

ORIGINAL RESEARCH

The Economics of Access to Youth Sports

Income Gradients, Geographic Variation, and the Participation Shortfall in U.S. National Survey Data

Zhengdao Xie

Xi'an Gaoxin No.1 High School International Department, Xi'an, China

Accepted: October 24, 2026

Abstract

Organized youth sports participation depends on families' ability to meet financial, transportation, and time costs. This study examines socioeconomic inequalities in U.S. youth sports participation using publicly available national survey estimates. The primary analysis uses the 2023–2024 National Survey of Children's Health (NSCH), covering youth ages 6–17. Supplementary analyses use the 2020–2021 NSCH and the 2020 National Health Interview Survey. In 2023–2024, 56.7% of youth participated in sports teams or lessons. Participation increased from 36.7% among families below the federal poverty threshold to 74.1% among those at or above 400% of that threshold, a 37.4-percentage-point gap and prevalence ratio of 2.02. State participation ranged from 46.3% to 70.9%. Using the 2020–2021 NSCH, Monte Carlo propagation of published uncertainty yielded a slope index of inequality of 46.6 percentage points (95% simulation interval: 43.5–49.6) and a relative inequality index of 2.59 (2.41–2.79). Comparing lower-income groups with the highest-income group's participation rate implied a descriptive shortfall of 8.23 million participants (7.66–8.80 million). Independent national survey estimates showed a similarly large income gap. These findings establish substantial economic stratification in youth sports access but do not identify causal effects of income or subsidies. They support considering policies that address participation fees alongside equipment, transportation, and program accessibility.

Keywords: youth sports; socioeconomic inequality; public policy; economics; National Survey of Children's Health; access; inequality index

1. Introduction

Organized youth sport sits at the intersection of health policy, education policy, household economics, and community infrastructure. A child may be interested in playing soccer, basketball, swimming, or another organized activity, yet interest alone does not determine participation. Participation can require registration fees, equipment, transportation, parental time, access to facilities, knowledge of available programs, and schedules that are compatible with schoolwork and family responsibilities. In economic terms, the observed decision to participate reflects both preferences and constraints. When these constraints vary systematically with household resources, access to sport can become stratified by income even if the underlying desire to participate is widely shared.

The policy importance of this issue follows from the role that organized sport can play in adolescents' lives. A systematic review by Eime et al. (2013) identified recurring associations between sport partici-

pation and psychosocial outcomes including self-esteem and social interaction. More recent federal data also show that U.S. youth who participate in sports are more likely to be physically active on most days of the week and differ from nonparticipants on several health and behavioral indicators. These observational relationships should not be interpreted as causal, because healthier or more advantaged youth may also be more likely to enter sport. Nevertheless, they establish that access to organized sport is connected to a set of outcomes that policymakers and schools already treat as socially important.

A large international literature documents socioeconomic differences in organized sport. Owen et al. (2022) synthesized 104 studies from high-income countries and found that children and adolescents in higher-socioeconomic-status households had 1.87 times the odds of sport participation compared with lower-status peers. The disparity was larger for organized sport than for physical activity overall, suggesting that institutions, fees, facilities, and other structured inputs matter above and beyond general willingness to be active. In a U.S. youth sample, Tandon et al. (2021) similarly found lower sport participation among youth from lower-affluence families and disproportionate reporting of barriers including cost, transportation, injury concern, and not feeling welcome on teams.

The same general relationship is visible outside the United States. Zheng (2025), analyzing 1,512 parents in Shanghai, reported a significant positive association between socioeconomic status and children's organized sport participation. Parental physical activity and parental support accounted for part of the relationship. This is important for interpreting income gradients: household income is unlikely to operate through a single mechanism. Money can cover fees and equipment, but socioeconomic advantage can also be bundled with transportation, schedule flexibility, information, parental experience with sport, social networks, and access to programs.

The federal policy environment recognizes these barriers. The U.S. National Youth Sports Strategy identifies cost, lack of accessible play spaces, social factors, lack of knowledge about opportunities, and time or competing demands as barriers to entry. It also notes that equipment and travel costs can create ongoing financial strain and that pay-to-play structures can disproportionately exclude youth from lower socioeconomic backgrounds. The policy question is therefore not only whether inequality exists, but how large it is, how it compares with other demographic gaps, how it varies geographically, and what scale of participation is potentially associated with the income divide.

This paper contributes a transparent secondary analysis of publicly available national survey estimates. Rather than claiming access to restricted microdata or estimating individual-level causal models that cannot be reproduced from published tables, it focuses on effect sizes and inequality measures that can be derived directly from federal survey products. The 2023–2024 National Survey of Children's Health (NSCH) is used for the primary contemporary description of income, age, sex, and state differences. A detailed 2020–2021 NSCH income table is used for supplementary rank-based inequality measures and a population shortfall calculation because it publishes subgroup confidence intervals and population estimates. The 2020 National Health Interview Survey (NHIS) provides an independent national replication of the income gradient.

Four questions guide the analysis. First, how large is the current income gradient in U.S. youth sports participation? Second, how does the magnitude of the income gap compare with age and sex differences reported in the same national data? Third, how much geographic variation exists across states and Census regions? Fourth, when a detailed national income table is used, what do rank-based inequality indices and a descriptive equal-access counterfactual imply about the scale of the participation divide?

2. Background and Conceptual Framework

2.1 Youth sport as an economic access problem

A useful economic model begins with the idea that organized sport has a full price that exceeds the posted registration fee. Let the household's effective cost of participation be $C = F + E + T + O$, where F denotes formal fees, E equipment and clothing expenses, T transportation and travel costs, and O the opportunity cost of time. A program can therefore be nominally inexpensive while remaining costly if it requires long travel, specialized equipment, frequent practices, or substantial parental coordination.

Conversely, a school-based program with shared equipment and immediately-after-school scheduling can have a low effective price even if a modest registration fee remains.

This framework explains why a pure income-participation association cannot reveal which policy lever is most effective. Income changes a household's capacity to pay F and E , but it may also be associated with car ownership, flexible work schedules, neighborhood facilities, information, prior sport experience, and the ability to absorb uncertain future costs such as tournaments or travel. The observed income coefficient in any cross-sectional model is therefore a reduced-form association that combines several channels.

The framework also predicts cumulative advantage. Children who can afford early participation acquire skills, confidence, social networks, and knowledge of the sport system. Those accumulated advantages can lower the nonfinancial cost of later participation. A child excluded at age eight because of fees or transportation may not simply enter at age fourteen when the family's finances improve, because competitive tryouts can then reward skills developed through earlier access. This dynamic makes youth sport policy relevant not only to current activity but also to the distribution of future opportunity.

2.2 Evidence on socioeconomic inequality

The strongest synthesis of the international evidence is the meta-analysis by Owen et al. (2022). Across 23 studies contributing to the sport-participation meta-analysis, higher-SES children and adolescents had higher odds of participating in sport (pooled OR 1.87, 95% CI 1.38–2.36). The authors concluded that socioeconomic differences were more pronounced for organized sport than for total physical activity. This distinction is consistent with an access-cost interpretation because organized sport requires institutional participation rather than only individual movement.

U.S. evidence points in the same direction. Tandon et al. (2021) surveyed 1,038 youth in grades 5–12 in King County, Washington. Low-affluence youth reported fewer days of physical activity, fewer sports sampled, and lower rates of sports participation. Cost and transportation were disproportionately reported barriers, but the study also identified nonfinancial constraints such as feeling unwelcome and fear of injury. The findings imply that economic disadvantage can operate through both material and institutional pathways.

National federal surveys provide a broader population perspective. Using the 2020 NHIS, Black et al. (2022) reported that organized sports participation rose from 31.2% among children below 100% of the federal poverty level (FPL) to 70.2% among children at 400% or more of FPL. Participation also rose strongly with parental education and declined as county social vulnerability increased. These associations were based on a nationally representative household survey and therefore demonstrate that the socioeconomic gradient is not limited to a particular city or sports system.

2.3 Policy interventions and the limits of fee-only explanations

If money is one important barrier, subsidies are an intuitive policy response. Evidence from real programs suggests that reducing price can increase participation, although effects depend on implementation and context. Meyers et al. (2023) evaluated the YES Card leisure-time voucher program in Franklin County, Kentucky. Across intervention and comparison schools, voucher recipients had 1.91 times the odds of sports participation over the study period (95% CI 1.08–3.38). Because the design was quasi-experimental rather than individually randomized, the estimate should not be interpreted as a universal causal parameter, but it provides policy-relevant evidence that reducing out-of-pocket costs can matter.

The ACTIVE trial in South Wales used a broader intervention combining £20 monthly activity vouchers with peer mentoring and support-worker engagement. Seven secondary schools were randomized, with 524 students in intervention schools and 385 in control schools. The intervention increased the likelihood of being classified as fit (OR 1.21, 95% CI 1.07–1.38), and qualitative findings emphasized demand for accessible, local, informal, and social activities (James et al. 2020). This result is especially relevant because the intervention did not treat price as the only barrier. It paired purchasing power with local access and social support.

These studies motivate a cautious policy interpretation of national income gaps. A large income gradient supports concern about economic access, but it does not identify a fee subsidy as a complete solution. The gradient may reflect transportation, facilities, social inclusion, parental time, and program design as well. The empirical analysis below therefore focuses on measuring the scale and structure of inequality, while the discussion evaluates policy mechanisms using evidence from intervention studies rather than inferring causality from cross-sectional national data.

3. Data and Methods

3.1 Primary data source: 2023–2024 NSCH

The primary data source is the Health Resources and Services Administration (HRSA) Maternal and Child Health Bureau's July 2026 data brief, *Youth Sports Participation, 2023–2024*, based on the National Survey of Children's Health. The NSCH is an annual household survey funded and directed by HRSA and conducted by the U.S. Census Bureau. A parent or guardian completes the questionnaire for one sampled child. The 2023–2024 combined data included questionnaires for 106,537 children ages 0–17, including 71,172 youth ages 6–17. Estimates are survey-weighted to represent children living in U.S. households (Sappenfield et al. 2026).

Youth sports participation is defined as participation during the past 12 months in a sports team or sports lessons after school or on weekends. The brief reports estimates by four age groups, sex, four family-income-to-poverty categories, and state. Family income categories are less than 100% FPL, 100–199% FPL, 200–399% FPL, and 400% or more of FPL. The primary analysis extracts these published estimates and derives effect-size contrasts from them.

3.2 Supplementary detailed source: 2020–2021 NSCH income table

A supplementary analysis uses the Data Resource Center for Child and Adolescent Health's 2020–2021 NSCH query for the same sports item, stratified by household income. The archived table publishes, for each of the four FPL groups, the weighted participation rate, 95% confidence interval, unweighted sample counts, and weighted population estimates. These additional quantities permit population-rank inequality calculations that cannot be reconstructed from the abbreviated 2023–2024 federal brief.

The 2020–2021 estimates are not interpreted as a formal time trend. HRSA and Census have revised historical NSCH weighting, and the 2020–2021 period also overlaps the COVID-19 pandemic. Accordingly, the older table is used only as a sensitivity analysis of the structure and scale of income inequality. The contemporary conclusion rests on the 2023–2024 estimates.

3.3 External replication: 2020 NHIS

An independent replication uses the National Center for Health Statistics Data Brief No. 441, based on the 2020 National Health Interview Survey. The NHIS reported sports participation of 31.2% (SE 2.9 percentage points) below 100% FPL, 50.8% (SE 1.5) at 100% to less than 400% FPL, and 70.2% (SE 1.5) at 400% or more of FPL. The survey, sampling frame, and income grouping differ from the NSCH, so the NHIS is used to test whether a similarly large high-versus-low income contrast appears in an independent federal data system rather than to estimate a pooled trend.

3.4 Effect-size calculations

For the highest- and lowest-income groups, three descriptive measures are calculated. The absolute participation gap is $RD = p_H - p_L$, where p_H and p_L are the high- and low-income participation proportions. The prevalence ratio is $PR = p_H / p_L$. The odds ratio is $OR = [p_H / (1 - p_H)] / [p_L / (1 - p_L)]$. These measures describe the same disparity on different scales. The risk difference is most intuitive for policy because it is expressed in percentage points, while the prevalence and odds ratios describe relative inequality.

To assess the substantive magnitude of income inequality relative to other demographic differences, the 2023–2024 income gap is divided by the male–female participation gap and by the difference between

the highest-participating and lowest-participating age groups. This is a descriptive scale comparison, not a causal decomposition.

3.5 Rank-based inequality analysis

For the 2020–2021 NSCH sensitivity analysis, each income group is assigned a ridity rank based on its weighted population share. The ridity is the midpoint of the cumulative population distribution occupied by each ordered income category. A population-weighted linear regression of participation prevalence on the ridity rank yields the slope index of inequality (SII). The SII can be interpreted as the model-implied absolute difference in participation between the bottom and top of the socioeconomic distribution. A logistic regression of the log-odds of participation on the same rank produces predicted endpoint prevalences; their ratio is reported as a relative inequality index (RII).

Uncertainty is propagated using 200,000 Monte Carlo draws. For each income group, the reported 95% confidence interval is converted to an approximate standard error by dividing the interval width by 3.92. Prevalence estimates are then sampled from normal distributions centered on the published point estimates, constrained to the 0–1 range. The SII, RII, high-versus-low gap, prevalence ratio, odds ratio, and population shortfall are recomputed for every draw. The 2.5th and 97.5th percentiles form simulation intervals. This procedure does not recover the full survey covariance matrix, so the intervals should be interpreted as approximate sensitivity intervals rather than replacements for design-based microdata inference.

3.6 Descriptive equal-access counterfactual

The detailed NSCH table also permits a population-scale calculation. For each lower-income group g , the shortfall relative to the highest-income participation rate is defined as $S_g = N_g(p_H - p_g)$, where N_g is the weighted number of youth in group g . Summing across the three lower-income categories gives the number of additional participants that would be observed if those groups had the same participation prevalence as the $\geq 400\%$ FPL group, holding group population sizes fixed. This quantity is explicitly descriptive. It is not an estimate of how many children a subsidy or any other specific policy would cause to participate, because the income gradient is not causal and the highest-income rate is not a policy treatment.

3.7 State-level analysis

The 2023–2024 HRSA brief reports sports participation for all 50 states and the District of Columbia. Those 51 values are summarized using the median, standard deviation, interquartile range, and range. States are grouped using the four U.S. Census regions. Regional means and medians are intentionally unweighted across jurisdictions. They therefore describe the typical state in a region, not the population-weighted participation rate for all children living in that region. No significance tests are performed on the regional summaries because the public brief does not provide state-specific standard errors in the map.

4. Results

4.1 Contemporary national participation and the income gradient

In 2023–2024, 56.7% of U.S. youth ages 6–17 participated in a sports team or sports lessons, representing almost 28 million youth. Participation differed modestly by age and sex but dramatically by family income. The highest age-group prevalence was 60.0% among youth ages 9–11, while the lowest was 50.8% among youth ages 15–17, a 9.2-percentage-point difference. Boys participated at 60.7% compared with 52.6% of girls, an 8.1-point difference.

The income gradient was much larger. Participation rose from 36.7% below 100% FPL to 44.4% at 100–199% FPL, 56.2% at 200–399% FPL, and 74.1% at 400% or more of FPL. The high-versus-low absolute gap was therefore 37.4 percentage points. On a relative scale, the highest-income group had

2.02 times the participation prevalence of the lowest-income group. The corresponding odds ratio was 4.93.

The size of the income gap is substantively notable even before formal modeling. It was 4.62 times the male–female participation gap and 4.07 times the difference between the highest- and lowest-participating age groups. The stepwise income increases were 7.7, 11.8, and 17.9 percentage points moving upward across the four income categories. Thus, the gradient was monotonic and the largest step occurred between the 200–399% and $\geq 400\%$ FPL groups.

Table 1. Selected 2023–2024 NSCH youth sports participation estimates.

Group	Participation (%)	Prevalence ratio vs. <100% FPL
Overall	56.7	N/A
Age 6–8	56.7	
Age 9–11	60.0	
Age 12–14	59.6	
Age 15–17	50.8	
Boys	60.7	
Girls	52.6	
<100% FPL	36.7	1.00
100–199% FPL	44.4	1.21
200–399% FPL	56.2	1.53
$\geq 400\%$ FPL	74.1	2.02

Source: Sappenfield et al. (2026). Ratios are author calculations from published percentages.

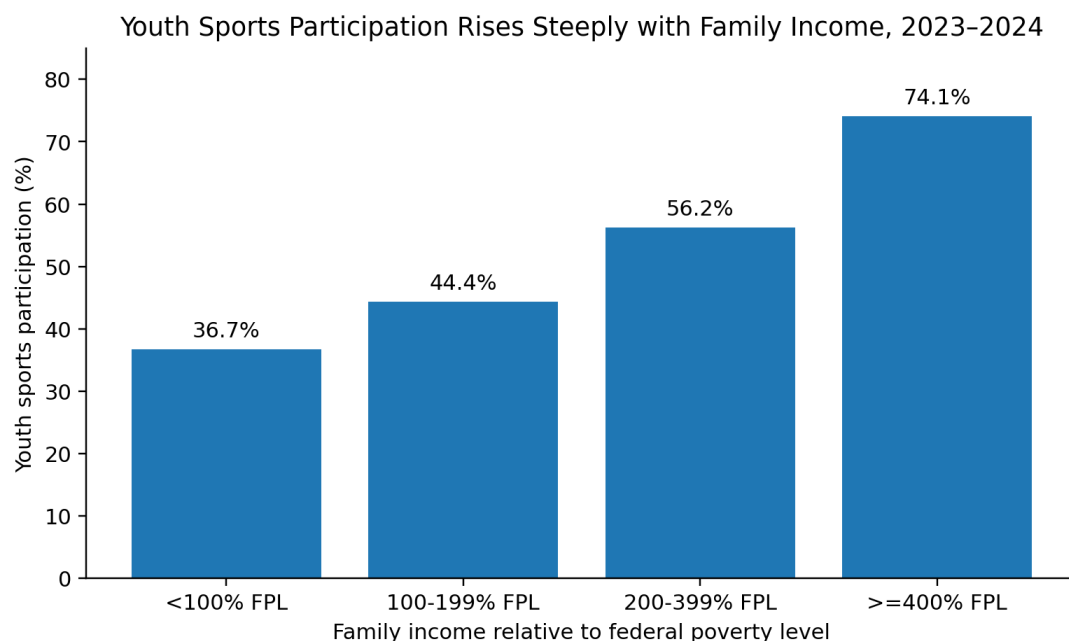


Figure 1. Youth sports participation by family income, 2023–2024. Percentages are reported in the HRSA NSCH Youth Sports Participation data brief; visualization is based on those published estimates.

4.2 Geographic variation

State-level participation varied by 24.6 percentage points, from 46.3% in Nevada to 70.9% in New Hampshire. Across the 50 states and District of Columbia, the unweighted median state estimate was 58.8%, the unweighted mean was 58.5%, the standard deviation was 6.0 percentage points, and the interquartile range was 53.1% to 63.5%. These state summaries should not be confused with the national weighted prevalence of 56.7% because each state receives equal weight in this descriptive calculation.

Regional grouping revealed a pronounced geographic pattern. The mean of state estimates was 63.1% in the Northeast and 63.0% in the Midwest, compared with 57.7% in the West and 53.6% in the South. On this equal-state basis, the South was approximately 9.5 percentage points below the Northeast and 9.4 points below the Midwest. This pattern is consistent with the independent 2020 NHIS report, which also identified the South as the lowest-participating Census region, although the present state summary is not a formal longitudinal comparison.

Table 2. Unweighted descriptive summaries of 2023–2024 state-level NSCH participation estimates by Census region.

Region	N	Mean %	Median %	Min %	Max %
Northeast	9	63.1	61.6	55.8	70.9
Midwest	12	63.0	63.7	56.9	66.3
South	17	53.6	52.7	48.2	65.8
West	13	57.7	58.8	46.3	64.5

Each state or DC receives equal weight.

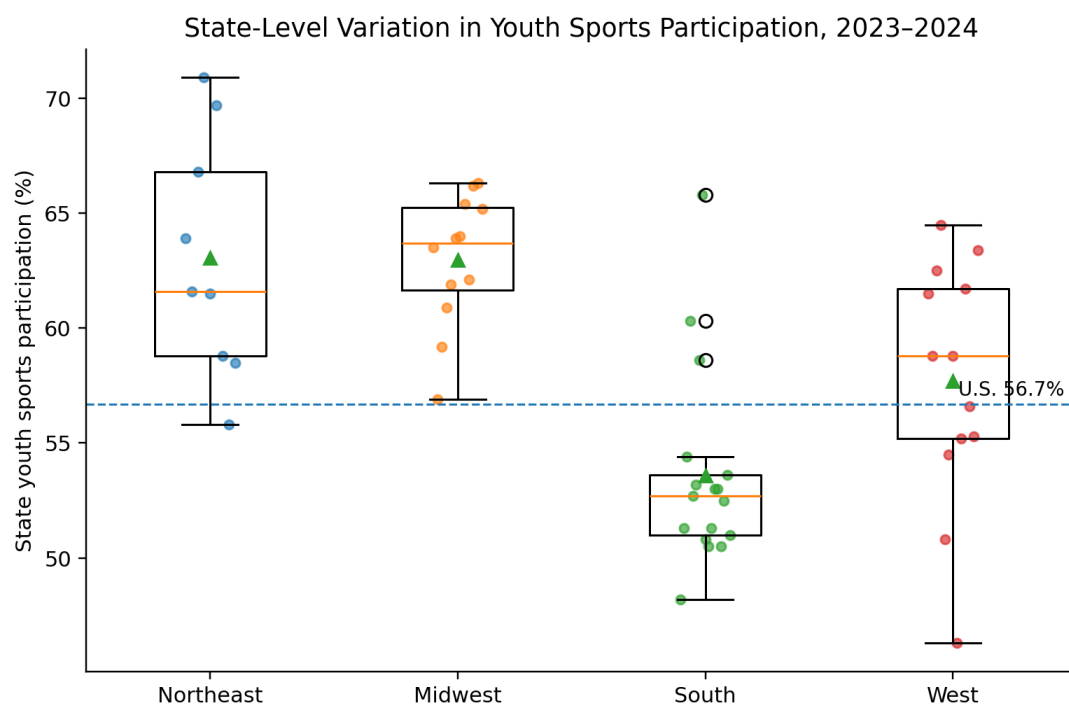


Figure 2. Distribution of state-level youth sports participation estimates by Census region, 2023–2024. Dots are state or District of Columbia estimates; triangles indicate equal-state regional means; the dashed line is the national survey-weighted estimate (56.7%).

4.3 Detailed income inequality in the 2020–2021 NSCH sensitivity analysis

The detailed 2020–2021 NSCH income table reproduced the same monotonic structure. Participation was 34.1% below 100% FPL, 38.8% at 100–199% FPL, 51.7% at 200–399% FPL, and 67.7% at $\geq 400\%$ FPL. Monte Carlo propagation of the published confidence intervals yielded a high-versus-low absolute gap of 33.6 percentage points (95% simulation interval 31.0–36.2), a prevalence ratio of 1.99 (1.86–2.13), and an odds ratio of 4.05 (3.61–4.57).

The rank-based indices showed an even larger distribution-wide contrast. The SII was 46.6 percentage points (43.5–49.6), meaning that the fitted linear socioeconomic rank gradient implied nearly a 47-point difference between the bottom and top of the income distribution. The logistic-rank RII was 2.59 (2.41–2.79), indicating that the model-predicted participation prevalence at the top of the socioeconomic

rank distribution was about 2.6 times the predicted prevalence at the bottom.

The descriptive population shortfall was 8.23 million youth (7.66–8.80 million). Of the point estimate, approximately 2.95 million came from the <100% FPL group, 3.01 million from the 100–199% group, and 2.28 million from the 200–399% group. This does not mean that 8.23 million children would enter sport if fees were eliminated. It means that, given the published population counts, that many additional participants would be observed if the lower-income strata displayed the same participation prevalence as the highest-income stratum.

Table 3. Derived inequality measures from the published 2020–2021 NSCH household-income table.

Measure	Point estimate	95% simulation interval
High-low absolute gap	33.6 pp	31.0 to 36.2 pp
Prevalence ratio	1.99	1.86 to 2.13
Odds ratio	4.05	3.61 to 4.57
Slope index of inequality	46.6 pp	43.5 to 49.6 pp
Relative inequality index	2.59	2.41 to 2.79
Equal-access descriptive shortfall	8.23 million	7.66 to 8.80 million

Simulation intervals use 200,000 Monte Carlo draws from the reported prevalence confidence intervals and do not reproduce the full complex-survey covariance structure.

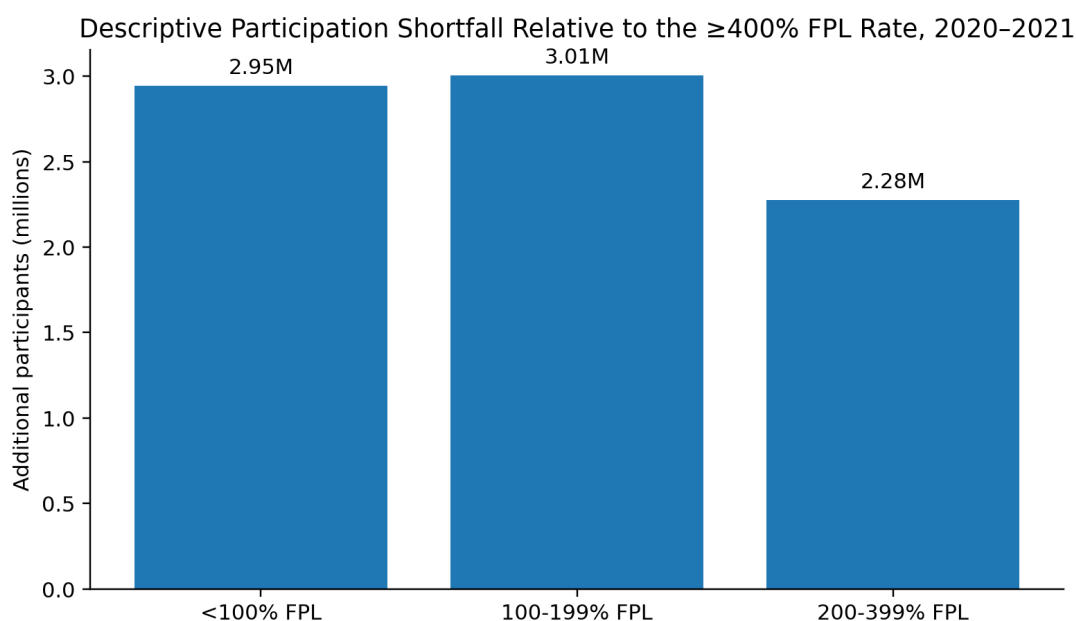


Figure 3. Descriptive participation shortfall by lower-income group relative to the $\geq 400\%$ FPL participation rate, 2020–2021. This is an inequality benchmark, not a causal policy forecast.

4.4 Independent replication in the 2020 NHIS

The independent NHIS produced a similarly large income gradient despite using a different federal survey. Sports participation was 31.2% below 100% FPL and 70.2% at $\geq 400\%$ FPL. The implied absolute gap was 39.0 percentage points. Using the published standard errors and Monte Carlo propagation, the approximate 95% simulation interval for the gap was 32.6–45.4 points. The corresponding prevalence ratio was 2.25 (1.89–2.76), and the odds ratio was 5.21 (3.89–7.12). The consistency of direction and magnitude across the NHIS and NSCH supports the conclusion that the income gradient is not an artifact of a single survey instrument.

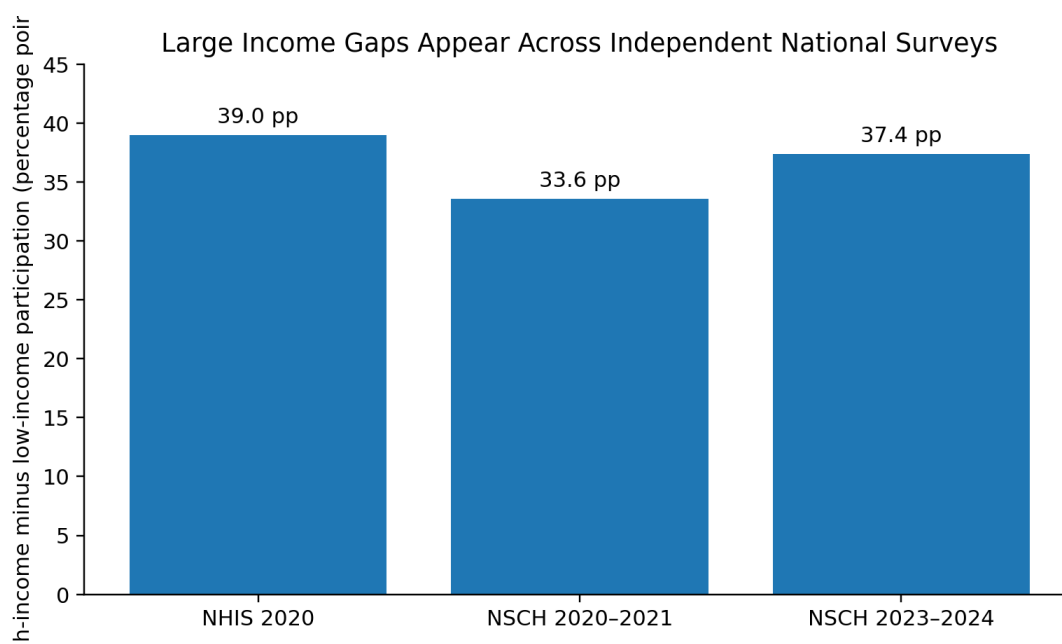


Figure 4. High-income minus low-income youth sports participation gaps across federal survey products. NHIS 2020 uses <100% vs. $\geq 400\%$ FPL; NSCH comparisons use the same endpoint categories. Values are descriptive and should not be read as a time series.

4.5 Sports participation and health-related correlates

The 2023–2024 HRSA brief also reports substantial cross-sectional differences between sports participants and nonparticipants. Participants were more likely to meet the 60-minutes-per-day physical activity guideline every day (21.7% vs. 17.0%; RD 4.7 points; PR 1.28) and were more than twice as likely to be active for at least 60 minutes on 4–6 days per week (37.4% vs. 16.7%; RD 20.7 points; PR 2.24). They were also more likely to have limited weekday recreational screen time (54.3% vs. 43.1%), adequate sleep (68.9% vs. 60.3%), and no difficulty making or keeping friends (81.3% vs. 67.8%).

Several health conditions were less common among participants: obesity (12.7% vs. 20.9%), current anxiety (10.9% vs. 16.4%), and current depression (3.3% vs. 7.6%). These associations increase the policy relevance of unequal access, but they cannot be interpreted as benefits caused by sport. Selection into sport is nonrandom, and health status itself can affect participation. The federal brief explicitly cautions against causal interpretation.

Table 4. Crude cross-sectional outcome contrasts from the 2023–2024 HRSA NSCH data brief.

Outcome	Sports %	No sports %	Difference (pp)	Prevalence ratio
Daily 60+ min physical activity	21.7	17.0	+4.7	1.28
60+ min activity on 4–6 days	37.4	16.7	+20.7	2.24
Limited weekday screen time	54.3	43.1	+11.2	1.26
Adequate sleep	68.9	60.3	+8.6	1.14
No difficulty making/keeping friends	81.3	67.8	+13.5	1.20
Obesity	12.7	20.9	−8.2	0.61
Current anxiety	10.9	16.4	−5.5	0.66
Current depression	3.3	7.6	−4.3	0.43

Ratios and differences are author calculations. These are associations, not causal effects of sports participation.

5. Discussion

5.1 Main finding: economic inequality is not a marginal feature of youth sports

The central finding is straightforward: family income is associated with a very large difference in organized youth sports participation. In the most recent 2023–2024 federal estimates, participation among

youth in the highest income category was roughly double the prevalence among youth below the poverty threshold. The 37.4-point income gap was more than four times larger than either the sex gap or the maximum age-group gap reported in the same data brief. Independent NHIS data from 2020 and the detailed 2020–2021 NSCH table produced similarly large endpoint contrasts.

This result matters because demographic discussions of youth sports often focus on sex, age, or athletic interest while treating household resources as one covariate among many. The scale comparison suggests the opposite framing may be warranted: economic access is a first-order feature of participation. The data do not imply that income is the only important determinant, but they show that participation is strongly sorted along an economic dimension.

The international literature supports this interpretation. Owen et al. (2022) found a pooled socioeconomic difference in sport participation across high-income countries, and the effect was larger for organized sport than for physical activity overall. The present national U.S. effect sizes are larger than the review's pooled OR of 1.87, although direct comparison is imperfect because studies use different socioeconomic measures, reference groups, ages, and sport definitions. The implication is not that the United States is necessarily more unequal than every other setting. Rather, it is that contemporary U.S. federal data sit comfortably within a broader evidence base in which structured sport is socioeconomically stratified.

5.2 Why income may matter: price, complements, and cumulative advantage

An economic interpretation helps avoid the overly simple statement that low-income families 'do not value sports.' Participation requires a bundle of complementary inputs. A registration subsidy has limited value if a family lacks transportation to practice; free equipment has limited value if practice conflicts with work schedules or school obligations; and a nearby program may remain inaccessible if selection is based on skills developed through years of prior paid participation. Income may therefore relax several such constraints at once, although these national data do not measure fees, transportation, or time costs directly.

The steepest step in the 2023–2024 income gradient occurred between the 200–399% and $\geq 400\%$ FPL groups, where participation increased by 17.9 percentage points. That pattern is consistent with a system in which upper-income households may be able to afford not only basic participation but also the complementary resources that make sustained participation easier. It would be premature to infer a threshold effect from four categories, yet the gradient does not look like a simple poverty-versus-nonpoverty divide. Inequality extends well into the middle of the income distribution.

Cumulative advantage provides a second mechanism. Early access may build sport-specific human capital. Skills, confidence, knowledge of teams and leagues, relationships with coaches, and parental familiarity with enrollment systems can all reduce the cost of later participation. This means that an inequality observed among teenagers may partly reflect access decisions made years earlier. From a policy perspective, low-cost introductory opportunities at younger ages may have long-run value even if later competitive programs remain selective.

5.3 Geographic variation and the possible role of local institutions

Income differences do not account for the entire pattern. State participation ranged by nearly 25 percentage points in 2023–2024. The equal-state regional summaries show that the South had substantially lower typical state estimates than the Northeast and Midwest. The 2020 NHIS independently reported lower sports participation in the South than in other Census regions. This repeated geographic pattern suggests that local sports infrastructure, climate, school systems, cultural norms, transportation, program availability, and other state or regional institutions may shape access alongside household resources.

The geographic findings should be interpreted cautiously. State estimates have different sampling precision, and the regional means in this paper are not population-weighted. The analysis therefore cannot rank regional policy performance or identify why particular states are high or low. It does, however, suggest that an exclusively household-level model may be incomplete. A youth sports market is produced

jointly by families, schools, nonprofit organizations, private clubs, parks departments, transportation systems, and local governments.

5.4 Policy implications: from equal fees to equal effective access

The most direct policy response to an income gradient is to lower the monetary price faced by families. Existing intervention evidence supports this mechanism. The Kentucky YES Card evaluation found substantially higher odds of sports participation among voucher recipients, and the Welsh ACTIVE trial showed benefits from a package combining vouchers with peer mentoring and local support. These studies are not identical to the U.S. national population and should not be mechanically applied to the national gap, but they suggest that price and access interventions can alter behavior.

A stronger policy principle is to target the effective price rather than only the posted fee. Four policy components follow from the economic framework and the National Youth Sports Strategy:

- Targeted or sliding-scale fee subsidies. Because the national gradient extends across several income categories, eligibility rules based only on extreme poverty may miss families that still face meaningful affordability constraints.
- Equipment libraries and low-cost starter programs. Shared equipment can reduce upfront costs and allow children to sample sports before a family commits to costly specialized equipment.
- School-based and neighborhood-based provision. Locating programs where youth already spend time can reduce transportation and parental-time costs. The ACTIVE trial's qualitative findings similarly emphasized local and informal opportunities.
- Open-entry recreational pathways alongside competitive teams. Competitive programs can serve performance goals, but a public-access objective requires options that do not condition entry on skills that themselves reflect prior opportunity.

The policy objective should also be stated carefully. Equal participation rates across income groups are not necessarily required for equity because preferences may differ. The problem arises when resource constraints prevent youth from acting on preferences they would otherwise hold. National survey data cannot identify that latent demand directly. For this reason, the 8.23-million shortfall should be treated as an inequality benchmark rather than a target number for government enrollment.

A more modest benchmark illustrates the scale of the challenge. In 2023–2024, halving the 37.4-point high-versus-low income gap would require the lowest-income participation rate to rise by 18.7 points, from 36.7% to 55.4%, which is close to the current national average. Whether that change is feasible, cost-effective, or desirable cannot be answered from prevalence data alone. It identifies the magnitude that an intervention portfolio would have to address.

5.5 Relevance beyond the United States

Although the empirical analysis is U.S.-based, the socioeconomic mechanism is relevant to other education systems. Zheng's (2025) Shanghai study found a significant association between SES and organized sports participation and showed that parental exercise and parental support mediated part of the relationship. The Chinese evidence cautions against exporting a U.S. fee-subsidy model without adaptation. In settings where academic time pressure, parental expectations, or school structure are especially important, the opportunity cost of time may be as important as direct monetary price.

For public policy research, this creates a useful comparative agenda. Future studies could ask whether income gradients in China, the United States, and other countries are produced by the same mechanisms. A similar prevalence gap can arise from very different institutions. Identifying those mechanisms is necessary before policy can move from documenting inequality to designing effective interventions.

5.6 Strengths and limitations

This study has several strengths. It centers the newest available federal NSCH data, triangulates the primary finding with an independent national survey, reports effect sizes rather than only statistical significance, quantifies inequality on both absolute and relative scales, and distinguishes descriptive coun-

terfactuals from causal effects. The analysis is also fully reproducible from public survey products; it does not depend on inaccessible proprietary data.

The limitations are equally important. First, the primary 2023–2024 HRSA data brief publishes subgroup point estimates but not the full covariance matrix or microdata model results. The current-year effect sizes are therefore descriptive. Second, the SII, RII, and population shortfall use an older 2020–2021 NSCH table because that source publishes confidence intervals and weighted population counts by income category. Weighting revisions and the pandemic period mean those results should not be read as a time trend. Third, Monte Carlo intervals assume approximate normality and do not recover correlations among survey estimates. Fourth, income is not randomly assigned. The observed gradient combines monetary resources with parental education, transportation, neighborhood resources, health, race and ethnicity, family structure, and other correlated factors. Fifth, organized sports participation is measured as any team or lesson in the past 12 months and does not capture intensity, cost, quality, duration, or the type of sport.

Finally, the state analysis is ecological. It describes where participation is high or low but cannot identify state policy effects. A state with high participation may have favorable demographics rather than superior sports policy, and a low-participation state may have barriers not observed in the public brief. Causal policy evaluation requires individual-level microdata, natural experiments, randomized interventions, or credible quasi-experimental designs.

6. Conclusion

U.S. youth sports participation is strongly stratified by family income. In 2023–2024, fewer than four in ten youth below the federal poverty threshold participated in a sports team or sports lessons, compared with nearly three in four youth in families at 400% or more of the poverty threshold. The 37.4-percentage-point income gap was more than four times larger than the sex and age-group gaps reported in the same national brief. The pattern also appeared in independent federal survey data, while the detailed NSCH sensitivity analysis indicated large absolute and relative inequality across the full income distribution.

These results do not prove that a particular subsidy would close the gap. They show that organized sport is not distributed independently of household resources. The policy implication is therefore broader than ‘make sports cheaper.’ Effective access depends on the combined price of fees, equipment, transportation, time, location, information, and program structure. Evidence from voucher and school-based interventions suggests that reducing those constraints can change participation, but future research should test which combinations generate the largest and most equitable effects.

From an economics perspective, the key issue is constrained choice. If children with similar interest face systematically different feasible sets because of family resources or local institutions, youth sports may become a mechanism through which economic inequality is reproduced in extracurricular life. Measuring that inequality is only the first step. The next is to identify which constraints are both modifiable and worth addressing with limited public and school resources.

Data and Reproducibility Statement

All numerical inputs used in this paper come from publicly available federal survey products listed in the references. The 2023–2024 primary estimates are from the HRSA Maternal and Child Health Bureau NSCH Youth Sports Participation data brief. The 2020–2021 income table is from the Child and Adolescent Health Measurement Initiative Data Resource Center. The 2020 replication estimates are from NCHS Data Brief No. 441. Derived calculations include risk differences, prevalence ratios, odds ratios, ridit ranks, SII/RII estimates, state-level descriptive summaries, and Monte Carlo uncertainty propagation. No individual-level records were accessed and no new human-subject data were collected.

References

- Black, L. I., Terlizzi, E. P., and Vahratian, A. 2022. *Organized Sports Participation among Children Aged 6–17 Years: United States, 2020*. NCHS Data Brief 441. National Center for Health Statistics. <https://doi.org/10.15620/cdc:119026>.
- Child and Adolescent Health Measurement Initiative. 2026. “2020–2021 National Survey of Children’s Health Data Query: Participation in Sports Teams or Sports Lessons after School or on Weekends, Age 6–17 Years, by Household Income Level.” Data Resource Center for Child and Adolescent Health. <https://www.childhealthdata.org/>.
- Eime, R. M., Young, J. A., Harvey, J. T., Charity, M. J., and Payne, W. R. 2013. “A Systematic Review of the Psychological and Social Benefits of Participation in Sport for Children and Adolescents: Informing Development of a Conceptual Model of Health through Sport.” *International Journal of Behavioral Nutrition and Physical Activity* 10: 98. <https://doi.org/10.1186/1479-5868-10-98>.
- James, M. L., Christian, D., Scott, S. C., et al. 2020. “Active Children Through Individual Vouchers Evaluation: A Mixed-Method RCT.” *American Journal of Preventive Medicine* 58 (2): 232–43. <https://doi.org/10.1016/j.amepre.2019.10.005>.
- Meyers, C. C. A., Mann, M. J., Thorisdottir, I. E., et al. 2023. “Examining the Impact of a Leisure Time Intervention on Participation in Organized Out-of-School Activities among Adolescents: A Quasi-Experimental Study in Franklin County, KY, USA.” *Health Education Research* 38 (4): 320–28. <https://doi.org/10.1093/her/cyad016>.
- Owen, K. B., Nau, T., Reece, L. J., et al. 2022. “Fair Play? Participation Equity in Organised Sport and Physical Activity among Children and Adolescents in High Income Countries: A Systematic Review and Meta-Analysis.” *International Journal of Behavioral Nutrition and Physical Activity* 19: 27. <https://doi.org/10.1186/s12966-022-01263-7>.
- Sappenfield, O., Kenney, M. K., and Lebrun-Harris, L. 2026. *National Survey of Children’s Health: Youth Sports Participation, 2023–2024*. Data brief, July. Health Resources and Services Administration, Maternal and Child Health Bureau.
- Tandon, P. S., Kroshus, E., Olsen, K., Garrett, K., Qu, P., and McCleery, J. 2021. “Socioeconomic Inequities in Youth Participation in Physical Activity and Sports.” *International Journal of Environmental Research and Public Health* 18 (13): 6946. <https://doi.org/10.3390/ijerph18136946>.
- U.S. Department of Health and Human Services. 2019. *National Youth Sports Strategy*. Office of Disease Prevention and Health Promotion. <https://health.gov/our-work/nutrition-physical-activity/national-youth-sports-strategy>.
- Zheng, N. 2025. “The Relationship between Socioeconomic Status and Organized Sports Participation among Chinese Children and Adolescents: The Chain-Mediated Role of Parental Physical Exercise and Parental Support.” *BMC Public Health* 25: 1945. <https://doi.org/10.1186/s12889-025-23187-0>.

A. Published Inputs for the 2020–2021 NSCH Inequality Analysis

The following values were transcribed from the Data Resource Center's national 2020–2021 NSCH query for K7Q30, stratified by household income. Population estimates are weighted to represent U.S. children ages 6–17 in the measure denominator.

Table 5. Published 2020–2021 NSCH inputs by household income.

Income group	Yes %	95% CI	Sample yes	Sample no	Weighted population	Ridit
<100% FPL	34.1	31.9–36.5	2,902	4,727	8,765,585	0.091
100–199% FPL	38.8	36.7–41.0	4,318	5,680	10,398,958	0.288
200–399% FPL	51.7	50.1–53.4	9,881	8,325	14,227,945	0.543
≥400% FPL	67.7	66.4–68.9	16,541	7,532	15,034,149	0.845

Note. Ridit ranks are author calculations from weighted group population shares. The DRC notes that income missing values for this archived subgroup were imputed within year and that displayed estimates are based on the first implicate.

B. State-Level Youth Sports Participation, 2023–2024

Values below are the published NSCH state estimates shown in the HRSA July 2026 youth sports data brief. They are included to make the regional descriptive analysis auditable.

Table 6. State-level youth sports participation, 2023–2024.

Jurisdiction	%	Jurisdiction	%
Alabama (AL)	50.5	Alaska (AK)	55.3
Arizona (AZ)	54.5	Arkansas (AR)	50.5
California (CA)	56.6	Colorado (CO)	64.5
Connecticut (CT)	63.9	Delaware (DE)	48.2
District of Columbia (DC)	65.8	Florida (FL)	51.3
Georgia (GA)	54.4	Hawaii (HI)	63.4
Idaho (ID)	58.8	Illinois (IL)	63.5
Indiana (IN)	56.9	Iowa (IA)	63.9
Kansas (KS)	64.0	Kentucky (KY)	51.3
Louisiana (LA)	52.5	Maine (ME)	61.6
Maryland (MD)	60.3	Massachusetts (MA)	66.8
Michigan (MI)	59.2	Minnesota (MN)	65.4
Mississippi (MS)	53.0	Missouri (MO)	62.1
Montana (MT)	61.5	Nebraska (NE)	66.2
Nevada (NV)	46.3	New Hampshire (NH)	70.9
New Jersey (NJ)	58.8	New Mexico (NM)	50.8
New York (NY)	55.8	North Carolina (NC)	52.7
North Dakota (ND)	66.3	Ohio (OH)	60.9
Oklahoma (OK)	53.6	Oregon (OR)	61.7
Pennsylvania (PA)	58.5	Rhode Island (RI)	61.5
South Carolina (SC)	53.2	South Dakota (SD)	65.2
Tennessee (TN)	53.0	Texas (TX)	51.0
Utah (UT)	58.8	Vermont (VT)	69.7
Virginia (VA)	58.6	Washington (WA)	55.2
West Virginia (WV)	50.8	Wisconsin (WI)	61.9
Wyoming (WY)	62.5		

C. Reproducibility Formulas

High-versus-low absolute gap: $RD = p_H - p_L$.

Prevalence ratio: $PR = p_H / p_L$.

Odds ratio: $OR = [p_H / (1 - p_H)] / [p_L / (1 - p_L)]$.

Ridit rank for group g : $r_g = \text{cumulative population share below group } g + 0.5 \times \text{group } g \text{ population share}$.

Slope index of inequality: coefficient on r_g from a population-weighted linear regression of group prevalence on ridit rank.

Relative inequality index used here: predicted prevalence at ridit = 1 divided by predicted prevalence at ridit = 0 from a population-weighted logistic-rank model.

Descriptive participation shortfall: $\sum_g N_g(p_H - p_g)$ over lower-income groups. Values below zero, if any, would be set to zero; none occurred in the observed table.

Monte Carlo intervals: 200,000 draws from normal distributions centered on each published prevalence, with standard errors approximated from reported 95% confidence intervals as $(\text{upper} - \text{lower}) / 3.92$. Population counts and ridit ranks were held fixed across draws.