

# Build It and Men Will Come: A Spatial Analysis of Sports Infrastructure and the Gender Play Gap in England

Michele Lovato Menin

michele.lovato-1@studenti.unitn.it

<https://github.com/MicheleLovatoMenin/Build-It-and-Man-Will-Come.git>

**Abstract**—This study challenges the prevailing supply-side narrative in sports policy by investigating whether the built environment acts as a catalyst for physical activity or merely reflects existing inequalities. Focusing on the entire territory of England ( $N = 6,856$  MSOAs), we employ a Spatial Durbin Error Model to disentangle the impact of sports infrastructure from socio-economic sorting and unobserved spatial heterogeneity. The analysis reveals three critical dynamics. First, we identify a “Local Fallacy”: once structural controls such as wealth and age are accounted for, the mere density of facilities shows no significant impact on participation rates. Second, quality matters more than quantity: while facility count is ineffective, a higher *Diversity Index* of sports provision is significantly associated with increased activity and reduced sedentary behavior, although the result is highly marginal. Third, and most crucially, the current infrastructure stock acts as an “Inequality Multiplier” regarding the *Gender Play Gap*. We find that both a higher density of facilities and a greater diversity of supply are associated with a widening of the gender divide, affecting both active participation and inactivity rates. Consequently, indiscriminate investment in physical provision disproportionately benefits men. These findings suggest that “spatial justice” cannot be achieved through building more facilities but requires intersectional policies accompanied by a profound cultural change to address the male-coded nature of sports spaces.

## I. INTRODUCTION

### A. From Individual Participation to Spatial Justice

Sport participation and physical activity are widely recognized as fundamental pillars for public health, social cohesion, and economic development in contemporary societies [1]. However, the distribution of these benefits is rarely uniform, neither socially nor geographically [2]. Traditionally, the academic literature has investigated these disparities by focusing primarily on individual determinants, such as income, age, and education, often treating the place of residence merely as a passive container or a fixed control variable.

This study aims to bridge the gap between individual behavior and the built environment, with a specific focus on the *Gender Play Gap*. As highlighted in sociological literature, the gender gap in sport is not solely defined by the quantity of practice but is deeply rooted in the types of activities available, the accessibility of spaces or the difference in pay for pro athletes, often conceptualized as complex inequality [3], [4]. Furthermore, sport has historically been constructed as a “masculine space”, creating structural barriers that go beyond simple personal motivation [5].

Furthermore, by adopting an intersectional perspective [6], it becomes evident that barriers to entry are not identical for all women; rather, they are compounded by ethnicity and socio-economic status.

### B. The Spatial Turn in Sports Economics

In recent years, sports economics and health geography have embraced a “spatial turn”, acknowledging that individual decisions to practice sport are not completely independent but are embedded in a dynamic spatial context. Tobler’s First Law of Geography posits that “everything is related to everything else, but near things are more related than distant things” [7]. This principle is particularly relevant for sports infrastructure: individuals cross administrative borders to access facilities implying that the effective service area of sports infrastructure extends beyond city borders, creating local spillovers of accessibility. Consequently, ignoring spatial interaction risks underestimating the true impact of infrastructure and providing biased policy recommendations [8].

## II. RELATED WORK

The literature on sports participation provides the context for understanding the interplay between infrastructure, society, and space. This section analyzes three key aspects: the “deprivation paradox” regarding facility access, the role of supply diversity in bridging the gender gap, and the methodological advantages of accounting for spatial dependence in health geography.

### A. Infrastructure vs. Socio-Economic Factors

A central debate in the literature concerns the extent to which the physical availability of facilities drives participation compared to socio-economic factors. Early studies demonstrated the importance of objective proximity [9], showing that swimming pools and fields are critical determinants for participation. This supports the supply-side argument “build it and they will come”.

However, more recent evidence suggests a more complex relationship. Hoekman et al. [10] found that physical distance explains only a small fraction of the urban-rural gap in the Netherlands, suggesting that socio-cultural factors and *habitus* often prevail over mere proximity. This phenomenon

is often described through the lens of “spatial sorting” or the “deprivation paradox”: wealthier individuals may choose to live in areas with better services or possess the financial means and mobility (e.g., private cars) to overcome distance, thereby making the local density of gyms less significant than the neighborhood’s wealth profile. Conversely, deprived areas might have public facilities nearby, but residents may face “time poverty” or economic barriers that prevent usage [11].

### B. The Gendered Infrastructure

When analyzing the Gender Gap, the definition of infrastructure requires nuance. It is not sufficient to merely count the number of facilities if the supply is dominated by male-coded spaces, such as football pitches. Noel et al. [12] introduced the concept of the *Sport Diversity Index*, arguing that a lack of variety in the sports offer creates a structural barrier for girls and women. Furthermore, qualitative studies on “spatial justice” highlight that the quality and safety of the environment are just as crucial as the physical existence of a facility [13]. Therefore, examining the heterogeneity of sports facilities, rather than just their density, is essential to understanding whether the built environment fosters equity or perpetuates exclusion.

### C. Methodological Approach and Research Contribution

To empirically investigate these dynamics in the English context, this study adopts a spatial approach. A significant portion of the existing literature on sports participation in the UK has effectively utilized Multi-Level Modelling approaches, typically analyzing individuals nested within Local Authorities (LAs) [2]. These models are excellent for disentangling individual and contextual effects at a macro scale.

However, recognizing that daily activities often transcend the boundaries of large administrative units, this research focuses on a more granular scale: the Middle Layer Super Output Areas (MSOA). By employing a Spatial Durbin Error Model (SDEM), we aim to complement existing hierarchical analyses by specifically capturing the local spillovers of contextual characteristics (such as neighboring infrastructure) while controlling for shared unobserved spatial heterogeneity. This specification allows us to test whether the availability of infrastructure in adjacent areas influences local participation rates, offering a different perspective on the localized ecology of sport.

## III. RESEARCH QUESTIONS

The analysis focuses on the entire territory of England, examining the relationship between sports infrastructure supply and adult participation. The population of interest comprises individuals aged 16 and over, consistent with the definitions adopted by Sport England in the Active Lives Survey [1].

To overcome the limitations of broad regional analyses and capture localized spatial dynamics, this study aggregates data at the level of Middle Layer Super Output Areas (MSOAs). Crucially, the analysis does not treat these areas as isolated units; it explicitly considers local spillovers, evaluating how

the density and diversity of facilities, derived from the Active Places Power database [14], in the neighboring areas influence local participation rates alongside internal provision.

Following this framework, the research adopts a “funnel approach”, moving from a general assessment of the infrastructure-participation relationship to a more granular analysis of supply diversity and gender inequalities. The study is guided by four sequential questions:

- 1) **The Baseline Spatial Correlation:** *To what extent is the spatial density of sports facilities associated with active participation rates?*

Based on supply-side theories [9], we initially hypothesize a positive spatial autocorrelation: areas with a higher density of facilities (and accessible neighboring supply) should exhibit higher participation rates due to reduced travel costs and increased visibility of sporting opportunities.

- 2) **The Socio-Economic Confounders:** *Does the positive association between infrastructure and participation persist when controlling for structural socio-economic variables (NS-SEC occupation, age, and ethnicity)?*

Drawing on the “deprivation paradox” and previous findings on the dominance of *habitus* over physical distance [10], we hypothesize that the direct effect of infrastructure will significantly diminish once socio-economic status is included. We expect wealth and social composition to be the primary drivers of participation, revealing that facility location may be less determinant than the neighborhood’s wealth profile.

- 3) **Beyond Quantity - The Role of Supply Diversity:** *Does the heterogeneity of the sports offer predict participation better than the mere count of facilities?*

Quantity does not equal quality. We hypothesize that a higher *Diversity Index*, indicating a wider range of options (e.g., pools, gym, courts), will have a stronger positive impact on general participation than the raw count of facilities, as it caters to broader preferences and reduces the dominance of mono-functional spaces like football pitches [12].

- 4) **The Gendered Dimension:** *How do facility density and supply diversity differentially impact the participation of men and women? Is the Gender Play Gap spatially correlated with the composition of the local infrastructure?*

Finally, the central issue of inequality is addressed. The hypothesis is that the *Gender Play Gap* is accentuated in areas with fewer facilities and, above all, with low diversity of supply. While male participation can be sustained even with a lack of diversity in facilities (due to the fact that most of the male population plays football), female participation should be more elastic with respect to the diversity of facilities and the quality of “spatial justice” [13], [4].

To answer these questions, we employ a Spatial Durbin Error Model (SDEM), which allows us to estimate the local spillovers of the surrounding infrastructure while accounting for spatially clustered unobserved heterogeneity in the error term.

#### IV. DESCRIPTION OF THE DATA

To investigate the spatial determinants of sports participation, this study integrates three distinct datasets: supply-side data on sports facilities, demand-side data on physical activity, and census data for socio-demographic controls. The integration of these sources allows for a comprehensive analysis of the relationship between the built environment and health behaviors.

##### A. Study Area and Unit of Analysis

The empirical analysis covers the entire territory of England. To capture localized spatial dynamics effectively, the chosen unit of analysis is the Middle Layer Super Output Area (MSOA). MSOAs are a standard statistical geography produced by the Office for National Statistics (ONS) designed to improve the reporting of small area statistics. With a minimum population of 5,000 and an average of approximately 7,200 residents, MSOAs offer a compromise between micro-level granularity, necessary to detect neighborhood effects, and statistical robustness. The final dataset comprises  $N = 6,856$  spatial units.

##### B. Dependent Variables: Sport Participation

Data on physical activity participation are derived from the *Active Lives Survey* [1], the official instrument used by Sport England to monitor the nation's activity levels. The survey collects data on the type, duration, and intensity of physical activities undertaken by adults (aged 16+). For this study, we focus on three key indicators:

**Active Adults:** The proportion of the adult population performing at least 150 minutes of moderate-intensity physical activity per week, in accordance with the Chief Medical Officer's guidelines.

**Inactive Adults:** The proportion of adults doing less than 30 minutes of physical activity per week. This variable is used to test the structural rigidity of sedentary behavior.

**Gender Play Gap:** A derived variable calculated as the percentage point difference between male and female active participation rates ( $Active_{Male} - Active_{Female}$ ). A positive value indicates higher male participation. We also calculated the inactive gender play gap ( $Inactive_{Male} - Inactive_{Female}$ ), where a negative value indicates greater inactivity among the female population.

It is important to note that these metrics capture overall engagement in physical activity rather than exclusive membership in formal sports clubs. Nevertheless, this broader scope aligns with our research objective, as we aim to assess whether a higher density or supply of sports activities correlates with an increase in the general population's propensity to exercise, regardless of where it is practiced.

##### C. Independent Variables: The Sports Infrastructure

The supply-side data is sourced from the *Active Places Power* database [14], which covers 42,924 distinct sports sites. A key distinction is made between *sites* (the physical location, such as a Leisure Centre) and *facilities* (the specific amenities

nested within a site, e.g., a swimming pool and a gym located at the same address).

**Facilities:** The total count of specific amenities (e.g., individual courts, pitches, pools) located within each MSOA. This is our primary measure of supply quantity.

**Diversity Index:** To assess the quality and variety of the offer, we computed a diversity score based on the number of unique facility types available in a specific MSOA (e.g., swimming pools, tennis courts and a gym, *Diversity Index* = 3). This index tests the hypothesis that a heterogeneous offer stimulates broader participation.

##### D. Control Variables: Socio-Demographic Context

To isolate the effect of infrastructure from compositional factors, we include a set of structural controls from the 2021 Census.

**Socio-Economic Status (NS-SEC):** We incorporate the full National Statistics Socio-economic Classification to capture the social gradient. The variable is categorized into: Higher managerial and professional (NS-SEC 1-2), Intermediate occupations (NS-SEC 3), Small employers (NS-SEC 4), Lower supervisory (NS-SEC 5), Semi-routine and routine (NS-SEC 6-7), Long-term unemployed (NS-SEC 8), and Students (NS-SEC 9). Intermediate occupations (NS-SEC 3) serve as the reference category.

**Ethnicity:** The proportion of residents identifying as White, Black, Asian, Mixed, or Other ethnic groups. The White population serves as the reference category.

**Age Structure:** The population is segmented into key age bands (16-34, 35-54, 55-74, 75+). The 16-34 age group serves as the reference category to control for life-cycle effects.

##### E. The Spatial Weights Matrix

A critical component of the Spatial Durbin Error Model is the definition of the spatial weights matrix ( $W$ ), which formalizes the connectivity between spatial units. Given the uneven size of MSOAs (smaller in urban centers, larger in rural areas), a simple contiguity matrix would be inappropriate. Instead, we adopted a Hybrid Distance-Based Matrix to model realistic accessibility:

- 1) **K-Nearest Neighbors (KNN):** The primary model uses  $k=10$  nearest neighbors to ensure connectivity. For robustness purposes, we also tested matrices with  $k=20$  and  $k=50$ , confirming that results are not sensitive to the choice of  $k$ .
- 2) **Distance Cutoff:** Interactions are restricted to a maximum radius of 15km. This threshold was selected to approximate a realistic maximum travel range for regular sports activity. Consequently, MSOAs in remote rural areas with no neighbors within this radius are treated as isolates.
- 3) **Inverse Distance Weighting (IDW):** The weights are standardized using an inverse distance function ( $1/d_{ij}$ ), assigning greater influence to geographically closer neighbors.

This specification allows the model to capture both the immediate neighborhood effects and the broader regional spillover of the sports ecosystem.

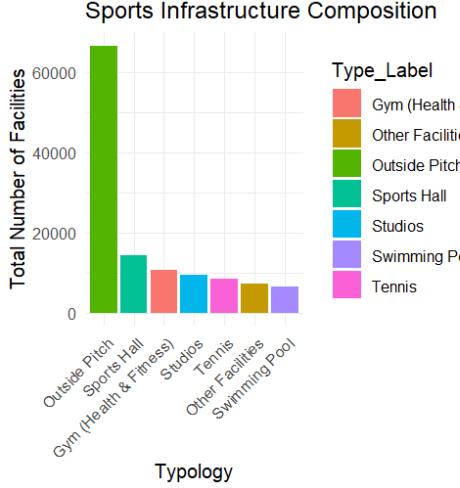


Fig. 1: **Composition of Sports Infrastructure.** The bar chart highlights the disproportionate prevalence of pitches compared to other facility types.

## V. DATA ANALYSIS AND RESULTS

### A. Descriptive Analysis: The Landscape of Sport in England

Before addressing the spatial modeling, it is essential to characterize the structural inequalities present in the raw data. The descriptive analysis reveals a sporting landscape characterized by a rigid dichotomy between the supply of infrastructure and the demand for participation.

1) *The Anatomy of Supply:* The analysis of the *Active Places Power* database reveals an imbalance in the composition of English sports infrastructure. As shown in Figure 1, the supply is overwhelmingly dominated by outdoor pitches (grass and synthetic), which account for the vast majority of facilities. It is worth noting that while this category technically encompasses pitches for other team sports such as rugby or cricket, football pitches constitute the overwhelming majority of this stock. Consequently, the prevalence of outdoor pitches serves as a reliable proxy for football-oriented provision. This predominance suggests a historical legacy of planning focused on team sports, particularly football, often at the expense of a more diverse portfolio of activities (e.g., swimming pools, studios, or tennis courts) that might appeal to a broader demographic.

To quantify this dominance, we identified “*Only-Football MSOAs*”, neighborhoods where the local sporting offer consists exclusively of football pitches. In addition to 113 neighborhoods that don’t even have a single facility, the data indicates that 406 MSOAs fall into this category. This lack of supply or mono-functional supply creates a structural barrier: for residents in these areas who do not play football, the effective local supply of sports infrastructure is zero, regardless of the physical number of pitches available.

2) *The Socio-Economic Gradient:* On the demand side, participation levels exhibit a steep gradient associated with socio-economic status. By stratifying MSOAs into quartiles based on the proportion of high managerial and professional (1-2 NS-SEC) in the neighborhood, a clear linear relationship emerges (Figure 2).

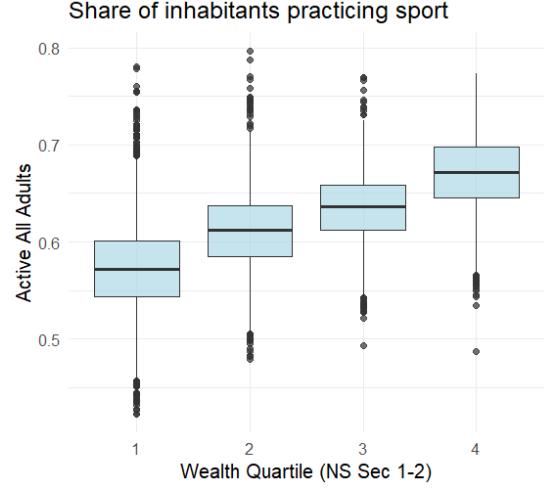


Fig. 2: **Participation by Wealth Quartile.** Active participation increases linearly with the socio-economic status of the neighborhood (NS-SEC 1-2).

Residents in the top quartile (High Wealth) report an active participation rate of 67.0%, significantly higher than the 57.3% observed in the bottom quartile (Low Wealth). This disparity is mirrored on the supply side: wealthy neighborhoods enjoy both a higher density of facilities (21.2 vs 14.8 on average) and a higher *Diversity Index* (5.69 vs 4.77).

3) *Gender Play Gap:* Finally, the Gender Play Gap proves remarkably persistent. On average, male participation exceeds female participation by 5.45 percentage points. Crucially, this gap remains constant across the socio-economic spectrum, showing negligible variation between the wealthiest (5.47%) and poorest (5.49%) quartiles. This suggests that while economic capital and infrastructure density increase overall participation, they fail to bridge the gender divide, indicating that the current infrastructure stock, biased towards male-coded team sports, is structurally unable to foster equity.

### B. Exploratory Spatial Data Analysis

To validate the necessity of a spatial approach, we first tested the hypothesis of spatial randomness using the Global Moran’s I statistic on the dependent variable (*Active Adults*). The test yields a highly significant index of 0.817 ( $p < 0.001$ , based on 999 Monte Carlo permutations). As illustrated in the Moran Scatterplot (Figure 3), there is a remarkably strong positive linear association between local participation rates and the average rates of neighboring areas. This confirms that physical activity is not spatially independent; rather, it is characterized by intense clustering, where active communities tend to be grouped together.

To inspect the local geography of this dependence, we computed the Local Indicators of Spatial Association (LISA). Figure 4 compares two specifications: the standard significance map ( $p < 0.05$ ) to identify broad regional trends, and a robust map corrected with the Bonferroni method to isolate the most statistically significant “hard core” clusters.

The Regional Trends map reveals a pervasive spatial structure, with 3,124 MSOAs (~45% of the sample) identified as significant clusters.

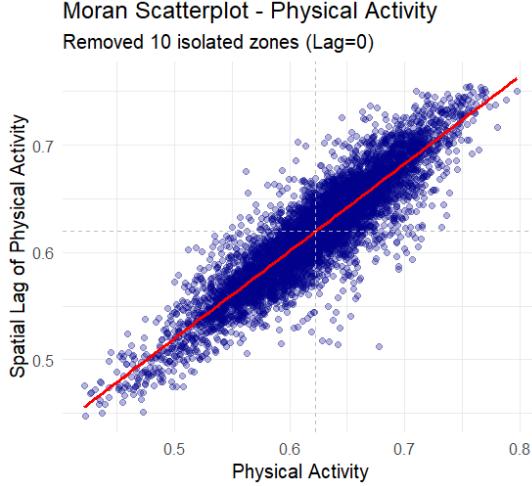


Fig. 3: **Moran Scatterplot for Active Adults.** The slope of the regression line (0.817) indicates strong positive spatial autocorrelation.

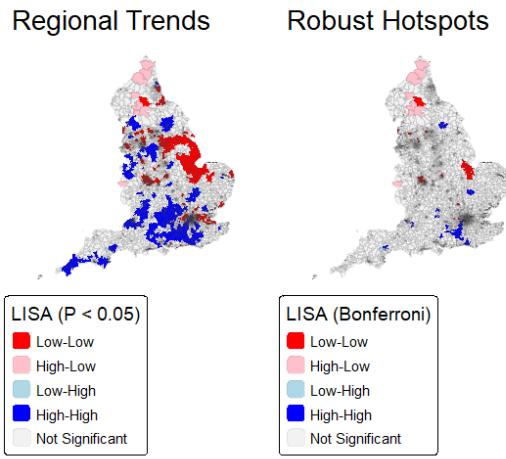


Fig. 4: **LISA Cluster Maps.** The panel compares standard significance (left) with Bonferroni-corrected significance (right). While the regional trends cover broad areas (3124 MSOAs), the robust correction isolates the structural core of the North-South divide (637 MSOAs).

**High-High Clusters:** A distinct “corridor of activity” dominates London, the South-East, and the South-West coast. In these regions, high participation is the spatial norm.

**Low-Low Clusters:** Conversely, inactivity is heavily concentrated in the major urban conurbations of the Midlands and the North. Specifically, the cities of Birmingham, Liverpool, and Manchester emerge as the epicenters of extensive Low-Low clusters. A large cluster is also visible in the central-eastern part of the state. Interestingly, a pocket of low activity is also visible immediately north of the Greater London boundary, suggesting that proximity to the capital does not guarantee participation.

Applying the Bonferroni correction imposes a severe penalty for multiple testing, reducing the number of significant clusters from 3124 to just 637. This drastic reduction acts as a spatial filter: the extensive red cluster in the Centre-East shrinks significantly, implying that while inactivity is present, it is less spatially cohesive than in the North-West.

However, the core polarization survives: the deep inactivity in Birmingham and Manchester and the intense activity belts in the Home Counties remain significant even under the strictest statistical conditions.

This persistent geographic fracture demonstrates that local participation rates are largely determined by the surrounding spatial context, validating the use of the Spatial Durbin Error Model.

### C. The Spatial Durbin Model: Unveiling the “Local Fallacy”

**1) The Baseline Infrastructure Model:** To answer the first research question, we initially estimated a *Naive Spatial Durbin Model* including only infrastructure variable  $N.$  of Facilities. In this simplified scenario, the supply-side hypothesis appears strongly supported. Both the local density of facilities ( $p < 0.001$ ) and the neighboring facilities ( $WX$ ) ( $p < 0.001$ ) show a positive and significant association with active participation. However, as subsequent models reveal, this relationship is largely spurious, driven by the omitted variable bias of socio-economic status.

TABLE I: Naive Spatial Durbin Error Model (Infrastructure Only)

Parameter	Estimate	Significance
<b>Local Infrastructure</b> N. of Facilities	0.00033	***
<b>Spillover Infrastructure</b> Lag.Facilities ( $W \times X$ )	0.00105	***
<b>Spatial Error</b> Lambda ( $\lambda$ )	0.903	***
<b>Model Fit (AIC)</b>	-30,546	

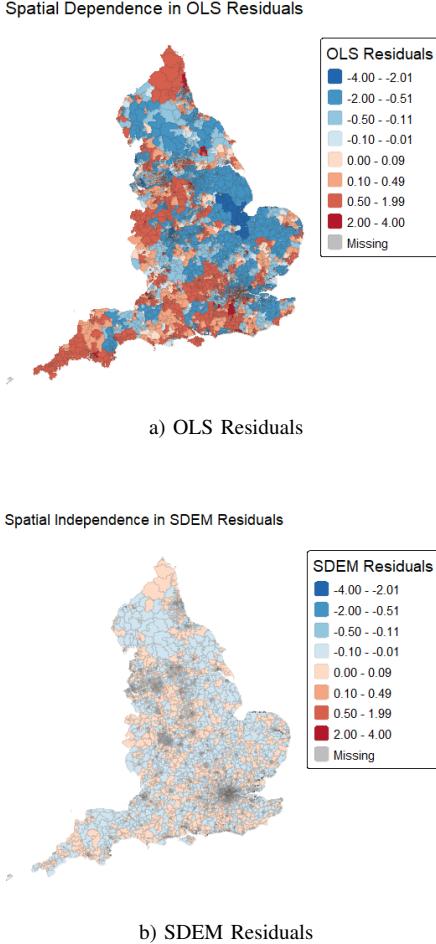
Significance codes: \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ .  $N = 6,856$ .

**2) Model Selection and Diagnostics:** We then introduced the full set of socio-economic and demographic controls. To verify the robustness of our spatial specification, we compared the Spatial Durbin Error Model (SDEM) against the non-spatial OLS and other spatial specifications (SAR, SEM, and SDM).

Table II reports the diagnostics. The SDEM exhibits the best fit among all specifications (AIC = -40,018), outperforming both the nested SEM and the SDM. This confirms that the phenomenon of physical activity is best described by local spillovers in covariates combined with spatially structured error dependence, rather than by global behavioral contagion.

Inspection of the residuals visualized in Figure 5 further validates this choice. While the OLS residuals show deep clusters of unexplained variation (Global Moran's I = 0.74), indicating omitted spatial processes, the SDEM residuals appear randomly distributed (Moran's I = 0.10). This demonstrates that the spatial error term ( $\lambda = 0.93$ ) has successfully captured the unobserved heterogeneity. The high magnitude of this coefficient suggests that physical activity is deeply embedded in localized *geographies of habitus*, shared cultural norms and environmental micro-factors that persist even after controlling for structural variables.

Consequently, the superior fit combined with the theoretical need to isolate local infrastructure spillovers from nuisance



**Fig. 5: Comparison of Residuals.** The maps share the same scale. The SDEM map appears visibly flatter, demonstrating that the spatial error model has effectively reduced the error magnitude and clustering compared to OLS.

spatial dependence leads us to retain the SDEM as the most appropriate specification for policy analysis.

*3) Full Model Impacts:* The transition to the Full Model radically alters the conclusions regarding infrastructure. As shown in Table III, the addition of socio-economic controls renders both the Direct ( $p = 0.160$ ) and the Indirect ( $p = 0.305$ ) Impact of facilities statistically non-significant. This finding extends the “Local Fallacy”: once socio-economic sorting is accounted for, neither the local facilities nor the neighboring sports ecosystem predicts higher participation rates. The model, however, detects massive unexplained clustering, captured by the spatial error parameter  $\lambda$  estimated at 0.93 ( $p < 0.001$ ), suggesting that omitted environmental or cultural factors share a strong spatial structure. The fact that  $\lambda$  is so high while infrastructure impacts vanish confirms that participation is clustered by territorial affinity rather than by the mere proximity of facilities. Consequently, the “regional ecosystem” effect is driven not just by the aggregate supply of facilities in the area, but probably by a self-reinforcing culture of movement.

Crucially, the table highlights the overwhelming dominance of structural factors over infrastructural ones. The “Class Gradient” is evident and reinforces the spatial clustering of

TABLE II: Model Selection Diagnostics

Statistic	OLS	SAR	SEM	SDM	SDEM
Log Lik.	14,480	16,586	19,962	19,518	20,040
AIC	-28,934	-33,138	-39,891	-38,974	-40,018
Param.	-	0.46 $\rho$	0.94 $\lambda$	0.90 $\rho$	0.93 $\lambda$

TABLE III: Impact on Physical Activity (SDEM)

Variable	Direct	Indirect	Total
<b>Infrastructure</b>			
N. of Facilities ( <i>p-value</i> )	0.00002 (0.160)	0.00006 (0.305)	0.00008 (0.241)
<b>Social Class (NS-SEC)</b>			
High Managerial (1-2)	0.071***	0.062**	0.133***
Small Employers (4)	0.039*	0.106**	0.146***
Lower Supervisory (5)	-0.073	-0.169	-0.242*
Routine (6-7)	-0.138***	-0.022	-0.160***
Unemployed (8)	-0.208***	0.106*	-0.102*
Students (9)	0.103***	-0.002	0.101***
<b>Age Structure</b>			
Age 35-54	0.011	-0.036	-0.025
Age 55-74	-0.040***	-0.116***	-0.156***
Age 75+	-0.266***	0.088*	-0.178***
<b>Ethnicity</b>			
Asian	-0.142***	-0.039***	-0.181***
Black	-0.120***	0.031	-0.089***
Mixed	-0.096***	-0.057	-0.153
Other	-0.049**	-0.035	-0.084

$$\lambda = 0.93.$$

Significance: \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ .

Reference categories: NS-SEC 3, Age 16-34, White. N = 6,856.

privilege. Moving residents from intermediate occupations to high managerial roles (NS-SEC 1-2) generates a significant increase in participation via the Direct Impact (+0.071), confirming class as a primary local driver. Furthermore, the indirect effects for high socio-economic status are positive and significant (+0.062). This indicates a synergistic effect: the benefits of high social status are cumulative, implying that living in a prosperous neighborhood with a high concentration of managerial professionals further stimulates individual physical activity, likely due to peer effects and better environmental amenities.

Finally, demographic variables act as powerful structural constraints, overshadowing the marginal benefits of infrastructure provision. The analysis reveals that ethnic barriers are robust negative predictors, with the proportion of Asian residents exhibiting significant negative coefficients for both direct (-0.142) and indirect (-0.039) impacts. Furthermore, aging acts as an insurmountable wall: the proportion of residents aged over 75 exhibits a massive negative direct impact (-0.266). This confirms that demographic aging acts as a generalized depressor of physical activity rates that strictly limits the effectiveness of supply-side interventions.

*4) Sensitivity Analysis:* To verify the stability of our results, we performed robustness checks utilizing alternative spatial specifications (detailed results in Appendix A).

First, we tested the sensitivity to the neighborhood definition. Re-estimating the model with  $K = 20$  and  $K = 50$  nearest neighbors (Tables IX and X) reveals that the impact of local infrastructure is not sensitive to the choice of the

spatial weights matrix. While the p-value for the Direct Impact improves slightly with  $K = 20$  ( $p = 0.101$ ) compared to the baseline  $K = 10$  ( $p = 0.160$ ), it reverts to similar levels with  $K = 50$  ( $p = 0.166$ ). The coefficients remain consistently negligible, confirming that the lack of a strong supply-side effect is a structural feature of the data rather than an artifact of the matrix definition.

Second, we substituted the count of facilities with the count of sites (Table XI). In this specification, the Direct Impact becomes statistically significant ( $p = 0.043$ ), although the Indirect and Total effects remain non-significant. This result is consistent with an aggregation effect: since there are 126,554 facilities but only 42,924 sites in England, one site aggregates on average 3 facilities. Consequently, a unit increase in sites represents a larger structural shock to the local supply than a unit increase in facilities.

However, even in this specification, the economic magnitude remains marginal (+0.0001). Therefore, these checks confirm the validity of our main specification (facilities,  $K = 10$ ) and reinforce the “Local Fallacy”: compared to the dominant drivers of wealth and demographics, the direct or indirect leverage of infrastructure provision remains practically irrelevant.

#### D. Quality vs. Quantity

Having established the limitations of mere infrastructure density, we now turn to the third research question: *does the variety of sports provision matter more than its quantity?* To test the hypothesis that a diverse portfolio of facilities fosters higher engagement, we estimated a Spatial Durbin Error Model replacing the simple count of facilities with the *Diversity Index*.

The results, summarized in Table IV, reveal a significant shift. Unlike the simple facility count, which confirmed the “Local Fallacy” by showing no significant effects in the full model, the *Diversity Index* exhibits a statistically significant positive impact. Specifically, the Direct effect is significant ( $p = 0.013$ ), indicating that a wider range of local activities promotes participation. The Indirect effect is marginally significant ( $p = 0.052$ ), contributing to a significant Total impact ( $p = 0.030$ ).

It is important to contextualize the magnitude of these coefficients. The *Diversity Index* operates on a much compressed scale (ranging from 0 to 14 unique types) compared to the raw facility count (which reaches 144 in the MSOA: City of London). Consequently, a one-unit increase in *Diversity Index* represents a substantial structural upgrade in the local offer (e.g., adding a totally new sport type), whereas a one-unit increase in facilities is marginal. Even accounting for this scale difference, the emergence of statistical significance validates the hypothesis: residents respond to a heterogeneous offer that caters to broader preferences in a way they do not to a mere accumulation of identical pitches.

#### E. The Inactivity Wall: Structural Determinants

Furthermore, we test whether supply-side interventions can effectively address the problem of physical inactivity, rather

TABLE IV: Comparison of Infrastructure Impacts: Quantity vs. Diversity (SDEM)

Specification	Direct	Indirect	Total
<b>1. Facilities</b> ( <i>p-value</i> )	0.00002 (0.161)	0.00006 (0.305)	0.00008 (0.241)
<b>2. Diversity</b> ( <i>p-value</i> )	0.00018* (0.013)	0.00061 (0.052)	0.00079* (0.030)

Significance codes: \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ .

than just boosting existing participation. We analyzed the determinants of *Physical Inactivity* (residents doing less than 30 minutes of activity per week). Table V presents the full impact estimates using the standard facility count. Comparing these results with the Active model reveals both inverse patterns and unique rigidities.

**1. The Limits of Supply-Side Interventions.** The analysis reveals the limitations of a quantity-based approach. The direct impact of local facilities on inactivity is negative but only marginally significant ( $p = 0.088$ ), while the indirect spillover effect is statistically non-significant ( $p = 0.223$ ). This suggests that simply increasing the density of sports venues is insufficient to convert sedentary populations. However, it is worth noting that the *quality* of the offer plays a more distinct role. Table VI, the *Diversity Index* exhibits a statistically significant direct impact ( $p = 0.022$ ) on reducing inactivity. This implies that while “more of the same” fails, a diverse portfolio catering to different needs can slightly reduce sedentary behavior, though the magnitude of this effect remains small compared to structural factors.

**2. Invers and Structural Drivers.** The socio-economic variables largely act as a mirror to the participation model, but with intensified magnitude, confirming the existence of an “Inactivity Wall”:

*Social Status (The Class Gradient):* Low social status is the strongest driver of inactivity. Long-term unemployment (NS-SEC 8) shows a massive positive direct impact (+0.261,  $p < 0.001$ ), followed by Routine occupations (+0.143,  $p < 0.001$ ). This confirms that economic deprivation and the associated lack of free time or resources constitute the primary barrier.

*Demographics:* The effect of aging is drastic. The proportion of residents aged 75+ exhibits a very large direct impact on inactivity (+0.237,  $p < 0.001$ ). While in the Active model age was a constraint, here it appears as an insurmountable wall, suggesting that inactivity in these cohorts is driven by health and mobility limitations that physical infrastructure alone cannot address.

#### F. The Gender Play Gap

Finally, we address the fourth research question regarding the gendered dimension of sports provision. By modeling the *Gender Play Gap* (calculated as the difference between male and female participation rates:  $Gap = Active_{Male} - Active_{Female}$ ) as the dependent variable, we test whether the current infrastructure stock acts as an equalizer or an exacerbator of inequality. A positive coefficient indicates factors that widen the gap favoring men, since no neighborhood shows a

TABLE V: Impact on Physical Inactivity (SDEM - Facilities)

Variable	Direct	Indirect	Total
<b>Infrastructure</b>			
N. of Facilities ( <i>p-value</i> )	-0.00002 (0.088)	-0.00006 (0.223)	-0.00009 (0.160)
<b>Social Class (NS-SEC)</b>			
High Managerial (1-2)	-0.046***	-0.031	-0.077***
Small Employers (4)	-0.022	-0.107***	-0.129***
Lower Supervisory (5)	0.103**	0.133	0.236*
Routine (6-7)	0.143***	0.025	0.168***
Unemployed (8)	0.261***	-0.031	0.230***
Students (9)	-0.061***	0.011	-0.051**
<b>Age</b>			
Age 35-54	-0.020**	0.002	-0.018
Age 55-74	0.034***	0.106***	0.139***
Age 75+	0.237***	-0.115***	0.122***
<b>Ethnicity</b>			
Asian	0.127***	0.029**	0.156***
Black	0.096***	-0.007	0.089***
Mixed	0.096***	0.008	0.103
Other	0.029*	0.015	0.044

$\lambda = 0.93$ .

Significance: \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ .

Reference categories: NS-SEC 3, Age 16-34, White. N = 6,856.

Note: Positive coefficients indicate an increase in inactivity levels.

TABLE VI: Impact on Physical Inactivity (SDEM - Diversity)

Variable	Direct	Indirect	Total
<b>Infrastructure</b>			
Diversity Index ( <i>p-value</i> )	-0.00015* (0.022)	-0.00051 (0.057)	-0.00066* (0.036)
<b>Social Class (NS-SEC)</b>			
High Managerial (1-2)	-0.046***	-0.029	-0.075***
Small Employers (4)	-0.021	-0.105***	-0.126***
Lower Supervisory (5)	0.103**	0.128	0.232*
Routine (6-7)	0.144***	0.028	0.172***
Unemployed (8)	0.261***	-0.030	0.231***
Students (9)	-0.061***	0.012	-0.049**
<b>Age</b>			
Age 35-54	-0.020**	0.002	-0.017
Age 55-74	0.033***	0.104***	0.137***
Age 75+	0.238***	-0.110***	0.128***
<b>Ethnicity</b>			
Asian	0.127***	0.029**	0.156***
Black	0.096***	-0.006	0.090***
Mixed	0.096***	0.008	0.104
Other	0.030*	0.016	0.046

$\lambda = 0.93$ .

Significance: \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ .

Reference categories: NS-SEC 3, Age 16-34, White. N = 6,856.

Note: Positive coefficients indicate an increase in inactivity.

percentage of women who practice sport greater than that of men, while a negative coefficient indicates a narrowing effect.

1) *Quantity - The Inequality Multiplier:* First, we analyze the impact of simple facility density. The results, presented in Table VII, provide compelling evidence of a “Multiplier Effect”. Controlling for socio-economic factors, the Direct Impact of *N. of Facilities* on the Gender Gap is positive and highly significant (+0.00003,  $p < 0.001$ ). This implies that, *ceteris paribus*, an increase in the local supply of sports

TABLE VII: Impact on Gender Play Gap (Facilities Model)

Variable	Direct	Indirect	Total
<b>Infrastructure</b>			
N. of Facilities ( <i>p-value</i> )	0.00003*** (< 0.001)	0.00005* (0.011)	0.00008*** (< 0.001)
<b>Social Class (NS-SEC)</b>			
High Managerial (1-2)	-0.048***	-0.012*	-0.060***
Small Employers (4)	-0.103***	0.059***	-0.044***
Lower Supervisory (5)	-0.151***	0.114***	-0.037
Routine (6-7)	-0.039***	-0.023***	-0.062***
Unemployed (8)	0.018**	-0.067***	-0.049***
Students (9)	-0.065***	-0.009*	-0.073***
<b>Age</b>			
Age 35-54	0.031***	0.007	0.037***
Age 55-74	-0.036***	-0.010	-0.046***
Age 75+	0.153***	-0.007	0.145***
<b>Ethnicity</b>			
Asian	0.026***	0.000	0.026***
Black	0.011***	0.003	0.014***
Mixed	-0.021*	-0.052**	-0.073***
Other	0.007	0.034**	0.042***

$\lambda = 0.53$ .

Significance: \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ .

Reference categories: NS-SEC 3, Age 16-34, White. N = 6,856.

Note: Positive coefficients indicate a widening of the gap.

facilities leads to a widening of the participation gap. As detailed in the disaggregated analysis in Appendix C, this counter-intuitive finding stems from differential elasticity: while male participation responds positively to supply expansion ( $p = 0.038$ ), female participation remains statistically inelastic ( $p = 0.766$ ). Consequently, any expansion of the current infrastructure stock predominantly benefits men, structurally increasing the divide. Finally, a comparison of the spatial diagnostics reveals a distinct feature of inequality. Unlike the general participation models where the spatial error was 0.93, in the gender play gap model  $\lambda$  drops to 0.53. This suggests that the gender divide is less driven by unobserved neighborhood *habitus* or cultural contagion, and is instead more rigidly determined by the explicit structural and infrastructural variables included in our equation.

2) *Quality - The Diversity Paradox:* We then hypothesized that a more varied offer would be the key to narrowing the gap. The results (Table VIII) reject this hypothesis. The *Diversity Index* also exhibits a statistically significant positive coefficient for the Direct Impact (+0.00015,  $p < 0.001$ ). This suggests a “Diversity Paradox”: even in neighborhoods with a high mix of amenities, the gender gap remains wider than in resource-poor areas. Once again, the gender-specific models reveal the mechanism: men are highly responsive to variety ( $p < 0.001$ ), while women show no significant direct response ( $p = 0.256$ ). This points to a failure of the “spatial justice” approach: without accompanying policies addressing time-poverty, safety, and care responsibilities, physical infrastructure, regardless of its variety, remains a subsidy that disproportionately benefits men.

Beyond infrastructure, both models highlight the deep demographic roots of inequality. High social status (NS-SEC 1-2) acts as a buffer, significantly narrowing the gap (Direct: -0.048). Conversely, structural constraints widen it: the propor-

TABLE VIII: Impact on Gender Play Gap (Diversity Model)

Variable	Direct	Indirect	Total
<b>Infrastructure</b>			
Diversity Index ( <i>p</i> -value)	0.00015*** (<0.001)	0.00023* (0.022)	0.00038*** (0.001)
<b>Social Class (NS-SEC)</b>			
High Managerial (1-2)	-0.047***	-0.012*	-0.059***
Small Employers (4)	-0.101***	0.057***	-0.044***
Lower Supervisory (5)	-0.149***	0.116***	-0.034
Routine (6-7)	-0.039***	-0.023***	-0.062***
Unemployed (8)	0.019**	-0.069***	-0.050***
Students (9)	-0.064***	-0.009*	-0.072***
<b>Age</b>			
Age 35-54	0.031***	0.007	0.037***
Age 55-74	-0.036***	-0.008	-0.044***
Age 75+	0.152***	-0.009	0.143***
<b>Ethnicity</b>			
Asian	0.026***	0.001	0.027***
Black	0.011***	0.003	0.014***
Mixed	-0.022*	-0.054**	-0.076***
Other	0.007	0.035**	0.042***

$\lambda = 0.54$ .

Significance: \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ .

Reference categories: NS-SEC 3, Age 16-34, White. N = 6,856.

Note: Positive coefficients indicate a widening of the gap.

tion of residents aged 75+ (+0.153) and Asian communities (+0.026) are associated with a significantly larger gender divide.

Finally, as a robustness check, we extended the analysis to the *Inactivity Gap* ( $Inactive_{Male} - Inactive_{Female}$ ). As detailed in Tables XVII and XVIII in Appendix C, the results corroborate the findings on active participation. Both facility density (Direct:  $-0.00002$ ,  $p < 0.001$ ) and diversity (Direct:  $-0.00012$ ,  $p < 0.001$ ) exhibit negative coefficients, which in the context of the negative baseline gap, indicates a further widening of the inequality.

## VI. CONCLUSION

This study challenged the prevailing supply-side narrative in sports policy, utilizing a Spatial Durbin Error Model to investigate whether the built environment acts as a catalyst for participation or merely reflects existing inequalities. The empirical evidence from the English context suggests that the relationship between infrastructure and behavior is far from linear, revealing a complex spatial structure, and highlighting the structural rigidity of the gender gap.

First, our findings demystify the impact of mere density. While a higher number of facilities is positively associated with participation in naive models, this effect vanishes once socio-economic structural factors are controlled for, also because there is a positive correlation between the number of facilities and the wealth of MSOA. This confirms the “Local Fallacy”: in a mature sports market, adding “more of the same” yields diminishing marginal returns. Conversely, the quality of the offer, measured through the *Diversity Index*, tells a different story. Unlike simple facility density, which proves ineffective against sedentary behavior, the diversity of the sports portfolio retains a statistically significant impact on

both boosting active participation and reducing inactivity. This suggests that “spatial justice” is better served by a heterogeneous portfolio of amenities catering to diverse preferences rather than a monoculture of standardized pitches. However, while diversity manages to slightly erode the “Inactivity Wall”, its magnitude remains secondary compared to the overwhelming dominance of structural drivers: age, poverty, ethnicity, creating a barrier that an increase of the number of facilities cannot fully breach.

Crucially, regarding the gender dimension, the current infrastructure stock acts as an “Inequality Multiplier”. Our analysis reveals a stark differential elasticity: while male participation responds positively to supply expansion and diversification, female participation remains largely inelastic to physical provision. Consequently, indiscriminate supply expansion creates a paradox where public investment disproportionately benefits male users, structurally widening the Gender Play Gap. This dynamic confirms that without an intersectional approach [6] addressing the compounding barriers of ethnicity, class, and gender, the built environment remains a male-coded space. Furthermore, the spatial diagnostics offer a distinct insight: unlike the general participation models where a near-saturated error term ( $\lambda = 0.93$ ) pointed to a dominant geography of *habitus*, the lower clustering in these models ( $\lambda = 0.53$ ) confirms that the gender gap is less driven by unobserved neighborhood effects and more by the explicit structural barriers identified in our analysis.

Moreover, this spatial clustering invites a reflection on reverse causality. It is possible that the relationship is bidirectional: infrastructure is not merely a generator of activity but is often developed in response to pre-existing demand from active communities. While our current analysis does not statistically disentangle this directionality, the overlap between wealth and facility density suggests that provision may tend to align with established demand patterns rather than addressing structural needs. Nonetheless, our analyses show a correlation between physical activity and the number of facilities that is only marginal.

Finally, two key limitations of this study point towards future research avenues. First, these results are deeply rooted in the specific English context, which represents a saturated scenario characterized by a historically high stock of infrastructure and high baseline participation. We hypothesize that in a scarcity scenario, such as in Southern European regions where baseline provision is lower, the marginal utility of new facilities might be significantly higher, potentially making physical access a primary driver.

Second, this analysis relied on an aggregate measure of infrastructure. A lack of participation in a specific area might stem from a mismatch between the physical container (e.g., a football pitch) and the social structure required to use it (e.g., the presence of a women’s team). Future research should aim for a sport-specific analysis, intersecting the geolocation of facilities with data on active associations, to distinguish between theoretical physical access and concrete social opportunity.

## REFERENCES

- [1] S. England, "Active lives survey - measuring the nation's activity levels." 2024.
- [2] P. Downward and C. Muniz, "The opportunities and challenges facing participation in different types of women's football in england," *Sport Management Review*, pp. 1–25, 2025.
- [3] J. Hargreaves, *Sporting females: Critical issues in the history and sociology of women's sport*. Routledge, 2002.
- [4] P. Russo, "La disegualanza complessa. tutte le dimensioni del gender gap nello sport," *Eraclio. Journal of Sport and Social Sciences*, vol. 3, pp. 6–17, 2020.
- [5] S. Scraton, "Reconceptualizing race, gender and sport: The contribution of black feminism," in *Race, sport and british society*. Routledge, 2002, pp. 192–209.
- [6] K. W. Crenshaw, "Mapping the margins: Intersectionality, identity politics, and violence against women of color," in *The public nature of private violence*. Routledge, 2013, pp. 93–118.
- [7] W. R. Tobler, "A computer movie simulating urban growth in the detroit region," *Economic geography*, vol. 46, no. sup1, pp. 234–240, 1970.
- [8] X. Feng and B. Humphreys, "Assessing the economic impact of sports facilities on residential property values: A spatial hedonic approach," *Journal of Sports Economics*, vol. 19, no. 2, pp. 188–210, 2018.
- [9] P. Wicker, K. Hallmann, and C. Breuer, "Analyzing the impact of sport infrastructure on sport participation using geo-coded data: Evidence from multi-level models," *Sport management review*, vol. 16, no. 1, pp. 54–67, 2013.
- [10] R. Hoekman, K. Breedveld, and G. Kraaykamp, "Sport participation and the social and physical environment: explaining differences between urban and rural areas in the netherlands," *Leisure studies*, vol. 36, no. 3, pp. 357–370, 2017.
- [11] J. E. Ruseski, B. R. Humphreys, K. Hallmann, and C. Breuer, "Family structure, time constraints, and sport participation," *European review of aging and physical activity*, vol. 8, no. 2, pp. 57–66, 2011.
- [12] K. Noel-London, E. Shacham, K. Ortiz, and R. BeLue, "Neighborhood context as a barrier to sport participation for girls: Evidence from a sport diversity index," *Journal of athletic training*, vol. 59, no. 5, pp. 465–473, 2024.
- [13] N. Bevan, R. Jeanes, and H. Truskewycz, "Spatial justice in the development of a women's football team in melbourne, australia; an ethnographic study," *Gender, Place & Culture*, vol. 31, no. 3, pp. 312–334, 2024.
- [14] Sport England, "Active places power," 2024, data set. [Online]. Available: <https://www.activeplacespower.com>

**APPENDIX A**  
**ROBUSTNESS CHECKS: FULL IMPACT TABLES**

This appendix presents the complete impact estimates for the sensitivity analyses discussed in the main text. Tables IX and X display results using spatial weight matrices based on the 20 and 50 nearest neighbors, respectively. Table XI details the model using the count of sites as the alternative infrastructure variable.

TABLE IX: SDEM with  $K = 20$  Nearest Neighbors

Variable	Direct	Indirect	Total
<b>Infrastructure</b>			
N. of Facilities ( <i>p-value</i> )	0.00003 (0.101)	0.00014 (0.089)	0.00017 (0.071)
<b>Social Class</b>			
High Managerial (1-2)	0.093***	0.100***	0.193***
Small Employers (4)	0.047**	0.095*	0.141**
Lower Supervisory (5)	-0.041	-0.260*	-0.300*
Routine (6-7)	-0.124***	-0.008	-0.132**
Unemployed (8)	-0.187***	0.197***	0.010
Students (9)	0.123***	-0.032	0.091***
<b>Age</b>			
Age 35-54	0.012	-0.065	-0.052
Age 55-74	-0.037***	-0.117**	-0.154***
Age 75+	-0.262***	0.073	-0.189***
<b>Ethnicity</b>			
Asian	-0.140***	-0.051***	-0.192***
Black	-0.116***	0.039	-0.077**
Mixed	-0.098***	-0.104	-0.201
Other	-0.047**	-0.142	-0.189*

$\lambda = 0.95$ .

Significance: \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ .

Reference categories: NS-SEC 3, Age 16-34, White. N = 6,856.

TABLE X: SDEM with  $K = 50$  Nearest Neighbors

Variable	Direct	Indirect	Total
<b>Infrastructure</b>			
N. of Facilities ( <i>p-value</i> )	0.00002 (0.166)	0.00007 (0.518)	0.00009 (0.420)
<b>Social Class</b>			
High Managerial (1-2)	0.101***	0.209***	0.310***
Small Employers (4)	0.063***	0.109	0.172**
Lower Supervisory (5)	-0.019	-0.270	-0.289
Routine (6-7)	-0.114***	0.038	-0.075
Unemployed (8)	-0.180***	0.267***	0.087
Students (9)	0.137***	0.026	0.162***
<b>Age</b>			
Age 35-54	0.029**	-0.140**	-0.111*
Age 55-74	-0.032***	-0.197***	-0.229***
Age 75+	-0.246***	0.082	-0.164*
<b>Ethnicity</b>			
Asian	-0.137***	-0.102***	-0.239***
Black	-0.111***	0.047	-0.063*
Mixed	-0.101***	-0.402**	-0.504***
Other	-0.064***	-0.014	-0.078

$\lambda = 0.96$ .

Significance: \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ .

Reference categories: NS-SEC 3, Age 16-34, White. N = 6,856.

TABLE XI: Alternative Infrastructure Variable (Sites)

Variable	Direct	Indirect	Total
<b>Infrastructure</b>			
N. of Sites ( <i>p</i> -value)	0.00010 (0.043)	0.00029 (0.163)	0.00039 (0.102)
<b>Social Class</b>			
High Managerial (1-2)	0.069***	0.059**	0.128***
Small Employers (4)	0.035*	0.099**	0.134**
Lower Supervisory (5)	-0.076*	-0.174	-0.250*
Routine (6-7)	-0.140***	-0.024	-0.164***
Unemployed (8)	-0.211***	0.104*	-0.107*
Students (9)	0.102***	-0.003	0.099***
<b>Age</b>			
Age 35-54	0.012	-0.033	-0.021
Age 55-74	-0.040***	-0.115***	-0.154***
Age 75+	-0.265***	0.091*	-0.174***
<b>Ethnicity</b>			
Asian	-0.142***	-0.038***	-0.180***
Black	-0.121***	0.030	-0.091***
Mixed	-0.094***	-0.047	-0.141
Other	-0.050**	-0.034	-0.084

$\lambda = 0.93$ .

Significance: \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ .

Reference categories: NS-SEC 3, Age 16-34, White. N = 6,856.

## APPENDIX B

### DIVERSITY MODEL: FULL IMPACT TABLE

This appendix provides the full impact estimates for the Diversity Model discussed in Section V-D. While the main text compares the infrastructure coefficients, Table XII details the complete set of social class and demographic drivers associated with the *Diversity Index* specification.

TABLE XII: Full Impacts: Diversity Model on Active Adults (SDEM)

Variable	Direct	Indirect	Total
<b>Infrastructure</b>			
Diversity Index ( <i>p</i> -value)	0.00018* (0.013)	0.00061 (0.052)	0.00079* (0.030)
<b>Social Class (NS-SEC)</b>			
High Managerial (1-2)	0.070***	0.060**	0.130***
Small Employers (4)	0.038*	0.103**	0.141***
Lower Supervisory (5)	-0.074	-0.163	-0.237*
Routine (6-7)	-0.140***	-0.026	-0.166***
Unemployed (8)	-0.209***	0.105*	-0.104*
Students (9)	0.102***	-0.004	0.098***
<b>Age</b>			
Age 35-54	0.011	-0.037	-0.026
Age 55-74	-0.040***	-0.114***	-0.153***
Age 75+	-0.267***	0.082*	-0.186***
<b>Ethnicity</b>			
Asian	-0.142***	-0.038***	-0.180***
Black	-0.121***	0.031	-0.090***
Mixed	-0.097***	-0.056	-0.153
Other	-0.050**	-0.036	-0.087

$\lambda = 0.93$ .

Significance: \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ .

Reference categories: NS-SEC 3, Age 16-34, White. N = 6,856.

**APPENDIX C**  
**GENDER ANALYSIS: FULL ESTIMATES**

This appendix presents the complete impact estimates for the gender-disaggregated analysis discussed in Section V-F. Appendix Section C-A details the separate models for Male and Female active participation, highlighting the differential elasticity to supply-side factors. Appendix Section C-B presents the robustness check performed on the *Inactivity Gap*, confirming the inequality multiplier effect.

*A. Active Participation by Gender*

TABLE XIII: Male Active Participation (Facilities Model)

Variable	Direct	Indirect	Total
<b>Infrastructure</b>			
N. of Facilities ( <i>p-value</i> )	0.00003* (0.038)	0.00008 (0.200)	0.00011 (0.126)
<b>Social Class</b>			
High Managerial (1-2)	0.041**	0.053*	0.093***
Small Employers (4)	-0.027	0.108**	0.082*
Lower Supervisory (5)	-0.174***	-0.146	-0.321**
Routine (6-7)	-0.158***	-0.041	-0.200***
Unemployed (8)	-0.203***	0.091*	-0.112*
Students (9)	0.072***	-0.004	0.068***
<b>Age</b>			
Age 35-54	0.037***	-0.033	0.004
Age 55-74	-0.052***	-0.096***	-0.148***
Age 75+	-0.176***	0.072*	-0.104*
<b>Ethnicity</b>			
Asian	-0.131***	-0.041***	-0.172***
Black	-0.114***	0.029	-0.084***
Mixed	-0.099***	-0.062	-0.161
Other	-0.052**	-0.016	-0.069

$\lambda = 0.93$ .

Significance: \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ .

Reference categories: NS-SEC 3, Age 16-34, White. N = 6,856.

TABLE XIV: Female Active Participation (Facilities Model)

Variable	Direct	Indirect	Total
<b>Infrastructure</b>			
N. of Facilities ( <i>p-value</i> )	0.00000 (0.766)	0.00005 (0.480)	0.00005 (0.503)
<b>Social Class</b>			
High Managerial (1-2)	0.094***	0.071**	0.165***
Small Employers (4)	0.084***	0.107**	0.191***
Lower Supervisory (5)	-0.006	-0.188	-0.194
Routine (6-7)	-0.115***	-0.008	-0.123***
Unemployed (8)	-0.220***	0.114**	-0.106*
Students (9)	0.140***	-0.005	0.135***
<b>Age</b>			
Age 35-54	0.006	-0.044	-0.037
Age 55-74	-0.020**	-0.128***	-0.149***
Age 75+	-0.325***	0.100**	-0.225***
<b>Ethnicity</b>			
Asian	-0.156***	-0.035	-0.190***
Black	-0.124***	0.035	-0.089***
Mixed	-0.081**	-0.024	-0.105
Other	-0.059***	-0.049	-0.108

$\lambda = 0.93$ .

Significance: \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ .

Reference categories: NS-SEC 3, Age 16-34, White. N = 6,856.

TABLE XV: Male Active Participation (Diversity Model)

Variable	Direct	Indirect	Total
<b>Infrastructure</b>			
Diversity Index ( <i>p-value</i> )	0.00025*** (0.001)	0.00075* (0.016)	0.00100** (0.006)
<b>Social Class</b>			
High Managerial (1-2)	0.039*	0.050*	0.090***
Small Employers (4)	-0.028	0.104**	0.076
Lower Supervisory (5)	-0.175***	-0.138	-0.313**
Routine (6-7)	-0.160***	-0.046	-0.206***
Unemployed (8)	-0.204***	0.090*	-0.114*
Students (9)	0.070***	-0.007	0.063**
<b>Age</b>			
Age 35-54	0.037***	-0.034	0.002
Age 55-74	-0.052***	-0.093***	-0.145***
Age 75+	-0.177***	0.065	-0.113**
<b>Ethnicity</b>			
Asian	-0.131***	-0.041***	-0.171***
Black	-0.114***	0.029	-0.085***
Mixed	-0.100***	-0.060	-0.160
Other	-0.054**	-0.018	-0.072

$\lambda = 0.93$ .

Significance: \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ .

Reference categories: NS-SEC 3, Age 16-34, White. N = 6,856.

TABLE XVI: Female Active Participation (Diversity Model)

Variable	Direct	Indirect	Total
<b>Infrastructure</b>			
Diversity Index ( <i>p-value</i> )	0.00009 (0.256)	0.00044 (0.171)	0.00053 (0.161)
<b>Social Class</b>			
High Managerial (1-2)	0.093***	0.070**	0.163***
Small Employers (4)	0.083***	0.105**	0.187***
Lower Supervisory (5)	-0.007	-0.183	-0.190
Routine (6-7)	-0.116***	-0.010	-0.127***
Unemployed (8)	-0.221***	0.114**	-0.107*
Students (9)	0.139***	-0.007	0.132***
<b>Age</b>			
Age 35-54	0.006	-0.045	-0.039
Age 55-74	-0.020*	-0.127***	-0.147***
Age 75+	-0.326***	0.095*	-0.230***
<b>Ethnicity</b>			
Asian	-0.156***	-0.034	-0.190***
Black	-0.124***	0.035	-0.089***
Mixed	-0.081**	-0.022	-0.103
Other	-0.059***	-0.051	-0.110

$\lambda = 0.93$ .

Significance: \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ .

Reference categories: NS-SEC 3, Age 16-34, White. N = 6,856.

## B. Inactivity Gap Analysis

TABLE XVII: Inactivity Gender Play Gap: Facilities Model

Variable	Direct	Indirect	Total
<b>Infrastructure</b>			
N. of Facilities ( <i>p-value</i> )	-0.00002*** (< 0.001)	-0.00003 (0.137)	-0.00005* (0.020)
<b>Social Class</b>			
High Managerial (1-2)	0.025***	0.004	0.029***
Small Employers (4)	0.077***	-0.054***	0.023**
Lower Supervisory (5)	0.125***	-0.090***	0.035
Routine (6-7)	0.017**	0.012*	0.029***
Unemployed (8)	-0.047***	0.054***	0.008
Students (9)	0.043***	0.006	0.049***
<b>Age</b>			
Age 35-54	-0.021***	0.004	-0.017**
Age 55-74	0.040***	0.007	0.047***
Age 75+	-0.149***	0.018*	-0.131***
<b>Ethnicity</b>			
Asian	-0.031***	0.001	-0.029***
Black	-0.008**	-0.006	-0.014***
Mixed	0.027**	0.059***	0.086***
Other	-0.003	-0.029**	-0.032***

$\lambda = 0.52$ .

Significance: \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ .

Reference categories: NS-SEC 3, Age 16-34, White. N = 6,856.

Note: Negative coefficients indicate a widening of the gap (further from 0).

TABLE XVIII: Inactivity Gender Play Gap: Diversity Model

Variable	Direct	Indirect	Total
<b>Infrastructure</b>			
Diversity Index ( <i>p-value</i> )	-0.00012*** (< 0.001)	-0.00014 (0.134)	-0.00026* (0.016)
<b>Social Class</b>			
High Managerial (1-2)	0.025***	0.004	0.029***
Small Employers (4)	0.076***	-0.052***	0.024**
Lower Supervisory (5)	0.125***	-0.092***	0.032
Routine (6-7)	0.017**	0.013*	0.030***
Unemployed (8)	-0.047***	0.055***	0.008
Students (9)	0.043***	0.006	0.049***
<b>Age</b>			
Age 35-54	-0.021***	0.004	-0.017**
Age 55-74	0.040***	0.006	0.046***
Age 75+	-0.149***	0.019*	-0.130***
<b>Ethnicity</b>			
Asian	-0.031***	0.001	-0.030***
Black	-0.008**	-0.006	-0.014***
Mixed	0.028**	0.059***	0.087***
Other	-0.002	-0.030**	-0.032***

$\lambda = 0.52$ .

Significance: \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ .

Reference categories: NS-SEC 3, Age 16-34, White. N = 6,856.

Note: Negative coefficients indicate a widening of the gap (further from 0).