Evidence on WIC using administrative data linkage

Charlotte E. Ambrozek

Evidence-Based Policymaking For Applied Economists

Agricultural and Applied Economics Association Annual Meeting
Virtual Post-conference Workshop

October 1st 2021

Research Question

How does a change in the set of WIC vendors available to a participant affect participation and use of benefits?

Context: Econ, public health, and nutrition literatures find positive effects of WIC benefit receipt on participants.

Access to vendors is a necessary condition for obtaining supplemental foods and their health/nutrition benefits. Prior evidence finds mixed effects on participant responses to WIC access.

Prior evidence is limited by an inability to observe which participants are directly affected by changes to WIC access. ▸ Literature

I deterministically link WIC participants to WIC vendors to resolve this limitation.

Data Linkage Framework

Participants' links to vendors come from redemptions.

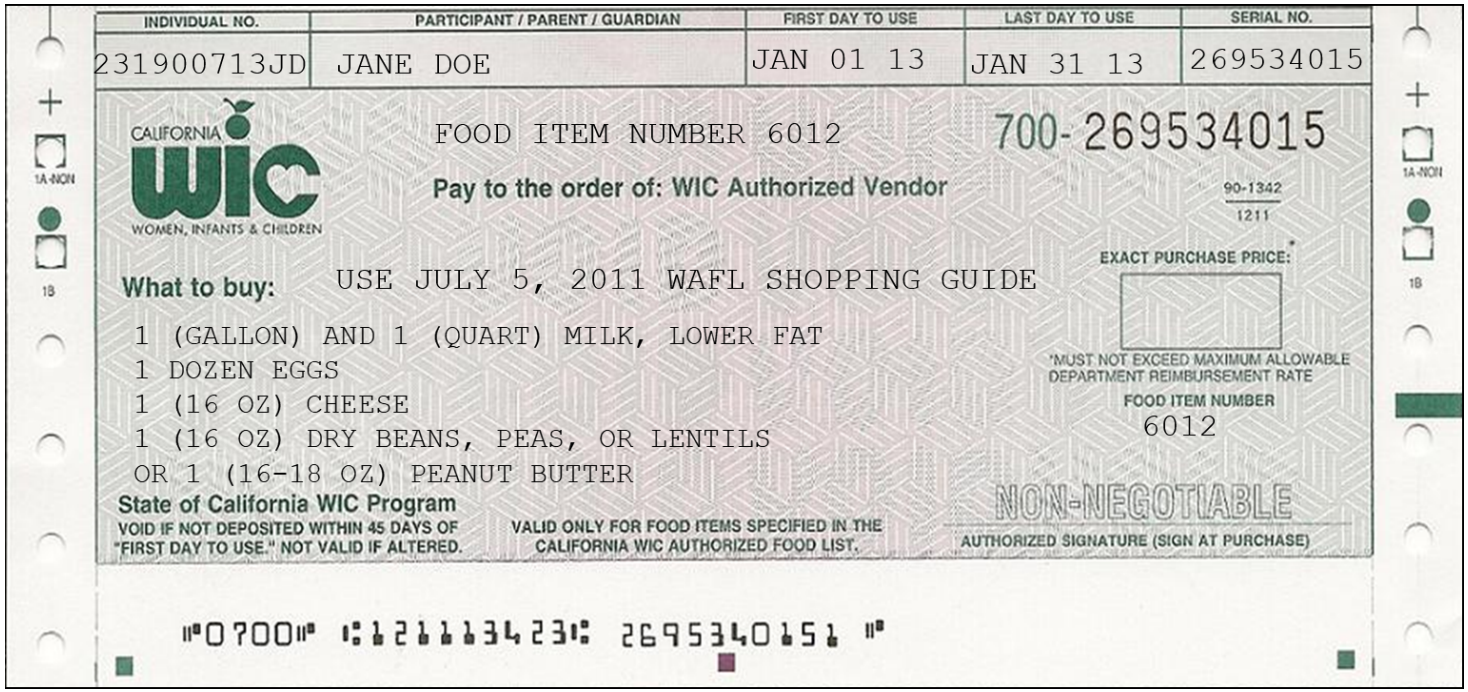


Figure: Sample WIC check

California WIC administrative data assigns participant, vendor, and food instrument specific IDs.

Treated participants redeem benefits at vendors that go on to be disqualified.

Data

I use California Department of Public Health (CDPH) WIC Division's administrative data from October 2015 to June 2019.

I use six internal data sources: participant certification and demographics, vendors, food instrument redemptions, food instrument issuance, food prices, and vendor sanctions.

Advantages:

- Completeness
- Internal consistency
- Policy relevance

Disadvantages:

- Computing constraints
- Limits on external data

Vendor, participant, and sanction data

These data sources provide time invariant or less frequently varying information on individuals and firms.

Disqualifications (DQs) are particularly important, because they remove the vendor from the participant's choice set.

- More than 300 DQs in the proposed sample period
- Type of sanction indicates duration and a check on anticipation and parallel trends assumptions
- I create a DQ specific identifier (DQ date, vendor ID, sanction type combinations).

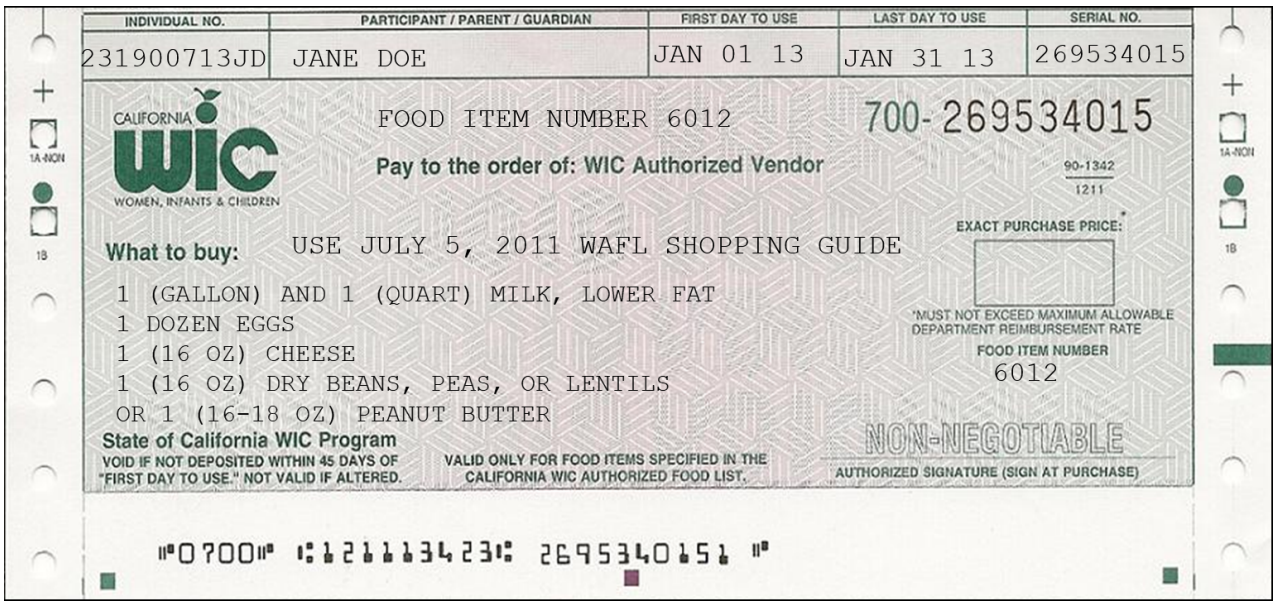
Participant data include certification, location, and demographic data, as well as unique individual and household IDs. Participant ID links to redemptions and issuance data.

Vendor data includes: vendor level ID, authorization dates, peer group, number of registers, and address. Vendor ID links to disqualifications. Peer group establishes some average prices. ▸ Sample vendor data

Redemptions and issuance data

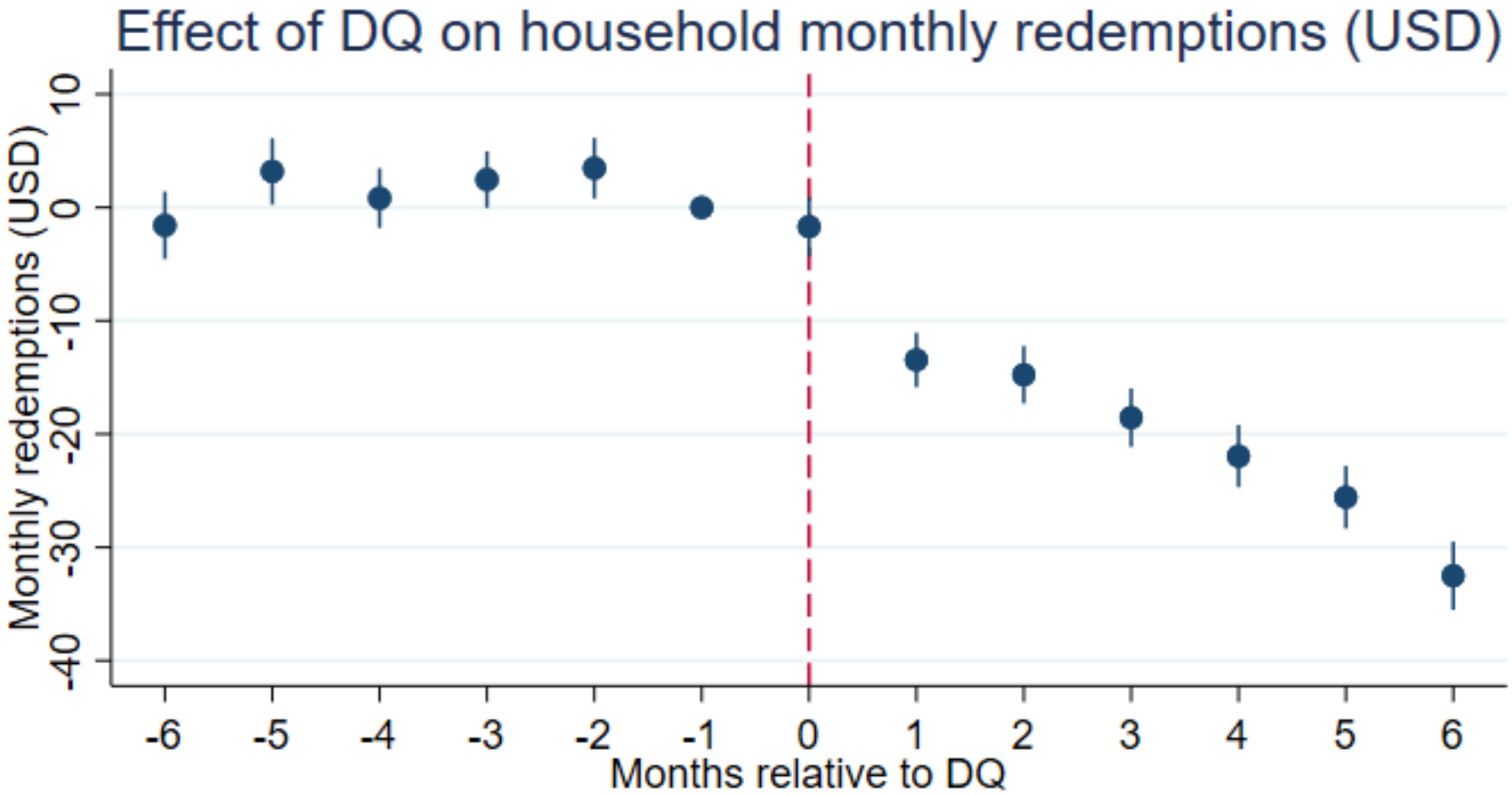
Redemptions data allow me to link vendors to participants.

Redemption occurs when participants exchange a food instrument for authorized foods. Vendor, participant, and food instrument ID are recorded at redemption (in addition to amount redeemed, type of benefit redeemed, and maximum allowable redemption value).



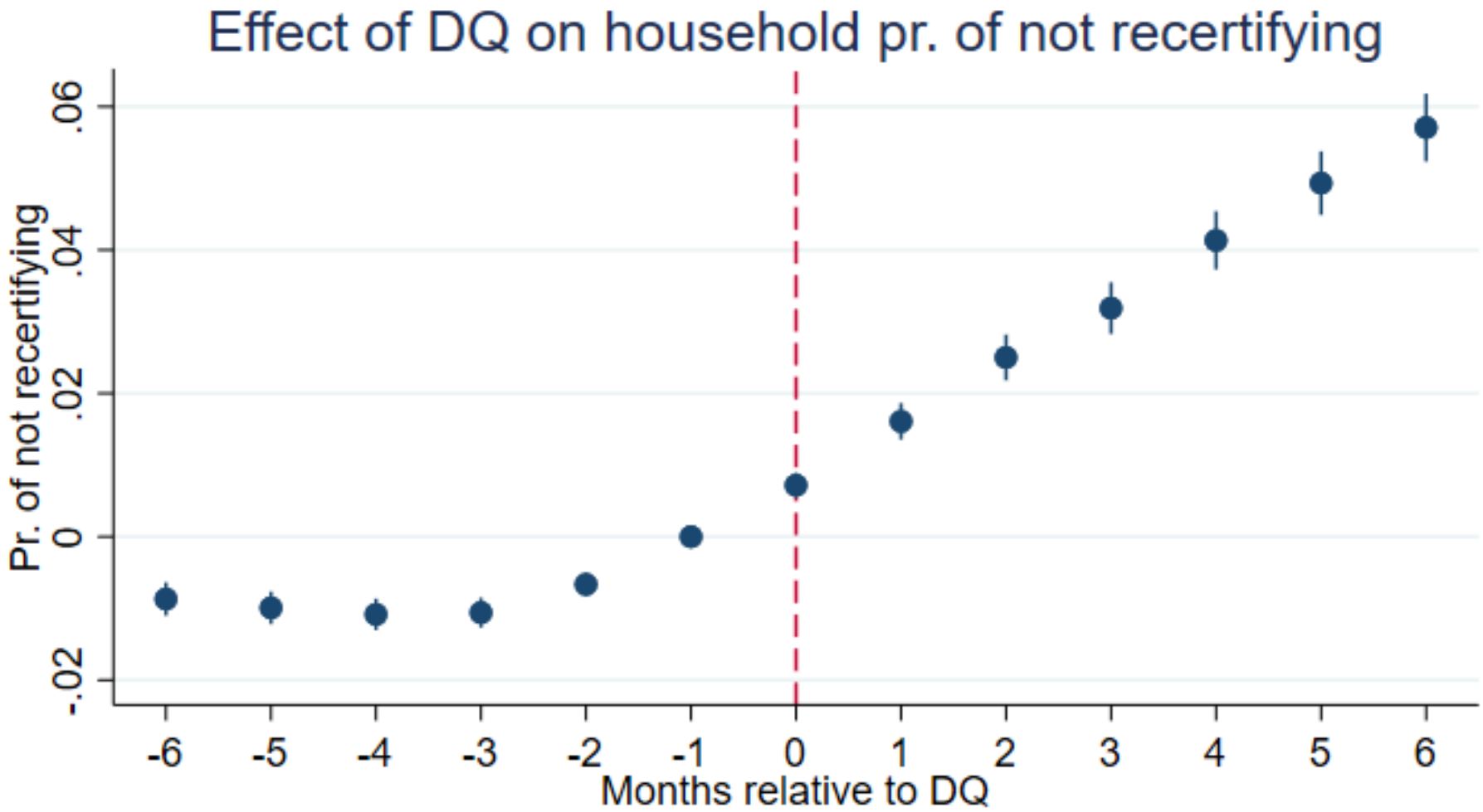
Issuance refers to the participant receiving food instruments from WIC. Issuance data provide more information on individual food instruments.

Preliminary results - redemptions



Pre-treatment (t=-6,...,-1) mean is 47.381 (SD = 67.854).
Post-treatment (t=0,..,6) mean is 33.615 (SD = 46.725).
Sample of households for which at least one WIC participating member experienced a salient DQ.
Sample only includes households for which at least one participant is issued FI in every period of the event wi
Sample restricted to households that experience at least 16 months between DQs
Also restricted to households that experience no other change in vendor authorization.
Event time zero indicates that in that month, a vendor that the household had shopped at within the previous 6
experienced a DQ.

Preliminary results - not recertifying



Pre-treatment (t=-6,...,-1) mean is .002 (SD = .048).
Post-treatment (t=0,..,6) mean is .135 (SD = .342).
Sample of households for which at least one WIC participating member experienced a salient DQ.
Sample restricted to households that experience at least 16 months between DQs
Also restricted to households that experience no other change in vendor authorization.
Event time zero indicates that in that month, a vendor that the household had shopped at within the previous 6
experienced a DQ.

Data Access

CDPH concerns:

- Data and internal communications secure
- Protect personally identifiable information

UCD concerns:

- Researcher maintains access to data
- Results from project are easily disseminated
- Minimize university liability

Lessons learned

R package for parallelization: `multidplyr`

- Creates n cores and partitions a data frame across cores to parallelize processing in R

R package for partitioning data: `arrow`

- Partitions data across an attribute you determine
- Saves n_g files in parquet format in a specified directory
- Loads partitioned data with minimal RAM
- Process select `dplyr` verbs on partitioned data, then collect

Preaching to the choir: basic coding and documentation practices (building code on a small random sample of data, maintaining code notes and external READ ME) make a difference.

Conclusion

This paper combines novel administrative data, a natural experiment that has not previously been studied, and an approach that weakens assumptions relative to previous models.

Using these contributions, this paper addresses a new question within the literature on how changes in the food retail environment affect households.

Preliminary results suggest a decrease in the amount of benefits redeemed after exposure to a DQ, as well as a decrease in the probability participants complete the recertification process after exposure to a DQ. These effects do not dissipate over the event window.

ceambrozek@ucdavis.edu

Evidence on WIC's benefits for participants

The literature documenting WIC's effects on birth outcomes, nutrition, is relatively large although often non-experimental. See Bitler and Seifoddini (2019) for an overview.

Khanani et al (2010) review the public health literature on WIC effectiveness.

Several papers have shown that WIC participation during pregnancy is associated with lower incidence of low birthweight (Currie and Rajani 2015; Figlio et al 2009; Joyce et al 2005, 2008; Rossin-Slater 2013).

Chorniy et al (2018) use administrative data from South Carolina to find that prenatal exposure to WIC is associated with lower rates of ADHD diagnosis and grade repetition.

Arteaga et al (2016) found that rates of food insecurity increase for children who age out of WIC, especially if these children have not yet started kindergarten.

◂ Question

Mixed evidence on participant responses to WIC access

Meckel, Uniat, Rossin-Slater (2020) find that policies which decrease the number of authorized (esp small) vendors decreases WIC take-up among first-time mothers, especially foreign-born mothers.

Rossin-Slater (2013) found that increased participant access to WIC clinics increased participation and improved birth outcomes.

Wu et al (2017) found that WIC participants in LA in food deserts did not redeem benefits differently than participants living outside of food deserts.

Hanks et al (2018) found that the eWIC implementation caused an increase in redemptions at a grocery chain in Ohio.

Meckel (2020) found that eWIC implementation increased probability that a ZIP code had zero authorized vendors and reduced participation and benefit redemption in poor neighborhoods in Texas. ◂ Context

Sample vendor data

_id	Vendor	Address	Second Address	City	Zip Code	County	Latitude	Longitude
418	SAFEWAY #1119	3747 GRAND AVE		OAKLAND	94610	ALAMEDA	37.817254	-122.24543
400	SAFEWAY #0951	867 ISLAND DR		ALAMEDA	94502	ALAMEDA	37.737543	-122.24025
2772	HOLIDAY QUALITY FOODS #40	9350 DESCHUTES RD		PALO CEDRO	96073	SHASTA	40.55959	-122.24021
1180	PRIME TIME NUTRITION #508	1530 SOLANO AVE	UNIT C	VALLEJO	94590	SOLANO	38.101754	-122.23901
1229	FOOD MAXX #489	1740 TUOLUMNE ST		VALLEJO	94589	SOLANO	38.124275	-122.23737
454	A & G MARKET	E 1849 21ST ST		OAKLAND	94606	ALAMEDA	37.791454	-122.23659
812	CHAVEZ MEAT MARKET	775 ARGUELLO ST		REDWOOD CITY	94063	SAN MATEO	37.488512	-122.2354
450	NOB HILL FOODS #632	2531 BLANDING AVE		ALAMEDA	94501	ALAMEDA	37.769347	-122.23448
458	EL RANCHITO MARKET	1536 23RD AVE		OAKLAND	94606	ALAMEDA	37.784675	-122.23423
3023	WAL-MART SUPERCENTER #1608	608 LUTHER RD		RED BLUFF	96080	TEHAMA	40.161082	-122.23357

◂ Data

Prior studies using WIC administrative data

Chorniy, Currie, and Sonchak (2020) use South Carolina participation, birth certificate, Medicaid, and education data.

Meckel (2020) uses Texas administrative data on vendors, combined with Nielsen HomeScan data and birth certificate data.

Similarly, Meckel, Rossin-Slater, and Uniat (2020) use publicly available lists of California authorized vendors, along with birth certificate data.

Saitone, Sexton, and Volpe (2015) use redemption and vendor data along with non-WIC price data.

Zhang, Tang, McLaughlin, and Diggs (2017) use Virginia EBT redemption data that link vendors and participants but do not examine participant responses to changes in vendor authorization.

◂ Data

Event study regression specification

Use an event study specification to track dynamics in treatment effects over an effect window $\left[\underline{j}, \bar{j}\right]$.

Observe the outcomes y over weeks $t = \underline{t}, \ldots, \bar{t}$.

The main outcomes of interest are participants' weekly redemption value and binary participation status.

Observe the events (DQs) over $\left[\underline{t} - \bar{j} + 1, \bar{t} + \left|\underline{j}\right| - 1\right]$.

Define an event for an individual i redeeming benefits at a store r disqualified at time t

$$d_{itr} = 1\left\{(DQ_{rt} = 1) \text{ and } (\exists\ k \in (t, t-1, ..., t-25)\ s.t.\ Red_{irk} > 0)\right\}$$

Assume that the treatment effect scales up additively with multiple events $d_{it} = \sum_{r=1}^{R} d_{itr}$.

◂ Empirics

Event study regression specification

Treatment indicators b_{it}^{j} for each week j in the effect window are defined using the event dummies.

$$b_{it}^{j} = \begin{cases} \sum_{s=t-\underline{j}}^{\bar{t}-\underline{j}-1} d_{is} & \text{if} \quad j = \underline{j} \\ d_{i,t-j} & \text{if} \quad \underline{j} < j < \bar{j} \\ \sum_{s=\underline{t}-\bar{j}+1}^{t-\bar{j}} d_{is} & \text{if} \quad j = \bar{j} \end{cases}$$

The empirical specification is

$$y_{it} = \alpha + \sum_{j=\underline{j}, j\neq -1}^{\bar{j}} \beta_j b_{it}^{j} + \gamma_t + \delta_i + \epsilon_{it}.$$

The coefficients β_j track the effect size over the effect window.

◂ Empirics

Policy implications of participant responses to vendor disqualification

Finding that disqualifications reduce participants' use of benefits or participation in the program suggests non-zero costs of disqualifications on participants.

A lower bound on the cost to participants (loss of social welfare from disqualifications) is the decrease in redemptions (adjusted for potential overcharging behavior of the vendor) as a result of a DQ aggregated over the number of DQs.

An approximation of the benefit from disqualifications is the decrease in the value of price discrimination (as addressed to some degree by Meckel, Uniat, and Rossin-Slater 2020). I may be able to say something about this empirically by looking at overcharging and/or trying to estimate counterfactual prices for overcharging stores.

Policy implications of participant responses to vendor disqualification (cont'd)

If the social welfare loss from disqualifications is close to zero (no statistically significant effects of DQs on redemptions), but the disqualifications reduce price discrimination and adverse experiences for participants, then program administrators should not reduce the current rate of sanctions. However, if the social welfare loss is comparable to the welfare gain from reducing fraud, then administrators may want to consider non-disqualification sanctions.