# Incentivizing Health: A Theoretical and Empirical Evaluation of the Effect of Wellness Policy on Health Outcomes

Yunxi Yang
Supervisor: Jakob Svensson
Stockholm University

Course Code: EC9901
Spring 2025

### Abstract

This thesis explores the effectiveness of wellness policies in promoting healthy behaviors by combining a theoretical model with empirical evidence from a U.S. wellness initiative: the Comprehensive School Physical Activity Program (CSPAP), which is designed to improve adolescent health and combat obesity. The model predicts that only individuals with intermediate preferences are likely to change their behavior in response to policy interventions when the primary mechanism is increased accessibility to wellness activities (e.g., physical activity). It also suggests that policies may have a greater impact on health outcomes when accessibility is combined with educational interventions. Empirically, the results reveal a statistically significant two-year-lagged reduction in obesity rates, particularly in states with higher adoption intensity. This finding is consistent with the model's prediction: pairing accessibility with education can amplify the policy's effect on health outcomes. The effect is estimated to have lowered national obesity among high school students by up to 2.7 percentage points in 2022. These findings draw attention to the importance of program intensity and sustained implementation in shaping public health outcomes.

# Acknowledgments

I would like to express my warmest thanks to my supervisor, Jakob Svensson, for his generosity with time and his patient guidance throughout the process of writing this thesis. I am also deeply grateful for the insightful discussions and encouragement from my dear friends Gustav Olsson, Quan Cheng Xie, Lukas Ma, and Matteo Pianella.

# Contents

# 1 Introduction

This thesis investigates the behavioral and structural obstacles that limit the effectiveness of public health and welfare policies aimed at promoting healthier behaviors. It seeks to find some of the answers to why wellness policies don't always lead to meaningful behavioral change. Specifically, using both a theoretical model and empirical analysis, I examine wellness policies that aim to increase access to physical activity and assess their impact on health outcomes. I begin by developing a structural behavioral model in which individuals with heterogeneous preferences face both internal (e.g., motivation) and external (e.g., access) barriers to physical activity. The model suggests that only individuals with intermediate preferences are likely to change their behavior in response to improved accessibility. It also offers insights into the potential for wellness policies to become more effective when educational components are integrated into their design. Empirically, I evaluate the Comprehensive School Physical Activity Program (CSPAP), a wellness policy introduced in 2013 to promote physical activity in U.S. schools. Using a dose-response panel model, I estimate the marginal effect of CSPAP adoption on adolescent obesity rates and validate the findings with the Callaway and Sant'Anna (2021) estimator.

In order to promote healthy behaviors, various programs have been developed to improve access to health-related activities. Examples include the Supplemental Nutrition Assistance Program (SNAP) in the United States, which is designed to promote healthier eating; and wellness allowances, such as Sweden's Friskvårdsbidrag, which aim to encourage participation in physical activity.

Despite the effort, financial subsidies have shown moderate effectiveness in promoting desired behaviors (Feng et al., 2024; Mancino and Guthrie, 2014; Holden, 2020). For instance, children from SNAP-eligible households face an increased risk of diet-related diseases compared to income-eligible nonparticipating households, primarily due to higher consumption of sugar-sweetened beverages (SSBs), high-fat dairy, and processed meats (Mande and Flaherty, 2023). As for alcohol and cigarette consumption, studies find no significant difference in both alcohol intake and cigarette consumption between SNAP participants and SNAP-eligible nonparticipants, suggesting that eligibility restrictions alone are not the primary factor influencing consumption behavior (Todd and Ver Ploeg, 2015).

Regarding wellness benefits, studies have found no significant effect of financial incentives on health outcomes, particularly in terms of reducing healthcare costs (Einav et al., 2019). Additionally, another study reports high attrition rates among participants and only a moderate impact on weight loss associated with financial incentives (Cawley and Price, 2013). According to a Sifo-study, as of 2023, 35 years after the wellness policy Frskvårdsbidrag was introduced in 1988, 16% of workers offered wellness allowance still choose to forgo this subsidy completely, and roughly 41% of those that use it don't redeem the full eligible amount (Epassi, 2023).

One possible explanation for the moderate effect of financial incentives in promoting preferred behavior is that they often come with restrictions on how the financial benefits can be used. For instance, SNAP promotes healthier behaviors by restricting the purchase of alcohol and cigarettes using the SNAP coupons, the Swedish wellness allowance ensures that funds can only be used for health-related activities. Thus, the financial incentives may fall short of their intended impact, as households can often achieve their preferred consumption bundle by strategically reallocating the subsidy alongside their own cash, especially when the incentive only covers a small fraction of total household spending (Schanzenbach, 2017).

If we understand financial incentives as mechanisms that improve access to a preferred behavior, then the limited effectiveness of such policies may stem from the heterogeneity in individual preferences, beliefs, and valuations that shape the responsiveness of individual behaviors. The remainder of this thesis seeks to unpack the underlying obstacles to behavioral change in response to wellness policies, such as those aimed at expanding public access to physical activity, and to explore how such policies may achieve greater impact when complemented by educational interventions. This prediction is supported by the empirical analysis, which finds that the implementation of CSPAP, a program that enhances both access to and education about physical activity, led to an estimated 2.7 percentage point decline in obesity rates in 2022. The analysis contributes to the broader discussion on the design and effectiveness of wellness policies.

The structure of the thesis is as follows, first, it provides background on the policy under analysis: the Comprehensive School Physical Activity Program (CSPAP), which is followed by a theoretical model, an empirical strategy and corresponding results, and finally, a discussion of policy implications and conclusions.

# 2 CSPAP

The Comprehensive School Physical Activity Program (CSPAP) was introduced in the United States by the Centers for Disease Control and Prevention (CDC) in 2013 to promote physical activity among students through a multi-component framework.[1] The program was designed to combat rising obesity rates among adolescents and includes five key components: high-quality physical education, physical activity during school, physical activity before and after school, staff involvement, and family and community engagement (Centers for Disease Control and Prevention, 2015).

The goal of CSPAP according to Centers for Disease Control and Prevention (2015) is

---

[1] https://www.cdc.gov/healthyschools/professional_development/e-learning/CSPAP/index.html

two fold: one is to encourage all students to engage in 60 minutes of moderate-to-vigorous physical activity everyday, by increasing accessibility to physical activity before, during and after school; the other is to equip them with the knowledge, skills, and confidence to maintain physically active lifestyles throughout life, achieved by improving the quality of physical education at school. To support implementation, the Centers for Disease Control and Prevention (2015) provides schools with planning guides, assessment and evaluation tools, as well as technical assistance.

In order to encourage and guide schools to provide accessible facilities and infrastructure for physical activities through CSPAP, Centers for Disease Control and Prevention (2019) emphasizes that parents, school staff, and community members should collaborate to expand opportunities for physical activity by establishing agreements with schools, enabling students and community members to utilize school facilities for physical activity programs or events outside of regular school hours. Furthermore, sports are organized both during and after school to increase accessibility to physical activity(Centers for Disease Control and Prevention, 2019).

To help students form better views and beliefs about the importance of physical activity, CSPAP promotes standards-based curriculum to help students build knowledge and skills related to physical activity, fitness, and movement (Centers for Disease Control and Prevention, 2019).

The program was initiated as a federal program, but states were free to choose whether to implement it. Further, within states, the decision to implement the program was delegated to school district boards and in the end schools. The degree of implementation, including starting year, varied greatly across states and year. But overall, the rate of implementation was relatively low the first years (2014-2018) but increased significantly after 2020.

# 3 Model

## 3.1 Model Setup

To understand how wellness policies affect physical activity, it is important to recognize that individuals face trade-offs when deciding the optimal amount of exercise per week: physical activities yield health returns but also involves costs. Specifically, costs of physical activity is generated from two types of barriers: internal and external.

External barriers refer to practical constraints, such as cost, time, access to facilities, or lack of school infrastructure. The higher the external barrier, the less accessible physical activities are, and the less convenient for people to acquire them. A unit external barrier parameter, denoted by $\lambda$ (dollars/hour) is introduced to capture the cost of exercise per hour/week induced by external barriers.

Internal barriers include both physical constraints and psychological or emotional perceptions of exercise. Physical constraint refers to the body's tolerance for exertion — for instance, exercising one hour per week may be manageable for most people, whereas 20 hours per week can be physically taxing. To represent physical constraints formally, I introduce a unit effort cost $\epsilon$ (measured in dollars/hour$^2$) along with a standard convex effort cost term $\frac{1}{2}\epsilon H_i^2$ to capture increasing marginal difficulty as hours of exercise increase, where $H_i$ is hours of exercise per week for an individual.

Psychological and emotional perceptions, on the other hand, relate to how individuals interpret the benefits of exercise. These perceptions are shaped by personal beliefs: while some believe that physical activity significantly improves health, others may view it as ineffective or even harmful. Moreover, individuals differ in how much they subjectively value health investments relative to other priorities such as work or leisure. While the physiological benefits of exercise are objectively present, heterogeneity in perceived payoff drives differences in behavior. To capture this variation, individuals are modeled as having a personal valuation of physical activity, denoted by $\alpha_i$ (measured in dollars/hour), which reflects the subjective payoff derived from exercising. Note that $\alpha$ is marginal benefit of physical activity per hour. It is assumed that $\alpha$ is continuously and uniformly distributed over the interval $[b, a]$, where $b \in (-\infty, 0]$ and $a \in (0, \infty)$. This assumption allows preference heterogeneity and incorporates negative values of $\alpha$ according to the nature of the preference distribution.

The model specifies that when deciding how many hours to exercise per week, individuals face a payoff function that captures both the benefits and costs of physical activity. The benefit is modeled as the product of an individual's marginal benefit of exercise and the number of hours exercised, $\alpha_i H_i$. The cost side includes two components: an internal effort cost due to physical limitations, represented by a convex term $\frac{1}{2}\epsilon H_i^2$, and an external cost arising from barriers such as time, transportation, or access fees, captured by a linear cost term $\lambda H_i$.

The linear benefit term $\alpha_i H_i$ is chosen to emphasize heterogeneity in how individuals perceive the value of an additional hour of exercise. The convex effort cost term reflects the increasing physical strain or fatigue associated with higher amount of exercise per week. Finally, the external barrier term $\lambda H_i$ accounts for the monetary or logistical costs required to access each hour of exercise, such as commuting to a gym, entry fees, or time lost to travel.

Each individual chooses their optimal weekly exercise level $H_i$ to maximize the following payoff function:

$$\pi_i = \alpha_i H_i - \frac{1}{2}\epsilon H_i^2 - \lambda H_i$$

The optimal exercise level $H_i$ (hours/week) is obtained by taking the derivative of

$\pi_i$ with respect to $H_i$, subject to the constraint that the payoff function is non-negative $\pi_i \geq 0$. The model, inspired by Banerjee et al. (2002), balances perceived benefits against internal and external costs, and allows us to clarify the mechanisms through which preferences and constraints jointly determine behavior.

It is worth nothing that this payoff function is a reduced form of a general consumer problem with budget constraint. The focus of the model is the individual choice of effort and health investment (how much time to allocate to physical activity), rather than a consumption trade-off between wellness and other goods. This model aims to provide insight into how individuals choose the level of effort they devote to maintaining or improving their health.

### 3.1.1 Two Benchmark Cases

Suppose there are two states of world when it comes to health investment, one realistic, where health investment (e.g., physical activities) have external barriers; and one fictional, where there is no external barriers at all.

In the realistic state of world, an individual gains from exercise depending on her valuation of exercise $\alpha_i H_i$, and her total payoff decreases due to effort cost $\frac{1}{2}\epsilon H_i^2$ as well as external barriers $\lambda H_i$. She gets a payoff function as follows:

$$\pi_i(H_i) = \alpha_i H_i - \frac{1}{2}\epsilon H_i^2 - \lambda H_i$$

She maximizes her payoff by choosing the optimal hours of exercise per week $H_i^*$, subject to the constraint that the total benefit is nonnegative $\pi_i \geq 0$.

Taking the first order condition:

$$\frac{d\pi_i}{dH_i} = \alpha_i - \lambda - \epsilon H_i = 0 \quad \Rightarrow \quad H_i^* = \frac{\alpha_i - \lambda}{\epsilon}$$

Plugging it back to the constraint suggests that $\alpha_i > \lambda$ is required for an individual to choose physical activity at all.

In the fictional case, her payoff function is instead given by:

$$\pi_i(H_i) = \alpha_i H_i - \frac{1}{2}\epsilon H_i^2$$

Taking the first order condition:

$$\frac{d\pi_i}{dH_i} = \alpha_i - \epsilon H_i = 0 \quad \Rightarrow \quad H_i^* = \frac{\alpha_i}{\epsilon}$$

This suggests that participation in physical activity increases once the external barriers are cleared.

Plugging it back to constraint $\pi_i \geq 0$ suggests that an individual will choose to spend time on physical activity that has no external barriers as long as $\alpha_i > 0$.

For individuals with $\alpha \leq 0$, there is no amount of exercise that will make their payoff function positive, therefore, they do not participate in physical activity with or without external barriers.

### 3.1.2 Wellness Policy Intervention

To model the wellness policy intervention explicitly, I modify the individual's payoff function to account for the subsidy $B$ (measured in dollars), which removes the external cost $\lambda$ for the first $\frac{B}{\lambda}$ hours of physical activity per week. Beyond this threshold, the individual must again incur the cost $\lambda$ for additional hours.

We define the cost function of external barriers as:

$$C(H_i) = \begin{cases} 0 & \text{if } H_i \leq \frac{B}{\lambda} \\ \lambda\left(H_i - \frac{B}{\lambda}\right) & \text{if } H_i > \frac{B}{\lambda} \end{cases}$$

Then the individual's total payoff becomes:

$$\pi_i(H_i) = \alpha_i H_i - \frac{1}{2}\epsilon H_i^2 - C(H_i)$$

This formulation captures the kink in the cost structure induced by the policy: for any individual, the first $\frac{B}{\lambda}$ hours are free of external cost, and subsequent hours are subject to the standard barrier $\lambda$. The individual chooses $H_i \geq 0$ to maximize this payoff.

Three cases arise depending on the size of the subsidy: the individual's optimal choice of hours of exercise per week (interior solution) can be larger than, smaller than or equal to the subsidized amount.

Case 1: interior solution below the subsidy cap ($H_i^* < \frac{B}{\lambda}$). In this case, the individual's optimal amount of physical activity falls below the subsidized limit. The external barrier does not bind, and the payoff function is:

$$\pi_i(H_i) = \alpha_i H_i - \frac{1}{2}\epsilon H_i^2$$

Taking the first-order condition:

$$\frac{d\pi_i}{dH_i} = \alpha_i - \epsilon H_i = 0 \quad \Rightarrow \quad H_i^* = \frac{\alpha_i}{\epsilon}$$

This solution is valid only if $H_i^* < \frac{B}{\lambda}$, which implies:

$$\frac{\alpha_i}{\epsilon} < \frac{B}{\lambda} \quad \Rightarrow \quad \alpha_i < \frac{\epsilon B}{\lambda}$$

The subsidized amount is more than her optimal choice in the fictional world without external barrier, since one's optional choice without external barrier is $\frac{\alpha_i}{\epsilon}$.

Case 2: interior solution above the subsidy cap ($H_i^* > \frac{B}{\lambda}$). Here, the optimal choice of exercise level is higher than the amount covered by the policy, and incurs external cost $\lambda$ for each additional hour beyond $\frac{B}{\lambda}$. The payoff function becomes:

$$\pi_i(H_i) = \alpha_i H_i - \frac{1}{2}\epsilon H_i^2 - \lambda(H_i - \frac{B}{\lambda}) = \alpha_i H_i - \frac{1}{2}\epsilon H_i^2 - \lambda H_i + B$$

Taking the first-order condition:

$$\frac{d\pi_i}{dH_i} = \alpha_i - \lambda - \epsilon H_i = 0 \quad \Rightarrow \quad H_i^* = \frac{\alpha_i - \lambda}{\epsilon}$$

This solution is valid only if $H_i^* > \frac{B}{\lambda}$, which implies:

$$\frac{\alpha_i - \lambda}{\epsilon} > \frac{B}{\lambda} \quad \Rightarrow \quad \alpha_i > \frac{\epsilon B}{\lambda} + \lambda$$

The subsidized amount is less than her optimal choice with external barrier, since one's optimal choice with external barrier is $H_i^* = \frac{\alpha_i - \lambda}{\epsilon}$.

Case 3: interior solution at the subsidy cap ($H_i^* = \frac{B}{\lambda}$). For intermediate values of $\alpha_i$, neither of the interior solutions is optimal. Instead, the maximum payoff occurs exactly at the point where the cost function kinks:

$$H_i^* = \frac{B}{\lambda}$$

This occurs when:

$$\alpha_i \in \left[\frac{\epsilon B}{\lambda}, \frac{\epsilon B}{\lambda} + \lambda\right]$$

$$\frac{\epsilon B}{\lambda} \leq \alpha_i \leq \frac{\epsilon B}{\lambda} + \lambda \implies \frac{\alpha_i - \lambda}{\epsilon} \leq \frac{B}{\lambda} \leq \frac{\alpha_i}{\epsilon}$$

The subsidized amount covers more than her optimal choice with external barrier, but not enough to cover her optimal choice without.

This classification follows directly from the first-order conditions of the maximization problem. It also provides a natural basis for partitioning the population in response to the wellness policy.

### 3.1.3 The Distribution of Preference and People's Choice under Wellness Policy

If we look at the distribution of $\alpha$ across the population, shown in figure 1. In the absence of wellness policy, people with $\alpha \in [b, \lambda]$ do not participate in any physical activity; and people with $\alpha > \lambda$ choose to exercise $\frac{\alpha_i - \lambda}{\epsilon}$ hours/week, the higher the $\alpha$, the more hours spent in physical activity.

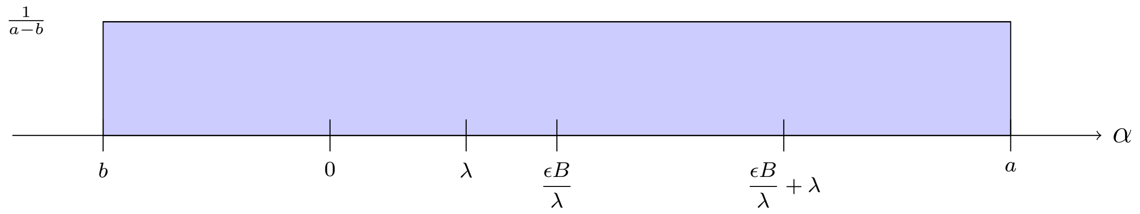


Figure 1: Distributaion of $\alpha$

Consider a wellness policy that provides external-barrier-free physical activity up to $\frac{B}{\lambda}$ hours/week for each individual. Focusing on the positive spectrum of $\alpha$, individuals with $\alpha < \frac{\epsilon B}{\lambda}$, whose interior solution is below what the wellness budget provides (case 1), will not take full advantage of the wellness policy, rather parts of it, up to the amount equivalent to their optimal choice $\frac{\alpha_i}{\epsilon}$. For people in the rest of the positive spectrum $\alpha \geq \frac{\epsilon B}{\lambda}$, they will take full advantage of the wellness policy, fully enjoying the subsided PA facilities $\frac{B}{\lambda}$, because their need for exercise (interior solution) exceeds what the policy covers.

The question is whether people with $\alpha \geq \frac{\epsilon B}{\lambda}$ exercise more than the amount covered by the policy? As discussed in section 3.1.2, optimal choice under wellness policy for individuals with $\alpha \in [\frac{\epsilon B}{\lambda}, \frac{\epsilon B}{\lambda} + \lambda]$ is exactly the same as the subsidy limit $\frac{B}{\lambda}$; and optimal choice for individuals with $\alpha > \frac{\epsilon B}{\lambda} + \lambda$ remains unchanged: $\frac{\alpha_i - \lambda}{\epsilon}$.

To further confirm this result, consider their payoff function when decising whether to exercise additional hours beyond the subsidy limit $\frac{B}{\lambda}$. Since they have already participated in $\frac{B}{\lambda}$ hours/week of external-barrier-free physical activity offered by the wellness policy, let $H_i^{ecs}$ be excess/additional amount of physical activity beyond $\frac{B}{\lambda}$, they now face the following problem:

$$\max_{H_i^{ecs}} \alpha_i(\frac{B}{\lambda} + H_i^{ecs}) - \frac{1}{2}\epsilon(\frac{B}{\lambda} + H_i^{ecs})^2 - \lambda(\frac{B}{\lambda} + H_i^{ecs})$$

Taking the first order condition:

$$H_i^{ecs} = \frac{\alpha_i - \lambda}{\epsilon} - \frac{B}{\lambda}$$

Where $\frac{\alpha_i - \lambda}{\epsilon}$ is the optimal choice of exercise in the absence of wellness budget, and $\frac{B}{\lambda}$ is the amount of physical activity that the budget provides. For an individual to participate in any excess amount of physical activity $H_i^{ecs} > 0$, it requires that $\frac{\alpha_i - \lambda}{\epsilon} > \frac{B}{\lambda}$, that is $\alpha_i > \frac{\epsilon B}{\lambda} + \lambda$. This implies, for people with $\alpha > \frac{\epsilon B}{\lambda} + \lambda$, they will take the full advantage and will exercise additional hours beyond subsidy limit, even if they have to go out of their way to acquire the exercises that are less accessible. However, they will not exercise more than what they would without the policy, since total amount of exercise for these people is $\frac{B}{\lambda} + H_i^{ecs} = \frac{\alpha_i - \lambda}{\epsilon}$. This confirms the result in case 2. Therefore, there is no behavioral change for this group incurred by the wellness policy. They will exercise the same amount as before the policy: $\frac{\alpha_1-\lambda}{\epsilon}$ hours/week, and enjoy the first $\frac{B}{\lambda}$ hours/week offered by the wellness policy.

People with $\alpha \in [\frac{\epsilon B}{\lambda}, \frac{\epsilon B}{\lambda} + \lambda]$ will increase their exercise to the full amount that the budget covers: $\frac{B}{\lambda}$ and not consume more, because $H_i^{ecs} = 0$ for all $\alpha \leq \frac{\epsilon B}{\lambda} + \lambda$. Again, this yields the same result as that of the payoff maximization problem case 3.

Figure 2 illustrates three different cases in which individuals 1, 2, and 3 have varying preferences for physical activity, denoted by $\alpha_1$, $\alpha_2$, and $\alpha_3$, respectively. Specifically, $\alpha_1 > \frac{\epsilon B}{\lambda} + \lambda$; $\alpha_2 \in [\frac{\epsilon B}{\lambda}, \frac{\epsilon B}{\lambda} + \lambda]$; $\alpha_3 < \frac{\epsilon B}{\lambda}$.

Individual 1, presented in panel (a), highly values physical activity and is typically very active. The wellness budget cannot fully cover her optimal level of exercise, i.e., $\frac{B}{\lambda} < \frac{\alpha_1-\lambda}{\epsilon}$ (case 2). Her optimal choice of physical activity, $H_1^* = \frac{\alpha_1-\lambda}{\epsilon}$, remains unchanged by the policy. However, she benefits from the first $\frac{B}{\lambda}$ hours/week of external-barrier-free activity provided by the policy due to increased convenience and accessibility. She belongs to the group of always-takers.

Individual 2, presented in panel (b), has a moderate preference for physical activity. She values it enough that the wellness budget offers more than her optimal choice in the presence of external barriers, but less than her choice in their absence, i.e., $\frac{\alpha_2-\lambda}{\epsilon} \leq \frac{B}{\lambda} \leq \frac{\alpha_2}{\epsilon}$ (case 3). Her optimal choice without the policy is $\frac{\alpha_2-\lambda}{\epsilon}$, which increases to $\frac{B}{\lambda}$ when the subsidy is available. She is a full complier.

Individual 3, shown in panel (c), places a relatively low value on physical activity. The wellness budget exceeds her interior solution without external barriers, i.e., $\frac{\alpha_3}{\epsilon} < \frac{B}{\lambda}$ (case 1). Her behavior without the policy would be either zero or limited at $\frac{\alpha_3-\lambda}{\epsilon}$. When the policy removes external barriers, her effort increases to $\frac{\alpha_3}{\epsilon}$, making her a partial complier.

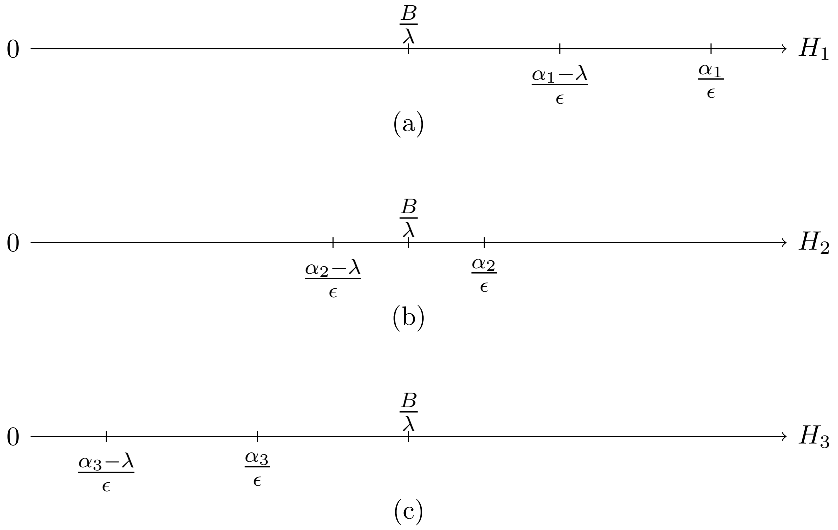


Figure 2: Three cases of individual responses to wellness policy, depending on $\alpha$

Table 1 summarizes the behavioral response groups implied by the model, based on individual values of $\alpha$. These categories: never-takers, partial compliers, full compliers, and always-takers, capture how different segments of the population are expected to respond to the wellness policy.

## 3.2 Implications of the Model

### 3.2.1 How different groups respond to wellness policy

One key implication of this model is that individuals vary widely in how they respond to wellness policies, depending on their internal valuation of physical activity, represented by $\alpha$. Those with relatively high values of $\alpha$, who place a strong intrinsic value on health and wellness, are unaffected by the policy in terms of their actual participation in physical activity. These individuals, known as the always-takers, are already active and motivated; they have incorporated physical activity into their lifestyle and are less likely to suffer from obesity or related health conditions. Although they benefit from improved convenience and accessibility, the policy does not alter their behavior, so its marginal effect on this group is negligible.

At the other end of the spectrum are individuals with negative or near-zero values of $\alpha$, who either perceive physical activity as harmful or face prohibitively high opportunity costs. This group represents the never-takers who reject the policy altogether.

The actual impact of the policy comes from individuals in the middle range: those with moderately positive values of $\alpha$, referred to as compliers. This group is heterogeneous:

- Partial compliers have a mild preference for physical activity, their optimal level of participation is lower than the amount subsidized by the policy. They respond by increasing their activity, but do not fully utilize the wellness budget.
- Full compliers are sufficiently motivated to fully use the subsidized amount of physical activity, though they do not go beyond it.

Because only this middle group changes behavior, the total increase in physical activity depends on the size of this group.

The behavioral responses to the policy illustrate how individual choices depend on the trade-off of perceived benefit from exercise against internal effort cost and external barriers.

### 3.2.2 The importance of education in wellness policies

Another implication of the model is that wellness policies that shift preferences, that is increase individuals' $\alpha$, can have a broader impact and are more likely to succeed. To illustrate this, let's consider two metrics: the total effect and the effectiveness ratio of the policy.

| **Group** | **Range of $\alpha$** | **PA without Policy** | **Response to Wellness Policy** | **Description** | **Payoff gain** |
|---|---|---|---|---|---|
| **Never-takers** | $\alpha < 0$ | No $H_i = 0$ | No response. $H_i = 0$ | Hold disbelief in PA or have high opportunity cost; policy does not change behavior | None |
| **Partial Compliers** | $0 \leq \alpha < \frac{\epsilon B}{\lambda}$ | No or limited $H_i = \frac{\alpha_i - \lambda}{\epsilon} \geq 0$ | Engage in PA, but below the subsidized amount. $H_i = \frac{\alpha_i}{\epsilon} < \frac{B}{\lambda}$ | Mild motivation; subsidy induces limited increase in PA, budget not fully used | Payoff gain from increased physical activity |
| **Full Compliers** | $\frac{\epsilon B}{\lambda} \leq \alpha < \frac{\epsilon B}{\lambda} + \lambda$ | No or limited. $H_i = \frac{\alpha_i - \lambda}{\epsilon} \geq 0$ | Engage in PA exactly up to the subsidized amount. $H_i = \frac{B}{\lambda}$ | Respond fully to policy but do not exceed what the budget covers | Payoff gain from increased physical activity |
| **Always-takers** | $\alpha \geq \frac{\epsilon B}{\lambda} + \lambda$ | Yes. $H_i = \frac{\alpha_i - \lambda}{\epsilon} > \frac{B}{\lambda}$ | No change in behavior; continue PA as before $H_i = \frac{\alpha_i - \lambda}{\epsilon} > \frac{B}{\lambda}$ | Already active and motivated; unaffected by policy but benefit from the convenience of accessibility | Payoff gain from improved accessibility and convenience |

Table 1: Behavioral Response Groups by $\alpha$ under the Wellness Policy Model

The total effect is measured by the aggregate increase in hours of physical activity per week across the population. This is calculated by subtracting the total hours of physical activity before policy implementation from the total hours after implementation.

$$\text{Total Effect} = \underbrace{\int_0^{\frac{\epsilon B}{\lambda}} \frac{\alpha}{\epsilon}\frac{1}{a-b}\,d\alpha}_{\text{partial compliers (after)}} + \underbrace{\int_{\frac{\epsilon B}{\lambda}}^{\frac{\epsilon B}{\lambda}+\lambda} \frac{B}{\lambda}\frac{1}{a-b}\,d\alpha}_{\text{full compliers (after)}} - \underbrace{\int_{\lambda}^{\frac{\epsilon B}{\lambda}+\lambda} \frac{\alpha-\lambda}{\epsilon}\frac{1}{a-b}\,d\alpha}_{\text{compliers total (before)}} = \frac{B}{a-b}$$

Total effect of the wellness policy depends positively on the budget per person $B$, and negatively on the span of the distribution $(a-b)$. Intuitively, the higher the budget per person, the more external barriers the policy is able to remove for each individual, leading to more exercised hours. Since $(a-b)$ is always positive ($a \in (0,\infty)$ and $b \in (-\infty, 0]$), the total effect is always positive. However, the larger the span, the smaller the effect of the policy, given any budget $B$. The reason is twofold. First, $(a-b)$ shows up in the density function $\frac{1}{a-b}$, the larger the span of $(a-b)$, the smaller the density. Second, the value of $\lambda$ is given by the nature of external barriers, $\epsilon$ is assumed to be constant and homogeneous, given physical constraints of human body, therefore the values of $\lambda$, $\frac{\epsilon B}{\lambda}$ and $\frac{\epsilon B}{\lambda} + \lambda$ are all exogenous and predetermined. Thus, a larger span between $a$ and $b$ means a smaller proportion of compliers with $\alpha \in (0, \frac{\epsilon B}{\lambda} + \lambda)$, and larger proportions of never takers with $\alpha \in (-\infty, 0]$ as well as always takers for which $\alpha \in [\frac{\epsilon B}{\lambda} + \lambda, \infty)$, leading to smaller policy effect. The idea is that a larger $(a-b)$ implies more people fall outside the effective treatment range.

The effectiveness ratio measures how efficiently the intended policy impact is translated to actual effect. The actual effect is what we have calculated above $\frac{B}{a-b}$, it is the total increase in hours of exercise per week. Intended effect is the expected increase in hours per week of exercise, given the budget per person $B$. For example, a school with $n$ students purchases a new set of facility for physical activity, has budget $B$ per student, and expects to increase on average $\frac{B}{\lambda}$ hours of exercise per week for a student, the total expected increase in hours of exercise per week for this school is $\frac{B}{\lambda} * n$ hours. Thus, the effectiveness ratio is given by:

$$\text{Effectivenss ratio} = \frac{\text{Total Effect}}{\text{Intended Effect}} = \frac{\frac{B}{a-b}}{\int_b^a \frac{B}{\lambda}\cdot\frac{1}{a-b}\,d\alpha} = \frac{\lambda}{a-b}$$

Surprisingly, the effectiveness ratio does not depend on the budget $B$, but rather on the unit external barrier $\lambda$. Why is that? Mathematically, $B$ cancels out. Intuitively, no matter how much is spent per person, the efficiency of that spending: the actual return in terms of increased physical activity, depends entirely on the conversion rate — the marginal increase in hours of exercise per week per unit of budget. This conversion rate is governed by the size of the external barrier removed by the policy, represented by $\lambda$.

In this sense, $\lambda$ captures how much additional activity a person gains when a unit of constraint is lifted, or equivalently, the marginal gain from an extra hour of exercise per week. Therefore, the larger the external barrier $\lambda$, the greater the gain per dollar spent, and hence the higher the effectiveness ratio. Unsurprisingly, the ratio also decreases with the span $(a - b)$, a larger $(a - b)$ span weakens the policy's impact across the population.

There are two ways to shorten the span of $(a - b)$: one is to lower $a$, effectively reducing the fraction of fitness enthusiasts in the population; the other is to raise $b$, which shrinks the group of never-takers. Since the second approach is more morally and politically justifiable, this highlights the importance of education, specifically, efforts aimed at increasing the value of $\alpha$ among never-takers. By encouraging these individuals to begin exercising, and by motivating existing compliers to increase their participation, such interventions expand the share of the population that responds to the policy. This transition from non-participation to partial or full compliance ultimately enhances the effectiveness of the overall policy.

In contrast, policies that fail to shift preferences rely entirely on the existing distribution of intrinsic motivation in the population. When a large share of the population consists of never-takers, the effectiveness of such policies will necessarily be limited.

Programs that combine increased accessibility with efforts to improve perceptions of physical activity are particularly promising. For example, the Comprehensive School Physical Activity Program (CSPAP) addresses both external barriers (e.g., lack of facilities or structured opportunities) and internal barriers (e.g., low motivation or limited awareness) by not only increasing access to physical activity but also fostering more positive attitudes toward active lifestyles among students.

# 4 Empirical Design

Since CSPAP works on both removing the external barriers by improving accessibility to fitness for students, and educating students as well as their parents so that they will acquire knowledge to better understand the important role of physical activity in health investment, it is interesting to find out whether such a comprehensive policy will have a detectable effect on health outcomes.

To assess whether exposure to CSPAP influences physical activity participation and obesity rates, I employ two differences-in-differences strategies. First, I use a dynamic dose-response model to estimate the marginal effect of CSPAP intensity on health outcomes. Second, I apply the Callaway and Sant'Anna (2021) method to examine whether states that reach a certain threshold of CSPAP adoption experience larger treatment effects compared to states that have not yet reached the threshold.

Given that CSPAP adoption rates (treatment intensity) vary across states and over time, the dynamic dose-response model is well-suited to capture how health outcomes respond to different levels of treatment exposure. The model is specified as:

$$y_{s,t} = \theta_s + \pi_t + \beta_1 x_{s,t} + \beta_2 x_{s,t-2} + \gamma m_{s,t} + u_{s,t} \tag{1}$$

where $y_{st}$ is obesity rate in state $s$ and year $t$. $\theta_s$ is state-fixed effect, $\pi_t$ is year-fixed effect. $x_{st}$ is the percentage of secondary schools in state $s$ that have fully adopted CSPAP that year. While $\beta_1$ estimates the instantaneous marginal effect of treatment intensity, $\beta_2$ estimates the two-year-lagged marginal effect. $\gamma$ is the coefficient of control variable real GDP per capita (in year 2017 dollars) $m_{s,t}$.

The model incorporates potential lagged effect, based on the assumption that it takes time for students' behaviors, and thus their health outcomes, to respond. However, due to the limitations of data, it is only reasonable to include a two-year lag [2]: the panel is short and covers just six post-treatment time periods in total, which constrains the ability of the model to reliably estimate longer-term effects.

While the model uses both time-series and cross-sectional variation in CSPAP adoption across U.S. states, credible identification also requires a discussion about the source of identifying variation, that is, whether the variation in adoption rate across states is exogenous and that it is not driven by underlying factors like culture and political priorities that shape the attitude towards physical activity and health. A key concern is that states with lower adoption rates might differ systematically from those with higher rates in unobserved ways that also affect obesity. For example, states that have less emphasis on health and physical activity may have higher obesity rate and lower adoption rate, this would necessarily endogenize adoption rate and the estimated effect will be biased. Since decision-making of CSPAP implementation is decentralized and at school level, it creates a idiosyncratic variation of state-level adoption rate that is unlikely to be correlated with factors that also affect obesity rate. To further address this concern, I first incorporate state and time fixed effects to capture any systematical differences among states. Then I test for parallel pre-treatment trends using both the event study estimator proposed by Callaway and Sant'Anna (2021) and visual inspection of obesity rate trajectories over time for states grouped by treatment intensity (presented in section 6) to take into account the possible pre-existing differences. Lastly, I add real GDP per capita as control variable to address the possible correlation between state-level GDP per capita with both CSPAP adoption rate and obesity rate, since real GDP per capital is time-variant and will not be captured by the state-fixed effect.

Another identification concern is the possibility of reverse causality—that is, states with higher obesity rates may be more likely to adopt CSPAP at higher intensities as a response to worsening public health outcomes. This concern is addressed in two ways. First, the parallel pre-treatment trends ensure that adoption intensity is not systemat-

---

[2] Variables are measured and reported biennially.

ically related to earlier trends in obesity. Second, the empirical model incorporates a lagged treatment effect, which helps to mitigate the possibility of reverse causality. Since the effect is measured two years after adoption, it is unlikely that current obesity rates are directly influencing the adoption rate two years ago.

Aside from parallel trend, another identifying assumption of a dose-response model is homogeneous marginal treatment effects, in other words, the marginal effect of treatment is constant across states. Given that the observed units are high school students, whose physiological characteristics and metabolic conditions are likely to be relatively similar across states, it is reasonable to assume that the average marginal effect of the program may be broadly comparable across settings. While the assumption of marginal treatment effect homogeneity is necessary for identification of a dose-response relationship, especially without a pure control group (CSPAP is a nationwide policy, no state can serve as a pure counterfactual group), it is not meant to rule out the possibility of treatment effect heterogeneity. If the treatment effect differs by state, then the average treatment effect estimated by a two-way fixed effect model might be biased (Sun and Shapiro, 2022). This is another reason why the employment of (Callaway and Sant'Anna, 2021) estimator is necessary for robustness checks.

For the Callaway and Sant'Anna (2021) estimator, I use a binary treatment defined at varying adoption thresholds. Due to limited effect of low treatment doses, I focus on treatment periods where adoption rates are generally higher than 15%, thus 2020 would be the first treatment year in Callaway and Sant'Anna (2021) setting. To operationalize treatment intensity, I define the "intensely treated group" as those states with CSPAP adoption rates above different percentiles of the 2020 distribution: the 10th, 30th, 50th (median), and 80th percentiles, respectively. I begin with the median threshold to test whether states in the upper half of the adoption distribution experienced significantly different Average Treatment Effect on the Treated (ATT) compared to those in the lower half. I then vary the threshold upward and downward to assess the robustness and persistence of the estimated effects. This approach enables comparison of ATT estimates over time for units that achieve higher levels of program adoption, while accounting for staggered treatment timing and allowing for heterogeneous treatment effects across groups. The regression employed is as follows:

$$\mathrm{y}_{s,t} = \theta_s + \pi_t + \sum_g \mathrm{ATT}(g) \cdot D^g_{s,t} + \gamma m_{s,t} + u_{s,t} \tag{2}$$

Where $y_{s,t}$ is outcome variable in state s and year t, $\theta_s$ is state-fixed effect, $\pi_t$ is year-fixed effect. $\sum_g \mathrm{ATT}(g) \cdot D^g_{s,t}$ is an interaction term of average treatment effect on treated (ATT) and treatment dummy $D^g_{s,t}$. $g$ is event time (years from treatment). $\gamma$ is the coefficient of control variable real GDP per capita $m_{s,t}$.

While the dynamic dose-response model estimates the marginal effect of CSPAP adoption intensity by using adoption rate as a continuous measure, the Callaway and Sant'Anna (2021) method enables testing whether reaching higher adoption rate generates larger effect, using each threshold as a discrete intervention. Thus, the dose-response model captures how incremental increases in CSPAP adoption affect health outcomes, while the Callaway and Sant'Anna (2021) estimator identifies the average treatment effect on treated (ATT) of achieving high program adoption rate relative to never or not yet reaching it. Using both approaches allows me to complement the marginal effect estimates with a richer comparison of health outcomes before and after crossing key adoption thresholds. Moreover, the Callaway and Sant'Anna (2021) method relaxes the assumption of homogeneous treatment effects across states and controls for staggered treatment timing, providing a useful robustness check for the findings from the dose-response analysis. Finally, parallel trend assumption is also tested using Callaway and Sant'Anna (2021). [3]

# 5 Data

## 5.1 Explanatory Variable

As individual or school specific data are not available, the primary explanatory variable in this thesis is adoption rate by state (percentage schools within each state that meet all criteria of CSPAP), which can be accessed through CDC's School Health Profiles website [4]. The CDC's School Health Profiles provides reports of surveys that assess policies and practices of school health, it provides data on topics about nutrition, tobacco-use prevention and physical activity (assessment regarding CSPAP), among other school health topics (Centers for Disease Control and Prevention, 2024b). According to Centers for Disease Control and Prevention (2024b), data are collected biennially in even-numbered years through self-administered questionnaires sent to principals and lead health education teachers in representative secondary schools. I use their school health profile data from 2014 to 2022 in which adoption rate of CSPAP is included.

## 5.2 Dependent Variable

For the dependent variable: obesity rate, I use CDCs Youth Risk Behavior Surveillance data (YRBS). [5] The Youth Risk Behavior Surveillance System (YRBSS) is a national health surveillance system established by the CDC in 1991 to monitor health-risk be-

---

[3] It is worth noting that although adoption rates generally increase over time, some states experience occasional temporary declines. Therefore, it is important to ensure that once a state is classified as treated at a given adoption threshold, it remains treated in all subsequent years. To address this, I implement coding procedures that enforce the "once treated, always treated" assumption at each threshold level.

[4] `https://www.cdc.gov/school-health-profiles/index.html`

[5] `https://www.cdc.gov/yrbs/index.html`

haviors among high school students in the U.S. The surveys is conducted biennially and assesses behaviors of adolescents (typically middle school and high school students), data are available in odd-numbered years, including information like dietary habits and physical activity as well as health conditions (Centers for Disease Control and Prevention, 2024a). According to Centers for Disease Control and Prevention (2024a), YRBSS uses a randomized, school-based survey methodology, ensuring a representative sample of high school and middle school students at the national, state, tribal, and local school district levels. The data can be used to detect health trends, identify rising health issues and analyze effect of health-promoting programs (Centers for Disease Control and Prevention, 2024a). Obesity rate in the data represents percentage ”high school students who were above 95th percentile for body mass index, based on sex- and age-specific reference data from the 2000 CDC growth charts” (Centers for Disease Control and Prevention, 2025).

### 5.3 Matching the Data

The CDC's School Health Profiles (SHP) data, which reports CSPAP adoption rates, is collected biennially during the spring semester of even-numbered years. In contrast, the Youth Risk Behavior Surveillance System (YRBS) data, which measures student health outcomes such as obesity rates, is collected during the spring of odd-numbered years.

I align adoption rates from year $t$ with obesity outcomes from year $t + 1$ by shifting the obesity rate data backward by one year (i.e., reassigning the obesity rate reported for year $t + 1$ to year $t$).[6] This adjustment ensures that the independent (adoption rate) and dependent (obesity rate) variables are properly synchronized.

Consequently, the dataset provides biennial information for even-numbered years from 1998 to 2022. It is worth noting that although CSPAP was introduced in 2013, adoption rate data for that year are unavailable, as the first recorded measurement appears in 2014, which therefore serves as the first treatment year in the dataset.

## 6 Descriptive Statistics

Table 2 shows the statistical description of obesity rate and adoption rate. Adolescent average obesity rate has been steadily increasing over the years with a slight decrease in 2022. Average adoption rate was low during the first years following the introduction of CSPAP, and rose substantially in 2020.

Figure 3 illustrates trends in average obesity rates over time, comparing states with relatively low CSPAP adoption rates (those whose maximum adoption rate never exceeded the 20th percentile of the adoption distribution) to states with progressively higher levels of adoption. Each panel compares the low-adoption group to a different threshold of

---

[6] Shifting adoption rate one year forward generates the same result.

high adoption, defined by whether a state's maximum CSPAP adoption rate exceeded the 20th, 50th, 70th, or 90th percentile, respectively.

Panels (a) through (d) correspond to comparisons with increasingly intensive treatment groups. Panel (a) compares states above the 20th percentile to those below it; panel (b) uses the 50th percentile; panel (c) the 70th; and panel (d) the 90th. In each panel, the vertical dashed line marks the year before introduction year of CSPAP (2013), which serves as the treatment onset.

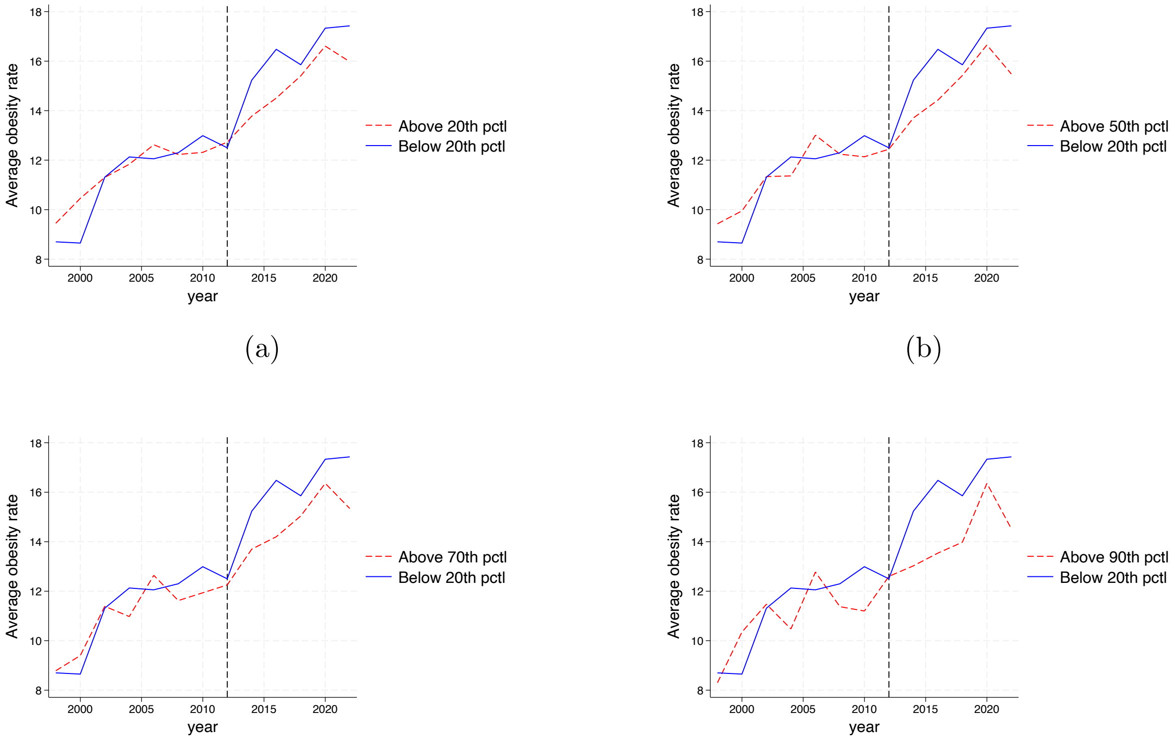
Across all panels, the two groups display relatively parallel pre-treatment trends in obesity rates, supporting the parallel trends assumption necessary for causal interpretation in a differences-in-differences framework. Looking at years following the introduction of CSPAP, the divergence between high- and low adoption groups becomes increasingly pronounced as the adoption threshold rises. This pattern suggests a potential dose-response relationship: states with higher levels of CSPAP adoption experienced smaller increases—or flatter trajectories—in obesity rates relative to states with minimal adoption. The effect is most visible in panels (c) and (d), where the obesity rates among states above the 70th and 90th percentiles remain consistently lower than those in the lowest adoption group.

Furthermore, in all four panels, states with higher adoption rates exhibit a notably sharper decline in obesity rates beginning around 2020. This coincides with a period when CSPAP adoption surged across states, suggesting that the effect of implementing CSPAP became more detectable once a critical threshold of adoption was reached, and highlighting the importance of distinguishing between different intensities of program exposure when evaluating policy effectiveness.

Table 2: Statistical Description of Variables

| Year | Obesity Rate | | | | | Adoption Rate | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Min | Max | N | Mean | SD | Min | Max | N |
| 2008 | 12.19 | 2.33 | 6.3 | 18.1 | 43 | . | . | . | . | . |
| 2010 | 12.37 | 2.36 | 7.9 | 17.0 | 44 | . | . | . | . | . |
| 2012 | 12.56 | 2.50 | 6.4 | 18.0 | 43 | . | . | . | . | . |
| 2014 | 14.00 | 2.44 | 10.3 | 18.9 | 37 | 3.57 | 2.71 | 0.0 | 14.3 | 44 |
| 2016 | 14.76 | 2.78 | 9.6 | 21.7 | 39 | 3.81 | 2.91 | 0.7 | 13.9 | 46 |
| 2018 | 15.30 | 3.12 | 7.7 | 23.4 | 45 | 4.40 | 3.00 | 0.0 | 12.3 | 40 |
| 2020 | 16.58 | 3.24 | 9.0 | 26.9 | 46 | 50.10 | 15.40 | 19.9 | 87.8 | 44 |
| 2022 | 16.00 | 2.77 | 8.7 | 22.1 | 38 | 49.50 | 16.80 | 15.6 | 95.5 | 41 |

Figure 3: Obesity Trend by Group And Year



(a)

(b)

(c)

(d)

# 7 Result

## 7.1 Results from Dose-response Model

Table 3 reports the results from regression model (1), which estimates the marginal effect of CSPAP treatment intensity on state-level obesity rates. The key coefficients of interest are $\beta_1$ , capturing the instantaneous effect of CSPAP adoption, and $\beta_2$, capturing the effect with a two-year lag. The first row present estimates for $\beta_1$, and the second row give estimates for $\beta_2$. Columns (1) and (2) give results for instantaneous effect only, while columns (2) and (4) give results including both instantaneous and two-year-lagged effect. Columns (1) and (3) show results from a baseline specification without controls, whereas columns (2) and (4) include real GDP per capita as a control to account for variation in economic conditions across states.

There is no clear evidence that CSPAP has an immediate impact on obesity rates in the same year of adoption, perhaps unsurprisingly, as behavioral change and health outcomes rarely shift overnight. Both column (1) and column (2) show statistically insignificant estimates for the instantaneous effect.

By contrast, the two-year-lagged effects reported in columns (3) and (4) tell a more compelling story. These coefficients capture the delayed impact of CSPAP as the program becomes embedded in school environments and begins to influence student activity levels and lifestyle habits. Both estimates are statistically significant at the 1% level and highly consistent across specifications: -0.052 without controlling for real GDP per capita (column 3) and -0.054 with the control (column 4), each with a standard error of 0.019. The robustness and magnitude of these effects suggest that CSPAP's influence takes time to materialize, aligning with broader evidence that the benefits of behavioral interventions typically accumulate gradually as exposure increases and habits shift.

While the program does not appear to move the needle in the short run, the evidence points to a meaningful reduction in obesity rates two years after higher adoption levels, highlighting the value of patience and persistence in school-based wellness policy. These findings are consistent with the visual trends shown in Figure 3, where the gap in obesity rates between high- and low-adoption states becomes increasingly visible several years after program implementation, especially in states with higher treatment intensity.

Table 3: Marginal Effect of CSPAP Intensity on Obesity Rate

| | (1) Obesity rate | (2) Obesity rate | (3) Obesity rate | (4) Obesity rate |
|---|---|---|---|---|
| Instantaneous effect | -0.095 | -0.012 | 0.018 | 0.015 |
| | (0.014) | (0.015) | (0.020) | (0.020) |
| After 2 yrs | | | -0.052*** | -0.054*** |
| | | | (0.019) | (0.019) |
| *N* | 469 | 448 | 439 | 438 |

*standard errors in parentheses*
*clustered at state level*
$^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

Although CSPAP adoption rates varied across states, 2020 marked a widespread increase in participation. Between 2018 and 2020, the average adoption rate rose from 4% to 50%, reflecting a significant national push toward broader implementation. While this expansion demonstrates the program's momentum, it also reduces cross-sectional variation in treatment intensity: variation that is essential for causal identification in difference-in-differences and panel data settings.

To address this concern, I re-estimate regression model (1), this time excluding post-2020 observations of adoption rates to focus on the program's early effects, before the collective adoption surge. The results are presented in Table 4.

All specifications in Table 4 include controls for real GDP per capita. Columns (1) and (3) reproduce the full-sample estimates from Table 3, capturing the instantaneous effect ($\beta_1$) and two-year-lagged effect ($\beta_2$) of CSPAP treatment intensity, respectively. Columns (2) and (4) re-estimate these effects using a restricted sample that excludes post-2020 adoption data, allowing us to isolate the relationship between early-stage CSPAP intensity and changes in obesity rates.

As in previous results, the instantaneous effects in columns (1) and (2) remain small and statistically insignificant, reinforcing the view that the CSPAP program does not generate immediate changes in obesity outcomes.

The two-year-lagged effects, however, continue to show meaningful and policy-relevant patterns. Column (3) shows a statistically significant effect of $-0.054$ ($p < 0.01$), replicating the earlier result. In the restricted sample, column (4) shows a larger lagged effect of $-0.084$, significant at the 10% level. Although the estimate in column (4) is less precise due to a smaller sample size and reduced within-group variation, the result supports the hypothesis that early adopters of CSPAP experienced measurable reductions in obesity over time.

Together, these findings confirm the delayed impact of CSPAP treatment intensity on obesity and suggest that the observed effects are not driven solely by the post-2020 national rollout. Instead, the core result, that higher CSPAP adoption leads to lower obesity rates after two years, remains robust to the exclusion of late-adopting states and to variation in the size and timing of program implementation.

Table 4: Marginal Effect of CSPAP Intensity on Obesity Rate

| | (1)<br>Obesity rate | (2)<br>Obesity rate | (3)<br>Obesity rate | (4)<br>Obesity rate |
|---|---|---|---|---|
| Instantaneous effect | -0.012 | -0.015 | 0.018 | 0.028 |
| | (0.015) | (0.066) | (0.02) | (0.075) |
| After 2 yrs | | | -0.054$^{***}$ | -0.084$^{*}$ |
| | | | (0.019) | (0.05) |
| $N$ | 448 | 371 | 438 | 368 |

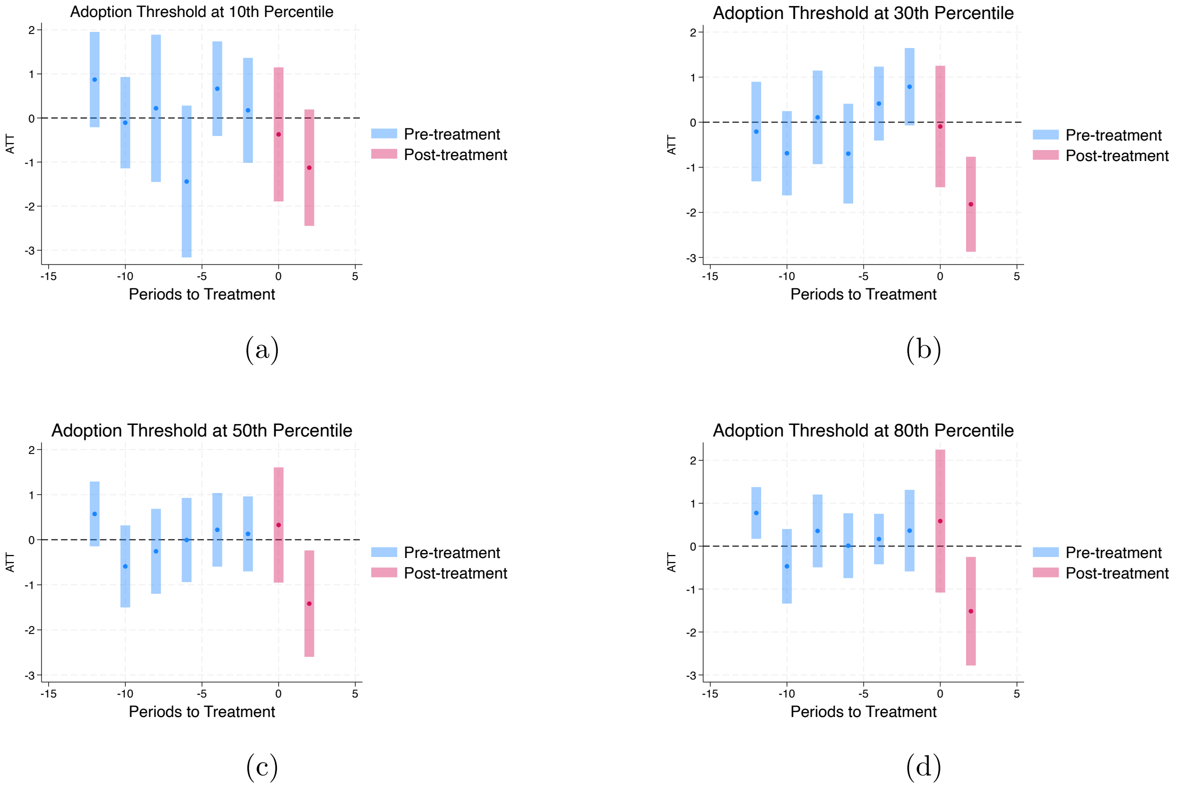
*standard errors in parentheses*
*clustered at state level*
$^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

## 7.2 Results from Callaway and Sant'Anna Estimator

Figure 4 presents estimates of the event study using the Callaway and Sant'Anna (2021) estimator, applied at various adoption thresholds, specifically, the 10th, 30th, 50th and 80th percentiles of the CSPAP adoption rate distribution. Each panel displays the estimated Average Treatment Effect on the Treated (ATT) over time, where period 0 marks the first year a state crosses the specified threshold. The vertical bars indicate 95% confidence intervals, and the dots represent point estimates for each period relative to treatment.

Panels (a) through (d) show the dynamic effects of CSPAP adoption across increasingly intensive thresholds. Across all panels, the pre-treatment periods (shown in blue) display estimates that cluster closely around zero with wide, overlapping confidence intervals, suggesting no clear evidence of anticipatory effects or significant differences in obesity trends prior to treatment. This pattern provides support for the parallel trends assumption and strengthens the identification of the dose-response model.

In contrast, the post-treatment periods (shown in pink) consistently show negative ATT estimates, indicating a decline in obesity rates among states that cross the adoption threshold. The results suggest a significant and consistent two-year-lagged Average Treatment Effect on the Treated (ATT) at different threshold, particularly at higher ones. As the threshold rises, the treatment group becomes smaller and more selective, while the control group increasingly includes states with moderate CSPAP exposure. This setup blurs the contrast between treated and control units, which would typically make it harder to detect an effect. Yet, the ATT remains negative and statistically significant even under these conditions. This confirms the policy effect on obesity rate even when the treatment effect is allowed to vary across units.



(a)

(b)

(c)

(d)

Figure 4: Event study estimates using the Callaway & Sant'Anna estimator at various adoption thresholds

# 8 Discussion and Implication

This section interprets the estimated effects of CSPAP adoption on adolescent obesity rates based on the two-year-lagged marginal effect obtained from the dose-response regression model (Regression 1).

## 8.1 Marginal Treatment Effect

The estimated two-year-lagged marginal effect is:

$$\beta_2 = -0.054$$

This coefficient can be interpreted as the marginal average treatment effect on the treated (ATT) per unit of treatment intensity, following the framework in Borusyak et al. (2024). In other words, a 1 percentage point increase in CSPAP adoption is associated with a 0.054 percentage point reduction in the adolescent obesity rate two years later.

This marginal treatment effect, in the view of Sun and Shapiro (2022), also measures the effect of implementing CSPAP (treatment intensity going from 0 to 1) on obesity rate, meaning a full implementation of CSPAP (all schools are exposed) would lead to a reduction in obesity rate of 5.4 percentage points

$$-0.054 \times 100 = -5.4$$

## 8.2 Calculating the Average Treatment Effect (ATE)

To estimate the actual population-level impact, we multiply the marginal effect by the average adoption rate in each year. The ATE in year $t$ is calculated using:

$$\text{ATE}_t = \beta_2 \cdot \overline{x}_{t-2}$$

where $\overline{x}_{s,t-2}$ is the average CSPAP adoption rate across states two years before year $t$.

For example, in 2020, the average adoption rate was 50.1%. Plugging this into the formula:

$$\text{ATE}_{2022} = -0.054 \times 50.1 = -2.7054$$

Thus, in 2022, CSPAP adoption is estimated to have reduced national adolescent obesity rates by approximately 2.7 percentage points.

## 8.3 Interpreting the Magnitude

To contextualize this result, the observed average obesity rate among high school students in 2022 was 16%. Without CSPAP, the predicted rate would have been:

$$16\% + 2.7\% = 18.7\%$$

This suggests that, on average, any randomly selected high school student had a 2.7 percentage point lower probability of being obese in 2022 due to CSPAP.

Table 5 presents the estimated average treatment effects in each year, alongside observed and counterfactual obesity rates (i.e., predicted rates in the absence of CSPAP).

| **Year** | **2016** | **2018** | **2020** | **2022** |
|---|---|---|---|---|
| Estimated ATE (percentage points) | 0.19 | 0.20 | 0.24 | 2.70 |
| Observed average obesity rate (%) | 14.76 | 15.30 | 16.58 | 16.00 |
| Predicted rate without CSPAP (%) | 14.95 | 15.50 | 16.82 | 18.70 |

Table 5: Estimated reduction in obesity rate by year

The empirical analysis shows that CSPAP adoption has had a sizable and meaningful effect on reducing obesity rates among high schools students in the U.S, especially in years with higher average adoption rate. This result suggests a possible significant policy impact in light of the rising concern of youth obesity.

# 9 Conclusion

By combining theoretical modeling and empirical analysis, this thesis analyzes and discusses the challenges faced by wellness policies when it comes to incentivizing healthy behaviors. It also explores possible improvements in design to enhance the effectiveness of such policies. The model developed in this thesis incorporates individuals with preference heterogeneity facing both internal and external barriers when it comes to choosing the optimal amount of health investment (in the context: physical activity). The model suggests that wellness policies aimed at reducing external barriers, such as subsidies or infrastructure investments, can encourage individuals with moderate preferences to increase their participation in physical activity, potentially leading to improved health outcomes. However, the overall effect of such policies depends heavily on the underlying distribution of preferences within the population. Moreover, the model points to the potential for greater impact when wellness policies are paired with educational components that shift internal preferences. The framework is generalizable and can be extended to evaluate other behavioral health interventions.

The empirical analysis investigated the effect of CSPAP on obesity rate among high school students using panel data from Centers for Disease Control and Prevention (CDC). My main identification is a dose-response model that estimates both instantaneous and two-year-lagged marginal effect of treatment intensity on obesity rate. The results suggest that a one percentage point increase in CSPAP adoption rate leads to 0.054 percentage points reduction in obesity rate, which translates to 2.7 percentage points reduction in obesity rate nationwide in 2022.

To check the robustness of the two-year-lagged marginal effect, I employed Callway

Sant'Anna (2021) and set the binary treatment at different thresholds. The consistent and significant two-year-lagged ATT results support and reinforce the estimated marginal effect from the dose-response model. Additionally, I re-estimated the dose-response model using only pre-2020 data to account for the timing of the nationwide surge in adoption, confirming the robustness of the treatment effect.

The theoretical and empirical findings suggest two key policy implications: (1) wellness policies may be more effective when they simultaneously reduce external barriers and reshape internal perceptions through education, and (2) both the intensity and duration of implementation play a critical role in determining policy success.

# References

Banerjee, A. V., Gertler, P. J., and Ghatak, M. (2002). Empowerment and efficiency: Tenancy reform in west bengal. *Journal of political economy*, 110(2):239–280.

Borusyak, K., Jaravel, X., and Spiess, J. (2024). Revisiting event-study designs: robust and efficient estimation. *Review of Economic Studies*, 91(6):3253–3285.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of econometrics*, 225(2):200–230.

Cawley, J. and Price, J. A. (2013). A case study of a workplace wellness program that offers financial incentives for weight loss. *Journal of health economics*, 32(5):794–803.

Centers for Disease Control and Prevention (2015). Comprehensive school physical activity program. Last reviewed September 25, 2015, available at: `https://archive.cdc.gov/#/details?url=https://www.cdc.gov/healthyschools/physicalactivity/cspap.htm`.

Centers for Disease Control and Prevention (2019). Increasing physical education and physical activity: A framework for schools. `https://www.cdc.gov/physical-activity-education/media/pdfs/2019_04_25_PE-PA-Framework_508tagged.pdf`. Accessed May 2025.

Centers for Disease Control and Prevention (2024a). About the youth risk behavior surveillance system (yrbss). `https://www.cdc.gov/yrbs/about/index.html`. Accessed May 2025.

Centers for Disease Control and Prevention (2024b). School health profiles. `https://www.cdc.gov/school-health-profiles/index.html`. Accessed May 2025.

Centers for Disease Control and Prevention (2025). High school youth risk behaviour survey questionairs. Accessed on April, 2025, available at: `http://yrbs-explorer.services.cdc.gov/`.

Einav, L., Lee, S., and Levin, J. (2019). The impact of financial incentives on health and health care: Evidence from a large wellness program. *Health Economics*, 28(2):261–279.

Epassi (2023). Friskvårdsförmåner för miljarder kronor förblev outnyttjade enligt ny sifo-undersökning. Press release, published on Via TT.

Feng, J., Fan, L., and Jaenicke, E. C. (2024). The distributional impact of snap on dietary quality. *Agricultural Economics*, 55(1):104–139.

Holden, L. (2020). Cost tops list of hurdles to eating healthy while receiving snap: A new usda study. Accessed: YYYY-MM-DD.

Mancino, L. and Guthrie, J. (2014). Snap households must balance multiple priorities to achieve a healthful diet. *Amber Waves: The Economics of Food, Farming, Natural Resources, and Rural America*, (10).

Mande, J. and Flaherty, G. (2023). Supplemental nutrition assistance program as a health intervention. *Current Opinion in Pediatrics*, 35(1):33–38.

Schanzenbach, D. W. (2017). Pros and cons of restricting snap purchases. *Brookings*. Acceessed on April, 2025, available at: `https://www.brookings.edu/articles/pros-and-cons-of-restricting-snap-purchases/`.

Sun, L. and Shapiro, J. M. (2022). A linear panel model with heterogeneous coefficients and variation in exposure. *Journal of Economic Perspectives*, 36(4):193–204.

Todd, J. and Ver Ploeg, M. (2015). Restricting sugar-sweetened beverages from snap purchases not likely to lower consumption. *Amber Waves: The Economics of Food, Farming, Natural Resources, and Rural America*, (2).