data Model Identification in this paper is based on the variation of time in distributing the stimulus checks to different Zip codes. Figure 1 reports the geographic variation in stimulus payments and spending. In panel (a), the treatment group includes the Zip codes in red that received deposits on April 10 (Friday), and the control group is those receiving deposits starting from April 13 (Monday). Therefore, we compare the consumption of Zip codes that receiving checks at a different time within the same county to estimate the average treatment effect of the stimulus on local consumption. As mentioned in the introduction, we address three identification problems: dependent variable construction, reverse causality, and omitted variable bias.

Since this data set does not tell us about the sampling process, we know little about some essential information such as the number of consumers in each Zip code, the ratio of consumers eligible for the stimulus, etc. Given that aggregated consumption depends on the population, we normalize the consumption to make the expenditure comparable across Zip codes. One approximation for the population would be the mean number of Zip code level daily transactions in January. The daily transaction is a better instrument than daily consumption because the price level varies across regions. Besides, using population might be problematic because the consumers in the sample are not representative enough of the regions. Since the first cases of Covid-19 in the United States were known in February, then the consumption in January should be independent of Covid-19 severity. Therefore, our normalized spending is defined as the Zip code level spending divided by the mean number of daily transactions in this Zip code in January.

Since fiscal payments are often countercyclical, it would be a problem if consumption induces the disbursement of the CARES ACT rebates. If our estimator picks up both the treatment effect of the stimulus and the reverse effect, then it is underestimated because the reverse effect is usually harmful. In our case, reverse causality is not a concern because the CARES Act was passed on March 25. As we observe Zip code level daily transactions, the establishment and eligibility of the CARES Act are unaffected by the consumption on April 10.

Another concern is the omitted variable bias, where some confounding variables that we do no observe as researchers can affect the treatment. We control the county fixed effect to remove the county level trend because people living in some counties shop more than others. Besides, since people often shop across Zip code but not across counties, we control the fixed effect at the county level instead of Zip code level. We also add the time fixed effect to absorb aggregate variation in consumption due to people's preference in shopping on different days-of-week and weeks. Therefore, the only possible confounding factors are local that means both time- and cross-sectional-variant. Because qualified households and people all over the countries will receive the stimulus eventually, the CARES Act did not aim at certain counties or states. What might be problematic is the timing of the deposit, i.e., whether some counties have the priority of receiving the rebates if their local economies are among the most affected. Strategies to reduce omitted variable bias related to this reason are twofold: 1) we add controls varying over time and counties; 2) we conduct robustness check by manifold placebo tests. Such controls include Covid-19 cases/deaths and unemployment rate, as the Covid-19 severity and unemployment affect the local economy, which might shift the ordering of the disbursement.

Heterogeneous Treatment Effect We are also interested in the Conditional Average Treatment Effect (CATE), where $CATE(x) = E[Y(1) - Y(0) \mid X = x]$. Rubin (1974) fomalizes the notions for observational studies and experiments, $Y(0)$ and $Y(1)$ are potential outcomes in the control state and treatment state [22]. Since only one of them can be observed, the estimation process requires us to solve the missing data problem. Following the framework in Hitsch and Misra (2018), the observation set is $D = (Y_{z,t}, X_z, W_z)_{z=1,t=0}^{N,T}$, where N is the total number of Zip codes. Besides, t represent the number of days after the treatment, and $T = \{0, 1, 2\}$ in the data set. The challenge in obtaining causal inference is that $Y_{z,t}(1) - Y_{z,t}(0)$ is never observed, then we need to derive the counterfactual outcomes $Y_{z,t}^{mis} = Y_{z,t}(1)(1 - W_z) + Y_{z,t}(0)(W_z)$. The CATE estimator is identified under three conditions: unconfoundedness, overlap assumption, and Stable Unit Treatment Value Assumption (SUTVA).

In this context, *unconfoundedness* allows different Zip codes to have different probabilities of treatment. After we control for the covariates, the treatment is as good as random. It also means that the treatment is uncorrelated with unobservables that can affect the observed outcomes. For example, one unobservable can be the price level that impacts the expending, and we require the geographic and temporal variation in the distribution of stimulus is independent of the price level.

$$Y_{z,t}(0), Y_{z,t}(1) \perp W \mid X \tag{1}$$

Under *overlap* assumption, $e(x)$ is defined as the propensity score or targeting probability. For general application, the propensity score has to be defined strictly between zero and one. If there is no treated or untreated observation for specific values of the feature vector x_i, then this assumption is violated.

$$0 < e(x) < 1 \tag{2}$$

The *SUTVA* assumption requires well-defined treatment and no spillover effects. As for well-defined treatment, the treatment scale should be identical in forms and magnitude across all units in the data set. Since the data is observed in the zip code level, the money received varies across different districts depending on the number of people qualified for the program. To normalize the spending and treatment, we take the mean number of transactions to approximate the number of people at the zip code level in this data set. This is appropriate because 1) the number of transactions reflects population and consumption habits for a district; 2) the United States remained unaffected by Covid-19 during January. Spillover effects mean that the stimulus for one zip code affects the spending in another. This becomes a concern if people spend their money in other zip codes. However, since the transactions are from a panel of consumers' cards other than that of merchants, then the spending of people who receive the stimulus will be recorded in the zip code they live in.

Similarly to the event study, we need to consider the dynamic trend of the treatment effect. The CATE estimator $\tau(x)$ for each t is identified as the following.

$$\tau(x) = E[Y_z \mid X_z, W_z = 1] - E[Y_z \mid X_z, W_z = 1] \tag{3}$$

III.II Empirical Specification

Two-way Fixed Effect Model The identification of the panel model is based on the variation in treatment for Zip codes within the same county. Our parsimonious specification follows a two-way fixed effect model

$$Y_{z,t} = W_{z,t}\beta + \alpha_{c(z)} + \delta_t + X_{z,t} + \epsilon_{z,t} \tag{4}$$

where $Y_{z,t}$ is the normalized spending in Zip code z on day t. The left-hand side variable has different measures: overall spending and spending on four types of goods and services. As mentioned in the introduction, these types are essential, non-essential, financial and others. $W_{i,t}$ indicates whether zip code i has received stimulus on day t. $\alpha_{c(z)}$ represents the county fixed effect, and δ_t is the time fixed effect. $X_{z,t}$ is a vector of cross-sectional and time varying covariates that we mutate in different specifications including Covid-19 cases and unemployment rate. In this model, parameter β is the main coefficient of interests, which reveals the difference in spending after the treatment. More than that, we are able to compare the changes for different types of goods, i.e. where households and people spend the rebates immediate after receiving the money.

However, the effect of treatment might not be homogeneous across Zip codes. As an extension to equation 4, we run the following specification

$$Y_{z,t} = W_{z,t}\beta + W_{z,t} \times U_z\gamma + \alpha_{c(z)} + \delta_t + X_{z,t} + \epsilon_{z,t} \tag{5}$$

where U_z is a Zip code level feature vector, and γ indicates how the effect size change for different levels of features. It is a preliminary estimation for the heterogeneity in response. The features include poverty ratio, white ratio, and college education ratio (see Appendix A.1). Most of the characteristics in the list are correlated, while these three are the most significant.

Heterogeneous Treatment Effect We consider causal KNN that is a method directly predicting CATE, with which we can compare the treatment effect across groups. For the vector of covariates X, we find the K nearest treated neighbors and K nearest untreated neighbors. Then, the CATE is estimated by the mean difference of outcomes between the treated and untreated units,

$$\hat{\tau_K}(X) = \frac{1}{K}\Big(\sum_{i \in N_K(X,W=1)} Y_i - \sum_{i \in N_K(X,W=0)} Y_i\Big) \tag{6}$$

where $N_K(X, W = 1)$ and $N_K(X, W = 0)$ is the sets of the K treated and K untreated nearest neighbors respectively. The estimator $\hat{\tau_K}(X)$ depends on the choice of K, and the value of K is based on the transformed outcome loss, which is the mean-squared difference between $\hat{\tau_K}(X)$ and the transformed outcome Y^*. Athey and Imbens (2016) [2] define Y^* as the proxy for the true CATE,

$$Y_i^* = W_i \frac{Y_i(1)}{e(X_i)} - (1 - W_i)\frac{Y_i(0)}{1 - e(X_i)} = \frac{W_i - e(X_i)}{e(X_i)(1 - e(X_i))} Y_i \tag{7}$$

where the propensity score $e(X) = Pr(W = 1 \mid X)$. Hitsch and Misra (2018) show that $E[Y_i^* \mid X_i = x] = \tau(x)$ under unconfoundedness, which means the transformed outcome is an unbiased estimate for CATE. The transformed outcome can be decomposed into infeasible loss and variance of the residual that is uncorrelated with K, then the K value minimizes the transformed outcome loss also minimizes infeasible loss - $E[(\tau(X)^* - \tau_K\hat{(}X_i))^2]$. In application, we choose K that minimizes sample asymptotics - $E_n[(Y_i^* - \tau_K\hat{(}X_i))^2]$. $\tau_K\hat{(}X_i)$ is determined in the training set, while the transformed outcome loss is assessed in the test set.

IV Empirical Results

IV.I Main Results

Two-way Fixed Effect Model Table 1 - 5 report the results for the panel data model. The baseline models are presented in panel (a) of these tables, while panel (b) - (d) address the heterogeneity across Zip codes that the parsimonious might miss. For the overall consumption, receiving the stimulus checks increases the spending by 53% ($e^{0.427} - 1$). The effect size ranging from 52% to 54% depending on the specifications, which means the estimation is stable and the specification is good. When the white ratio of a Zip code increase by 1 standard error, the overall treatment effect decreases by 3%. It is worth mentioning that the white ratio negatively correlates black ratio, which means the treatment effect increases with black ratio. Besides, a 1 standard error increase in poverty ratio lifts the treatment effect by 4%. Thus, when the poverty ratio of a district is high, they tend to spend more than usual once they receive the stimulus. College education ratio corresponds to factors such as median income and median gross rent. Zip codes with 1 standard error higher college ratio than mean would obtain 4% lower effect size.

In Table 2 - 5, we explore the heterogeneous treatment effect on decomposed classification of goods by re-estimating the models for overall spending. We note that the effect size is larger for essential (62%) and financial (76%) goods and services, while it is smaller for non-essential (20%) and other (28%) goods and services. It indicates that after receiving the stimulus, people proportionally increase the spending more on essential category such as food and supplies, and financial category including money order, cash disbursement, etc. As for the cross-sectional heterogeneity on different categories, they exhibit similar patterns as the overall spending with some variation in the effect size. These three features play significant roles in affecting essential and financial spending. One standard error higher white ratio and would lead to 2% and 7% less spending on essential and financial spending respectively, while Zip codes with 1 standard error higher poverty ratio will spend 4% more on essential goods and 8% more on financial goods after the treatment. Additionally, Zip codes would end up with 5% and 6% less spending on essential and financial goods when they have 1 standard error higher college education ratio than the average. The impact of these features on non-essential and other spending is either less economically significant or statistically significant.

Heterogeneous Treatment Effect Figure 3 displays the CATE estimator in causal KNN model with optimal k value in the test data set. To extract valuable insights from the data, we calculate CATE by categorizing the observations with respect to three dimensions (race, poverty, and education) into three groups: below 20%, 20%-80%, and above 80%. After finding the most accurate estimation of the CATE, it is possible to extract valuable insights from the data. Again, CATE indicates that the difference between the treatment group and control group given the covariates. The overall CATE is 0.55, which indicates a 73% increase in spending because of the treatment. The effects on essential and financial goods are larger than overall spending, i.e. 82%, and 120%; in other words, the effects on non-essential and other goods are respectively 32% and 49% that are smaller than overall spending. This corresponds to the expectation of the policy makers, since their aim is to prompt people to spend money on the most needed goods and services. The categorized groups exhibit identical heterogeneity in overall spending and spending across the four types of goods. Higher poverty ratio result in larger effect size, but higher education level and white ratio lead to smaller effect size. After receiving the stimulus, top 20% poverty districts increase 60% more on spending than the bottom 20% poverty districts, while bottom 20% education level and white ratio districts respectively enhance 24% and 19% more on spending than top 20% education level and white ratio districts.

Among the three features, poverty ratio generates the most striking difference in CATE. As for the poverty ratio, top 20% Zip codes increase essential and financial spending by 136% and 215%, but bot 20% Zip codes only increase essential and financial spending by 46% and 49%. However, the top 20% districts only increase the spending 35% ($\frac{1+58\%}{1+17\%}$) than the bottom 20% districts on non-essential goods and services. Besides, the bottom 20% education level districts increases the essential spending by 114%, and the top 20% education level districts expand their the essential spending by 55%. For the rest types of goods and services, the education level does not product such a contrasting comparison. Similarly, white ratio plays an important role in financial goods and services with an disparity of 200% for the bottom 20% versus 57% for the top 20%, but the effect sizes are close in the else categories of spending. Other goods and services are those are either essential or non-essential heavily depending on the circumstances, so it is not surprising that the CATE estimation also lies in between that of essential goods and that of non-essential goods.

Therefore, for the purpose of stimulus expenditures in essential and financial goods and services, the future policies should target districts with low education level, high poverty ratio, and low white ratio. Plus these three groups of Zip codes share the same preference on different categories of spending, with the rank $other > essential > others > non-essential$.

IV.II Robustness Tests

Dynamic Controls In panel (f) of Table 1 - 5, we add additional time- and cross-sectional varying covariates that might have impact on both the treatment and the spending. Potential confoundings can be factors that change overtime within the same states, and probably affect IRS decision on the order of disbursing the stimulus checks. The variables we select are about local employment and Covid-19 severity including unemployment rate, Covid-19 cases, and Covid-19 deaths. It turns out that the estimated treatment effect is hardly affected by these controls, since the coefficients are close, and the standard errors remain unchanged. For the covariates, their coefficients are close to zero, and most of which are marginally significant or insignificant. Therefore, the potential time- and cross-sectional varying confounding does affect much of our estimation.

Alternative Sample It is still possible that the estimated effect is driven by certain Zip codes or days of the week. I consider a subset of Zip codes that does not receive the stimulus payments on April 10. Then the Zip codes receiving stimulus on April 13 becomes the treatment group, and the rest of them are the control group. Table 6 reports the estimates on this subsample, and the effect size does not vary much. The treatment rises the overall consumption by 50%, compared with 53% in the first specification. The estimates for the poverty-stimulus and education-stimulus interaction is close in terms of coefficients and standard error, except that the white-stimulus interaction becomes insignificant. However, given that switching the sample has little effect on the overall estimation, plus race is not an economically influential feature, we evaluate the estimation being stable. Besides, while we include observations of Friday, Saturday, and Sunday in the original sample and those of Monday and Tuesday in the alternative sample, which further amplify the robustness of our estimation. However, we do not compare the Zip codes receiving the stimulus payments on April 15 and April 17, because there is no enough variation between Zip codes within the same county. Moreover, the estimation remains unaffected when we add potential confounding variables.

Placebo Effect Karger and Rajan (2020) [17] and Misra, Singh, and Zhang (2021) [18] reports no evidence for anticipatory effect on spending in leading days up to the payments. We still worry about the effect of the stimulus during the days around the payments. Following Pennings (2021) [20], we conduct two placebo tests on the overall effect. First, because the government passed the CARES Act on March 15, we run a series of regression on observations from March 2 to March 19 where the stimulus payments were claimed but not deposited. Then, March 15 is the day of treatment in this sample, so we aim to explore that whether people in Zip codes receiving stimulus on April 10 started to spend more once the act was notified. The result is displayed in Table 7, and we observe small but significant treatment effect. This might be because we do not perfectly control the participants in each Zip code with the transactions in January. However, the effect size is around 7%, much smaller compared to 53% in the two-way fixed effect model.

Second, we counterfactually move the payments backward and forward to detect the potential spurious effect by a placebo regression. We include both leads ($t < 0$) and lags ($t > 0$) of the stimulus in the two-way fixed effect model like Dasgupta, Gawande and Kapur (2017) [9], and Fouirnaies and Mutlu-Eren (2015)[10]. We expect to have small and insignificant anticipatory effect($t < 0$), but positive and significant delayed effect ($t > 0$). The delay effect is permitted because people usually do not spend up money once they receive it. Figure 4 shows the results of placebo tests on the full sample and alternative sample. The number of leads and lags depends the number of days each week for that data set, where full sample covers Friday, Saturday, and Sunday, and alternative sample include Monday and Tuesday. All lags are significant, while the leads are close to zero and marginally significant in full sample and insignificant in alternative sample. The chance of no false positive in leads is 77%, which is a high probability. Then, the marginally significance is more likely to be because of other confoundings such as data failure in properly control number of participants. The largest effect size occurs on the second day of treatment (Saturday) in the original sample, and on the day of treatment (Monday) in the alternative sample.

Above all, the main results of the two-way fixed effect model are robust. Since we only apply one machine learning method in this paper, we ignore the comparison part of performance.

V Conclusion

In this paper, we analyze the immediate causal effect of the first round CARES Act stimulus in the United States. Utilizing the geographic and temporal variation in the disbursement of the deposit due to the procedures of IRS and different banks, we obtain a natural experiment setting such that we are able to estimate the immediate average treatment effect (ATE) and conditional average treatment effect (CATE). By exploring Factus data, while Misra, Singh, and Zhang (2021) and Kager and Rajan (2020) focus on the marginal propensity to consume (MPC), we aim to explore heterogeneity in Zip codes characteristics and expenditure categories. We show that the overall spending increases by 53% in the first three days after the treatment. As for the four categories, essential and financial spending increase 62% and 76% respectively, which indicates people spend a large proportion of the payments on non-durable goods such as grocery and utilities, and financial services like money orders and wire transfer. Regarding the heterogeneity in three dimensions of the Zip code level characteristics, the estimated treatment effect exhibits the same pattern in the overall spending and spending of the four categories. The effect size is highest for Zip codes with top 20% poverty ratio, bottom 20% college education ratio, and bottom 20% white ratio. Poverty is the dominant factor, where the top 20% poverty ratio Zip codes enhance the overall, essential, and financial spending by 101%, 136%, and 215%. Zip codes with bottom 20% college education ratio and white ratio group increase 90% and 93% overall, 114% and 105% in essential category, and 153% and 200% in financial category. Besides, it seems that there is no group of Zip codes that has a uncommon preference towards particular categories of goods and services. For the purpose of subsidizing the spending on essential and financial products of people and households, policy makers are supposed to focus on the individuals who are low-income, less-educated, and living in black or minority neighborhoods. Since poverty ratio is a key driver of heterogeneity in this research, it corresponds to the CARES Act that the qualified households or individuals are those with income lower than the threshold.

There are some policy implications related to our findings in the heterogeneous treatment effect on spending. First, the most dominant spike in spending is on financial goods and services, which indicates that people suffer from constraints in liquid balances. Besides, the increase in essential spending (necessary during the pandemic) is dramatic. While most of the economic stimulus payments disbursed in mid-2008 were spent on durable goods, especially vehicles (Parker et al. 2013 [19]), people have a stronger intention to spend on non-durable and financial goods after receiving stimulus in 2020. It might reflect the difference in the macroeconomic situations in 2008 and 2020, where low-skilled service workers bore the most of the economic impact from the Covid-19 pandemic (Althoff et al. 2020 [1]). Since the stimulus was distributed to individuals and households whose income is lower than a threshold, vulnerable workers who might lose their jobs badly needed the stimulus to pay for food, supplies, and mortgage. Besides cash rebates, the government could consider disbursement of daily necessities, plus forbearance in mortgages and rents, in which way the non-essential spending would be avoided. Second, the difference in poverty ratio generates the most striking disparity in effect size. The top 20% poverty Zip codes have incremental effects 62% and 112% higher than the bottom 20% poverty Zip codes in essential and financial spending. This finding amplifies the decision to make income as the eligibility standard because the government's motivation is to relieve the cash flow pressure of households and individuals and guarantee their necessary needs. Third, we also consider whether the consumers are rational in spending the stimulus rebates. Ideally, hand-to-mouth people should spend as much money on the essential and financial categories as possible during a hard time. When we compare the incremental effect ratio of essential/financial and non-essential spending, education plays the most significant role in distinguishing people's rationality. This finding is related to projection bias. Projection bias means people falsely project current preferences onto a future event that usually exists in consumer behaviors. For example, people might have spent in the non-essential category such as liquor stores and variety stores because they have more budget after receiving the rebates. Our results suggest that the relative effect size of essential and financial to non-essential top 20% educated Zip codes is 30% and 50% higher than that of the bottom 20% educated Zip codes. Our result is following that of Chen, Moskowitz, and Shue (2006)[6] that education can reduce biases in decisions. Policymakers can show the recipients essays or videos about the rational way of spending the stimulus in economics perspective on the IRS website, which might have a nudge effect on reducing the unnecessary expenditure.

To sum up, we state the caveats and the future direction of this research. The analysis is based on the debit card transaction of relatively young and lower- or mid-income cardholders. Thus, the sample is not representative concerning the characteristics of the Zip codes and the population in the United States. Besides, because the individual-level transaction is not observed, the inference derived from Zip code level transactions might bring aggregation bias. In future research, we hope to explore the individual or household level transaction for more precise inference. In addition, we do not include the pre-treatment observations in the causal KNN model, and then it means we ignore the pre-event trend. The CATE is consistently overestimated than the two-way fixed effect estimator, probably because treatment Zip codes always spend more than the control Zip codes. Thus, the instrument, average January transactions, does not fully capture the missing information about the number of participants in each Zip codes. Given more time, we would consider models that can adequately embed the pre-treatment trend. Despite the importance of dynamic effect and general equilibrium to economists, we cannot address it in this paper because of the data constraint. Access to the data spanning more extended periods allows us to estimate the temporary and permanent effect of the CARES Act on the local and national economies.

References

[1] Lukas Althoff, Fabian Eckert, Sharat Ganapati, and Conor Walsh. The city paradox: Skilled services and remote work. 2020.

[2] Susan Athey and Guido Imbens. Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences*, 113(27):7353–7360, 2016.

[3] Scott R Baker, Robert A Farrokhnia, Steffen Meyer, Michaela Pagel, and Constantine Yannelis. How does household spending respond to an epidemic? consumption during the 2020 covid-19 pandemic. *The Review of Asset Pricing Studies*, 10(4):834–862, 2020.

[4] Scott R Baker, Robert A Farrokhnia, Steffen Meyer, Michaela Pagel, and Constantine Yannelis. Income, liquidity, and the consumption response to the 2020 economic stimulus payments. Technical report, National Bureau of Economic Research, 2020.

[5] Kerwin Kofi Charles, Erik Hurst, and Nikolai Roussanov. Conspicuous consumption and race. *The Quarterly Journal of Economics*, 124(2):425–467, 2009.

[6] Daniel L Chen, Tobias J Moskowitz, and Kelly Shue. Decision making under the gamblerâs fallacy: Evidence from asylum judges, loan officers, and baseball umpires. *The Quarterly Journal of Economics*, 131(3):1181–1242, 2016.

[7] Raj Chetty, John Friedman, Nathaniel Hendren, Michael Stepner, et al. How did covid-19 and stabilization policies affect spending and employment? a new real-time economic tracker based on private sector data. *NBER working paper*, (w27431), 2020.

[8] Olivier Coibion, Yuriy Gorodnichenko, and Michael Weber. The cost of the covid-19 crisis: Lockdowns, macroeconomic expectations, and consumer spending. Technical report, National Bureau of Economic Research, 2020.

[9] Aditya Dasgupta, Kishore Gawande, and Devesh Kapur. (when) do antipoverty programs reduce violence? india's rural employment guarantee and maoist conflict. *International organization*, 71(3):605–632, 2017.

[10] Alexander Fouirnaies and Hande Mutlu-Eren. English bacon: Copartisan bias in intergovernmental grant allocation in england. *The Journal of Politics*, 77(3):805–817, 2015.

[11] Andreas Fuster, Greg Kaplan, and Basit Zafar. What would you do with $500? spending responses to gains, losses, news and loans. Technical report, National Bureau of Economic Research, 2018.

[12] Joshua K Hausman. Fiscal policy and economic recovery: The case of the 1936 veterans' bonus. *American Economic Review*, 106(4):1100–1143, 2016.

[13] Günter J Hitsch and Sanjog Misra. Heterogeneous treatment effects and optimal targeting policy evaluation. *Available at SSRN 3111957*, 2018.

[14] IRS.GOV. *159 million Economic Impact Payments processed*, 2020. [Online].

[15] IRS.GOV. *Treasury, IRS deliver 89.5 million Economic Impact Payments in first three weeks, release state-by-state Economic Impact Payment figures*, 2020. [Online].

[16] David S Johnson, Jonathan A Parker, and Nicholas S Souleles. Household expenditure and the income tax rebates of 2001. *American Economic Review*, 96(5):1589–1610, 2006.

[17] Ezra Karger and Aastha Rajan. Heterogeneity in the marginal propensity to consume: evidence from covid-19 stimulus payments. 2020.

[18] Kanishka Misra, Vishal Singh, and Qianyun Poppy Zhang. Impact of stay-at-home-orders and cost-of-living on stimulus response: Evidence from the cares act. *Available at SSRN 3631197*, 2021.

[19] Jonathan A Parker, Nicholas S Souleles, David S Johnson, and Robert McClelland. Consumer spending and the economic stimulus payments of 2008. *American Economic Review*, 103(6):2530–53, 2013.

[20] Steven Pennings. Cross-region transfer multipliers in a monetary union: Evidence from social security and stimulus payments. *American Economic Review*, 111(5):1689–1719, 2021.

[21] Christina D Romer and David H Romer. Transfer payments and the macroeconomy: The effects of social security benefit increases, 1952-1991. *American Economic Journal: Macroeconomics*, 8(4):1–42, 2016.

[22] Donald B Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, 66(5):688, 1974.

[23] Stefan Wager and Susan Athey. Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523):1228–1242, 2018.

Figures

Figure 1

Panel (a) shows the distribution of Zip codes in the treatment group and control group. Panel (b) shows the sum of spending and number of transactions across Zip codes in January 2020. Panel (c) shows the sum of spending and number of transactions across Zip codes from March 6, 2020 to April 12, 2020. Panel (d) shows the difference in weekend spending across states, where the difference indicates the difference of spending in week of April 10 and spending in week of April 3.

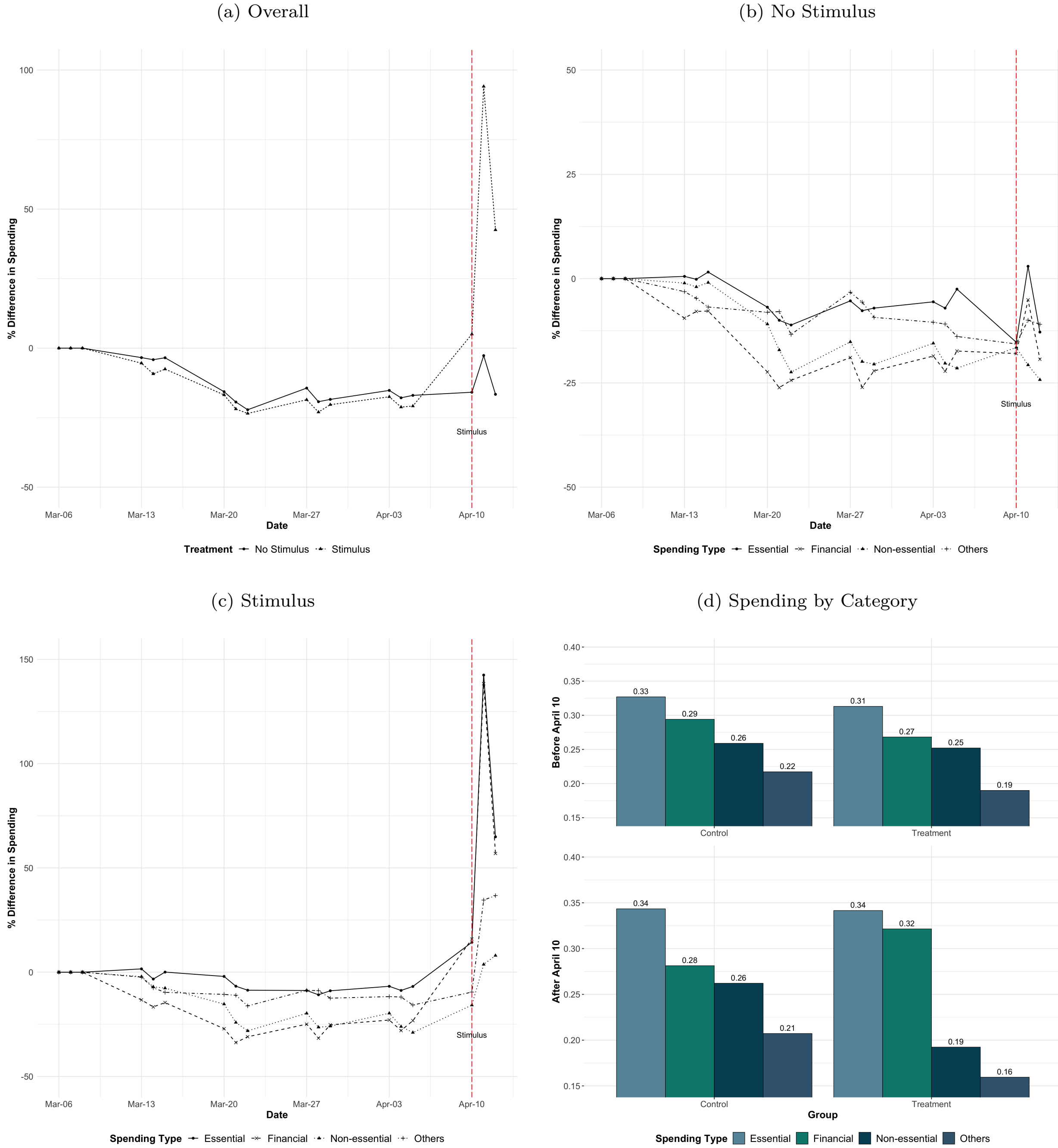
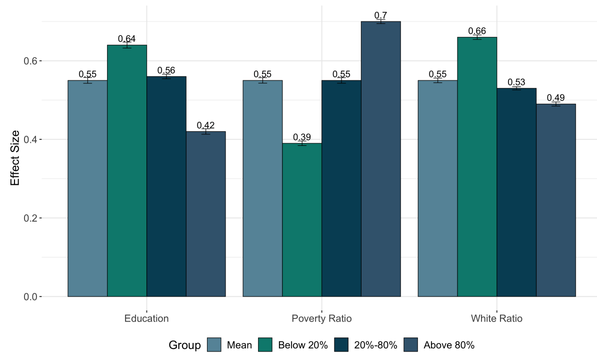


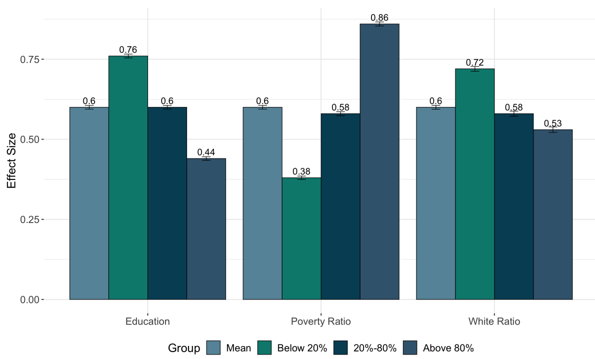
Figure 2

Panel (a) shows the overall spending change from March 6 to April 12 for the treatment group and the control group. Panel (b) shows spending change for each type of goods and services for the treatment group. Panel (c) shows spending change for each type of goods and services for the control group. Panel (d) shows the fraction of spending on each type of goods and services for the per-treatment period and post-treatment period for the treatment group and the control group.

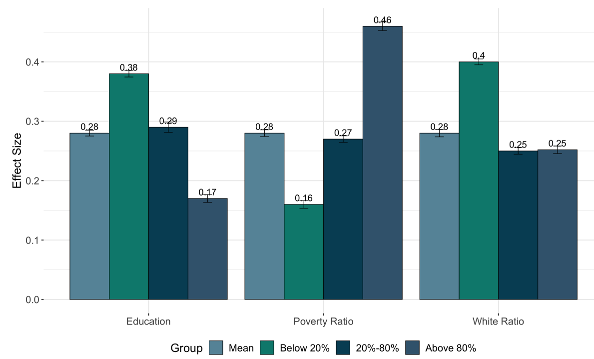
(a) Overall



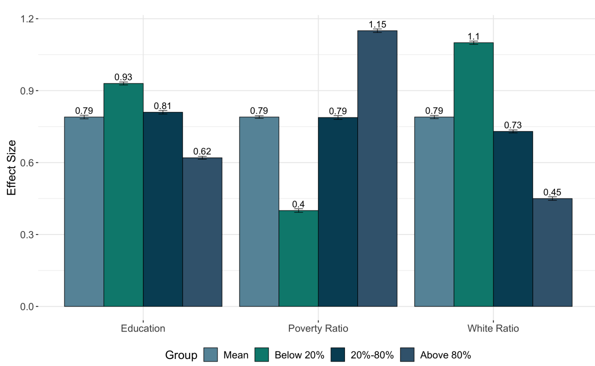
(b) Essential Goods and Services



(c) Non-essential Goods and Services



(d) Financial Goods and Services



(d) Other Goods and Services

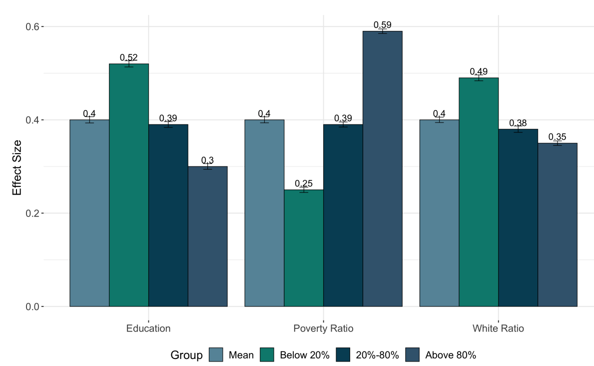


Figure 3

This figure shows the causal KNN estimated CATE with the test data set for (a) overall, (b) essential, (c) non-essential, (d) financial, and (e) other goods and services. The observations are grouped by the percentiles 1) below 20%, 2) 20% - 80% and 3) above 80% of three dimensions: education, poverty ratio, and white ratio.

(a) Full Sample Placebo Test

(b) Alternative Sample Placebo Test

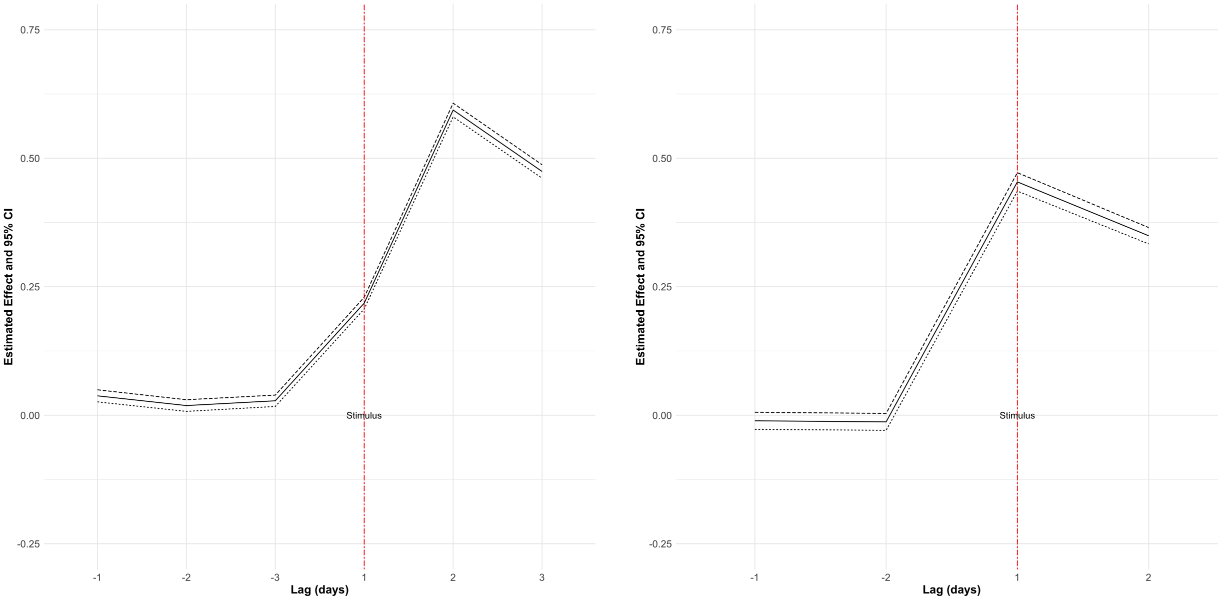


Figure 4

Panel (a) shows the coefficients of the benchmark two-way fixed effect regression with placebo timing on the full sample. Panel (b) shows the coefficients of the benchmark two-way fixed effect regression with placebo timing on the alternative sample.

Tables

Table 1: Overall Estimation

	(a)	(b)	(c)	(d)	(e)	(f)
Stimulus	0.427***	0.422***	0.424***	0.432***	0.426***	0.3957***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
White Ratio ×Stimulus		-0.021***			-0.016**	
		(0.005)			(0.007)	
Poverty Ratio ×Stimulus			0.036***		0.012	
			(0.006)		(0.009)	
College Education Ratio ×Stimulus				-0.041***	-0.036***	
				(0.008)	(0.009)	
Covariates	No	No	No	No	No	Yes
Time and County FE	Apply to all specifications					
Observations	151376 to all specifications					
Standard Error	Cluster by Zip code to all specifications					

Table 2: Essential Goods and Services

	(a)	(b)	(c)	(d)	(e)	(f)
Stimulus	0.481***	0.477***	0.476***	0.482***	0.478***	0.4466***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
White Ratio ×Stimulus		-0.019***			-0.005	
		(0.003)			(0.004)	
Poverty Ratio ×Stimulus			0.044***		0.024***	
			(0.003)		(0.005)	
College Education Ratio ×Stimulus				-0.054***	-0.040***	
				(0.004)	(0.005)	
Covariates	No	No	No	No	No	Yes
Time and County FE	Apply to all specifications					
Observations	437889 to all specifications					
Standard Error	Cluster by Zip code to all specifications					

Table 3: Non-essential Goods and Services

	(a)	(b)	(c)	(d)	(e)	(f)
Stimulus	0.183***	0.181***	0.181***	0.183***	0.180***	0.1748***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
White Ratio ×Stimulus		-0.011***			0.002	
		(0.002)			(0.003)	
Poverty Ratio ×Stimulus			0.018***		0.023***	
			(0.002)		(0.003)	
College Education Ratio ×Stimulus				-0.004	0.009***	
				(0.003)	(0.003)	
Covariates	No	No	No	No	No	Yes
Time and County FE	Apply to all specifications					
Observations	662280 to all specifications					
Standard Error	Cluster by Zip code to all specifications					

Table 4: Financial Goods and Services

	(a)	(b)	(c)	(d)	(e)	(f)
Stimulus	0.565***	0.549***	0.557***	0.569***	0.552***	0.5205***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
White Ratio ×Stimulus		-0.076***			-0.061***	
		(0.005)			(0.007)	
Poverty Ratio ×Stimulus			0.077***		0.024***	
			(0.006)		(0.008)	
College Education Ratio ×Stimulus				-0.060***	-0.042***	
				(0.007)	(0.008)	
Covariates	No	No	No	No	No	Yes
Time and County FE	Apply to all specifications					
Observations	248909 to all specifications					
Standard Error	Cluster by Zip code to all specifications					

Table 5: Other Goods and Services

	(a)	(b)	(c)	(d)	(e)	(f)
Stimulus	0.250***	0.251***	0.251***	0.252***	0.253***	0.2313***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
White Ratio ×Stimulus		0.001			-0.017***	
		(0.003)			(0.003)	
Poverty Ratio ×Stimulus			-0.005*		-0.036***	
			(0.003)		(0.004)	
College Education Ratio ×Stimulus				-0.031***	-0.050***	
				(0.003)	(0.004)	
Covariates	No	No	No	No	No	Yes
Time and County FE	Apply to all specifications					
Observations	497027 to all specifications					
Standard Error	Cluster by Zip code to all specifications					

Table 6: Overall Estimation on Alternative Sample

	(a)	(b)	(c)	(d)	(e)	(f)
Stimulus	0.403***	0.407***	0.387***	0.402***	0.399***	0.3931***
	(0.014)	(0.016)	(0.015)	(0.014)	(0.016)	(0.016)
White Ratio ×Stimulus		0.009			0.028***	
		(0.008)			(0.010)	
Poverty Ratio ×Stimulus			0.045***		0.048***	
			(0.010)		(0.016)	
College Education Ratio ×Stimulus				-0.049***	-0.025*	
				(0.011)	(0.014)	
Covariates	No	No	No	No	No	Yes
Time and County FE	Apply to all specifications					
Observations	83016 to all specifications					
Standard Error	Cluster by Zip code to all specifications					

Table 7: Overall Estimation of Placebo Regression

	(a)	(b)	(c)	(d)	(e)	(f)
Stimulus	0.074***	0.064***	0.066***	0.064***	0.054***	0.070***
	(0.015)	(0.015)	(0.014)	(0.014)	(0.016)	(0.014)
White Ratio ×Stimulus		-0.006			-0.018**	
		(0.005)			(0.008)	
Poverty Ratio ×Stimulus			-0.006		-0.019*	
			(0.006)		(0.011)	
College Education Ratio ×Stimulus				0.009	0.001	
				(0.008)	(0.011)	
Covariates	No	No	No	No	No	Yes
Time and County FE	Apply to all specifications					
Observations	16282 to all specifications					
Standard Error	Cluster by Zip code to all specifications					

Appendix

A.1 Spending Categories

We consider the following goods in this study. We classify them as categories including essential, non-essential, financial, and others.

- Taxicabs and Limousines: Non-essential
- Fax services: Others
- Money Orders –Wire Transfer: Financial
- Cable and other pay television (previously Cable Services): Others
- Electric, Gas, Sanitary and Water Utilities: Essential
- Discount Stores: Essential
- Variety Stores: Non-essential
- Grocery Stores: Essential
- Misc. Food Stores –Convenience Stores and Specialty Markets: Non-essential
- Automotive Parts, Accessories Stores: Others
- Service Stations (with or without ancillary services): Others
- Automated Fuel Dispensers: Others
- Eating places and Restaurants: Non-essential
- Fast Food Restaurants: Non-essential
- Drug Stores and Pharmacies: Essential
- Package Stores –Beer, Wine, and Liquor: Non-essential
- Book Stores: Others
- Miscellaneous and Specialty Retail Stores: Non-essential
- Financial Institutions –Manual Cash Disbursements: Financial
- Financial Institutions –Merchandise and Services: Financial
- Non-Financial Institutions –Foreign Currency, Money Orders (not wire transfer) and Travelers Cheques: Non-essential
- Insurance Sales, Underwriting, and Premiums: Essential
- Government Services (Not Elsewhere Classified): Essential
- Automobile and Truck Dealers (Used Only): Non-essential

A.2 Zip Codes Characteristics

These are the characteristics of Zip codes that we use to find k nearest neighbors for Zip codes when estimating the heterogeneous treatment effect. All statistics are normalized.

- *per_white* ratio of white in population of a Zip code area
- *per_black* ratio of black in population of a Zip code area
- *per_asian* ratio of Asian in population of a Zip code area
- *per_hispanic* ratio of Hispanic in population of a Zip code area
- *median_income* median income of a Zip code area
- *per_poverty* ratio of people with income lower than poverty limit of a Zip code area
- *per_college_above* ratio of people with degree of college or higher education of a Zip code area
- *median_age* median age of a Zip code area
- *commute_car* ratio of people commute with cars of a Zip code area
- *ADI* acceptable daily intake of a Zip code area
- *per_foreign_born* ratio of foreign born people in population of a Zip code area
- *commute_public_tran* ratio of people commute with public transportation of a Zip code area
- *median_gross_rent* median gross residential rent of a Zip code area