Online Appendices

for

Spending Responses to High-Frequency Shifts in Payment Timing: Evidence from the Earned Income Tax Credit

Aditya Aladangady, Shifrah Aron-Dine, David Cashin, Wendy Dunn, Laura Feiveson, Paul Lengermann, Katherine Richard, and Claudia Sahm

Appendix A: Additional Details on Datasets

Fiserv Spending Indexes

Central to our study is a set of novel daily, state-level indexes of spending at retail stores and restaurants, as developed by Aladangady et al. (2019). These indexes are available at a state, 3-digit NAICS level for 2010 forward and have been used to study the spending effects of several recent hurricanes (Aladangady et al., 2016, 2019), as well as state sales tax holidays (Aladangady et al., 2017). The data are key to our ability to identify spending responses to refunds while controlling for a very rich set of location-specific trends and fixed effects.

The indexes were constructed using anonymized merchant-level credit, debit, and other electronic transactions (including electronic benefit transfers, or EBT) from Fiserv, one of the largest payment processors in the United States.[1] Movements in the raw merchant-level transactions data partly reflect changes in Fiserv's client base or acquisition of other payment processing firms. In addition, the transactions data have no sample frame, and consequently do not match the distribution of merchants in the US. To address these issues, we follow the methodology laid out in Aladangady et al. (2019) to filter the data to merchants with stable relationships with Fiserv, reweight the data to match the 2012 Economic Census at a 3-digit NAICS and state level, and aggregate to NAICS-by-state level indexes.[2]

---

[1] Because our spending measures are based on merchant-level data, our estimates will capture spending responses occurring in the same state where the EITC household files taxes.

[2] Our data contract does not allow us to directly use merchant-level data, only anonymized aggregates. As such, our methodology queries and aggregates merchant-level data to provide a measure of retail spending that tracks official statistics at a national level but can be disaggregated to smaller geographies or NAICS categories.

In this paper, we focus on spending at retail stores and restaurants, which comprises about 30 percent of consumer spending and is well measured by card transactions within the Fiserv data. These data roughly overlap with the coverage of the Census Monthly Retail Trade Survey (MRTS) and comprise a large share of discretionary spending categories that are likely to respond to changes in cash flows. Notably, the measure excludes sales at motor vehicle dealers and certain services categories such as entertainment, healthcare, or transportation. In addition, because the location of the transaction is unknown for e-commerce sales, we exclude sales at non-store retailers and merchants identified as likely online retailers.

What we lack in household-level information, we make up for with the broader coverage of our source data in constructing spending measures. The coverage statistics, validation exercises, and applications presented in Aladangady et al. (2019) provide us with confidence that the spending indexes measure consumer spending well. Nationally, the coverage ratio – computed as total Fiserv sales used in the creation of the index divided by total estimated sales – has increased from 5.3 percent in 2010 to 9.5 percent in 2019, averaging about 7.4 percent over our baseline sample period from 2014 to 2018. Coverage is not uniform across states (Figure A.1), but always exceeds 3 percent of total spending.[3] Notably, spending growth in the Fiserv series compares favorably to official statistics, and in particular the Census MRTS. As illustrated in Figure A.2, 12-month growth rates from the spending index (at a monthly frequency) closely track the MRTS since 2012, with a correlation of 0.94.

For this paper, we utilize data from 2014 to 2018 to provide a symmetric window around the implementation of the PATH Act. Our baseline estimates include only data from weeks around tax season, which we define as January through June of each sample year, though our results are robust to including all months of data as well.

While daily, state-level measures provide statistical power necessary to estimate spending responses in our setting, we also provide results using Metropolitan Statistical Area-level (MSA) versions of these measures in Appendix E. The MSA spending measures are constructed in an identical manner to the state-level counterparts but are noisier. Furthermore, these measures are

---

[3] Differences in coverage across states and industries are adjusted because we benchmark all spending to the 2012 Economic Census. Coverage variation simply provides a measure of the relative sample size, and therefore the likely precision, of our estimates. Because sample sizes are substantial even in relatively low-coverage states and industries, these differences are unlikely to cause problems. Moreover, our estimates all account for potential measurement error correlations by state since standard errors are clustered at the state level.

subject to occasional suppression at the daily, 3-digit NAICS level, limiting our ability to use them to study the composition of spending responses across categories of expenditures.[4]

IRS Tax Refund Data

The tax refund data used in this study come from the Internal Revenue Service's Research, Applied Analytics, & Statistics group, and form the data underlying Statistics of Income aggregates provided publicly. The data provide a measure of the number of tax refunds that were issued by IRS on a given day in each state, as well as the total refund dollar amount. "Issuance" refers to the day on which the Treasury made a withdrawal from its operating cash balance in order to send out a refund. It does not imply that the refund appeared in an EITC-claiming household's bank account on that day. The vast majority of tax filers—83 percent in 2017—received refunds as direct deposits, and those refunds generally appear in a tax filer's bank account within a few days of refund issuance.[5] The time between issuance and receipt is slightly longer for those who receive their refunds in other forms, such as paper check.

Because the actual receipt of an EITC refund tends to lag its issuance by the Treasury by a few days and daily EITC refund issuance by the Treasury tends to be very lumpy (Appendix Figure A.3), we use a 7-day trailing moving average of daily refund dollars in our main regressions. While this lumpiness is likely exogenous, smoothing in this manner allows us to better isolate variation from the PATH Act, and also estimate a model in which responses to refunds over a full week are reflected on a given day.[6]

Refund data also split out tax filers who claimed either a refundable EITC or Additional Child Tax Credit (ACTC). Recipients of both credits were subject to the shift in the timing of refund receipt induced by the PATH Act. Because the overwhelming majority of this group (over 90 percent) are EITC recipients, we refer to these as "EITC refunds" in the paper for brevity.

---

[4] While our spending measures can be constructed at an MSA level (though with additional noise), our daily tax refund data are only available at a state level. We allocate daily state returns to MSAs using annual shares of EITC returns from the SOI. More granular geographic breakdowns may provide additional variation and power, but there is no clear sense in which results at a state level would be biased. As such, we prefer the state-level specifications as our baseline, though the MSA-level results are very similar.

[5] Direct deposit statistics retrieved from https://www.efile.com/efile-tax-return-direct-deposit-statistics/. E-file.com is an IRS authorized e-file provider.

[6] If spending typically occurs on a Saturday, refunds received at any point over the prior week are likely to have a similar response. As such, our estimator utilizes a distributed lag of 7-day averages of issuance to cover receipts over the past week, the week prior, etc. without considering about when in the week the refund was issued.

Appendix B: Framework for Analyzing Alternative Counterfactuals

Because we utilize state (or MSA) data, our unit of observation contains a mixture of households that are impacted differently and respond differently to the identifying variation we exploit— namely, the two-week shift in refund timing induced by the PATH Act. As such, our estimate recovers the average spend out from EITC refunds across various types of households. This section provides a conceptual framework to formalize the effect we recover and better compare it to other counterfactuals used in the literature.

The table below provides a taxonomy of relevant household types:[7]

	Non-Compliers – Late filers and non-EITC households	Compliers – early filers for whom EITC payments are shifted
Non-Smoother – Spending altered by payment timing	Nothing shifted directly, all variation captured by controls	**Spending *and* EITC refund shifted**
Smoother – Spending not altered by payment timing	Nothing shifted directly, all variation captured by controls	**EITC refund shifted, *but spending not*.**

Our fixed-effects estimator compares the shift in spending and refund issuance across the time period between "high" and "low" EITC states. Specifically, a "High-EITC state" is one with a lot of compliers (early-filing EITC recipients), and a "Low EITC State" is one with few compliers. The PATH Act shifts a larger amount of per-capita refunds in the "High EITC states" and differences between the "High" and "Low" states provide identification. Our regression can be thought of in terms of a simpler difference-in-differences framework comparing a "High EITC" state with a "Low EITC" state between years before and after the PATH Act. For simplicity of exposition, we lay out a conceptual framework in a two-period, two-state setting where our estimator is identical to the difference-in-difference version. In the High EITC state, a larger amount of tax refunds are shifted on a given day, $refund.shift^{hi} > refund.shift^{lo}$. Spending in the High EITC state is shifted by…

---

[7] Following the PATH Act, federal tax refunds—the dependent variable of interest—are shifted for early filers claiming EITC. In the terminology of the treatment effects literature, these households can be considered "compliers" and our estimate recovers the average treatment effect on spending among these compliers.

$$spend.shift^{hi} = \lambda^{hi} * MPC^{hi}_{nonsmoother} * local.mult^{hi} * refund.shift^{hi}$$

where λ^{hi} gives the fraction of compliers that are non-smoothers in the high state, $MPC^{hi}_{non-smoother}$ is the direct impact on spending for non-smoothers in the high state, and $local.mult^{hi}$ is the short-run local multiplier to arrive at the local aggregate effect. The spending in the Low EITC state is shifted in a similar fashion, providing a diff-in-diff corollary for θ, given by

$$\frac{\lambda^{hi}MPC^{hi}_{nonsmoother}local.mult^{hi} * refund.shift^{hi} - \lambda^{low}MPC^{low}_{nonsmoother}local.mult^{low} * refund.shift^{low}}{refund.shift^{hi} - refund.shift^{low}}$$

Under the assumption that local multipliers, shares of non-smoothers, and MPCs for non-smoothers are similar across states, the formula reduces to

$$\theta = \lambda * MPC_{nonsmoother} * local.mult.$$

This simple framework allows us to consider some important questions when comparing our results to those in the existing literature. Specifically, it provides us with a means to understand the various ways in which different counterfactuals and data generating processes influence the estimated spending response. Not surprisingly, shocks that raise the fraction of non-smoothers impacted by the shock (λ) or the spending response among non-smoothers ($MPC_{nonsmoother}$) would deliver larger estimates.

The variation in our paper primarily exploits a roughly two-week shift in a very large, infrequent, but regular lump-sum payment. If we instead consider a setting with a longer delay, we may expect a higher estimated response.[8] Because, at the limit, an infinitely long delay would be comparable to removing refunds altogether, our estimate may be a lower bound for the overall effect of refunds on EITC recipient households. Alternatively, if we consider taking the same two-week shift in EITC refunds, but estimating the spending response during a period of low liquidity or credit supply, we may expect a larger estimate than in our data, which largely covers

---

[8] In particular, a longer delay would likely lead to a higher λ, as fewer households are able to weather an extended delay in liquidity. Of course, a longer delay may also lead households to tap into alternative forms of credit or insurance mechanisms to smooth through the shock, thereby lowering the overall response.

a late-stage expansion period. Results from Gross et al (2020) suggest this may be the case in practice.

The framework above highlights nuanced differences between existing results in the literature. Our paper is the first to analyze the impact of a short-lived shift in a large, infrequent lump-sum transfer to low-income households. In closely related work, Parker et al (2013) use survey data to exploit variation in $300-600 stimulus payments over a three-month period. Compared to theirs, our paper considers a much larger lump-sum transfer, averaging $4,250 per household, which is shifted by a shorter time period. It also utilizes newer administrative data that may provide more precise estimates of spending responses. More recently, Baker and Yannelis (2017) and Gelman et al (2019) use data from a financial planning app to study the impact of delaying regular paychecks by exploiting variation induced by the 2013 government shutdown. Our study expands on these, providing an estimate of how shifts in large, infrequent lump-sums impact spending.

How does this framework inform excess sensitivity to refund receipt?

In this setting, we used the term "smoother" to refer to a household that smooths through a shift in refund timing and "non-smoother" to refer to a household whose spending responds to the shift. Naturally, spending is unlikely to respond to a shift in refund timing if a household does not exhibit excess sensitivity to refund timing (ie, a spike in spending upon refund receipt). However, the opposite is possible: a household may exhibit excess sensitivity to refunds generally, but use short-term credit to avoid shifting the timing of spending when a small transitory shock to refund timing occurs. This type of household drives a wedge between the estimate we recover and a measure of how sensitive spending is to refund receipt. In our setting, such a household is unlikely to exist because the PATH Act shifted refund timing permanently. It seems unlikely that households responding sharply to refund receipt persistently used short-term credit to pull forward spending by two weeks every year following the implementation of the PATH Act. Indeed, our results suggest the shift in spending was persistent. As such, it seems reasonable to assume all households exhibiting a spike in spending also shift that spending when refunds are shifted, and we interpret our result as equivalent to a measure of excess sensitivity.

6

How does this framework inform how households would respond to the introduction of a new lump-sum payment, such as fiscal stimulus?

It is also useful to consider a counterfactual that varies the *magnitude* of refunds rather than their *timing*. Such a counterfactual can answer the question of how households respond to additional lump-sum transfers relative to a world without these additional transfers. In such a setting, permanent income is altered, and even households that would smooth through a shock to timing are likely to respond. As such, we should expect a larger overall response than we would in the counterfactual where only timing is altered, and our estimate thus provides a lower bound on the spending response to a change in EITC refund magnitude.

However, households that smooth through timing shocks are unlikely to exhibit excess sensitivity to refund timing, and probably smooth the permanent income increase over a long period. Therefore, the predominant source of variation in spending in the short-term is among non-smoothers, and the per-dollar spending response to the timing shock may provide a reasonable proxy for the overall effect of refunds in the short-run. In other words, the lower bound is likely quite tight when the response of smoothers is small.

Appendix C: Residual Variation in Refunds and PATH Act Instrumental Variables Estimates

As discussed in Section 2, beginning in 2017 the PATH Act shifted the timing of refunds for early-filer EITC claimants by about 2 weeks, generating plausibly exogenous variation in the timing and magnitude of state-level EITC refund issuance. We use this type of variation in the identification of the OLS fixed effects estimates because a high-EITC state will have later-than-normal disbursements in 2017 and 2018 relative to prior years, whereas disbursements in a low-EITC state will be relatively little affected by the policy change. However, the baseline results also include other variation in refund disbursement timing and magnitudes beyond that induced by the PATH Act. As we will show, the residual variation in refunds, after conditioning on our baseline controls, is largely driven by shifts in timing for early-filing EITC households, particularly in 2017 following the PATH Act. In addition, a policy instrumental variable (IV) isolating variation in EITC refunds from the PATH Act yields very similar results to our baseline

exercise, suggesting much of the residual variation in our baseline estimate is driven by idiosyncrasies of IRS's processing of returns and Treasury's disbursement timing, and is largely exogenous to consumption.

Residual refund variation in baseline fixed-effects model

Before turning to our IV approach isolating refund timing variation from the PATH Act explicitly, we first consider what drives the variation in our baseline fixed-effects model. To do so, we can look at the variation in EITC refunds conditional on our baseline controls (Frisch and Waugh, 1933), as shown in Figure A.6.

Because our fixed-effects approach compares each state's per-capita refunds to the average across states, it is natural to expect that refunds are higher-than-average in 2015 and 2016 in the "High EITC" states with larger fractions of early filing EITC claimants, and lower in "Low EITC" states. In addition, because we include state-by-week-of-year fixed effects, each year's refunds are compared to the average over the sample. As a result, refunds are lower-than-normal in 2017 and 2018 in High EITC states after the PATH Act in early-to-mid February, and then spike up in late February and early March when refunds are paid out. Naturally, the opposite occurs in Low EITC states. Finally, because our regression includes state-by-year fixed effects, the residual in each year and in each state cumulate to zero. Importantly, residuals are nearly zero following early-filing periods (with a handful of EITC claimants having more significant delays due to the PATH Act). As such, the variation being exploited is short-lived, transitory shifts in refund timing and not shifts in refund magnitudes from year to year.

While our approach is not explicitly a difference-in-difference approach, it does compare spending before and after the PATH Act in high and low EITC states differing in exposure to the PATH Act, similar to a difference in difference. Taking the difference in difference of the refund residuals above yields Figure A.7, which clearly shows the transitory shock to refund primarily driving the variation in our data.

This variation between states and over time is what identifies our effect on spending and appears largely driven by the PATH Act itself. Notably, we see very small variation in refunds

across states and years outside of early-filing time, and year-to-year variation is *largely* driven by the pre- vs post-PATH Act variation. The IV approach we describe next explicitly isolates this variation.

PATH Act instrumental variables estimate

To isolate variation in refund timing and magnitudes driven by the PATH Act, we construct the following instrument that interacts the fraction of EITC recipients in a state with a dummy that takes a value of 1 after the implementation of the PATH Act in 2017. The variable is further interacted with week-of-year dummies for the weeks in which the PATH Act affected disbursement of EITC refunds, and is given by the following formula:[9]

$$Z_{st} = EITC.Fraction_s * \mathbf{1}(t \in Post.Path.Act) * \mathbf{1}(week.of.year = w)$$

The strength of this instrument in predicting refunds is evident from Figures 1 and A.4. In the years following the PATH Act's implementation (2017 and 2018), EITC refunds were lower than the pre-PATH Act years in the first couple weeks of February and then higher for roughly the three subsequent weeks. The level of EITC refund issuance returned roughly to the pre-PATH Act level by mid-March. This sharp, short-lived shift in EITC refund issuance timing is captured by the second two terms of the instrument above. Furthermore, as shown in Figure A.4 there was little to no change in refund issuance timing for non-EITC households pre- and post-PATH Act; as such, we know that refund timing in states with higher shares of EITC recipients will be affected to a greater extent than states with lower shares specifically because of the PATH Act. We capture this PATH Act-induced cross-state variation in refund timing and magnitudes by interacting the post-PATH Act and week-of-year dummies with the state's EITC share. Indeed, first-stage regressions show the IV is quite strong, and the joint F-test on excluded instruments is well above the critical values.

In addition to predicting EITC refunds, for this instrument to identify the causal effect of refunds on spending, the following exclusion restriction must hold:

---

[9] The EITC share is set to the ratio of filers claiming EITC in the state to the total number of filers in 2015. This share is very stable over time, and is highly correlated with per-capita EITC refund disbursements in a state. The week of year dummy includes weeks 4-12 of the year, starting in late Jan (Jan 22-27) through late March (March 19-24). These dates roughly correspond to the period over which refund issuance timing for early filers was impacted by the PATH Act, as shown in Figures 1 and A.4.

$$E[\ EITC.Fraction_s * \mathbf{1}(t \in Post.Path.Act) * \mathbf{1}(week.of.year = w) * u_{st} | X_{st}] = 0$$

where u_{st} is the error term in Equation (1). In words, we assume that spending in states with higher EITC shares does not respond to the PATH Act except through the impact the legislation had on refund timing.

As in our baseline estimates, we continue to control for state-specific year and week-of-year effects, and winter weather controls, along with daily time dummies. These controls address concerns that aggregate shocks and state-specific trends or seasonality in spending are spuriously correlated with refund timing and exposure to the PATH Act. Any threat to identification would have to be a shock that occurs concurrently with the PATH Act implementation and is correlated with the EITC share in a state.

Of course, one such concern is that the PATH Act itself caused changes in tax refund magnitudes for households that claim the EITC. Aside from imposing limits on when the IRS would begin issuing EITC refunds, the PATH Act extended certain deductions that were set to expire. For example, the earned income threshold for the Child Tax Credit was maintained at $3,000, whereas it had been previously set to rise to $10,000. In addition, the Act extended limits on tuition deductions that had been set to fall under previous law. Both of these changes may have raised the expected magnitude of refunds for certain households, such that the PATH Act not only shifted the timing of refunds but also their expected size. Our inclusion of state-year fixed effects addresses this concern. In particular, states with differing shares of CTC claimants or those claiming tuition deductions are allowed to have different spending levels in 2017 and 2018 after the legislation is enacted. As such, spending responses reflecting these changes are absorbed by the fixed effects.

The IV approach laid out here is more explicitly a difference-in-differences approach in that it compares states before and after the PATH Act and across EITC shares, which provide a proxy for exposure. In doing so, it isolates only the timing shift due to EITC refunds being held until February 15 due to the PATH Act. Moreover, the IV removes state-time variation that isn't explicitly correlated with the PATH Act and EITC shares. For example, refund timing differences between states driven by idiosyncrasies of IRS issuance are removed, reducing the overall variation in the data, and therefore power.

The results from the IV exercise, shown in column 6 of Appendix Table A.2, appear quite similar to our baseline results, though standard errors are a wider as one may expect with this IV approach. Notably, the cumulative spending per EITC refund dollar over the several weeks immediately around refund receipt is 0.27, very close to our baseline OLS fixed effects estimate. The IV estimate suggests a slightly sharper response when refunds are received, but the overall timing is not significantly different from the baseline results.

What do the similarities between the IV and OLS results tell us? As mentioned previously, variation in refund issuance driven by the PATH Act is a large part of the residual variation in our baseline OLS estimate after controlling for state-specific trends, seasonality, weather, and national shocks. However, the OLS estimate also includes other variation in refund timing between states and over time that isn't correlated with state-specific trends, seasonality, weather, or national shocks. The similarity of the IV result with the baseline suggests that this additional (non-PATH Act induced) variation is either small relative to the variation generated by the PATH Act or is exogenous with respect to consumption. Indeed, discussions with IRS suggest that much of this variation is the result of idiosyncrasies in processing and disbursing refunds at the IRS and Treasury.

Appendix D: Treatment Randomization Placebo Tests

Our baseline specification relies on a key assumption that, conditional on our controls, spending in high- and low-EITC states follows parallel trends. As such, any variation in spending we observe that is correlated with EITC refunds can be interpreted as a causal effect of refunds on state spending. Under our assumption of parallel trends, it should also be the case that the effect of refunds on spending at points in time outside of tax season is zero, since in reality no refunds were issued at that time. Systematically significant results may suggest that high EITC states have spending growth that differs from low EITC states more generally, and correlations with EITC refund issuance are spurious. In addition, scrambling EITC refunds across states in our data should shut down meaningful cross-sectional variation and lead to a null result. Systematically significant results would suggest spending and EITC refunds are simply coincident, but may not truly be related.

These arguments motivate two placebo exercises.[10] First, we re-estimate our state-level baseline specification after shifting the timing of EITC refunds in each state to be outside of tax season, repeating the exercise for all possible offsets that ensure the distributed lag of EITC refunds falls outside of tax season. Second, we re-estimate the model using a random reassignment of EITC refunds across states, repeating the exercise for 1,000 estimates. These two distributions of spending estimates provide both a check on the validity of our specification assumptions, as well as a robust, small-sample distribution of our estimator under two alternative randomization assumptions. We use these distributions to construct a Fisher exact test for each randomization (Fisher, 1922).The histograms in Figure A.8 display the distribution of estimates for the total spending response from each of our two placebo exercises along with our baseline estimate in red. The left panel shows results from the exercise scrambling dates are firmly centered around zero, with our baseline estimate clearly in the tail. In fact, only 2.4 percent of placebo estimates fall above our baseline, suggesting our baseline estimate remains significantly different from zero.[11] These results, along with the fact that our estimated spending responses fall to zero outside the window around refund disbursement, suggest our controls sufficiently establish parallel trends, and that observed co-movement between spending and refunds is likely driven by a causal relationship.[12]

The state randomization (right panel) also shows a similar story, with estimates clearly centered around zero, and our baseline estimate falling at the upper tail. In this case, only .1 percent of placebo draws (only one estimate in our 1,000 replicates) fall above our baseline estimate, suggesting our estimate is not driven by an idiosyncratic allocation of EITC refunds across states.

---

[10] Because conditional on our controls, variation in tax refunds occurs primarily around the PATH Act delay, the key assumption is that we have established parallel trends *at this point in time*. This allows us to say that cross-state differences in spending are driven by refunds and not differential trends. While we cannot check this assumption directly, the randomized date placebo test provides evidence that parallel trends exist at various *other* points in the year, suggesting our flexible specification establishes similarly parallel trends at the moment of interest.

[11] In the main body of the paper, our baseline estimates show standard errors clustered at the state level. Using these standard errors, we see only modest over-rejection in the placebos, with 1.6 percent significant at $\alpha = .01$ in the date randomization placebo and 1.4 percent in the state randomization placebo. The Fisher exact test, of course, corrects for this over-rejection by using the distribution of placebos themselves to construct the test thresholds. As such, our baseline remains significant in both randomized inference tests, and confidence bands presented in the paper are only modestly changed if constructed using either of the placebo distributions.

[12] The result also suggests our result is robust to potential serial correlation (Bertrand, et al, 2004), though our baseline results do cluster at the state level, allowing arbitrary correlations over time. Similarly, the state randomization shows our result is robust to cross-state covariances.

Appendix E: Alternative Specification Using MSA-Level Variation

In order to estimate the spending response to EITC refunds using MSA-level variation, we construct daily spending indexes at the MSA level following the same methodology of Aladangady et al. (2019). We use Economic Census data from 2012 at the county level to reweight 3-digit NAICS codes and aggregate up to the MSA level. Otherwise, filtering and benchmarking methods are the same as that used for state data.

Because we only have data on daily, state-level EITC refund magnitudes, we must impute daily refunds at the MSA level using daily, state refund magnitudes and annual data at a more disaggregated level. Specifically, we allocate daily, state refunds to MSAs using annual IRS Statistics of Income data on each MSA's share of EITC recipients within a state. Consequently, daily EITC refund magnitudes attributed to each MSA are an estimate rather than the true value.[13] We then repeat our baseline specification shown in Equation (1), replacing all state-level fixed effects with MSA-level fixed effects to allow for MSA-specific levels, trends, and seasonality in spending. The results of this exercise are displayed in column 7 of Appendix Table A.2. The cumulative spending response per EITC refund dollar within a few weeks of refund receipt is 0.25, very similar to our baseline estimate of 0.27.[14]

Relative to state-level data, MSA data also provide additional variation in the credit profiles of households that could offer insight into why their spending exhibits sensitivity to EITC refunds. A likely reason is that households are liquidity or credit constrained, such that they can only spend when funds arrive. Figure A.9 provides additional suggestive evidence that this is the case. In each panel, we estimate the spending response for subsets of MSAs that vary in their levels of credit utilization rates and shares of subprime borrowers in the FRBNY/Equifax CCP data. While we cannot observe credit constraints directly, these measures are likely correlated with credit constraints.

---

[13] Notably, our imputation assumes IRS issuance timing across MSAs is identical within state, such that time-series variation is largely identical to the state-level estimates. Variation in refunds across MSAs is only driven by EITC shares. As such, we prefer our state-level measures, which offer largely the same variation in refunds and likely more precise measures of spending.

[14] The MSA and state-level measures are not directly comparable since the concept of "local aggregate effects" differs between the two. In particular, local spillovers within a state may be much larger than within an MSA or smaller geographic unit, attenuating the response as we disaggregate spatially. Moreover, spending outside one's own geographic area may become more likely, further attenuating responses in our merchant-based data. As such, it is reasonable to expect a slightly smaller response in the MSA data than in the state data.

Indeed, MSAs in the bottom third of (population-weighted) credit utilization rates and subprime shares do not exhibit a cumulative spending response that is statistically different from zero (light blue bars), while those in the top-third exhibit a sizeable and significant spending response (dark blue bars). The results suggest that constrained households may exhibit somewhat higher spend outs than unconstrained ones. However, for each credit outcome the difference in the cumulative spending response between the top third and bottom third MSAs is not statistically different from zero. This result could reflect a couple of things. First, it may be that most households claiming the EITC are fairly constrained, such that there is little variation across credit measures. Second, MSA data may mask heterogeneity within MSAs. In fact, most cities are quite varied in terms of income and credit use such that much of the heterogeneity in credit constraints may occur across parts of a city. Unfortunately, our data do not allow us to easily explore more granular geographies easily, and we leave this question for further research.

Appendix Figures and Tables

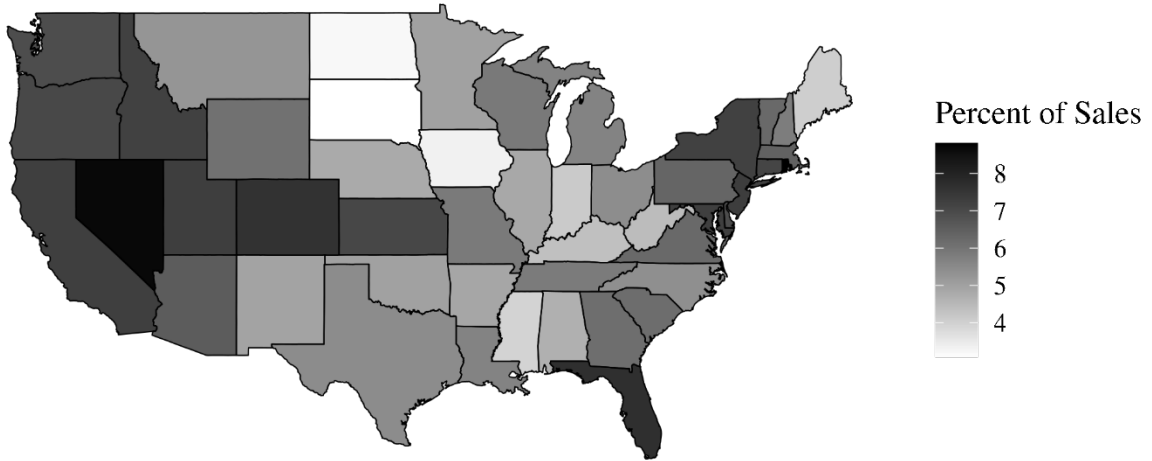**Figure A.1. Fiserv coverage of Economic Census Retail Sales Group by state, 2018**

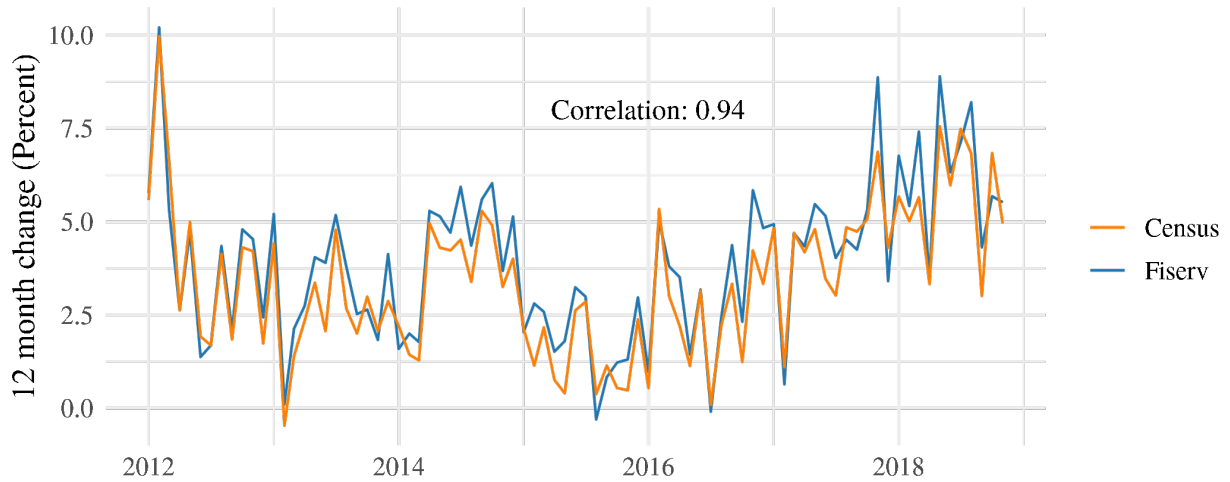**Figure A.2. National Retail Sales, 12-month percent change**

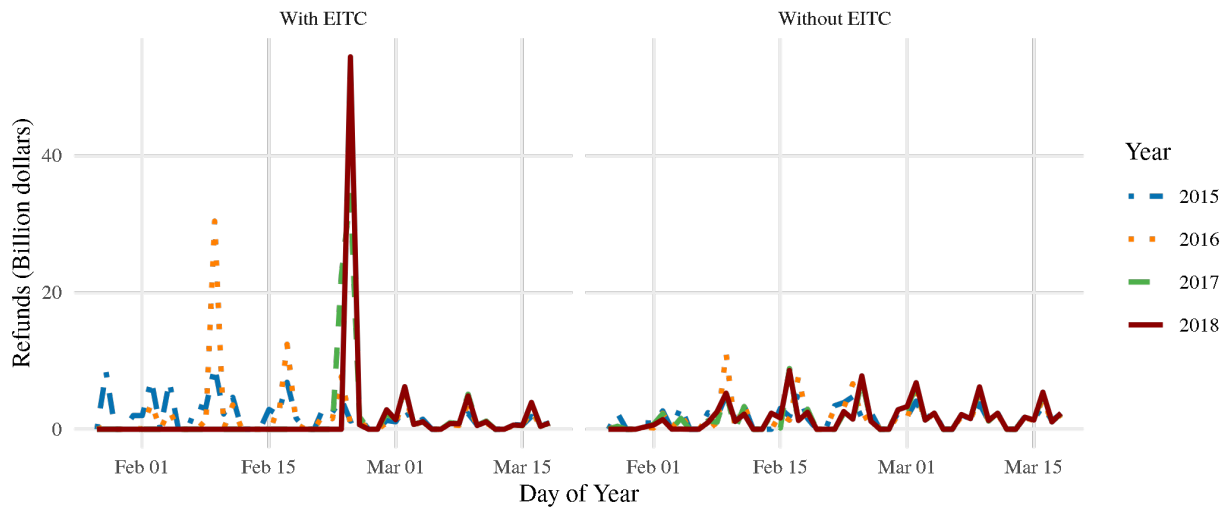**Figure A.3. Daily issuance of federal tax refunds with and without EITC**

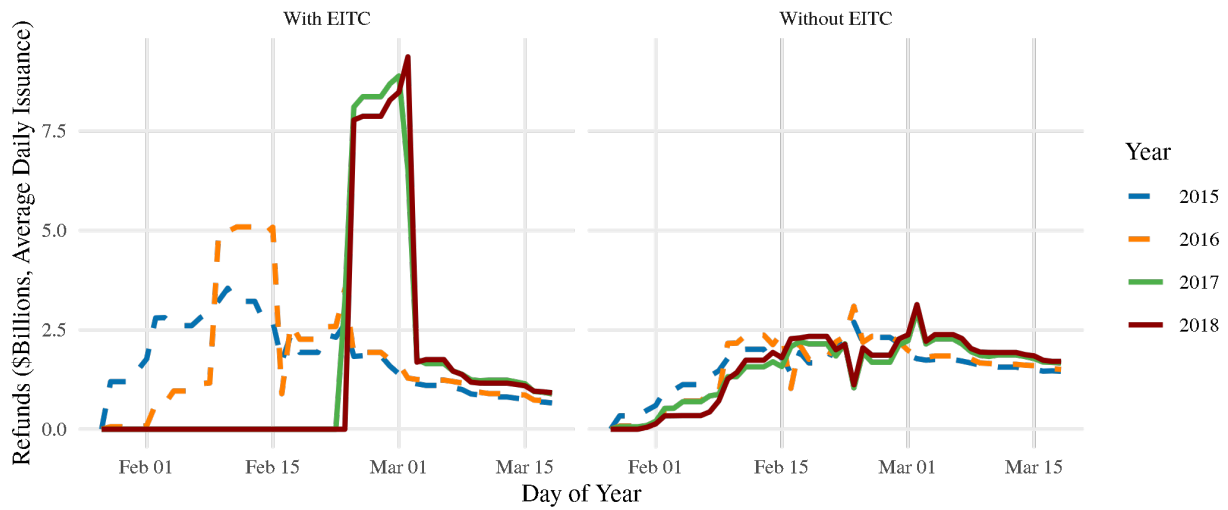**Figure A.4. 7-day trailing moving average of federal tax refund issuance**
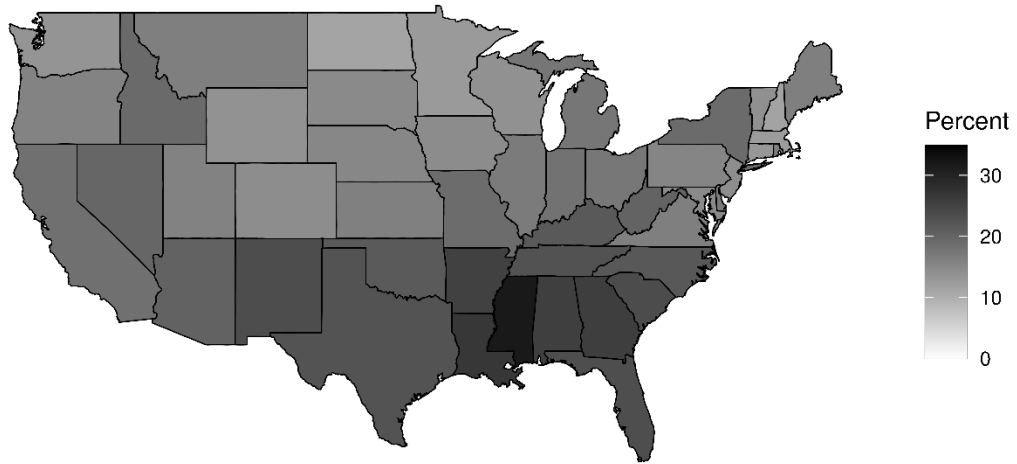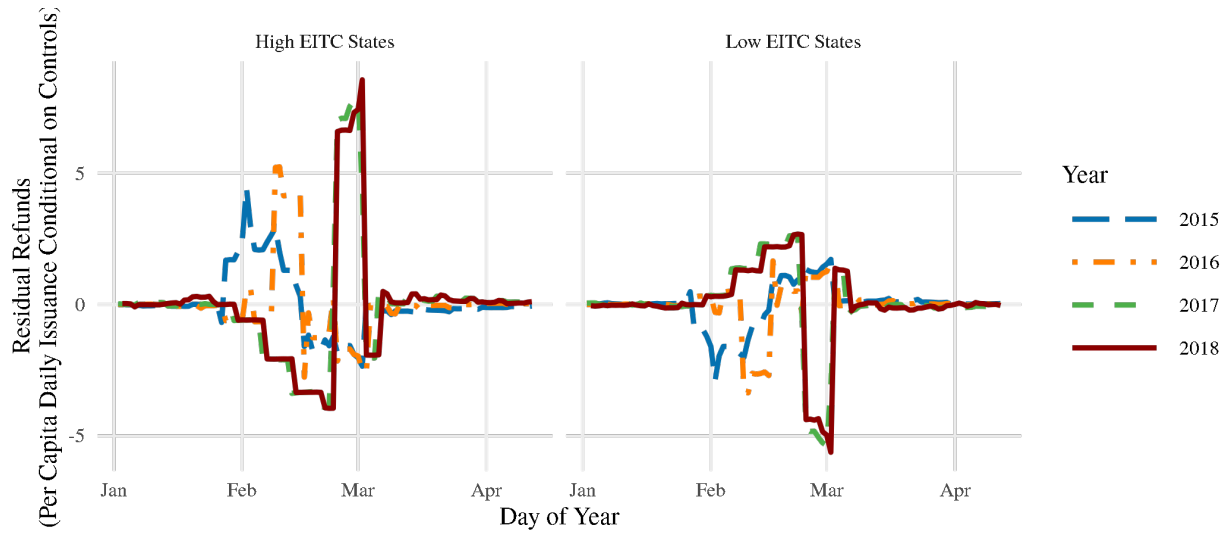
Figure A.5. Fraction of federal returns receiving EITC by state, 2016

Source: Internal Revenue Service

Figure A.6: Residual variation in refunds conditional on baseline controls

Figure A.7: Difference-in-Difference of Residual Refunds
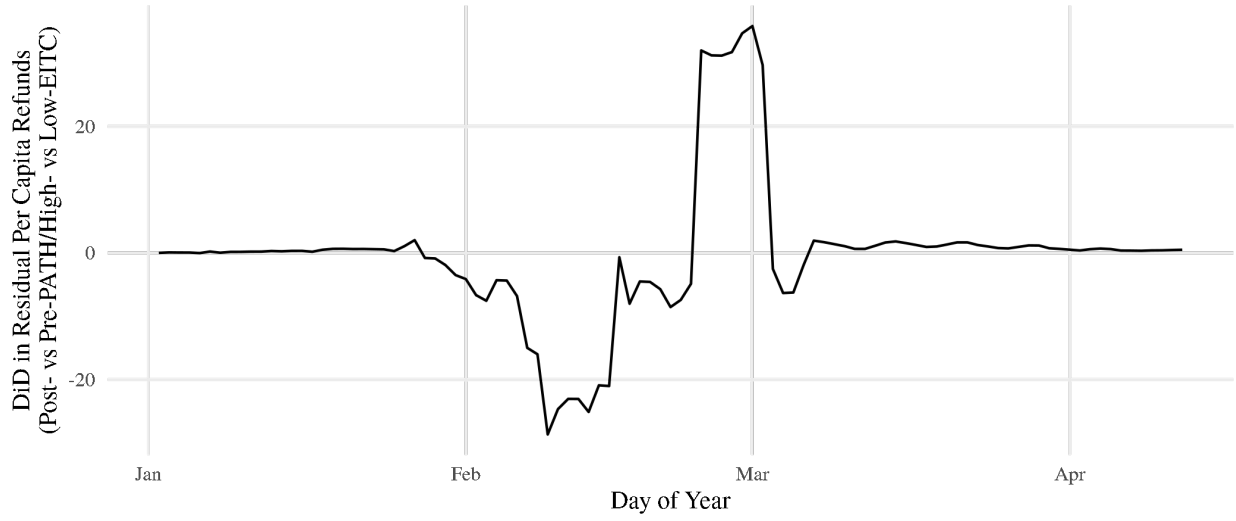## High vs Low EITC states, pre- vs post-PATH Act

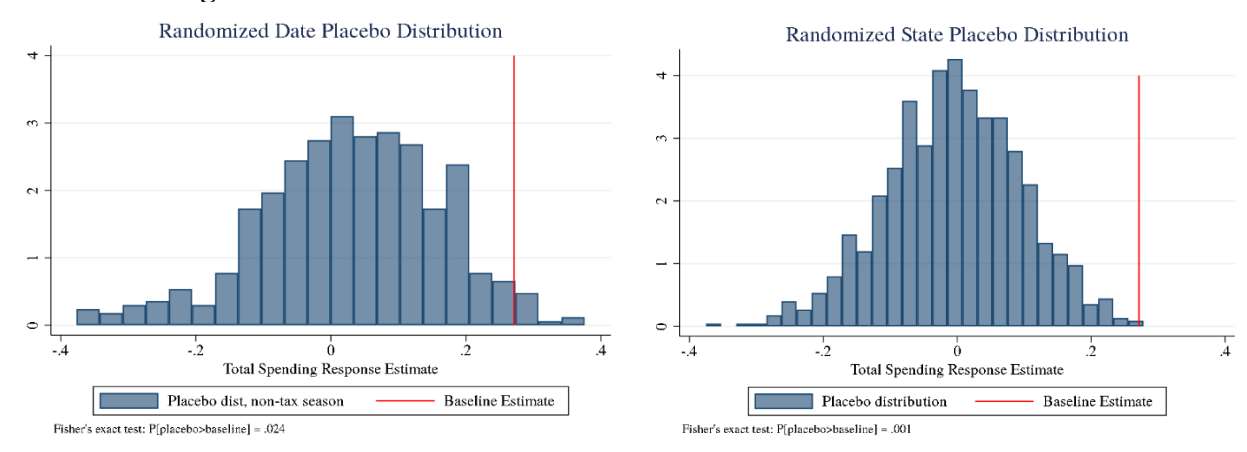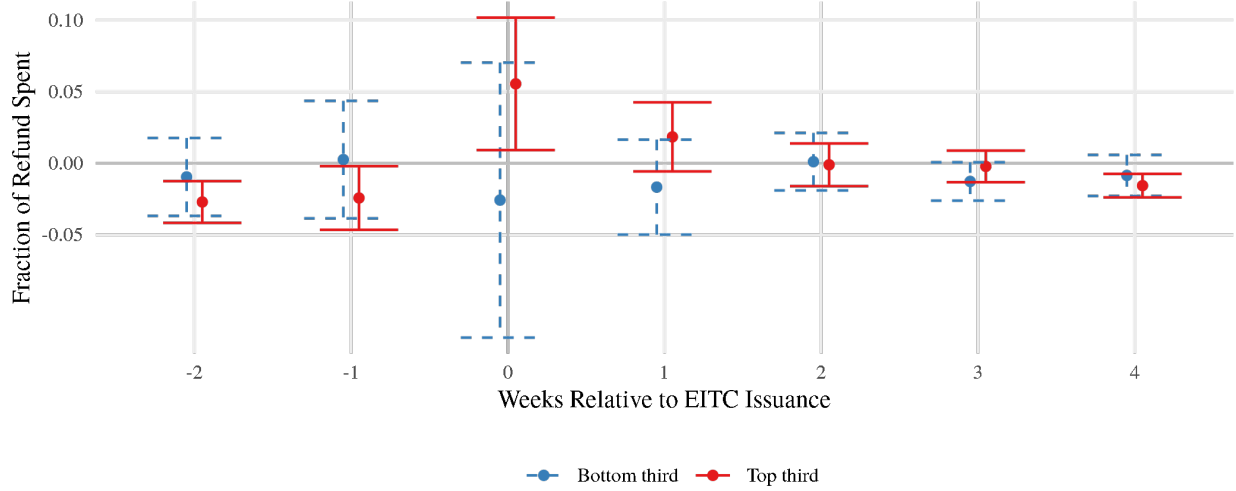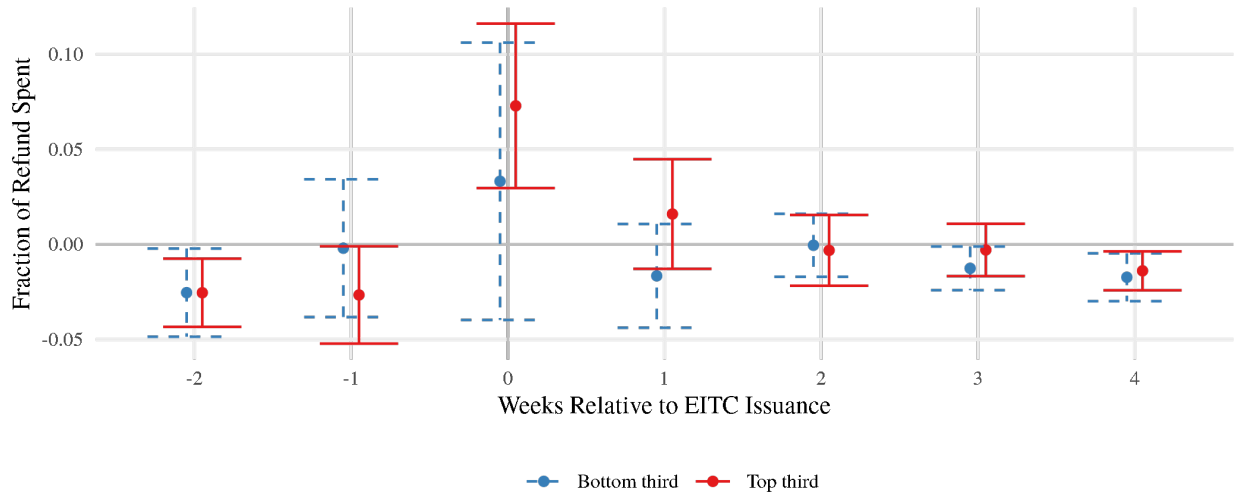## Figure A.8: Randomized Placebo Estimates and Fisher Exact Tests

Figure A.9: Spending heterogeneity across MSA-level credit outcomes

Average credit utilization rate in MSA

Table A.1. Spending per EITC refund dollar
at selected establishment types by week relative to refund issuance

	Sales / Population				
	Furniture	Electronics	Grocery	Restaurants	General Merchandise
	(1)	(2)	(3)	(4)	(5)
EITC Shock Lead 2 Weeks	0.001	0.001	0.0002	0.004	0.006
	(0.001)	(0.001)	(0.004)	(0.002)	(0.006)
EITC Shock Lead 1 Week	-0.001	0.0002	0.002	-0.001	0.020**
	(0.001)	(0.002)	(0.006)	(0.003)	(0.008)
EITC Shock	0.008***	0.012***	0.005	0.011***	0.039***
	(0.002)	(0.003)	(0.004)	(0.002)	(0.009)
EITC Shock Lag 1 Week	0.004***	0.005***	0.006*	0.010***	0.019**
	(0.001)	(0.002)	(0.003)	(0.003)	(0.009)
EITC Shock Lag 2 Weeks	0.002*	0.003**	0.012**	0.006**	0.006
	(0.001)	(0.001)	(0.005)	(0.002)	(0.011)
EITC Shock Lag 3 Weeks	0.001	0.002*	0.008**	0.004	-0.006
	(0.001)	(0.001)	(0.004)	(0.002)	(0.011)
EITC Shock Lag 4 Weeks	-0.001	-0.001	0.005	-0.006***	-0.003
	(0.001)	(0.001)	(0.005)	(0.001)	(0.011)
Total MPC	0.013***	0.022***	0.039*	0.026***	0.082*
	(0.005)	(0.008)	(0.020)	(0.009)	(0.046)

Note: Table A.1 displays point estimates and standard errors for the θ_ℓ coefficients in Equation (1) when the dependent variable is daily, state-level per capita spending for establishment-type (3-digit NAICS) subsets of the retail sales group category. We interpret the coefficients as yielding spending per EITC refund dollar at a particular establishment type ℓ weeks from refund issuance. All specifications cluster standard errors by state. *p<.1, **p<.05, ***p<0.01

Table A.2. Alternate Specifications

	Baseline	No Snow	Add. Leads/Lags	Excl. 2017	Add 2019	PATH IV	MSA Data
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
EITC Shock Lead 4 Weeks			0.003				
			(0.015)				
EITC Shock Lead 3 Week			0.001				
			(0.020)				
EITC Shock Lead 2 Weeks	0.011	0.018	0.012	-0.013	0.010	0.065	-0.010
	(0.017)	(0.023)	(0.021)	(0.016)	(0.016)	(0.044)	(0.009)
EITC Shock Lead 1 Week	0.065***	0.085***	0.067***	0.076***	0.060***	0.024	0.003
	(0.017)	(0.024)	(0.020)	(0.022)	(0.018)	(0.034)	(0.009)
EITC Shock	0.096***	0.101***	0.098***	0.116***	0.088***	0.159**	0.131***
	(0.015)	(0.020)	(0.021)	(0.020)	(0.021)	(0.060)	(0.015)
EITC Shock Lag 1 Week	0.095***	0.096***	0.098***	0.104***	0.094***	0.056	0.087***
	(0.014)	(0.015)	(0.020)	(0.021)	(0.015)	(0.053)	(0.011)
EITC Shock Lag 2 Weeks	0.035*	0.007	0.038	0.019	0.038*	0.031	0.018**
	(0.018)	(0.018)	(0.024)	(0.027)	(0.020)	(0.083)	(0.009)
EITC Shock Lag 3 Week	-0.011	-0.034**	-0.007	-0.009	-0.004	0.030	0.023*
	(0.015)	(0.016)	(0.018)	(0.019)	(0.016)	(0.045)	(0.012)
EITC Shock Lag 4 Weeks	-0.019	-0.037*	-0.018	-0.040*	-0.016	-0.094	-0.001
	(0.024)	(0.019)	(0.026)	(0.022)	(0.026)	(0.078)	(0.008)
EITC Shock Lag 5 Week			0.022				
			(0.027)				
EITC Shock Lag 6 Weeks			-0.016				
			(0.018)				
Total MPC	0.271***	0.238***	0.297*	0.253**	0.269***	0.271	0.252***
	(0.070)	(0.067)	(0.163)	(0.107)	(0.083)	(0.224)	(0.028)
Weather Controls	Yes	No	Yes	Yes	Yes	Yes	Yes
Storm States Excluded	No	Yes	No	No	No	No	No
Lags	4	4	6	4	4	4	4
Leads	2	2	4	2	2	2	2

Note: Column 1 replicates baseline result from Figure 3. Column 2 excludes several northern low-EITC-share states (ME, VT, NH, MA, RI, CT, NY, NJ, DE) which experienced a winter storm during tax season in 2016. Column 3 includes additional leads and lags to ensure parallel trends outside the immediate window of refund receipt. Columns 4 and 5 include alternative choices of years. Note that 2019 is excluded from the baseline specification because the Tax Cuts and Jobs Act (TCJA) likely influenced tax refunds across the income distribution that year, though it does provide additional variation in timing from the 2019 government shutdown. Columns 6 uses an IV strategy to isolate variation from the PATH Act as described in Appendix C, and column 7 utilizes MSA-level data as described in Appendix E. All specifications cluster standard errors at state level. *p<.1, **p<.05, ***p<0.01