Master Thesis Strategy Economics

Stimulating Participation in Sports: An Empirical Analysis of the Role of Dutch Municipalities

Abstract

Dutch municipalities spend approximately 1.5 billion EUR on sports per year. They aim to keep sports affordable and accessible for everyone and hope to stimulate inhabitants to start exercising. Yet, some municipalities spend more on sports than others. Using data on municipal expenditures from 2017 to 2020, I investigate whether sports spending is positively related to the sports participation rate in 2020. Ordinary Least Squares regression estimates do not provide evidence for a significant association. Additionally, the estimates of a Two-Stage Least Squares regression with the presence of a historic city center in the municipality as the instrument are insignificant too. These results are robust against the chosen time frame and a different measure of sports participation.

Author:
Tygo Beekman (453540)

Supervisor:
S.D.T. Hoey

Second Assessor:

Erasmus School of Economics

Contents

1 Introduction

Physical activity is widely recognized to have substantial health benefits (World Health Organization, 2020). Moreover, as described in the Human Capital Model of Bailey et al. (2013), exercising also adds to people's emotional, social, and financial capital. On an aggregate level, these benefits can be of great value to society. Therefore, governments often are involved in providing opportunities for people to engage in sports.

This is certainly the case in the Netherlands, as the total public sports expenditures added up to almost 2.5 billion EUR in 2018 alone (KPMG, 2019). The vast majority of these expenditures come from municipalities. They spend it on sports facilities, activation policies, and public parks and open-air recreation. Yet, local governments can freely choose how much of their budget they allocate to sports. This leads to heterogeneity in the local public sports expenditures per capita. Given that higher per capita spending likely leads to more and better opportunities to exercise, and this, in turn, increases the probability someone is physically active, one would therefore expect to see differences in local sports participation accordingly.

The goal of this thesis is to investigate the legitimacy of this proposition. Dutch municipalities spend large sums of public funds on sports with the justification that they yield large societal benefits because of a more active and healthy population. Hence, it is important to know whether these public sports expenditures indeed contribute to more participation in sports. This is especially relevant in the current debate about the organization and financing of the Dutch sports sector. A prominent topic in this discussion is whether sports should be financed publicly or privately (Netherlands Sports Council, 2020). Additional insight into the impact of public sports expenditures thus may help to make an informed decision on this matter. Lastly, with the rising popularity of unorganized outdoor sports, it is of interest to policymakers to gain an understanding of how public spending on parks and open-air recreation compares to sports expenditures in terms of activating citizens. The research question I aim to answer is, therefore, as follows:

What is the influence of local public expenditures on regular physical activity in Dutch municipalities?

The existing empirical literature finds no solid evidence for a positive influence of public expenditures on sports participation. Earlier research on the association between municipal sports expenditures and sports participation in the Netherlands reports a significant but weak correlation (Van den Dool & Hoekman, 2017). Another study finds that municipal spending on sports is positively related to the sports participation of Dutch children, but not to those of Dutch adults (Hoekman et al., 2017). In Germany, Dallmeyer et al. (2017) identify a significant relationship between spending on sports facilities by the regional government and regular sports participation. Yet, Steckenleiter et al. (2019) do not find evidence for an effect of local governments' sports expenditures on the propensity to participate in sports among German adults. Likewise, Kokolakakis et al. (2014) show that local governments' sports expenditures do not correlate with adult sports participation in England. However, a common shortcoming in these studies is that they do not take into account the potential simultaneous effect of sports participation on sports expenditures.

This thesis addresses the issue of simultaneity by employing a Two-Stage Least Squares (2SLS) regression. I use the presence of a historical city center in a municipality as an instrument for municipal sports expenditures. If a municipality has a historical center, it receives more money from the municipal fund. This creates seemingly random variation in cities' sports budgets. Hence, I am able to obtain unbiased estimations if the instrument assumptions hold. I present the 2SLS estimates alongside the outcomes of Ordinary Least Squares (OLS) regressions.

I perform the analysis for both sports and exercise participation. The influence of municipal sports spending on the propensity to engage in active non-sports activities has not been extensively studied in the empirical literature. Yet, it is interesting to see whether public sports expenditures contribute to a more active population regardless of their participation in sports. Therefore, I use the share of the population with an active lifestyle as an alternative dependent variable to estimate the relationship between municipal sports expenditures and exercise participation.

In the analysis, I distinguish between different types of sports expenditures. Besides the total sports expenditures per capita, I consider spending on sports facilities and sports activation policies. This allows me to assess whether spending on policies aimed at stimulating people to play sports has a different effect on sports participation than investing in sports infrastructure. Additionally, I include municipal spending on public parks and recreation to investigate how the quality of public sports facilities affects sports participation.

Using municipal sports expenditure data from 2017 to 2020, I find no significant relationship between the average per capita sports expenditures and the rate of sports participation in Dutch municipalities in 2020. Similarly, the analysis suggests there is no correlation between spending on activation policies and sports participation. I do find a significant association between sports facilities expenditures and participation in sports, especially when I exclude data on the expenditures in 2020. Yet, the economic impact of this association is negligible. Spending on public parks and recreation does not appear to significantly correlate with sports participation. These results are robust to the expenditure period I consider and to a different measure of sports participation. The analysis for exercise participation shows similar results. Therefore, I conclude that there is no evidence of a relationship between municipal sports expenditures and the rate of sports participation in Dutch municipalities.

The structure of this thesis is as follows. First, I construct a theoretical framework by discussing the relevant theoretical and empirical literature. Second, I describe the data that I use to perform the analysis. Next, I explain the empirical approach. Then, I present the outcomes of my analysis and provide a discussion on the main results. The last section concludes.

2 Theoretical Framework

In this section, I set up a theoretical framework for understanding why municipal sports expenditures may influence sports and exercise participation. To do so, I first discuss the sports participation decision of individuals. I then explain the role of the government in the sports sector and how this may affect sports participation. Lastly, I review the extant empirical literature regarding the relationship between public sports expenditures and sports participation.

2.1 The Sports Participation Decision

According to Downward et al. (2009), a consumer's decision to exercise or to take part in sports depends on the interaction between their motives and their available resources. First, the consumer's motives are formed by a combination of personal preferences and objectives and so provide a set of possible actions for the consumer. Then, their resources, such as income and time, identify the feasibility of these actions and eventually determine the behavior of the consumer.

A popular way to formalize this decision-making process is Becker's (1965) Household Production Function. In this model, individuals are part of a household that generates commodities Z through activities performed by the household (Heckman, 2015). Examples of such activities are baking a cake, child-rearing, or participating in sports. Consumption of the commodities yields utility to the household through the following function:

$$U = U(Z_1, Z_2, ..., Z_I) \tag{2.1}$$

Production of a commodity Z_i is a function of a vector of goods purchased on the market X_i and time t_i, and can be expressed as:

$$Z_i = Z(X_i, t_i), \text{ with } i = 1, ..., I \tag{2.2}$$

The price of X_i is a composite price and depends on the underlying prices of the necessary goods, as shown in the next equation:

$$P_i = p_1x_1 + p_2x_2 + ... + p_Jx_J, \text{ with } j = 1, ..., J \tag{2.3}$$

where P_i represents the composite price of goods required for the production of commodity Z_i, x_j is a single market good, and p_j stands for the price of a good. A consumer can buy goods with earnings obtained from working outside the household. Yet, time is finite and the consumer thus faces a time constraint:

$$T = W + H \tag{2.4}$$

Here, T is total time, W denotes the hours worked outside the household, and H represents the time spent on household production. Provided the wage rate w, it follows that a consumer's earnings E are given by:

$$E = w(T - H) \tag{2.5}$$

Hence, in maximizing utility households are constrained by equation 2.2, equation 2.4, and by:

$$\sum_{i=1}^{I} P_i = E \tag{2.6}$$

From this, we can derive four determinants of the demand for sports (Taylor & Gratton, 2002). Firstly, the composite price of sports affects the participation decision. This composite price consists of all costs involved in playing sports, like sports club membership costs, entrance fees for sports facilities, costs of sports clothing and equipment, and transportation costs. Keeping all else constant, a lower composite price implies playing sports becomes cheaper relative to other commodities and is thus expected to increase sports participation.

Secondly, there is a relationship between wage and sports participation. A rise in the real wage leads to an increase in the price of time relative to the price of goods. This can cause a change in household production patterns through substitution and income effects. The former implies that an individual starts working more when the wage increases because of the higher opportunity costs of time spent on leisure activities. Consequently, household production of time-intensive commodities will fall and shift to goods-intensive commodities. The latter effect states that an individual will work less as the wage rises, as fewer hours of work are required to obtain the same income level. This induces the household to produce more time-intensive commodities instead of goods-intensive ones. Due to these conflicting effects, it is unclear how household production patterns react to a higher wage rate. However, Taylor and Gratton (2002) hypothesize that for sports the income effect likely dominates, since research suggests sports is a superior good and many sports activities take up a fixed amount of time. Hence, a higher real wage is expected to increase sports participation.

Thirdly, the demand for sports is affected by the prices of other goods. In the case of a complementary good, like healthy food, this implies that sports participation goes up as the price of the complementary good decreases. For substitution goods, such as cultural activities, sports participation declines when the substitution good becomes cheaper.

Lastly, consumer preferences play a role in the demand for sports. As preferences are unobservable, it is impossible to predict how they affect the sports participation decision at an individual level. However, we can infer general information about preferences based on personal characteristics because these often correlate with preferences for sports. For instance, gender (Van Tuyckom et al., 2010; Fridberg, 2010; Kokolakakis et al., 2012), age (Hovemann & Wicker, 2009; Hoekman et al., 2017), and educational background (Studer et al., 2011; Hoekman et al., 2017) have all been found to be associated with sports participation. Hence, it is important to take into account such variables when explaining differences in sports participation.

A noteworthy extension to the household production approach is Grossman's (1972) model for the demand for health. In this model, health status is seen as a part of human capital and, therefore, a determinant of income (Taylor & Gratton, 2002). Since exercising improves health, playing sports can thus be seen as an investment in human capital that yields return in the form of higher earnings. So, individuals may not only participate in sports because exercising itself generates utility but also because they want to invest in their human capital.

2.2 The Government's Role in the Sports Sector

2.2.1 Rationale for Government Intervention

Governments intervene in the sports sector for two main reasons. Firstly, governments try to improve efficiency by tackling the market failures that arise in the market for sports. A known market failure in the sports industry is the presence of positive externalities (Downward et al., 2009). These externalities arise because playing sports does not only yield a benefit to the consumer but also has a positive effect on society. For instance, exercising contributes to a healthier population and so decreases health care costs. It can also increase labor productivity, leading to more economic growth. Yet, individual consumers do not reap the benefits of these externalities and will thus not take them into account in the sports participation decision. Therefore, the amount of sports consumed will be lower than the social optimum.

Another market failure present in the market for sports is imperfect information, as consumers cannot fully recognize all the personal benefits of exercising (Downward et al., 2009). For example, it is impossible to know the net present value of the effect of playing sports on future income. Hence, the perceived benefits of exercising are lower than the actual benefits. This results in a lower demand for sports than what would be optimal for the consumer. A paternalistic government can thus aim to promote sports participation if they believe that this is in the consumer's best interest.

Secondly, governments may desire a more equal distribution of the benefits of sports to compensate for the distribution of income (Taylor & Gratton, 2002). This argument assumes that there is a group that does not have the resources to participate in sports, but would be playing sports if they did. This implies that they do not have access to the private benefits of exercising, which can exacerbate inequalities in the existing income distribution. Therefore, governments can adopt policies that make sports available for individuals with a low income to ensure that they have the same access to sports as individuals with a higher income.

2.2.2 Types of Government Intervention in the Netherlands

Governments intervene on both the demand and supply side to stimulate sports participation. On the demand side, two main kinds of policies can be identified. First, governments address the issue of imperfect information by raising awareness about the personal benefits of exercising. In the Netherlands, this is mainly done by the national government as a part of a broader strategy to encourage a healthy lifestyle (Ministry of Health, Welfare and Sport, 2018).

Second, governments help to overcome resource constraints by providing subsidies to targeted groups. Dutch municipalities primarily do this through a local or regional Youth Sports and Culture Fund. Such a fund provides in-kind subsidies for sports club memberships and sports gear to children living in low-income households (Van den Berg, 2021). Similar funds exist for adults with an income close to the social minimum, but these funds are present only in a limited amount of municipalities. Additionally, municipalities may offer subsidies to groups that need extra assistance to be able to exercise. They, therefore, face a higher cost of playing sports which may prevent them to exercise. Examples of such groups are the disabled or the elderly.

There are also two types of policies that aim to promote sports participation through the supply side. First, sports facilities can be funded, maintained, or managed by the government. In the Netherlands, local governments to large extent responsible take up this responsibility, even though there is no legal obligation to do (Hoekman, 2018). Municipalities own the majority of sports halls, a vast amount of swimming pools, and a large share of outdoor sports facilities. By renting these facilities out to sports clubs for a price lower than the cost (Schadenberg & Van Eldert, 2020), they ensure that sports clubs have access to affordable facilities. Sports clubs can then lower their membership fees, which makes becoming a member of a sports club more accessible for individuals with low incomes. Furthermore, by providing a sufficient amount of well-maintained sports facilities governments also make exercising more appealing in general. Besides the sports accommodation expenditures, Dutch municipalities also spend a considerable amount on public parks and open-air recreation. This, for instance, includes the construction and maintenance of bike paths and public gym equipment. Hence, on top of the sports accommodation expenditures, also spending on public parks and recreation can make exercising more attractive and accessible.

Second, sports participation can be stimulated by developing delivery staff such as sports coaches. Dutch municipalities do this by funding so-called Neighborhood Sports Coaches (NSCs). These NSCs help people that have difficulties starting to play sports or increase their frequency of exercising (Association of Sports and Municipalities, n.d.). In addition, they are usually affiliated with an organization outside of the sports sector as well. This way they can encourage individuals to take up sports and identify what barriers to exercise they experience.

2.3 Sports Participation and Public Expenditures in Empirical Literature

A growing body of literature empirically analyzes the effect of public sports expenditures on sports participation. To date, two studies have examined this relationship in the Netherlands. First, Van den Dool and Hoekman (2017) investigated whether municipal sports spending is related to the share of the municipality's population that plays sports in municipality-funded sports facilities. Using data from 2010 to 2016, their OLS regression identifies a statistically significant positive correlation between sports expenditures per capita and sports participation. Yet, sports expenditures only explain a minor part of the variation in participation rates, suggesting that there is only a weak relationship. Also, the analysis does not account for a potential simultaneous effect of sports participation on sports spending, so the results cannot be interpreted as causal. Additionally, the study examines how sports expenditures in 2010 and 2011 relate to participation rates over time. By categorizing the municipalities based on their per capita sports expenditures and comparing how sports participation changes within each group, they find that the high expenditure group has a higher average participation rate than the low expenditure group in 2012 and 2013. However, after 2014 this difference disappears, suggesting that the influence of municipal sports expenditures may diminish over time.

Second, Hoekman et al. (2017) studied whether local public sports expenditures affect sports participation and sports club membership of individuals in the Netherlands. The authors employ a 3-level logistic regression model to estimate these relationships, as they argue that sports behavior depends on individual and environmental factors through a hierarchical structure. Their results indicate that municipal sports spending has a significant positive impact on the sports participation and sports club membership of children aged 6 to 17, but has no significant effect on the sports behavior of adults between the ages of 25 and 79. A plausible explanation for these contrasting outcomes is that municipal sports policies in Dutch municipalities are often aimed at children.

The relationship between public sports spending and sports participation has also been investigated outside of the Netherlands. In Germany, Dallmeyer et al. (2017) have looked at the influence of various types of federal governments' expenditures on regular sports participation of individuals between 2003 and 2011. The results of their probit model point towards a significant association between spending on sports facilities and pools and the propensity to participate, but do not provide evidence that spending on sports promotion leads to more sports participation. Moreover, they find that public transport and street infrastructure expenditures positively influence sports participation, whereas spending on education, culture, and the natural environment does not have a significant effect. However, the analysis fails to account for the fact that individuals who like to play sports may prefer to live in states with good sports facilities. Hence, these results are potentially biased.

Another German study by Steckenleiter et al. (2019) examines whether municipal sports expenditures increase individual sports participation. They do control for self-selection into municipalities that satisfy people's preferences by measuring personal and municipality characteristics before the sports budget is determined and the participation decision is made. Also, they include sports spending by neighboring municipalities through a distance-weighted function to account for potential spillover effects. Using a sample of individuals between the ages of 17 and 65, Steckenleiter et al. (2019) then estimate dose-response relationships to assess the effect of sports facilities expenditures for different levels of expenditures. Their results suggest there is no significant effect of sports facilities expenditures on the probability to play sports at any expenditure level.

Lastly, Kokolakakis et al. (2014) include local governments' sports spending as a variable in their research on adult sports participation in England. By estimating a beta regression they find no significant association between the total sports expenditures between 2007 and 2010 and the sports participation rate in 2011. Likewise, the density of sports facilities in the area is not significantly correlated with the proportion of the population that regularly plays sports. Instead, they conclude socio-economic variables are more important determinants of sports participation at a local level in England. Yet, again these findings may be biased as the authors do not address the potential influence of sports participation on local governments' sports spending.

Overall, the empirical literature points towards no or a weak effect of local governments' sports expenditures on sports participation. However, the results do appear to depend on the research context. Especially important is the type of sports expenditures that is investigated and the distinction between youth and adult sports participation. Furthermore, the existing literature demonstrates that identifying causal relationships is difficult, implying that findings should be interpreted carefully.

2.4 Hypotheses

The theoretical literature suggests that the price of sports and income level are two relevant factors in the decision to participate in sports. Sports expenditures of Dutch local governments influence both of these factors. They aim to keep sports club membership fees affordable by providing sports facilities, and grant subsidies to low-income households. Both policies make it more likely that an individual will participate in sports. Therefore, my first hypothesis is:

H1. *There exists a positive relationship between municipal spending on sports and the sports participation rate in Dutch municipalities.*

Municipal efforts that stimulate people to play sports also increase the likelihood that inhabitants engage in exercise activities that are not considered sports, such as dancing. Additionally, Dutch municipalities spend money on public parks and other open-air recreation facilities. This not only provides a better environment for outdoor sports but also encourages recreational activities like walking and cycling. Hence, my second hypothesis is:

H2. *There exists a positive relationship between municipal spending on sports and exercise participation rate in Dutch municipalities.*

Lastly, I hypothesize that sports participation is more heavily influenced by municipal sports expenditures than exercise participation. Everyday activities like cleaning and walking are also considered exercising. Because of this, exercising is less constrained by income level than sports participation. Therefore, I expect policies that take away financial barriers are less effective in encouraging exercise participation as opposed to sports participation. My third hypothesis thus is:

H3. *The relationship between municipal sports expenditures and sports participation is stronger than the relationship between municipal sports expenditures and exercise participation.*

3 Data

Next, I discuss the data used to perform the analysis. First, I describe the data sources and briefly discuss their credibility. Subsequently, I explain how the data is merged and examine whether the loss of observations during this process could affect the results. Lastly, I discuss the descriptive statistics of the sample.

3.1 Data Collection and Dataset Construction

As a first step in constructing the dataset, I collect data from Statistics Netherlands, the Dutch national statistical office. Statistics Netherlands provides a wide range of freely accessible municipality-level data, making it a suitable data source. Within this data, I obtain information on sports and exercise participation from the Health Monitor Adults and Elderly 2020. The monitor is carried out every four years to gather information on the health status, social situation, and lifestyle of the adult Dutch population on a national, regional, and local level. Close to 540,000 individuals participated in the monitor in 2020, so a reasonable sample size in each municipality is ensured. Furthermore, data is weighted based on individual- and municipality characteristics to correct for differences in the composition of the sample and the population (Statistics Netherlands, n.d.-a). Overall, the Health Monitor Adults and Elderly 2020 is thus considered to be a credible source for participation rate statistics.

For information on municipal expenditures, I make use of Financial Information for Third Parties (Iv3) data. Iv3 is an information system used by Dutch municipalities to report their financial statements to Statistics Netherlands. In the records, expenditures must be allocated to the policy category they are incurred in. Sports spending can be assigned to two policy categories: 'Sports Policy and Activation' and 'Sports Facilities'. However, sometimes expenditures relate to both categories while municipalities can only choose one. This can lead to inaccuracies in the recording of sports spending on the specific categories (Van den Dool & Van Eldert, 2021). Nevertheless, this matters little for the total amount of recorded sports expenditures. A more serious issue arises when expenditures can be assigned to both a sports and non-sports category. It is expected that this mainly concerns sports-related spending in the public space, which is often reported as spending on 'Public Parks and (open-air) Recreation' (Van Eldert & Beekman, 2021). Therefore, the expenditures of this category are included as a separate variable. Furthermore, due to a change in the way spending is recorded, expenditure data from before and after 2017 is not comparable. Henceforth, only spending from 2017 to 2020 is taken into account.

Municipality characteristics are primarily obtained from a Statistics Netherlands database with local key figures. These do not include income and health statistics, so I gather these from a household income dataset of Statistics Netherlands (for disposable income data) and the Health Monitor Adults and Elderly 2020 (for information on the prevalence of long-standing illnesses). I use 2020 data for all background variables, as participation rates are measured in 2020 as well.

The remaining data required for the analysis is obtained from two sources. Firstly, whether a municipality has a historic city center is checked by examining the specification of the distribution of the municipal fund, available on the website of the Dutch government. Secondly, sports federation membership rates are collected from Volksgezondheidenzorg.info, a website of the Dutch National Institute for Public Health and the Environment.

Combining the data mentioned above into one dataset leads to a loss of observations for two reasons. Firstly, information on participation rates is not available for every municipality. In the province of Utrecht, participation data is only available for the city of Utrecht due to an error in the data collection of the Health Monitor Adults and Elderly 2020, resulting in the loss of 25 observations. Furthermore, 2 municipalities lack participation rate data because of a too small sample of respondents. Secondly, since I compute average sports spending, municipalities that do not have expenditure data for all years between 2017 and 2020 are dropped. This is the case for 30 observations as at some point during this period they were involved in a municipal merger. Consequently, the final sample contains 298 out of the 355 municipalities in the Netherlands in 2020.

3.2 Descriptive Statistics

Table 1 shows the descriptive statistics of all variables in the sample. Regarding the dependent variables, both the average sports participation and average active lifestyle rate in Dutch municipalities are 49.8%. Furthermore, the sports participation rate is spread over a slightly wider interval, but both variables range from approximately 30% to 70%. To see whether municipalities with high participation rates are clustered in certain regions, I create Figure 1. The left side of the figure displays that municipalities with a high sports participation rate are mainly concentrated in the south and west of the Netherlands. On the other hand, it follows from the map on the right that municipalities with a relatively active population seem to be more scattered. Therefore, a low sports participation rate does not necessarily imply that a municipality's population is inactive or vice versa.

The descriptive statistics of the expenditure variables also require further discussion. On average, the municipalities in the sample spend about 80% of the total sports expenditures per capita on sports facilities and use the remaining 20% for sports policy and activation. This ratio is in line with how sports expenditures are divided over the two categories on an aggregate level, which can be derived from the numbers presented by Van Eldert and Beekman (2021).

However, the minimum and maximum values suggest that average yearly sports expenditures vary widely between municipalities. So, to gain more insight into how sports expenditures per capita are distributed, I plot histograms for each expenditure variable in Figure A1. The figure displays each distribution to have a

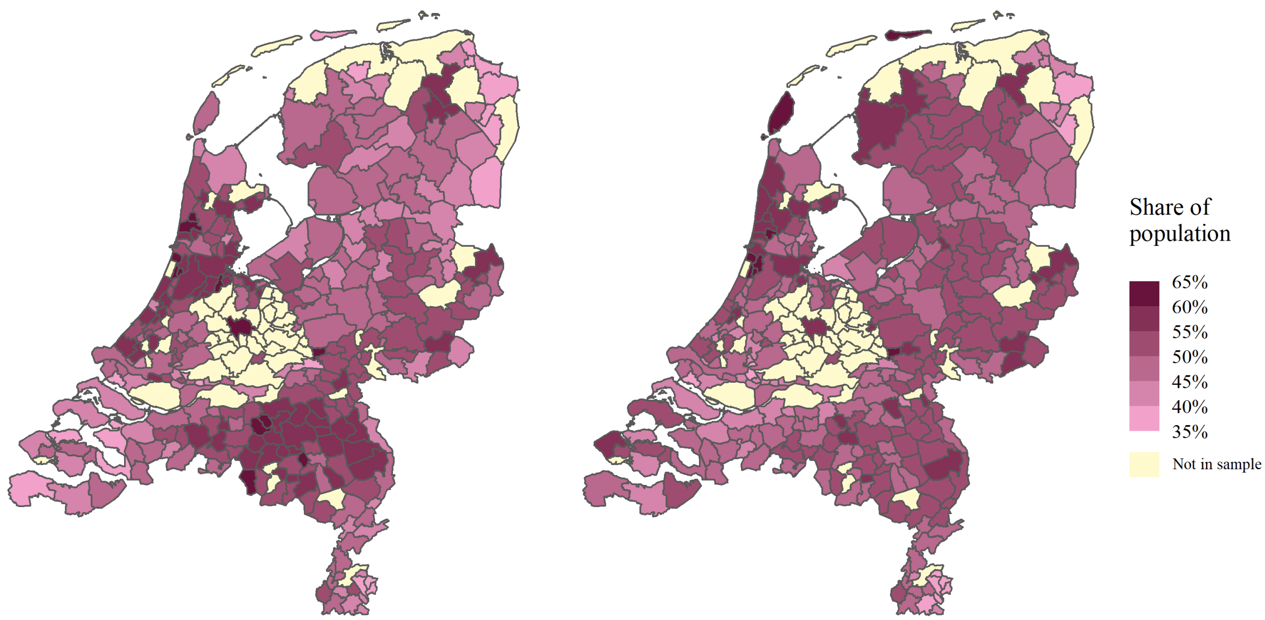
Table 1: Sample descriptive statistics *(n = 298)*

Variable	Mean	St. Dev.	Min	Max
Sports participation^s	0.498	0.062	0.328	0.672
Active lifestyle^s	0.498	0.049	0.362	0.664
Sports federation membership^s	0.195	0.039	0.104	0.307
Sports expenditures (in €)	80.692	28.524	14.488	224.461
Sports policy and activation expenditures (in €)	16.412	14.123	0.965	120.360
Sports facilities (in €)	64.280	27.752	8.826	219.561
Public parks and Recreation expenditures (in €)	88.009	30.048	4.695	304.899
Historic city center	0.362	0.482	0	1
Male^s	0.498	0.008	0.471	0.515
Age 15-24^s	0.138	0.020	0.102	0.256
Higher education^s	0.270	0.077	0.108	0.525
Long-standing illness^s	0.327	0.032	0.232	0.449
Migration background^s	0.168	0.091	0.042	0.556
Average household size	2.248	0.171	1.700	3.300
Median disposable income	28.538	2.386	21.600	37.800
Drenthe	0.040	0.197	0	1
Flevoland	0.020	0.141	0	1
Friesland	0.044	0.205	0	1
Gelderland	0.151	0.359	0	1
Groningen	0.027	0.162	0	1
Limburg	0.097	0.297	0	1
North Brabant	0.201	0.402	0	1
North Holland	0.148	0.355	0	1
Overijssel	0.077	0.267	0	1
South Holland	0.151	0.359	0	1
Utrecht	0.003	0.058	0	1
Zeeland	0.040	0.197	0	1

Note: ^{s} indicates that the variable is measured as the share of a municipality's population. The 'Historic city center' and province variables are dummies that take a value of 1 if a municipality has at least one historic city center or a municipality is located in that province, and 0 otherwise. Means are not population weighted and thus cannot be interpreted as representative for the Netherlands as a whole.

rather long right tail due to remarkably high per capita expenditures in some municipalities. Furthermore, the minimum value of the Public Parks and Recreation expenditures appears to be an anomaly given its distance from the rest of the distribution. These outliers can have a considerable impact on OLS estimates due to the large weight their residuals receive in minimizing the sum of squared residuals (Wooldridge, 2015). Therefore, as the expenditure data approximately follows a log-normal distribution, I take the natural logarithms of the expenditure variables to reduce the influence of outliers. Figure A2 shows the distribution of the log-transformed expenditure variables.

Yet, there are still several outlying observations after the log transformation. A solution to this is to winsorize the data. In winsorization, values in the highest and lowest percentiles are replaced by values of less extreme percentiles. Compared to dropping outliers, the advantage of this approach is that some information about an outlier can still be used (Ghosh & Vogt, 2012). In this research, for example, a winsorized



(a) Sports participation

(b) Active lifestyle

Figure 1: Sports participation rate and share of the population with an active lifestyle in Dutch municipalities

observation still signals that a municipality has high per capita sports expenditures even if the winsorized expenditures are lower than the actual expenditures, whereas this information is lost when an outlier is dropped. However, a caveat is that winsorizing a too large share of the data means that values of non-outlier observations are replaced as well, leading to an unnecessary loss of information. Therefore, I only winsorize observations in the 1^{st} and 100^{th} percentile and set their values to the lowest value in the 2^{nd} or the highest value in the 99^{th} percentile, respectively. Lastly, it is important to note that the expenditure variables are winsorized before taking natural logarithms. Figure A3 presents the distributions of the expenditure variables after winsorization.

Additionally, I compare the characteristics of the sample municipalities and the missing municipalities to examine whether the sample is representative. As reported in Table 2, for most variables the means of both groups are nearly identical. Yet, the results of the t-tests show that the average share of the population with a long-standing illness and the average median disposable income do differ significantly at a 95% confidence level.

Further investigation reveals that this is likely due to the missing municipalities in the province of Utrecht, as their population is relatively young and high-educated. Furthermore, Table A1 reports a significantly

Table 2: Descriptive statistics of the background variables for the sample and omitted municipalities

	Sample (*n = 298*)		*Omitted municipalities* (*n = 57*)		
Variable	Mean	St. Dev.	Mean	St. Dev.	t-stat.
Sports federation membership^s	0.195	0.039	0.198	0.041	0.596
Male^s	0.498	0.008	0.499	0.009	0.750
Age 15-24^s	0.138	0.020	0.142	0.019	1.535
Higher education^s	0.270	0.077	0.283	0.085	1.059
Long-standing illness^s	0.327	0.032	0.313	0.027	-3.287^{***}
Migration background^s	0.168	0.091	0.149	0.073	-1.705^{*}
Average household size	2.248	0.171	2.289	0.201	1.446
Median disposable income	28.538	2.386	29.539	2.401	2.822^{***}

Note: ^{s} *indicates that the variable is measured as the share of a municipality's population. For the variables 'Long-standing illness' and 'Median disposable income' data is missing for two and three omitted municipalities, respectively.* $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

higher average sports federation membership rate in municipalities in Utrecht. As the sports federation membership rate correlates strongly with sports participation (Figure A4), this raises concern that the results may be affected by the missing data in municipalities in Utrecht. Therefore, using the sports federation membership rate as the dependent variable, I will estimate an additional regression for a sample including the municipalities in Utrecht to assess the influence of the missing data in the province of Utrecht.

4 Methodology

Subsequently, I present the empirical strategy for analyzing the relationship between municipal sports expenditures and sports participation. I first explain how this relationship can be investigated using Ordinary Least Squares regression. Thereafter, I propose using the presence of a historic city center in a municipality as an instrumental variable in a Two Stages Least Squares regression to resolve the endogeneity issue that may arise in the OLS regression. Lastly, I describe how I use a similar approach to examine the relationship between municipal sports expenditures and exercise participation.

4.1 City Sports Expenditures and Sports Participation

4.1.1 OLS Regression

A standard approach for analyzing the relationship between two variables is to use Ordinary Least Squares regression. Therefore, to examine how local governments' sports expenditures relate to sports participation, I estimate the following regression:

$$SP_i = \beta_0 + \beta_1 * EXP_i + \gamma * X_i' + \varepsilon_i \tag{4.1}$$

In this model, SP_i denotes the sports participation rate in municipality i, EXP_i represents municipality i's average yearly sports expenditures per capita, X_i' is a vector of control variables for municipality i, and ε_i is the error term.

The dependent variable in the regression, sports participation, is defined as the municipality's share of the population that reports to exercise at least once a week in the 2020 Health Monitor of Statistics Netherlands. It is important to note that the dependent variable is a rate and hence is bounded between 0 and 1. This can complicate statistical inference in OLS regression through non-normally distributed error terms and heteroscedasticity. In addition, linear regression may predict values for the bounded dependent variable outside of the realistically possible [0, 1] interval (Paolino, 2001). However, these issues are more severe when large parts of the data are close to either one or both of the boundaries. Since the sports participation rates in our sample range from 32.8% to 67.2% and are concentrated around a mean of 49.8%, the aforementioned problems are not considered to hinder the analysis. To validate this, I will check for normality and homoscedasticity of the errors and predict the values of the dependent variable to confirm that these are within the possible [0, 1] range.

Another comment regarding the proposed dependent variable is that the Covid-19 pandemic may have negatively affected sports participation in 2020. Sports facilities in the Netherlands were closed for several months during the year, hindering people in their ability to take part in the sports they usually play. Nevertheless, research by Duijvestijn et al. (2021) and Van den Dool (2021) shows that sports participation in the Netherlands did not appear to be significantly lower in 2020 than in 2019. It is thus unlikely that the results are influenced by an effect of the Covid-19 pandemic on sports participation.

As the independent variable of interest, I use an average of the yearly amount of municipal sports expenditures per capita. To calculate this, I first add up municipalities' spending on the two designated sports accounts in the municipal financial records for each year in the period 2017-2020. Subsequently, per year I divide this by the number of residents to control for the fact that municipalities with a larger population likely have higher expenditures. I then compute the average of the per capita sports expenditures to account for fluctuations in sports spending. This is necessary as Van Eldert and Beekman (2021) show sports expenditures within a municipality can vary substantially between years, for instance, due to a large one-off investment in a new sports hall. As a final step, I take the natural logarithm of this average to reduce the impact of outliers.

In addition to the analysis with total sports spending as the main independent variable, I also perform OLS regressions where I incorporate the two sports accounts as separate variables. This allows us to distinguish between the effects of expenditures aimed at activating citizens (allocated to the 'Sports Policy and Activation' account) and spending that is related to sports facilities (allocated to the 'Sports Facilities' account).

I also include municipalities' Public Parks and (open-air) Recreation expenditures as a separate variable.

Humphreys and Ruseski (2007) report that federal states' spending on parks and recreation has a positive effect on participation in outdoor activities and individual sports in the United States. Moreover, research on the municipal sports expenditures by Van Eldert and Beekman (2021) reveals that a considerable part of Public Parks and (open-air) Recreation spending can be considered sports expenditures. Hence, there may be a relationship between municipal spending on the natural environment and sports participation. I transform the additional expenditure variables in a similar way as the total sports expenditures variable.

Furthermore, a set of covariates is included in the regression model. It is important to do so, as failing to control for factors that are related to both the dependent and independent variable leads to omitted variable bias in the regression coefficients (Wooldridge, 2015). Because earlier research suggests that factors such as population demographics and economic prosperity are likely to be associated with both the propensity to participate in sports and municipal expenditures (Kokolakakis et al., 2014; Steckenleiter et al., 2019), we must account for them in our model to obtain unbiased results.

The socio-economic municipality characteristics I include as covariates are gender, age, education, health status, migration background, household size, and income. With regards to gender, various studies found that men participate in sports more frequently than women (Van Tuyckom et al., 2010; Fridberg, 2010; Kokolakakis et al., 2012). Therefore, the proportion of males in the population is controlled for. Age is another widely recognized determinant of sports participation, with younger people being more likely to play sports (Hovemann & Wicker, 2009; Hoekman et al., 2017). I account for this by incorporating the share of a municipality's population between the age of 15 and 24 in the regression. Existing literature also finds that sports participation rates are higher among highly educated individuals (Kokolakakis et al., 2014; Hoekman et al., 2017). On top of that, Pulles et al. (2021) find that the Covid-19 pandemic has increased the difference in sports participation rates between lower- and higher-educated people in the Netherlands. Therefore, I include the share of the population with a higher vocational or university degree as a variable too. Furthermore, Kokolakakis et al. (2014) show that having a chronic disease is negatively correlated with sports participation, so I control for the share of the population with a long-standing illness. Kokolakakis et al. (2014) and Breuer et al. (2011) also report a correlation between sports participation and ethnicity. They find that people with a migration background are significantly less likely to participate in sports in England and Germany, presumably due to cultural barriers associated with migrating to a new country. Hence, the share of the population with a migration background is included as a control as well. With respect to household size, several studies present evidence that individuals in larger households participate in sports less frequently (Downward, 2007; García et al., 2011; Kokolakakis et al., 2014). Additionally, married people are found to have a lower propensity to participate than single or divorced individuals (Eberth & Smith, 2010). I aim to capture these effects by using the average household size in a municipality as a control variable.

Prior research also reports sports participation to be related to economic variables. There is consensus that a higher income is positively associated with the likelihood of exercising (Downward et al., 2009), a relationship that is identified in the Netherlands too Hoekman et al. (2017). Therefore, I include municipalities' median disposable household income as a covariate as well. Lastly, I account for potential region-specific factors by including dummy variables for each Dutch province. A complete overview of all variables and their definitions are reported in Table 3.

Table 3: Overview of variables

Variable	Type	Description
Sports participation	Continuous*	Share of adult population regularly participating in a sports activity (at least once per week)
Active lifestyle	Continuous*	Share of adult population meeting the requirements of the Dutch government's active lifestyle guidelines
Sports federation membership	Continuous*	Share of population that is a member of a Dutch sports federation (excluding members of the golf, equestrian sport, or sport fishing federations)
Sports expenditures	Continuous	Average yearly municipal spending on sports, per capita (from 2017 to 2020)
Public parks and Recreation expenditures	Continuous	Average yearly municipal spending on public parks and (open-air) recreation, per capita (from 2017 to 2020)
Historic city area	Binary	Variable indicating whether a historic city area is present in the municipality (takes value 1 if yes, 0 otherwise)
Male	Continuous*	Share of population that is male
Age 15-24	Continuous*	Share of population that is between the age of 15 to 24
Higher education	Continuous*	Share of population that has obtained a higher vocational or university degree
Long-standing illness	Continuous*	Share of population suffering from a long-standing illness
Migration background	Continuous*	Share of population with a migration background
Average household size	Continuous	Average household size in a municipality
Median disposable income	Continuous	Median disposable household income in a municipality (after standardization to account for differences in household size and composition (Statistics Netherlands, n.d.-b))
Province	Categorical	Variable indicating in which province of the Netherlands a municipality is located

*Note: * indicates the variable is continuous only within the bounded interval [0, 1].*

To assess the robustness of the results, I estimate models with alternative spending periods. Previous research shows that the strength of the relationship between municipal spending and sports participation declines over time (Van den Dool & Hoekman, 2017). Taking only recent years into account may thus affect the results. On the other hand, it may take time for investments in sports facilities to start influencing sports participation. Moreover, the Covid-19 pandemic may have changed municipalities' spending behavior.

It is thus worthwhile to also consider specifications with expenditures in earlier years only. Therefore, the alternative spending periods I calculate average municipal sports expenditures for are 2017-2019, 2018-2020, 2017-2018, 2018-2019, and 2019-2020.

Additionally, I test the robustness of the results against a different dependent variable by using municipalities' sports federation membership rate as the dependent variable. The share of the population that is a member of a sports federation is thought to specifically capture participation in sports that are often played in sports clubs, as being a member of a sports club usually implies that a person is also a member of a sports federation. Because sports clubs receive a large share of the municipal sports budget (KPMG, 2019), the association between sports expenditures and sports federation membership is expected to be stronger than the one between sports expenditures and sports participation.

4.1.2 2SLS Regression

Although OLS regression is a convenient starting point for our analysis, there are two reasons why the coefficients of the expenditure variables are possibly endogenous. Firstly, despite including a set of observable control variables, it could be that we fail to account for unobservable characteristics of municipalities (e.g., the general preference for sports as a leisure activity) that are correlated with sports participation and municipal sports expenditures. In that case, the independent variables will not be uncorrelated with the error term and the regression coefficients will suffer from omitted variable bias (Wooldridge, 2015).

A second issue is that sports expenditures may not only affect sports participation, but sports participation is also likely to influence municipal sports spending. For instance, a higher sports participation rate may imply more usage of municipal sports facilities, causing them to wear faster and requiring municipalities to spend more on sports facilities. This simultaneity also causes coefficients to be biased due to the correlation between the independent variables and the error term (Wooldridge, 2015).

A suitable approach to tackle these problems is to perform Two-Stage Least Squares regression. 2SLS uses an instrumental variable (z) to estimate the unbiased effect of an endogenous independent variable (x) on the dependent variable (y) (Angrist & Pischke, 2008). This is done in two stages. In the first stage, x is regressed on z. Then, using the predicted values of the endogenous dependent variable ($\hat{x}$) obtained from the first stage, a second stage regression is run of y on $\hat{x}$. Provided that the effect of the instrument on the endogenous variable is sufficiently strong (relevance assumption), any relationship between the instrument and the dependent variable happens only through the independent variable (exclusion restriction), and the instrument is uncorrelated with other factors influencing the dependent variable (independence assumption), this will yield unbiased estimates as $\hat{x}$ is uncorrelated with the error term.

In the context of this thesis, I use the presence of a historic city center in a municipality to instrument municipal sports expenditures. The nature of this relationship stems from the way Dutch municipalities are financed. Almost half of the income municipalities receive comes from the municipal fund of the Dutch government. The allocation of the money in this fund is based on over sixty benchmarks, such as population size or the number of inhabitants receiving social assistance. These benchmarks are categorized into thirteen clusters suggesting the policy area the money is supposed to be used for.

One of the benchmarks of the Culture and Recreation cluster, the policy area sports belongs to, is the presence of one or more historic city centers in the municipality. Moreover, another benchmark assigns funds based on the number of inhabited historic houses within these centers. In 2019 alone, this caused municipalities with a historic city center to, on average, receive over €600.000 more from the municipal fund than municipalities without one. Having a historic city center thus provides a considerable increase in the municipalities' Culture and Recreation budget and is, therefore, expected to lead to higher municipal sports spending. Figure 2 displays a graphical representation of this relationship.

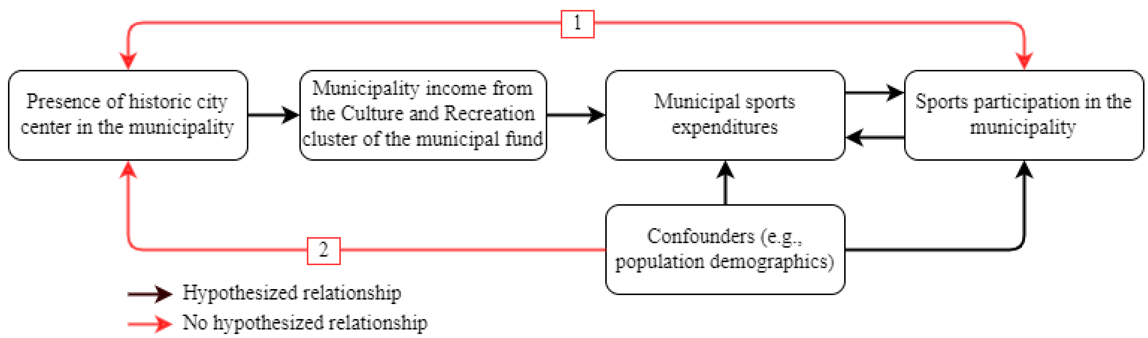


Figure 2: Diagram of (expected) causal relationships

The presence of a historic city center is also assumed to satisfy the exclusion restriction and independence assumption (illustrated in Figure 2 by arrows 1 and 2, respectively). Regarding the exclusion restriction, a historic city center is improbable to induce people to play more sports, so any effect on sports participation must happen through the increase in the municipal budget. Because it seems unlikely that non-sports-related municipal expenditures would affect sports participation, the higher budget is thought to influence sports participation exclusively through municipal sports expenditures.

The proposed instrument is also presumed to be independent due to its random component. One of the requirements for a city center to be historic is that the area had city rights obtained between 1600 and 1850. Because characteristics of municipalities today cannot influence this, the assignment of historic city centers can thus be considered essentially random. Yet, a reason for concern is that as cities increase in size they become more likely to have a historic center. If larger cities have features that correlate with sports participation, this would violate the independence assumption. Therefore, I include control variables in the 2SLS regression to make the instrument independent conditional on the covariates.

To estimate the 2SLS regression, I make use of the same variables as in the OLS regression, plus a dummy variable for the presence of a historic city center. The regression equation for the first stage thus can be written as follows:

$$EXP_i = \delta_0 + \delta_1 * HC_i + \zeta * X_i' + \eta_i \tag{4.2}$$

In this equation, EXP_i denotes the average yearly sports expenditures of municipality i, HC_i is a binary variable that takes a value of 1 if municipality i has a historic city center and is 0 otherwise, X_i' represents a vector of control variables for municipality i, and η_i is the error term.

Subsequently, the second stage regression equation for examining the relationship between municipal sports spending and sports participation is as follows:

$$SP_i = \beta_0 + \beta_1 * \widehat{EXP}_i + \gamma * X_i' + \varepsilon_i \tag{4.3}$$

Here, SP_i stands for the sports participation rate in municipality i, $\widehat{EXP}_i$ represents municipality i's average yearly sports expenditures per capita as predicted by the first stage, X_i' is a vector of controls for municipality i, and ε_i denotes the error term.

To test the assumptions of the 2SLS regression, I carry out several robustness checks. Firstly, the relevance assumption can be verified by computing the instrument's F statistic in the first stage regression. As a general rule of thumb, the instrument is considered strong when this statistic is greater than 10 (Angrist & Pischke, 2008).

Confirming the validity of the exclusion restriction and independence assumption is not as straightforward because the error term is unknown, making it impossible to check whether the error term and the predicted sports expenditures are correlated. However, by using falsification strategies, we can examine the strength of these assumptions. To falsify the exclusion restriction, I investigate whether sports participation is related to municipal spending on culture and media, two policy areas that are also part of the Culture and Recreation cluster of the municipal fund. I do so using the following regression equation:

$$SP_i = \beta_0 + \beta_1 * EXP_i + \beta_2 * CUL_i + \beta_3 * MED_i + \gamma * X_i' + \varepsilon_i \tag{4.4}$$

Here, denotes SP_i the sports participation rate in municipality i, and EXP_i, CUL_i, and MED_i represent average yearly per capita expenditures on sports, culture, and media, respectively. Furthermore, X_i' is a vector of controls and ε_i stands for the error term. If no relationship between sports participation and culture or media spending is identified, this strengthens our belief that an effect of having a historic city centers occurs only through higher municipal sports expenditures.

The primary threat to the independence assumption is the positive association between the size of a city and the probability of having a historic city center, which makes highly urbanized municipalities more likely to have a historic city center. If such municipalities have unobserved characteristics related to sports participation, this would imply that the instrument is not independent. Therefore, I estimate an additional

OLS regression to examine whether urbanization correlates with sports participation, where I use population size as a proxy of urbanization. The corresponding regression equation is as follows:

$$SP_i = \beta_0 + \beta_1 * POP_i + \gamma * X_i' + \varepsilon_i \tag{4.5}$$

In this equation, SP_i again is the sports participation rate in municipality i, POP_i is a categorical variable indicating the population size of municipality i, X_i' represents a vector of control variables, and ε_i is the error term. Subsequently, by performing an F-test for the joint significance of the coefficients of the population size categories, it can be assessed whether urbanization is significantly related to sports participation after controlling for municipality characteristics. If this is the case this would be a strong indication that the independence assumption is violated. Otherwise, it suggests that the relationship between urbanization and the probability of having a historic city center does not threaten the validity of the instrument.

Additionally, I compare the background characteristics of municipalities with and without historic city centers by performing t-tests. If cities with historic centers appear to be different from cities without historic centers in terms of demographics and economy, this hints that the two groups may also differ in other, unobservable characteristics. In that case, the independence assumption may be violated even after controlling for the observable differences. On the other hand, if the groups of municipalities do have similar characteristics it is plausible that they also have comparable unobservable characteristics. However, due to the unobservable nature of these characteristics insignificant t-test results cannot provide proof for this.

Furthermore, I test the general robustness of the 2SLS results. This implies that I consider different spending periods for the calculation of the instrumented variable. The periods I use are similar to the ones used in the robustness checks of the OLS regression. Also, I again use the sports federation membership rate to assess the sensitivity of the results to the dependent variable.

4.2 City Sports Expenditures and Exercise Participation

4.2.1 OLS Regression

In addition to examining the relationship between municipal sports expenditures and sports participation, I also investigate how spending on sports affects exercise participation. I do this using a similar approach as described for sports participation. The OLS regression equation thus looks as follows:

$$AL_i = \beta_0 + \beta_1 * EXP_i + \gamma * X_i' + \varepsilon_i \tag{4.6}$$

Here, AL_i is the share of municipality i's population that is considered to have an active lifestyle, EXP_i stands for the average yearly sports expenditures of municipality i, X_i' is a vector of controls for municipality i, and η_i denotes the error term.

Regarding the dependent variable, I use having an active lifestyle as a measure of exercise participation. A person is considered to be active when he or she meets the active lifestyle guidelines of the Dutch government. For adults, this means that someone must exercise moderately intensively for at least 2.5 hours per week and do muscle- or bone-strengthening exercises at least twice a week. Activities such as walking, cycling, and some forms of housework also contribute towards meeting these guidelines (Kenniscentrum Sport en Bewegen, n.d.). The variable thus captures a wide range of physical activity, making it a suitable way to measure exercise participation.

The independent variables in the exercise participation regression are identical to those in the regression for sports participation. This means that average municipal sports expenditures are calculated in the same way. Also, I again use the total sports expenditures, the two individual sports spending accounts, and the spending on Public Parks and (open-air) Recreation to estimate the relationship between expenditures and participation. Furthermore, I compute the average sports expenditures using different sets of years to assess the robustness of the results with respect to the period considered.

I also use the same control variables in the sports and exercise participation regressions. A literature review by Trost et al. (2002) shows that men are more likely to exercise than women, but age is negatively related to exercise participation. Education has a positive impact on the propensity to exercise, just like income. Better self-perceived health status is also positively associated with exercising. Furthermore, white individuals tend to exercise more than individuals with a different ethnic background. Lastly, being married and having children have a negative association with exercising, suggesting that household size has a negative influence on exercise participation. Exercise participation and sports participation thus appear to be affected by the same factors, allowing us to use similar control variables to reduce the chances of omitted variable bias entering the model.

4.2.2 2SLS Regression

As the OLS regression for exercise participation may suffer from omitted variable bias and simultaneity bias, I estimate a 2SLS regression as well. The first-stage regression equation is already given by Equation 4.2, as the instrumented variable, the instrument, and the control variables are identical to those of the sports participation 2SLS. However, due to the different dependent variables the second-stage regression does differ and can be written as:

$$AL_i = \beta_0 + \beta_1 * \widehat{EXP}_i + \gamma * X_i' + \varepsilon_i \tag{4.7}$$

In the equation, AL_i represents the exercise participation rate in municipality i, $\widehat{EXP}_i$ is municipality i's average yearly sports expenditures per capita as predicted by the first stage, X_i' stands for a vector of control variables for municipality i, and ε_i denotes the error term.

To test the validity of the OLS assumptions I repeat the robustness checks for 2SLS regression for sports participation. Hence, I attempt to falsify the exclusion restriction by estimating the relationship between municipal spending on culture and media and exercise participation. Furthermore, I examine whether exercise participation is correlated with population size. Lastly, to assess the sensitivity of the results, I estimate the 2SLS regressions using different sets of years to compute the average municipal expenditures.

5 Results

In the upcoming section, I present the results of the analysis. Firstly, I compare regression estimates across samples handling outliers in different ways to decide which one is the most appropriate. Then, I discuss the OLS and 2SLS regression estimates for sports participation and assess the robustness of these results. Lastly, I report the outcomes of the OLS and 2SLS regressions for exercise participation and briefly examine their robustness as well.

5.1 Relationship between City Sports Expenditures and Sports Participation

I begin the analysis by estimating baseline regressions for both sports- and exercise participation using the full sample, a sample where outliers are removed, and a sample with winsorized outliers to assess the influence of outliers. The estimates in Table A2 show that the coefficient for sports expenditures substantially increases when the lowest and highest percentiles are removed from the data. Moreover, the coefficient is significant at a 90% confidence level when outliers are excluded, whereas it is insignificant at all conventional significance levels when the complete sample is used. Dropping outliers thus appears to have a considerable impact on the results.

Yet, disregarding the outliers altogether implies that information on municipalities with very low or very high sports expenditures is lost. Columns 5 and 6 in Table A2 illustrate this, as the coefficient for sports expenditures becomes smaller and turns insignificant when the 1^{st} and 100^{th} percentile are winsorized. So, the magnitude and significance of the sports expenditure coefficients are similar to those for the sample with untreated outliers, but without the worry of outliers affecting the estimates.

Additionally, Table A3 presents a similar pattern regarding the magnitude of the coefficients for different datasets when using the share of the municipality population with an active lifestyle as the dependent variable. Therefore, hereafter I only present the results of regressions using winsorized data.

5.1.1 OLS Regression Results

In Table 4 I report the OLS estimates of the relationship between municipal sports expenditures and sports participation rate. From columns 1 and 2 we derive that there does not appear to be a significant association between the average yearly sports expenditures per capita and sports participation. We do observe a significant relationship at a 90% confidence level between sports facilities expenditures and sports participation in column 4, where the sports expenditures are split up into sports policy and activation spending and spending on sports facilities. However, the economic significance of the coefficient is marginal, as the coefficient of

0.008 implies that, all else equal, a 10% increase in the average yearly sports facilities expenditures per capita is associated with only a 0.08 percentage point increase in the sports participation rate. Additionally, F-tests show that the coefficients of the sports spending categories in columns 3 and 4 are jointly not significantly different from 0 either (F-statistics are 1.343 (277 and 2 degrees of freedom) and 1.758 (276 and 2 degrees of freedom), respectively).

The estimated coefficients for Public Parks and Recreation expenditures have a negative sign. This is surprising, as a considerable part of these expenditures relate to sports and, therefore, were expected to positively associate with sports participation. A possible explanation is that the non-sports-related spending on Public Parks and Recreation induce a substitution effect because people rather spend time doing other leisure activities in the park than play sports. Yet, the coefficients are statistically insignificant, so the results should be interpreted carefully.

Nearly all control variables are significant at a 95% confidence level. In line with prior literature, a larger proportion of males in a municipality is associated with a higher sports participation rate. Sports participation also appears to be higher in municipalities with a more youthful population and in municipalities with a larger share of high-educated individuals, which is consistent with earlier findings. The average household size in a municipality is negatively associated with the sports participation rate, suggesting that having kids or a partner may decrease the time available for leisure activities like sports. A higher median disposable income in a municipality correlates with a higher sports participation rate. This is consistent with the prediction of Taylor and Gratton (2002) that the income effect dominates the substitution effect in the demand for sports. Furthermore, as expected the sports participation rate is lower in municipalities where a large share of inhabitants suffers from a long-standing illness. The proportion of the population with a migration background is the only variable that is not significantly related to sports participation. This is surprising, as on a national level migrants are reported to play less sports compared to people without a migration background in the Netherlands (National Institute for Public Health and the Environment, 2022). An explanation for this result is that other variables already capture the effect of having a migration background. Hoogendoorn and De Hollander (2016) state that a low income and health issues are important barriers to exercise for immigrants. It may be that after controlling for these factors, having a migration background is not as important anymore in the decision to participate in sports.

To test the robustness of the OLS estimates, I regress a set of models where the average municipal sports expenditures are computed using different sets of years. Table A4 shows the estimates of these alternative specifications. An interesting finding is that for specifications without expenditure data from 2020 the total sports expenditures coefficients are significant at a 90% confidence level. The models with the sports expenditures per policy category reveal that this is largely due to spending on sports facilities, as their coefficients are significant at the 95% level. Furthermore, when only spending in 2019 and 2020 is considered the coefficient for both total sports expenditures and sports facilities expenditures halves. The

Table 4: OLS estimates of the relationship between municipal sports expenditures and the sports participation rate in Dutch municipalities

	Dependent variable:			
	Sports participation			
	(1)	(2)	(3)	(4)
Sports expenditures (log)	0.007	0.008		
	(0.005)	(0.005)		
Sports policy and activation expenditures (log)			0.001	0.002
			(0.002)	(0.002)
Sports facilities expenditures (log)			0.007	0.008^{*}
			(0.004)	(0.004)
Public parks and Recreation expenditures (log)		−0.012		−0.012
		(0.008)		(0.008)
Male^{s}	1.087^{***}	1.064^{***}	1.093^{***}	1.073^{***}
	(0.335)	(0.335)	(0.336)	(0.335)
Age 15-24^{s} (log)	0.073^{***}	0.070^{***}	0.074^{***}	0.071^{***}
	(0.020)	(0.020)	(0.020)	(0.020)
Higher education^{s} (log)	0.122^{***}	0.118^{***}	0.122^{***}	0.117^{***}
	(0.014)	(0.015)	(0.014)	(0.015)
Average household size (log)	-0.112^{**}	-0.137^{**}	-0.112^{**}	-0.139^{**}
	(0.055)	(0.057)	(0.055)	(0.057)
Median disposable income	0.008^{***}	0.008^{***}	0.008^{***}	0.008^{***}
	(0.002)	(0.002)	(0.002)	(0.002)
Long-standing illness^{s}	-0.334^{***}	-0.357^{***}	-0.332^{***}	-0.356^{***}
	(0.088)	(0.089)	(0.088)	(0.089)
Migration background^{s} (log)	−0.004	−0.004	−0.003	−0.004
	(0.006)	(0.006)	(0.006)	(0.006)
Constant	0.220	0.275	0.214	0.270
	(0.176)	(0.179)	(0.176)	(0.179)
Province dummies	Yes	Yes	Yes	Yes
Observations	298	298	298	298
R^2	0.789	0.791	0.790	0.792

Note: The table reports the estimates of OLS regressions with the share of a municipality's population that exercises at least once per week as the dependent variable. The municipal expenditure variables are measured as the average yearly per capita expenditures on an expenditure category between 2017 and 2020. Variables denoted with ^{s} are measured as a share of the municipality population. Standard errors are reported in parentheses. Significance is indicated as follows: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

data thus suggests that it takes time before investments in sports facilities start to have an impact on sports participation. An alternative explanation is that the expenditures are higher because the facilities are used more if many people play sports. Yet, this does not explain why the relationship becomes insignificant once expenditure data from 2020 is included. One could argue that this a result of different usage of sports facilities or different spending behavior by municipalities due to the Covid-19 pandemic. However, Van Eldert and Beekman (2021) show that municipal spending on sports facilities actually increased in 2020, making this explanation unlikely.

Regarding the magnitude of the coefficients, the size of the estimates in Table A4 are comparable to the size of those in Table 4. Hence, it must still be concluded that the relationship between municipal sports expenditures and sports participation is weak even when recent expenditures are not taken into account.

Additionally, the results of OLS regressions with the sports federation membership rate in a municipality as the dependent variable are reported in Table A5. Although the relationship between municipal sports expenditures and sports federation membership was expected to be stronger than the association between sports expenditures and sports participation, the magnitude of the sports expenditure coefficients is smaller. Also, the sports expenditure estimates remain insignificant in the sports federation membership regressions. The absence of a significant relationship between sports participation and municipal sports expenditures thus does not seem contingent on the dependent variable.

The control variables inform us that the sports federation membership rate and weekly sports participation rate have different determinants. The first difference is that the coefficient for the share of the population between the ages of 15 to 24 does not have significant explaining power for sports federation membership. A possible reason for this is that the sports federation membership rate also includes the membership of children, who account for a large share of federation membership. The proportion of inhabitants between 15 and 24 may, therefore, be relatively less important. Secondly, having a relatively large immigrant population is significantly associated with lower federation membership rates. This is in line with the idea that cultural or language barriers prevent immigrants from becoming a member of a sports club, and thus instead prefer to exercise in gyms or public spaces. A final difference is that there is no significant relationship between the proportion of people with a long-standing illness and sports federation membership.

Lastly, by comparing columns 3 and 4 to columns 1 and 2 in Table A5, it can be seen that the regression estimates of a sample including municipalities from the province of Utrecht are nearly identical to the estimates of the sample I use for the main analysis. Hence, the omission of municipalities in the province of Utrecht is unlikely to influence the results of our main regressions.

A discussion of the validity of the OLS assumptions is provided in Appendix B.

5.1.2 2SLS Regression Results

Next, I present the results of the 2SLS regressions in Table 5. Compared to the OLS estimates, the coefficient for sports expenditures is considerably larger but remains insignificant. Additionally, I examine the difference between the OLS and 2SLS estimates of the sports policy and activation expenditures variable. Again, the size of the coefficient increases but is insignificant in both regressions. Performing a 2SLS regression for sports facilities expenditures is not possible, as the instrument does not sufficiently explain variation in spending on sports facilities in the first stage (Table A6).

The signs and significance levels of the control variables in 2SLS regression are similar to those in the OLS regressions. Only the household size and long-standing illness variables become insignificant at the 95% confidence level in the regression for total sports expenditures. It appears that this is caused by a weak instrument, as standard errors increase when the instrument does not have sufficient explaining power.

I do several checks to investigate the validity of the instrument. Firstly, I examine the relevance assumption by looking at the F-statistic of the instrument in the first stage. As shown at the bottom of Table 5, the F-statistic is significant at a 95% confidence level for total sports expenditures and significant at a 99% confidence level for sports policy and activation expenditures. However, the F-statistics of 5.122 and 7.497 are below the suggested threshold of 10, indicating that the instrument is rather weak. Therefore, the 2SLS estimates may be biased towards the OLS estimates (Angrist & Pischke, 2008), which implies that the effect of sports expenditures is underestimated in the 2SLS. Consequently, to allow for valid inference I perform Anderson-Rubin (AR) tests, which can be used to estimate weak instrument robust confidence intervals (Andrews & Stock, 2005). The confidence intervals given by the AR tests are [-0.065, 0.306] for the sports expenditures variable and [-0.019, 0.070] for the sports policy and activation expenditures variable. It must thus be concluded that, after accounting for the weakness of the instrument, the effect on the sports participation rate is not significantly different from 0 for both variables.

Subsequently, to test for a potential violation of the exclusion restriction, I regress the sports participation rate on non-sports-related municipal expenditures. The OLS estimates in column 1 in Table A7 show that neither spending on culture nor media is significantly associated with the sports participation rate in Dutch municipalities after controlling for background characteristics. Assuming that extra government proceedings from having a historic city center are not used for expenditure categories outside the Culture and Recreation cluster, this result indicates there is no violation of the exclusion restriction through an effect of non-sports-related municipality expenditures on sports participation.

Next, I compare the characteristics of municipalities with and without a historic city center. Assuming that observable characteristics are indicative of unobservable characteristics, this allows us to examine how likely it is that the instrument is indeed independent. Table 6 shows that the two groups of municipalities

Table 5: OLS and 2SLS estimates of the relationship between municipal sports expenditures and the sports participation rate in Dutch municipalities

	Dependent variable:			
	Sports participation			
	OLS		*2SLS*	
	(1)	(2)	(3)	(4)
Sports expenditures (log)	0.007		0.034	
	(0.005)		(0.041)	
Sports policy and activation expenditures (log)		0.0005		0.012
		(0.002)		(0.015)
Male^s	1.087***	1.096***	1.063***	1.175***
	(0.335)	(0.337)	(0.353)	(0.366)
Age 15-24^s (log)	0.073***	0.072***	0.075***	0.080***
	(0.020)	(0.020)	(0.021)	(0.023)
Higher education^s (log)	0.122***	0.122***	0.123***	0.124***
	(0.014)	(0.014)	(0.015)	(0.015)
Average household size (log)	−0.112**	−0.116**	−0.097	−0.126**
	(0.055)	(0.055)	(0.062)	(0.059)
Median disposable income	0.008***	0.007***	0.009***	0.008***
	(0.002)	(0.002)	(0.003)	(0.002)
Long-standing illness^s	−0.334***	−0.351***	−0.265*	−0.335***
	(0.088)	(0.088)	(0.139)	(0.094)
Migration background^s (log)	−0.004	−0.004	−0.005	−0.007
	(0.006)	(0.006)	(0.007)	(0.008)
Constant	0.220	0.258	0.052	0.178
	(0.176)	(0.175)	(0.312)	(0.208)
Province dummies	Yes	Yes	Yes	Yes
Observations	298	298	298	298
R^2	0.789	0.788	0.769	0.768
Instrument F-Stat (1, 278)			5.122**	7.497***

Note: The table reports the estimates of OLS and 2SLS regressions with the share of a municipality's population that exercises at least once per week as the dependent variable. The municipal expenditure variables are measured as the average yearly per capita expenditures on an expenditure category between 2017 and 2020, and are instrumented by a dummy variable for having at least one historic city center in the municipality in the 2SLS regressions. Variables denoted with ^{s} are measured as a share of the municipality population. Standard errors are reported in parentheses. Significance is indicated as follows: $^*p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

are comparable in terms of the proportion of males, higher-educated individuals, and individuals with a long-standing illness in the population. However, on average municipalities with a historic city center have a significantly larger share of youth in the population, a lower average household size, a lower median disposable income, and a relatively larger immigrant population. Hence, there appear to be considerable differences between municipalities with and without a historic city center, shedding doubt on the validity of the independence assumption.

Table 6: Descriptive statistics for municipalities with and without a historic city center

	No historic city center (n = *190*)		*Historic city center* (n = *108*)		
Variable	Mean	St. Dev.	Mean	St. Dev.	t-stat.
Male^s	0.498	0.008	0.497	0.007	1.137
Age 15-24^s	0.136	0.016	0.142	0.024	-2.087^{**}
Higher education^s	0.271	0.078	0.270	0.074	0.119
Long-standing illness^s	0.326	0.033	0.328	0.030	-0.0667
Average household size	2.288	0.158	2.178	0.171	5.511^{***}
Median disposable income	29.169	2.322	27.428	2.081	6.653^{***}
Migration background^s	0.152	0.082	0.195	0.100	-3.859^{***}

*Note: * indicates that the variable is measured as the share of a municipality's population.*

A possible way for unobserved differences to lead to biased estimates is through the relationship between the degree of urbanization and sports participation. Therefore, I attempt to falsify this potential violation of the independence assumption by regressing the sports participation rate on population size. Population size is a proxy for the degree of urbanization in a municipality, which is positively correlated with the probability of having a historic city center and may relate to the rate of sports participation through unobservable characteristics of highly urbanized municipalities. A significant result would thus imply that having a historic city center is associated with factors that influence sports participation other than municipal sports expenditures, and, therefore, indicate that the instrument is not independent. Yet, the left column in Table A8 shows that the sports participation rates in municipalities with 50,000 to 100,000 and more than 100,000 inhabitants are not significantly different from the sports participation rate in municipalities with a population of 20,000 to 50,000 people. On the other hand, municipalities with less than 20,000 inhabitants do have a significantly lower sports participation rate at the 95% confidence level. However, an F-test for the joint significance of the population size coefficients is not significant.[1] So, I find no evidence for a possible violation of the independence assumption through the relationship between the population size of a municipality and the probability of having a historic city center.

Table A9 shows that the 2SLS results are robust against the period used to calculate the average municipal spending. Unlike for the OLS regressions, omitting the 2020 expenditures does not lead to larger or significant

[1] The F-statistic is 2.090 with 3 and 276 degrees of freedom.

coefficients. Due to a weak first stage, this cannot be confirmed for sports facilities. Yet, as these expenditures are also taken into account in the total sports spending, it is unlikely that the outcome would be different for sports facilities expenditures.

In addition, I run a 2SLS regression using the sports federation membership rate as the dependent variable to show that the absence of a significant relationship does not stem from the dependent variable that is used. Table A10 and Table A11 show that the coefficients remain insignificant and are smaller in size than the coefficients for the sports participation regressions. Moreover, the estimated 95% AR confidence intervals are not significantly different from 0 either.[2] This is consistent with the OLS results, providing additional evidence that the relationship between sports federation membership and municipal sports expenditures is weaker than the one between sports participation and sports expenditures.

Furthermore, columns 3 and 4 in both Table A10 and Table A11 show that the estimated relationship between sports federation membership and sports expenditures changes little when municipalities in the province of Utrecht are included in the sample. It is worth noting that the strength of the instrument does decrease when Utrecht municipalities are taken into account, so these results should be interpreted with caution. This is especially the case for the total sports expenditures reported in Table A10.

5.2 Relationship between City Sports Expenditures and Exercise Participation

5.2.1 OLS Regression Results

I now turn to the analysis of the relationship between municipal sports expenditures and exercise participation, which I measure as the share of the municipality population with an active lifestyle. As Table 7 shows, sports expenditures are not significantly associated with exercise participation. Likewise, the coefficients for sports policy and activation expenditures and sports facilities expenditures are insignificant as well, both independently and jointly.[3] Furthermore, in contrast with the results for sports participation, the sign of the coefficient for Public parks and Recreation expenditures is positive. Yet, again this outcome is not statistically significant. The OLS estimates thus do not provide evidence for a relationship between municipal sports expenditures and exercise participation.

All control variables have the expected sign. Yet, the significance levels show that the determinants of the sports participation rate are not necessarily determinants of the exercise participation rate as well. Firstly, a youthful population is not associated with a higher level of exercise participation. A plausible explanation for this is that activities commonly done by older adults, such as cycling or walking, are also part of an active lifestyle. Secondly, the average household size does not have a significant influence on exercise participation, perhaps because larger households require more physical activity within the household. A third finding is that income is not related to exercise participation. This seems surprising, but may be explained by the

[2] The 95% AR confidence interval is [-0.103, 0.146] for the coefficient in model 2 of Table A10 and [-0.028, 0.036] for the coefficient in model 2 of Table A11

[3] The F-statistic for the joint significance of the sports policy and activation expenditures and sports facilities expenditures variables is 1.407 with 277 and 2 degrees of freedom for the coefficients in column 3 of Table 7 and 0.973 with 276 and 2 degrees of freedom in column 4.

Table 7: OLS estimates of the relationship between municipal sports expenditures and share of the population with an active lifestyle in Dutch municipalities

	Dependent variable:			
	Active lifestyle			
	(1)	(2)	(3)	(4)
Sports expenditures (log)	0.008	0.006		
	(0.006)	(0.006)		
Sports policy and activation expenditures (log)			0.004	0.003
			(0.003)	(0.003)
Sports facilities expenditures (log)			0.007	0.006
			(0.005)	(0.005)
Public parks and Recreation expenditures (log)		0.014		0.013
		(0.009)		(0.009)
Male^s	0.980**	1.006**	1.003**	1.023**
	(0.454)	(0.451)	(0.454)	(0.451)
Age 15-24^s (log)	0.037	0.041	0.039	0.042
	(0.026)	(0.026)	(0.026)	(0.026)
Higher education^s (log)	0.094***	0.098***	0.093***	0.098***
	(0.019)	(0.019)	(0.019)	(0.019)
Average household size (log)	−0.076	−0.046	−0.078	−0.050
	(0.072)	(0.073)	(0.072)	(0.073)
Median disposable income	0.001	−0.0002	0.001	0.0001
	(0.003)	(0.003)	(0.003)	(0.003)
Long-standing illness^s	−0.298***	−0.273**	−0.297***	−0.272**
	(0.109)	(0.112)	(0.109)	(0.112)
Migration background^s (log)	−0.018**	−0.017**	−0.017**	−0.017**
	(0.008)	(0.008)	(0.008)	(0.008)
Constant	0.302	0.238	0.288	0.229
	(0.221)	(0.227)	(0.222)	(0.229)
Province dummies	Yes	Yes	Yes	Yes
Observations	298	298	298	298
R^2	0.508	0.512	0.510	0.514

Note: The table reports the estimates of OLS regressions with the share of a municipality's population with an active lifestyle (according to Dutch guidelines) as the dependent variable. The municipal expenditure variables are measured as the average yearly per capita expenditures on an expenditure category between 2017 and 2020. Variables denoted with ^{s} are measured as a share of the municipality population. Robust standard errors are reported in parentheses. Significance is indicated as follows: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

fact that low-income individuals cannot afford luxury products that diminish the amount of physical activity necessary for everyday tasks. For instance, people who cannot afford a car are more likely to use a bike as a means of transport. Lastly, the share of individuals with a migration background has a significant negative relationship with exercise participation. As the analysis for sports participation showed that immigrants do not necessarily play less sports, this could be due to the fact that immigrants often live in cities. Therefore, they have to travel less far to school or work, reducing the amount of exercise they require for traveling. The remaining coefficients all have a significant influence on exercise participation rates, consistent with the findings in prior literature (Trost et al., 2002).

In a similar vein to the analysis for sports participation, I test the robustness of the results to the period the municipal spending was done. Table A12 reports that the coefficients of the expenditure variables remain insignificant when different sets of years are used. Furthermore, the magnitude of the estimates corresponds with the size of the coefficients for the 2017-2020 period. A notable exception is a specification where only spending in 2019 and 2020 is considered. The coefficients for total sports expenditures and sports facilities expenditures then drop, suggesting that recent sports expenditures are less influential for exercise participation. On the other hand, the Public Parks and Recreation spending coefficient more than doubles in size and becomes significant at the 95% level when only 2019 and 2020 are taken into account. This could be caused by the fact that recent spending more accurately reflects how well public parks or walking and cycling routes are maintained at the time exercise participation is measured. More spending then indicates that facilities required for exercising in public spaces are more attractive and, therefore, are used more often.

5.2.2 2SLS Regression Results

As the final part of the analysis, I estimate the relationship between municipal sports expenditures and the share of the municipality population with an active lifestyle using 2SLS regression. The results in Table 8 show that the magnitude of the expenditure variables substantially increases compared to the OLS estimates, but the variables remain insignificant. Hence, no evidence is found for a relationship between municipal sports expenditures and exercise participation. The estimated coefficients of the control variables are largely similar to those in the OLS regression.

The first stage regression results in Table A6 indicate that the coefficients may be insignificant due to a weak instrument, as this can result in large standard errors. Therefore, I again compute weak instrument robust confidence intervals using the Anderson-Rubin test. For the total sports expenditures variable this yields a 95% confidence interval of [-0.043, 0.491], and for the sports policy and activation expenditures variable one of [-0.014, 0.100]. So, also after accounting for the weakness of the instrument, no significant relationship between municipal sports expenditures and exercise participation is identified.

Table 8: OLS and 2SLS estimates of the relationship between municipal sports expenditures and share of the population with an active lifestyle in Dutch municipalities

	Dependent variable:			
	Active lifestyle			
	OLS		*2SLS*	
	(1)	(2)	(3)	(4)
Sports expenditures (log)	0.008		0.060	
	(0.006)		(0.052)	
Sports policy and activation expenditures (log)		0.003		0.022
		(0.003)		(0.018)
Male^s	0.980**	1.006**	0.935*	1.131**
	(0.454)	(0.452)	(0.508)	(0.528)
Age 15-24^s (log)	0.037	0.038	0.042	0.050
	(0.026)	(0.027)	(0.028)	(0.031)
Higher education^s (log)	0.094***	0.094***	0.096***	0.098***
	(0.019)	(0.019)	(0.021)	(0.020)
Average household size (log)	−0.076	−0.082	−0.047	−0.099
	(0.072)	(0.072)	(0.082)	(0.079)
Median disposable income	0.001	0.0005	0.003	0.002
	(0.003)	(0.003)	(0.004)	(0.003)
Long-standing illness^s	−0.298***	−0.315***	−0.167	−0.289**
	(0.109)	(0.112)	(0.179)	(0.129)
Migration background^s (log)	−0.018**	−0.018**	−0.018**	−0.023**
	(0.008)	(0.008)	(0.009)	(0.010)
Constant	0.302	0.332	−0.017	0.204
	(0.221)	(0.226)	(0.427)	(0.300)
Province dummies	Yes	Yes	Yes	Yes
Observations	298	298	298	298
R^2	0.508	0.507	0.392	0.427
Instrument F-Stat (1, 278)			5.122**	7.497***

*Note: The table reports the estimates of OLS and 2SLS regressions with the share of a municipality's population with an active lifestyle (according to Dutch guidelines) as the dependent variable. The municipal expenditure variables are measured as the average yearly per capita expenditures on an expenditure category between 2017 and 2020, and are instrumented by a dummy variable for having at least one historic city center in the municipality in the 2SLS regressions. Variables denoted with ^{s} are measured as a share of the municipality population. Robust standard errors are reported in parentheses. Significance is indicated as follows: *p<0.1; **p<0.05; ***p<0.01.*

I also examine the robustness of the instrument with respect to the exclusion restriction and independence assumption for the exercise participation regressions. Firstly, Table A7 displays that there is no significant association between non-sports-related municipal expenditures and the share of the municipality population with an active lifestyle. It thus seems unlikely that the presence of a historic city center affects exercise participation through municipal spending on non-sports-related expenditure categories within the Culture and Recreation cluster.

Secondly, the comparison between the characteristics of municipalities with and without a historic city center in Section 5.1.2 revealed that the instrument may not be independent. Therefore, I regress the share of the municipality population with an active lifestyle on the size of the municipality population to assess whether the instrument could affect exercise participation through its association with urbanization. Yet, Table A8 reports no significant differences in exercise participation across municipalities with different population sizes after controlling for other municipality characteristics. Moreover, the F-test for joint significance of the population size categories is insignificant as well.[4] So, as population size is not linked to the rate of exercise participation, the independence assumption is unlikely to be violated by the relationship between the probability of having a historic city center and the degree of urbanization.

I end the analysis with assessing the sensitivity of the 2SLS regression estimates against the years in which the sports expenditures are incurred. Table A13 shows that the coefficients remain insignificant regardless of the period considered. As in the OLS regressions, a slight drop in the size of the coefficient for sports expenditures is noticeable when only 2019 and 2020 are taken into account, but the difference seems marginal. Therefore, it can be concluded that the results of the 2SLS regressions for exercise participation are robust.

6 Discussion

The results presented in the previous section do not provide evidence in support of the first hypothesis. The OLS and 2SLS regression estimations indicate that there is no significant relationship between the level of municipalities' per capita sports expenditures and the rate of sports participation. Likewise, I do not identify an association between the spending on sports activation policies and sports participation. Spending on sports facilities is significantly correlated with sports participation, especially in specifications that do not include expenditure data from 2020. Nevertheless, the strength of the correlation indicates that the economic impact is marginal. The expenditures on public parks and recreation exhibit a negative relationship with sports participation, although again this relationship is not significant. Altogether, the data thus suggests that there is no positive association between municipal expenditures on sport and sports participation.

There are several explanations for the absence of a relationship between municipal sports expenditures and

[4]The F-statistic is 0.136 with 3 and 276 degrees of freedom.

sports participation. Firstly, it generally takes little time to travel to a sports facility in the Netherlands. This not only applies to facilities within the municipality an individual lives in but also to facilities in surrounding municipalities. Consequently, it is relatively easy for people to play sports in a different municipality than they live in if their local government does not provide adequate sports facilities.

Secondly, spending on activation policies predominantly targets the elderly, disabled persons, and children. Regarding the elderly, it is likely that only a small group will start exercising. The impact this has on the municipal participation rate is slight. Similar reasoning applies to the disabled, as they are only a relatively small part of the population. An effect of stimulating children is not observed in this study, as I only consider the sports participation of adults. Nevertheless, I still do not find a significant relationship when I use the sports association membership rate as a dependent variable, even though children are included in this measure.

Thirdly, higher sports facilities expenditures are unlikely to create a large impact on municipal sports participation. Membership fees of sports clubs are not necessarily lower in municipalities that spend more on sports facilities (Davids et al., 2020). A sports club membership is thus not always more accessible for low-income individuals in municipalities with high sports facilities expenditures. Moreover, even if municipal sports spending leads to lower membership fees it would only target a small part of the population. It can be assumed that the majority of people who play sports would also be willing to pay a little extra for their membership. Furthermore, a large share of the people that do not exercise will remain inactive even if the price of sports is lower. Lower sports club membership fees thus only affect the sports participation decision of a limited number of people.

Lastly, the estimates of the 2SLS regressions may be biased. The robustness checks revealed that there are considerable differences in the background characteristics of municipalities with and without a historic city center. Therefore, the instrument may not be independent. Additional analysis showed that the independence assumption is not violated through the degree of urbanization. However, there could be other (unobservable) characteristics of municipalities with a historic city center that correlate with sports participation and are not included as covariates. This can lead to estimates that are even more biased than OLS estimates, especially because the instrument is weak. Nevertheless, it is impossible to know whether there is bias and, if so, the direction and magnitude of the bias. Therefore, the results should be interpreted carefully.

Based on the outcomes of the regression analysis, I also reject the second hypothesis. None of the sports expenditure variables displays a significant association with the rate of exercise participation. The coefficient for public park and recreation expenditures is positive but also insignificant. Hence, there appears to be no relationship between the amount of municipal sports spending and the rate of exercise participation.

The explanations for the lack of this relationship are likely to be similar to those discussed for sports participation. Additionally, it is plausible that having an active lifestyle is already more accessible than playing sports. Therefore, policies that encourage people to have an active lifestyle have an even smaller target population than policies promoting sports participation. Furthermore, people are likely willing to travel to neighboring municipalities for well-maintained parks or cycling paths. This can explain why I do not observe a significant correlation between expenditures on public parks and recreation and exercise participation.

Lastly, the regression results do not support the third hypothesis. As both the sports and exercise participation regressions do not yield significant coefficients for the expenditure variables, it is impossible to determine whether the municipal sports expenditures have more effect on sports participation or exercise participation. However, it seems that the relationship with sports participation is stronger because of the significant correlation with sports facilities expenditures. The size of the coefficients for the spending on sports facilities is in the same order of magnitude in both analyses though. Hence, instead, the most plausible conclusion is that municipal sports expenditures are not related to sports or exercise participation.

7 Conclusion

In summary, this thesis examined the relationship between local public sports expenditures and sports participation in Dutch municipalities. Economic theory suggests that such expenditures can increase sports participation as they lower the barriers to playing sports. However, most of the previous empirical research finds no evidence for a significant effect of municipal expenditures on sports participation. Nevertheless, the existing literature generally does not account for simultaneity bias. I have attempted to address this issue by employing Two-Stage Least Squares regressions with a binary variable for whether there is a historic city center in a municipality as an instrument. To estimate the 2SLS regressions, I use municipal expenditure data from 2017 to 2020. I consider both the total sports expenditures and spending earmarked for two specific purposes: spending on activation policies and spending on sports facilities. Additionally, I take into account public parks and recreation expenditures. Furthermore, I use information on local participation rates from 2020 to assess how the expenditures relate to sports participation and exercise participation.

The results of the OLS and 2SLS regressions suggest there is no relationship between municipal sports expenditures and sports participation. I only find spending on sports facilities to be correlated with sports participation, especially when I disregard expenditure data from 2020. However, the economic size of this effect is marginal. These results are robust against the considered period of municipal spending and an alternative measure of sports participation. I also do not find proof of an association between sports expenditures and exercise participation. Finally, municipal spending on public parks and recreation also does not correlate with the level of sports or exercise participation. Hence, as an answer to the research question, I must conclude there is no evidence for an influence of municipal expenditures on regular physical activity in the Netherlands.

The main policy implication of this thesis is that increasing municipal sports expenditures is unlikely to be an effective tool for raising mass sports participation rates. Nevertheless, this does not imply that municipal sports expenditures are unnecessary or a misuse of public funds. Local governments' sports spending may, for instance, do have an impact on the sports participation of children in low-income households, the elderly, and the disabled. If governments or society place value on these groups having access to sports, then the government spending is still justifiable. Investigating the relationship between municipal sports expenditures and the sports participation of these specific groups thus provides an interesting avenue for further research. Furthermore, I do not take into account the frequency of sports participation. It may be that people who exercise regularly already start exercising even more if their local government provides good sports facilities. In that case, additional municipal sports spending is still valuable.

There are several limitations to this study. Most importantly, the differences in observable characteristics of municipalities with and without a historic city center indicate that the instrument may not be independent, leading to biased estimates. Moreover, the instrument correlates only weakly with sports expenditures and is thus unlikely to correct for the bias induced by the violation of the independence assumption. As good instruments are scarce, future literature thus may consider alternative empirical approaches to identify an unbiased relationship between public sports expenditures and sports participation.

The data exhibits several shortcomings as well. Regarding the expenditure data, municipalities do not consistently record similar types of spending in the same category. For instance, some municipalities see the salaries of sports facility personnel as sports facilities expenditures, whereas others would record this as sports policy and activation expenditures. Moreover, this might even change within a municipality from year to year. This is not an issue for the total sports expenditures. Yet, it may affect the results of the regressions where spending on individual categories is considered.

The main threat to the validity of the participation data is that participation for a large part was measured during the Covid-19 pandemic. Although other research states that Covid-19 has not had a great impact on sports participation, it is still possible that the pandemic has affected the way people exercise. Therefore, it may be interesting to repeat this study in the future with new data.

Finally, future literature on local public expenditures and sports participation could aim to incorporate the sports spending of surrounding municipalities. This way, it can be investigated how mobile people are when it comes to the place they exercise. Furthermore, further research can distinguish between municipalities that manage the municipal sports facilities themselves and those that employ a commercial party for the management and exploitation of sports facilities. An increasing number of municipalities opt for the latter option, so it is interesting to see if these municipalities can obtain the same or higher participation rates with lower per capita sports spending.

References

Andrews, D., & Stock, J. H. (2005). *Inference with weak instruments.* National Bureau of Economic Research Cambridge, Mass., USA.

Angrist, J. D., & Pischke, J.-S. (2008). *Mostly harmless econometrics.* Princeton university press.

Association of Sports and Municipalities. (n.d.). *Wat is een buurtsportcoach?* Retrieved from `https://sportindebuurt.nl/buurtsportcoach/wat-is-een-buurtsportcoach/`

Bailey, R., Hillman, C., Arent, S., & Petitpas, A. (2013). Physical activity: an underestimated investment in human capital? *Journal of physical activity and health*, *10*(3), 289–308.

Becker, G. S. (1965). A theory of the allocation of time. *The economic journal*, *75*(299), 493–517.

Breuer, C., Hallmann, K., & Wicker, P. (2011). Determinants of sport participation in different sports. *Managing Leisure*, *16*(4), 269–286.

Dallmeyer, S., Wicker, P., & Breuer, C. (2017). Public expenditure and sport participation: An examination of direct, spillover, and substitution effects. *International Journal of Sport Finance*, *12*(3).

Davids, A., Hover, P., de Boer, W., Schoemaker, J., Berfelo, B., van Bezooijen, B., & van Rooijen, M. (2020). *Contributiemonitor 2019/2020.* Utrecht: Mulier Instituut. Retrieved from `https://www.kennisbanksportenbewegen.nl/?file=10191&m=1607512339&action=file.download`

Dodge, Y. (2008). *The concise encyclopedia of statistics.* Springer Science & Business Media.

Downward, P. (2007). Exploring the economic choice to participate in sport: Results from the 2002 general household survey. *International review of applied economics*, *21*(5), 633–653.

Downward, P., Dawson, A., & Dejonghe, T. (2009). *Sports economics.* Routledge.

Duijvestijn, M., Schurink, T., Van den Berg, S., Euser, S., & Wendel-Vos, W. (2021). *Sport- en beweeggedrag 2020.* Utrecht: National Institute for Public Health and the Environment. Retrieved from `https://www.sportenbewegenincijfers.nl/sites/default/files/2021-08/2021-0117%20met%20omslag%20beveiligd.pdf`

Eberth, B., & Smith, M. D. (2010). Modelling the participation decision and duration of sporting activity in scotland. *Economic modelling*, *27*(4), 822–834.

Everitt, B. S., & Skrondal, A. (2010). The cambridge dictionary of statistics.

Fridberg, T. (2010). Sport and exercise in denmark, scandinavia and europe. *Sport in Society*, *13*(4), 583–592.

García, J., Lera-López, F., & Suárez, M. J. (2011). Estimation of a structural model of the determinants of the time spent on physical activity and sport: Evidence for spain. *Journal of Sports Economics*, *12*(5), 515–537.

Ghosh, D., & Vogt, A. (2012). Outliers: An evaluation of methodologies. In *Joint statistical meetings* (Vol. 2012).

Grossman, M. (1972). On the concept of health capital and the demand for health. *The Journal of Political Economy*, *80*(2), 223–255.

Heckman, J. J. (2015). Introduction to a theory of the allocation of time by gary becker. *The Economic Journal*, *125*(583), 403–409.

Hoekman, R. (2018). *Sport policy, sport facilities and sport participation: a socio-ecological approach* (Unpublished doctoral dissertation). [Nijmegen: Radboud Universiteit].

Hoekman, R., Breedveld, K., & Kraaykamp, G. (2017). Sport participation and the social and physical environment: explaining differences between urban and rural areas in the netherlands. *Leisure studies*, *36*(3), 357–370.

Hoogendoorn, M. P., & De Hollander, E. L. (2016). *Belemmeringen en drijfveren voor sport en bewegen bij ondervertegenwoordigde groepen.* Bilthoven: National Institute for Public Health and the Environment. Retrieved from `https://www.rivm.nl/bibliotheek/rapporten/2016-0201.pdf`

Hovemann, G., & Wicker, P. (2009). Determinants of sport participation in the european union. *European journal for sport and society*, *6*(1), 51–59.

Humphreys, B. R., & Ruseski, J. E. (2007). Participation in physical activity and government spending on parks and recreation. *Contemporary economic policy*, *25*(4), 538–552.

Kenniscentrum Sport en Bewegen. (n.d.). *Beweegrichtlijnen.* Retrieved from `https://www.kenniscentrumsportenbewegen.nl/producten/beweegrichtlijnen/`

Kokolakakis, T., Lera-López, F., & Castellanos, P. (2014). Regional differences in sports participation: The case of local authorities in england. *International journal of sport finance*, *9*(2), 149–171.

Kokolakakis, T., Lera-Lopez, F., & Panagouleas, T. (2012). Analysis of the determinants of sports participation in spain and england. *Applied Economics*, *44*(21), 2785–2798.

KPMG. (2019). *Brancherapport sport.* Amstelveen: KPMG Advisory N.V. Retrieved from `https://www.kennisbanksportenbewegen.nl/?file=9752&m=1573120314&action=file.download`

Ministry of Health, Welfare and Sport. (2018). *Nationaal preventieakkoord: Naar een gezonder nederland.* The Hague: Ministry of Health, Welfare and Sport. Retrieved from `https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/convenanten/2018/11/23/nationaal-preventieakkoord/nationaal-preventieakkoord.pdf`

National Institute for Public Health and the Environment. (2022). *Sportakkoord: Inclusief sporten en bewegen.* Retrieved from `https://www.sportenbewegenincijfers.nl/beleid/sportakkoord/inclusief-sporten#Sportdeelname-wekelijks`

Netherlands Sports Council. (2020). *Het speelveld van de sport. discussienota over de organisatie en financiering van de sport in de toekomst.* Retrieved from `https://www.nederlandse-sportraad.nl/adviezen/documenten/publicaties/2020/07/10/het-speelveld-van-de-sport.-discussienota-over-de-organisatie-en-financiering-van-de-sport-in-de-toekomst`

Paolino, P. (2001). Maximum likelihood estimation of models with beta-distributed dependent variables. *Political Analysis*, *9*(4), 325–346.

Pulles, I., Van Eldert, P., & Van de Poel, H. (2021). *Monitor sport en corona iv: De gevolgen* *van coronamaatregelen voor de sportsector.* Utrecht: Mulier Instituut. Retrieved from https://www.kennisbanksportenbewegen.nl/?file=10539&m=1637758894&action=file.download

Schadenberg, B., & Van Eldert, P. (2020). *Kostendekkendheid van gemeentelijke sportaccommodaties.* Utrecht: Mulier Instituut. Retrieved from https://www.kennisbanksportenbewegen.nl/?file=10089&m=1597060572&action=file.download

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*(3/4), 591–611.

Statistics Netherlands. (n.d.-a). *Gezondheidsmonitor volwassenen en ouderen 2020.* Retrieved from https://www.cbs.nl/nl-nl/onze-diensten/methoden/onderzoeksomschrijvingen/korte-onderzoeksbeschrijvingen/gezondheidsmonitor-volwassenen-en-ouderen-2020

Statistics Netherlands. (n.d.-b). *Standardised income.* Retrieved from https://www.cbs.nl/en-gb/onze-diensten/methods/definitions/standardised-income

Steckenleiter, C., Lechner, M., Pawlowski, T., & Schüttoff, U. (2019). *Do local public expenditures on sports facilities affect sports participation in germany?* (Discussion Paper no. 2019-05). University of St. Gallen, School of Economics and Political Science.

Studer, F., Schlesinger, T., & Engel, C. (2011). Socio-economic and cultural determinants of sports participation in switzerland from 2000 to 2008. *European Journal for Sport and Society*, *8*(3), 147–166.

Taylor, P., & Gratton, C. (2002). *The economics of sport and recreation: an economic analysis.* Routledge.

Trost, S. G., Owen, N., Bauman, A. E., Sallis, J. F., & Brown, W. (2002). Correlates of adults' participation in physical activity: review and update. *Medicine & science in sports & exercise*, *34*(12), 1996–2001.

Van den Berg, L. (2021, September 6). *Vind de weg naar de juiste regelingen voor meedoen met sport en bewegen.* Retrieved from https://www.allesoversport.nl/thema/beleid/vind-de-weg-naar-de-juiste-regelingen-voor-meedoen-met-sport-en-bewegen/

Van den Dool, R. (2021). *Sportdeelname en lidmaatschap 2012-2020.* Utrecht: Mulier Instituut. Retrieved from https://www.kennisbanksportenbewegen.nl/?file=10493&m=1634128676&action=file.download

Van den Dool, R., & Hoekman, R. H. A. (2017). *Monitor sportuitgaven gemeenten: een overzicht van de ontwikkelingen (2010-2017).* Utrecht: Mulier Instituut. Retrieved from https://www.kennisbanksportenbewegen.nl/?file=8617&m=1518787329&action=file.download

Van den Dool, R., & Van Eldert, P. (2021). *Monitor sportuitgaven gemeenten 2019: Een overzicht van de uitgaven in 2019 plus samenhangende vraagstukken.* Utrecht: Mulier Instituut. Retrieved from https://www.kennisbanksportenbewegen.nl/?file=10263&m=1612955656&action=file.download

Van Eldert, P., & Beekman, T. L. (2021). *Monitor sportuitgaven gemeenten 2020: Gemeentelijke uitgaven aan sport in 2020 en verdieping op specifieke vraagstukken.* Utrecht: Mulier Instituut. Retrieved from https://www.kennisbanksportenbewegen.nl/?file=10608&m=1642161342&action=file.download

Van Tuyckom, C., Scheerder, J., & Bracke, P. (2010). Gender and age inequalities in regular sports participation: A cross-national study of 25 european countries. *Journal of sports sciences*, *28*(10), 1077–1084.

Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach.* Cengage learning.

World Health Organization. (2020, November 26). *Physical activity.* Retrieved from `https://www.who.int/news-room/fact-sheets/detail/physical-activity`

Appendix A

Table A1: Descriptive statistics for municipalities in the sample and municipalities in the province of Utrecht

	Sample (n = *298*)		*Utrecht* (n = *25*)		
Variable	Mean	St. Dev.	Mean	St. Dev.	t-stat.
Sports federation membership^s	0.195	0.039	0.208	0.028	2.116**
Sports expenditures (in €)	80.692	28.524	75.483	28.533	−0.844
Sports policy and activation expenditures (in €)	16.412	14.123	14.297	17.807	−0.556
Sports facilities expenditures (in €)	64.280	27.752	61.186	22.451	−0.625
Public parks and Recreation expenditures (in €)	88.009	30.048	76.285	13.153	−3.609***
Male^s	0.498	0.008	0.494	0.006	−3.208***
Age 15-24^s	0.138	0.020	0.145	0.014	2.445**
Higher education^s	0.270	0.077	0.316	0.085	2.618**
Long-standing illness^s	0.327	0.032	0.310	0.023	−3.295***
Migration background^s	0.168	0.091	0.159	0.052	−0.698
Average household size	2.248	0.171	2.332	0.146	2.707**
Median disposable income	28.538	2.386	30.784	1.451	6.986***

*Note: ^{s} indicates that the variable is measured as the share of a municipality's population. Two municipalities in Utrecht were part of a municipal merger between 2017 and 2020, so the averages of the expenditure variables are calculated using the remaining 23 municipalities. Significance is indicated as follows: *p<0.1; **p<0.05; ***p<0.01.*

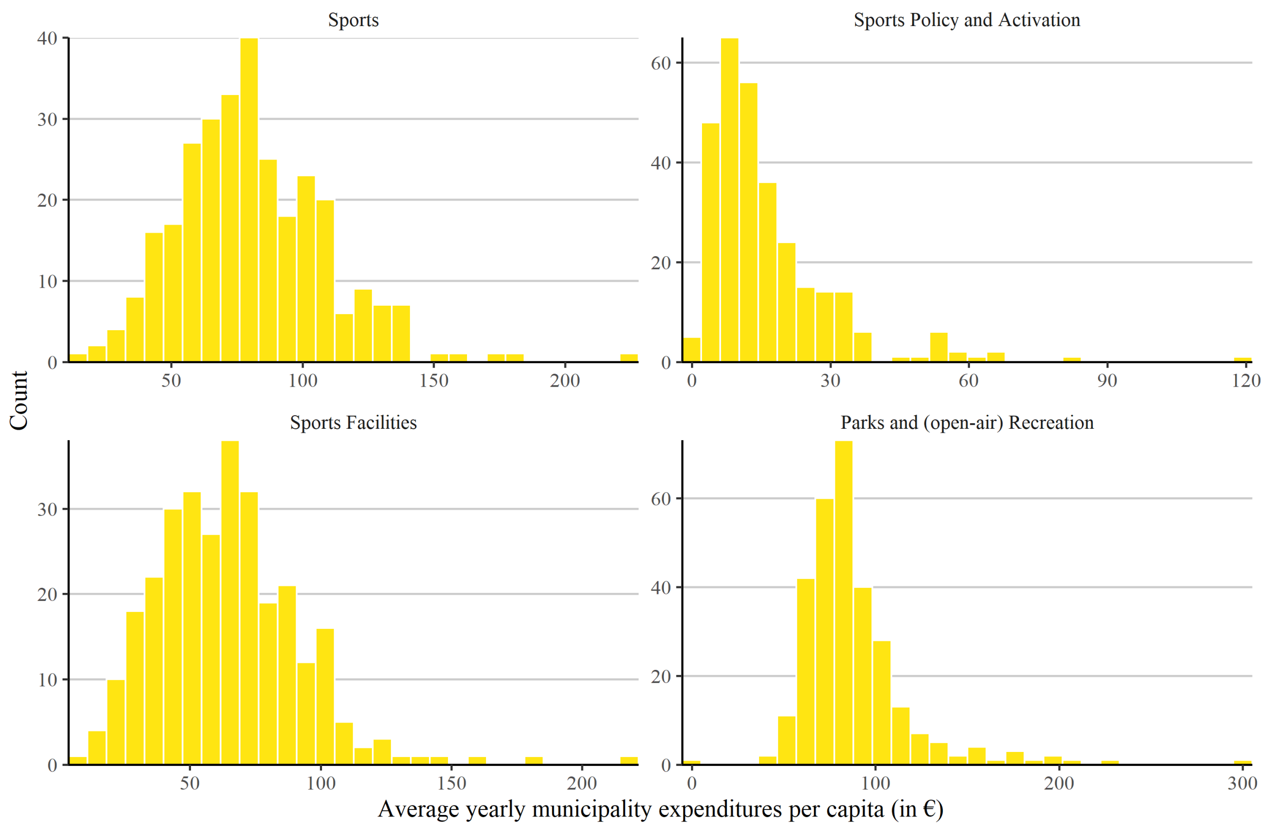


Figure A1: Distribution of sports-related municipal expenditures per capita in the Netherlands

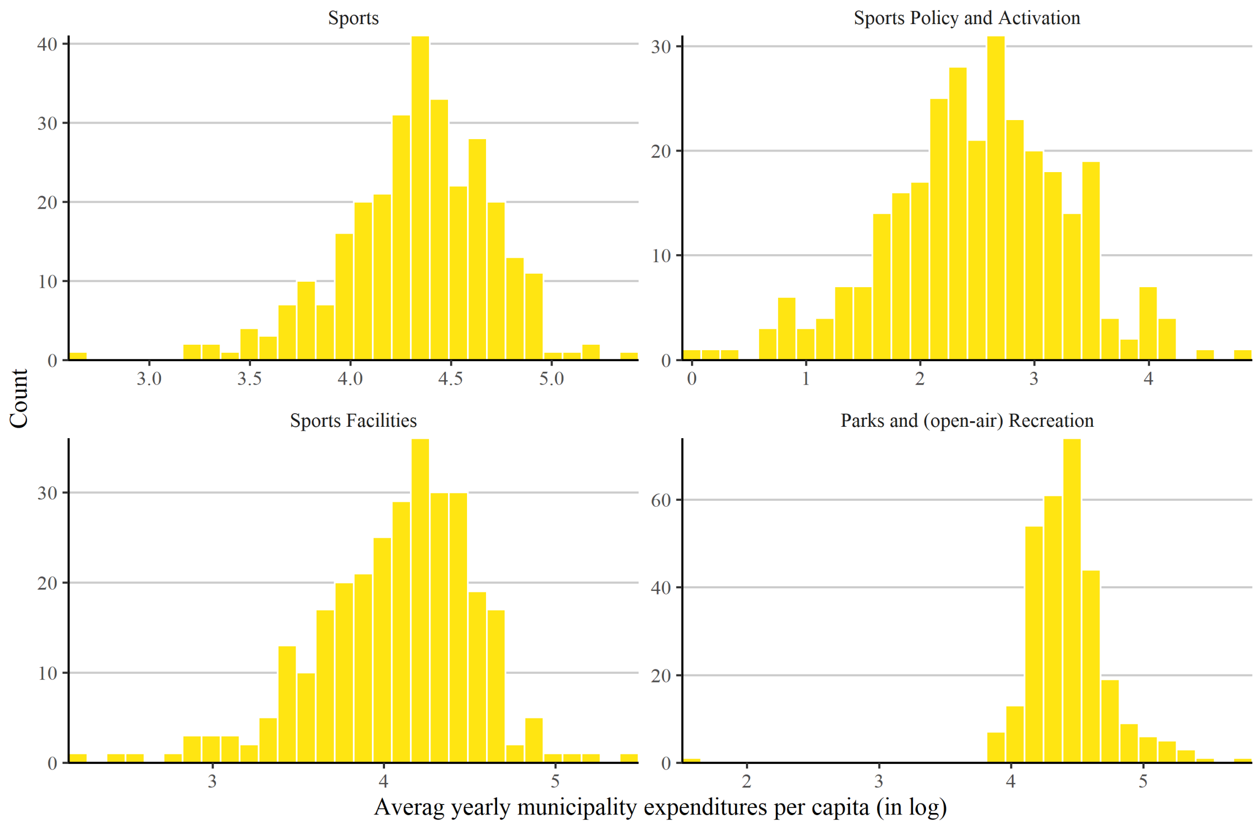


Figure A2: Distribution of sports-related municipal expenditures per capita in the Netherlands after log-transformation

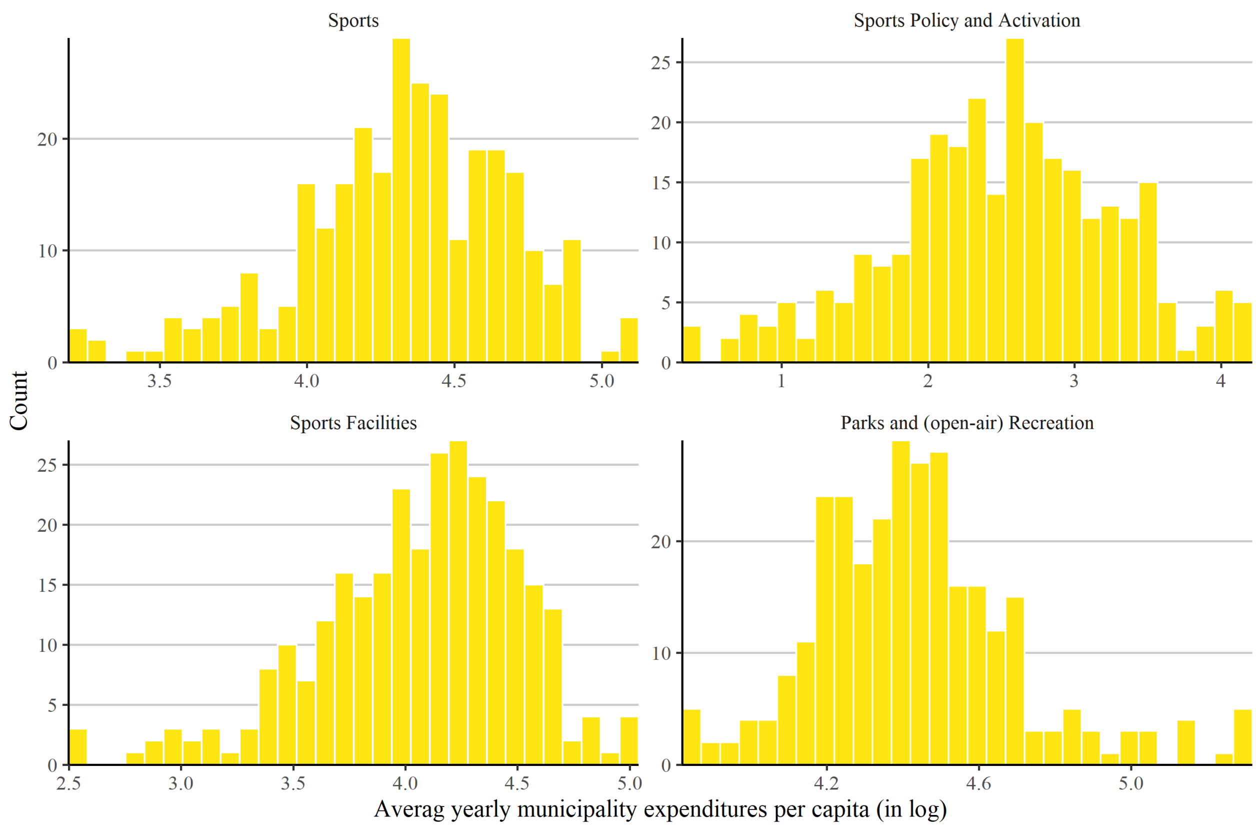


Figure A3: Distribution of sports-related municipal expenditures per capita in the Netherlands after log-transformation and 98% winsorization

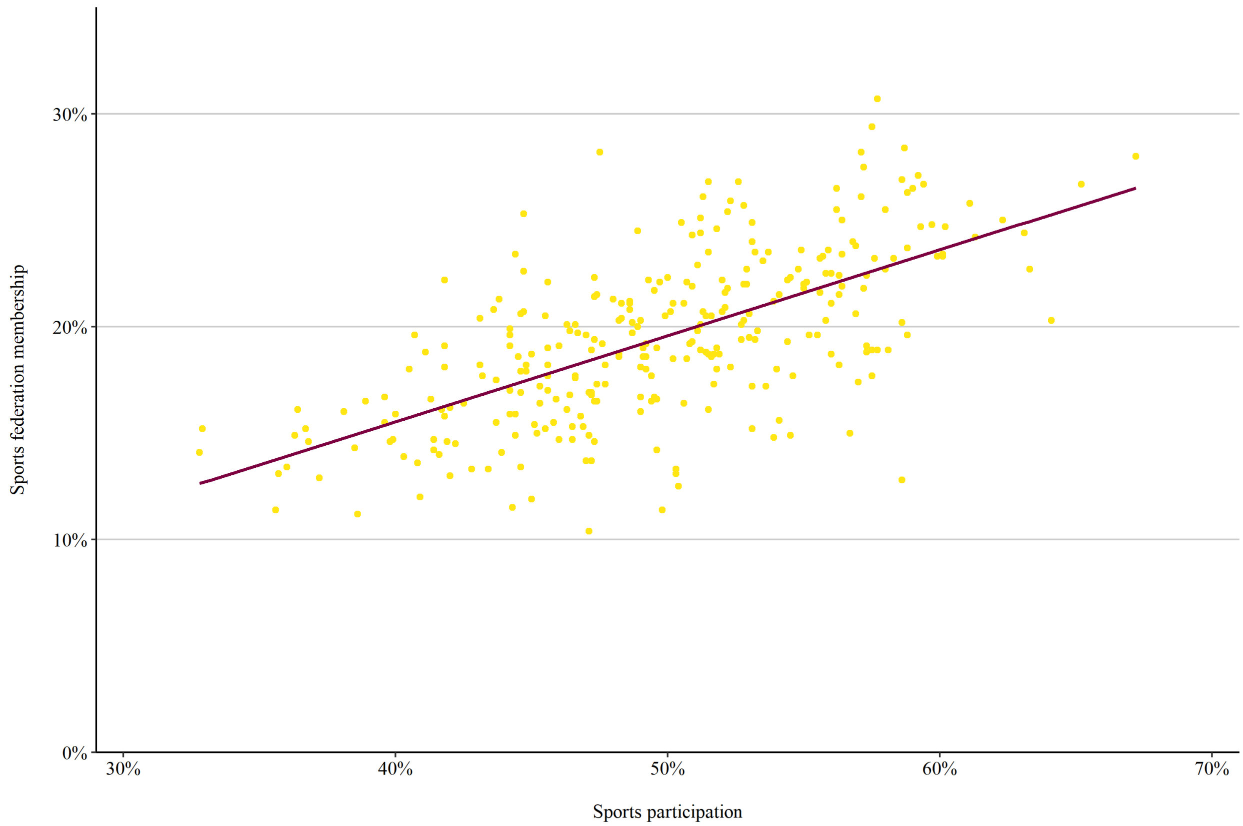


Figure A4: Correlation between sports participation rates and sports federation membership rates in Dutch municipalities

Table A2: OLS estimates of the relationship between municipal sports expenditures and the sports participation rate in Dutch municipalities for samples with different treatment of outliers

	Dependent variable:					
	Sports participation					
	Unadjusted data		*No outliers*		*Winsorized*	
	(1)	(2)	(3)	(4)	(5)	(6)
Sports expenditures (log)	0.006	0.007	0.010*	0.011*	0.007	0.008
	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)
Public parks and Recreation expenditures (log)		−0.007		−0.011		−0.012
		(0.006)		(0.008)		(0.008)
Male^s	1.089***	1.091***	1.046***	1.025***	1.087***	1.064***
	(0.336)	(0.335)	(0.338)	(0.338)	(0.335)	(0.335)
Age 15-24^s (log)	0.073***	0.071***	0.071***	0.068***	0.073***	0.070***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
Higher education^s (log)	0.122***	0.120***	0.123***	0.120***	0.122***	0.118***
	(0.014)	(0.015)	(0.015)	(0.015)	(0.014)	(0.015)
Average household size (log)	−0.112**	−0.127**	−0.111*	−0.134**	−0.112**	−0.137**
	(0.055)	(0.057)	(0.057)	(0.059)	(0.055)	(0.057)
Median disposable income	0.008***	0.008***	0.008***	0.008***	0.008***	0.008***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Long-standing illness^s	−0.336***	−0.350***	−0.332***	−0.353***	−0.334***	−0.357***
	(0.088)	(0.089)	(0.089)	(0.090)	(0.088)	(0.089)
Migration background^s (log)	−0.004	−0.004	−0.004	−0.004	−0.004	−0.004
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Constant	0.223	0.249	0.234	0.282	0.220	0.275
	(0.176)	(0.177)	(0.177)	(0.180)	(0.176)	(0.179)
Province dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	298	298	293	293	298	298
R^2	0.789	0.790	0.792	0.793	0.789	0.791

Note: The table reports the estimates of OLS regressions with the share of a municipality's population that exercises at least once per week as the dependent variable. The municipal expenditure variables are measured as the average yearly per capita expenditures on an expenditure category between 2017 and 2020. Variables denoted with ^{s} are measured as a share of the municipality population. For the regression in columns 3 and 4, I drop the municipalities at the top and bottom 1% of the sports expenditures distribution. For the regression in columns 5 and 6, I replace the values of these municipalities with the value of the municipality with the highest sports expenditures in the 98^{th} percentile and the lowest sports expenditures in the 2^{nd} percentile, respectively. Standard errors are reported in parentheses. Significance is indicated as follows: *p<0.1; **p<0.05; ***p<0.01.

Table A3: OLS estimates of the relationship between municipal sports expenditures and share of the population with an active lifestyle in Dutch municipalities for samples with different treatment of outliers

	Dependent variable:					
	Active lifestyle					
	Unadjusted data		*No outliers*		*Winsorized*	
	(1)	(2)	(3)	(4)	(5)	(6)
Sports expenditures (log)	0.006	0.005	0.008	0.006	0.008	0.006
	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)
Public parks and Recreation expenditures (log)		0.012^{*}		0.015		0.014
		(0.007)		(0.009)		(0.009)
Male^s	0.983^{**}	0.981^{**}	0.903^{*}	0.934^{**}	0.980^{**}	1.006^{**}
	(0.453)	(0.449)	(0.461)	(0.456)	(0.454)	(0.451)
Age 15-24^s (log)	0.037	0.040	0.035	0.038	0.037	0.041
	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)
Higher education^s (log)	0.094^{***}	0.097^{***}	0.094^{***}	0.099^{***}	0.094^{***}	0.098^{***}
	(0.019)	(0.019)	(0.020)	(0.020)	(0.019)	(0.019)
Average household size (log)	−0.076	−0.051	−0.076	−0.044	−0.076	−0.046
	(0.072)	(0.071)	(0.077)	(0.077)	(0.072)	(0.073)
Median disposable income	0.001	−0.0002	0.001	−0.00005	0.001	−0.0002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Long-standing illness^s	-0.302^{***}	-0.279^{**}	-0.308^{***}	-0.279^{**}	-0.298^{***}	-0.273^{**}
	(0.110)	(0.111)	(0.112)	(0.114)	(0.109)	(0.112)
Migration background^s (log)	-0.017^{**}	-0.018^{**}	-0.018^{**}	-0.018^{**}	-0.018^{**}	-0.017^{**}
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Constant	0.310	0.268	0.332	0.262	0.302	0.238
	(0.221)	(0.222)	(0.223)	(0.228)	(0.221)	(0.227)
Province dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	298	298	293	293	298	298
R^2	0.507	0.512	0.511	0.516	0.508	0.512

Note: The table reports the estimates of OLS regressions with the share of a municipality's population with an active lifestyle (according to Dutch guidelines) as the dependent variable. The municipal expenditure variables are measured as the average yearly per capita expenditures on an expenditure category between 2017 and 2020. Variables denoted with ^{s} are measured as a share of the municipality population. For the regression in columns 3 and 4, I drop the municipalities at the top and bottom 1% of the sports expenditures distribution. For the regression in columns 5 and 6, I replace the values of these municipalities with the value of the municipality with the highest sports expenditures in the 98^{th} percentile and the lowest sports expenditures in the 2^{nd} percentile, respectively. Standard errors are reported in parentheses. Significance is indicated as follows: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table A4: OLS estimates of the relationship between sports participation and municipal sports expenditures in various spending periods

	Dependent variable:									
	Sports participation									
	2017-2019		2018-2020		2017-2018		2018-2019		2019-2020	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Sports expenditures (log)	0.009*		0.008		0.009*		0.010*		0.004	
	(0.005)		(0.005)		(0.005)		(0.005)		(0.005)	
Sports policy and activation expenditures (log)		0.002		0.001		0.002		0.001		0.002
		(0.002)		(0.002)		(0.002)		(0.002)		(0.002)
Sports facilities expenditures (log)		0.009**		0.007*		0.009**		0.009**		0.004
		(0.004)		(0.004)		(0.004)		(0.004)		(0.004)
Public parks and Recreation expenditures (log)	−0.013*	−0.014*	−0.012	−0.012	−0.012*	−0.013*	−0.014*	−0.014**	−0.006	−0.007
	(0.007)	(0.007)	(0.007)	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Constant	0.273	0.269	0.279	0.274	0.275	0.276	0.277	0.274	0.270	0.260
	(0.178)	(0.178)	(0.179)	(0.178)	(0.179)	(0.178)	(0.177)	(0.176)	(0.180)	(0.180)
Province dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	298	298	298	298	298	298	298	298	298	298
R^2	0.792	0.793	0.791	0.792	0.792	0.793	0.793	0.794	0.789	0.790

*Note: The table reports the estimates of OLS regressions with the share of a municipality's population that exercises at least once per week as the dependent variable. The municipal expenditure variables are measured as the average yearly per capita expenditures on an expenditure category between the years indicated above the columns. Variables denoted with ^{s} are measured as a share of the municipality population. Standard errors are reported in parentheses. Significance is indicated as follows: *p<0.1; **p<0.05; ***p<0.01.*

Table A5: OLS estimates of the relationship between sports participation and municipal sports expenditures for different measures of sports participation

	Dependent variable:					
	Sports federation membership				Sports participation	
	Sample		Sample + Utrecht		Sample	
	(1)	(2)	(3)	(4)	(5)	(6)
Sports expenditures (log)	0.004	0.004	0.005	0.006	0.007	0.008
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
Public parks and Recreation expenditures (log)		−0.003		−0.004		−0.012
		(0.006)		(0.005)		(0.008)
Male^s	0.940***	0.933***	0.816***	0.821***	1.087***	1.064***
	(0.263)	(0.264)	(0.252)	(0.252)	(0.335)	(0.335)
Age 15-24^s (log)	0.019	0.019	0.019	0.018	0.073***	0.070***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.020)	(0.020)
Higher education^s (log)	0.041***	0.040***	0.039***	0.038***	0.122***	0.118***
	(0.011)	(0.012)	(0.010)	(0.010)	(0.014)	(0.015)
Average household size (log)	−0.090**	−0.097**	−0.091**	−0.098**	−0.112**	−0.137**
	(0.043)	(0.045)	(0.041)	(0.041)	(0.055)	(0.057)
Median disposable income	0.009***	0.009***	0.009***	0.009***	0.008***	0.008***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Long-standing illness^s	−0.114	−0.120*	−0.130**	−0.136**	−0.334***	−0.357***
	(0.069)	(0.070)	(0.065)	(0.066)	(0.088)	(0.089)
Migration background^s (log)	−0.028***	−0.028***	−0.028***	−0.028***	−0.004	−0.004
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)
Constant	−0.385***	−0.369***	−0.331**	−0.319**	0.220	0.275
	(0.139)	(0.141)	(0.133)	(0.133)	(0.176)	(0.179)
Province dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	298	298	321	321	298	298
R^2	0.670	0.670	0.672	0.673	0.789	0.791

*Note: The table reports the estimates of OLS regressions with the share of a municipality's population that is a member of a sports federation as the dependent variable in columns 1 to 4, and the share of a municipality's population that exercises at least once per week as the dependent variable in columns 5 and 6. The municipal expenditure variables are measured as the average yearly per capita expenditures on an expenditure category between 2017 and 2020. Variables denoted with ^{s} are measured as a share of the municipality population. The sample for the regressions in columns 3 and 4 consist of the sample used in the main analysis and a sample of 23 municipalities in the province of Utrecht for which only data on sports federation membership is available. Standard errors are reported in parentheses. Significance is indicated as follows: *p<0.1; **p<0.05; ***p<0.01.*

Table A6: First stage estimates of the 2SLS regression of sports participation on municipal sports expenditures

	Dependent variable:		
	Sports expenditures	Policy and activation expenditures	Facilities expenditures
	(1)	(2)	(3)
Historic city area	0.103**	0.282***	0.025
	(0.045)	(0.103)	(0.059)
Male^s	1.094	−6.014	1.182
	(3.763)	(8.535)	(4.918)
Age 15-24^s (log)	−0.096	−0.654	−0.095
	(0.220)	(0.499)	(0.287)
Higher education^s (log)	−0.039	−0.193	0.071
	(0.162)	(0.368)	(0.212)
Average household size (log)	−0.480	1.065	−0.701
	(0.619)	(1.405)	(0.809)
Median disposable income	−0.038*	−0.058	−0.051*
	(0.022)	(0.050)	(0.029)
Long-standing illness^s	−2.290**	−0.672	−2.537**
	(0.985)	(2.234)	(1.287)
Migration background^s (log)	−0.005	0.221	−0.141
	(0.071)	(0.161)	(0.093)
Constant	5.642***	5.324	5.812**
	(1.957)	(4.439)	(2.557)
Province dummies	Yes	Yes	Yes
Observations	298	298	298
R^2	0.205	0.123	0.140
Instrument F-Stat (1, 278)	5.122**	7.497***	0.177

*Note: The table reports the estimates of the first stage OLS regressions. The dependent variables are the total municipal sports expenditures, municipal expenditures on sports policy and activation and municipal expenditures on sports facilities. These variables are measured as the average yearly per capita expenditures between 2017 and 2020. The main independent variable is a dummy that takes a value of 1 if there is a historical city center in a municipality, and 0 otherwise. Variables denoted with ^{s} are measured as a share of the municipality population. Significance is indicated as follows: *p<0.1; **p<0.05; ***p<0.01.*

Table A7: OLS estimates of the relationship between sports- and exercise participation rates and non-sports related municipal expenditures

	Dependent variable:	
	Sports participation	Active lifestyle
	(1)	(2)
Sports expenditures (log)	0.005	0.005
	(0.005)	(0.006)
Culture expenditures (log)	0.003	0.004
	(0.002)	(0.003)
Media expenditures (log)	0.001	0.004
	(0.003)	(0.003)
Constant	0.165	0.208
	(0.182)	(0.235)
Control variables	Yes	Yes
Observations	298	298
R^2	0.791	0.514

Note: The table reports the estimates of OLS regressions with the share of a municipality's population that exercises at least once per week as the dependent variable in column 1, and the share of a municipality's population with an active lifestyle (according to Dutch guidelines) as the dependent variable in column 2. The municipal expenditure variables are measured as the average yearly per capita expenditures on an expenditure category between 2017 and 2020. Standard errors are reported in parentheses. Significance is indicated as follows: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table A8: OLS estimates of the relationship between sports- and exercise participation rates and municipality population size

	Dependent variable:	
	Sports participation	Active lifestyle
	(1)	(2)
Population $< 20,000$	-0.011^{**}	-0.0004
	(0.005)	(0.006)
Population 50,000 - 100,000	-0.002	0.003
	(0.006)	(0.006)
Population $> 100,000$	-0.005	0.005
	(0.008)	(0.010)
Constant	0.173	0.241
	(0.175)	(0.238)
Control variables	Yes	Yes
Observations	298	298
R^2	0.792	0.497
F-stat Population (3, 276)	2.090	0.136

Note: The table reports the estimates of OLS regressions with the share of a municipality's population that exercises at least once per week as the dependent variable in column 1, and the share of a municipality's population with an active lifestyle (according to Dutch guidelines) as the dependent variable in column 2. The size of municipality is a categorical variable with 4 categories. The reference category is municipalities with a population size of 20,000 to 50,000. Standard errors are reported in parentheses. Significance is indicated as follows: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table A9: 2SLS estimates of the relationship between sports participation and municipal sports expenditures in various spending periods

	Dependent variable:									
	Sports participation									
	2017-2019		2018-2020		2017-2018		2018-2019		2019-2020	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Sports expenditures (log)	0.036		0.033		0.037		0.034		0.030	
	(0.043)		(0.039)		(0.045)		(0.041)		(0.036)	
Sports policy and activation expenditures (log)		0.013		0.012		0.014		0.012		0.011
		(0.016)		(0.014)		(0.017)		(0.014)		(0.014)
Constant	0.047	0.188	0.078	0.190	0.044	0.181	0.084	0.205	0.068	0.178
	(0.318)	(0.204)	(0.286)	(0.201)	(0.322)	(0.209)	(0.280)	(0.195)	(0.298)	(0.207)
Province dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	298	298	298	298	298	298	298	298	298	298
R^2	0.769	0.766	0.771	0.769	0.767	0.761	0.772	0.767	0.767	0.772
Instrument F-stat. (1, 278)	4.666**	6.824***	5.404**	8.273***	4.297**	5.654**	4.929**	7.908***	5.671**	7.757***

*Note: The table reports the estimates of 2SLS regressions with the share of a municipality's population that exercises at least once per week as the dependent variable. The municipal expenditure variables are measured as the average yearly per capita expenditures on an expenditure category between the years indicated above the columns. Variables denoted with ^{s} are measured as a share of the municipality population. Standard errors are reported in parentheses. Significance is indicated as follows: *p<0.1; **p<0.05; ***p<0.01.*

Table A10: OLS and 2SLS estimates of the relationship between sports participation and municipal sports expenditures for different measures of sports participation

	Dependent variable:					
	Sports federation membership				Sports participation	
	Sample		Sample + Utrecht		Sample	
	OLS	*2SLS*	*OLS*	*2SLS*	*OLS*	*2SLS*
	(1)	(2)	(3)	(4)	(5)	(6)
Sports expenditures (log)	0.004	0.008	0.005	0.014	0.007	0.034
	(0.004)	(0.031)	(0.004)	(0.039)	(0.005)	(0.041)
Male^s	0.940***	0.936***	0.816***	0.817***	1.087***	1.063***
	(0.263)	(0.265)	(0.252)	(0.255)	(0.335)	(0.353)
Age 15-24^s (log)	0.019	0.020	0.019	0.020	0.073***	0.075***
	(0.015)	(0.016)	(0.015)	(0.016)	(0.020)	(0.021)
Higher education^s (log)	0.041***	0.041***	0.039***	0.039***	0.122***	0.123***
	(0.011)	(0.011)	(0.010)	(0.011)	(0.014)	(0.015)
Average household size (log)	−0.090**	−0.087*	−0.091**	−0.085*	−0.112**	−0.097
	(0.043)	(0.047)	(0.041)	(0.048)	(0.055)	(0.062)
Median disposable income	0.009***	0.009***	0.009***	0.010***	0.008***	0.009***
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.003)
Long-standing illness^s	−0.114	−0.103	−0.130**	−0.106	−0.334***	−0.265*
	(0.069)	(0.105)	(0.065)	(0.121)	(0.088)	(0.139)
Migration background^s (log)	−0.028***	−0.028***	−0.028***	−0.028***	−0.004	−0.005
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.007)
Constant	−0.385***	−0.411*	−0.331**	−0.393	0.220	0.052
	(0.139)	(0.234)	(0.133)	(0.291)	(0.176)	(0.312)
Province dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	298	298	321	321	298	298
R^2	0.670	0.669	0.672	0.665	0.789	0.769
Instrument F-Stat		5.122** (1, 278)		2.826* (1, 301)		5.122** (1, 278)

Note: The table reports OLS and 2SLS regression estimates with the share of a municipality's population that is a member of a sports federation as the dependent variable in columns 1 to 4, and the share of a municipality's population that exercises at least once per week as the dependent variable in columns 5 and 6. Municipal expenditures are measured as the average yearly per capita expenditures on sports between 2017 and 2020. The expenditure variables are instrumented by a dummy that takes a value of 1 if there is a historical city center in the municipality, and 0 otherwise. Variables denoted with ^{s} are measured as a share of the municipality population. The sample for the regressions in columns 3 and 4 consist of the sample used in the main analysis and a sample of 23 municipalities in the province of Utrecht for which only data on sports federation membership is available. Standard errors are reported in parentheses. Significance is indicated as follows: **p<0.1;* ***p<0.05;* ****p<0.01.*

Table A11: OLS and 2SLS estimates of the relationship between sports participation and municipal sports policy and activation expenditures for different measures of sports participation

	Dependent variable:					
	Sports federation membership				Sports participation	
	Sample		Sample + Utrecht		Sample	
	OLS	*2SLS*	*OLS*	*2SLS*	*OLS*	*2SLS*
	(1)	(2)	(3)	(4)	(5)	(6)
Sports policy and activation	0.002	0.003	0.002	0.005	0.0005	0.012
expenditures (log)	(0.002)	(0.011)	(0.002)	(0.014)	(0.002)	(0.015)
Male^s	0.956***	0.962***	0.824***	0.838***	1.096***	1.175***
	(0.264)	(0.274)	(0.253)	(0.261)	(0.337)	(0.366)
Age 15-24^s (log)	0.020	0.021	0.020	0.021	0.072***	0.080***
	(0.015)	(0.017)	(0.015)	(0.017)	(0.020)	(0.023)
Higher education^s (log)	0.041***	0.041***	0.039***	0.039***	0.122***	0.124***
	(0.011)	(0.012)	(0.010)	(0.011)	(0.014)	(0.015)
Average household size (log)	−0.093**	−0.094**	−0.096**	−0.100**	−0.116**	−0.126**
	(0.043)	(0.044)	(0.041)	(0.044)	(0.055)	(0.059)
Median disposable income	0.009***	0.009***	0.009***	0.009***	0.007***	0.008***
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
Long-standing illness^s	−0.120*	−0.119*	−0.140**	−0.135**	−0.351***	−0.335***
	(0.069)	(0.070)	(0.065)	(0.068)	(0.088)	(0.094)
Migration background^s (log)	−0.029***	−0.029***	−0.028***	−0.029***	−0.004	−0.007
	(0.005)	(0.006)	(0.005)	(0.007)	(0.006)	(0.008)
Constant	−0.375***	−0.382**	−0.309**	−0.328**	0.258	0.178
	(0.137)	(0.156)	(0.131)	(0.155)	(0.175)	(0.208)
Province dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	298	298	321	321	298	298
R^2	0.670	0.670	0.671	0.667	0.788	0.768
Instrument F-Stat		7.497*** (1, 278)		4.728** (1, 301)		7.497*** (1, 278)

Note: The table reports OLS and 2SLS regression estimates with the share of a municipality's population that is a member of a sports federation as the dependent variable in columns 1 to 4, and the share of a municipality's population that exercises at least once per week as the dependent variable in columns 5 and 6. The expenditure variable is measured as the average yearly per capita expenditures on sports policy and activation between 2017 and 2020. The expenditure variables are instrumented by a dummy that takes a value of 1 if there is a historical city center in the municipality, and 0 otherwise. Variables denoted with ^{s} are measured as a share of the municipality population. The sample for the regressions in columns 3 and 4 consist of the sample used in the main analysis and a sample of 23 municipalities in the province of Utrecht for which only data on sports federation membership is available. Standard errors are reported in parentheses. Significance is indicated as follows: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table A12: OLS estimates of the relationship between exercise participation and municipal sports expenditures in various spending periods

	Dependent variable:									
	Exercise participation									
	2017-2019		2018-2020		2017-2018		2018-2019		2019-2020	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Sports expenditures (log)	0.007		0.006		0.006		0.007		0.004	
	(0.006)		(0.006)		(0.006)		(0.006)		(0.006)	
Sports policy and activation expenditures (log)		0.003		0.003		0.003		0.004		0.003
		(0.003)		(0.003)		(0.003)		(0.003)		(0.003)
Sports facilities expenditures (log)		0.007		0.005		0.006		0.006		0.003
		(0.005)		(0.005)		(0.005)		(0.005)		(0.005)
Public parks and Recreation expenditures (log)	0.010	0.009	0.014	0.013	0.011	0.010	0.009	0.008	0.021**	0.020**
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)
Constant	0.257	0.248	0.246	0.237	0.252	0.247	0.270	0.263	0.203	0.192
	(0.228)	(0.229)	(0.228)	(0.229)	(0.227)	(0.229)	(0.228)	(0.229)	(0.224)	(0.226)
Province dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	298	298	298	298	298	298	298	298	298	298
R^2	0.511	0.513	0.512	0.514	0.511	0.513	0.510	0.513	0.518	0.520

Note: The table reports the estimates of OLS regressions with the share of a municipality's population with an active lifestyle (according to Dutch guidelines) as the dependent variable. The municipal expenditure variables are measured as the average yearly per capita expenditures on an expenditure category between the years indicated above the columns. Variables denoted with ^{s} are measured as a share of the municipality population. Standard errors are reported in parentheses. Significance is indicated as follows: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table A13: 2SLS estimates of the relationship between exercise participation and municipal sports expenditures in various spending periods

	Dependent variable:									
	Exercise participation									
	2017-2019		2018-2020		2017-2018		2018-2019		2019-2020	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Sports expenditures (log)	0.063		0.057		0.064		0.060		0.052	
	(0.055)		(0.049)		(0.056)		(0.052)		(0.045)	
Sports policy and activation expenditures (log)		0.023		0.021		0.025		0.021		0.020
		(0.019)		(0.017)		(0.022)		(0.017)		(0.017)
Constant	−0.027	0.221	0.028	0.226	−0.032	0.210	0.038	0.251	0.011	0.204
	(0.435)	(0.293)	(0.395)	(0.286)	(0.443)	(0.306)	(0.385)	(0.274)	(0.414)	(0.296)
Province dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	298	298	298	298	298	298	298	298	298	298
R^2	0.381	0.416	0.397	0.440	0.364	0.389	0.388	0.436	0.397	0.439
Instrument F-stat. (1, 278)	4.666**	6.824***	5.404**	8.273***	4.297**	5.654**	4.929**	7.908***	5.671**	7.757***

Note: The table reports the estimates of 2SLS regressions with the share of a municipality's population with an active lifestyle (according to Dutch guidelines) as the dependent variable. The municipal expenditure variables are measured as the average yearly per capita expenditures on an expenditure category between the years indicated above the columns. The expenditure variables are instrumented by a dummy that takes a value of 1 if there is a historical city center in the municipality, and 0 otherwise. Variables denoted with ^{s} are measured as a share of the municipality population. Standard errors are reported in parentheses. Significance is indicated as follows: *$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.*

Appendix B

I assess the validity of the OLS assumptions by performing several (statistical) tests. Firstly, I check whether the predicted values of the dependent variable fall within the [0,1] interval.

Secondly, I use a regression specification error test (RESET) to detect if a model suffers from functional form misspecification. To execute the test, first a regression including the squared and cubed terms of the independent variable is estimated. Then, it is tested whether the coefficients of the polynomials are jointly significant by computing an F statistic. A significant F statistic implies we must reject the null hypothesis of no joint significance and would thus be indicative of a functional form misspecification issue (Wooldridge, 2015).

Thirdly, to confirm that the error terms follow a normal distribution I carry out the Shapiro-Wilk test. This test uses a W statistic, which is calculated as "dividing the square of an appropriate linear combination of the sample order statistics by the usual symmetric estimate of variance" (Shapiro & Wilk, 1965, p. 591). If the W statistic is significantly large, the null hypothesis of normal distributed should be rejected and thus violate the assumption of normally distributed error terms.

Fourthly, I perform the studentized Breusch-Pagan test to verify the error terms are homoscedastic. The first step in this test is to obtain the squared residuals of the OLS regression. Subsequently, a regression with the squared residuals as dependent variable and the independent variables in the main OLS model must be estimated. The R^2 of this regression can then be used to compute the LM statistic. If this statistic is significant, it must be concluded that the errors are heteroscedastic (Wooldridge, 2015).

Lastly, I compute variance inflation factors (VIFs) to assess the degree of multicollinearity. A VIF is calculated by regressing a variable on all other independent variables in the model and plugging the corresponding R^2 into the following formula (Dodge, 2008):

$$VIF = 1/(1 - R^2)$$

If the VIF is 1, a variable is not multicollinear at all. A VIF above 1 indicates there is some form of multicollinearity, with higher VIFs suggesting multicollinearity is a greater problem.

OLS Assumptions in the Sports Participation Regressions

Firstly, I check that the predicted values of the dependent variable fall within the [0, 1] interval. The predicted sports participation rates in all regressions in Table 4 range from approximately 29% to 65%, so the bounded properties of the dependent variable do not appear to be an issue. Secondly, as presented in Table B1, the RESET F-statistic is insignificant in all estimations. Hence, I find no evidence that the models are misspecified. Thirdly, the insignificant Shapiro-Wilk tests show that in every regression the error terms are approximately normally distributed. Fourthly, the Breusch-Pagan test results are also insignificant, suggesting that the errors are homoscedastic. Finally, to examine the degree of multicollinearity I compute

Table B1: Tests for the internal validity of the main OLS models

Table	Model	RESET F-stat. (df1, df2)	Shapiro-Wilk W-stat.	Breusch-Pagan χ^2-stat. (df)
5	1	1.335 (2, 276)	0.996	26.685 (19)
5	2	1.642 (2, 275)	0.995	28.563* (20)
5	3	1.284 (2, 275)	0.996	26.661 (20)
5	4	1.617 (2, 274)	0.995	28.208 (21)
7	1	0.060 (2, 276)	0.997	48.585*** (19)
7	2	0.156 (2, 275)	0.997	54.115*** (20)
7	3	0.153 (2, 275)	0.997	51.503*** (20)
7	4	0.267 (2, 274)	0.997	56.461*** (21)
A7	1	1.116 (2, 276)	0.992	42.867*** (19)
A7	2	1.145 (2, 275)	0.992	44.708*** (20)
A7	3	1.209 (2, 299)	0.993	48.135*** (19)
A7	4	1.282 (2, 298)	0.993	50.650*** (20)

*Note: The table reports the test statistics of three tests that examine the validity of a certain OLS assumption. The RESET test checks for functional form misspecification, the Shapiro-Wilk test assesses whether errors are normally distributed, and the Breusch-Pagan test examines whether the errors are homoscedastic. A significant test statistic indicates that the assumption is violated. Significance is indicated as follows: *p<0.1; **p<0.05; ***p<0.01.*

variance inflation factors. Table B2 shows these are all below the threshold of 10 as proposed by Everitt and Skrondal (2010), so multicollinearity is unlikely to be a problem. Together, these results suggest that the estimates do not suffer from bias due to functional form misspecification or multicollinearity and that the assumptions for valid hypothesis testing hold.

OLS Assumptions in the Exercise Participation Regressions

First, predictions of the share of the population with an active lifestyle range from approximately 37% to 60% in each regression and thus all fall in the realistically possible [0, 1] interval. Second, the insignificant RESET F-statistics in Table B1 indicate that no functional form misspecification is identified. Third, the error terms appear approximately normally distributed as the Shapiro-Wilk test does not yield significant test statistics. Fourth, the χ^2-statistics of the Breush-Pagan are significant at the 99% confidence level. This implies that the error terms are heteroscedastic, which I account for by reporting heteroscedasticity robust standard errors. So, given that we use heteroscedasticity robust standard errors, valid hypothesis testing is possible. Finally, the variance inflation factors reported in Table B2 are well below the threshold of 10. It is thus unlikely that functional form misspecification or multicollinearity introduce bias in the regression coefficients.

Table B2: Variance inflation factors of the independent variables

Variable	VIF	
Sports expenditures (log)	1.268	
Sports policy and activation expenditures (log)		1.169
Sports facilities (log)		1.221
Public parks and Recreation expenditures (log)	1.408	1.418
Male^s	2.157	2.160
Age 15-24^s (log)	2.213	2.222
Higher education^s (log)	5.199	5.206
Average household size (log)	6.504	6.536
Median disposable income	7.855	7.956
Long-standing illness^s	2.765	2.754
Migration background^s (log)	3.489	3.542
Drenthe	1.425	1.424
Flevoland	1.356	1.363
Friesland	1.515	1.523
Groningen	1.359	1.393
Limburg	1.980	1.985
North Brabant	2.558	2.574
North Holland	2.635	2.635
Overijssel	1.635	1.622
South Holland	2.577	2.600
Utrecht	1.098	1.100
Zeeland	1.416	1.446