Passing on the Flame: Do Mega Sports Events Promote Health Behaviours?*

Christian Krekel[†], Georgios Kavetsos[‡], Nicolas R. Ziebarth[§]

January 2025

Abstract

To justify billion-dollar public expenditures on mega sports events, proponents often suggest lasting improvements in health behaviours among the general public. To estimate the returns to health behaviours from hosting the 2012 London Olympics, we collected panel data on more than 19,000 respondents across two European capitals, London and Paris, between 2011 and 2013. Using a difference-in-differences design with Paris as counterfactual, we find an increase in physical activity by six percentage points among the inactive, from a baseline of 34%. Activation, however, lasts only for about 100 days. Although we also find suggestive evidence for reduced alcohol and tobacco consumption during the event, a cost-benefit analysis suggests that staging mega sports events is not a cost-effective policy to promote lasting health behaviour change.

Keywords: Mega sports events, Olympics, health behaviours, physical activity, quasi-experiment, cost-benefit analysis
JEL: I12, I18, D90

*We are grateful to Joanna Coast and two anonymous referees, as well as Richard Cookson, Paul Dolan, Vassil Girginov, Laura Kudrna, Richard Layard, Sabrina Lenzen, Ekaterina Oparina, Robert Metcalfe and participants at the 8th workshop of the Erasmus Centre for Applied Sports Economics, for their helpful comments and suggestions. We thank Larissa Davies and Andrew Spiers for insightful discussions throughout this project. Nils Mallock provided excellent research assistance. Support from the Economic and Social Research Council (ESRC: RES 360-25-0036), the German Science Foundation (Deutsche Forschungsgemeinschaft, DFG: WA 547/5-1), CEPREMAP, the French National Research Agency, through the programme Investissements d'Avenir (ANR-10-LABX-93-01), and the Open Research Area in Europe for the Social Sciences (ORA-10-36) is gratefully acknowledged. This study is not part of a for-pay consultancy. The authors state that they do not have any financial interest in this study. There are no known conflicts of interest. Statement of Contributions: Krekel: Conceptualization, Methodology, Software, Validation, Formal Analysis, Lead Writing – Original Draft, Lead Writing – Review Editing, Visualization, Project Administration; Kavetsos: Lead Conceptualization, Methodology, Resources, Writing - Original Draft, Project Management, Funding Acquisition; Ziebarth: Lead Conceptualization, Methodology, Resources, Writing - Original Draft, Funding Acquisition

[†]Corresponding Author, Centre for Economic Performance (CEP), London School of Economics (LSE); Department of Psychological and Behavioural Science, LSE (c.krekel@lse.ac.uk)

[‡]Queen Mary University of London, School of Business and Management; CEP, LSE; Lee Kum Sheung Center for Health and Happiness, Harvard University (g.kavetsos@qmul.ac.uk)

[§]ZEW Mannheim and University of Mannheim (nicolas.ziebarth@zew.de)

1 Introduction

For decades, there have been intensive debates about the worthwhileness of hosting mega sports events like the Olympics (cf. IMF, 2010). Given increasing costs and taxpayer money spent on such events[1], a growing literature estimates the returns to hosting, focusing mostly on economic indicators such as income, wages, investments, or tourism (Baade and Matheson, 2016). Most studies find little to no tangible impacts on the economy. In their review, Baade and Matheson (2016) conclude that *"in most cases, the Olympics are a money-losing proposition for host cities."* Economists have thus started to study the returns to *intangible* outcomes such as the wellbeing of residents in the host country or city, documenting significant impacts (Kavetsos and Szymanski, 2010; Dolan et al., 2019). Dolan et al. (2019), in particular, find that staging the 2012 London Olympics raised the happiness of Londoners, strong enough that these could possibly justify the billion-dollar public expenditures.

In this paper, we estimate the impact of hosting the Olympics on health behaviours among residents in the host city, using the example of London 2012. Lasting improvements in health behaviors are often put forward to justify bidding for mega sports events (Kavetsos and Szymanski, 2009; London Bid, 2012), yet causal evidence evaluating such claims is missing. We ask: *Is hosting the Olympics an "intervention" capable of encouraging residents to improve their health behaviours? If so, does it "stick"?* To our knowledge, we are the first to estimate the causal returns from hosting the Olympics on health behaviours among residents in the host city and to present such evidence in a quasi-experimental setting.

To estimate the causal effects of hosting the Games, we exploit a unique feature of London 2012 as a quasi-experiment: both London and Paris bid to host the Olympics, yet Paris – considered among the front runners to win the bid – lost by a close 54 to 50 vote in favour of London. Both cities were, therefore, similarly interested and invested in staging the event. We show that, in our data, residents of both cities had similar observable characteristics and followed a common trend in health behaviours prior to the event. We exploit this quasi-random allocation into treatment and control groups in a difference-in-differences (DiD) design, comparing within-individual changes in health behaviours between Londoners

[1] The Olympic Summer Games last two weeks. Atlanta 1996 had estimated costs of 4.2$ billion, Sydney 2000 of 8.1$ billion, Athens 2004 of 18.7$ billion, and Beijing 2008 of 52.7$ billion. Tokyo 2020, which was confounded by Covid-19, had costs of 35$ billion. In contrast, London 2012 and Rio 2016 were relatively modest affairs, with estimated costs between 12$ and 14$ billion (Council on Foreign Relations, 2021). Paris 2024 had a budget of 7.7$ billion (France 24, 2020).

(our treatment group) with those of Parisians (our control group) over time. We use our own, uniquely designed and collected panel data on more than 19,000 respondents living in London and Paris (about 9,500 per city) during the summers of 2011 (before), 2012 (during), and 2013 (after the Olympics). Our surveys, which were incentivised, are broadly representative of the population in each city and should thus score high in terms of external validity.

Conceptually, we interpret the Olympics as an intervention that aims to, among others, raise people's appreciation of sport and physical activity, fair competition, diversity of human culture, and human excellence (IOC, 2022). This could change people's attitudes, beliefs, and (ultimately) behaviours. Hosting the Olympics could make health behaviours and active lifestyles seem more attractive, fashionable, "cool", or even patriotic (Kavetsos and Szymanski, 2009; Baumann et al., 2021). While the Olympics are followed worldwide, they are particularly salient in the host city. In the case of London 2012, the event summoned 10,820 athletes from 204 nations to compete in 302 tournaments. London hosted more than 1,000 accompanying sports or cultural events targeted at the general public (Mayor of London, 2012). Thirty-one percent of Londoners have attended some event (Greater London Authority, 2013). Importantly, the host city stages the opening and closing ceremonies, which attract peak media attention, with 27 million UK viewers in the case of London 2012 (Ofcom, 2012). In a post-Olympics survey, 65% of Londoners reported that the summer of 2012 was a *"summer like no other"*, and 69% thought that the Olympics would increase sports participation among adults (81% among children) (Greater London Authority, 2013).

The bid from London 2012, in particular, argued that hosting the Olympics would *"inspire [...] to greater sporting activity"* (p.19). It also mentioned legacy effects prominently as part of its vision, suggesting that hosting would leave *"a legacy for sport in Britain"* (p.17) and that *"the UK will build on the sporting momentum of the expected successes coming out of the* 2012 *Games"* (p.19). More locally in London, the Olympic Park would provide *"people with significant improvements in health and well-being"* (p.23) (London Bid, 2012). The UK Department for Culture, Media and Sport (DCMS)'s foremost aim for the event was to *"make the UK a world-class sporting nation"*, with a particular focus on *"mass participation"* (Department for Culture, Media and Sport, 2007). Such claims are not unique to London 2012. More recently, those of Paris 2024 argued that hosting would be a *"powerful catalyst for participation"* (p.13), given the *"inspiring role that sport champions can play in driving sport* *participation and positive lifestyles*" (p.23) (Paris 2024 Candidate City, 2016). Moreover, claims about lasting improvements in health behaviours are not exclusive to the Olympics. In its recent, united bid to host the 2026 FIFA World Cup, Canada, Mexico, and the US argue that hosting the event would see *"a surge in enthusiasm for football, which can be used to further improve people's access and ability to play the game*" (p.10), and that it would *"inspire increased registration in youth and adult recreational football*" (p.11) (United 2026, 2017). We thus expect our findings to yield generalisable insights into the impacts of mega sports events.

Despite all these claims, previous reviews find little systematic, causal evidence for lasting improvements in health behaviours (McCartney et al., 2010; Weed et al., 2012; Mahtani et al., 2013). Smith et al. (2014), in a prospective cohort study among East Londoners, found that adolescents who were positively engaged with London 2012 reported a higher frequency of physical activity and less sedentary time than those who were less engaged. However, impacts were not persistent (cf. Cummins et al., 2018). In another study, using repeated cross-sectional data from the Active People Survey in England, Kokolakakis et al. (2019, 2020) find positive associations between London 2012 and increased sports participation. However, the authors cannot establish a causal link due to the lack of a valid counterfactual. The mixed results and lack of causality in these studies highlight the need for more causal evidence on alleged pro-health behavioural claims.

Applying our quasi-experimental methods, we obtain three results. First, on average, we find no changes in physical activity among residents of London. However, we do find a significant increase among those who were previously inactive, by six percentage points, from a baseline of 34%. This is a large effect of 18%. The inactive are arguably the most policy-relevant group: medical evidence shows that even modest physical activity carries large health benefits (cf. Myers, 2008). Hence, the UK Chief Medical Officer's *Physical Activity Guidelines* suggests that interventions targeted at this subpopulation should have the largest health benefits (UK Chief Medical Officer, 2019). Further, we find no crowding out among the previously active; rather, the previously active in London are, on average, more likely to *remain* physically active during the Olympics relative to the previously active in Paris. Second, we also find suggestive evidence for reduced alcohol and tobacco consumption, yet these effects are relatively small in size, are temporally limited to the duration of the main event, and do not withstand a correction for multiple hypotheses testing using stepdown-adjusted P-values (Romano and Wolf, 2005b,a). Third, activation of the previously inactive lasts only for about 100 days, suggesting no long-term improvements in health behaviours. Once the main event ends, the impact fades. Finally, a back-of-the-envelope cost-benefit analysis, comparing potential healthcare cost savings from increased physical activity among the previously inactive with the allocated budget for activation, suggests that hosting the Olympics is not a cost-effective policy to promote health behaviour change.

Using DiD models with sharp cut-off dates, we show that the increase among those Londoners who were previously inactive is strongest during the main event (between the opening and closing ceremonies, when infrastructure and information were largely fixed). While we do not detect an additional impact of medals won, we find that Londoners who were previously inactive are more likely to think about their health during the event. Together, this suggests that the act of hosting, and potentially the salience of the Olympics or physical activity more generally, rather than national sporting success is a likely mechanism behind our results.

Our findings align well with experimental evidence in the economics literature. Research has shown that even small monetary incentives can encourage people to change their health behaviours and exercise, for example by going to the gym. However, effects are often short-lived, suggesting that most people find it difficult to permanently change their habits, possibly due to limited self-control (DellaVigna and Malmendier, 2006; Charness and Gneezy, 2009; Akee et al., 2013; Royer et al., 2015; Stutzer and Meier, 2016; Rohde and Verbeke, 2017; Carrera et al., 2018, 2019; Aggarwal et al., 2020; Milkman et al., 2021). Notable exceptions are Reichert (2015) and Augurzky et al. (2018), who find long(er)-lasting effects on weight loss and employment prospects in a randomised controlled trial among health plan enrollees with obesity.[2]

Only a few studies exploit quasi-experiments to evaluate the effectiveness of policies aimed at changing health behaviours (Cawley et al., 2013; Sarma et al., 2014; Nakamura et al., 2021). Marcus et al. (2022) study a policy that distributed vouchers for sports club memberships among all 33,000 third-graders in a German federal state in 2009. While policy awareness was still high even after a decade, neither in the short- nor the long-run did the policy increase the share of students who were sports club members or physically active. We contribute to

[2] Homonoff et al. (2020) find that rebate-framed incentives for gym attendance significantly and strongly increase gym attendance among university students, with half of the effect lasting even one year after the incentives have been taken away.

this limited quasi-experimental literature by studying the case of hosting mega sports events.

Finally, our paper adds to a much broader literature on the benefits of physical activity, or, rather, the adverse consequences of physical *in*activity such as obesity, see Cawley (2011) for an interdisciplinary overview. Obesity rates in the UK and around the world are high and rising, particularly among deprived groups (Griffith, 2022). The UK Government estimates that a lack of physical activity costs the UK about £7.4 billion per year, including £0.9 billion to the National Health Service (NHS) alone (Public Health England, 2016). It is well documented that regular physical activity reduces health risks associated with premature mortality and chronic health conditions, including high blood pressure, coronary heart disease, type-II diabetes, cancer, and osteoporosis (Batty, 2002; Stamatakis et al., 2007; Humphreys et al., 2014), resulting in substantial healthcare cost savings (Sari, 2009; Cawley and Meyerhoefer, 2012). Health behaviours and active lifestyles also contribute towards improved mental health and wellbeing (Dimeo et al., 2001; Dolan et al., 2014), social outcomes (Puhl and Heuer, 2009), and labour market prospects (Barron et al., 2000; Cawley, 2004; Morris, 2006, 2007; Lechner, 2009; Kavetsos, 2011; Rooth, 2011; Hyytinen and Lahtonen, 2013).

Despite the benefits of physical activity, modern life has led people to adopt sedentary lifestyles, contributing to rising obesity rates around the world (Costa-Font and Mas, 2016; World Health Organization, 2021). A rich literature in economics looks at the causal driving forces behind this trend, including the growing availability of (fast food) restaurants (Currie et al., 2010; Dunn, 2010; Anderson and Matsa, 2011), increased portion sizes, and consumption of soft drinks (Jeitschko and Pecchenino, 2006; Fletcher et al., 2010), gluttony (Griffith et al., 2015), changes in income and relative (food) prices (Courtemanche, 2009; Wehby and Courtemanche, 2012; Akee et al., 2013; Courtemanche et al., 2014; Dubois et al., 2014; Grossman et al., 2014; Griffith et al., 2015; Dragone and Ziebarth, 2017), or family and social influences (Strulik, 2014; Cawley et al., 2019), as well as interventions on how to reverse this trend, including calorie labelling (Cawley et al., 2020, 2021), sin taxes (Fletcher et al., 2010; Muller et al., 2017), healthy school meals (Belot and James, 2011), interventions targeting type of food as well as timing and frequency of healthy food intake (Belot et al., 2018), early childhood programmes (Conti et al., 2016), or education more generally (Kamhöfer et al., 2019). Our paper adds to this literature by studying mega sports events as large-scale interventions to promote physical activity among the general public.

2 Data and Methods

2.1 Data Collection

We use our own, uniquely designed and collected panel data. They are large-scale and cover two European capitals over three years. In particular, we surveyed more than $19,000$ unique respondents in London (the host city, treatment group) and Paris (control group) during the summers of 2011 (8 August – 30 September, 2011, i.e. the year before the Olympics); 2012 (20 July – 2 October, 2012, i.e. the year of the event); and 2013 (23 July – 12 September, 2013, i.e. the year after). Importantly, our survey was not framed as being related to the Olympics, health behaviours, or active lifestyles to avoid selection into the survey, social desirability bias, or attitude expression.[3]

We chose Paris as counterfactual because it bid to host the 2012 Olympics but lost by a close vote. It was thus similarly interested and invested in staging the event. Just like London, Paris had hosted the Olympics several times in the past, most recently in 2024. London and Paris are the two largest metropolitan areas in Europe in terms of population size and broadly comparable in terms of socio-economic structure, culture, and wealth. In Section 2.2, we show that individuals in our data are comparable between both cities and that they followed a common trend in outcomes before the event. Hence, Paris can be seen as a valid counterfactual.

Survey Mode. Ipsos MORI collected the data on our behalf using a mix of online and phone surveys. There were three survey waves: one in the summer of 2011, one in the summer of 2012, and one in the summer of 2013. The same respondent was exposed to the same survey mode in all waves. It was randomised when exactly in the summers of 2011, 2012, and 2013 the invitation to participate in the survey was sent, i.e. either before, during, or after the equivalent of the Olympics period in the summer of 2012. This applied to both London and Paris.[4] In each city, Ipsos MORI collected a sample that is representative of either the population with broadband (London) or the general population (Paris), in terms of age, gender, and employment status. To minimise attrition, we incentivised respondents

[3]Our data collection was funded by an Open Research Area in Europe for the Social Sciences (ORA) grant on evaluating the intangible impacts of the 2012 London Olympics. The data, including surveys, are freely available from the UK Data Service: https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=8267

[4]The online sample was recruited using the Ipsos MORI Interactive Services Panel (IIS) and was released on a rolling basis each week. The phone sample was recruited via random digit dialling.

through inclusion in a lottery.[5] Note that broadband take-up was already 78% in London in 2011 (Ofcom, 2012). Thus, the population with and without broadband in London, and by extension, between our London and Paris samples should be similar.[6]

Sample Size. Our raw sample includes 36,607 observations on 19,144 unique individuals, of which 9,483 are located in London and 9,661 in Paris. With, on average, 235 observations per day over an observation period of 126 days in our preferred specification, our data are much higher-powered than existing secondary data, enabling us to study event dynamics using sharp cut-off dates. Due to attrition, our sample reduced from 19,144 individuals in 2011 to 10,820 (57%) in 2012 and 6,643 (35%) in 2013. Attrition was slightly larger in London than in Paris (e.g. out of the remaining 10,820 individuals in 2012, 4,762 are in London and 6,058 are in Paris). We turn to selection into the follow-up survey in Section 3.2, where we use matching and inverse-probability weighting to account for differential attrition. As we will see, our results remain robust.

Outcomes. We are primarily interested in physical activity. Our outcome stems from a single-item, six-point Likert scale asking respondents: "How often do you usually participate in sport or physical activity?" Answer possibilities include: 1 (">5 Times a Week"), 2 ("$3-4$ Times a Week"), 3 ("$1-2$ Times a Week"), 4 ("$1-3$ Times a Month"), 5 ("Less Often" than once a month), and 6 ("Never"). We dichotomise this item such that categories 1 to 4 are one and 5 to 6 zero. That is, we define being active as engaging in sport or physical activity at least monthly (66% in our estimation mean) as opposed to less often (34%). We label this outcome *Physical Activity*.

We chose this definition because we are primarily interested in the extensive margin of physical activity (i.e. being active as opposed to inactive). Arguably, from a public health perspective, encouraging inactive people to get active is more policy-relevant than encouraging already active people to exercise even more (cf. Myers, 2008). Moreover, our definition of being inactive lines up with benchmark surveys according to which 25% of the UK population were

[5] Prizes were £/€500 in each city in 2011, £/€1,000 in 2012 and £/€1,500 in 2013.

[6] As we will see, there are indeed little differences in pre-treatment observable characteristics between Londoners and Parisians in our data (Appendix Table A1). Controlling for observables or not produces robust results (Tables A2 and A3), while an Oster (2019) bounding analysis suggests that selection on unobservables is no major threat to identification (Appendix Section B). Note that simple level differences in observables (or unobservables) should be no major threat to identification either as our models elicit *relative changes* in outcomes pre- and post-treatment. Finally, selective attrition seems to be of little concern too (Table 5).

inactive at the time (Public Health England, 2016). Apart from the extensive margin, we also exploit the different categories of physical activity as outcomes to look at the intensive margin and to show where in the distribution the movement, if any, is coming from.

Beyond physical activity, we also study other health behaviours as secondary outcomes, in particular *alcohol* and *tobacco consumption*. For alcohol, our outcome is the self-reported number of glasses of alcoholic drinks on the previous day; for tobacco, it is the self-reported number of cigarettes.[7] While self-reports referring to consumption *on the previous day* may include reporting biases due to, for example, social desirability, such elicitation is very common in surveys (cf. DeSimone et al., 2023). In fact, Stockwell et al. (2004) conclude that "Recent recall methods encourage fuller reporting of volumes plus more accurate estimates of unrecorded consumption [...]." Further, eliciting consumption on the previous day allows us to study consumption as an outcome at the daily level. To the extent that misreporting still exists but is unrelated to treatment and similar among Londoners and Parisians, it should be no major threat to identification as we study *relative changes* over time. The average number of alcoholic drinks in our estimation sample was 0.93 the day before the interview, and the average number of cigarettes smoked was 2.7. To our knowledge, there were no supply-side constraints regarding alcohol or tobacco products during our observation period. Table 1 shows summary statistics for our outcomes and covariates.

Table 1 about here

On average, respondents in our estimation sample are 45 years old (standard deviation of 15); 14% are between 16 and 27, 23% between 28 and 37 years, 21% between 38 and 47, 18% between 48 and 57, 18% between 58 and 67, and 6% above 68 years old. We thus cover a wide age range, including young, middle-aged, and older respondents.

2.2 Estimation and Identification

Our DiD design compares changes in health behaviours of respondents living in London (our treatment group) with those living in Paris (our control group) – from before, to during, to after the Olympics. We use three models. The first solely focuses on the year of the Olympics, 2012:

[7] The former is obtained from an open-text question asking respondents: "How many glasses of an alcoholic drink (e.g. beer, wine, spirit, etc.) did you have yesterday?". The latter comes from an open-text question asking: "How many cigarettes did you smoke yesterday?".

$$
\begin{aligned}
y_i = \beta_0 &+ \beta_1 London \times Olympics + \beta_2 London \times PostOlympics \\
&+ \beta_3 London + \beta_4 Olympics + \beta_5 PostOlympics \\
&+ \beta_6' X_i + t_m + t_{wd} + \epsilon_i
\end{aligned}
\tag{1}
$$

where y_i is the health behaviour of individual i; *London* is a dummy that is one if the individual lives in London; *Olympics* and *PostOlympics* are dummies that are one if the respondent is interviewed during the Olympics (27 July – 12 August, 2012), and zero if interviewed after (13 August – 2 October, 2012). t_m and t_{wd} are month and weekday fixed effects. The vector X_i includes demographics, including age and age squared, gender, marital status, whether the individual is still in education or not, the highest educational degree obtained, and the type of accommodation. Note that demographics are surveyed in each wave, i.e. in 2011, 2012, and 2013.[8] Equation (1) compares *between*-individual changes in health behaviours of Londoners with those of Parisians in 2012. *London* nets out systematic differences in unobservables between Londoners and Parisians at the city level, and *Olympics* and *PostOlympics* between individuals interviewed in different periods in 2012.

Our second model exploits the panel dimension and eliminates time-invariant unobservables via the inclusion of individual fixed effects u_i:

$$
y_{it} = \beta_1 London \times 2012 + \beta_2 2012 + \beta_3' X_{it} + t_m + t_{wd} + u_i + \epsilon_{it} \tag{2}
$$

where 2012 is a year dummy and $t = \{2011, 2012\}$ denotes the temporal dimension of our panel data. In Section 3.3, we also study long-term (legacy) effects by making use of our 2013 survey.

Equation (2) now compares the *within*-individual changes in health behaviours of Londoners with those of Parisians from 2011 to 2012. In doing so, it takes the whole period in 2012 (20 July – 2 October, 2012) as the relevant treatment period. This is a conservative approach, and hence our preferred specification: while the Olympics are officially constrained by their

[8]In this specification, we omit potentially endogenous controls (i.e. employment and income). In Section 3.2, we include them along with controls for meteorological conditions (i.e. local temperature and precipitation). These account for potential differences in outdoor weather conditions between London and Paris possibly linked to our outcomes. As we will see, our results remain robust.

opening and closing ceremonies (27 July and 12 August, 2012), many associated events take place throughout the summer and autumn (including the Paralympics, which took place between 29 August and 9 September, 2012). On the one hand, such events may drive health behaviours, be it because they directly target these or, for example, because of their relative salience. On the other hand, we expect any salience of the Olympics or physical activity more generally to be strongest during the main event.

To test this empirically, our third model splits the 2012 dummy into three dummies that capture distinct sub-periods: before, during, and after the main event. This model also offers a first insight into an immediate decay effect, if present:

$$\begin{aligned} y_{it} = &+ \beta_1 London \times PreOlympics_{2012} + \beta_2 London \times Olympics_{2012} \\ &+ \beta_3 London \times PostOlympics_{2012} \\ &+ \beta_4 PreOlympics_{2012} + \beta_5 Olympics_{2012} + \beta_6 PostOlympics_{2012} \\ &+ \beta_7' X_{it} + t_m + t_{wd} + u_i + \epsilon_{it} \end{aligned} \tag{3}$$

where $PreOlympics_{2012}$, $Olympics_{2012}$, and $PostOlympics_{2012}$ are dummies that are one if the individual is interviewed, respectively, before (20 July – 26 July), during (27 July – 12 August), or after the main event (13 August – 2 October, 2012).

We estimate all three models for alcohol and tobacco consumption, yet only the second and third for physical activity. While our questions on alcohol and tobacco consumption ask about consumption *on the previous day*, our question on physical activity asks about *usual* participation, which should be less susceptible to meaningful change in our first model, as it focuses only on the closely spaced time period of the Olympics in 2012.

We estimate linear models using OLS with robust standard errors clustered at the interview date level, which is justified by the daily variation in exposure to the Olympics. In Section 3.2, we use wild-cluster bootstrapped standard errors at the city level (the level of quasi-randomisation) as a robustness check. Our results for physical activity as our primary outcome become, if anything, more significant.[9] We prefer linear over probit or logit models for two reasons. First, neither probit nor logit is readily applicable to panel data and individual fixed effects (due to the incidental parameter problem). Second, linear models are easier

[9] Our results are also robust to clustering at the individual level.

to interpret: in linear models, a one-unit increase in a regressor directly translates into a one percentage-point increase in the probability that the outcome is one, whereas, say, logit models yield coefficients in log-odds (or odds-ratios), which are less intuitive. Note that, in many situations without extreme probabilities (which are characterised by highly non-linear relationships between probabilities and log-odds), linear and logit models produce results that are very similar (Angrist and Pischke, 2009). When estimating logit instead of linear models for our binary outcomes, our results remain robust.[10]

Note that our DiD design avoids potentially biased estimates due to treatment effect heterogeneity and dynamics in staggered DiD designs (Goodman-Bacon, 2021).

Identifying Assumptions. Our parameters of interest are β_1 and β_2 in Equation (1), β_1 in Equation (2), and β_2 and β_3 in Equation (3). We interpret them as the average treatment effects on the treated (ATT) – the average causal effects of hosting the Olympics on the health behaviours of host-city residents – if two identifying assumptions are satisfied, conditional on time-varying observables X_{it}, time fixed effects t_m and t_{wd}, and individual fixed effects u_i: *(i)* exogeneity of treatment and *(ii)* a common trend pre-treatment.

Regarding *(i)* exogeneity of treatment, we exploit that both London and Paris bid to host the 2012 Olympics. Paris lost by a small margin of 54 to 50. Moreover, the entire city hosts, which is exogenous from the perspective of a single individual. Finally, we only consider respondents who reported living in London or Paris during the entire observation period, omitting movers.[11]

Appendix Table A1 shows balancing properties of observables by calculating scale-free normalised differences to account for large group sizes. According to Imbens and Wooldridge (2009); Imbens and Rubin (2015), a normalised difference greater than 0.25 suggests covariate imbalance. As seen, we only find imbalances for some educational variables and thus routinely control for these. We study coefficient stability in Section 3.2, where we follow the bounding argument by Oster (2019) to elicit the relative importance of unobservable selection. As we will see, our results remain robust.

Regarding *(ii)* a common trend pre-treatment, Appendix Figure A1 plots the share of

[10] See Section 3.2 for these results.

[11] We include both citizens and non-citizens. Our results are robust to omitting non-citizens (available upon request). Contrary to phone interviews, respondents who have been interviewed via survey may not have been in the host city at the time of the Olympics. While likely a minor issue, this would yield a lower-bound estimate.

the physically active during the pre-treatment period (2011), averaged by the interview date, shown separately for London and for Paris. As seen, while the share of the physically active has a slightly higher baseline level in London, both cities show similar changes over time, suggesting a common trend. Note that simple level differences are no major threat to identification, as we are primarily interested in *relative changes* during a closely spaced observation period. Appendix Figure A2 then tests more formally for a common trend by plotting the coefficients from a regression of being physically active on our treatment dummy (being located in London in 2012) interacted with the calendar week in 2011. As seen, there are no meaningful differences in physical activity between London and Paris in the pre-treatment period.

To our knowledge, no confounding events other than the Olympics occurred in either London or Paris that could explain changes in health behaviours.[12] Importantly, the use of DiD models with sharp cut-off dates around the event (i.e. Equation 3) allows us to relax this unconfoundedness assumption by being able to directly attribute changes in health behaviours to the Olympics.

3 Results

3.1 Main Results

Graphical Evidence. First, we look at graphical evidence from our raw data without manipulation. Our binary outcome is *Physical Activity*. We focus on respondents who were inactive before the Olympics. Once activated, we refer to them as *Newly Physically Active*. That is, we stratify on not being physically active pre-treatment (i.e. in 2011), so that any change can be interpreted as respondents moving from inactivity to activity.

Figure 1 plots the share of *Newly Physically Active* in 2012 during our observation period from mid July to early September, averaged by randomised interview dates and shown separately for London and Paris. We split 2012 into three sub-periods: the pre-Olympics (20 – 26 July), Olympics (27 July – 12 August), and post-Olympics period (13 August – 2 September, 2012). Overlaying the daily raw means in each subperiod are non-parametric (Epanechnikow-kernel-weighted) local quadratic polynomials.

[12]The 2012 Tour de France finished one week before the opening ceremony of the 2012 Olympics.

Figure 1 about here

We see a stronger increase in the share of *Newly Physically Active* in London relative to Paris during the Olympics period, whereby the strongest relative increase is seen around the opening and closing ceremonies, which are the two events that attract the largest audiences. After the Olympics, the relative increase among Londoners decays slowly. In contrast, the share of *Newly Physically Active* remains essentially flat in Paris during both the Olympics and the post-Olympics period.

Figure 2 about here

Figure 2 now exploits the panel structure and plots *within-individual changes* in the share of *Newly Physically Active* between 2011 and 2012, averaged by randomised interview dates and separately for London and for Paris; as the values for Paris are subtracted from those for London, we see net changes. This figure is essentially a non-parametric, graphical representation of our DiD model using sharp cut-off dates around the event (Equation (3)).

In line with Figure 1, we find a stronger increase in the share of *Newly Physically Active* in London relative to Paris during the Olympics, whereby the strongest relative increase is again found around the opening and closing events. We then observe a gradual return to (almost) baseline during the post-Olympics period, suggesting a temporary effect.

Regression Results. Table 2 shows the results from our second and third models in Equations (2) and (3), for the previously inactive (first column of each model) and for all respondents on average (second column). Our binary outcome is *Physical Activity*. We are thus estimating linear probability models.[13] For the previously inactive, we are again stratifying on not being physically active in the pre-treatment period (i.e. in 2011), so that any change can be interpreted as respondents moving from inactivity to activity.

Table 2 about here

Model 2 Column 1 suggests an effect size of about six percentage points. This is our preferred estimate, as it is conservative and takes all events related to the Olympics into account, including those happening before and after the opening and closing ceremonies. In terms

[13] Appendix Table A2 shows that our results are robust to omitting controls, Table A3 shows the full set of controls.

of effect size, about a third (34%) of Londoners were inactive in the pre-treatment period, yielding an increase in the share of *Newly Physically Active* by about 18%.

Model 3 Column 3 employs precise cut-off dates of the pre-Olympics, Olympics, and post-Olympics periods. The results show an effect size of roughly eight percentage points during the Olympics, with a marked reduction during the post-Olympics period. Assuming a linear decay, this suggests a return to baseline within three months (about 100 days). We look at legacy in more detail in Section 3.3.

Appendix Table A4 shows that there is no crowding out among the previously active, rather the opposite: the previously active in London are, on average, significantly more likely to remain physically active during the Olympics relative to the previously active in Paris, though the effect size is rather small.

When studying all residents pooled together, not just the inactive, we find precisely estimated null effects (Columns 2 and 4). The Olympics do not encourage the *average* respondent to become more active, but those who were previously inactive and presumably already at the margin to becoming active at least somewhat in the short-term. [14]

Potential Mechanism. We find that the increase among those Londoners who were previously inactive is strongest during the main event, i.e. between the opening and closing ceremonies, when infrastructure and information were largely fixed. In Appendix Table A6, we study the role of national sporting success by re-estimating Equation (2) and including an interaction between our treatment dummy and the number of gold medals (current or lagged, in levels or cumulative) won by the respondent's country. As seen, we do not detect any additional impact of national sporting success.[15] In Appendix Table A7, we replicate and extend our analysis in Table 2 for *thoughts about health* and *satisfaction with health* as exploratory outcomes. We find that the *Newly Physically Active* in London are significantly more likely to think about their health (and are also significantly more satisfied with it) during the event than Parisians. Although our research design does not allow us to conclusively prove mechanisms, together, this suggests that the act of hosting, and potentially the salience

[14] In an exploratory analysis, we looked at heterogeneity by age, splitting our estimation sample by the mean age (i.e. 45 years). In line with our previous results, we did not find heterogeneous treatment effects for residents on average, but stronger effects for younger residents during the main event (Appendix Table A5). For both age groups, we were again unable to detect long-term impacts (results available upon request).

[15] We do not find any additional impact regardless of whether we measure medals as gold medals or medals in total. The results are available upon request.

of the Olympics or physical activity more generally, rather than national sporting success is a likely mechanism behind our results.

We now study the intensive exercising margin and ask: how active have the previously inactive become? Table 3 re-estimates our preferred specification by using different intensities of physical activity as outcomes: Column 1 uses a binary indicator for > 5 *Times a Week* of physical activity, and zero else; Column 2 a binary indicator for > 3 *Times a Week*, Column 3 for *At Least Weekly*, and Column 4 for *At Least Monthly*.[16]

Table 3 about here

Unsurprisingly, we find a declining gradient in the intensive margin: most respondents who have become active engage in sport or physical activity at least weekly (Column 3), some even several times a week (Column 2). However, only a small share moved from 'no' into 'daily' activity. The effect size of two percentage points in Column 1 translates into an increase of about six percent from a baseline of 34%.[17]

Drinking and Smoking. We also study other health behaviours. Table 4 shows the results from our three models in Equations (1) to (3) using drinking and smoking as outcomes. These were originally recorded as the number of glasses of alcoholic drinks and the number of cigarettes on the previous day. We standardise them to have a mean of zero and a standard deviation of one (i.e. z-scores) for ease of interpretation.[18]

Table 4 about here

Our first model (Columns 1 and 2) suggests that drinking and smoking decreased during the Olympics period by about 0.09 and 0.08 standard deviations, respectively. However, only smoking turns out significant at the 10% level. Our second model also shows negative effects on drinking and smoking, but effect sizes are much attenuated and insignificant. Finally, our third model shows that drinking and smoking decreased by about 0.06 and 0.02 standard

[16]Note that the different intensities are not mutually exclusive: respondents who report to be physically active > 5 *Times a Week*, by definition, are also physically active *At Least Weekly*, for example.

[17]Instead of a linear model, we have also estimated a probit model, though with random as opposed to fixed effects (due to the incidental parameters problem). Given the quasi-random nature of our empirical setting, fixed effects should, in principle, not matter. As expected, the results from our panel data random-effects probit model almost perfectly corroborate those from our baseline specification (see Appendix Table A14).

[18]We obtain similar results when using binary indicators as outcomes, taking a threshold of, respectively, more than one alcoholic drink and more than five cigarettes on the previous day, as is often done in the related literature (results available upon request).

deviations, respectively. These effects are significant at the 5% level during the Olympics period yet turn insignificant during the post-Olympics period.[19]

Taken together, there is evidence for positive spillover effects on other risky health behaviours, namely alcohol and tobacco consumption. However, these are small in size and temporarily limited to the main event. Moreover, they do not withstand a correction for multiple hypotheses testing, as we will show below, which is why they should be interpreted as suggestive only.[20]

3.2 Robustness Checks

Extended Controls. So far, we have omitted potentially endogenous controls that may be influenced by hosting the Olympics. In particular, hosting the event may raise opportunities for local employment and hence income in London relative to Paris via multiplier effects, for example via an increase in local spending through tourism. This, in turn, may increase working hours and, thereby, reduce time that could otherwise be spent on sport or physical activity. At the same time, there may be additional confounders, in particular differences in outdoor weather conditions between London and Paris that may result in differences in the extent to which people engage in these activities.

Appendix Table A9 Column 1 adds additional controls for employment and household income from our surveys.[21] Column 2 adds additional controls for daily average temperature and precipitation in London and Paris during our observation period, obtained from the National Oceanic and Atmospheric Administration (NOAA) in the US.[22]
As seen, adding these additional economic and meteorological controls has negligible impacts on our coefficients. Note that the literature by and large does not identify economic multiplier effects by hosting mega sports events (cf. Baade and Matheson, 2016).

Selection Into Follow-Up Survey. A potential concern could be that respondents who answer our survey in London in 2012 may be different from those who answer it in Paris

[19] Appendix Table A8 shows that our results are robust to removing outliers (defined as more than 2.5 standard deviations below or above the mean); effect sizes are moderated somewhat downwards.

[20] Recall that there may be measurement errors in these outcomes due to imperfect recall of what happened on the previous day, which may attenuate our estimates.

[21] The question about annual gross household income asks about income in bands. We take the midpoint of each band and log it.

[22] The weather data can be downloaded from `https://www.ncei.noaa.gov/`. We chose the weather stations of London-Heathrow and Paris-Orly airports.

(selective attrition, leading to selection into the follow-up survey), and in particular, that this difference may be driving our results rather than the Olympics. As our preferred specification includes individual fixed effects, selection into the follow-up survey is a concern only in case that it is correlated with health behaviours.

Appendix Table A9 checks the robustness of our results regarding selective attrition in two ways. Column 3 weights observations by the inverse probability of taking part in our 2012 survey; the probability is obtained from predicting the presence of a 2012 interview from observables in an auxiliary regression. Column 4 then matches respondents in our 2011 and 2012 surveys in London and Paris. More specifically, to make respondents more similar in each city and to control for potential selection into the follow-up survey, we first match respondents using one-to-one nearest-neighbour matching without replacement based on all observables at our disposal. That is, a person in London is matched to his or her closest "neighbour" in Paris, based on what we can observe about them. We then consider only matched respondents (statistical twins) in each city in our sample (to compare only like-with-like) and re-run our models. N drops accordingly.

Using inverse probability weights yields very similar coefficients (Column 3), while matching yields coefficients that are slightly larger in size (Column 4). As another robustness check, we regressed the likelihood of being interviewed in our 2012 survey on different intensities of physical activity in our 2011 survey: we did not find that different intensities of physical activity are a significant predictor of panel attrition (results available upon request). Taken together, this suggests that selection into the follow-up survey is, if anything, only a minor concern.

Another concern may be that people who are less likely to be physically inactive (e.g. who are older, or who have lower income) may be more likely to be interviewed before the Olympics in London but not in Paris. To check this, we looked at differences in observable characteristics between Londoners and Parisians in 2012, by calculating scale-free normalised differences for the entire summer period (Appendix Table A10) and for the pre-Olympics, Olympics, or post-Olympics periods (Table A11). Recall that a normalised differences greater than 0.25 suggests covariate imbalance (Imbens and Wooldridge, 2009; Imbens and Rubin, 2015). We find little evidence for meaningful differences.

Placebo and Confirmation Tests. Next, we conduct placebo and confirmation tests to check that we indeed pick up the Olympics (as opposed to a confounding event). Column 5 of Appendix Table A9 uses *thoughts about finances* as a placebo outcome.[23] Hosting the Olympics should not have any impact on this outcome. Conversely, Column 6 uses *national pride* as a confirmation outcome.[24] Hosting the Olympics should have a positive impact on this outcome.

As expected, we do not find that hosting has a significant impact on thoughts about finances of previously inactive Londoners (Column 5), our placebo outcome. This also suggests that our results are unlikely to be caused by different economic developments between London and Paris. On the contrary, hosting has a significant, positive, and sizeable impact (about 0.18 standard deviations) on national pride among previously inactive Londoners (Column 6), our confirmation outcome.

When estimating, instead of our second model in Equation 2 our first model in Equation 1, we find the increase in national pride to be strongest (about 0.28 standard deviations) during the Olympics period, after which it decreases to about 0.20 standard deviations during the post-Olympics period (Column 7), in line with the return to baseline for physical activity. These patterns are not detectable when estimating the same model as a placebo during the pre-treatment time period, i.e. the year 2011 (Column 8).[25] Taken together, these tests using alternative outcomes and time periods strongly suggest that we do indeed identify the effect of the Olympics and not some confounding event.

Further Robustness. We conduct a range of further robustness checks. In Appendix Table A12, we check our results for robustness to non-linear models, showing that logit models yield very similar results. In Appendix Table A13, we use wild-cluster bootstrapped standard errors clustered at the city level (the level of quasi-randomisation) to account for the small number of clusters (Cameron et al., 2008), showing that our results for physical activity as our primary

[23]This binary indicator comes from the Likert scale question: "How often did you think about your finances yesterday?". Answer possibilities include 1 ("Not at all"), 2 ("A few times"), 3 ("Many times"), and 4 ("Continually"). We dichotomise this item such that category 1 is one and 2 to 4 are zero. The results are robust to alternative ways of dichotomising (available upon request).

[24]This outcome comes from the Likert scale question: "Overall, how proud are you to be (British/French)?". Answer possibilities range from 0 ("Not at all proud") to 10 ("Very proud"). We standardise this item to have a mean of zero and a standard deviation of one (i.e. a z-score) for ease of interpretation. The results are robust to using this item in natural units (available upon request).

[25]Note that, due to a shorter observation period in the pre-treatment period, we cannot estimate separate effects for the placebo Olympics and post-Olympics periods.

outcome become, if anything, more significant.

Appendix Section B looks at unobservable selection and coefficient stability in line with the bounding argument by Oster (2019), showing that selection on unobservables and potentially resulting omitted variable bias is, if anything, only a minor concern. Appendix Section C looks at multiple hypotheses testing, showing that stepdown-adjusted P-values still yield statistical significance at conventional levels for all of our estimates for physical activity, yet not for those for drinking and smoking, in neither of our three models. Finally, Appendix Section D shows that our results remain similar when modifying our preferred specification in Equation (2) to a difference-in-differences-in-differences (DDD) design to account for potentially time-varying confounders that may differ between both cities during the treatment period (i.e. the summer of 2012).

3.3 Legacy

The bid from London 2012 mentioned legacy effects prominently as part of its vision, suggesting that hosting would leave *"a legacy for sport in Britain"* (p.17) and that *"the UK will build on the sporting momentum of the expected successes coming out of the 2012 Games"* (p.19) (London Bid, 2012). Do mega sports events like the Olympics promote lasting health behaviour change?

To answer this question, we exploit the third wave of our panel data. In particular, we extend our preferred specification in Equation (2) by adding another interaction term, between *London* and 2013, to test for long-term effects. Our outcome is, as before, a binary indicator that is equal to one if an individual reports to be physically active (i.e. engaged in sport or physical activity at least monthly as opposed to less often or never) and zero else. To account for attrition, we estimate three models: Table 5 Column 1 does not adjust for attrition and just takes the data as given, Column 2 enforces a balanced panel (i.e. respondents have to be present in all three waves), and Column 3 weights observations by the inverse probability of taking part in the third wave; the probability is obtained from predicting the presence of an interview in the third wave from observables in an auxiliary regression. As before, our focus is on those respondents who were physically inactive prior to the Olympics, for whom we have identified short-term effects.

Table 5 about here

As seen, while we continue to identify short-term effects on physical activity for the previously inactive between five (Column 1) and seven percentage points (Column 2), we do not find evidence for long-term effects in any of our models. This is in line with the results from Column 3 of Table 2, where we find a reduction in physical activity during the post-Olympics period, suggesting a return to baseline within three months (about 100 days). Clearly, one year after the event, we find no evidence for lasting health behaviour change.

4 Cost-Benefit Analysis

We find that hosting London 2012 had a significant, positive effect on physical activity, yet only among those who were physically inactive prior to the event and only for about three months. To put this effect into perspective, we monetise it by looking at potential healthcare cost savings as one (closely related and clearly defined) category of benefits from increased physical activity. We then compare the monetised benefits with the Olympic budget aimed at increasing participation in physical activity among the general public. Our calculations can be found in Appendix Section E.

We arrive at healthcare cost savings of about £4.2 million from hosting London 2012, which stand against costs of £40 million, dedicated to increasing participation in grassroots sports, sports competitions, and physical activity among the general public, UK-wide (Girginov, 2016). This yields net benefits of about £40 − 4.2 = –35.8 million.

This estimate, however, may be too conservative: it is likely that hosting London 2012 also had a positive effect on physical activity outside of London, among those who were physically inactive prior to the event. We calculate that, to break even, the impact of London 2012 on residents in the rest of the UK would have had to be more than four times as strong as its impact on Londoners, which amounts to mobilising almost *all* of the previously inactive in the rest of the UK. Such a spillover does not seem realistic. At least when looking at potential healthcare cost savings, hosting London 2012 does not seem to have been worth it.

5 Discussion and Conclusion

This paper studied the short and long-term effects of hosting the Olympics on the health behaviours of host-city residents. Improvements in health have become a major selling point of proponents as economic study after economic study has failed to put forward meaningful economic benefits. In particular, we studied the impacts of hosting the Olympics on physical activity, as well as alcohol and tobacco consumption.

Overall, our findings show significant increases in physical activity among those who were previously inactive, by six percentage points, from a baseline of 34%, which is a large effect in percent (18%). Given that obesity, overweight, and health issues in general are concentrated among entirely inactive individuals, this group is arguably a priority for health economists and policy-makers (cf. Myers, 2008). London 2012 seems to have played an activating role for this group. Hence, the Olympics are an intervention capable of motivating *some* people to improve their health behaviours, most likely those who are already at the margin of improving them. This is consistent with experimental evidence in economics, for example Charness and Gneezy (2009), who show that the effect of small monetary incentives on physical activity is entirely driven by people who were previously inactive.

However, we also found that activation only lasts for about 100 days, suggesting an absence of long-term, positive changes to health behaviours, again consistent with evidence in economics (Royer et al., 2015; Carrera et al., 2018, 2019; Milkman et al., 2021; Butera et al., 2022; Marcus et al., 2022). We detected no crowding out of "good" behaviours among the previously active. Using our own, uniquely designed and collected panel data from London and Paris between 2011 and 2013 and DiD models with sharp cut-off dates, we showed that effects are strongest during the main event, i.e. between the opening and closing ceremonies, when infrastructure and information were largely fixed. We did not find an additional impact of medals won, but that Londoners who were previously inactive were more likely to think about their health during the event. This suggests that the act of hosting, and potentially the salience of the Olympics or physical activity more generally, rather than national sporting success is a likely mechanism behind our results.

Finally, we found suggestive evidence for reduced alcohol and tobacco consumption. However, these effects are relatively small in size, are temporally limited to the duration of the main event, and, notably, do not withstand a correction for multiple hypotheses testing. They should, therefore, be interpreted with caution. Whilst standard in the literature, our outcomes for drinking and smoking are based on self-reported quantities on the previous day. They hence may be subject to attenuation and other biases. How to systematically, and precisely, capture these outcomes using observed behaviours is an important avenue for future research.

A notable limitation of our study is our focus on London 2012, and that our research design relies on the comparison of only one treatment with one control city. This raises questions of the external validity and generalisability of our findings. While this should be caveated, it should also be noted that the Olympics are an infrequent event that takes place only every four years. The lack of comparable international micro data makes studying our research question inherently difficult. At the same time, we expect our findings to provide valuable insights for host cities and hosting concepts that are comparable to London 2012, such as that of Paris 2024 or Los Angeles 2028.

Besides promoting physical activity and an active life style, hosting London 2012 followed other objectives such as national as well as economic regeneration. Yet, improvements in health behaviours were one important argument to justify parts of the expenses of hosting. Although London 2012 achieved the objective of improving health behaviours partly, our back-of-the-envelope cost-benefit analysis suggests that staging mega sports events like the Olympics is not a cost-effective policy to promote lasting health behaviour change, at least as they are currently implemented. Future research and project planning should thus focus on more specific outreach and on elements to make the intervention "stick" in the long-run.

References

Aggarwal, S., R. Dizon-Ross, and A. D. Zucker (2020). Incentivizing Behavioral Change: The Role of Time Preferences. NBER Working Paper 27079, National Bureau of Economic Research.

Akee, R., E. Simeonova, W. Copeland, A. Angold, and E. J. Costello (2013). Young Adult Obesity and Household Income: Effects of Unconditional Cash Transfers. *American Economic Journal: Applied Economics 5*(2), 1–28.

Anderson, M. L. and D. A. Matsa (2011). Are Restaurants Really Supersizing America? *American Economic Journal: Applied Economics 3*(1), 152–188.

Angrist, J. D. and J.-S. Pischke (2009). *Mostly Harmless Econometrics*. Princeton: Princeton University Press.

Augurzky, B., T. K. Bauer, A. R. Reichert, C. M. Schmidt, and H. Tauchmann (2018). Habit formation, obesityand cash rewards. Ruhr Economic Papers 750.

Baade, R. A. and V. R. Matheson (2016). Going for the Gold: The Economics of the Olympics. *Journal of Economic Perspectives 2*(30), 201–218.

Barron, J. M., B. T. Ewing, and G. R. Waddell (2000). The Effects of High School Athletic Participation on Education and Labor Market Outcomes. *Review of Economics and Statistics 82*(3), 409–421.

Batty, G. D. (2002). Physical activity and coronary heart disease in older adults: A systematic review of epidemiological studies. *European Journal of Public Health 12*(3), 171–176.

Baumann, A. E., M. Kamada, R. S. Seis, R. P. Troiano, D. Ding, K. Milton, N. Murphy, and P. C. Hallal (2021). An evidence-based assessment of the impact of the Olympic Games on population levels of physical activity. *The Lancet 398*(10298), 456–464.

Belot, M., N. Berlin, J. James, and V. Skafida (2018). The Formation and Malleability of Dietary Habits: A Field Experiment with Low Income Families. IZA Discussion Paper, Institute for the Study of Labor (IZA).

Belot, M. and J. James (2011). Healthy school meals and educational outcomes. *Journal of Health Economics 30*(3), 489–504.

Butera, L., R. Metcalfe, W. Morrison, and D. Taubinsky (2022). Measuring the Welfare Effects of Shame and Pride. *American Economic Review 112*(1), 122–168.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008). Bootstrap-Based Improvements for Inference with Clustered Errors. *Review of Economics and Statistics 90*(3), 414–427.

Carrera, M., H. Royer, M. Stehr, and J. Sydnor (2018). Can financial incentives help people trying to establish new habits? Experimental evidence with new gym members. *Journal of Health Economics 58*, 202–214.

Carrera, M., H. Royer, M. Stehr, and J. Sydnor (2019). The Structure of Health Incentives: Evidence from a Field Experiment. *Management Science 66*(5), 1783–2290.

Cawley, J. (2004). The Impact of Obesity on Wages. *Journal of Human Resources 39*(2), 451–474.

Cawley, J., D. Frisvold, and C. Meyerhoefer (2013). The impact of physical education on obesity among elementary school children. *Journal of Health Economics 32*(4), 743–755.

Cawley, J., E. Han, J. Kim, and E. C. Norton (2019). Testing for family influences on obesity: The role of genetic nurture. *Health Economics 28*(7), 937–952.

Cawley, J. and C. Meyerhoefer (2012). The medical care costs of obesity: An instrumental variables approach. *Journal of Health Economics 31*(1), 219–230.

Cawley, J., A. Susskind, and B. Willage (2020). The Impact of Information Disclosure on Consumer Behavior: Evidence from a Randomized Field Experiment of Calorie Labels on Restaurant Menus. *Journal of Policy Analysis and Management 39*(4), 1020–1042.

Cawley, J., A. M. Susskind, and B. Willage (2021). Does Information Disclosure Improve Consumer Knowledge? Evidence from a Randomized Experiment of Restaurant Menu Calorie Labels. *American Journal of Health Economics 7*(4), 427–456.

Cawley, J. a. (2011). The Economics of Risky Health Behaviors. *Handbook of Health Economics 2*, 95–199.

Charness, G. and U. Gneezy (2009). Incentives to Exercise. *Econometrica 77*(3), 909–931.

Conti, G., J. J. Heckman, and R. Pinto (2016). The Effects of Two Influential Early Childhood Interventions on Health and Healthy Behaviour. *Economic Journal 126*(596), F28–F65.

Costa-Font, J. and N. Mas (2016). 'Globesity'? The effects of globalization on obesity and caloric intake. *Food Policy 64*, 121–132.

Council on Foreign Relations (2021). The Economics of Hosting the Olympic Games. https://www.cfr.org/backgrounder/economics-hosting-olympic-games.

Courtemanche, C. (2009). Rising cigarette prices and rising obesity: Coincidence or unintended consequence? *Journal of Health Economics 28*(4), 781–798.

Courtemanche, C., G. Heutel, and P. McAlvanah (2014). Impatience, Incentives and Obesity. *Economic Journal 125*(582), 1–31.

Cummins, S., C. Clark, D. Lewis, N. Smith, C. Thompson, M. Smuk, S. Stansfeld, S. Taylor, A. Fahy, T. Greenhalgh, and S. Eldridge (2018). The effects of the London 2012 Olympics and related urban regeneration on physical and mental health: the ORiEL mixed-methods evaluation of a natural experiment. *Public Health Research 6*(12).

Currie, J., S. DellaVigna, E. Moretti, and V. Pathania (2010). The Effect of Fast Food Restaurants on Obesity and Weight Gain. *American Economic Journal: Economic Policy 2*(3), 32–63.

DellaVigna, S. and U. Malmendier (2006). Paying Not to Go to the Gym. *American Economic Review 96*(3), 694–719.

Department for Culture, Media and Sport (2007). Our Promise for 2012: How the UK will benefit from the Olympic and Paralympic Games. https://assets.publishing.service.gov.uk/media/5a78a93140f0b62b22cbbc69/DCMSLeafletAdobev5andlaterTPL.pdf.

DeSimone, J., D. Grossman, and N. R. Ziebarth (2023). Regression Discontinuity Evidence on the Effectiveness of the Minimum Legal E-cigarette Purchasing Age. *American Journal of Health Economics 9*(3), 461–485.

Dimeo, F., M. Bauer, I. Varahram, and U. Halter (2001). Benefits from aerobic exercise in patients with major depression: a pilot study. *British Journal of Sports Medicine 35*, 114–117.

Dolan, P., G. Kavetsos, C. Krekel, D. Mavridis, R. Metcalfe, C. Senik, S. Szymanski, and N. R. Ziebarth (2019). Quantifying the intangible impact of the Olympics using subjective well-being data. *Journal of Public Economics 177*, 104043.

Dolan, P., G. Kavetsos, and I. Vlaev (2014). The Happiness Workout. *Social Indicators Research 119*, 1363–1377.

Dragone, D. and N. R. Ziebarth (2017). Non-separable time preferences, novelty consumption and body weight: Theory and evidence from the East German transition to capitalism. *Journal of Health Economics 51*, 41–65.

Dubois, P., R. Griffith, and A. Nevo (2014). Do Prices and Attributes Explain International Differences in Food Purchases? *American Economic Review 104*(3), 832–867.

Dunn, R. A. (2010). The Effect of Fast-Food Availability on Obesity: An Analysis by Gender, Raceand Residential Location. *American Journal of Agricultural Economics 92*(4), 1149–1164.

Fletcher, J. M., D. E. Frisvold, and N. Tefft (2010). The effects of soft drink taxes on child and adolescent consumption and weight outcomes. *Journal of Public Economics 94*(11–12), 967–974.

France 24 (2020). Paris 2024 budget to be 'studied' by end of year. https://www.france24.com/en/20200708-paris-2024-budget-to-be-studied-by-end-of-year.

Fujiwara, D., L. Kudrna, T. Cornwall, K. Laffan, and P. Dolan (2015). Further analysis to value the health and educational benefits of sport and culture. Technical report, Department for Culture, Media and Sport.

Girginov, V. (2016). Has the London 2012 Olympic Inspire programme inspired a generation? A realist view. *European Physical Education Review 22*(4), 490–505.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics 225*(2), 254–277.

Greater London Authority (2013). London 2012 Opinion Research. https://data.london.gov.uk/dataset/london-2012-opinion-research.

Griffith, R. (2022). Obesity, Poverty and Public Policy. *Economic Journal 132*(644), 1235–1258.

Griffith, R., R. Lluberas, and M. Lührmann (2015). Gluttony and Sloth? Calories, Labor Market Activity and the Rise of Obesity. *Journal of the European Economic Association 14*(6), 1253–1286.

Griffith, R., M. O'Connell, and K. Smith (2015). Relative prices, consumer preferencesand the demand for food. *Oxford Journal of Economic Policy 31*(1), 116–130.

Grossman, M., E. Tekin, and R. Wada (2014). Food prices and body fatness among youths. *Economics & Human Biology 12*, 4–19.

Homonoff, T., B. Willage, and A. Willén (2020). Rebates as incentives: The effects of a gym membership reimbursement program. *Journal of Health Economics 70*, 102285.

Humphreys, B. R., L. McLeod, and J. E. Ruseski (2014). Physical Activity and Health Outcomes: Evidence From Canada. *Health Economics 23*(1), 33–54.

Hyytinen, A. and J. Lahtonen (2013). The effect of physical activity on long-term income. *Social Science & Medicine 96*, 129–137.

Imbens, G. W. and D. B. Rubin (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences*. Cambridge: Cambridge University Press.

Imbens, G. W. and J. M. Wooldridge (2009). Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature 47*(1), 5–86.

IMF (2010). Is It Worth It? https://www.elibrary.imf.org/abstract/journals/022/0047/001/article-A004-en.xml, retrieved May 15, 2024.

IOC (2022). Beyond the Games. https://olympics.com/ioc/beyond-the-games.

Jeitschko, T. D. and R. A. Pecchenino (2006). Do You Want Fries with That? An Exploration of Serving Size, Social Welfare and Our Waistlines. *Economic Inquiry 44*(3), 442–450.

Kamhöfer, D. A., H. Schmitz, and M. Westphal (2019). Heterogeneity in Marginal Non-Monetary Returns to Higher Education. *Journal of the European Economic Association 17*(1), 205–244.

Kavetsos, G. (2011). The impact of physical activity on employment. *Journal of Socio-Economics 40*(6), 775–779.

Kavetsos, G. and S. Szymanski (2009). From the Olympics to the grassroots: What will London 2012 mean for sport funding and participation in Britain? *Public Policy Research 16*(3), 192–196.

Kavetsos, G. and S. Szymanski (2010). National well-being and international sports events. *Journal of Economic Psychology 31*(2), 158–171.

Kokolakakis, T., F. Lera-Lopez, and G. Ramchandani (2019). Did London 2012 deliver a sports participation legacy? *Sport Management Review 22*(2), 276–287.

Kokolakakis, T., F. Lera-Lopez, and G. Ramchandani (2020). Sport promotion through sport mega-events: An analysis for types of Olympic sports in London 2012. *International Journal of Environmental Research and Public Health 17*(17), 6193.

Lechner, M. (2009). Long-run labour market and health effects of individual sports activities. *Journal of Health Economics 28*(4), 839–854.

London Bid (2012). Volume 1 – Theme 1 Concept and Legacy. https://webarchive.nationalarchives.gov.uk/ukgwa/20080107221051mp_/http://www.london2012.com/documents/candidate-files/theme-1-olympic-games-concept-and-legacy.pdf.

Mahtani, K. R., J. Protheroe, S. P. Slight, M. M. Piva Demarzo, T. Blakeman, C. A. Barton, B. Brijnath, and N. Roberts (2013). Can the London 2012 Olympics 'inspire a generation' to do more physical or sporting activities? An overview of systematic reviews. *BMJ Open 3*(1), e3002058.

Marcus, J., T. Siedler, and N. R. Ziebarth (2022). The Long-Run Effects of Sports Club Vouchers for Primary School Children. *American Economic Journal: Economic Policy forthcoming*.

Mayor of London (2012). Millions enjoy Mayor's cultural events for 2012. https://www.london.gov.uk/press-releases-4614.

McCartney, G., S. Thomas, H. Thomson, J. Scott, V. Hamilton, P. Hanlon, and L. Bond (2010). The health and socioeconomic impacts of major multi-sport events: systematic review (1978-2008). *BMJ 340*, c2369.

Milkman, K. L., D. Gromet, H. Ho, J. S. Kay, T. W. Lee, P. Pandiloski, Y. Park, A. Rai, M. Bazerman, J. Beshears, L. Bonacorsi, C. Camerer, E. Chang, G. Chapman, R. Cialdini, H. Dai, L. Eskreis-Winkler, A. Fishbach, J. J. Gross, S. Horn, A. Hubbard, S. J. Jones, D. Karlan, T. Kautz, E. Kirgios, J. Klusowski, A. Kristal, R. Ladhania, G. Loewenstein, J. Ludwig, B. Mellers, S. Mullainathan, S. Saccardo, J. Spiess, G. Suri, J. Talloen, J. Taxer, Y. Trope, L. Ungar, K. G. Volpp, A. Whillans, J. Zinman, and A. L. Duckworth (2021). Megastudies improve the impact of applied behavioural science. *Nature 600*, 478–483.

Morris, S. (2006). Body mass index and occupational attainment. *Journal of Health Economics 25*(2), 347–364.

Morris, S. (2007). The impact of obesity on employment. *Labour Economics 14*(3), 413–433.

Muller, L., A. Lacroix, J. L. Lusk, and B. Ruffieux (2017). Distributional Impacts of Fat Taxes and Thin Subsidies. *Economic Journal 127*(604), 2066–2092.

Myers, J. (2008). The Health Benefits and Economics of Physical Activity. *Current Sports Medicine Reports 7*(6), 314–316.

Nakamura, R., A. Albanese, E. Coombes, and M. Suhrcke (2021). Do Economic Incentives Promote Physical Activity? Evidence from the London Congestion Charge. GLO Discussion Paper, Global Labor Organization (GLO).

National Audit Office (2012). The London 2012 Olympic Games and Paralympic Games: Post-Games review. https://www.nao.org.uk/report/the-london-2012-olympic-games-and-paralympic-games-post-games-review/.

Ofcom (2012). The London 2012 Olympic Games: media consumption. https://www.ofcom.org.uk/__data/assets/pdf_file/0019/55360/report.pdf.

Office for National Statistics (2012). Families and households in the UK: 2012. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/families/bulletins/familiesandhouseholds/2012-11-01#household-size.

Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business & Economic Statistics 37*(2), 187–204.

Paris 2024 Candidate City (2016). Official bid file of Paris for the Summer Olympic Games in 2024. https://library.olympics.com/Default/doc/SYRACUSE/171681/candidature-file-paris-candidate-city-olympic-games-2024.

Public Health England (2016). Health matters: getting every adult active every day. https://www.gov.uk/government/publications/health-matters-getting-every-adult-active-every-day/health-matters-getting-every-adult-active-every-day.

Puhl, R. M. and C. A. Heuer (2009). The Stigma of Obesity: A Review and Update. *Obesity 17*(5), 941–964.

Reichert, A. R. (2015). Obesity, Weight Lossand Employment Prospects: Evidence from a Randomized Trial. *Journal of Human Resources 50*(3), 759–810.

Rohde, K. I. M. and W. Verbeke (2017). We like to see you in the gym – A field experiment on financial incentives for short and long term gym attendance. *Journal of Economic Behavior & Organization 134*, 388–407.

Romano, J. P. and M. Wolf (2005a). Exact and Approximate Stepdown Methods for Multiple Hypothesis Testing. *Journal of the American Statistical Association 100*(469), 94–108.

Romano, J. P. and M. Wolf (2005b). Stepwise Multiple Testing as Formalized Data Snooping. *Econometrica 73*(4), 1237–1282.

Romano, J. P. and M. Wolf (2016). Efficient computation of adjusted p-values for resampling-based stepdown multiple testing. *Statistics & Probability Letters 113*, 38–40.

Rooth, D. (2011). Work out or out of work – The labor market return to physical fitness and leisure sports activities. *Labour Economics 18*(3), 399–409.

Royer, H., M. Stehr, and J. Sydnor (2015). Incentives, Commitments and Habit Formation in Exercise: Evidence from a Field Experiment with Workers at a Fortune-500 Company. *American Economic Journal: Applied Economics 7*(3), 51–84.

Sari, N. (2009). Physical inactivity and its impact on healthcare utilization. *Health Economics 18*(8), 885–901.

Sarma, S., G. S. Zaric, M. K. Campbell, and J. Gilliland (2014). The effect of physical activity on adult obesity: Evidence from the Canadian NPHS panel. *Economics & Human Biology 14*, 1–21.

Smith, N. R., D. L. Lewis, A. Fahy, C. Thompson, C. Clark, S. Stansfield, and S. Cummins (2014). Changes in physical activity in East London's adolescents following the 2012 Olympic Games: findings from the prospective Olympic Regeneration in East London (ORiEL) cohort study. *European Journal of Public Health 24*(S2), cku162–062.

Stamatakis, E., U. Ekelund, and N. J. Wareham (2007). Temporal trends in physical activity in England: The Health Survey for England 1991 to 2004. *Preventive Medicine 45*(6), 416–423.

Stockwell, T., S. Donath, M. Cooper-Stanbury, T. Chikritzhs, P. Catalano, and C. Mateo (2004). Under-reporting of alcohol consumption in household surveys: a comparison of quantity–frequency, graduated–frequency and recent recall. *Addiction 99*(8), 1024–1033.

Strulik, H. (2014). A mass phenomenon: The social evolution of obesity. *Journal of Health Economics 33*, 113–125.

Stutzer, A. and A. N. Meier (2016). Limited Self-control, Obesity and the Loss of Happiness. *Health Economics 25*(11), 1409–1424.

UK Chief Medical Officer (2019). UK Chief Medical Officer's Physical Activity Guidelines. https://assets.publishing.service.gov.uk/media/5d839543ed915d52428dc134/uk-chief-medical-officers-physical-activity-guidelines.pdf.

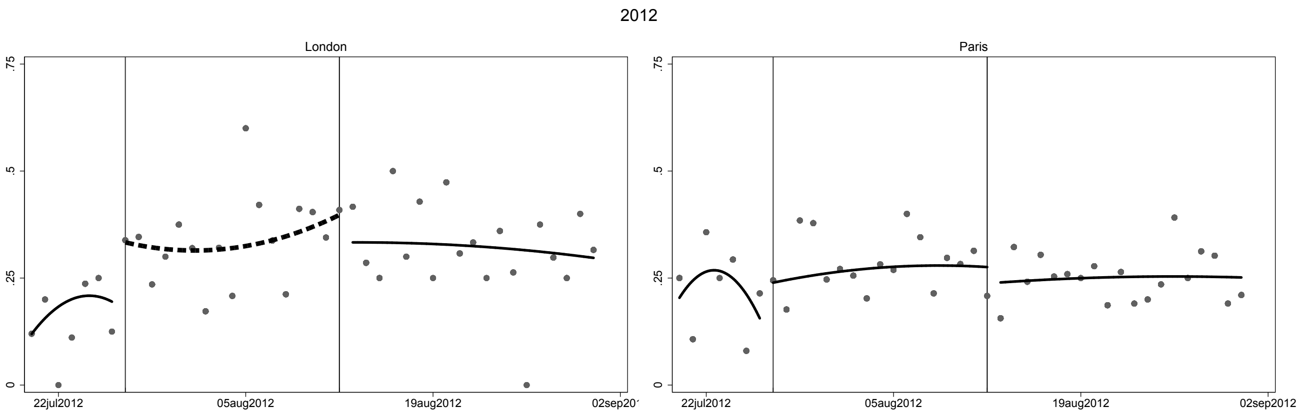
United 2026 (2017). Canada, Mexico, and the United States United Bid to Host the 2026 FIFA World Cup. https://digitalhub.fifa.com/m/3c077448dcd5c0ab/original/w3yjeu7dadt5erw26wmu-pdf.pdf.

Weed, M., E. Coren, J. Fiore, I. Wellard, L. Mansfield, D. Chatziefstathiou, and S. Dowse (2012). Developing a physical activity legacy from the London 2012 Olympic and Paralympic Games: A policy-led systematic review. *Perspectives in Public Health 132*(2), 75–80.

Wehby, G. L. and C. Courtemanche (2012). The heterogeneity of the cigarette price effect on body mass index. *Journal of Health Economics 31*(5), 719–729.

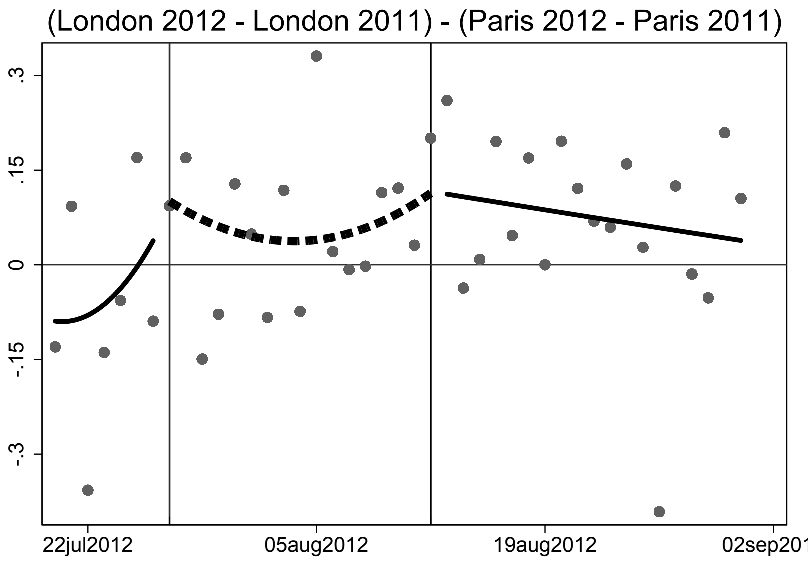
World Health Organization (2021). Fact Sheet 311: Obesity and Overweight. http://www.who.int/mediacentre/factsheets/fs311/en/.

Figure 1: Share of Newly Physically Active in London and Paris in Treatment Period (2012)



Notes: Scatter plot shows raw responses for *physical activity* (on a zero-to-one scale, whereby one denotes at least monthly and zero denotes less often or never) in 2012 of those who reported to be physically inactive in 2011, averaged by the interview date, shown separately for London and for Paris. The vertical lines are the opening (July 27) and closing (August 12, 2012) ceremonies. Each of the three periods in each graph shows non-parametric Epanechnikow-kernel-weighted local quadratic polynomial regressions of physical activity on interview date fixed effects.
Sources: Own data, own calculations.

Figure 2: Change in Share of Newly Physically Active Between London and Paris From 2011 to 2012



Notes: Scatter plot shows raw changes for *physical activity* (on a zero-to-one scale, whereby one denotes at least monthly and zero denotes less often or never) from 2011 to 2012 of those who reported to be physically inactive in London in 2011, averaged by interview date, minus the corresponding changes in Paris. The graph thus resembles the difference-in-differences model in Equations 2 and 3 (estimated without controls). The vertical lines are the opening (July 27) and closing (August 12, 2012) ceremonies. Each of the three periods in each graph shows non-parametric Epanechnikow-kernel-weighted local quadratic polynomial regressions of physical activity on interview date fixed effects.
Sources: Own data, own calculations.

Table 1: Summary Statistics

	Mean	Standard Deviation	Minimum	Maximum	Number of Observations
Outcomes					
Physical Activity	0.6585	0.4742	0	1	29,548
Drinking	0.9310	1.6652	0	25	29,606
Smoking	2.7324	6.4034	0	100	29,618
Covariates					
Age	44.5140	14.9991	16	97	29,548
Female	0.5570	0.4967	0	1	29,548
Male	0.4430	0.4967	0	1	29,548
Single	0.2985	0.4576	0	1	29,549
Married	0.3940	0.4886	0	1	29,550
Partnered	0.1770	0.3817	0	1	29,551
Separated	0.0213	0.1442	0	1	29,552
Divorced	0.0799	0.2711	0	1	29,553
Widowed	0.0293	0.1687	0	1	29,554
Graduate	0.9406	0.2364	0	1	29,555
Studies	0.0594	0.2364	0	1	29,556
Secondary School Degree	0.2232	0.4164	0	1	29,557
Vocational Degree	0.1249	0.3307	0	1	29,558
Undergraduate Degree	0.3078	0.4616	0	1	29,559
Graduate Degree	0.3440	0.4751	0	1	29,560
Employed Full-Time	0.5098	0.4999	0	1	29,561
Employed Part-Time	0.0936	0.2913	0	1	29,562
Self-Employed	0.0626	0.2422	0	1	29,563
Unemployed and Looking	0.0495	0.2169	0	1	29,564
Unemployed and Not Looking	0.0601	0.2377	0	1	29,565
Retired	0.1649	0.3711	0	1	29,566
Log Income	10.3716	0.7285	8.2940	11.5129	29,567
Owns Dwelling	0.5488	0.4976	0	1	29,568
Rents or Shares Dwelling	0.3717	0.4833	0	1	29,569
Lives with Relatives	0.0614	0.2401	0	1	29,570
Other Tenancy Type	0.0147	0.1203	0	1	29,571

Notes: We provide summary statistics for our estimation sample for *all residents* on average (Table 3 Column 4), for London and for Paris pooled. The question on income asks about "total household pre-tax annual income from all sources" in 5,000 bands, i.e. "Under 5,000", "5,000 - 9,999", "10,000 - 14,999", … "45,000 - 54,999", "55,000 - 99,999", and "100,000 or more". It asks about GBP for Londoners and EUR for Parisians. We convert EUR into GBP using the exchange rate at the time, then take the midpoint of each band (leaving it at 100,000 for "100,000 or more"), and finally convert it into log. See Section 2.1 for a detailed description of our data.
Sources: Own data, own calculations.

Table 2: Impact of Hosting the Olympics on Physical Activity (Extensive Margin)

	Physical Activity (1: At Least Monthly = Being Active, 0: Less Often or Never = Being Inactive)			
	Model 2, 2011-2012		**Model 3, 2011-2012**	
	(1)	(2)	(3)	(4)
	Previously Inactive	All Residents	Previously Inactive	All Residents
London x Olympics			0.0794***	0.0009
			(0.0186)	(0.0078)
London x PostOlympics			0.0570***	-0.0074
			(0.0179)	(0.0090)
Olympics			0.2710***	0.0121
			(0.0231)	(0.0091)
PostOlympics			0.2470***	0.0172*
			(0.0256)	(0.0092)
London x 2012	0.0559***	-0.0053		
	(0.0151)	(0.0060)		
2012	0.2630***	0.0156*		
	(0.0238)	(0.0080)		
N	10,082	29,548	10,082	29,548
Individuals	6,400	19,046	6,400	19,046
Controls	Yes	Yes	Yes	Yes
Day-of-Week FE	Yes	Yes	Yes	Yes
Calendar-Month FE	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
R Squared	0.283	0.002	0.286	0.002

Notes: Our outcome is a binary indicator that indicates whether a respondent reports to be *physically active* (i.e. engaged in sport or physical activity at least monthly as opposed to less often or never). We estimate linear probability models separately for the *previously inactive* (i.e. respondents who were not physically active in the pre-treatment year 2011) and for *all residents* on average. Treatment is defined as being resident in *London* as opposed to Paris. Model 2, which is our preferred specification, is based on Equation 2 and estimates a difference-in-differences model in the years 2011 and 2012. It takes the entire year 2012 as the relevant treatment period. The baseline category is the pre-treatment year 2011. Model 3 is based on Equation 3 and estimates a difference-in-differences model in the years 2011 and 2012. It again takes the entire year 2012 as the relevant treatment period but splits it into a pre-Olympics (suppressed for brevity), Olympics, and post-Olympics period. The baseline category is the pre-treatment year 2011. Models 2 and 3 look at within-individual changes between Londoners and Parisians between 2011 and 2012. See Section 2.1 for a detailed description of our data and Section 2.2 for our models. Robust standard errors clustered at interview date level in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Sources: Own data, own calculations.

Table 3: Impact of Hosting the Olympics on Physical Activity (Intensive Margin)

	Physical Activity (1: Yes, 0: No) **Model 2, 2011-2012**			
	(1) > 5 Times a Week	(2) > 3 Times a Week	(3) At Least Weekly	(4) At Least Monthly
Panel A: Previously Inactive				
London x 2012	0.0158***	0.0410***	0.0634***	0.0559***
	(0.0056)	(0.0073)	(0.0129)	(0.0151)
2012	0.0159**	0.0377***	0.1590***	0.2630***
	(0.0072)	(0.0136)	(0.0204)	(0.0238)
N	10,082	10,082	10,082	10,082
Individuals	6,400	6,400	6,400	6,400
R Squared	0.038	0.073	0.178	0.283
Panel B: All Residents				
London x 2012	0.0083**	-0.0063	-0.0058	-0.0053
	(0.0039)	(0.0068)	(0.0072)	(0.0060)
2012	-0.0144***	0.0045	0.0191**	0.0156*
	(0.0043)	(0.0058)	(0.0082)	(0.0080)
N	29,548	29,548	29,548	29,548
Individuals	19,046	19,046	19,046	19,046
R Squared	0.003	0.002	0.002	0.002
Controls	Yes	Yes	Yes	Yes
Day-of-Week FE	Yes	Yes	Yes	Yes
Calendar-Month FE	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes

Notes: Our outcomes are binary indicators for different intensities of physical activity, including > 5 *Times a Week*, > 3 *Times a Week*, *At Least Weekly*, and *At Least Monthly*. These different intensities are not mutually exclusive: respondents who report to be physically active > 5 *Times a Week*, by definition, are also physically active *At Least Weekly*, for example. We estimate linear probability models separately for the *previously inactive* (i.e. respondents who were not physically active in the pre-treatment year 2011, Panel A) and for *all residents* on average (Panel B). Treatment is defined as being resident in *London* as opposed to Paris. Model 2, which is our preferred specification, is based on Equation 2 and estimates a difference-in-differences model in the years 2011 and 2012. It takes the entire year 2012 as the relevant treatment period. The baseline category is the pre-treatment year 2011. Model 2 looks at within-individual changes between Londoners and Parisians between 2011 and 2012. See Section 2.1 for a detailed description of our data and Section 2.2 for our models. Robust standard errors clustered at interview date level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Sources: Own data, own calculations.

Table 4: Impact of Hosting the Olympics on Other Health Behaviours

	Other Health Behaviours (z-Scores)					
	Model 1, 2012		**Model 2, 2011-2012**		**Model 3, 2011-2012**	
	(1)	(2)	(3)	(4)	(5)	(6)
	Drinking	Smoking	Drinking	Smoking	Drinking	Smoking
London x Olympics	-0.0855	-0.0813*			-0.0634**	-0.0196**
	(0.0789)	(0.0421)			(0.0266)	(0.0098)
London x PostOlympics	0.0231	-0.0430			-0.0075	0.0021
	(0.0798)	(0.0360)			(0.0226)	(0.0103)
Olympics	0.0025	0.0883***			-0.0063	-0.0170
	(0.0427)	(0.0287)			(0.0184)	(0.0107)
PostOlympics	-0.0282	0.1170***			0.0021	-0.0386***
	(0.0424)	(0.0277)			(0.0154)	(0.0079)
London	0.1480**	-0.0577*				
	(0.0723)	(0.0299)				
London x 2012			-0.0245	-0.0049		
			(0.0193)	(0.0076)		
2012			-0.0024	-0.0307***		
			(0.0147)	(0.0078)		
N	10,591	10,599	29,606	29,618	29,606	29,618
Individuals	10,591	10,599	19,071	19,072	19,071	19,072
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Day-of-Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Calendar-Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	No	No	Yes	Yes	Yes	Yes
R Squared	0.058	0.051	0.028	0.006	0.028	0.006

Notes: Our outcomes are the number of alcoholic beverages drunk and the number of cigarettes smoked on the previous day, respectively, standardised to have a mean of zero and a standard deviation of one (i.e. z-scores). We estimate linear probability models for *all residents* on average. Treatment is defined as being resident in *London* as opposed to Paris. Model 1 is based on Equation 1 and estimates a difference-in-differences model in the year 2012 only. The baseline category is the pre-Olympics period in the same year. Model 2, which is our preferred specification, is based on Equation 2 and estimates a difference-in-differences model in the years 2011 and 2012. It takes the entire year 2012 as the relevant treatment period. The baseline category is the pre-treatment year 2011. Model 3 is based on Equation 3 and estimates a difference-in-differences model in the years 2011 and 2012. It again takes the entire year 2012 as the relevant treatment period but splits it into a pre-Olympics (suppressed), Olympics, and post-Olympics period. The baseline category is again the pre-treatment year 2011. Models 2 and 3 look at within-individual changes between Londoners and Parisians between 2011 and 2012, Model 1 looks at between-individual changes in 2012. See Section 2.1 for a detailed description of our data and Section 2.2 for our models. Robust standard errors clustered at interview date level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Sources: Own data, own calculations.

Table 5: Legacy Effects

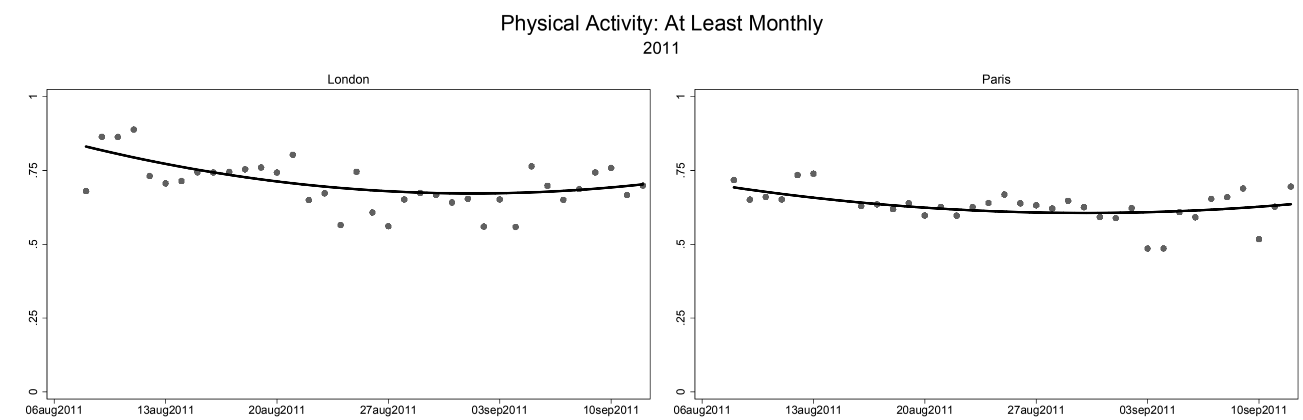
	Physical Activity by Previously Inactive (1: Yes, 0: No) Model 2, 2011-2013		
	(1) Unadjusted	(2) Balanced Panel	(3) IPW
London x 2012	0.0533***	0.0697***	0.0395**
	(0.0158)	(0.0198)	(0.0166)
2012	0.2750***	0.2920***	0.2960***
	(0.0104)	(0.0199)	(0.0121)
London x 2013	0.0134	0.0249	0.0159
	(0.0194)	(0.0206)	(0.0211)
2013	0.3190***	0.3520***	0.3360***
	(0.0113)	(0.0306)	(0.0137)
N	12,334	6,653	12,334
Individuals	6,402	2,249	6,402
Controls	Yes	Yes	Yes
Day-of-Week FE	Yes	Yes	Yes
Calendar-Month FE	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes
R Squared	0.226	0.207	0.245

Notes: Our outcome is a binary indicator that indicates whether a respondent reports to be *physically active* (i.e. engaged in sport or physical activity at least monthly as opposed to less often or never). We estimate linear probability models, restricting our estimation sample to the *previously inactive* (i.e. respondents who were not physically active in the pre-treatment year 2011). Treatment is defined as being resident in *London* as opposed to Paris. Model 2, which is our preferred specification, is based on Equation 2 and estimates a difference-in-differences model in the years 2011, 2012 (i.e. short-term effects), and 2013 (i.e. long-term effects). It takes the entire years 2012 and 2013 as the relevant treatment periods. The baseline category is the pre-treatment year 2011. Model 2 looks at within-individual changes between Londoners and Parisians from 2011 to 2012 to 2013. Column 1 is unadjusted, whereas Column 2 uses a balanced panel, and Column 3 weights observations by the inverse probability of taking part in our 2013 survey. See Section 2.1 for a detailed description of our data and Section 2.2 for our models. Robust standard errors clustered at interview date level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Sources: Own data, own calculations.

Appendix

A Additional Figures & Tables

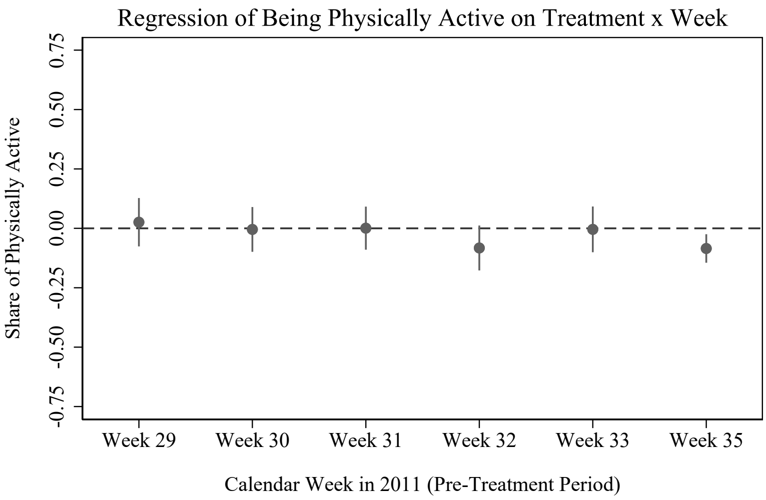
Figure A1: Share of Physically Active in London and Paris in Pre-Treatment Period (2011)



Notes: Scatter plot shows raw responses for *physical activity* (on a zero-to-one scale, whereby one denotes at least monthly and zero denotes less often or never) in 2011, averaged by the interview date, shown separately for London and for Paris. Each graph shows non-parametric Epanechnikow-kernel-weighted local quadratic polynomial regressions of physical activity on interview date fixed effects.

Sources: Own data, own calculations.

Figure A2: Formal Common Trend Test



Notes: Confidence intervals are 95%.

Notes: Coefficient plot shows coefficients from regression of being physically active on our treatment dummy (being located in London in 2012) interacted with the calendar week in 2011 (the pre-treatment period).
Sources: Own data, own calculations.

Table A1: Balancing Properties (Pre-Treatment Year, 2011)

	Previously Inactive			All Residents		
	Mean		Norm. Difference	Mean		Norm. Difference
	London	Paris		London	Paris	
Age	47.7181	45.4041	0.1117	45.1011	43.9841	0.0527
Female	0.6017	0.5692	0.0466	0.5909	0.5264	0.0920
Male	0.3983	0.4308	0.0466	0.4091	0.4736	0.0920
Single	0.2802	0.2738	0.0101	0.3010	0.2963	0.0073
Married	0.4179	0.3704	0.0688	0.4294	0.3620	0.0978
Partnered	0.1351	0.2056	0.1333	0.1421	0.2086	0.1241
Separated	0.0284	0.0192	0.0431	0.0221	0.0205	0.0080
Divorced	0.0935	0.0976	0.0100	0.0742	0.0851	0.0284
Widowed	0.0449	0.0334	0.0421	0.0312	0.0276	0.0148
Graduate	0.9718	0.9502	0.0791	0.9578	0.9250	0.0988
Studies	0.0282	0.0498	0.0791	0.0422	0.0750	0.0988
Secondary School Degree	0.3325	0.2450	0.1373	0.2270	0.2198	0.0122
Vocational Degree	0.1384	0.1338	0.0096	0.1455	0.1064	0.0835
Undergraduate Degree	0.2015	0.3926	0.3023	0.1933	0.4112	0.3454
Graduate Degree	0.3275	0.2287	0.1569	0.4342	0.2626	0.2589
Employed Full-Time	0.3956	0.5562	0.2303	0.4574	0.5572	0.1419
Employed Part-Time	0.1241	0.0690	0.1324	0.1219	0.0681	0.1304
Self-Employed	0.0798	0.0293	0.1582	0.0947	0.0336	0.1776
Unemployed and Looking	0.0622	0.0549	0.0218	0.0527	0.0466	0.0198
Unemployed and Not Looking	0.1317	0.0565	0.1838	0.0855	0.0372	0.1430
Retired	0.1783	0.1843	0.0109	0.1456	0.1823	0.0702
Log Income	10.1835	10.2656	0.0783	10.4032	10.3431	0.0582
Owns Dwelling	0.5625	0.4958	0.0946	0.5894	0.5121	0.1102
Rents or Shares Dwelling	0.3665	0.4377	0.1030	0.3304	0.4090	0.1155
Lives With Relatives	0.0562	0.0490	0.0228	0.0673	0.0560	0.0333
Other Tenancy Type	0.0148	0.0120	0.0171	0.0128	0.0164	0.0207
N	4,183	5,899	-	14,019	15,529	-

Notes: We provide means and normalised differences separately for our estimation sample for the *previously inactive* (i.e. respondents who were not physically active in the pre-treatment year 2011) and for *all residents* on average (Table 3 Columns 3 and 4), for London and for Paris separately. Scale-free normalised differences are calculated as $\Delta x = (\bar{x_t} - \bar{x_c})/\sqrt{(\bar{\sigma_t^2} - \bar{\sigma_c^2}}$, where $\bar{x_t}$ and $\bar{x_c}$ is the sample mean of the covariate for the treatment and control group, respectively. σ denotes the respective variance. As a rule of thumb, a normalised difference greater than 0.25 indicates a non-balanced covariate (Imbens and Wooldridge, 2009; Imbens and Rubin, 2015). See Section 2.1 for a detailed description of our data. *Sources:* Own data, own calculations.

Table A2: Impact of Hosting the Olympics on Physical Activity (Extensive Margin) – No Controls

	Physical Activity (1: At Least Monthly = Being Active, 0: Less Often or Never = Being Inactive)			
	Model 2, 2011-2012		**Model 3, 2011-2012**	
	(1)	(2)	(3)	(4)
	Previously Inactive	All Residents	Previously Inactive	All Residents
London x Olympics			0.0706***	0.0003
			(0.0187)	(0.0082)
London x PostOlympics			0.0536***	-0.0054
			(0.0154)	(0.0081)
Olympics			0.2660***	0.0018
			(0.0099)	(0.0058)
PostOlympics			0.2510***	0.0058
			(0.0096)	(0.0061)
London x 2012	0.0527***	-0.0050		
	(0.0138)	(0.0055)		
2012	0.2540***	0.0032		
	(0.0075)	(0.0046)		
N	10,082	29,548	10,082	29,548
Individuals	6,400	19,046	6,400	19,046
Controls	No	No	No	No
Day-of-Week FE	Yes	Yes	Yes	Yes
Calendar-Month FE	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
R Squared	0.277	0.000	0.280	0.000

Notes: Our outcome is a binary indicator that indicates whether a respondent reports to be *physically active* (i.e. engaged in sport or physical activity at least monthly as opposed to less often or never). We estimate linear probability models, separately for the *previously inactive* (i.e. respondents who were not physically active in the pre-treatment year 2011) and for *all residents* on average. Treatment is defined as being resident in *London* as opposed to Paris. Model 2, which is our preferred specification, is based on Equation 2 and estimates a difference-in-differences model in the years 2011 and 2012. It takes the entire year 2012 as the relevant treatment period. The baseline category is the pre-treatment year 2011. Model 3 is based on Equation 3 and estimates a difference-in-differences model in the years 2011 and 2012. It again takes the entire year 2012 as the relevant treatment period but splits it into a pre-Olympics (suppressed), Olympics, and post-Olympics period. The baseline category is again the pre-treatment year 2011. Models 2 and 3 look at within-individual changes between Londoners and Parisians between 2011 and 2012. See Section 2.1 for a detailed description of our data and Section 2.2 for our models. Robust standard errors clustered at interview date level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Sources: Own data, own calculations.

Table A3: Impact of Hosting the Olympics on Physical Activity (Extensive Margin) – Full Table

	Physical Activity by Previously Inactive (1: Yes, 0: No)	
	Model 2, 2011-2012	**Model 3, 2011-2012**
	(1)	(2)
London x Olympics		0.0794***
		(0.0186)
London x PostOlympics		0.0570***
		(0.0179)
Olympics		0.2710***
		(0.0231)
PostOlympics		0.2470***
		(0.0256)
London x 2012	0.0559***	
	(0.0151)	
2012	0.2630***	
	(0.0238)	
Controls		
Age	0.0677**	0.0744***
	(0.0271)	(0.0272)
Age Squared	-0.0011***	-0.0011***
	(0.0002)	(0.0002)
Female	Reference Category	Reference Category
Male		
Single	Reference Category	Reference Category
Married	-0.0153	-0.0209
	(0.0409)	(0.0411)
Partnered	-0.0110	-0.0098
	(0.0290)	(0.0290)
Separated	-0.0126	-0.0141
	(0.0496)	(0.0500)
Divorced	0.0144	0.0114
	(0.0474)	(0.0467)
Widowed	0.1080*	0.1150*
	(0.0624)	(0.0629)
Graduate	Reference Category	Reference Category
Studies	0.0349	0.0360
	(0.0504)	(0.0501)
Secondary School Degree	Reference Category	Reference Category
Vocational Degree	0.0273	0.0260
	(0.0181)	(0.0179)
Undergraduate Degree	-0.0074	-0.0069
	(0.0144)	(0.0143)
Graduate Degree	-0.0095	-0.0071
	(0.0179)	(0.0180)
Owns Dwelling	Reference Category	Reference Category
Rents or Shares Dwelling	0.0704***	0.0709***
	(0.0248)	(0.0247)
Lives With Relatives	0.0193	0.0206
	(0.0241)	(0.0242)
Other Tenancy Type	-0.0163	-0.0172
	(0.0423)	(0.0422)
N	10,082	10,082
Individuals	6,400	6,400
Day-of-Week FE	Yes	Yes
Calendar-Month FE	Yes	Yes
Individual FE	Yes	Yes
R Squared	0.283	0.286

Notes: Our outcome is a binary indicator that indicates whether a respondent reports to be *physically active* (i.e. engaged in sport or physical activity at least monthly as opposed to less often or never). We estimate linear probability models separately for the *previously inactive* (i.e. respondents who were not physically active in the pre-treatment year 2011) and for *all residents* on average. Treatment is defined as being resident in *London* as opposed to Paris. Model 2, which is our preferred specification, is based on Equation 2 and estimates a difference-in-differences model in the years 2011 and 2012. It takes the entire year 2012 as the relevant treatment period. The baseline category is the pre-treatment year 2011. Model 3 is based on Equation 3 and estimates a difference-in-differences model in the years 2011 and 2012. It again takes the entire year 2012 as the relevant treatment period but splits it into a pre-Olympics (suppressed), Olympics, and post-Olympics period. The baseline category is again the pre-treatment year 2011. Models 2 and 3 look at within-individual changes between Londoners and Parisians between 2011 and 2012. See Section 2.1 for a detailed description of our data and Section 2.2 for our models. Robust standard errors clustered at interview date level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Sources: Own data, own calculations.

Table A4: Impact of Hosting the Olympics on Physical Activity (Previously Active)

	Physical Activity (1: At Least Monthly = Being Active, 0: Less Often or Never = Being Inactive)	
	Model 2, 2011-2012	**Model 3, 2011-2012**
	(1)	(2)
	Previously Active	Previously Active
London x Olympics		0.0203**
		(0.0094)
London x PostOlympics		0.0202**
		(0.0098)
Olympics		-0.1509***
		(0.0092)
PostOlympics		-0.1576***
		(0.0096)
London x 2012	0.0203***	
	(0.0067)	
2012	-0.1549***	
	(0.0076)	
N	19,430	19,430
Individuals	12,610	12,610
Controls	Yes	Yes
Day-of-Week FE	Yes	Yes
Calendar-Month FE	Yes	Yes
Individual FE	Yes	Yes
R Squared	0.151	0.151

Notes: Our outcome is a binary indicator that indicates whether a respondent reports to be *physically active* (i.e. engaged in sport or physical activity at least monthly as opposed to less often or never). We estimate linear probability models for the *previously active* (i.e. respondents who were physically active in the pre-treatment year 2011). Treatment is defined as being resident in *London* as opposed to Paris. Model 2, which is our preferred specification, is based on Equation 2 and estimates a difference-in-differences model in the years 2011 and 2012. It takes the entire year 2012 as the relevant treatment period. The baseline category is the pre-treatment year 2011. Model 3 is based on Equation 3 and estimates a difference-in-differences model in the years 2011 and 2012. It again takes the entire year 2012 as the relevant treatment period but splits it into a pre-Olympics (suppressed for brevity), Olympics, and post-Olympics period. The baseline category is the pre-treatment year 2011. Models 2 and 3 look at within-individual changes between Londoners and Parisians between 2011 and 2012. See Section 2.1 for a detailed description of our data and Section 2.2 for our models. Robust standard errors clustered at interview date level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Sources: Own data, own calculations.

Table A5: Heterogeneity by Age

	Physical Activity (1: At Least Monthly = Being Active, 0: Less Often or Never = Being Inactive)			
	Model 2, 2011-2012		Model 3, 2011-2012	
	(1)	(2)	(3)	(4)
	Previously Inactive		Previously Inactive	
	Below 45 Years	45 Years and Above	Below 45 Years	45 Years and Above
London x Olympics			0.1270***	0.0602**
			(0.0357)	(0.0244)
London x PostOlympics			0.0461**	0.0692**
			(0.0218)	(0.0279)
Olympics			0.3230***	0.2450***
			(0.0494)	(0.0310)
PostOlympics			0.3240***	0.2080***
			(0.0462)	(0.0311)
London x 2012	0.0581***	0.0609***		
	(0.0210)	(0.0214)		
2012	0.3300***	0.2260***		
	(0.0451)	(0.0295)		
N	4,693	5,389	4,693	5,389
Individuals	3,131	3,345	3,131	3,345
Controls	Yes	Yes	Yes	Yes
Day-of-Week FE	Yes	Yes	Yes	Yes
Calendar-Month FE	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
R Squared	0.316	0.269	0.322	0.273

Notes: Our outcome is a binary indicator that indicates whether a respondent reports to be *physically active* (i.e. engaged in sport or physical activity at least monthly as opposed to less often or never). We estimate linear probability models, restricting our estimation sample to the *previously inactive* (i.e. respondents who were not physically active in the pre-treatment year 2011) and splitting it by mean age (i.e. 45 years). Treatment is defined as being resident in *London* as opposed to Paris. Model 2, which is our preferred specification, is based on Equation 2 and estimates a difference-in-differences model in the years 2011 and 2012. It takes the entire year 2012 as the relevant treatment period. The baseline category is the pre-treatment year 2011. Model 3 is based on Equation 3 and estimates a difference-in-differences model in the years 2011 and 2012. It again takes the entire year 2012 as the relevant treatment period but splits it into a pre-Olympics (suppressed), Olympics, and post-Olympics period. The baseline category is again the pre-treatment year 2011. Models 2 and 3 look at within-individual changes between Londoners and Parisians between 2011 and 2012. See Section 2.1 for a detailed description of our data and Section 2.2 for our models. Robust standard errors clustered at interview date level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Sources: Own data, own calculations.

Table A6: National Sporting Success

	Physical Activity by Previously Inactive (1: Yes, 0: No) Model 2, 2011-2012			
	(1)	(2)	(3)	(4)
	GoldSum	lGoldSum	GoldCum	lGoldCum
London x 2012 x GoldSum	-0.0048			
	(0.0127)			
GoldSum	-0.0005			
	(0.0113)			
London x 2012 x lGoldSum		0.0012		
		(0.0114)		
lGoldSum		-0.0040		
		(0.0111)		
London x 2012 x GoldCum			-0.0009	
			(0.0020)	
GoldCum			0.0011	
			(0.0019)	
London x 2012 x lGoldCum				-0.0024
				(0.0018)
lGoldCum				0.0022
				(0.0018)
London x 2012	0.0620***	0.0579***	0.0581***	0.0650***
	(0.0160)	(0.0156)	(0.0189)	(0.0187)
2012	0.2630***	0.2640***	0.2590***	0.2550***
	(0.0241)	(0.0240)	(0.0258)	(0.0256)
N	10,082	10,082	10,082	10,082
Individuals	6,400	6,400	6,400	6,400
R Squared	0.284	0.284	0.283	0.284
Controls	Yes	Yes	Yes	Yes
Day-of-Week FE	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
Calendar-Month FE	Yes	Yes	Yes	Yes

Notes: We estimate linear models, restricting our estimation sample to the *previously inactive* (i.e. respondents who were not physically active in the pre-treatment year 2011). Treatment is defined as being resident in *London* as opposed to Paris. Model 2, which is our preferred specification, is based on Equation 2 and estimates a difference-in-differences model in the years 2011 and 2012. It takes the entire year 2012 as the relevant treatment period. The baseline category is the pre-treatment year 2011. Column 1 interacts our treatment dummy with the number of gold medals won by a respondent's country, Column 2 with the lagged number of gold medals, Column 3 with the cumulative number of gold medals, and Column 4 with the lagged cumulative number of gold medals. See Section 2.1 for a detailed description of our data and Section 2.2 for our models. Robust standard errors clustered at interview date level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Sources: Own data, own calculations.

Table A7: Impact of Hosting the Olympics on Thoughts About and Satisfaction With Health (Previously Active)

	Thoughts About and Satisfaction With Health					
	Model 1, 2012		Model 2, 2011-2012		Model 3, 2011-2012	
	(1)	(2)	(3)	(4)	(5)	(6)
	Thoughts About Health (0-1)	Satisfaction With Health (0-10)	Thoughts About Health (0-1)	Satisfaction With Health (0-10)	Thoughts About Health (0-1)	Satisfaction With Health (0-10)
London x Olympics	0.0097	0.0343			0.0522**	0.2900***
	(0.0402)	(0.2030)			(0.0243)	(0.0899)
London x PostOlympics	0.0311	-0.0439			0.0459**	0.1720**
	(0.0327)	(0.1870)			(0.0178)	(0.0744)
Olympics	-0.0145	0.1890			-0.0270	-0.2840**
	(0.0350)	(0.2430)			(0.0335)	(0.1170)
PostOlympics	-0.0079	0.0564			-0.0096	-0.2970***
	(0.0326)	(0.2290)			(0.0277)	(0.1110)
London	-0.0466*	-0.5990***				
	(0.0276)	(0.1570)				
London x 2012			0.0450***	0.2080***		
			(0.0148)	(0.0647)		
2012			-0.0110	-0.2950***		
			(0.0279)	(0.1080)		
N	3,164	3,164	9,551	9,551	9,551	9,551
Individuals	3,164	3,164	6,393	6,393	6,393	6,393
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Day-of-Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Calendar-Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	No	No	Yes	Yes	Yes	Yes
R Squared	0.011	0.062	0.009	0.008	0.009	0.009

Notes: Our outcomes are *thoughts about health*, which is a binary indicator that indicates whether a respondent reports to have thought about their health continually or many times on the previous day as opposed to only a few times or not at all; and *satisfaction with health*, which is recorded on a zero-to-ten Likert scale, where zero denotes "not at all" and ten "completely satisfied". We estimate linear probability models, restricting our estimation sample to the *previously inactive* (i.e. respondents who were not physically active in the pre-treatment year 2011). Treatment is defined as being resident in *London* as opposed to Paris. Model 1 is based on Equation 1 and estimates a difference-in-differences model in the year 2012 only. The baseline category is the pre-Olympics period in the same year. Model 2, which is our preferred specification, is based on Equation 2 and estimates a difference-in-differences model in the years 2011 and 2012. It takes the entire year 2012 as the relevant treatment period. The baseline category is the pre-treatment year 2011. Model 3 is based on Equation 3 and estimates a difference-in-differences model in the years 2011 and 2012. It again takes the entire year 2012 as the relevant treatment period but splits it into a pre-Olympics (suppressed), Olympics, and post-Olympics period. The baseline category is again the pre-treatment year 2011. Models 2 and 3 look at within-individual changes between Londoners and Parisians between 2011 and 2012, Model 1 looks at between-individual changes in 2012. See Section 2.1 for a detailed description of our data and Section 2.2 for our models. Robust standard errors clustered at interview date level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Sources: Own data, own calculations.

Table A8: Impact of Hosting the Olympics on Other Health Behaviours (No Outliers)

	Other Health Behaviour (z-Scores)					
	Model 1, 2012		**Model 2, 2011-2012**		**Model 3, 2011-2012**	
	(1)	(2)	(3)	(4)	(5)	(6)
	Drinking	Smoking	Drinking	Smoking	Drinking	Smoking
London x Olympics	-0.0749	-0.0881**			-0.0420***	-0.0111*
	(0.0497)	(0.0426)			(0.0143)	(0.0067)
London x PostOlympics	-0.0162	-0.0747**			-0.0195	-0.0033
	(0.0513)	(0.0371)			(0.0151)	(0.0082)
Olympics	-0.0065	0.0821***			0.0178	
	(0.0286)	(0.0289)			(0.0135)	
PostOlympics	-0.0286	0.1168***			0.0260**	
	(0.0304)	(0.0250)			(0.0126)	
London	0.0854*	-0.0371				
	(0.0463)	(0.0329)				
London x 2012			-0.0242**	-0.0029		
			(0.0117)	(0.0056)		
2012			0.0214*	-0.0203***		
			(0.0114)	(0.0058)		
N	10,355	10,417	28,898	29,069	28,898	29,069
Individuals	10,355	10,417	18,784	18,798	18,784	18,798
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Day-of-Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Calendar-Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	No	No	Yes	Yes	Yes	Yes
R Squared	0.060	0.050	0.027	0.004	0.027	0.005

Notes: Outliers are defined as observations more than 2.5 standard deviations below or above the mean. Our outcomes are the number of alcoholic beverages drunk and the number of cigarettes smoked on the previous day, respectively, standardised to have a mean of zero and a standard deviation of one (i.e. z-scores). We estimate linear probability models for *all residents* on average. Treatment is defined as being resident in *London* as opposed to Paris. Model 1 is based on Equation 1 and estimates a difference-in-differences model in the year 2012 only. The baseline category is the pre-Olympics period in the same year. Model 2, which is our preferred specification, is based on Equation 2 and estimates a difference-in-differences model in the years 2011 and 2012. It takes the entire year 2012 as the relevant treatment period. The baseline category is the pre-treatment year 2011. Model 3 is based on Equation 3 and estimates a difference-in-differences model in the years 2011 and 2012. It again takes the entire year 2012 as the relevant treatment period but splits it into a pre-Olympics (suppressed), Olympics, and post-Olympics period. The baseline category is again the pre-treatment year 2011. Models 2 and 3 look at within-individual changes between Londoners and Parisians between 2011 and 2012, Model 1 looks at between-individual changes in 2012. See Section 2.1 for a detailed description of our data and Section 2.2 for our models. Robust standard errors clustered at interview date level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Sources: Own data, own calculations.

Table A9: Robustness Checks

	Physical Activity by Previously Inactive (1: Yes, 0: No)			
	Extended Controls Model 2, 2011-2012		Selection Into Follow-Up Survey Model 2, 2011-2012	
	(1) Economic	(2) Meteorological	(3) IPW	(4) Matching
London x 2012	0.0577***	0.0575***	0.0571***	0.0662***
	(0.0149)	(0.0152)	(0.0157)	(0.0190)
2012	0.2650***	0.2630***	0.2970***	0.1930***
	(0.0236)	(0.0238)	(0.0263)	(0.0288)
N	10,082	10,082	10,082	7,117
Individuals	6,400	6,400	6,400	5,128
Controls	Yes	Yes	Yes	Yes
Day-of-Week FE	Yes	Yes	Yes	Yes
Calendar-Month FE	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
R Squared	0.285	0.284	0.291	0.301

	Placebo and Confirmation Tests Model 2, 2011-2012		Placebo and Confirmation Tests (Including Physically Active)	
			Model 1, 2012	Model 1, 2011
	(5) Thought About Finance (0-1)	(6) National Pride (z-Score)	(7) National Pride (z-Score)	(8) National Pride (z-Score)
London x 2012	-0.0015	0.1830***		
	(0.0148)	(0.0324)		
2012	0.0362	0.0251		
	(0.0251)	(0.0690)		
London x Olympics			0.2760***	0.0359
			(0.0881)	(0.0687)
London x PostOlympics			0.1980**	
			(0.0820)	
Olympics			0.0356	0.0388
			(0.0597)	(0.0415)
PostOlympics			-0.0056	
			(0.0527)	
London			-0.1110	-0.0678***
			(0.0732)	(0.0161)
N	9,551	7,632	4,864	17,318
Individuals	6,393	5,905	4,864	17,318
Controls	Yes	Yes	Yes	Yes
Day-of-Week FE	Yes	Yes	Yes	Yes
Calendar-Month FE	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	No	No
R Squared	0.018	0.036	0.034	0.026

Notes: We estimate linear models, restricting our estimation sample to the *previously inactive* (i.e. respondents who were not physically active in the pre-treatment year 2011). Treatment is defined as being resident in *London* as opposed to Paris. Model 1 is based on Equation 1 and estimates a difference-in-differences model in the year 2012 only. The baseline category is the pre-Olympics period in the same year. Model 2, which is our preferred specification, is based on Equation 2 and estimates a difference-in-differences model in the years 2011 and 2012. It takes the entire year 2012 as the relevant treatment period. The baseline category is the pre-treatment year 2011. Column 1 includes additional economic controls (i.e. employment status and annual gross household income). Column 2 includes additional meteorological controls (i.e. daily average temperature and precipitation). Column 3 weights observations by the inverse probability of taking part in the 2012 survey. Column 4 matches respondents in the 2011 and 2012 surveys using one-to-one nearest-neighbour matching without replacement based on observables and then includes only statistical twins. Column 5 uses *thoughts about finance* and Column 6 *national pride* as confirmation outcomes. Column 7 estimates Model 1 for *national pride* as a confirmation outcome in the treatment year 2012, whereas Column 8 estimates the same model as a placebo test in the pre-treatment year 2011. See Section 2.1 for a detailed description of our data and Section 2.2 for our models. Robust standard errors clustered at interview date level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Sources: Own data, own calculations.

Table A10: Differences in Observables Between London and Paris in 2012

	Mean London	Paris	Norm. Difference
Age	47.5065	45.3835	0.1023
Female	0.5861	0.5243	0.0881
Male	0.4139	0.4757	0.0881
Single	0.2770	0.2899	0.0202
Married	0.4507	0.3715	0.1142
Partnered	0.1354	0.2019	0.1260
Separated	0.0202	0.0185	0.0086
Divorced	0.0814	0.0885	0.0182
Widowed	0.0353	0.0297	0.0224
Graduate	0.9793	0.9403	0.1412
Studies	0.0207	0.0597	0.1412
Secondary School Degree	0.2344	0.2032	0.0534
Vocational Degree	0.1414	0.0325	0.2784
Undergraduate Degree	0.1808	0.2422	0.1065
Graduate Degree	0.4434	0.5221	0.1117
Employed Full-Time	0.4628	0.5686	0.1506
Employed Part-Time	0.1257	0.0634	0.1514
Self-Employed	0.0921	0.0295	0.1868
Unemployed and Looking	0.0413	0.0428	0.0053
Unemployed and Not Looking	0.0880	0.0361	0.1531
Retired	0.1694	0.1998	0.0555
Log Income	10.4356	10.3965	0.0390
Owns Dwelling	0.6388	0.5390	0.1443
Rents or Shares Dwelling	0.3009	0.3887	0.1312
Lives With Relatives	0.0484	0.0533	0.0156
Other Tenancy Type	0.0118	0.0020	0.0838
N	4,646	5,930	-

Notes: Scale-free normalised differences are calculated as $\Delta x = (\bar{x}_t - \bar{x}_c)/\sqrt{(\bar{\sigma_t^2} - \bar{\sigma_c^2}}$, where $\bar{x}_t$ and $\bar{x}_c$ is the sample mean of the covariate for the treatment and control group, respectively. σ denotes the respective variance. As a rule of thumb, a normalised difference greater than 0.25 indicates a non-balanced covariate (Imbens and Wooldridge, 2009; Imbens and Rubin, 2015). See Section 2.1 for a detailed description of our data.
Sources: Own data, own calculations.

Table A11: Differences in Observables Between London and Paris by Period in 2012

	Pre-Olympics Period in 2012 Mean			Olympics Period in 2012 Mean			Post-Olympics Period in 2012 Mean		
	London	Paris	Norm. Difference	London	Paris	Norm. Difference	London	Paris	Norm. Difference
Age	50.1901	47.6148	0.1228	48.3452	46.4599	0.0893	46.3786	44.2126	0.1068
Female	0.6061	0.5259	0.1146	0.5767	0.5268	0.0710	0.5909	0.5221	0.0981
Male	0.3939	0.4741	0.1146	0.4233	0.4732	0.0710	0.4091	0.4779	0.0981
Single	0.3058	0.3815	0.1129	0.2800	0.3226	0.0658	0.2700	0.2502	0.0320
Married	0.4132	0.2833	0.1944	0.4523	0.3319	0.1758	0.4552	0.4157	0.0565
Partnered	0.0964	0.1704	0.1545	0.1224	0.1849	0.1229	0.1525	0.2197	0.1226
Separated	0.0193	0.0185	0.0040	0.0173	0.0207	0.0173	0.0228	0.0170	0.0295
Divorced	0.1019	0.1093	0.0168	0.0898	0.1008	0.0266	0.0711	0.0760	0.0134
Widowed	0.0634	0.0370	0.0853	0.0382	0.0392	0.0034	0.0284	0.0215	0.0314
Graduate	0.9780	0.9963	0.1152	0.9827	0.9621	0.0887	0.9767	0.9147	0.1955
Studies	0.0220	0.0037	0.1152	0.0173	0.0379	0.0887	0.0233	0.0853	0.1955
Secondary School Degree	0.2397	0.1463	0.1683	0.2244	0.2095	0.0255	0.2420	0.2085	0.0569
Vocational Degree	0.1763	0.0000	0.4620	0.1418	0.0629	0.1854	0.1357	0.0160	0.3279
Undergraduate Degree	0.1901	0.1574	0.0610	0.1749	0.2870	0.1897	0.1843	0.2242	0.0700
Graduate Degree	0.3939	0.6963	0.4501	0.4589	0.4406	0.0261	0.4380	0.5513	0.1613
Employed Full-Time	0.4463	0.6111	0.2365	0.4661	0.5511	0.1206	0.4625	0.5741	0.1588
Employed Part-Time	0.1047	0.0389	0.1815	0.1203	0.0616	0.1451	0.1335	0.0690	0.1522
Self-Employed	0.0744	0.0204	0.1810	0.0938	0.0348	0.1715	0.0935	0.0273	0.1984
Unemployed and Looking	0.0386	0.0222	0.0673	0.0408	0.0506	0.0333	0.0422	0.0407	0.0052
Unemployed and Not Looking	0.0992	0.0111	0.2777	0.0765	0.0330	0.1357	0.0960	0.0427	0.1494
Retired	0.2149	0.2926	0.1267	0.1851	0.2311	0.0802	0.1490	0.1610	0.0234
Log Income	10.3243	10.4046	0.0854	10.4422	10.3593	0.0807	10.4473	10.4221	0.0254
Owns Dwelling	0.6391	0.5389	0.1446	0.6680	0.5233	0.2107	0.6141	0.5504	0.0916
Rents or Shares Dwelling	0.3223	0.4259	0.1521	0.2794	0.4278	0.2221	0.3157	0.3538	0.0571
Lives With Relatives	0.0275	0.0315	0.0164	0.0413	0.0357	0.0208	0.0577	0.0699	0.0353
Other Tenancy Type	0.0110	0.0000	0.1054	0.0112	0.0035	0.0637	0.0125	0.0013	0.0960
N	363	540	-	1,961	2,272	-	2,322	3,118	-

Notes: Scale-free normalised differences are calculated as $\Delta x = (\bar{x_t} - \bar{x_c})/\sqrt{(\bar{\sigma_t^2} - \bar{\sigma_c^2}}$, where $\bar{x_t}$ and $\bar{x_c}$ is the sample mean of the covariate for the treatment and control group, respectively. σ denotes the respective variance. As a rule of thumb, a normalised difference greater than 0.25 indicates a non-balanced covariate (Imbens and Wooldridge, 2009; Imbens and Rubin, 2015). See Section 2.1 for a detailed description of our data.
Sources: Own data, own calculations.

Table A12: Impact of Hosting the Olympics on Physical Activity (Extensive Margin) – Logit

	Physical Activity (1: At Least Monthly = Being Active, 0: Less Often or Never = Being Inactive)			
	Model 2, 2011-2012		**Model 3, 2011-2012**	
	(1)	(2)	(3)	(4)
	Previously Inactive		Previously Inactive	
	Linear RE	Logit RE	Linear RE	Logit RE
London x Olympics			0.0737***	0.0290***
			(0.0205)	(0.0089)
London x PostOlympics			0.0594***	0.0266***
			(0.0221)	(0.0087)
Olympics			0.2646***	0.1055***
			(0.0162)	(0.0072)
PostOlympics			0.2509***	0.1030***
			(0.0154)	(0.0069)
London	0.0010		0.0002	-0.0181**
	(0.0065)		(0.0061)	(0.0081)
London x 2012	0.0563***	0.0688***		
	(0.0163)	(0.0158)		
2012	0.2583***			
	(0.0127)			
N	10,082	10,082	10,082	10,082
Individuals	6,400	6,400	6,400	6,400
Controls	Yes	Yes	Yes	Yes
Day-of-Week FE	Yes	Yes	Yes	Yes
Calendar-Month FE	Yes	Yes	Yes	Yes
Individual FE	No	No	No	No
(Pseudo) R Squared	0.277	-	0.280	-

Notes: Our outcome is a binary indicator that indicates whether a respondent reports to be *physically active* (i.e. engaged in sport or physical activity at least monthly as opposed to less often or never). We estimate (random-effects) linear and logit models, restricting our estimation sample to the *previously inactive* (i.e. respondents who were not physically active in the pre-treatment year 2011). Treatment is defined as being resident in *London* as opposed to Paris. Model 2, which is our preferred specification, is based on Equation 2 and estimates a difference-in-differences model in the years 2011 and 2012. It takes the entire year 2012 as the relevant treatment period. The baseline category is the pre-treatment year 2011. Model 3 is based on Equation 3 and estimates a difference-in-differences model in the years 2011 and 2012. It again takes the entire year 2012 as the relevant treatment period but splits it into a pre-Olympics (suppressed), Olympics, and post-Olympics period. The baseline category is again the pre-treatment year 2011. Models 2 and 3 look at within-individual changes between Londoners and Parisians between 2011 and 2012. See Section 2.1 for a detailed description of our data and Section 2.2 for our models. Robust standard errors clustered at interview date level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Sources: Own data, own calculations.

Table A13: Impact of Hosting the Olympics on Physical Activity, Drinking, and Smoking (Wild-Cluster Bootstrapped Standard Errors)

	Physical Activity Model 3, 2011-2012		Other Health Behaviour (z-Scores) Model 3, 2011-2012	
	(1) Previously Inactive	(2) All Residents	(3) Drinking	(4) Smoking
London x Olympics	0.0794***	0.0009	-0.0634	-0.0196
	(0.0186)	(0.0078)	(0.0266)	(0.0098)
London x PostOlympics	0.0570***	-0.0074	-0.0075	0.0021
	(0.0179)	(0.0090)	(0.0226)	(0.0103)
Olympics	0.2710***	0.0121	-0.0063***	-0.0170***
	(0.0231)	(0.0091)	(0.0184)	(0.0107)
PostOlympics	0.2470***	0.0172***	0.0021	-0.0386***
	(0.0256)	(0.0092)	(0.0154)	(0.0079)
N	10,082	29,548	29,606	29,618
Individuals	6,400	19,046	19,071	19,072
Controls	Yes	Yes	Yes	Yes
Day-of-Week FE	Yes	Yes	Yes	Yes
Calendar-Month FE	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
R Squared	0.286	0.002	0.028	0.006

Notes: Our outcome for physical activity is a binary indicator that indicates whether a respondent reports to be *physically active* (i.e. engaged in sport or physical activity at least monthly as opposed to less often or never). Our outcomes for drinking and smoking are the number of alcoholic beverages drunk and the number of cigarettes smoked on the previous day, respectively, standardised to have a mean of zero and a standard deviation of one (i.e. z-scores). We estimate linear probability models. Treatment is defined as being resident in *London* as opposed to Paris. Model 3 is based on Equation 3 and estimates a difference-in-differences model in the years 2011 and 2012. It takes the entire year 2012 as the relevant treatment period but splits it into a pre-Olympics (suppressed), Olympics, and post-Olympics period. The baseline category is again the pre-treatment year 2011. Model 3 looks at within-individual changes between Londoners and Parisians between 2011 and 2012. See Section 2.1 for a detailed description of our data and Section 2.2 for our models. Wild-cluster bootstrapped standard errors clustered at city level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Sources: Own data, own calculations.

Table A14: Impact of Hosting the Olympics on Physical Activity (Intensive Margin) – RE Probit (Marginal Effects)

	Physical Activity (1: Yes, 0: No) Model 2, 2011-2012			
	(1)	(2)	(3)	(4)
	¿ 5 Times a Week	¿ 3 Times a Week	At Least Weekly	At Least Monthly
Panel A: Previously Inactive				
London x 2012	0.0155***	0.0432***	0.0580***	0.0665***
	(0.0051)	(0.0081)	(0.0129)	(0.0154)
2012				
N	3,690	3,690	3,690	3,690
Individuals	3,690	3,690	3,690	3,690
R Squared	-	-	-	-
Panel B: All Residents				
London x 2012	0.0091	0.0058	0.0072	0.0146
	(0.0058)	(0.0082)	(0.0095)	(0.0090)
2012	-0.0137***	-0.0121*	-0.0130*	-0.0171***
	(0.0047)	(0.0063)	(0.0069)	(0.0065)
N	29,548	29,548	29,548	29,548
Individuals	19,046	19,046	19,046	19,046
R Squared	-	-	-	-
Controls	Yes	Yes	Yes	Yes
Day-of-Week FE	Yes	Yes	Yes	Yes
Calendar-Month FE	Yes	Yes	Yes	Yes
Random FE	Yes	Yes	Yes	Yes

Notes: Our outcomes are binary indicators for different intensities of physical activity, including > 5 *Times a Week*, > 3 *Times a Week*, *At Least Weekly*, and *At Least Monthly*. These different intensities are not mutually exclusive: respondents who report to be physically active > 5 *Times a Week*, by definition, are also physically active *At Least Weekly*, for example. We estimate panel data random-effects probit models separately for the *previously inactive* (i.e. respondents who were not physically active in the pre-treatment year 2011, Panel A) and for *all residents* on average (Panel B). The presented coefficients are marginal effects. Treatment is defined as being resident in *London* as opposed to Paris. Model 2, which is our preferred specification, is based on Equation 2 and estimates a difference-in-differences model in the years 2011 and 2012. It takes the entire year 2012 as the relevant treatment period. The baseline category is the pre-treatment year 2011. See Section 2.1 for a detailed description of our data and Section 2.2 for our models. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Sources: Own data, own calculations.

B Unobservable Selection and Coefficient Stability

Implicit in our argument about the coefficient stability of our treatment effects regardless of whether we include controls or not is the idea that coefficient movements are informative about relative bias due to unobservables (i.e. omitted variable bias). However, this is the case only if observables are correlated with unobservables, and as Oster (2019) shows, both coefficient movements *and* R Squared movements need to be taken into account to make meaningful statements about the degree of unobservable selection. Note that our R Squared values move only slightly when including our main set of controls, i.e. from 0.277 to 0.284.

Oster (2019) develops a bounding argument that helps make more informative statements about coefficient stability of treatment effects based on two parameters: the maximum attainable R Squared (R^2_{max}) and the degree of selection on unobservables relative to observables (δ). In particular, the author suggests calculating the δ that would be necessary to explain away the treatment effect obtained in the full model, i.e. $\beta_1 = 0$ in Equation 2, which is our preferred specification. If we follow this approach, assuming $R^2_{max} = 1$, we obtain $\delta = 0.028$. This implies that selection on unobservables is considerably *less* important than selection on observables.[26] An alternative approach is to calculate bounds on β_1, by varying δ and R^2_{max}. If we set $\delta = 0$ (i.e. unobservables are irrelevant for selection) and $R^2_{max} = 1$, we obtain $\beta_1 = 0.056$. If we set $\delta = 1$ (i.e. unobservables are as important as observables for selection) and $R^2_{max} = 0.284$ (which is the R Squared in our full model), we obtain $\beta_1 = 0.039$.[27] This gives us an interval of $[0.039; 0.056]$ for β_1, whereby the lower bound excludes zero at the 95% significance level given a standard error of 0.015 in our full model, i.e. $0.039 - 1.96 \times 0.015 = 0.01$. We conclude that selection on unobservables and potentially resulting omitted variable bias is, if anything, minor.

[26] For reference, $\delta = 2$ implies that selection on unobservables is *twice* as important as selection on observables.

[27] Oster (2019) considers $\delta = 1$ a sensible seed value, as observables should, in theory, be at least as important as unobservables.

C Multiple Hypotheses Testing

Taken together, we test 24 hypotheses in our main analysis, i.e. eight hypotheses related to physical activity, drinking, and smoking in our first model in Equation 1, four in our second model in Equation 2, and twelve in our third model in Equation 3. To account for multiple hypotheses testing, we used the stepdown multiple testing procedure by Romano and Wolf (2005b,a), with the four-step algorithm outlined in Romano and Wolf (2016). In essence, the algorithm constructs a null distribution for each of our 24 hypotheses tests based on a set of null resampling test statistics (in our case, using a bootstrap with 100 repetitions and cluster-robust standard errors at the interview date level in both the original regression and during the resampling procedure). We find that our stepdown-adjusted P values (corresponding to the significance of a hypothesis test where 24 tests were implemented) continue to indicate significance at conventional levels for all of our coefficient estimates for physical health (where our original P values indicated significance), yet those for other health behaviours (i.e. drinking and smoking) turn insignificant (Appendix Tables C1 and C2).

Table C1: Impact of Hosting the Olympics on Physical Activity (Extensive Margin) – Multiple Hypotheses Testing

	Physical Activity (1: At Least Monthly = Being Active, 0: Less Often or Never = Being Inactive)			
	Model 2, 2011-2012		**Model 3, 2011-2012**	
	(1)	(2)	(3)	(4)
	Previously Inactive	All Residents	Previously Inactive	All Residents
London x Olympics			0.0794***	0.0009
			(0.0186)	(0.0078)
Original P Value			*0.00*	*0.91*
Stepdown-Adjusted P Value			*0.01*	*0.99*
London x PostOlympics			0.0570***	-0.0074
			(0.0179)	(0.0090)
Original P Value			*0.00*	*0.41*
Stepdown-Adjusted P Value			*0.07*	*0.93*
London x 2012	0.0559***	-0.0053		
	(0.0151)	(0.0060)		
Original P Value	*0.00*	*0.38*		
Stepdown-Adjusted P Value	*0.01*	*0.70*		
N	10,082	29,548	10,082	29,548
Individuals	6,400	19,046	6,400	19,046
Controls	Yes	Yes	Yes	Yes
Day-of-Week FE	Yes	Yes	Yes	Yes
Calendar-Month FE	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
R Squared	0.283	0.002	0.286	0.002

Notes: The coefficients *Olympics*, *PostOlympics*, *London*, and 2012 are not shown. Our outcome is a binary indicator that indicates whether a respondent reports to be *physically active* (i.e. engaged in sport or physical activity at least monthly as opposed to less often or never). We estimate linear probability model, separately for the *previously inactive* (i.e. respondents who were not physically active in the pre-treatment year 2011) and for *all residents* on average. We report, in addition to original P values, stepdown-adjusted P-values from the stepdown multiple testing procedure suggested by Romano and Wolf (2005b,a). Treatment is defined as being resident in *London* as opposed to Paris. Model 2, which is our preferred specification, is based on Equation 2 and estimates a difference-in-differences model in the years 2011 and 2012. It takes the entire year 2012 as the relevant treatment period. The baseline category is the pre-treatment year 2011. Model 3 is based on Equation 3 and estimates a difference-in-differences model in the years 2011 and 2012. It again takes the entire year 2012 as the relevant treatment period but splits it into a pre-Olympics (suppressed), Olympics, and post-Olympics period. The baseline category is again the pre-treatment year 2011. Models 2 and 3 look at within-individual changes between Londoners and Parisians between 2011 and 2012. See Section 2.1 for a detailed description of our data and Section 2.2 for our models. Robust standard errors clustered at interview date level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Sources: Own data, own calculations.

Table C2: Impact of Hosting the Olympics on Other Health Behaviours – Multiple Hypotheses Testing

	Other Health Behaviour (z-Scores)					
	Model 1, 2012		Model 2, 2011-2012		Model 3, 2011-2012	
	(1)	(2)	(3)	(4)	(5)	(6)
	Drinking	Smoking	Drinking	Smoking	Drinking	Smoking
London x Olympics	-0.0855	-0.0813*			-0.0634**	-0.0196**
	(0.0789)	(0.0421)			(0.0266)	(0.0098)
Original P Value	*0.28*	*0.06*			*0.02*	*0.05*
Stepdown-Adjusted P Value	*0.44*	*0.11*			*0.37*	*0.57*
London x PostOlympics	0.0231	-0.0430			-0.0075	0.0021
	(0.0798)	(0.0360)			(0.0226)	(0.0103)
Original P Value	*0.77*	*0.24*			*0.74*	*0.84*
Stepdown-Adjusted P Value	*0.73*	*0.44*			*0.99*	*0.99*
London x 2012			-0.0245	-0.0049		
			(0.0193)	(0.0076)		
Original P Value			*0.21*	*0.52*		
Stepdown-Adjusted P Value			*0.58*	*0.70*		
N	10,591	10,599	29,606	29,618	29,606	29,618
Individuals	10,591	10,599	19,071	19,072	19,071	19,072
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Day-of-Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Calendar-Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	No	No	Yes	Yes	Yes	Yes
R Squared	0.058	0.051	0.028	0.006	0.028	0.006

Notes: The coefficients *Olympics*, *PostOlympics*, *London*, and 2012 are not shown. Our outcomes are the number of alcoholic beverages drunk and the number of cigarettes smoked on the previous day, respectively, standardised to have a mean of zero and a standard deviation of one (i.e. z-scores). We estimate linear probability models for *all residents* on average. We report, in addition to original P values, stepdown-adjusted P-values obtained from the stepdown multiple testing procedure suggested by Romano and Wolf (2005b,a). Treatment is defined as being resident in *London* as opposed to Paris. Model 1 is based on Equation 1 and estimates a difference-in-differences model in the year 2012 only. The baseline category is the pre-Olympics period in the same year. Model 2, which is our preferred specification, is based on Equation 2 and estimates a difference-in-differences model in the years 2011 and 2012. It takes the entire year 2012 as the relevant treatment period. The baseline category is the pre-treatment year 2011. Model 3 is based on Equation 3 and estimates a difference-in-differences model in the years 2011 and 2012. It again takes the entire year 2012 as the relevant treatment period but splits it into a pre-Olympics (suppressed), Olympics, and post-Olympics period. The baseline category is again the pre-treatment year 2011. Models 2 and 3 look at within-individual changes between Londoners and Parisians between 2011 and 2012, Model 1 looks at between-individual changes in 2012. See Section 2.1 for a detailed description of our data and Section 2.2 for our models. Robust standard errors clustered at interview date level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Sources: Own data, own calculations.

D Triple Differencing

In our surveys, it was randomised when exactly a respondent was interviewed in each wave, in both London and Paris. We can thus implement a difference-in-differences-in-differences design (i.e. triple differencing) to account for potentially time-varying confounders that may differ between both cities during the treatment period (i.e. the summer of 2012). Equation (4) builds on our preferred specification in Equation (2), but now compares within-individual changes in physical activity between Londoners and Parisians from 2011 to 2012, between those who were randomly interviewed after and those who were randomly interviewed before the start of the Olympics in 2012:

$$\begin{aligned} y_{it} = \beta_0 + \beta_1 London \times 2012 \times Post + \beta_2 London \times Post + \beta_3 2012 \times Post + \beta_4 2012 + \beta_5 Post + \\ + \beta_6' X_{it} + t_m + t_{wd} + u_i + \epsilon_{it} \end{aligned} \tag{4}$$

where *Post* is a dummy that is one if the individual is randomly interviewed after the start of the Olympics in 2012 (i.e. on or after 27 July, 2012), and zero if randomly interviewed before (i.e. before 27 July, 2012). All other variables are as in Equation (2). Note that β_0, *London*, $London \times Post$, and $2012 \times Post$ are collinear and hence drop out.

If we estimate this triple-differencing specification, we again find that hosting London 2012 had no effect on the physical activity of residents on average, yet significantly increased physical activity among the previously inactive. In particular, we obtain a $\hat{\beta}_1 = 0.1170$ for this group, suggesting that the share of the active among the previously inactive increased by 11.7 percentage points, against a baseline of 34%, a strong effect (34%). The full table is available upon request.

E Cost-Benefit Analysis

Benefits. Hosting London 2012 increased the share of the physically active by about six percentage points (Table 2 Column 3). This yields an effect size of about 18%, or an increase of about 492,000 individuals, given a population of about 8.2 million in London in 2012 (Office for National Statistics, 2012).[28] However, this effect lasted only for three months (about 100 days, cf. Table 2 Column 5). Taking potential physical and mental healthcare cost savings from another UK study that uses data from around the same time as the event (Fujiwara et al., 2015), we estimate average physical healthcare cost savings of about $492,000 \times £13.25 \times (100 / 365) = £1,786,027$ and average mental healthcare cost savings of about $492,000 \times £17.86 \times (100 / 365) = £2,407,430$, hence average total cost savings of about £4.2 million.[29]

Costs. The costs of hosting London 2012 were about £9.3 billion (National Audit Office, 2012). Only about £40 million, however, were dedicated to the so-called *Inspire* programme, which aimed at increasing participation in grassroots sports, sports competition, and physical activity among the general public, UK-wide (Girginov, 2016).[30]

Net Benefits. The *Inspire* programme's total costs of about £40 million clearly exceed the total benefits of about £4.2 million from potential healthcare cost savings. In particular, we obtain net benefits of about £40 – 4.2 = –35.8 million. This estimate, however, may be overly conservative: for one, the costs are likely an upper bound, as the *Inspire* programme also included some elements unrelated to physical activity, or elements aimed at professional athletes. At the same time, our benefits are likely a lower bound, as physical activity brings with it more benefits than just healthcare cost savings, for example improved wellbeing (cf. Dolan et al., 2014), social outcomes (cf. Puhl and Heuer, 2009), or labour market prospects (cf. Rooth, 2011). Most importantly, it is likely that hosting London 2012 also had a positive effect on physical activity outside of London, in the rest of the UK, among those who were physically inactive prior to the event. This would be a more appropriate comparison, as the *Inspire* programme was UK-wide. What spillover would it take to break even?

Break-Even Spillover. According to the nationally representative UK Household Longitudinal Survey ("Understanding Society"), about 27.1% of the 59 million residents in the rest of the UK (about 16 million residents) were physically inactive in 2012. To break even, they would have had to experience a treatment effect of at least 0.2627, implying a break-even spillover from London to the rest of the UK of at least 433%.[31]

[28] $8.2 \text{ million} \times 0.34 - (8.2 \text{ million} \times (0.34 - 0.06)) = 492,000$

[29] In a report commissioned by the UK Department for Digital, Culture, Media and Sport (DCMS), Fujiwara et al. (2015) estimate that doing *any* sport, on average, reduces annual GP visit costs by £13.25 and annual psychotherapy visit costs by £17.86.

[30] The *Inspire* programme also had volunteering, sustainability, and business components, but it is widely acknowledged that the sports component made up the bulk of the programme (Girginov, 2016).

[31] The break-even spillover x can be calculated as: $x > 35,800,000 / (15,989,000 \times (13.25 + 17.86) \times (100 / 365))$. Solving for x yields $x = 0.2627$. This is about 433% of our identified treatment effect (Table 2 Column 3).