

# Long-Term Effect of Mathematical Literacy on Consumer Behavior: Evidence from Curricular Reforms

[Omitted]\*

June 26, 2025

## Abstract

This study investigates the impact of high school mathematics education on the financial behavior of individuals in later life. To assess this relationship, we employ a staggered treatment robust event study design that exploits variations in the requirements of high school math coursework across several US states following a series of curricular reforms in the late 1980s. Our findings indicate that increasing the required level of mathematical education has a significant impact on various aspects of financial behavior in adulthood. Specifically, we observe a reduction of up to 15.4% in monthly spending on packaged goods and a corresponding increase of over 25% in the use of coupons. These effects are robust to potential biases arising from staggered adoption or temporal and spatial correlations. In addition, our analyses demonstrate that the results vary between demographic groups, with a more pronounced effect among minorities. We also observe that the impact of high school math education on financial behavior persists with age, even though the magnitude of the effects may decrease.

---

\***Disclaimer:** Researcher's own analyses are calculated based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

Further, our study provides evidence of the role of mathematical literacy in shaping consumers' responses to economic crises. Treated households exhibited an increase in the uptake of coupons and price promotions following the Great Recession of 2008. Overall, our study highlights the important role of high school mathematics education in shaping financial behavior in adulthood and sheds light on the potential benefits of enhancing mathematical literacy among the population.

## 1 Introduction

Education exerts a persistent and multifaceted influence on consumer behavior, shaping both long-term financial trajectories and everyday consumption patterns. In particular, the formative years—especially during high school and college—play a critical role in molding individuals' identities and decision-making processes. Mathematical education, a cornerstone of most K-12 curricula, is central to this development. Engagement with mathematics not only equips consumers with quantitative skills but also fosters essential cognitive and behavioral competencies such as long-term planning, cost–benefit analysis, and self-regulation (see, e.g., [Mullainathan 2013](#)). The extensive and enduring benefits of mathematical education have been well documented in experimental and longitudinal studies. For instance, [Zacharopoulos et al. \(2021\)](#) report that students with deficiencies in mathematical education exhibit reduced neurotransmitter concentrations in brain regions pivotal for reasoning and cognitive learning, as revealed by advanced neuroimaging techniques. Similarly, [Ritchie and Bates \(2013\)](#) use a representative panel from the United Kingdom to show that mathematical achievement at age 7 robustly predicts socio-economic status, academic motivation, and career aspirations at age 42—even after controlling for predicted socio-economic status at birth. These findings underscore the critical role of mathematical literacy in shaping both cognitive capacities and socio-economic outcomes.

The influence of mathematical education extends beyond academic performance to encompass broader economic behavior. The emerging literature in consumer finance and labor economics has increasingly examined the long-term impacts of math and finance coursework during K-12 on later-life consumer decision-making (e.g., [Lusardi and Tufano 2015](#), [Cole et al. 2016](#)). While earlier research has largely concentrated on traditional financial outcomes—such as debt repayment, student loan management, and investment decisions—or macroeconomic indicators like employment and health (see, e.g., ?), less attention has been devoted to the implications for everyday retail decisions. These high-

frequency consumer choices, which include retail shopping behaviors and the use of price promotions, account for a substantial share of total consumer expenditures. For instance, data from the USDA indicate that U.S. consumers allocate, on average, 10.3 percent of their disposable personal income to food purchases alone, highlighting the significance of retail consumption in overall financial planning. Ignoring these behaviors potentially underestimates the holistic impact of mathematical education on consumer welfare—a gap that our research aims to bridge.

In this study, we leverage quasi random variation in state-level high school mathematical education requirements that emerged following a series of curricular reforms in the United States, initiated in response to a seminal 1983 report on K-12 education. By integrating Nielsen Household Panel data with survey information regarding panelists' locations at age eighteen (the typical high school graduation age), we identify whether individuals attended high school before or after their state implemented these reforms. Our primary empirical strategy involves both a classical Difference-in-Differences design with two-way fixed effects and a staggered-treatment-robust event study design to account for potential biases due to heterogeneity in treatment timing ([Callaway and Sant'Anna 2021](#)). The evidence from our analysis indicates that the reforms significantly affect later-life consumer spending behaviors. Specifically, we document robust effects on monthly expenditures for packaged goods, coupon utilization rates, and the frequency with which price promotions are exploited. These patterns are consistently observed across both state-cohort-level and household-level analyses, reinforcing the robustness of our findings.

The granular nature of the household-level data further enables us to explore heterogeneity in the treatment effects across diverse demographic groups. Our analyses reveal that, after controlling for other demographic factors, the augmented math requirement predominantly impacts minority groups, including Black and Asian consumers. Moreover, the observed effects are more pronounced among married and widowed individuals compared to their single counterparts, for whom the impacts are either attenuated or even

reversed. Although the magnitude of the treatment effects decreases with age, they remain statistically significant well into later adulthood. We also find notable differences when comparing middle-income households with those at the high- and low-income extremes, suggesting that the benefits of mathematical education are unevenly distributed across the socio-economic spectrum. Extending our analysis, we employ a Triple Differences framework to examine whether the reform effects vary across different macroeconomic conditions. Our results indicate significant differences in consumer behavior before and after the Great Recession of 2008: post-2008, the impact on coupon usage and the proportion of items purchased on sale intensifies, while the ratio of private label products to total purchases declines. These findings highlight the dynamic role that mathematical literacy can play in enhancing consumer resilience to economic shocks (cf. ?).

Our research contributes to several strands of literature and carries important implications. First, it extends the understanding of how childhood education, particularly mathematical proficiency, influences economic behavior later in life. Prior research has demonstrated that enhanced mathematical skills are associated with higher investment assets and lower debt levels (e.g., ?). Our findings suggest that the benefits of mathematical education are not confined to long-term financial planning; they also manifest in routine, day-to-day decisions, such as those observed in retail transactions. Given that retail consumption may account for as much as 40–50% of total expenses for many households (see, e.g., [Dubé et al. 2018](#); [Nevo and Wong 2019](#)), the implications of our results are particularly far-reaching.

Second, our study underscores the potential of mathematical education to bolster economic resilience. By fostering skills that lead to more judicious financial decisions—such as increased coupon usage and strategic purchasing of discounted products—mathematical education can serve as a buffer against economic volatility, thereby enhancing individual financial stability in uncertain times (cf. ?). Finally, the insights derived from our analysis offer practical guidance for retailers and brands. A deeper understanding of how

educational background influences consumer responses to price promotions and coupon strategies can inform more effective targeting and segmentation strategies, ultimately leading to improved marketing performance and consumer satisfaction (cf. ?).

In summary, the present research not only reinforces the long-held notion that education is a key determinant of economic outcomes but also elucidates the specific pathways through which mathematical education shapes everyday consumer behavior. Our findings advocate for policies that promote mathematical literacy as a means of enhancing long-term economic well-being and consumer resilience. Future research should continue to explore the interaction between educational interventions and financial decision making, particularly in light of evolving economic challenges and technological advances in data collection and analysis.

## 2 Related Literature

Our research makes several significant contributions by bridging multiple strands of literature that inform both theory and practice in marketing, consumer finance, and economics.

First, our study deepens the extensive literature on financial and mathematical literacy (Hastings et al. 2013, Lusardi and Mitchell 2014). Prior investigations have predominantly focused on the pathways through which financial and mathematical education shape financial literacy and influence outcomes such as credit scores, wealth accumulation, debt management, and savings behavior. Notably, the empirical evidence on financial literacy training remains mixed. For example, studies such as Bernheim et al. 2001, Skimmyhorn 2016, Lusardi and Tufano 2015, and Kaiser et al. 2022 document positive associations between financial education and improved financial outcomes, whereas others (e.g., Cole et al. 2016, Fernandes et al. 2014) report non-significant effects. In contrast, the impact of mathematical literacy appears to be more robust, with research by Cole et al. (2016), Brown et al. (2016), and Goodman (2019) revealing that additional math coursework

significantly enhances debt repayment behaviors, credit card usage patterns, student loan management, and even labor market performance. Our work builds on these insights by shifting the focus toward an outcome that has received comparatively little attention in prior studies—future consumption behavior. In doing so, we not only broaden the scope of financial and mathematical literacy research but also propose a mechanism by which enhanced numeracy skills may reduce cognitive biases and improve long-term financial decision-making.

Second, our research contributes to the burgeoning literature on consumer habit formation, an area that increasingly underscores the lasting influence of early-life experiences on adult consumption patterns. A growing body of evidence suggests that factors such as childhood geography, economic shocks, and formative educational experiences can profoundly shape long-term consumer preferences. For instance, [Bronnenberg et al. \(2012\)](#) employs the Nielsen Panel along with supplementary survey data to demonstrate that persistent effects of childhood location account for over 40 percent of the geographic variation in market shares. Similarly, [Severen and Van Benthem \(2022\)](#) capitalizes on exogenous variations in gasoline prices stemming from oil shocks in the 1970s, revealing that price fluctuations during formative years have enduring effects on travel behavior and commuting choices. Complementing these findings, [Binder and Makridis \(2020\)](#) shows that consumers exposed to the oil crisis tend to develop a more pessimistic economic outlook, while [Malmendier and Shen \(2018\)](#) finds that high unemployment during childhood is linked to increased coupon usage, a higher propensity for purchasing sale items, and a greater preference for generic brands later in life. Our study extends this literature by introducing the dimension of mathematical education as a formative experience that may shape consumption habits. In particular, we argue that the analytical skills and problem-solving abilities acquired through rigorous math training can have enduring effects on consumers' preferences, influencing not only their financial decisions but also their broader consumption patterns in adulthood.

Third, our research intersects with the marketing literature on promotional channel utilization, providing new insights into how consumers process and respond to price promotions and cost-saving opportunities. Recent studies have underscored the importance of cognitive abilities in navigating complex promotional offers. For example, Venkatesan and Farris (2012) identifies a strong correlation between coupon take-up and consumers' utility maximization skills, suggesting that numeracy plays a pivotal role in decision-making. Moreover, Lalwani and Wang (2019) employs a multi-method approach to illustrate that effective coupon usage hinges on a high degree of self-regulation—a trait that is often influenced by cultural and socioeconomic factors. Research by Suri et al. (2013) further highlights how math anxiety, or the fear of engaging in numerical calculations, can shape consumers' preferences for different promotional formats, while Viswanathan et al. (2005) observes that less numerate consumers tend to avoid discounted products in order to minimize the cognitive burden associated with deciphering complex price information. These findings align with the rationally inattentive consumer framework (e.g., Matějka and McKay 2015, Jerath and Ren 2021), which posits that consumers allocate limited attention across various alternatives in order to maximize utility. By integrating these perspectives, we hypothesize that enhanced mathematical education not only alleviates math-related anxiety but also reduces the need for rational inattention. As a result, consumers become better equipped to identify and exploit cost-saving opportunities, such as coupons and price promotions, thereby optimizing their purchasing decisions.

Lastly, our paper extends the nascent literature on private-label products in marketing and economics by examining the role of mathematical and financial education in mitigating informational barriers. Existing studies, such as Dubé et al. (2018), leverage the differential impacts of the Great Recession to reveal that income has a modest negative effect on private-label demand, while also challenging the notion that these products are inherently inferior. In addition, Bronnenberg et al. (2020) documents an informational gap that predisposes consumers to perceive private labels as lower quality compared to established

national brands. We explore whether improved mathematical education can serve as a countervailing force by enhancing consumers' analytical abilities and reducing cognitive biases. Such improvements may enable consumers to assess product value better, thereby potentially increasing the demand for private-label products when they represent superior value for money. This line of inquiry not only adds a novel dimension to the literature on private labels but also offers practical insights for retailers seeking to enhance the competitiveness of their private-label offerings.

Collectively, our study provides a comprehensive examination of how educational experiences—particularly in mathematics—can shape financial behaviors, consumer habits, and promotional engagement. By synthesizing insights across these interrelated domains, we advance both theoretical understanding and empirical evidence in a manner that holds significant implications for marketing science and practice.

### 3 Context and Data

#### 3.1 "A Nation at Risk" and Curricular Reform in the United States

The main exogenous source of variation used in this research arises from successive waves of curricular reforms that increased mathematics coursework requirements for high school graduation during the late 1980s. States across the United States adopted these reforms following the April 1983 publication of "A Nation at Risk," the final report of the National Commission on Excellence in Education (Gardner et al., 1983). The report opened with a stark warning: "Our Nation is at risk. Our once unchallenged preeminence in commerce, industry, science, and technological innovation is being overtaken by competitors throughout the world." It highlighted concerns that countries like Japan, South Korea, and Germany were making technological advances in industries where the United States had historically been dominant but was beginning to fall behind. The commission concluded that "learning is the indispensable investment required for success in the 'information age' we are entering."

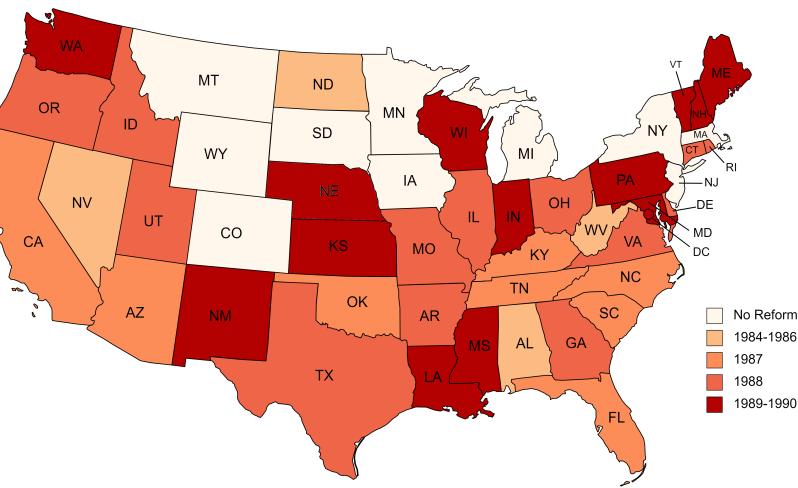


Figure 1: Timing of Math Curricular Reforms by State

One of the primary causes identified for the perceived educational decline was that "secondary school curricula have been homogenized, diluted, and diffused to the point that they no longer have a central purpose." The commission noted that U.S. high school students earned 25% of their credits in "physical and health education, work experience outside the school, remedial English and mathematics, and personal service and development courses." As a remedy, the commission proposed that state and local graduation requirements be extensively strengthened. Specifically, it recommended that high school students be required to take at minimum three years of mathematics and three years of science during their four years of high school. Prior to the report, no state had a three-year mathematics requirement, and the majority required only one full (year-long) course or less.

Figure 1 illustrates the differential timing of math curricular reforms across states. Only six states enacted reforms that applied to cohorts of high school graduates prior to 1987. The bulk of the reforms were roughly evenly split among cohorts graduating in 1987, 1988, and 1989, with only one state—New Mexico—enacting reforms after that period. This timing stemmed from state policymakers' swift responses to "A Nation at Risk" by legislating increased graduation requirements in year  $T$  (mostly 1983, 1984, or 1985) to

apply to students entering high school that year and thus graduating in year  $T + 4$ .

Importantly, variations in the timing of reforms are not closely related to geography, economic conditions, or political leaning. Every region contains both early and late reforming states, as well as states that did not implement the reforms at all. The timing varies similarly between Republican-led and Democrat-led states. This quasi-random variation in the timing of math curricular reforms, coupled with the fact that some states did not carry out the reforms and that the variation is not spatially concentrated, forms the basis of our main identification strategy to assess the causal impact of these reforms on later-in-life consumption-related savings behavior.

One might have concerns about other unobserved policy changes during this period that could confound our analysis. However, "A Nation at Risk" was the only major source of education reform of that era, at the state and national level, and it is unlikely that other policy changes would follow the exact same staggered adoption pattern. Nonetheless, in the robustness section, we address this concern by conducting placebo tests and discussing several methods to correct for any potential temporal or spatial correlation that may exist.

It should be noted that the main variation utilized for identification here is at the level of reforms, not the number of additional courses required by the reforms—that is, we treat the treatment as binary rather than continuous. While variations exist in the number of mandatory math courses needed for high school graduation before and after the reforms across states, the marginal impact of exactly one additional required course cannot be precisely identified. Each state and school district within a state differed in how many math courses they voluntarily "recommended" students take before the reforms. For example, [Goodman \(2019\)](#) points out that in California, even though the state had no formal math requirements pre-reform, very few students in pre-reform cohorts graduated without completing any math courses. Thus, it is unclear whether, for instance, two additional required math courses in one state are more substantial than one additional required math course in another state, and vice versa. Therefore, using the number of math courses as a

continuous treatment variable may lead to misleading interpretations. Instead, following the strategy used by [Goodman \(2019\)](#) and [Dursun et al. \(2021\)](#), we treat the treatment as binary, focusing on whether a state implemented the reform or not.

Additionally, our identification strategy benefits from the fact that the reforms were implemented independently of individual student characteristics or pre-existing trends in educational attainment. The staggered timing across states provides a natural experiment setting, allowing us to control for time-invariant state characteristics and common time effects. By employing a difference-in-differences framework, we can isolate the effect of increased math requirements on later-life consumer behavior, attributing changes to the reforms rather than other factors.

Our data sources include the Nielsen Household Panel, which provides detailed information on consumer purchases, and a survey indicating where panelists resided when they were eighteen years old (high school graduation age). This allows us to link individual consumer behavior to the state-level timing of the curricular reforms. By comparing cohorts who graduated before and after the reforms within each state, and across states with different implementation timings, we can estimate the causal impact of increased math education on consumption-related savings behavior.

In summary, the unique timing and implementation of the math curricular reforms following "A Nation at Risk" offer a valuable opportunity to study the long-term effects of mathematical education on consumer behavior. By leveraging this quasi-experimental setting, we aim to provide robust evidence on how enhanced math education influences individuals' financial decisions later in life, particularly in the context of consumption and savings.

## 3.2 Data

The primary data used in this study is NielsenIQ's Consumer Panel Data (provided through a partnership with the Kilts Center for Marketing at the University of Chicago), a panel of 40,000-60,000 households from 2004 - 2018, who, through the use of in-home

scanners, record all of their purchases (from any outlet) intended for personal, in-home use. These panelists provide information about their households, what products they buy, and when and where they make purchases. Transactions are recorded by households using an optical scanner in their homes. Members use barcodes on each of the packaged goods items they purchase during trips to supermarkets, convenience stores, mass merchandisers, and so on. Details on the stratified sampling methodology employed by Nielsen to promote the representativeness of the HomeScan panel can be found in Kilts Center for Marketing (2022). A major advantage of the household panel over retail scanner data is that it covers all retailers in the United States, including those without a contractual data-sharing agreement with Nielsen.

Conversely, one concern about these data is that many consumers do not reside in the same state as where they went to school; thus, if we were to use their current state to determine the treatment status, we would get imprecise estimates of the treatment effects due to incorrect treatment assignment. Utilizing the Supplemental Survey conducted in 2008, which asked panelists about their state of origin and when they moved, first used in Bronnenberg et al. (2012), we alleviate this concern by performing the analysis on both the full sample and a restricted sample consisting of only households where the head of household is still residing in the same state in which they were born (“Stayers”). The survey provides information on when and if a panelist moves out of their home state and how long they have been domiciled in their current state. A four-step matching procedure was performed to find which household members in the 2008 Survey are the heads of household<sup>1</sup>. We then use the head of household’s childhood state and year of high school graduation (defined as 18 years after the year of birth) to assign them to the correct treatment groups.

The main dataset consists of 23,376 households and more than 2,100,000 monthly observations. We narrow down the results to households where the head of household

---

<sup>1</sup>The detailed panelist - the household matching procedure can be found in Bronnenberg et al. (2012)'s appendix.

was born between 1960 and 1975, as the reforms were staggered introduced to cohorts born between 1965 and 1972. This helps us avoid other education reforms we are not observing (e.g., Cole et al. (2016) identify a new wave of reform starting in the early 2000s), reducing the sample to around 680,000 observations of 8,471 households. We also construct the main outcomes of interest: coupon usage, the ratio of purchases made on sale, share of private label, as well as control variables such as demographics, age, education, income, etc. We further supplement the household data with ZIP code level home price index as a proxy measure of household's wealth, to use as a control variable, following Dubé et al. (2018), as even though we can observe household's income, their total asset is not reported.

Our use of "Stayers" households has the main advantage of eliminating any contamination by the "mover effect" when movers partially adopt the consumption and spending patterns of residents of the local area they moved into, as shown in Bronnenberg et al. (2012). Additionally, the "Stayers" sample affords a more parsimonious interpretation of aggregated state-level results. Due to the variation in sample size between states and across cohorts, there are over 1,400 state-cohort pairs with only households who have moved out of their birth states. It is infeasible to account for this at an aggregated level. With the household level analyses, the inclusion of both sets of current and birth state fixed effects may alleviate some of this concern, but not entirely.

This restriction could raise some concerns about the generalizability of the results, as there may be unobserved factors that influence both the "staying vs. moving" decision and consumption behavior, rendering the results of the "Stayers" group not representative of the general population. We address this concern by replicating the main analyses with an extended sample consisting of both "stayers" and "movers" who moved out of their birth state after their eighteenth birthday<sup>2</sup>, increasing the sample size to approximately 33,000 households. The results of this robustness check, presented in Section 7 in the Appendix, are largely similar to the main results discussed in this paper, thus demonstrating

---

<sup>2</sup>As we only have information of a panelist's birth state and current state, we cannot reliably assign the treatment status to people who move out of their birth state before turning 18.

that the choice to focus on the “Stayers” sample does not lead to substantial threats to generalizability.

### 3.3 Variables Construction

To quantify the spending behavior of households, we examine the following quantities: (1) How often they buy discounted products, (2) How often do they use coupons, (3) How often do they buy private label products, and (4) What is their monthly expenditure. The first measure is calculated as:

$$PurchaseOnSale_{it} = \frac{\sum^s \mathbf{1}\{OnSale_s\}}{N_{it}}$$

That is, the ratio of number of transactions where a price promotion is active over the total number of transactions within a month  $t$  of a household  $i$ . This is representative of the *short-term saving behavior* of households, as the information about price promotion is often not known well in advance, hence the decision to purchase “on sale” products is often a just-in-time decision if they are price-conscious. The second outcome is calculated as:

$$CouponRate_{it} = \frac{\sum^s CouponValue_{sit}}{\sum^s TotalValue_{sit}}$$

This is relatively similar to the first measure, however here we are looking at monetary values instead of the numbers of transactions. Specifically, it is the ratio between the total amount of coupons used in month  $t$  by household  $i$ , in dollar terms, and the total value of their consumption (including coupons). On contrary to the previous measure, this is more representative of *long-term saving behavior* of households, as acquiring and redeeming coupons typically require longer-term planning and higher involvement by the consumer. Prior research in consumer behavior has demonstrated that coupon usage is a multi-stage and highly involved process (Bagozzi et al. 1992). To redeem coupons, users first need to be actively on the lookout for them in newspapers, in the mail, or through the Internet. They

then need to store and mentally catalog the acquired coupons. Additionally, the usage of coupons may require planning multiple trips to different stores. These long-term planning requirements mean that coupons are more enticing to consumers who are pure utility maximizers (Venkatesan and Farris 2012), and have stronger self-regulation (Lalwani and Wang 2019). Another difference is that coupon usage also signals higher variety seeking and lower brand loyalty (Bawa and Shoemaker 1987), as it is often tied to specific brands or products.

Another measure that may reflect the trade-off between price and brand loyalty is *Private label share*, i.e., the percentage of products purchased by a household  $i$  in month  $t$  that are private-labeled products. This is calculated as:

$$PLShare_{it} = \frac{\sum^s \mathbf{1}\{PrivateLabel_s\}}{N_{it}}$$

Dubé et al. (2018) examine the same measure, and find that income and wealth have a negative relationship with private label share, and thus private label usage may be one of the cost-saving measures available to consumers who are more price-conscious. Last but not least, we also examine the total monthly expenditure of each household, calculated as the total value of transactions minus the total value of coupons, which gives us a general overview of how much they spend, and whether the treated consumers are more frugal:

$$LogExpenditure_{it} = \log(\sum^s TotalValue_{sit} - \sum^s CouponValue_{sit})$$

If consumers with higher mathematical literacy tend to be more price-conscious, we should expect to see that, after controlling for household income and other demographic variables, the average monthly (grocery) expenditure of the treated households should be lower than that of the households in the control group, since they would utilize more cost-saving options.

## 4 Econometric Models

### 4.1 State-Cohort Level Analysis

Our primary identification strategy leverages the plausibly exogenous timing of mathematics curricular reforms across different states. In an ideal experimental setting, we would ascertain the causal treatment effect by randomly assigning individuals to varying levels of mathematics education during childhood and subsequently observing their consumption behavior over the ensuing decades. However, such an ideal experiment is neither realistic nor feasible. Alternatively, if we had access to data on the mathematical ability of each consumer in our panel, we could employ an instrumental variable strategy to identify the effects causally, perhaps using the curricular reform or local school district curriculum as an instrument. Unfortunately, such individual-level data are not available.

In this paper, we instead utilize a generalized Difference-in-Differences (DiD) approach, examining the changes in differences in average household outcomes between states with and without curricular reforms for cohorts before and after the reforms. It is important to note that this setup differs slightly from the traditional DiD framework, as the two indices are Unit (i.e., states) and Cohort (i.e., the year of high school graduation) instead of Unit and Time. Nonetheless, the main assumptions and identification strategy remain consistent with the standard DiD methodology. The basic regression equation is specified as follows:

$$Y_{sct} = \beta \cdot \text{Treat}_s \times \text{Post}_{sc} + \gamma_s + \gamma_c + \gamma_{st} + \varepsilon_{sct} \quad (1)$$

In this equation,  $Y_{sct}$  represents the average outcome of interest (e.g., percentage of purchases made on sale, coupon utilization, private label share, and total monthly expenditure) for cohort  $c$  in state  $s$  during period  $t$  (which corresponds to the panel year in our setting), weighted by the household projection factors provided by Nielsen. The variable  $\text{Treat}_s$  is

an indicator of whether state  $s$  implemented a mathematics curricular reform, and  $\text{Post}_{sc}$  is an indicator of whether cohort  $c$  graduated from high school after the reform was enacted in state  $s$ . Thus, the coefficient  $\beta$  captures the main effect of interest—the impact of the mathematics curricular reform on the outcomes of interest.

The terms  $\gamma_s$ ,  $\gamma_c$ , and  $\gamma_{st}$  denote sets of fixed effects at the state level, fixed effects at the cohort level, and fixed effects at the state by time, respectively. The set of state-level fixed effects accounts for time-invariant differences across states, such as cultural and geographical characteristics that may influence consumers' cost-saving behaviors. Cohort fixed effects control for cohort-specific unobserved factors, such as generational differences in tastes and age-related preferences. For instance, prior research has indicated that younger generations are more receptive to private-label products than older ones, partly due to greater familiarity with premium private-label offerings from retailers like Whole Foods or Trader Joe's. The state-by-time fixed effects,  $\gamma_{st}$ , control for time-varying unobserved factors at the state level, including changes in economic conditions, state-wide policy shifts, and the entry and exit of retailers. Additionally, this set of fixed effects accounts for differences in state composition across panel years, as households may join or leave the Nielsen panel over time.

Following [Allegretto et al. \(2017\)](#), we also estimate an alternative specification of equation (1) that incorporates state-specific linear trends:

$$Y_{sct} = \beta \cdot \text{Treat}_s \times \text{Post}_{sc} + \sum_{i \in S} \alpha_i (\mathbf{1}\{s = i\} \times c) + \gamma_s + \gamma_c + \gamma_{st} + \varepsilon_{sct} \quad (2)$$

In this specification, the term  $\sum_{i \in S} \alpha_i (\mathbf{1}\{s = i\} \times c)$  captures state-specific linear trends over cohorts, where  $c$  is treated as a continuous variable representing the cohort number. This addition accounts for variations in cohort trends between states, such as differences in demographic shifts or changes in consumer preferences. For example, some states may

experience a faster-growing minority population than others, or younger generations in more urbanized states may exhibit more pronounced changes in consumption behavior compared to those in rural states.

By including state-specific linear trends, we aim to control for potential confounding factors that could bias our estimates of the treatment effect. This approach allows us to more accurately isolate the impact of the mathematics curricular reforms on the outcomes of interest.

## 4.2 Staggered Treatment Robust Event Study Design

Recent advances in econometric literature (see, for example, [Goodman-Bacon 2021](#); [Imai and Kim 2021](#)) have highlighted potential biases in the classical two-way fixed effects model when applied to settings with staggered treatment adoption across units. In such settings, where different units receive the treatment at different times, the traditional Difference-in-Differences estimator may yield biased estimates of the treatment effect. This bias arises because units treated in later periods are compared not only with untreated units but also with units already treated in earlier periods. Consequently, the estimated Average Treatment Effect (ATE) becomes a weighted average of comparisons between treated and untreated units, as well as comparisons among treated units themselves, with some weights potentially being negative.

If the treatment effect is heterogeneous across different "treatment waves"—that is, the effect varies depending on when units receive the treatment—the estimated ATE may be biased, with the direction and magnitude of the bias depending on the proportion of early versus late treated units. This presents a serious threat to identification in our context, as the mathematics curricular reforms were implemented in different years across states, and heterogeneity in treatment effects is likely due to differences in pre-reform baseline math requirements and socio-demographic characteristics.

To address this identification challenge, several methodological corrections have been proposed in recent literature, notably those by [De Chaisemartin and d'Haultfoeuille \(2020\)](#),

[Callaway and Sant'Anna \(2021\)](#), and [Sun and Abraham \(2021\)](#). A common theme among these approaches is the aggregation of individual treatment effect estimates at the treatment wave and period levels, with differences arising in the estimation procedures. For our analysis, we adopt the method proposed by [Sun and Abraham \(2021\)](#) for two main reasons. First, this approach is straightforward to interpret, as it extends the classical event study or dynamic Difference-in-Differences design and relies on simple Ordinary Least Squares (OLS) regression, as opposed to more complex semi- or non-parametric estimators in alternative methods. Second, the [Sun and Abraham \(2021\)](#) estimator is readily applicable to unbalanced panel settings, which is pertinent in our context due to gaps in high school graduation cohorts in smaller states with insufficient household data.

In the robustness section, we further validate our findings by estimating the model using other staggered treatment robust methods, ensuring that our results are robust under a broader range of assumptions.

The staggered Difference-in-Differences event study approach begins with a modified event study regression:

$$Y_{sct} = \sum_g \sum_{k \in \mathcal{K}, k \neq -1} \delta_{gk} \mathbf{1}\{c - C_s = k\} \mathbf{1}\{C_s = g\} + \gamma_s + \gamma_c + \gamma_{st} + \varepsilon_{sct} \quad (3)$$

In this equation,  $\mathbf{1}\{c - C_s = k\}$  is an indicator for the relative treatment period  $k$ , representing the number of cohorts since the first treated cohort in state  $s$  (with negative values indicating cohorts before the treatment). The term  $\mathbf{1}\{C_s = g\}$  is an indicator for the "treatment wave," identifying the cohort  $g$  that was first treated in state  $s$  (which is set to zero for never-treated states). The coefficients  $\delta_{gk}$  capture the treatment effect for observations in states where the first treated cohort is  $g$ , for the  $k$ -th cohort relative to the reform implementation.

The fixed effects  $\gamma_s$ ,  $\gamma_c$ , and  $\gamma_{st}$  are the same as in equation (1), controlling for state-level,

cohort-level, and state-by-time unobserved heterogeneity, respectively.

Comparing equation (3) to the classical event study design:

$$Y_{sct} = \sum_{k \in \mathcal{K}, k \neq -1} \beta_k \mathbf{1}\{c - C_s = k\} + \gamma_s + \gamma_c + \gamma_{st} + \varepsilon_{sct} \quad (4)$$

we observe that the modified equation (3) disaggregates the relative period-specific effects (often termed "dynamic treatment effects") by treatment wave through the interaction terms. This disaggregation allows us to account for treatment effect heterogeneity across different treatment adoption times.

Once we estimate the treatment wave-specific dynamic treatment effects  $\hat{\delta}_{gk}$ , we proceed to compute the weights for each coefficient, reflecting the proportion of units in each treatment wave that have experienced at least  $k$  cohorts since treatment. Specifically, let  $N_g = \sum_s \mathbf{1}\{C_s = g\}$  denote the number of states in treatment wave  $g$ , and let  $\mathcal{G}_k$  represent the set of treatment waves with at least  $k$  cohorts since treatment. The corresponding weight for each estimator is calculated as:

$$w_{gk} = \frac{N_g}{\sum_{j \in \mathcal{G}_k} N_j} \quad (5)$$

Using these weights, we aggregate the individual  $\hat{\delta}_{gk}$  estimates to obtain the overall relative period-specific effects akin to those in the classical event study design:

$$\hat{\beta}_k = \sum_{g \in \mathcal{G}_k} w_{gk} \hat{\delta}_{gk} \quad (6)$$

The Average Treatment Effect on the Treated (ATT) is then estimated as the weighted average of the  $\hat{\beta}_k$  coefficients:

$$\text{ATT} = \sum_{k \in \mathcal{K}, k \neq -1} \left( \frac{\sum_{j \in \mathcal{G}_k} N_j}{\sum_{k' \in \mathcal{K}, k' \neq -1} \sum_{j \in \mathcal{G}_{k'}} N_j} \right) \hat{\beta}_k \quad (7)$$

[Sun and Abraham \(2021\)](#) demonstrate that this aggregated ATT estimator is asymptotically consistent and normally distributed, allowing for standard inference procedures.

By employing this staggered Difference-in-Differences approach, we effectively address the potential biases associated with heterogeneous treatment effects and varying treatment adoption times, enhancing the validity of our causal inferences regarding the impact of the mathematics curricular reforms.

### 4.3 Household-Level Regression and Treatment Heterogeneity

In addition to the state-level aggregated regressions, we also estimate the Difference-in-Differences model using household-level monthly data:

$$Y_{isct} = \beta \cdot \text{Treat}_s \times \text{Post}_{sc} + \mathbf{X}'_{it} \boldsymbol{\Gamma} + \gamma_t + \gamma_s + \gamma_c + \gamma_{st} + \varepsilon_{isct} \quad (8)$$

In this specification,  $Y_{isct}$  denotes the outcome of interest for household  $i$  in state  $s$ , cohort  $c$ , and time  $t$ . The vector  $\mathbf{X}_{it}$  includes time-varying covariates such as the current age of the household head, education level, income, wealth, race, household size, and marital status, as well as state-level covariates like the unemployment rate, population, and GDP growth. The inclusion of these covariates allows us to control for individual and state-level factors that may influence consumption behavior. The term  $\gamma_t$  represents month fixed effects, controlling for temporal shocks such as seasonality. The fixed effects  $\gamma_s$ ,  $\gamma_c$ , and  $\gamma_{st}$  retain their previous interpretations.

We estimate two versions of equation (8): one weighted by the household projection factors provided by Nielsen and one unweighted. It is important to recognize that, despite

the richer set of controls afforded by household-level regressions, not all households in treated states and cohorts are directly affected by the curricular reforms. Since the legislation specifies only the minimum number of required math courses for graduation, some panelists may have attended high schools where the number of math courses already exceeded the new minimum requirements, thereby not experiencing a substantive change in their education. Consequently, the estimated effects may be attenuated, and we should exercise caution in interpreting the results as causal effects of the reform on each household.

An additional advantage of the household-level model is the capacity to investigate treatment effect heterogeneity across different consumer groups. We employ a Triple Differences (Differences-in-Differences-in-Differences) approach by interacting the  $\text{Treat}_s \times \text{Post}_{sc}$  term with demographic variables to examine how the effects vary across subpopulations. The regression equation is specified as follows:

$$Y_{isct} = \beta_1 \cdot \text{Treat}_s \times \text{Post}_{sc} + \beta_2 \cdot \text{Treat}_s \times \text{Post}_{sc} \times \text{Demographic}_{it} + \mathbf{X}'_{it} \boldsymbol{\Gamma} + \gamma_t + \gamma_s + \gamma_c + \gamma_{st} + \varepsilon_{isct} \quad (9)$$

In this equation,  $\text{Demographic}_{it}$  represents the demographic characteristic of interest for household  $i$  at time  $t$ , such as race, age group, marital status, education level, or household income. The coefficient  $\beta_2$  captures the differential impact of the curricular reforms across different demographic groups.

One potential concern is that these demographic variables may themselves be affected by the treatment. For example, enhanced math education could lead to higher income or greater educational attainment and therefore might introduce endogeneity into the model. However, due to the stratified random sampling design of the Nielsen Household Panel, which samples households based on socio-demographic factors, this issue should be mitigated. Additionally, by using the projection factors as weights in the Triple Differences regression, we ensure that the sample remains representative of the broader population,

further reducing potential biases arising from the indirect effects of the curricular reforms.

Through this analysis, we aim to uncover whether the impact of increased math education on consumption-related savings behavior differs across various demographic groups, providing insights into the mechanisms through which education influences economic outcomes.

By integrating both state-level and household-level analyses, our econometric models are designed to robustly estimate the causal effects of the mathematics curricular reforms on consumer behavior. The combination of these approaches allows us to control for unobserved heterogeneity at multiple levels and to explore the nuances of treatment heterogeneity, thereby contributing to a more comprehensive understanding of the long-term impacts of educational policy changes.

## 5 Results

### 5.1 Descriptive Statistics

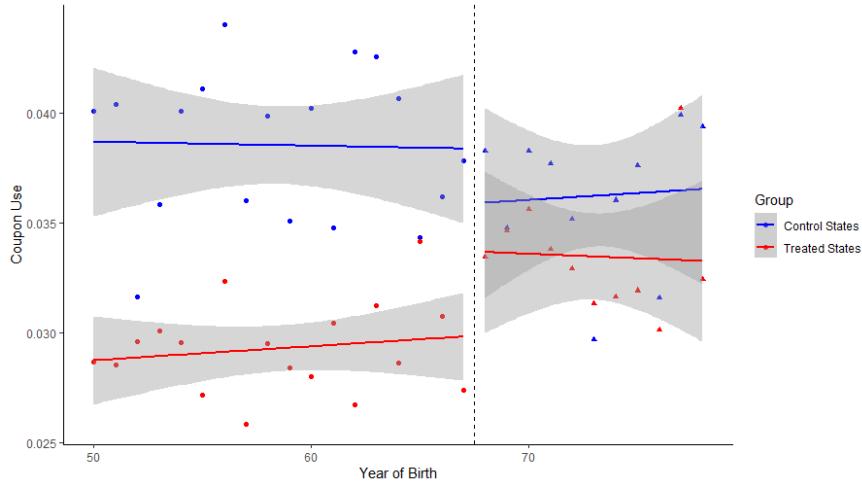


Figure 2: Model Free Evidence

The descriptive statistics of the Nielsen household panel are summarized in Table 1, broken down by Treated and Control states. From the table, the control variables are largely similar between the two conditions. Both groups have an average age of around

51, an average year of birth of around 1960 (treatment first started for the cohort born in 1966), and roughly similar education, income, household size, or marital status. The main difference here is between the share of minority households, with the treated states having a higher number of minority households (19.71% vs. 12.33%). This is addressed through the Projection Factor provided by Nielsen, which can be used as aggregation weights in state-level analyses and regression weights in household-level analyses to balance out the demographic factors. To further ensure the validity, we also rerun the analysis through a restricted sample of matched states by their demographic similarities and geographical closeness as a robustness check.

Table 1: Descriptive Statistics

Statistic	N	Control		N	Treat	
		Mean	Median		Mean	Median
Age	381,105	51.408	52	1,087,745	50.985	52
Cohort YOB	381,105	60.472	59	1,087,745	60.914	60
Minorities	381,105	0.123	0	1,087,745	0.197	0
Female Head Education	381,105	3.849	4	1,087,745	3.758	4
Male Head Education	381,105	3.158	4	1,087,745	3.161	4
HH Income Level	381,105	21.000	23	1,087,745	20.517	21
HH Size	381,105	2.503	2	1,087,745	2.491	2
Marital Status	381,105	1.860	1	1,087,745	1.824	1

Next, we also check the graphical evidence of the parallel trend assumption, as well as a model-free look at potential treatment effects. Figure 1 above and additional figures in the Appendix show the plots of the average outcome variables for each cohort of the treatment and control groups, along with the linear spline models, break at cohort born in 1966, when the first state started to roll out reform. From the plot of the coupon utilization rate, the parallel trend assumption appears to be satisfied, and the treatment group witnesses a rapid rise in coupon utilization rate for cohorts posttreatment. As for other outcomes, the model-free plots also show evidence of parallel pre-treatment trends between treated and control groups. However, while there appear to be post-treatment gaps (an indicator of

an effect), the changes are not as straightforward as that of coupon utilization. Naturally, we have to interpret these plots with caution, as the differential timing of treatment and differences between states and across cohorts are not yet taken into account, but overall it appears that the Difference-in-Differences strategy is appropriate in this setting.

## 5.2 Main Results

Dependent Variables: Model:	% Sale (1)	Coupon (2)	Private Label (3)	log(Spending) (4)
<b>A. Basic Difference-in-Differences</b>				
Treat × Post	0.0121 ** (0.0047)	0.0074 *** (0.0010)	0.0013 (0.0023)	-0.0083 (0.0143)
<b>B. Sun and Abraham (2021) aggregated ATT</b>				
ATT	0.0254 ** (0.0102)	0.0077 *** (0.0021)	0.0012 (0.0037)	-0.1677 *** (0.0301)
State-Year FEs	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Observations	17,478	17,478	17,478	17,478
Clustered Robust Standard Errors in parentheses				
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1				

Table 2: State-Year Level Main Results

**Table 2** presents the main state-level regression results, both of the naive Difference-in-Differences regressions, in Panel A, and of the [Sun and Abraham \(2021\)](#)'s staggered adoption robust models in Panel B, with four main outcomes of interest: (1) *Purchase-on-sale Rate*, (2) *Coupon Utilization Rate*, (3) *Private Label Share*, and (4) log-transformed *Monthly Expenditure*. We can see that the classical Difference-in-Differences results indicate precisely estimated and substantial positive effects of mathematical curricular reforms on Purchase-on-sale and Coupon Utilization Rates. Specifically, from column (1), the reforms led to an average 1.21 percentage point increase in the share of purchases made on sale.

This is approximately a 5.2% increase over the average Purchase-on-sale rate of 23% of the whole sample. As for the Coupon Utilization Rate in column (2), the naive DiD result indicates a 0.74 percentage point effect. While this may appear modest at first glance, given the average coupon usage rate of just 2.8% in the whole sample, this represents a considerable 26.2% increase over the average. On the other hand, from columns (3) and (4), we see that the estimated effects on Private Label Share and (log-transformed) Monthly Expenditure are negligible and cannot be accurately estimated.

As discussed above, however, these results may be biased due to the curricular reforms' staggered adoption nature. To address this, in Panel B of **Table 2**, we present the aggregated Average Treatment effects on the Treated of [Sun and Abraham \(2021\)](#)'s staggered adoption robust design, using the procedure from **Section 4.2**. Furthermore, the effects aggregated by relative periods (i.e., number of cohorts since the first treated ones in each state) are illustrated in event study plots in Figure 3. Additionally, in the Appendix, we also report the treatment effects broken down by each treatment wave to illustrate the treatment heterogeneity that may have caused the differences.

From column (1) in Panel B, we see that after accounting for staggered adoption, the estimated effect of mathematical curricular reforms on Purchase-on-sale rate is more than doubled, from 1.21 to 2.54 percentage points. Taking a closer look at the effect-by-treatment wave estimates in the Appendix, the downward bias in the naive model, in comparison to the robust estimate, appears to be driven by a large negative estimate (-16.68 percentage points) in the 1986 Treatment Wave, which consists of only the state of Nevada, which can be idiosyncratic due to its small, rapidly changing, and highly transient population. On the other hand, the results for Coupon Utilization and Private Label Share are very similar to what is observed with the naive model, indicating largely homogeneous effects between treatment waves, as confirmed by the treatment wave-specific results.

Perhaps the most remarkable divergence between the naive DiD results and the event study results is that between the estimates of curricular reforms' effect on log-transformed

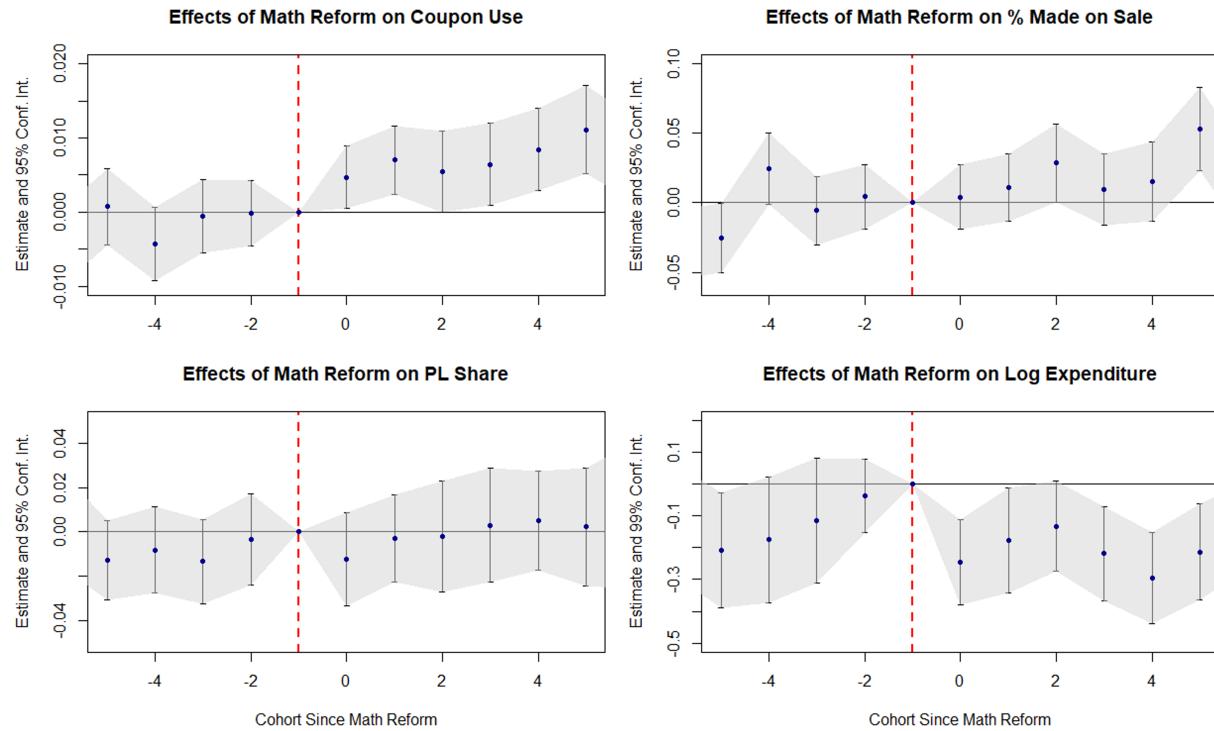


Figure 3: [Sun and Abraham \(2021\)](#) Staggered Treatment Robust Event Study Estimates

Monthly Expenditure. While the naive DiD estimate in Panel A suggests a negligible effect on monthly expenditure, after applying the correction for staggered adoption, the result in Panel B indicates a statistically and numerically significant relationship. Specifically, the curricular reforms led to a 15.4 percentage point reduction in monthly packaged goods expenditure, which translates to \$ 62 decrease per month out of an approximately \$ 400 average monthly expenditure. Similar to the Purchase-on-sale rate case, the bias in the naive DiD estimate appears to be driven mostly by idiosyncratic outliers, here it is the last treatment wave of Class of 1990, which consists of only the state of New Mexico<sup>3</sup>. This further illustrates the importance of accounting for staggered adoption in evaluating the effect of policies with a Difference-in-Differences design.

Figure 3 plots the treatment effects estimated using [Sun and Abraham \(2021\)](#)'s approach,

<sup>3</sup>Center for Urban Research at the City University of New York and US Census put New Mexico as the most difficult state to get reliable data from. Source: [UNM Newsroom](#).

aggregated by the number of cohorts since the first treated ones. The event study plots map out the “dynamic treatment effects”, or “treatment paths”, of curricular reform on each outcome variable. From the figure, except for Private Label Share, the general trends appear to increase in effect magnitude the further away a cohort is from the first treated one. This is an anticipated observation, as policies, especially education policies, often take many years before they can be implemented in full. The event study plots also serve as visual evidence of the parallel trends assumption, an important assumption for interpreting the Event study and Difference-in-Differences design results as valid causal quantities. As illustrated in the plots, the effects of pre-treatment cohorts are not statistically different from zero. Thus there are no diverging or converging trends between the treated and control states in the pre-treatment periods.

Dependent Variables: Model:	% Sale (1)	Coupon (2)	Private Label (3)	log(Spending) (4)
<b>A. Unweighted</b>				
Treat × Post	0.0108** (0.0044)	0.0054*** (0.0009)	-0.0029* (0.0017)	-0.0539*** (0.0105)
<b>B. Weighted</b>				
Treat × Post	0.0090*** (0.0018)	0.0028*** (0.0004)	0.0014* (0.0008)	-0.0468*** (0.0059)
Control Variables	Yes	Yes	Yes	Yes
State-Year FEs	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Month-Year FEs	Yes	Yes	Yes	Yes
Observations	1,468,850	1,468,844	1,468,844	1,468,844

Clustered robust standard-errors in parentheses

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 3: Household Level Fixed Effects Results

Table 3 below reports the estimates from our household-level models. In both the unweighted and weighted versions, we observe qualitatively similar effects to those found

in the state-level models. Specifically, there are positive and statistically significant impacts of the curricular reforms on the purchase-on-sale and coupon utilization rates, as well as negative and significant effects on monthly expenditure. The impact on private label share are only marginally significant and quantitatively small compared to the average, consistent with previous findings.

In the household-level estimations, the magnitudes of the effects are somewhat smaller, likely due to variation being absorbed by the wide range of control variables and more granular fixed effects. The impact on purchase-on-sale ranges from 0.90 to 1.08 percentage points, on coupon utilization from 0.28 to 0.54 percentage points (still representing a 10–20% increase over the average rate), and on monthly expenditure from a 4.7% to 5.4% reduction. The statistical significance of the effects in both unweighted and weighted models suggests that potential biases from covariate imbalance do not pose a substantial threat to the validity of the results.

Overall, from the main results, we can conclude that the mathematical curricular reforms lead to strong and persistent effects on their overall monthly packaged goods expenditure, even after controlling for demographic factors such as income, education, geographical location, and so on, indicating an impact on the overall saving tendency. This effect is the strongest in the long-term saving action of coupon usage, which is not surprising, as mathematical literacy is an important factor in determining planning ability. As for short-term saving action, we also observe significant effects of the curricular reforms on the share of purchases made on sale. However, this is less precisely estimated, perhaps due to a more significant number of potential confounders.

The effect on Private Label Share, however, is muted and is only statistically significant (yet still modest in magnitude) in the specification with state-specific linear time trends. While this may appear surprising, these results align with previous findings of [Dubé et al. \(2018\)](#), which show that generally, consumers do not consider private-label products inferior goods. The same paper further finds that the effects of income and wealth on

private-label demand are overall modest. Additionally, through a blind taste test, [Bronnenberg et al. \(2020\)](#) provides evidence of a sizeable informational barrier to demand for private-label products. Taken together, it is possible that private label products are not perceived as a cost-saving option by consumers, even the more mathematically savvy ones. The increased mathematical ability does not negate the branding effect of national brands, thus explain the null effect observed in our analyses.

### 5.3 Treatment Heterogeneity

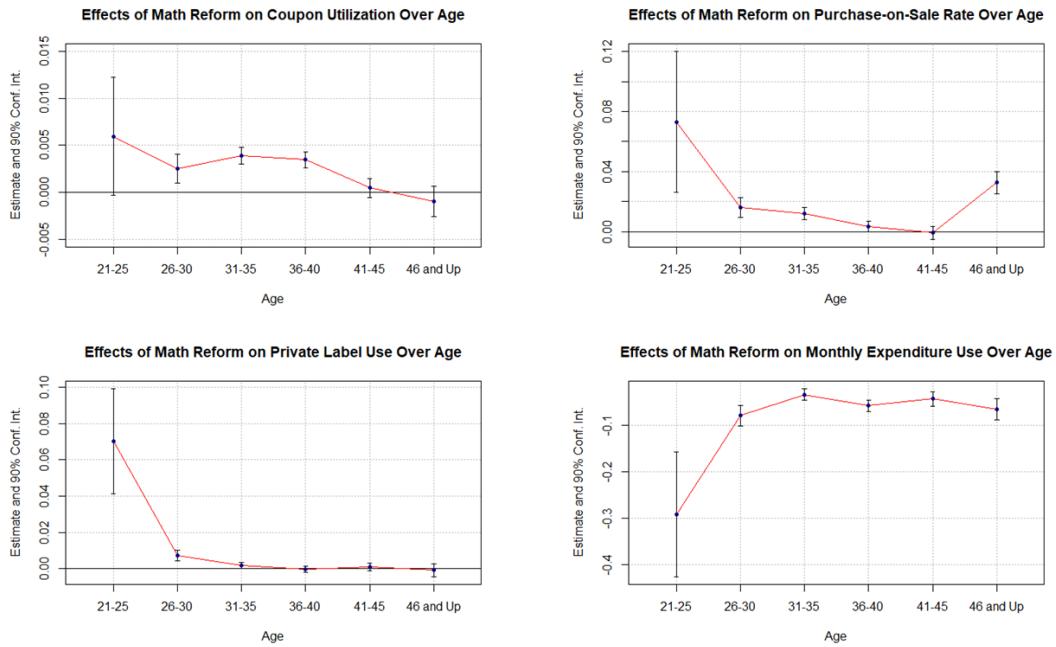


Figure 4: Treatment Effect Heterogeneity by Age

To obtain a more granular view of the effects, we examine the treatment effect heterogeneity across several demographic variables through Triple Differences models, while control for other demographic factors, as discussed in the previous section. These heterogeneities in treatment effects are illustrated through a series of plots as presented here and in the section 9.3 of the Appendix, along with the detailed regression tables. In this section, we will discuss the most interesting findings.

First, we can see in Figure 4 that the treatment effect appears to be attenuated with

age across all outcomes. The effects are strong around early to middle adulthood and gradually decline toward zero at around 45. This overall trend can be explained through the lens of “Neuroplasticity”, a framework in neuroscience that postulates that the way our brain functions is significantly shaped through experience during childhood and early adulthood and gradually changes over time as we get older. This means the saving tendencies and abilities from the increased mathematical literacy can be gradually replaced with age, or perhaps as not-treated consumers gain additional experience, their saving behavior catches up with the early boost from math education. Another explanation for this decreasing trend is that as consumers get older, their need for consumption-related savings decreases as they gain access to other savings channels such as investment accounts, real estate, and retirement planning. Additionally, the attenuation of effects over age may result from strong consumption smoothing by treated households, as they choose to give up more consumption and save more during their younger years to prevent future adverse outcomes. Finally, another notable observation with regards to age heterogeneity is that the effect on Private Label Share is particularly strong for younger panelists while being indistinguishable from zero for older ones. This may be evidence of the previously discussed informational barrier to private label products. It is possible that younger customers are more familiar with private label products (which are generally newer than national brands) and thus are more likely to utilize those as a cost-saving measure.

Moving on to socio-demographic factors, we find strong heterogeneities in treatment effects across races. The impact of math curricular reforms on Coupon Utilization is much more substantial for minorities such as Asian or Black households in comparison to White households. Additionally, the effect on private label shares, which is muted overall, appears stronger in minority demographics. This pattern is in line with [Goodman \(2019\)](#)’s finding that the curricular reforms have a more pronounced impact on minorities, as these tend to reside in school districts with lower or non-existent math requirements before the reforms. As for purchase-on-sale and total expenditure, however, while the effects remain

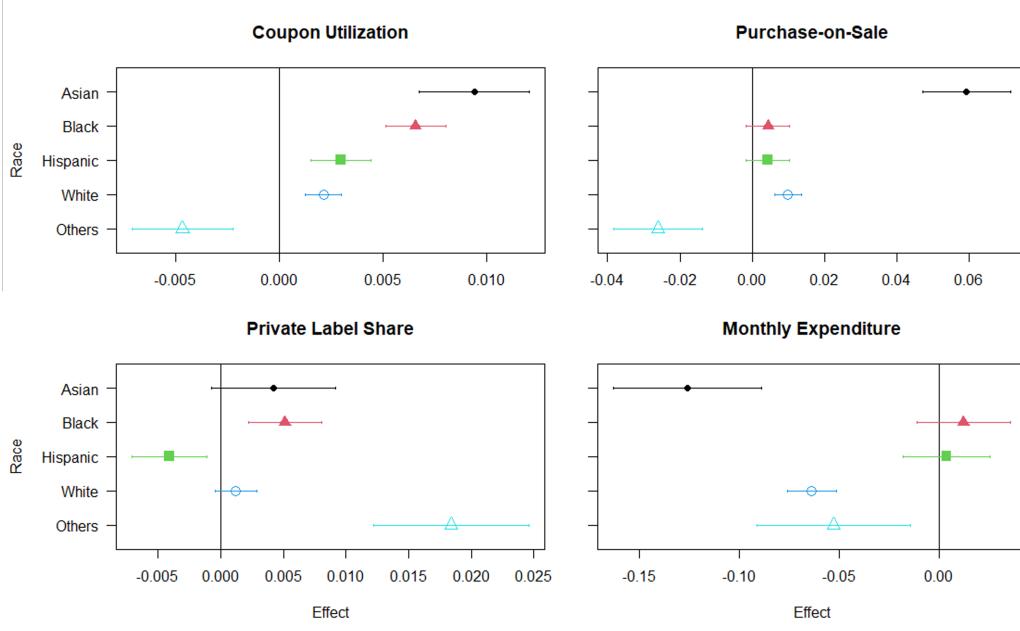


Figure 5: Treatment Effect Heterogeneity by Race

significant for Asian and White households, they are negligible for Black and Hispanic households. This may be due to differences in shopping habits or access to promotional offers<sup>4</sup>.

Another interesting result is that the effects on Coupon Utilization, Purchase-on-sale rate, and Monthly Expenditure appear to be the most significant for middle-class income households (household income of around \$60-100,000/year). This contradicts the more intuitive assumption that the effects would be more substantial for lower-income households. It is possible that for households in this group, the use of cost-saving measures is obligatory, out of necessity, so the incremental effect of mathematical literacy is not as substantial. Curiously, the opposite is observed for Private Label share. The effect is only statistically significant for households with lower than \$30,000 more than \$100,000 annual income. This is potentially due to the lower priority of private labels in terms of cost-saving measures, as discussed previously, and it is possible that buying private

<sup>4</sup>Regressing the Purchase-on-sale rate on Race, with the same set of fixed effects and control variables and only *untreated* observations, shows that Black households on average have 32.36% lower rate than White households. In comparison, Asian households have 39% higher rate, suggesting a pre-existing disparity.

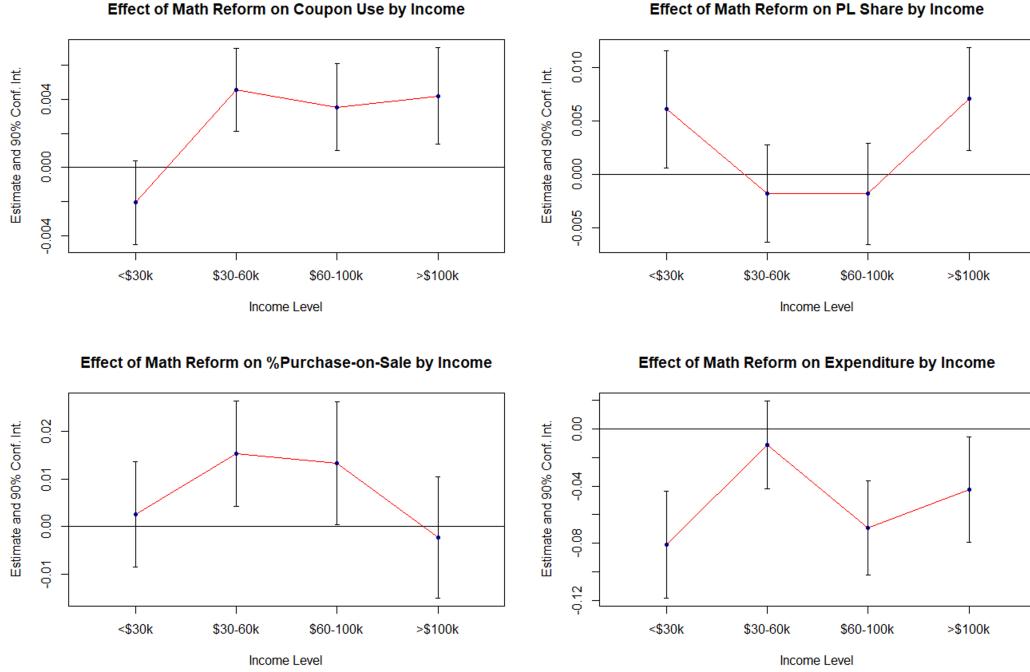


Figure 6: Treatment Effect Heterogeneity by Income Level

label products is mainly a substitute for buying discounted products (as opposed to using coupons).

## 6 Robustness Checks

**(1) Placebo Tests:** To ensure the validity of the results discussed in previous sections, we perform several robustness checks. First, even though the event study plots in Figure 3 demonstrate evidence of parallel trends, recent literature has pointed out that visually examining the plots alone may not be sufficient. To test the robustness of this assumption, as well as whether the treatment effects follow parallel paths post-treatment, and to ensure the estimates are not spurious, we carry out two “in-time” Placebo tests. The Placebo tests are performed by shifting the treatment indicators forward and backward for five relative periods, creating “placebo treatments” that either start five years earlier or later. If the observed results are not due to some other unobserved confounded shocks (e.g., other policy changes), we expect to observe the null effects of these placebo treatments.

Section 9.5 in the Appendix offers further detail on the procedure. As expected, the null results in tables 18 and 19 provide further evidence of the satisfaction of the parallel trends assumption, as well as attenuating the concern about other unobserved shocks in pre or post-treatment periods.

**(2) State Specific Time Trends:** Another concern about the validity of the results is the potential spatial heterogeneity between states that decided to adopt the reform at different times or not reform at all. We test this using two methods. First of all, following Allegretto et al. (2017), we allow for state-level parametric time trends by adding an interaction term between state dummies and the (continuous) cohort variable. From Table 4 in the Appendix, we can see that the results are largely qualitatively similar in direction and magnitude. Even though the effect of the reforms on Private Label Share is more precisely estimated in this specification, its overall magnitude is still small (0.9 percentage point increase on a 15% average Private Label Share). It could be the case that the differences in the inter-cohort dynamic of Private Label preference lead to considerable variation and thus the null results in the main analysis, so when we control for the trends, the effect becomes more identifiable.

**(3) Generalized Synthetic Control:** We further account for spatial correlation between states through the Generalized Synthetic Control method (Xu 2017). This is a generalization of the latent factor model with two-way fixed effects, the main motivating model behind the Synthetic Control method, an increasingly common causal inference technique in policy evaluation (Abadie and Gardeazabal 2003). The approach adds a latent factor structure, or *interactive fixed effects* (Bai 2009), to the model error, which accounts for unobserved time-varying heterogeneities. The Generalized Synthetic Control method calculates the individual factor loading (i.e., unobserved characteristics of the state) by eigen-decomposition of the squared residuals matrix in the pre-treatment period, then uses the complete set of data to calculate the time-varying factors (i.e., time-varying effects of the said characteristics). This is, in essence, similar to performing a Principal Component

Analysis on the pre-treatment residuals of the classical two-way fixed effects model. The model can be written in regression form as follows:

$$Y_{scy} = \beta Treat_s \times Post_{sc} + \sum_{n=1} f_c^n \lambda_s^n + \gamma_s + \gamma_c + \gamma_{ys} + \varepsilon_{scy} \quad (10)$$

With  $\sum_{n=1} f_c^n \lambda_s^n$  being the interaction between some unobserved characteristics of the state  $\lambda_s^n$  and its time-varying effect  $f_c^n$ . These could be some regional characteristics that differ by cohort, which will absorb potential spatial correlation. Additionally, the Generalized Synthetic Control approach uses only never-treated and not-yet-treated observations to estimate the factors then imputes those for the treated units, so in the basic case without additional factors other than time and unit fixed effects, this reduces to the imputation or two-stage Difference-in-Differences estimator (Borusyak et al. 2021, Gardner 2022). This imputation approach has been shown to be robust to staggered treatment timing, as the counterfactual construction process does not involve information from already treated units. Therefore, this also serves as an alternative to the Sun and Abraham (2021)'s correction used in our main analysis.

Through cross-validation, we find that the Generalized Synthetic Control model with only time and unit fixed factors fits the model best. However, we also estimate another version forcing the model to take at least one additional factor to exhaustively control for spatial correlation. The results of the two models are shown in Table 4, along with block-bootstrapped (at State-Panel Year level) standard errors. Overall, the estimates are qualitatively consistent with those of the main analysis, with statistically significant results for Purchase-on-sale rate and Coupon Utilization. Here, the effect on monthly expenditure is also substantial and statistically significant, similar to estimates with Sun and Abraham (2021)'s method, and differs from the naive Difference-in-Differences results.

**(4) Callaway and Sant'Anna (2021) Doubly Robust Estimator:** Aside from the main analysis and Generalized Synthetic Control, we also estimate another Staggered Treatment

Dependent Variables: Model:	% Sale (1)	Coupon (2)	Private Label (3)	log(Spending) (4)
GSC - No Add. Factor	0.0283*** (0.0067)	0.0085*** (0.0013)	0.0017 (0.0030)	-0.0661*** (0.0213)
GSC - 1 Add. Factor	0.0265*** (0.0094)	0.0085*** (0.0016)	-0.0002 (0.0022)	-0.0288* (0.0153)
Callaway and Sant'Anna (2021)	0.0107* (0.0088)	0.0064*** (0.0020)	0.0004 (0.0035)	-0.1211*** (0.0247)
State-Year FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes

Block-Bootstrapped Robust Standard Errors in parentheses

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 4: Generalized Synthetic Control & Callaway and Sant'Anna (2021) Results

robust estimator proposed by Callaway and Sant'Anna (2021). The main underlying idea behind this estimator is similar to that of Sun and Abraham (2021), with each treatment wave-cohort effect being estimated separately, and then aggregated into the ATT estimate. The main difference lies in the choice of treatment wave-cohort effect estimator and accompanying assumptions. While our primary analysis uses a modification of the classical event study estimator, this relies on a “doubly robust” estimator, which can be more robust to functional form assumptions than classical regression-based estimators. Additionally, Callaway and Sant'Anna (2021)’s estimator only assumes conditional parallel trends, and thus is less restricted than other, more classical, Difference-in-Difference methods. The estimates, as presented in Table 4, are close and qualitatively similar to those of the primary analysis and Generalized Synthetic Control models, further confirming the robustness of our findings.

**(5) Alternative Data Samples:** In section 9.4. in the Appendix, we present further robustness checks by estimating the main equations using alternative data samples. In the first alternative sample, we restrict the time window to a shorter period of 15 cohorts (Classes of 1978 to 1993), which still encompasses the reform waves, in order to reduce

further the possibility of spurious results due to other unobserved policy changes. The second alternative sample is restricted to only “eventually treated” states, and the cohorts since the last reform wave (Class of 1990) are dropped. This addresses the concern that the treated states may be inherently different from the control states by using only the “not yet treated” states in each cohort as the control. Finally, we also perform a robustness check with a matched sample of states based on their demographic variables (in 1982) and proximity to address potential biases due to covariates imbalance. The effects estimated using these alternative samples are qualitatively similar to the reported main results, indicating the robustness of the observed effects.

## 7 Extension: Response to 2008 Financial Crisis

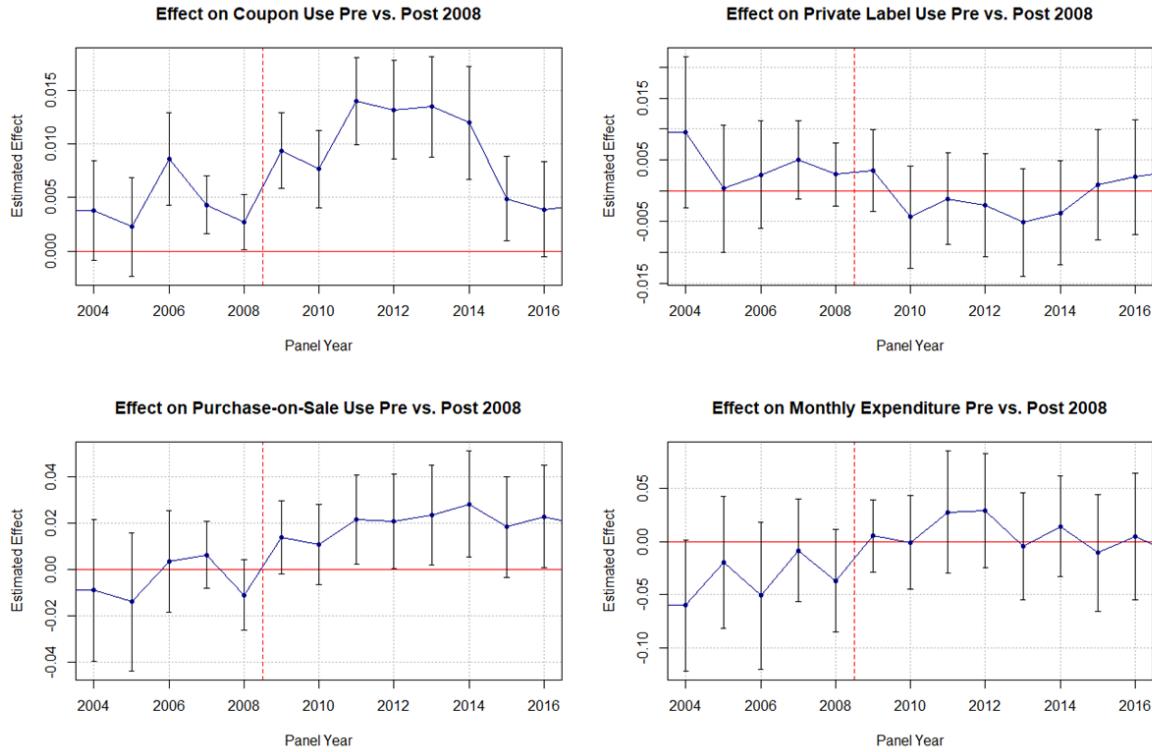


Figure 7: Treatment Effect Heterogeneity by Panel Year

In addition to the previous discussion on treatment heterogeneity, the wide temporal range of the household panel, from 2004 to 2018, also allows us to explore the dynamic of

the treatment effects over panel years. One interesting question is how the effects change in response to the financial crisis of 2008 and, in turn, how mathematical literacy affects the strategies consumers employ to respond to the period of uncertainty and economic downturn post-2008. This is done with a Trip Differences model as follows:

$$Y_{scy} = \beta_1 Treat_s \times Post_c + \beta_2 Treat_s \times Post_c \times Post2008_y + \gamma_s + \gamma_c + \gamma_y + \varepsilon_{scy} \quad (11)$$

Dependent Variables: Model:	% Sale (1)	Coupon (2)	Private Label (3)	log(Spending) (4)
Treat $\times$ Post	-0.0040 (0.0064)	0.0042*** (0.0012)	0.0039 (0.0027)	-0.0322* (0.0190)
Treat $\times$ Post $\times$ Post 2008	0.0233*** (0.0065)	0.0046*** (0.0012)	-0.0037 (0.0024)	0.0345** (0.0171)
State-Year FEs	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Observations	17,478	17,478	17,478	17,478

Clustered Robust Standard-Errors in parentheses

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 5: State Level Triple Differences Model Pre vs. Post 2008

Additionally, we also estimate a similar Triple Differences model at the household level, as well as the event study design as in the main analysis with the additional interaction term. Figure 7 illustrates the Panel Year specific treatment effect by interacting the  $Treat_s \times Post_c$  term with panel year dummies. From the plots, there are discernible differences in treatment effects pre- and post-2008. Specifically, the effect on Coupon Utilization increases substantially after 2008, and the effect on the Purchase-on-sale rate also increases from negligible to substantial in the post-2008 period. On the other hand, the effects on Private Label Shares and Monthly Expenditure show a reduction in magnitude, albeit it cannot be precisely estimated.

These visual observations are quantitatively confirmed by the regression results presented in Table 5. From Table 5, the treatment effect on the Purchase-on-Sale ratio rises from not statistically different from zero in the Pre-2008 period to 2.3 percentage points in Post 2008, while the effect on Coupon usage doubles from 0.42 percentage points to 0.88 percentage points. On the other hand, the effect on Monthly Expenditure, a statistically significant 3.2 % decline, vanishes after 2008. The household level estimates, as well as the estimates of staggered adoption robust ATTs by applying [Sun and Abraham \(2021\)](#)'s method separately to pre- and post-2008 samples (See Section 9.6 in the Appendix) also exhibit similar trends, albeit with a smaller decrease of the effect on Monthly Expenditure.

From the results above, the main picture that emerges is the curricular reforms exert substantial influence on how consumers respond to economic shocks such as the Great Recession. The differences in effects on Purchase-on-Sale and Coupon Utilization rates pre-and post-2008 indicate that the treated households are more responsive to the economic shocks and amplify their consumption-related saving behaviors such as collecting more coupons or hunting for more discounted products post-2008. This difference can be explained by the finding of [Zacharopoulos et al. \(2021\)](#), which shows that mathematical education during adolescence can increase neuroplasticity, thus enhancing the ability to adapt to changing situations. This stronger response of treated households may also be a product of better financial planning skills due to better numeracy or a better understanding of macroeconomic conditions. On the other hand, the lack of change in the effect on Private Labels confirms our previous conjecture that consumers likely do not view private label products as a feasible means of saving. Overall, these findings are testimonies of the importance of mathematical literacy and childhood education in understanding consumers' resilience to economic shocks.

## 8 Discussion

In this paper, we study the long-term effect of changes in required mathematical education due to curricular reforms on consumer saving-related behavior. Making use of the variation in state-level curricular reforms post “A Nation At Risk” report in 1983, we show that the reforms, which raised the amount of required mathematics coursework, led to long-lasting impacts on coupon usage, purchase-on-sale rate, and total grocery expenditure of consumers. Specifically, consumers who graduated from high school after math curricular reforms see a 15% decrease in total monthly packaged goods expenditure, an equivalence of more than \$60 a month, 0.5 - 0.8 percentage point effect on coupon utilization rate, which translates to roughly 25% higher rate than average, and a more modest yet still substantial increase of 5% higher Purchase-on-sale rate. These results are robust to the staggered adoption nature of the reforms and hold up to a wide range of robustness checks.

Additionally, we observe a sizable degree of heterogeneity in treatment effects across socio-demographic groups, with minorities, and middle-class, consumers seeing more substantial effects. Furthermore, the effects appear to dissipate with age gradually but still remain observable up to the 40-50 age range, and are more noticeable amongst married consumers. Last but not least, we also show that the effects on coupon use and purchase-on-sale intensified after the Great Recession of 2008. This is evidence of a more responsive reaction to adverse economic conditions amongst the treated households and speaks to the importance of mathematical education in enhancing the economic resilience of consumers.

These findings have several important implications. First of all, it contributes a marketing perspective to the large stream of research on the effects of mathematical abilities and the importance of mathematical education. Prior studies have demonstrated that math education and math abilities are predictive of future economic growth ([Hanushek and Woessmann 2010](#)), job prospects and income ([Goodman 2019](#)), or even infant health. The

results of this study further stress the importance of mathematical education, as it can further lead to more financially savvy, price-conscious, and less crisis-prone consumers. This is important for policy decision-makers, as it demonstrates another outcome one should consider in performing a cost-benefit analysis of investments in math education. It also has important managerial implications, providing managers with additional insights into planning their marketing decisions based on contextual information about consumer education. For example, promotion or coupon mail campaigns would be more effective if targeting more mathematically savvy consumers. The findings also serve as the basis for forming an expectation of which consumers are more resilient to crises, such as rising inflation or a recession.

The analyses in this paper are, of course, not without flaws. First, the use of curricular reforms as the primary identifying variation means we cannot recover the exact effect of each additional math course or an additional hour of math training. Additionally, the treatment is examined at the state level as the available data do not allow us to have an exact identification of who was affected by the increases in required coursework. Thankfully, these are unlikely to invalidate the conclusions, and we can reasonably expect the actual returns to math education to be qualitatively similar. A better view of the effects may be possible with more granular administrative data. Another limitation of the paper is that we can only look at behaviors related to packaged goods transactions, and generalization to other types of consumption is not guaranteed. Other consumer decisions such as housing or car purchases, college choices, etc. would be interesting questions to look at in the future.

## Funding and Competing Interests

All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. The authors have no funding to report.

## References

- Abadie, A. and Gardeazabal, J. (2003). The economic costs of conflict: A case study of the basque country. *American Economic Review*, 93(1):113–132.
- Allegretto, S., Dube, A., Reich, M., and Zipperer, B. (2017). Credible research designs for minimum wage studies: A response to neumark, salas, and wascher. *ILR Review*, 70(3):559–592.
- Bagozzi, R. P., Baumgartner, H., and Yi, Y. (1992). State versus action orientation and the theory of reasoned action: An application to coupon usage. *Journal of consumer research*, 18(4):505–518.
- Bai, J. (2009). Panel data models with interactive fixed effects. *Econometrica*, 77(4):1229–1279.
- Bawa, K. and Shoemaker, R. W. (1987). The coupon-prone consumer: some findings based on purchase behavior across product classes. *Journal of marketing*, 51(4):99–110.
- Bernheim, B. D., Skinner, J., and Weinberg, S. (2001). What accounts for the variation in retirement wealth among u.s. households? *American Economic Review*, 91(4):832–857.
- Binder, C. and Makridis, C. (2020). Stuck in the Seventies: Gas Prices and Consumer Sentiment. *The Review of Economics and Statistics*, pages 1–42.
- Borusyak, K., Jaravel, X., and Spiess, J. (2021). Revisiting event study designs: Robust and efficient estimation. *arXiv preprint arXiv:2108.12419*.
- Bronnenberg, B. J., Dubé, J.-P., and Sanders, R. E. (2020). Consumer misinformation and the brand premium: A private label blind taste test. *Marketing Science*, 39(2):382–406.
- Bronnenberg, B. J., Dubé, J.-P. H., and Gentzkow, M. (2012). The evolution of brand preferences: Evidence from consumer migration. *American Economic Review*, 102(6):2472–2508.
- Brown, M., Grigsby, J., van der Klaauw, W., Wen, J., and Zafar, B. (2016). Financial Education and the Debt Behavior of the Young. *The Review of Financial Studies*, 29(9):2490–

2522.

- Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.
- Cole, S., Paulson, A., and Shastry, G. K. (2016). High school curriculum and financial outcomes: The impact of mandated personal finance and mathematics courses. *Journal of Human Resources*, 51(3):656–698.
- De Chaisemartin, C. and d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–96.
- Dubé, J.-P., Hitsch, G. J., and Rossi, P. E. (2018). Income and wealth effects on private-label demand: Evidence from the great recession. *Marketing Science*, 37(1):22–53.
- Dursun, B., Eren, O., and Nguyen, M. (2021). Curriculum reforms and infant health. *The Review of Economics and Statistics*, pages 1–46.
- Fernandes, D., Lynch, J. G., and Netemeyer, R. G. (2014). Financial literacy, financial education, and downstream financial behaviors. *Management Science*, 60(8):1861–1883.
- Gardner, J. (2022). Two-stage differences in differences. *arXiv preprint arXiv:2207.05943*.
- Goodman, J. (2019). The labor of division: Returns to compulsory high school math coursework. *Journal of Labor Economics*, 37(4):1141–1182.
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*.
- Hanushek, E. A. and Woessmann, L. (2010). *The high cost of low educational performance: The long-run economic impact of improving PISA outcomes*. ERIC.
- Hastings, J. S., Madrian, B. C., and Skimmyhorn, W. L. (2013). Financial literacy, financial education, and economic outcomes. *Annual Review of Economics*, 5(1):347–373.
- Imai, K. and Kim, I. S. (2021). On the use of two-way fixed effects regression models for causal inference with panel data. *Political Analysis*, 29(3):405–415.
- Jerath, K. and Ren, Q. (2021). Consumer rational (in) attention to favorable and unfavorable product information, and firm information design. *Journal of Marketing Research*,

- 58(2):343–362.
- Kaiser, T., Lusardi, A., Menkhoff, L., and Urban, C. (2022). Financial education affects financial knowledge and downstream behaviors. *Journal of Financial Economics*, 145(2):255–272.
- Lalwani, A. K. and Wang, J. J. (2019). How do consumers' cultural backgrounds and values influence their coupon proneness? a multimethod investigation. *Journal of Consumer Research*, 45(5):1037–1050.
- Lusardi, A. and Mitchell, O. S. (2014). The economic importance of financial literacy: Theory and evidence. *Journal of Economic Literature*, 52(1):5–44.
- Lusardi, A.-M. and Tufano, P. (2015). Debt literacy, financial experiences, and overindebt-edness. *Journal of Pension Economics and Finance*, 14(4):332–368.
- Malmendier, U. and Shen, L. S. (2018). Scarred consumption. Working Paper 24696, National Bureau of Economic Research.
- Matějka, F. and McKay, A. (2015). Rational inattention to discrete choices: A new founda-tion for the multinomial logit model. *American Economic Review*, 105(1):272–98.
- Mullainathan, S. (2013). Scarcity: Why having too little means so much.
- Nevo, A. and Wong, A. (2019). The elasticity of substitution between time and market goods: Evidence from the great recession. *International Economic Review*, 60(1):25–51.
- Ritchie, S. J. and Bates, T. C. (2013). Enduring links from childhood mathematics and reading achievement to adult socioeconomic status. *Psychological science*, 24(7):1301–1308.
- Severen, C. and Van Benthem, A. A. (2022). Formative experiences and the price of gasoline. *American Economic Journal: Applied Economics*, 14(2):256–84.
- Skimmyhorn, W. (2016). Assessing financial education: Evidence from boot camp. *American Economic Journal: Economic Policy*, 8(2):322–43.
- Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199. Themed

Issue: Treatment Effect 1.

- Suri, R., Monroe, K. B., and Koc, U. (2013). Math anxiety and its effects on consumers' preference for price promotion formats. *Journal of the Academy of Marketing Science*, 41(3):271–282.
- Venkatesan, R. and Farris, P. W. (2012). Measuring and managing returns from retailer-customized coupon campaigns. *Journal of marketing*, 76(1):76–94.
- Viswanathan, M., Rosa, J. A., and Harris, J. E. (2005). Decision making and coping of functionally illiterate consumers and some implications for marketing management. *Journal of Marketing*, 69(1):15–31.
- Xu, Y. (2017). Generalized synthetic control method: Causal inference with interactive fixed effects models. *Political Analysis*, 25(1):57–76.
- Zacharopoulos, G., Sella, F., and Cohen Kadosh, R. (2021). The impact of a lack of mathematical education on brain development and future attainment. *Proceedings of the National Academy of Sciences*, 118(24):e2013155118.

## 9 Appendix

### 9.1 Additional Illustrations

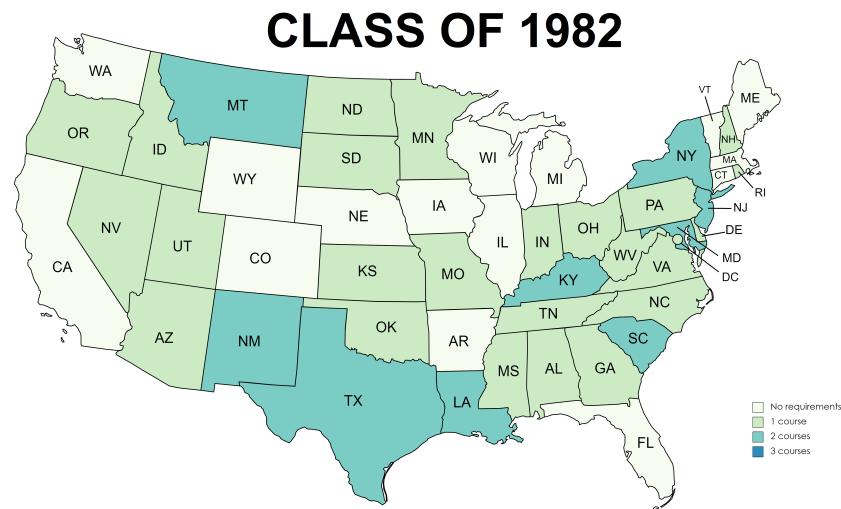


Figure 8: Math Requirements of Class 1982

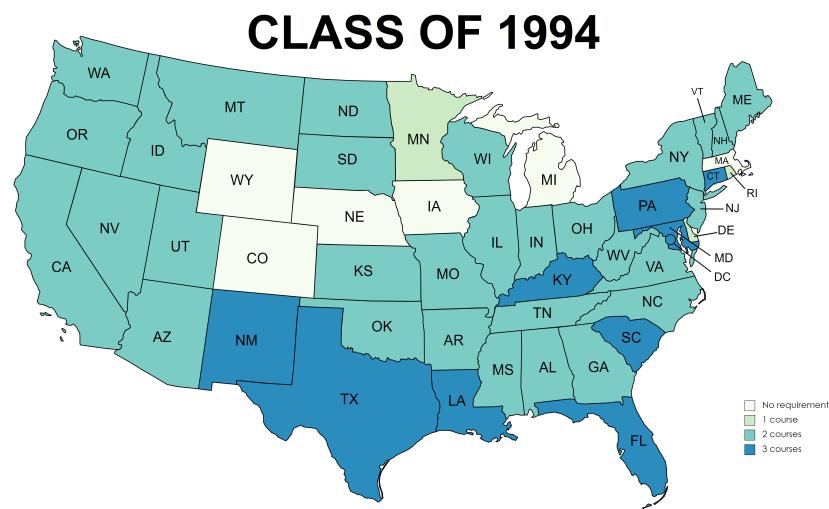


Figure 9: Math Requirements of Class 1994

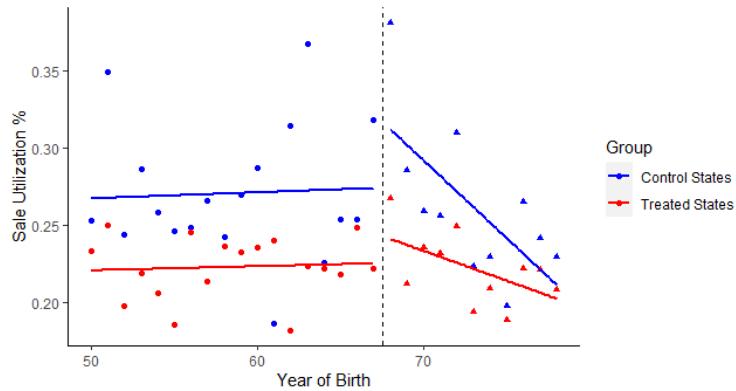


Figure 10: Model Free Evidence: Sale

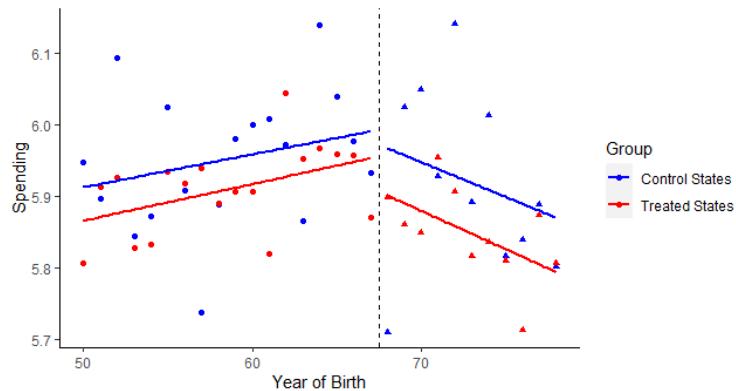


Figure 11: Model Free Evidence: Spending

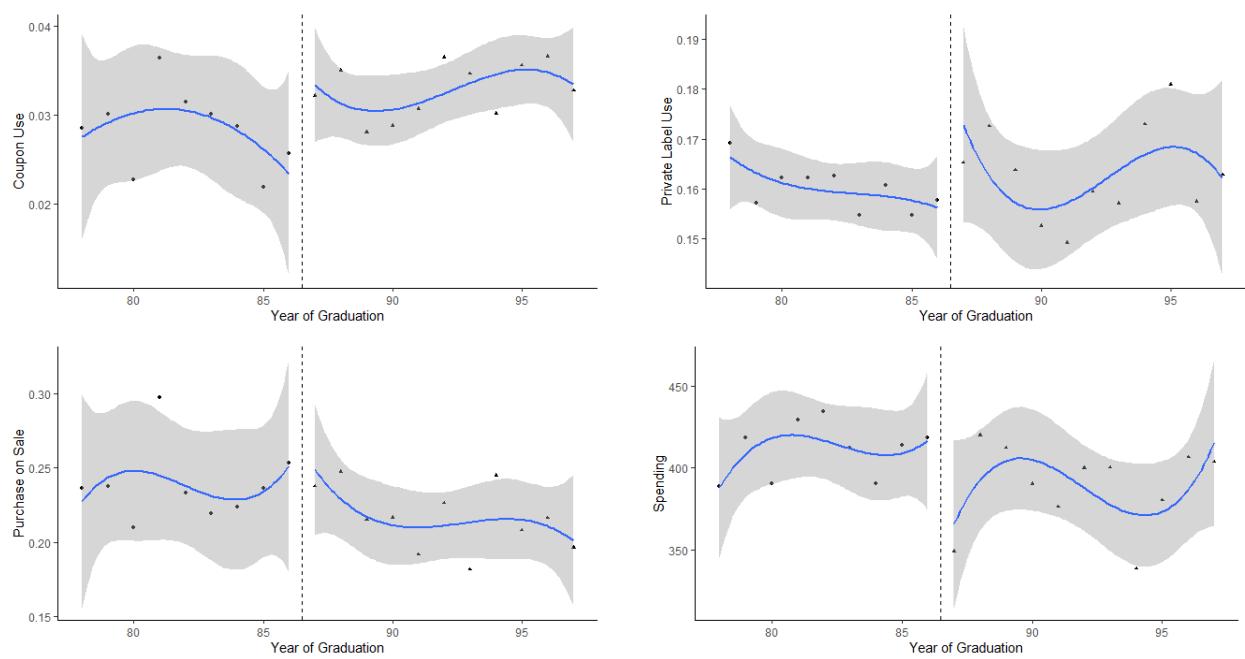


Figure 12: Model Free Evidence: Discontinuity in Time - Treatment Cohort 1987. Clockwise: Coupon Use; Private Label Share; % Purchase-on-Sale; Log Expenditure.

## 9.2 Additional Results

Dependent Variables: Model:	% Sale (1)	Coupon (2)	Private Label (3)	log(Spending) (4)
Treat × Post	0.0092 (0.0070)	0.0067*** (0.0014)	0.0088*** (0.0031)	-0.0086 (0.0217)
State linear trends	Yes	Yes	Yes	Yes
State-Year FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	17,478	17,478	17,478	17,478

Clustered Robust Standard-Errors in parentheses  
Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 6: Aggregated State Level Fixed Effects Model With State-specific Linear Trends

Dependent Variables: Model:	% Sale (1)	Coupon (2)	Private Label (3)	log(Spending) (4)
<i>Treated Cohort:</i>				
1984	-0.0684 (0.0587)	0.0201*** (0.0051)	0.0270** (0.0115)	0.2568* (0.1523)
1985	0.0654*** (0.0202)	0.0179*** (0.0037)	0.0006 (0.0188)	-0.0413 (0.0875)
1986	-0.1668*** (0.0510)	-0.0311*** (0.0113)	0.0530*** (0.0086)	-0.0809 (0.0839)
1987	-0.0155 (0.0223)	0.0035 (0.0054)	-0.0029 (0.0057)	-0.0287 (0.0682)
1988	0.0779*** (0.0203)	0.0083*** (0.0029)	-0.0067 (0.0069)	-0.3792*** (0.0619)
1989	0.0136 (0.0172)	0.0108** (0.0043)	0.0127* (0.0077)	-0.1453*** (0.0443)
1990	0.0513*** (0.0173)	0.0103* (0.0058)	-0.0599** (0.0263)	0.4050*** (0.1412)
State-Year FEs	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Observations	17,410	17,410	17,410	17,410

Clustered Robust Standard Errors in parentheses

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 7: Sun and Abraham (2021)'s ATT-by-cohort Aggregated Results

### 9.3 Treatment Heterogeneity

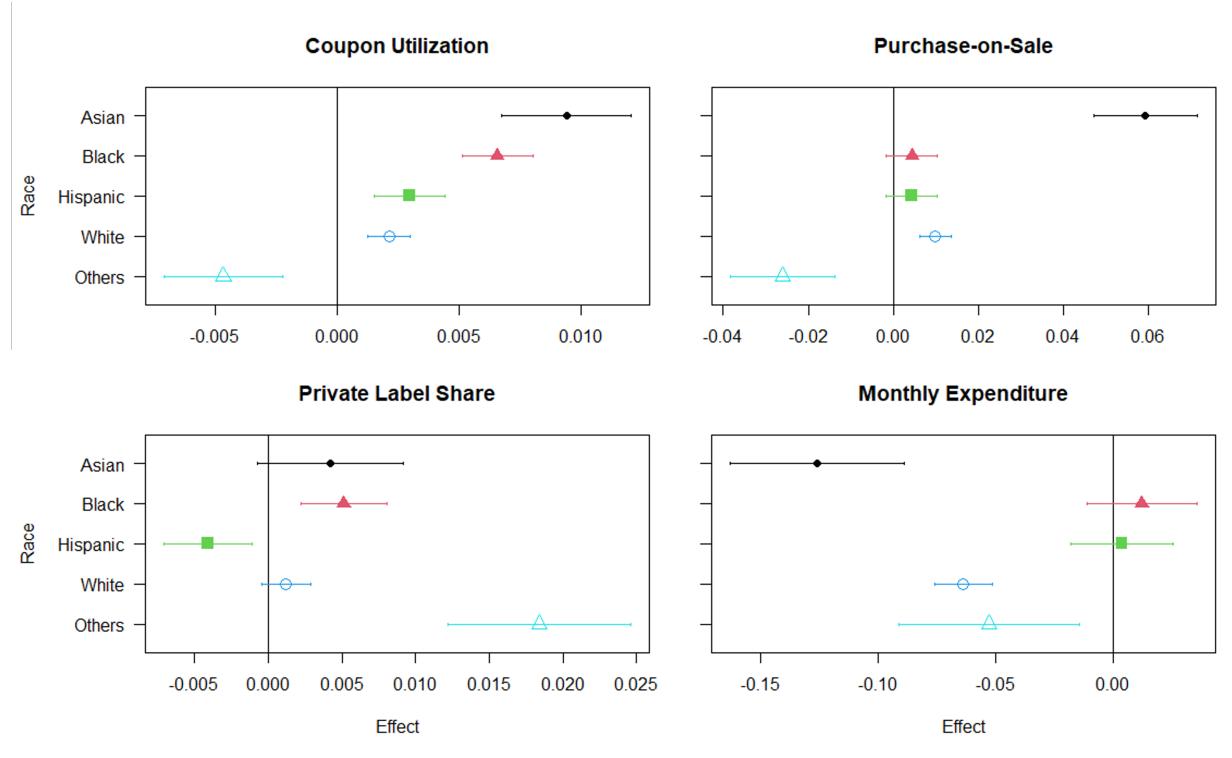


Figure 13: Treatment Effect Heterogeneity by Race

Dependent Variables: Model:	% Sale (1)	Coupon (2)	Private Label (3)	log(Spending) (4)
Treat × Post × White	0.0098*** (0.0019)	0.0021*** (0.0004)	0.0012 (0.0008)	-0.0638*** (0.0062)
Treat × Post × Black	0.0044 (0.0031)	0.0066*** (0.0007)	0.0051*** (0.0015)	0.0120 (0.0119)
Treat × Post × Asian	0.0592*** (0.0062)	0.0094*** (0.0014)	0.0042* (0.0025)	-0.1258*** (0.0189)
Treat × Post × Hispanic	0.0043 (0.0031)	0.0030*** (0.0007)	-0.0041*** (0.0015)	0.0035 (0.0111)
Treat × Post × Other	-0.0261*** (0.0063)	-0.0047*** (0.0012)	0.0184*** (0.0032)	-0.0528*** (0.0196)
Control Variables	Yes	Yes	Yes	Yes
State-Year FEs	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Month-Year FEs	Yes	Yes	Yes	Yes
Observations	1,468,844	1,468,850	1,468,844	1,468,844

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 8: Treatment Heterogeneity by Race

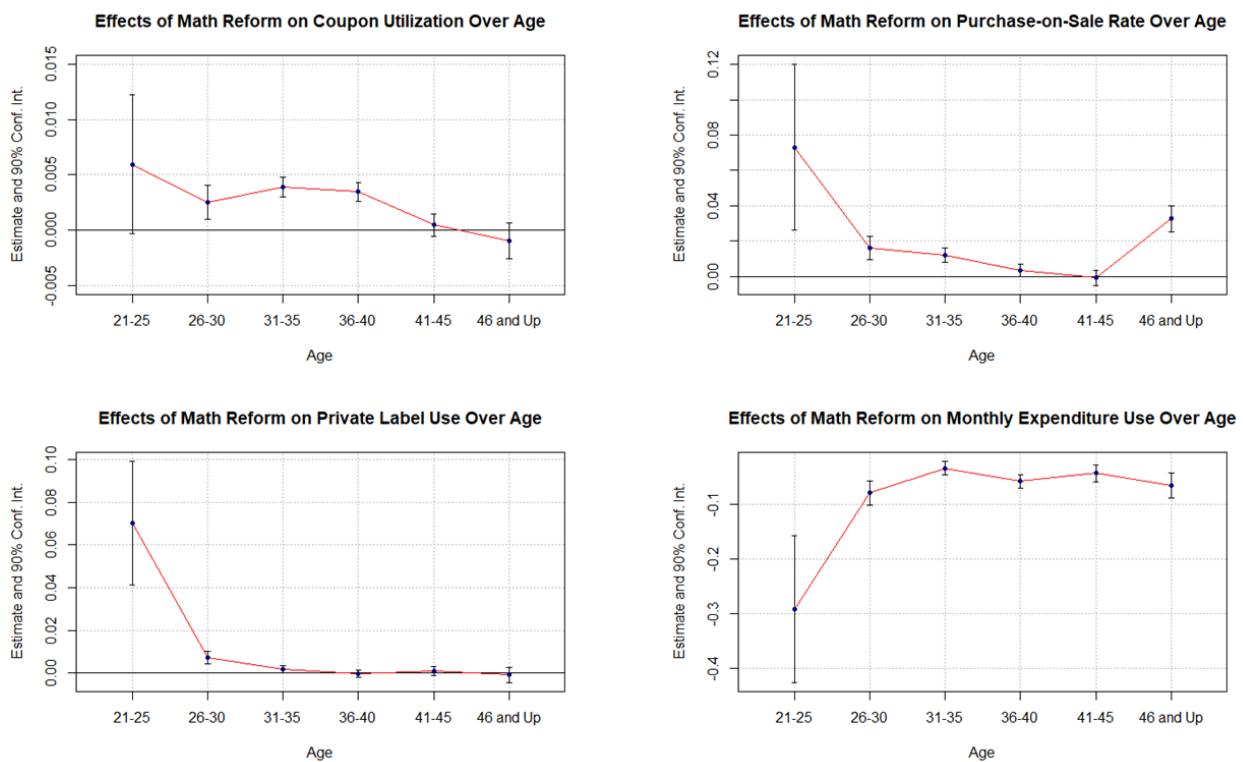


Figure 14: Treatment Effect Heterogeneity by Age

Dependent Variables: Model:	% Sale (1)	Coupon (2)	Private Label (3)	log(Spending) (4)
Treat $\times$ Post $\times$ 21-25	0.0731** (0.0285)	0.0060 (0.0038)	0.0703*** (0.0175)	-0.2913*** (0.0817)
Treat $\times$ Post $\times$ 26-30	0.0161*** (0.0040)	0.0025*** (0.0009)	0.0074*** (0.0018)	-0.0786*** (0.0132)
Treat $\times$ Post $\times$ 31-35	0.0121*** (0.0024)	0.0039*** (0.0005)	0.0020* (0.0010)	-0.0332*** (0.0076)
Treat $\times$ Post $\times$ 36-40	0.0035 (0.0022)	0.0035*** (0.0005)	-0.0004 (0.0010)	-0.0568*** (0.0074)
Treat $\times$ Post $\times$ 41-45	-0.0008 (0.0026)	0.0005 (0.0006)	0.0010 (0.0013)	-0.0425*** (0.0094)
Treat $\times$ Post $\times$ $\geq 46$	0.0326***	-0.0010	-0.0008	-0.0648***
Control Variables	Yes	Yes	Yes	Yes
State-Year FEs	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Month-Year FEs	Yes	Yes	Yes	Yes
Observations	1,468,844	1,468,850	1,468,844	1,468,844

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 9: Treatment Heterogeneity by Age

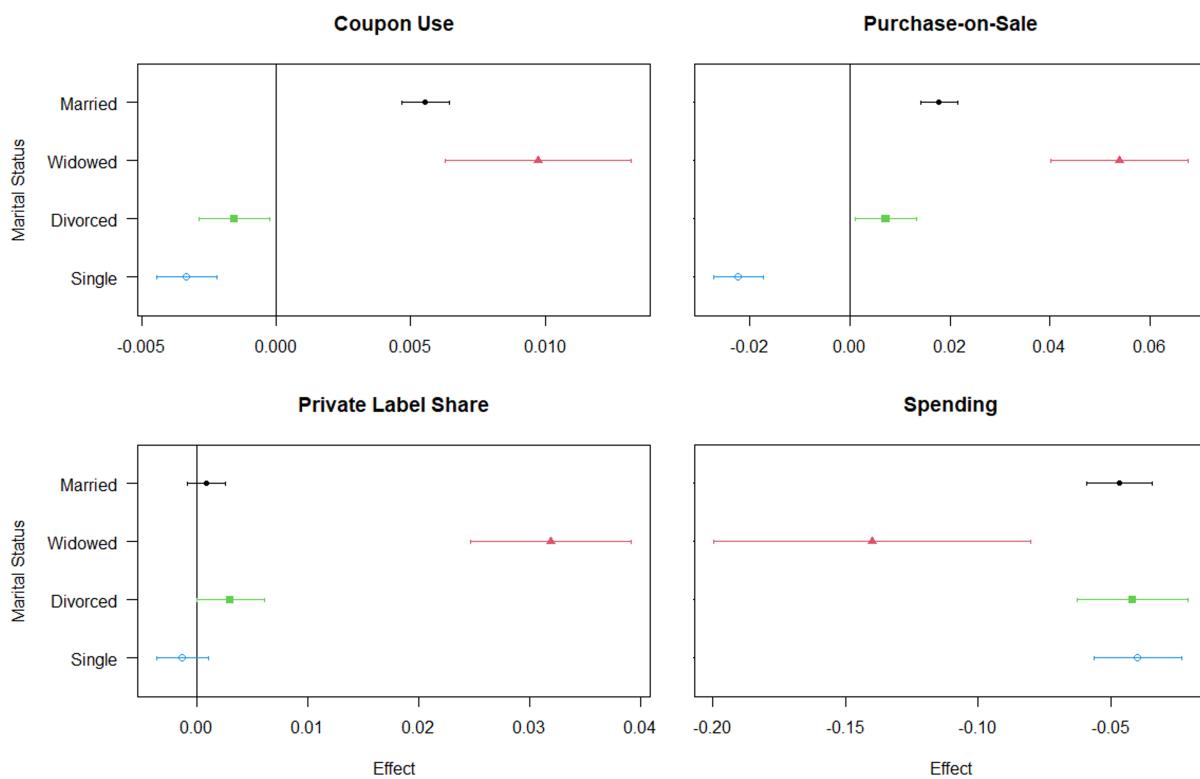


Figure 15: Treatment Effect Heterogeneity by Marital Status

Dependent Variables: Model:	% Sale (1)	Coupon (2)	Private Label (3)	log(Spending) (4)
Treat $\times$ Post $\times$ Married	0.0178*** (0.0019)	0.0055*** (0.0004)	0.0009 (0.0009)	-0.0467*** (0.0063)
Treat $\times$ Post $\times$ Widowed	0.0539*** (0.0070)	0.0098*** (0.0018)	0.0319*** (0.0037)	-0.1399*** (0.0305)
Treat $\times$ Post $\times$ Divorced	0.0072** (0.0031)	-0.0016** (0.0007)	0.0030* (0.0016)	-0.0417*** (0.0106)
Treat $\times$ Post $\times$ Single	-0.0223*** (0.0025)	-0.0033*** (0.0006)	-0.0013 (0.0012)	-0.0399*** (0.0085)
Control Variables	Yes	Yes	Yes	Yes
State-Year FEs	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Month-Year FEs	Yes	Yes	Yes	Yes
Observations	1,468,844	1,468,850	1,468,844	1,468,844

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 10: Treatment Heterogeneity by Marital Status

## 9.4 Robustness Check

### 9.4.1 Matched Sample

Dependent Variables: Model:	% Sale (1)	Coupon (2)	Private Label (3)	log(Spending) (4)
ATT	0.0296** (0.0128)	0.0131*** (0.0021)	0.0101* (0.0056)	-0.1681*** (0.0369)
State-Year FEs	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Observations	8,543	8,543	8,543	8,543

Clustered Robust Standard Errors in parentheses

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 11: Robustness Check: ATT of State Matched Sample

Dependent Variables: Model:	% Sale (1)	Coupon (2)	Private Label (3)	log(Spending) (4)
Treat × Post	0.0086*** (0.0016)	0.0049*** (0.0004)	0.0005 (0.0007)	-0.0736*** (0.0050)
Control Variables	Yes	Yes	Yes	Yes
State-Year FEs	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Month-Year FEs	Yes	Yes	Yes	Yes
Observations	865,321	865,316	865,316	865,316

Clustered robust standard-errors in parentheses

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 12: Robustness Check: Matched States Sample

#### 9.4.2 Restricted Sample Cohorts Born between 1960 and 1975

Dependent Variables: Model:	% Sale (1)	Coupon (2)	Private Label (3)	log(Spending) (4)
ATT	0.0231** (0.0106)	0.0072*** (0.0022)	-0.0003 (0.0039)	-0.1937*** (0.0296)
State-Year FEs	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Observations	9,084	9,084	9,084	9,084

Clustered Robust Standard Errors in parentheses

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 13: Robustness Check: ATT of Restricted Sample of Cohorts Born between 1960 and 1975

Dependent Variables: Model:	% Sale (1)	Coupon (2)	Private Label (3)	log(Spending) (4)
Treat $\times$ Post	0.0084* (0.0044)	0.0032*** (0.0009)	0.0002 (0.0018)	-0.0793*** (0.0112)
Control Variables	Yes	Yes	Yes	Yes
State-Year FEes	Yes	Yes	Yes	Yes
State FEes	Yes	Yes	Yes	Yes
Cohort FEes	Yes	Yes	Yes	Yes
Month-Year FEes	Yes	Yes	Yes	Yes
Observations	676,643	676,639	676,639	676,639

Clustered robust standard-errors in parentheses

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 14: Robustness Check: Restricted Sample of Cohorts Born between 1960 and 1975

#### 9.4.3 Restricted Sample of Eventually Treated States Only

Dependent Variables: Model:	% Sale (1)	Coupon (2)	Private Label (3)	log(Spending) (4)
ATT	0.0295* (0.0170)	0.0091** (0.0044)	0.0200*** (0.0074)	-0.4177*** (0.0501)
State-Year FEes	Yes	Yes	Yes	Yes
State FEes	Yes	Yes	Yes	Yes
Cohort FEes	Yes	Yes	Yes	Yes
Year FEes	Yes	Yes	Yes	Yes
Observations	13,673	13,673	13,673	13,673

Clustered Robust Standard Errors in parentheses

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 15: Robustness Check: ATT of Restricted Sample of Eventually Treated States Only (Last Treatment Wave is the baseline)

Dependent Variables: Model:	% Sale (1)	Coupon (2)	Private Label (3)	log(Spending) (4)
Treat $\times$ Post	0.0100*** (0.0021)	0.0019*** (0.0005)	-0.0028*** (0.0010)	-0.0574*** (0.0072)
Control Variables	Yes	Yes	Yes	Yes
State-Year FEs	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Month-Year FEs	Yes	Yes	Yes	Yes
Observations	1,087,745	1,087,741	1,087,741	1,087,741

Clustered robust standard-errors in parentheses

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 16: Robustness Check: Restricted Sample of Eventually Treated States Only

#### 9.4.4 Generalized Synthetic Control

Dependent Variables: Model:	% Sale (1)	Coupon (2)	Private Label (3)	log(Spending) (4)
Generalized Synth	0.0265*** (0.0094)	0.0085*** (0.0016)	-0.0002 ( 0.0022)	-0.0288* (0.0153)
State-Year FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Interactive Factor	1	1	1	1

Block-Bootstrapped Robust Standard Errors in parentheses

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 17: Generalized Synthetic Control Results

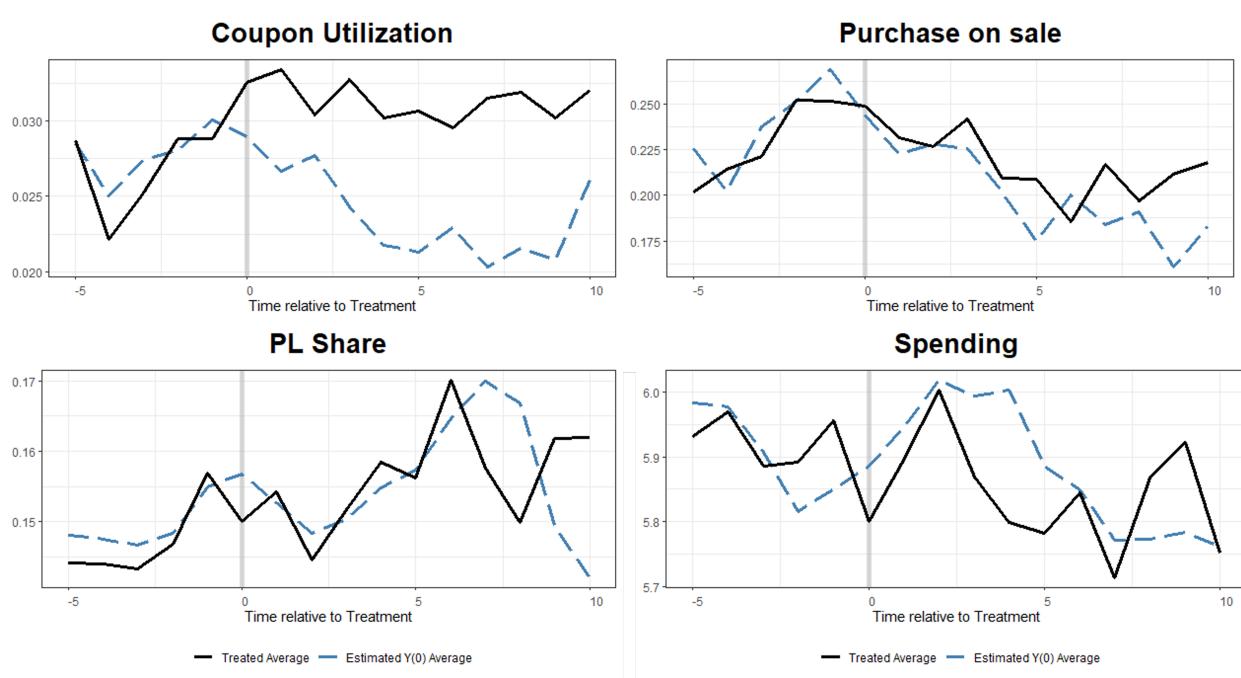


Figure 16: Real vs. Counterfactual of Generalized Synthetic Control Result

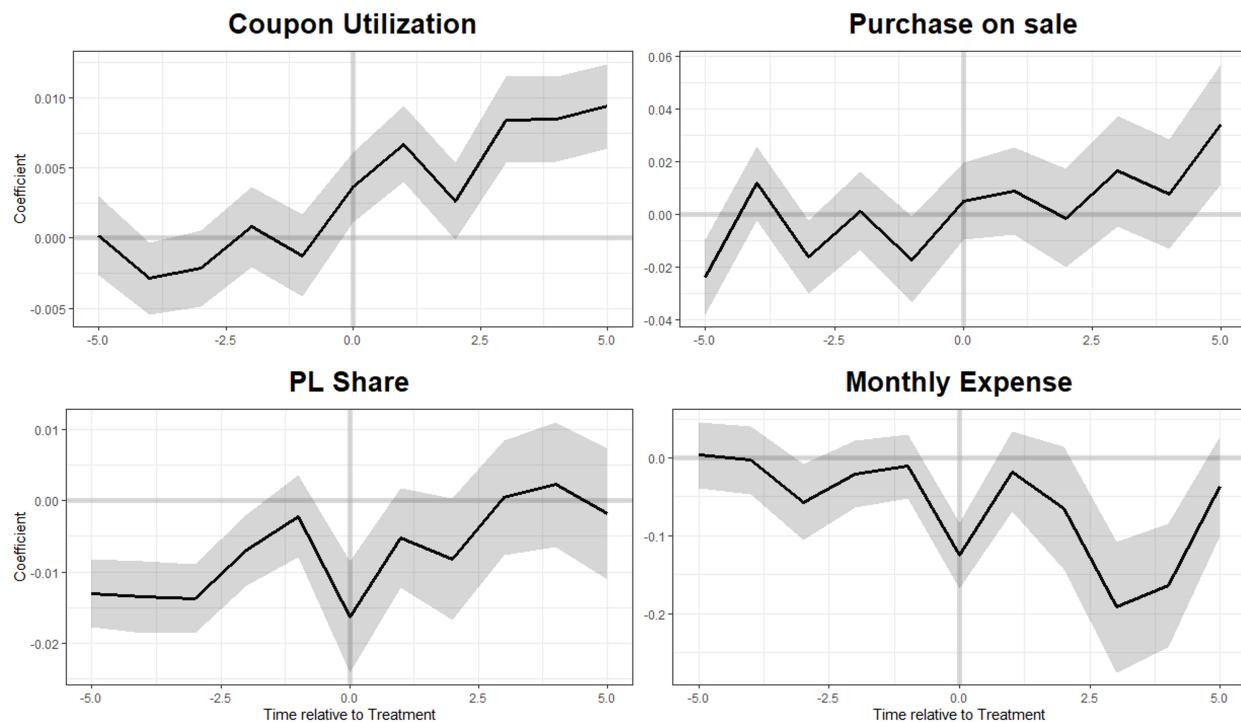


Figure 17: Gap Plot of Generalized Synthetic Control Result

## 9.5 Placebo Test

Dependent Variables: Model:	% Sale (1)	Coupon (2)	Private Label (3)	log(Spending) (4)
ATT	-0.0226* (0.0119)	-0.0025 (0.0024)	0.0028 (0.0043)	0.0428 (0.0327)
State-Year FEs	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Observations	17,410	17,410	17,410	17,410

Clustered Robust Standard Errors in parentheses  
Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 18: Placebo Test: Moving-forward (5 cohorts) Placebo Treatment

We perform two placebo tests for the [Sun and Abraham \(2021\)](#) event study estimations.

In Table 18, we create new treatment indicators to denote a “Placebo” treatment that begins 5 relative periods in the future in comparison to the real treatments. For example, in states where the real first treated cohort is Class of 1985, the first treated cohort would be Class of 1990 for the placebo treatment. Specifically, the first step regression equation is now:

$$Y_{scy} = \sum_g \sum_{k \in \{-T, T\}, k \neq 4} \delta_{gk} \mathbf{1}\{c - C_s - 5 = k\} \mathbf{1}\{C_s + 5 = g\} + \gamma_s + \gamma_c + \gamma_{sy} + \varepsilon_{scy}$$

This placebo specification tests whether the observed results are truly the effects of the curricular reforms, or from another coincidental shock that happens after the curricular reforms. Additionally, this placebo procedure also provides evidence on whether the treatment effects of different treatment waves are on the same “treatment path” (i.e. if their post-treatment dynamic effects are roughly parallel). If this is satisfied, the analytical inference of the ATT is robust, otherwise, [Callaway and Sant'Anna \(2021\)](#) suggests that “simultaneous inference” may be necessary. The null results we get in Table 18 indicate

that there are no coincidental shocks after the reform that may contaminate our estimates, and further there is no evidence of diverging or converging treatment paths.

Similarly, in Table 19, we perform a different placebo test by shifting the treatment five relative periods into the past (instead of the future). This also tests for pre-treatment parallel trends. The null results once again show that the parallel trends assumption is reasonable, and there is no pre-treatment divergence or convergence in trends.

Dependent Variables:	% Sale	Coupon	Private Label	log(Spending)
Model:	(1)	(2)	(3)	(4)
ATT	0.0314 (24.08)	0.0045 (5.233)	-0.0143 (14.07)	-0.0586 (85.93)
State-Year FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	13,357	13,357	13,357	13,357

Clustered Robust Standard Errors in parentheses

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 19: Placebo Test: Moving-backward (5 cohorts) Placebo Treatment

## 9.6 Pre vs. Post 2008

Dependent Variables: Model:	% Sale (1)	Coupon (2)	Private Label (3)	log(Spending) (4)
ATT (Pre-2008)	-0.0129 (0.0179)	-0.0013 (0.0036)	-0.0066 (0.0068)	-0.1667*** (0.0505)
ATT (Post-2008)	0.0432*** (0.0122)	0.0119*** (0.0026)	0.0052 (0.0044)	-0.1658*** (0.0376)
State-Year FEs	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Observations	17,410	17,410	17,410	17,410

Clustered Robust Standard Errors in parentheses

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 20: ATT Triple Differences Model Pre vs. Post 2008

Dependent Variables: Model:	% Sale (1)	Coupon (2)	Private Label (3)	log(Spending) (4)
Treat × Post	-0.0040 (0.0064)	0.0042*** (0.0012)	0.0039 (0.0027)	-0.0322* (0.0190)
Treat × Post × Post 2008	0.0233*** (0.0065)	0.0046*** (0.0012)	-0.0037 (0.0024)	0.0345** (0.0171)
State-Year FEs	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Observations	17,478	17,478	17,478	17,478

Clustered Robust Standard-Errors in parentheses

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 21: State Level Triple Differences Model Pre vs. Post 2008

Dependent Variables: Model:	% Sale (1)	Coupon (2)	Private Label (3)	log(Spending) (4)
Treated $\times$ Post	0.0015 (0.0055)	0.0027** (0.0011)	-0.0009 (0.0021)	-0.0636*** (0.0138)
Treated $\times$ Post $\times$ Post 2008	0.0142** (0.0055)	0.0041*** (0.0010)	-0.0031* (0.0018)	0.0162 (0.0120)
Control Variables	Yes	Yes	Yes	Yes
State-Year FEs	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Month-Year FEs	Yes	Yes	Yes	Yes
Observations	1,468,850	1,468,844	1,468,844	1,468,844

Clustered Robust Standard-Errors in parentheses

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 22: Household Level Triple Differences Model Pre vs. Post 2008

## 9.7 Data With Movers

Dependent Variables: Model:	% Sale (1)	Coupon (2)	Private Label (3)	log(Spending) (4)
Treated $\times$ Post	0.0040 (0.0036)	0.0066*** (0.0007)	-0.0014 (0.0014)	-0.0284*** (0.0086)
Control Variables	Yes	Yes	Yes	Yes
Current State-Year FEs	Yes	Yes	Yes	Yes
High School State FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Month-Year FEs	Yes	Yes	Yes	Yes
Observations	1,947,052	1,947,046	1,947,046	1,947,006

Clustered (Two-Way Current  $\times$  HS State) Robust Standard-Errors in parentheses

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 23: Household Level Results of Data With Movers

Dependent Variables: Model:	% Sale (1)	Coupon (2)	Private Label (3)	log(Spending) (4)
<b>A. Naive DiD</b>				
Treat × Post	0.0095** (0.0048)	0.0073*** (0.0009)	0.0041* (0.0022)	-0.0158 (0.0146)
<b>B. Sun and Abraham (2021)'s Aggregated ATT</b>				
ATT	0.0089 (0.0097)	0.0069*** (0.0019)	-0.0005 (0.0035)	-0.1819*** (0.0278)
State-Year FEs	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Observations	18,832	18,832	18,832	18,832

Clustered Robust Standard-Errors in parentheses

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 24: State Level Results of Data With Movers

Dependent Variables: Model:	% Sale (1)	Coupon (2)	Private Label (3)	log(Spending) (4)
Treat × Post	0.0014 (0.0062)	0.0055*** (0.0012)	0.0062** (0.0027)	-0.0403** (0.0189)
Treat × Post × Post2008	0.0118** (0.0059)	0.0027** (0.0012)	-0.0031 (0.0024)	0.0356** (0.0165)
State-Year FEs				
State FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Observations	18,832	18,832	18,832	18,832

Clustered Robust Standard-Errors in parentheses

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 25: State Level Results With Movers - Pre/Post 2008

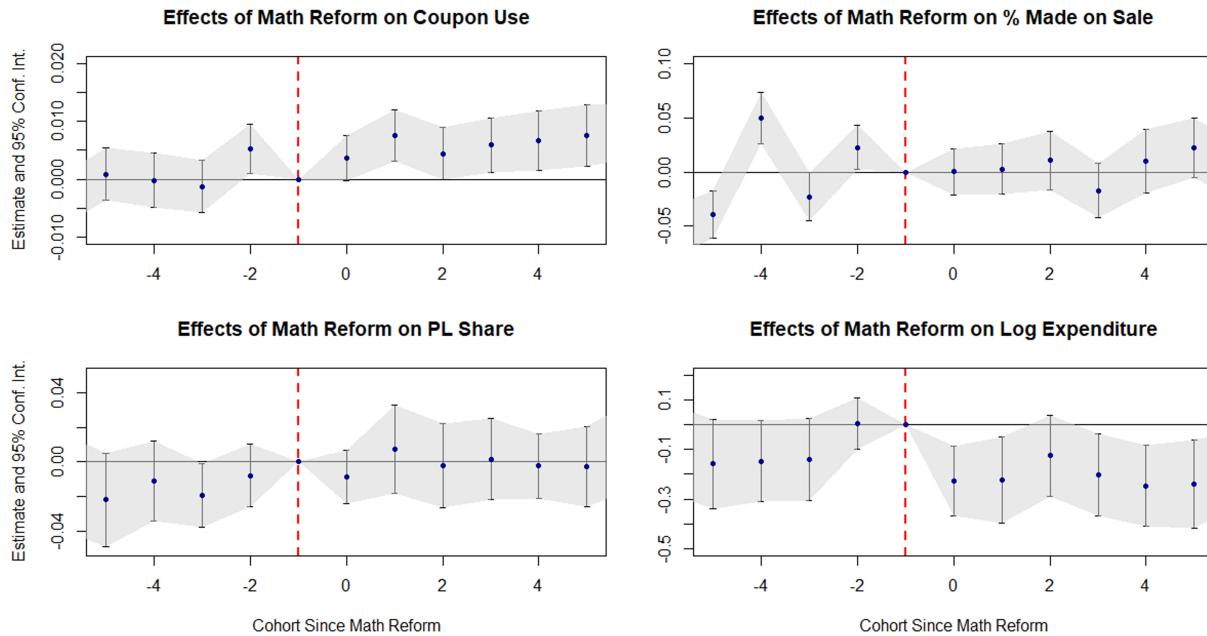


Figure 18: Event Study Results

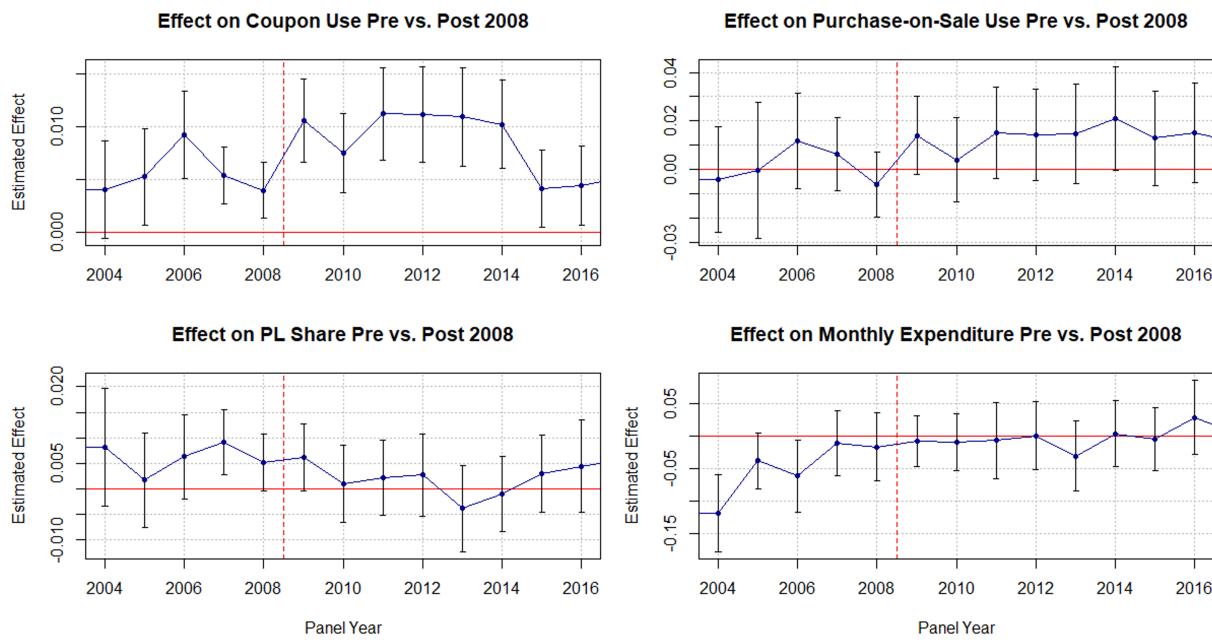


Figure 19: Pre- vs. Post 2008 Estimates

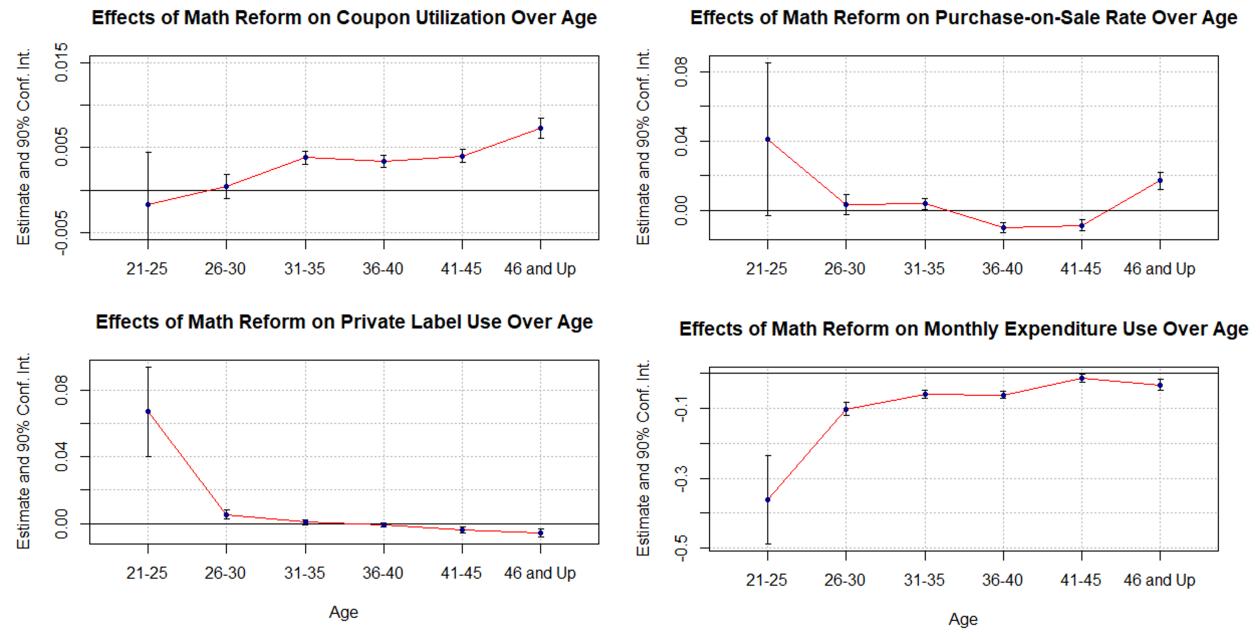


Figure 20: Effects by Age Group

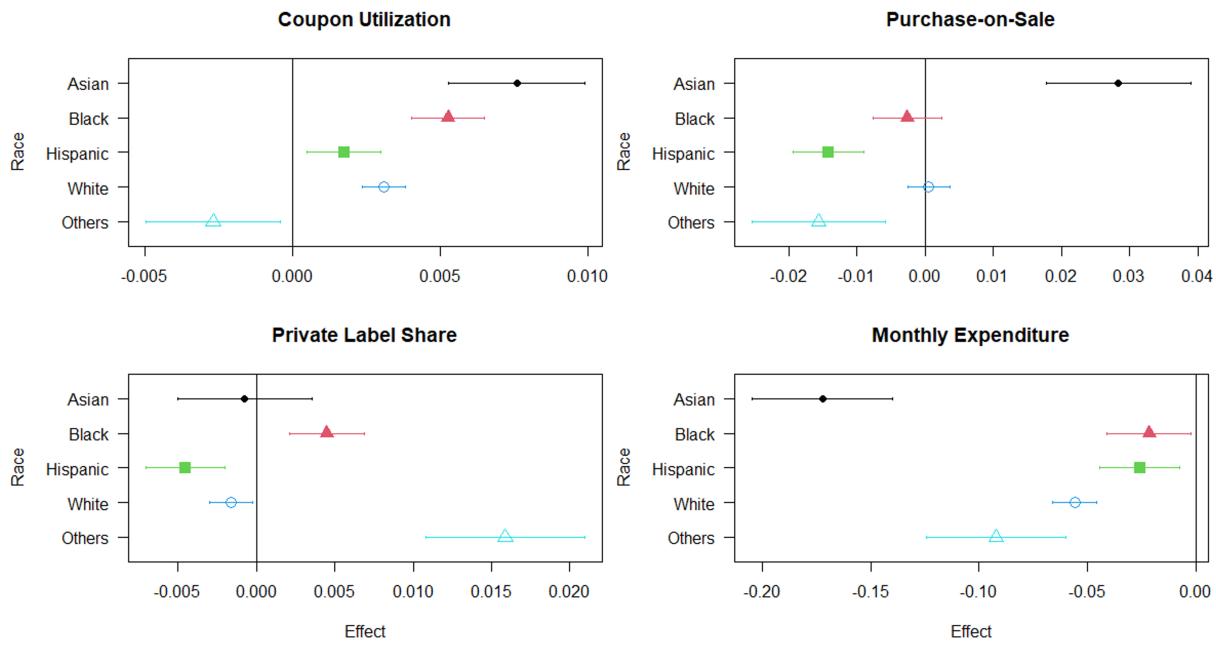


Figure 21: Effects by Race

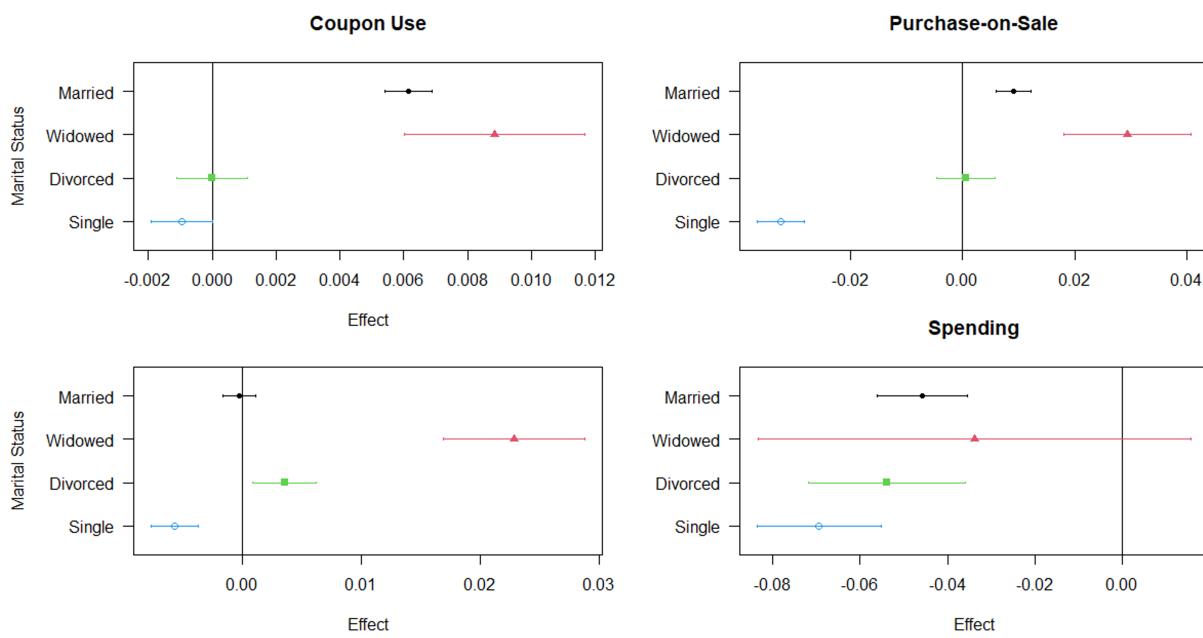


Figure 22: Effects by Marital Status