Team Visibility and City Travel
Evidence from the UEFA Champions League (Random) Draw

By BRUNO CAPRETTINI *

Does hosting a sports team boost the visibility of a city among tourists? I test this proposition by looking at the effect of playing soccer's UEFA Champions League on air travel. I compare routes across cities that had their teams randomly drawn into the same group in the first phase of the competition to routes across cities hosting teams randomly allocated to different groups. The average effect of being drawn into the same group is between 5 and 8 percent more arrivals for the three months following the group stage, a period which coincides with a break in the competition.

* Caprettini: Universitat Pompeu Fabra, Ramon Trias Fargas, 25-27, Barcelona, Spain 08005 (e-mail: bruno.caprettini@upf.edu). Special thanks to Antonio Ciccone for his constant feedback and encouragement. Thanks also to Francesco Amodio, Paula Bustos, Albrecht Glitz, Stephan Litschig and Alessandro Tarozzi for patiently reading and commenting various versions of the paper. Joshua Angrist, Gabrielle Fack, Marcel Fafchamps, Patricia Funk, Maria Paula Gerardino, Marc Goñi, Libertad Gonzalez, Andrea Matranga, Luigi Pascali, Jacopo Ponticelli, Andrei Potlogea, Tom Schmitz and Jag Tripathy discussed the paper at various stages and gave very valuable suggestions. All errors are mine.

I. Introduction

Professional teams are often taken to increase their home city's visibility among tourists. As teams get extensive media coverage, their home cities make the news, and this is thought to turn anonymous places into potential tourism destinations. Such potential benefits are often behind sports teams' demands for public subsidies. Yet, so far no analysis has shown that the increased visibility generated by a team brings more visitors into town.

Estimating the effect of team visibility on the attractiveness of a city as a destination is difficult, as both the location of teams and their decision to move are likely to be related to some of the factors that also make cities attractive to tourists (Siegfried and Zimbalist, 2000). I present the first evidence of a positive causal effect of team visibility on city travel by exploiting the random draw of the Union of European Football Associations (UEFA) Champions League soccer tournament, a European competition that takes place once a year. In the first phase of the tournament, participating teams are randomly divided into groups of 4, and have to play a 3-months round-robin tournament with the teams of their group. During this phase, the teams – and their home cities – are likely to get more mention in cities hosting group rivals, and the visibility of a city should therefore increase more in cities hosting teams in the same group than in other Champions League cities. As the groups are formed randomly, this setup allows to compare air travel between cities hosting teams in the same group (*treated routes*), to air travel between Champions League cities hosting teams assigned to different groups (*control routes*). The control routes represent a valid counterfactual for treated routes because the teams in these cities could have met, and did not by pure chance.

The group phase of the UEFA Champions League ends in early December and is followed by a break in the competition that lasts until early March. Looking at monthly data of intra European flights, I find that over the months from January to March cities hosting rivals in the group phase see an increase in monthly air traffic of 5 to 8 percentage points relative to routes across cities whose teams played in different groups. At the mean, this effect implies that between January and March around 2700 more people travel on treated routes. The effect fades away as the tournament proceeds to its final phase, and teams that were initially assigned to the same group no longer play against each other.

The UEFA Champions League brings enormous media attention to the 32 participating soccer teams. Football is by far the most popular sport in Europe, and the UEFA Champions League is the major continental competition for European soccer teams. The competition is covered widely by national and local media, and the games are watched by millions of people across the continent. Participating clubs do not play any other continental competition during the season, and Champions League games are never played on days when domestic league or other continental competitions take place, which means that no other game competes for attention. The first phase of the tournament takes place during fall and lasts for around three months. During this phase participating teams are randomly divided into groups of 4 and teams in the same group play against each other twice. My estimation strategy exploits the random formation of these groups and compares city travel between cities hosting teams in the same group with city travel between cities hosting teams in different groups. The advantage of this approach is that it eliminates the effect of participation in the UEFA Champions League, which is unlikely to be random.

Although the formation of groups is inherently random, I find that treated routes tend to be busier 9 to 3 months *before* the group phase starts (figure 1). I clean the residual heterogeneity across treated and control routes by estimating all regressions with route fixed effects. Figure 2 shows that air traffic has identical distribution across treated and control routes once route fixed effects are accounted for. I also estimate dyadic standard errors on incomplete networks of cities (Fafchamps and Gruber, 2007a and 2007b) and find that results are robust to dyadic inference.

The paper proceeds as follows. The next section reviews the related literature. Section III explains the identification strategy and presents the data. Section IV shows that observable variables are balanced across treated and control routes.

Section V presents the main results and section VI proves their robustness. Section VII discusses the results and concludes.

II. Related Literature

This paper relates to the empirical literature that has analyzed the impact of professional teams on local communities (Coates, 2007). This literature has examined the effect of sport franchises on the local economy and their contribution to the recovery of depressed urban areas (Baade, 1996; Coates and Humphreys, 2003). A recent strand of the literature has also examined the intangible benefits that teams bring to the local population. Carlino and Coulson (2004) argue that higher house rents in cities with NFL franchises compensate for better quality of life, and Coates and Humphreys (2006) show that, when asked to vote on the public funding of a new stadium, people living in areas closer to the proposed site of the facility support the project more than people living farther away.

With the exception of Coates and Humphreys (2006), these papers study a panel of U.S. cities. In this context, identification comes from a difference-in-difference approach that exploits the relocation of some professional teams across U.S. cities. Since the decision to relocate might be related to other time-varying determinants of cities' economic success, researchers control for existing city-specific linear trends. Coates and Humphreys (2006) on the other hand, analyze the voting behavior across different areas of a city during referendums to introduce subsidies for local sports facilities. Their empirical results are essentially descriptive, and they do not claim to uncover any causal relation. To the best of my knowledge, this is the first paper in the literature that exploits a natural experiment to establish the causal effects of having a sports team in town.

III. Identification and Data

A. Identification and Sample

If a team increases the visibility of its home town, then we should observe visitors travel to the city after the club appears on high-profile international games. Moreover, the greatest number of visitors should come from cities where rival teams reside, and where the matches are arguably more salient.

The Champions League is the most important showcase for European football teams, and the group phase of the competition offers the opportunity to be under the spotlight repeatedly over a period of three months. If the visibility effect is truly greater in the cities where opposing teams reside, it is possible to test its existence by regressing visitor arrivals from city j to city i in month m ($V_{ij,m}$) on an indicator of whether the two cities had their team matched in the previous group phase of the Champions League (G_{ij}):

$$V_{ij,m} = \beta_0 + \beta_1 G_{ij} + e_{ij,m} . \tag{1}$$

In equation (1) β_1 is identified consistently because, within the population of all routes across cities with at least 1 team in the Champions League, G_{ij} is randomly assigned. Notice that even observing $V_{ij,m}$ with no error, equation (1) is a conservative test for the visibility effect, because taking part to the Champions League is likely to increase the visibility of a city in all participating cities, regardless of the group in which they play. Since equation (1) compares visitors from cities with a team in the same group to visitors from cities with teams in different groups, it tests for the presence of an effect of being in the same group *on top of* the simple effect of taking part in the same edition of the Champions League.

Before discussing the data, it is useful to explain the format and timing of the Champions League. As of the 2012-13 edition, the group phase is played by 32 European teams divided in 8 groups of 4 teams each. Access to the group phase is reserved to the teams that performed best in their respective national leagues during the previous season. Once admitted to the group phase, every team is seeded into one of four pots according to its international standing: the eight strongest teams are seeded into pot number 1, the next eight teams into pot number 2, and so on. After seeding, each of the 8 groups is made of exactly one team randomly drawn from every pot, with the only provision that teams from the same Football Federation should not play in the same group. The random draw is performed publicly in front of the press at the end of August. Once the groups are formed, between September and December each club has to play twice against each of the other 3 teams in its group: once at home and the another time as visitor.

Games are always played on Tuesday or Wednesday, and they are broadcasted live on national televisions in prime time. These matches, and the media attention that they create in the two cities where they take place, is my treatment. Since teams assigned to different groups have no occasion to play against each other between September and December, and since only one group phase match was forfeited in the past 15 years, compliance is always perfect[1]. After the conclusion of the group phase in early December, the first two teams of every group advance to the final phase: this has the knock-out format and proceeds from the round of sixteen in March to the final in May[2].

[1] The forfeited match is A.S. Roma versus Dynamo Kyiv F.C., that was scheduled for the 15th of September 2004. For reasons that I will explain shortly, I exclude all routes that connect cities in Ukraine, and for this reason this observation is never part of the sample analyzed.

[2] The interested reader may refer to appendix C for additional details on the structure and history of the competition.

The rules of the random draw imply that not all routes across Champions League cities are valid controls for treated routes. Teams in the same pot and teams from the same country never play in the same group: since routes across the cities of these teams could not have been treated, I exclude them throughout. Teams that have been seeded in the same pot tend to be of similar strength, and so this approach makes sure that routes across cities with two very strong (or two very weak) teams are not over represented in the control group. Since the strength of a team might be correlated with the economic performance of its home town, excluding these routes makes sure that results are not biased by the different composition of treated and control groups. On the other hand, teams coming from the same country never play in the same group in the Champions League, but have to play against each other in their national league. Since also these national games have the potential to increase the visibility of a city, dropping these routes makes sure that results are not biased downward by the inclusion of these routes among the controls.

In addition to these exclusions, I always omit treated and control routes to and from UEFA countries that require a passport and/or a visa to enter (these are: Israel, Serbia, Russia, Turkey and Ukraine). I do so because travel to these countries requires significantly more time and effort than travel within the Schengen Area, and these costs are likely to offset any boost coming from the Champions League. Although randomization was performed using teams from these countries too, between 1998-99 and 2010-11 only 13 percent of participants came from these nations, and their exclusion makes no difference in terms of balance of the remaining treated and control routes. Inclusion of these routes has two consequences on the results shown later. First, the distribution of variables across treatment and control routes is more balanced when these routes are included. Second, coefficients from all regressions are less precise and somewhat smaller (but still significant).

B. Data

High frequency data on city-to-city number of visitors is scarce, and $V_{ij,m}$ in equation (1) can be observed only imperfectly. I proxy $V_{ij,m}$, the number of visitors from city j to city i in month m, with (the logarithm of) $P_{ij,m}$, the number of arrivals from all airports serving city j to all airports serving city i in month m. These data are available from Eurostat at monthly frequency since 1998, so I will focus on all Champions League editions between the 1998-99 and the 2010-11. Data on all Champions League games comes from the UEFA official website. See appendix A for further details.

Studying the visibility effect on air traffic alone would not be a limitation if visitors traveling with other modes of transport are affected in similar ways. In practice, the relative importance of different modes of transport depends on the relative position of two cities, so that the visibility effect is likely to have heterogeneous effects across different routes and modes of transport. Random assignment implies that estimated coefficients are consistent estimates of the Average Treatment Effect (ATE) of the Champions League on air traffic. Since most Europeans travel abroad by plane, this is a relevant effect to estimate[3]. Moreover, if we were willing to assume that on average the Champions League effect is the same for other modes of transport, then the estimated coefficient

[3] Among the countries considered, across any pair of countries for which both air and train traffic is available, there was a median of 11.2 air travelers for every passenger arriving by train between 2004 and 2010. For countries that are connected via sea, the median ratio of air to boat arrivals over the same period was 2.9. There are no similar statistics on intra European road traffic, but given that *within* the countries considered the median ratio of passenger-Km transported by car (by coach) to those transported by train is 11.2 (1.6), air traffic is likely to be at least as important as car travel, and several times more relevant than either train, coach or boat. Notice moreover that these numbers are likely to be lower bounds of the relevance of air traffic within Europe, because they are computed only for country pairs for which a direct connection is active (either by train or via sea). For several country pairs in the sample no such link exists, and on many routes airplanes are simply the only practical mode of transport available.

would be a consistent estimate of the proportional effect of the Champions League on overall travel[4].

Finally, notice that the definition of the dependent variable is likely to give a lower bound for the total effect of Champions League on air travel. Not all routes have equal capacity, and affected visitors might find it convenient to fly across airports that do not serve directly the cities in the Champions League (for instance, tourists going to Turin may land in the larger airport of Milan for convenience). Since I only look at the airports directly serving cities with teams in the Champions League, and do not control for potential stopovers or airports in nearby cities, my estimates would be biased downward if the effect spills over to routes that are classified as controls, or if some of the effect travels on routes that I do not consider.

IV. Balancedness

This section documents the balance between treated and control pairs. The first line of table 1 shows that the baseline value of the variable of interest is *not* balanced across treated and control routes. Average arrivals between January and June are 0.21 log points greater on routes that will be treated the following September (p-value = 0.016). Figure 1 shows that also the *distribution* of this variable is different across treated and controls (the Kolmogorov-Smirnov test rejects the null of identical distributions at the 0.1 percent level). Since between January and June most participants of the Champions League edition that starts in September are still to be decided, these tests suggest that treated routes tend to be busier than control routes *always*.

[4] In order to estimate the ATE of the Champions League on the overall *number* of travelers one would still need to multiply the estimated proportional change by the average overall travel across all routes in the sample. Since city-level information on non airborne travel is not available, these computations are not feasible.

Although the formation of groups is inherently random, cities that send their teams more often to the Champions League are more likely to have their teams matched together, and at the same time might be richer and have busier routes. In general, if unobservable characteristics of two cities affect both the average air traffic and the likelihood of their teams to meet during the group phase of the Champions League, the estimates of the Champions League effect will be inconsistent. This is especially true when the dependent variable is air arrivals, because air traffic is extremely persistent[5], and in these cases Bruhn and McKenzie (2009) insist that consistency of estimates is warranted only when also the baseline value of the outcome of interest is balanced across treated and controls.

I address the heterogeneity in the baseline value of air arrivals across treated and controls by exploiting the panel structure of my data. In figure 2 I plot the distributions of the residuals of a regression of air arrivals between January and June *before* the group phase on route fixed effects (FE). The figure shows that the distribution of these residuals in treated routes is very similar to the distribution in control routes, and the Kolmogorov-Smirnov tests can not reject the null of identical distributions (p-value = 0.194). Figure 2 suggests that including route FE in my specifications allows to estimate consistently the effect on air arrivals of playing in the same Champions League group: for this reason I will only present results from regressions that include routes FE.

The rest of table 1 shows that treated and control pairs are very balanced also with respect to other economic, geographic and demographic observables that should correlate with air arrivals. Notice that the units of observation are *pairs* of cities, since both air travel and the treatment are defined over a network of European cities. Network analysis leads to dyadic regression, and appendix B

[5] In the sample considered, a univariate regression of log arrivals on its value 12 months before explains 93 percent of total variability.

discusses issues of identification and inference that arise in this context. Here it is sufficient to note that the treatment G_{ij} does not vary within a route: when a team from city *i* plays in the same group of a team from city *j* the opposite is also true. In this case the balance of treated and control routes with respect to city-specific characteristics must be tested both for the average and for the absolute difference of the variables across the two cities on a route. The intuition is that both the *levels* and the *difference* of these variables may correlate with air travel across the two cities, and consistency is warranted when treatment is uncorrelated with both. Appendix B deals with the details.

Overall treated and control routes are balanced in terms of all observables. With the exception of the test on air arrivals before the group phase starts (first two lines in table 1) almost all of the other tests in table 1 can not reject the null of identical means across treated and control routes. Out of 60 tests, three are significant at the 10 percent level and only one at the 5 percent level. Also a joint test that allows correlation across these variables cannot reject the null of identical means: when I standardize all variables and run a single regression on the treatment dummy, the coefficient of this dummy is not significant (*p*-value = 0.284). Since in this stacked regression I use standard errors clustered at the route-year level, this procedure tests for significant differences across treatment and control routes while allowing errors of different variables to be correlated within a single route-year. Although the test is not perfect, estimation of 60 Seemingly Unrelated Regressions (SUR) is not feasible in this context because the covariance matrix of errors is singular.

By and large these results are reassuring. They prove that treatment is randomly assigned with respect to both time invariant and time varying characteristics, and that also pre trends do not differ significantly across treated and control routes. Notice moreover that not observing all air routes should not be source of concern. Panel B of table 1 shows that out of all the routes across cities that could have met but did not, 14.9 percent have non missing air data for both ways of the route during the months of the group phase. This proportion is very similar to the proportion of routes across cities that had two teams playing in the same group (p-value = 0.248).

V. Results

A. Effect on the Month of the Match

I start by documenting the effect of being drawn into the same Champions League group on air arrivals in the month in which the game is played. Groups of supporters usually follow their team when this plays international competitions, and visitors coming to watch the game are likely to be attracted by the match, rather than by the increased visibility of the city. For this reason, the evidence presented here is not especially interesting *per se* and it does not prove that cities whose teams appear on international competition attract more visitors after the event. The objective of this section is to show that a group phase game is a relevant shock for city-to-city air travel, and that monthly air traffic picks this shock very precisely. This seems a necessary "sanity check" before showing that playing in the same group increases the flow of visitors also *after* the group phase concludes.

Figure 3 shows that a single group phase games increases monthly air traffic by 7.5 percent on average (p-value < 0.000). The figure plots the coefficients from the regression of log $P_{ij,m}$: the logarithm of monthly arrivals, on $G_{ij,m}$: a dummy variable that is equal to 1 if in that month teams from the two cities play a game of the group phase of the Champions League, and 0 otherwise. I add to the regression and plot in the figure 6 leads and 12 lags of this dummy ($G_{ij,m-l}$: l = –

6,...,+12) in order to have a visual idea of the dynamic effect of this shock[6]. The complete specification of the regression is:

$$
(2) \qquad \begin{aligned} \log P_{ij,m} &= \alpha_{ij,m} + \delta_t + \mu_m \times t + \sum_{l=-6}^{12} \gamma_l G_{ij,m-l} + \\ &+ \sum_c \phi^c (c_i + c_j) \times t + \sum_c \psi^c (c_i - c_j) \times t + e_{ij,m} \end{aligned},
$$

where δ_t is a year fixed effect, $\mu_m \times t$ is a trend specific to month m and $\sum_c (c_i + c_j) \times t$ and $\sum_c (c_i - c_j) \times t$ are trends specific to every country of origin and of destination. As explained in the previous section, consistent identification of the effect of $G_{ij,m-l}$ on log $P_{ij,m}$ requires the inclusion of route fixed effects (α_{ij}). In this specification I include a different route fixed effect for every month m ($\alpha_{ij,m}$) because seasonality of air arrivals is very correlated with route fixed effects. Since all games are played during the fall, inclusion of route fixed effects biases the estimates when all months are included. When in the next section I estimate a separate regression for every month of the year, I include only route fixed effects (α_{ij}).

The way country of origin and country of destination trends enter this and all the following regressions requires some explanations. Equation (2) is a *directional* dyadic regression (i.e. log $P_{ij,m} \neq$ log $P_{ji,m}$), and for this reason I insert country trends $c_i \times t$ as suggested by Fafchamps and Gruber (2007b). This procedure is intended to impose symmetry on the effect that country trends have on air traffic. To see how this works, take all arrivals to and from Italy ($c_i = IT_i$): for this country, ϕ^{IT} captures the average trend of passengers to and from all Italian airports in the sample, while ψ^{IT} picks the average trend of passengers for

[6] The number of leads and lags was chosen in order to span half a year before and one year after the event, but figures similar to figure 3 can be produced with any number of leads and lags. Standard errors used to compute the confidence interval shown are clustered at the route-month level as in the regressions (3) and (4) below. See the discussion therein for details.

routes whose *origin* is in Italy. Symmetry in this context requires that the average trend of passengers for routes that have *destination* in Italy be equal to *minus* the average trend for routes with origin there. The terms $\sum_c \phi^c (c_i + c_j) \times t$ and $\sum_c \psi^c (c_i - c_j) \times t$ in equation (2) impose this condition.

Since treatment is randomly assigned once route treated effects are controlled for, all γ_l plotted in figure 3 have causal interpretation. On average, each Champions League match increases city-to-city monthly air travel by 7.5 percent. On a monthly average of 16545 arrivals, this equals 1238 passengers more for every match, 0.06 standard deviations or almost 7 full Airbus A320. Note moreover that not only the spike is exactly on the month of the match, but that all coefficients before the match are not significantly different from 0, neither alone nor jointly (F = 1.26; p-value = 0.271).

Figure 3 also suggests that treated routes stay relatively busier during the months following the game, although the effect seems to dim overtime. A joint test of the first 8 lags of $G_{ij\ m-l}$ rejects the null of no effect (F = 3.09; p-value = 0.002), but only the fourth lag is individually significant at the 5 percent level. Notice however that regression (2) is a good test for the effect on the month of the match, but it is not appropriate to test for the existence of an effect after the group phase concludes. To see why, note that when teams from cities i and j are in the same group, they play twice between September and December. The problem with specification (2) is that it treats these two matches as separate events with identical effects after a fixed number of periods, while the "treatment" that matters is likely to be the whole group phase, from September to December. To make an example, if A.S. Rome plays against F.C. Barcelona in October and November, equation (2) imposes the effect of the first match in January (three months after it) to be the same as the effect of the second match in February (three months after the second game). However, it seems more meaningful to estimate the joint effect of the two matches in January and February separately, without imposing to the two effects to be identical. This is the approach that I adopt in the next section.

B. The Effect after the Group Phase

In order to test more carefully the hypothesis that air travel increases more on treated routes after the group phase ends, I focus on one month at a time, and for every month between the end of the group phase and start of following edition I estimate a separate regression of the form:

$$\begin{aligned} \log P_{ij,m} &= \alpha_{ij} + \delta_t + \beta \cdot G_{ij} + \\ &+ \sum_c \phi^c (c_i + c_j) \cdot t + \sum_c \psi^c (c_i - c_j) \cdot t + e_{ij,m}, \end{aligned} \tag{3}$$

where $G_{ij} = 1$ if the teams from the two cities played in the same group in the last edition of the Champions League and the meaning of other symbols is the same as in regression (2). Consistency of β is warranted by random assignment of G_{ij} and the inclusion of route fixed effects α_{ij}.

Since the format of the competition has changed over the past 12 years, and since the likelihood of teams to access the group phase evolves overtime according to the UEFA ranking of their home country, both year and country trends could be correlated with the likelihood that a specific team accesses the group phase (see appendix C for a discussion of how country rankings affect the likelihood of accessing the group phase, and how UEFA computes them). In practice neither years nor country trends are correlated with the treatment: in a regression of G_{ij} on route-fixed effects, years fixed effects and country trends, it is not possible to reject the null of joint insignificance of year fixed effects and country trends ($F = 0.55$; p-value = 0.979). I include these controls to improve precision, but point estimates of β are barely affected when only route fixed effects (α_{ij}) are included.

Table 2 shows estimates of (3) for the month of the match and for all months from January through August. In all regressions standard errors are clustered at the route-month level[7]. The sample consists of all international routes across cities that could have met in the Champions League group phase between the 1998-99 and 2010-11 editions but it excludes routes across cities that have their teams playing a match during the Champions League knock-out stage either that year or the year before. I exclude these routes because knock-out phase matches are a bigger shock to air traffic than group phase matches and teams playing in the same group are less likely to meet again at later stages: as a result, inclusion of these routes biases estimates downwards[8].

The pattern in table 2 is clear: the effect of playing in the same group of the Champions League is large and significant both on the month of the match (row 1) and in the first three months following the end of the group phase. The effect is smaller and not significant by April, when the knock-out phase reaches its most important games (quarter of finals and semi finals), and it disappears before the start of the new season in September. The effect is also economically relevant: in the first 3 months of the year, the estimates imply 996 more passengers in January, 775 in February and 921 in March at the mean number of arrivals for these three months (11794, 12089 and 14798 respectively). Overall, two matches

[7] The correct unit of clustering is the route-month and not the route as a whole because here the "Moulton problem" is more troublesome than the persistence of the treatment. The Moulton problem arises because the dependent variable varies within a route-month while the treatment does not (Moulton, 1986 and Angrist and Pischke, 2009, p.313). On the other hand, the treatment is not persistent at all (the coefficient of a regression of treatment status on its lag on yearly data gives a coefficient of 0.01, p-value = 0.75): for this reason serial correlation is not an issue as in DD studies (Bertrand, Duflo and Mullainathan, 2004) and the correct unit of clustering is route-month rather than the whole route.

[8] The probability to meet a team from the same group is 0 for the round of sixteen, and very low afterwards. Excluding routes across cities that met in the knock-out phase one year before avoids that mean reversion 12 months after the event results in downward bias of the β, because meeting in a knock-out phase in a year is correlated with the probability of ending in the same group the year after. Inclusion of these routes drives estimates downward and worsen precision, but is not crucial for any of the results shown. Moreover, balance of treated and control routes holds also after excluding these routes.

played during the fall attract 2692 visitors during the first three months of the year.

Notice that the positive effects shown in the three months after the end of the group phase is not confounded by other matches played during the same period: the last group phase game is played on the first week of December, and I exclude all routes across cities that played a knock-out phase match during the spring. Moreover, teams are not allowed to take part in more than one international competition a year, and once they compete in the Champions League group phase they have no opportunity to meet until the next season (unless they advance in the Champions League and are pitched in a knock-out game, in which case they are dropped from my sample).

VI. Robustness Checks

In this section I show that all results shown in section V are robust to different strategies of inference. Traditional standard errors estimated in regression (3) are likely to be biased for several reasons. First, since the dependent variable varies within a route while the treatment does not, the "Moulton problem" (Moulton, 1986) is a serious issue. Second, standard errors in (3) can be biased because the regression is specified on a network of cities. Finally, the panel structure of data might complicate further the structure of standard errors, to the point that no single formula is appropriate to estimate them. I deal with these three issues in turn.

The Moulton problem arises here because, for every route *ij* in the sample, the dependent variable is observed twice: once as the number of arrivals from city *i* to city *j* in month *m* (log $P_{ij,m}$) and another time as number of arrivals from city *j* to city *i* in the same month (log $P_{ji,m}$). However, when a team from city *i* plays in the same group as a team from city *j* the opposite is also true, which means that the treatment variable does not vary within the route (i.e. $G_{ij} = G_{ji}$ always). Moulton (1986) shows that in these cases traditional standard errors are severely downward biased, and Angrist and Pischke (2009) propose several solutions to the problem. In figure 3 and in table 2 I estimate standard errors clustered at the route level; here, I follow another of the solutions proposed by Angrist and Pischke (2009) and use a single observation for every route-month. Panel A of table 3 shows estimates of (3) when I use as dependent variable the logarithm of the average number of arrivals across both ways of the route ($\log Avg\ P_{ij,m} \equiv \log [0.5 \times (P_{ij,m} + P_{ji,m})]$). Standard errors reported in this table are robust to heteroschedasticity. Although the point estimates are estimated less precisely than in table 2, all results go through.

Equation (3) specifies a dyadic regression because the units of observation are *pairs* of cities. This type of regression creates a complex variance-covariance matrix of errors, because shocks hitting a specific city can in principle affect air traffic between this city and all others destinations in the network. As a result, $E(e_{ij},e_{kl}) \neq 0$ whenever $i = k$ or $i = l$ or $j = k$ or $j = l$, and traditional standard errors are inconsistent. Fafchamps and Guber (2007a and 2007b) propose a formula to correct standard error in dyadic regressions and panel B of table 3 reports estimates of dyadic standard errors for the estimates presented in table 2[9]. These estimates are very close to the clustered standard errors estimated for table 2 and results with dyadic standard errors are, if anything, stronger than those with standard errors clustered at route-month level.

The panel structure of the data introduces additional complications, because the number of air arrivals is serially correlated overtime. I exploit the known structure of the Champions League random draw to assess the goodness of standard error shown in table 2 with a bootstrap-like procedure. More precisely, given the set of

[9] See appendix B for details on dyadic standard errors formula.

all routes across cities that had at least one team playing in the Champions League group phase between 1998 and 2010, I randomly create groups of four teams following the same rules UEFA applies to create its groups. Every year, each of my group is formed with exactly one team randomly drawn from each of the four pots. I also make sure that no two teams from the same football federation ever end up in the same group. Since these "placebo treatments" are not associated with any actual treatment, if standard errors are correct, estimates of equation (3) should on average be significant at the 5 percent level only 5 percent of the times. I draw 1000 such placebo treatments and for each draw I estimate regression (3) and store the *p*-value of the coefficient β. The first column of table 4 shows for each months in which regression (3) is estimated, the percentage of simulations that had a *p*-value smaller than 0.05. The results suggest that standard errors are slightly downward biased, as the likelihood to make a type I error at the 5 percent confidence is greater than 5 percent. However, the last column in table 4 also suggests that the bias in the standard errors is not so large as to invalidate all results. The last column in the table shows the share of simulations that estimated a β larger than the coefficient reported in table 2: overall they support the conclusion that a visibility effect exists at least for the three months following the group phase.

VII. Summary

Do professional teams make their home towns more visible among tourists? Using a natural experiment embedded in the European Champions League competition, I have shown that cities hosting teams receive more visitors from cities where Champions League games are more salient. My findings provide the first causal evidence that teams have the potential to increase the visibility of their home towns. To the extent that my results carry over to the US, they may help explain Carlino and Coulson's (2004) finding that house rents are significantly higher in U.S. cities with NFL franchises. The greater visibility of these cities may increase the flow of visitors, which would bring direct advantages for retail business and hotels for example, and ultimately should increase rents.

In any case, it is important to stress that the visibility effect I find appears to require continuous media exposure. The effect of playing in the same group of the Champions League disappears soon after teams stop playing against each other. This implies that the visibility effect depends on the structure of the competitions. For example, teams participating in leagues with no turnover (such as major American leagues) and teams that have to play more games per season may be more valuable for the visibility of a city.

Appendix A. Data description

I source air traffic data from Eurostat's "Detailed air passenger transport by reporting country and routes" tables[10]. For every airport in each European country, this database contains information on monthly air travel on every route from 1998 to 2010. On every route both arrivals and departures are available. Moreover, two different variables are available both for arrivals and for departures: total number of passengers carried and total number of passengers onboard. Passengers onboard equals passengers carried plus passengers that stop over and proceed to a different destination on the same aircraft, but in practice the two measures are almost identical (between January and August the correlation is 0.9994, p-value < 0.0001). Since for some countries only one between passengers carried and passengers onboard is reported, I pool all available information as follows. There are 4 measures of number of arrivals on each direction of a route: passengers carried and passengers onboard recorded as arrivals in the airport of destination, and passengers carried and passengers onboard recorded as departure in the airport of origin. The dependent variable used in the paper is the simple average of all measures available on every direction of a route. In the regressions shown in table 2, 67.9 percent of routes have all 4 measures available between January and August. Using information from the other end of a route cleans some of the noise present at the end of every month, when some passengers are recorded as flying on one month in one airport and on the following one at the other end[11]. In all regressions I use the natural logarithm of the dependent

[10] Data are available online at: http://epp.eurostat.ec.europa.eu/portal/page/portal/transport/data/database.

[11] Although also arrivals from j to i are very correlated with departures from j to i (0.9942, p-value < 0.0001) every month during which the former are greater than the latter are followed by a month in which the opposite happens, by exactly the same number of passengers.

variable so defined, and I winsorize the top and bottom 0.005 percent of observations to avoid extreme values to drive results.

Both tourism data (night spent in every NUTS 2 region) and demographic and economic data for European cities come from Eurostat[12]. Data on rail, maritime and road travel across and within European countries also come from Eurostat, and refer only to the countries used in the regressions[13]. Data on geographic coordinates of every European city used to compute distances come from Wikipedia. I hand-collected every match of the UEFA Champions League from the 1997-98 to the 2010-11 edition from the UEFA official website[14].

Appendix B. Dyadic Regression

This appendix is based on Fafchamps and Gubert (2007a and 2007b): refer to these papers for details. Both air traffic (log $P_{ij,m}$) and the Champions League treatment (G_{ij}) are observed on networks in which observations are city-pairs, and every city appears on several different pairs. Regression analysis on network data requires to specify a dyadic model, and both identification and inference need to be adjusted: I discuss these issues in turn.

Identification.—A simpler version of the dyadic regression analyzed in the text takes the form:

(B1) $$\log P_{ij} = \alpha + \beta \cdot G_{ij} + e_{ij} ,$$

where time subscripts are omitted for simplicity. In (B1) the treatment G_{ij} is specific to the route across city i and city j: in this case identification does not

[12] Tourism data is available at http://epp.eurostat.ec.europa.eu/portal/page/portal/tourism/data/database; data on cities at http://epp.eurostat.ec.europa.eu/portal/page/portal/region_cities/city_urban/data_cities/database_sub1.

[13] All available at: http://epp.eurostat.ec.europa.eu/portal/page/portal/transport/data/database.

[14] Online at: http://www.uefa.com/uefachampionsleague/history/.

require any correction. When characteristics specific to the two cities enter a dyadic regression however, it is important that they affect both ways of the route symmetrically: this means that the effect of city characteristic *c* on air travel must be such that the effect of c_i and c_j on log P_{ij} is the same as the effect of c_j and c_i on log P_{ji}. In order to impose this symmetry, Fafchamps and Gubert (2007a and 2007b) propose two different solutions, depending on whether the dyadic relationship is directional (as with air traffic, for which log $P_{ij} \neq$ log P_{ji}) or un-directional (as the average arrivals across a route log $Avg\ P_{ij}$ = log $Avg\ P_{ji}$). When the relationship is directional, symmetry is imposed by specifying the model:

(B2) $$\log P_{ij} = \alpha + \beta \cdot G_{ij} + \phi(c_i + c_j) + \psi(c_i - c_j) + e_{ij},$$

When the dyadic relationship is un-directional, symmetry is satisfied with:

(B3) $$\log AvgP_{ij} = \alpha + \beta \cdot G_{ij} + \phi(c_i + c_j) + \psi \,|\, c_i - c_j \,|+ e_{ij},$$

Regressions (2) and (3) are *directional* dyadic regressions (since arrivals from *j* to *i* in month *m* need not be equal to arrivals from *i* to *j* during the same period): in these regressions country-trends must enter the equation as in (B2). The dependent variables in the regressions on data collapsed at the route-month level and the treatment G_{ij} define *un*-directional relationships (log $Avg\ P_{ij}$ = log $Avg\ P_{ji}$ and $G_{ij} = G_{ji}$). For this reason country dummies enter regressions on collapsed data as in (B3), and for every city-specific variable for which I test the equality of means in table 1 I do so both for the sum and for the absolute difference across the two cities on a route.

Inference.— Standard errors in model (B1) need to take into account that shocks affecting city *i* will have an impact on all routes connecting *i*, and that this is true for all cities on all routes. This implies that in general $E(e_{ij}, e_{kl}) \neq 0$ whenever $i = k$ or $i = l$ or $j = k$ or $j = l$, and that the structure of the errors in regression (B1) has a form similar to that of a regression with clusters. Fafchamps and Gubert (2007a and 2007b) propose to correct the variance-covariance matrix of coefficients in a dyadic regression with a formula similar to the one proposed by Conley (1999) for spatially correlated errors:

$$\text{(B4)} \qquad \mathrm{AVar}(\hat{\beta}) = \frac{1}{N-K}(X'X)^{-1}\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{l=1}^{N}\frac{m_{ijkl}}{2N}X_{ij}e_{ij}e_{kl}'X_{kl}\right)(X'X)^{-1},$$

where $\hat{\beta}$ is the $K\times 1$ vector of estimated coefficient, N is the number of observations X is the matrix of all regressors, e_{ij} is the error in equation (B1) and $m_{ijkl} = 1$ if either $i = k$ or $i = l$ or $j = k$ or $j = l$.

Fafchamps and Gubert (2007a and 2007b) estimate (B4) on complete networks, i.e. networks in which every node is connected to every other node in the network. However, the air traffic network analyzed here is not complete: first, not every city has a direct connection to every other city taking part in the Champions League (there might exist both a Rome-Lille and a Rome-Valencia route, but no Lille-Valencia connection). Second, even if all routes existed, some routes are not valid “controls” for my treated routes, because cities that had teams in the same pot could not meet, and teams from the same country can not end up in the same group. In order to estimate (B4) on an incomplete network I coded a new option in the Stata program provided by Fafchamps: this is available upon request.

Appendix C. The UEFA Champions League

The first edition of the UEFA Champions League was held in 1955-56. Since then the number of participating teams and the general format of the competition have both changed many times. The group phase was introduced in the 1991-92 edition, and since 1994-95 all the matches in the group phase have been played between September and the first week of December. The number of groups has grown overtime, but these have always been formed randomly. Since the 1999-2000 edition there have always been 8 groups.

The rules to admit teams to the group phase vary by country and by year. “Major” leagues send the first 2 or 3 teams of the previous season directly to the group stage. Teams that ended first and second in one of the “minor” leagues, and teams that ended third or fourth in one of the major leagues take part to a “preliminary phase”. Official country rankings determine the number of teams that every country can send to the group phase or to the preliminary phase. These rankings are updated by UEFA every season according to 5-year moving average of the performance of national teams in all European competitions. The preliminary phase, played between July and August, consist of a series of knock-out matches that selects 10 of the 32 teams participating in the group phase. These games are not very popular and, since UEFA does not manage directly the TV rights for these games, they are only occasionally broadcasted, even in interested countries (European Commission, 2003).

Note that the format of the competition implies that both the year and the countries of team pairs might be correlated with the treatment. Year matters because the rules to qualify changed overtime (most notably in 1999, when the number of participating teams became 32, and in 2009, when the rules to access the group phase were renewed). These changes might have affected the probability of any 2 particular teams to meet, even conditional on reaching the group stage. Country specific trends are important because the number of participating teams from any country, and the pots where these teams are seeded depend on national UEFA coefficients, which in turn are updated every year, according to the current and past performance of national teams in UEFA competitions. Country rankings have evolved very differently over the last decade, often trailing domestic economic growth. For this reason, they might correlate with both the probability of treatment and the evolution of air traffic. In order to control for this confounders I include in every specification a set of dummies for both years and country of origin of the two teams.

REFERENCES

Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. "Mostly Harmless Econometrics - An Empiricist's Companion." Princeton, NJ: Princeton University Press.

Baade, Robert A. 1996. "Professional Sports As Catalysts For Metropolitan Economic Development." *Journal of Urban Affairs*, 18(1): 1–17.

Bertrand, Marianne, Ester Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics*, 119(1): 249–275.

Bruhn, Miriam, and David McKenzie. 2009 "In Pursuit of Balance: Randomization in Practice in Development Field Experiments." *American Economic Journal: Applied Economics*, 1(4): 200–232.

Carlino, Gerald, and Edward N. Coulson. 2004. "Compensating Differentials and the Social Benefits of the NFL." *Journal of Urban Economics*, 56(1): 25–50.

Coates, Dennis. 2007. "Stadiums And Arenas: Economic Development Or Economic Redistribution?," *Contemporary Economic Policy*, 25(4): 565–577.

Coates, Dennis, and Brad R. Humphreys. 2003 "The effect of professional sports on earnings and employment in the services and retail sectors in US cities," *Regional Science and Urban Economics*, 33(2): 175–198.

Coates, Dennis, and Brad R. Humphreys. 2006. "Proximity benefits and voting on stadium and arena subsidies." *Journal of Urban Economics*, 59(2): 285–299.

Conley, Timothy G. 1999. "GMM estimation with cross sectional dependence," *Journal of Econometrics*, 92(1): 1–45.

European Commission. 2003. "Commission Decision of 23 July 2003 relating to a proceeding pursuant to Article 81 of the EC Treaty and Article 53 of the EEA Agreement." 2003/778/EC. (COMP/C.2-37.398 — Joint selling of the commercial rights of the UEFA Champions League).

Eurostat. 1990-2011. "Occupancy in collective accommodation establishments: domestic and inbound tourism. Nights spent in tourist accommodation establishments by NUTS 2 regions - annual data." European Commission, Eurostat, Tourism statistics. http://epp.eurostat.ec.europa.eu/portal/page/portal/tourism/data/database (accessed February 25, 2013).

Eurostat. 1998-2010. "Air transport measurement - passengers. Detailed air passenger transport by reporting country and routes." European Commission, Eurostat, Transport statistics. http://epp.eurostat.ec.europa.eu/portal/page/portal/transport/data/database (accessed February 25, 2013).

Eurostat. 1999-2012. "Urban audit. Reduced set of derived indicators for 570 cities." European Commission, Eurostat, General and regional statistics. http://epp.eurostat.ec.europa.eu/portal/page/portal/region_cities/city_urban/data_cities/database_sub1 (accessed February 25, 2013).

Fafchamps, Marcel, and Flore Gubert. 2007a. "Risk Sharing and Network Formation," *American Economic Review Papers & Proceedings*, 97(2): 75–79.

Fafchamps, Marcel and Flore Gubert. 2007b. "The formation of risk sharing networks," *Journal of Development Economics*, 83(2): 326–350.

Moulton, Brent, R. 1986. "Random group effects and the precision of regression estimates," *Journal of Econometrics*, 32(3): 385–397.

Siegfried, John, and Andrew Zimbalist. 2000. "The Economics of Sports Facilities and Their Communities." *Journal of Economic Perspectives*, 14(3): 95–11.

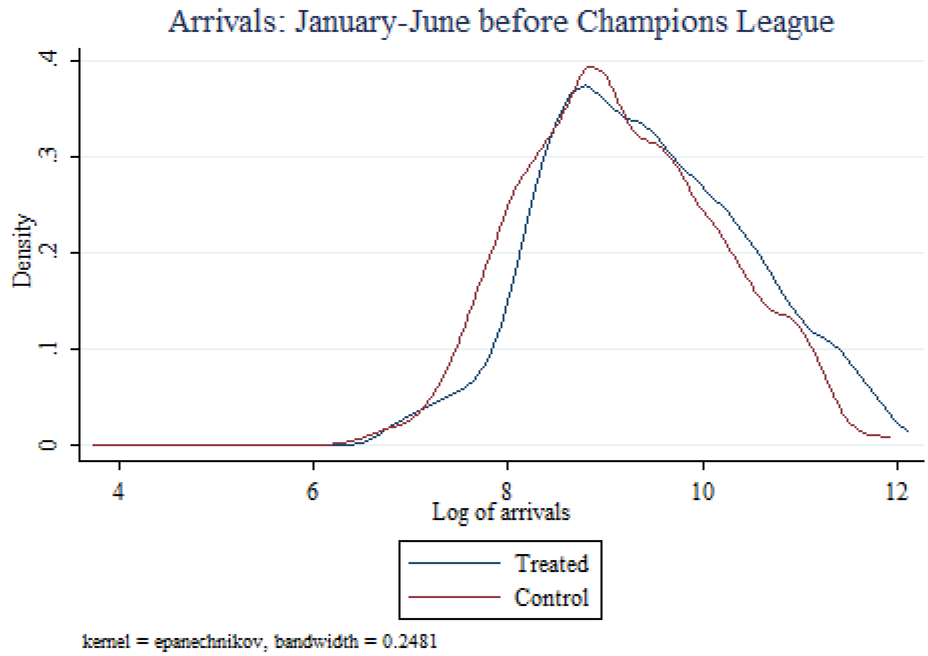

FIGURE 1. DISTRIBUTION OF ARRIVALS BETWEEN JANUARY AND JUNE ON ROUTES ACROSS CITIES TAKING PART TO THE CHAMPIONS LEAGUE GROUP PHASE BETWEEN 1998 AND 2010

Notes: There is one observation per route: number of arrivals is the average on the 2 directions. Treated routes have their team playing in the same group from September to December, control routes are routes across cities that could have met but eventually had their teams playing in different groups. I exclude all routes that: (i) could not have been treated given the seeding structure of the random draw and (ii) connect cities in Israel, Russia, Serbia, Turkey or Ukraine.

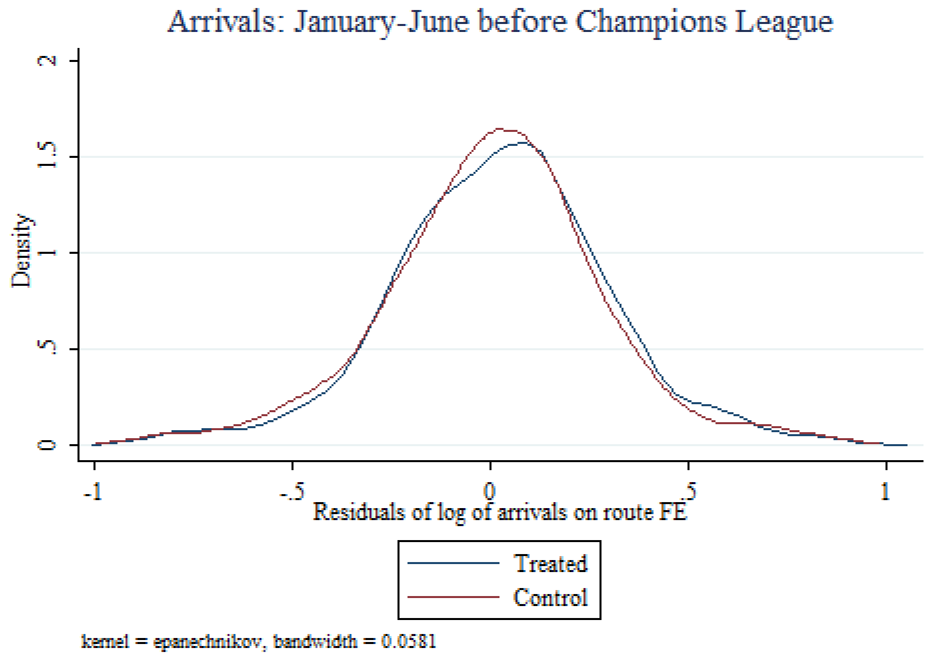

FIGURE 2. DISTRIBUTION OF RESIDUALS OF ARRIVALS BETWEEN JANUARY AND JUNE ON ROUTES ACROSS CITIES TAKING PART TO THE CHAMPIONS LEAGUE GROUP PHASE BETWEEN 1998 AND 2010 ON ROUTE FIXED EFFECTS

Notes: There is one observation per route, number of arrivals is the average on the 2 directions. Treated routes have their team playing in the same group from September to December, control routes are routes across cities that could have met but eventually had their teams playing in different groups. The figure plots the distribution of residuals from a regression on route fixed effects. I exclude all routes that: (i) could not have been treated given the seeding structure of the random draw and (ii) connect cities in Israel, Russia, Serbia, Turkey or Ukraine. The figure excludes observations with absolute change greater than 1 (for aesthetic reasons).

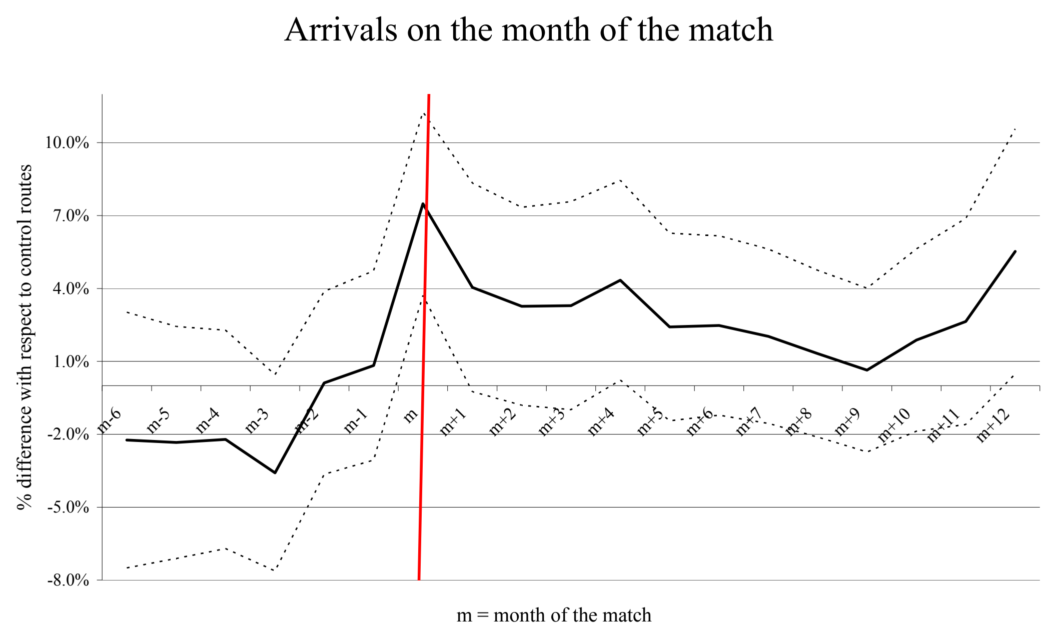

FIGURE 3. THE EFFECT ON THE MONTH OF THE MATCH

Notes: The figure plots estimates of γ_l from regression (2) ($l = -6,-5,\ldots,+12$). These coefficients represent the proportional difference in arrivals from city *j* to city *i* on treated routes relative to control routes. The coefficient in *m* (γ_0) is the effect on the month in which a group phase match is played. The figure also plots effects 6 months before the game (*m*–6) until 12 months after it (*m*+12). 95 percent confidence intervals calculated using standard errors clustered at the route-month level are reported around the estimates. The number of observations is 17922: these are all routes across cities that had at least 1 team taking part in the Champions League group phase either in the current or in the previous year, but excludes routes that: (i) could not have been treated given the seeding structure of the random draw; (ii) connect cities in Israel, Russia, Serbia, Turkey or Ukraine and (iii) had their teams met in the later stages of the competition either in the current or in the previous edition of the Champions League. The dependent variable ($\log P_{ij,m}$) has the top and bottom 0.5 percent of observations winsorized. Additional controls are year fixed effects, and trends specific to every month and to every country of origin and of destination. See text for details.

TABLE 1— BALANCE OF TREATED AND CONTROL ROUTES.

	Observations		Mean		
Variable	Treated	Control	Treated	Control	*p*-value (T=C)
PANEL A – Dependent variable (log *P*)					
Arrivals (January-June before group phase, logs)					
Average	162	1061	11.11	10.9	0.016**
Absolute difference	162	1061	6.6	6.36	0.008***
Route FE regression residuals (January-June before group phase, logs)					
Average	162	1061	0.01	0.00	0.761
Absolute difference	162	1061	0.01	0.00	0.696
Change in arrivals (January-June before group phase, logs)					
Average	137	933	6.80%	4.40%	0.230
Absolute difference	137	933	10.30%	2.50%	0.081*
PANEL B – Selection					
Routes with non missing air-traffic (September-December)					
	3544	23312	15.60%	14.90%	0.248
PANEL C - Tourism					
Touristic nights by residents (year of the match, logs)					
Average	112	720	15.73	15.66	0.286
Absolute difference	112	720	1.26	1.29	0.746
Touristic nights by residents (1 year before the match, logs)					
Average	109	703	15.7	15.63	0.351
Absolute difference	109	703	1.28	1.3	0.809
Change in touristic nights by residents (year of the match)					
Average	109	703	2.50%	2.20%	0.658
Absolute difference	109	703	9.90%	10.00%	0.964
Touristic nights by non residents (year of the match, logs)					
Average	112	720	15.89	15.85	0.621
Absolute difference	112	720	1.44	1.52	0.465
Touristic nights by non residents (1 year before the match, logs)					
Average	109	703	15.84	15.81	0.586
Absolute difference	109	703	1.41	1.53	0.311
Change in touristic nights by non residents (year of the match)					
Average	109	703	4.20%	3.70%	0.537
Absolute difference	109	703	6.50%	7.60%	0.138

Notes: The sample include all treated and control routes for which information on reported variable is available but excludes all routes to and from Israel, Russia, Serbia, Turkey and Ukraine.

Source: Author calculations.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

TABLE 1(CONTINUED) — BALANCE OF TREATED AND CONTROL ROUTES.

Variable		Observations: Treated	Observations: Control	Mean: Treated	Mean: Control	*p*-value (T=C)
PANEL D - Economy						
Income per capita (year of the match, logs)						
	Average	8	49	9.8	9.77	0.577
	Absolute difference	8	49	0.28	0.28	0.887
Income per capita (3 years before the match, logs)						
	Average	17	105	9.81	9.78	0.630
	Absolute difference	17	105	0.36	0.41	0.454
Income per capita growth (during the match)						
	Average	5	33	1.80%	0.40%	0.548
	Absolute difference	5	33	10.20%	13.10%	0.436
Unemployment rate (year of the match)						
	Average	54	294	9.10%	8.90%	0.636
	Absolute difference	54	294	5.20%	4.30%	0.135
Unemployment rate (3 years before the match)						
	Average	69	387	8.60%	8.30%	0.485
	Absolute difference	69	387	4.80%	4.40%	0.434
Change in unemployment rate (year of the match)						
	Average	35	190	1.10%	0.60%	0.323
	Absolute difference	35	190	5.50%	4.20%	0.106
PANEL E - Geography						
Distance (logs)		113	668	6.76	6.76	1.000
Cities are capital						
	Both	139	860	18.00%	17.20%	0.823
	Only 1	139	860	50.40%	46.50%	0.399
Cities are on an island (e.g. Great Britain, Cyprus)						
	Both	139	860	0.00%	0.50%	0.421
	Only 1	139	860	29.50%	31.40%	0.654
Cities are in a landlocked country						
	Both	139	860	0.00%	0.10%	0.688
	Only 1	139	860	10.10%	10.00%	0.979
Cities are in a Mediterranean country						
	Both	139	860	14.40%	11.30%	0.290
	Only 1	139	860	48.20%	47.60%	0.888

Notes: The sample include all treated and control routes for which information on reported variable is available but excludes all routes to and from Israel, Russia, Serbia, Turkey and Ukraine.

Source: Author calculations.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

TABLE 1(CONTINUED) — BALANCE OF TREATED AND CONTROL ROUTES.

Variable	Observations Treat.	Observations Control	Mean Treat.	Mean Control	*p*-value (T=C)
PANEL F – Demography					
Population (year of the game, logs)					
Average	95	525	14.11	14.07	0.546
Absolute difference	95	525	1.33	1.33	0.993
Population (3 year before the game, logs)					
Average	91	530	14.1	14.06	0.629
Absolute difference	91	530	1.49	1.43	0.643
Population growth (year of the match)					
Average	75	426	1.70%	2.00%	0.292
Absolute difference	75	426	3.80%	3.50%	0.516
Percent population aged 20 to 35 (year of the match)					
Average	81	475	23.40%	23.30%	0.706
Absolute difference	81	475	4.10%	4.70%	0.116
Percent population aged 20 to 35 (3 years before the match)					
Average	63	358	23.20%	23.30%	0.701
Absolute difference	63	358	3.80%	4.00%	0.538
Change in percent population aged 20 to 35 (year of the match)					
Average	46	269	-0.70%	-0.70%	0.976
Absolute difference	46	269	1.30%	1.40%	0.577
Percent non national, EU residents (year of the match)					
Average	40	205	4.60%	4.40%	0.603
Absolute difference	40	205	5.70%	5.30%	0.644
Percent non national, EU residents (3 years before the match)					
Average	45	225	4.00%	3.60%	0.371
Absolute difference	45	225	5.00%	4.60%	0.636
Change in percent non national, EU residents (year of the match)					
Average	28	153	0.90%	0.70%	0.313
Absolute difference	28	153	1.00%	0.70%	0.063*
Cities speak a romance language (e.g. Italian, Spanish,...)					
Both	139	860	31.70%	26.20%	0.176
Only 1	139	860	41.70%	41.50%	0.962
Cities speak a germanic language (e.g. German, English,...)					
Both	139	860	11.50%	15.30%	0.237
Only 1	139	860	51.80%	43.30%	0.060*

Notes: The sample include all treated and control routes for which information on reported variable is available but excludes all routes to and from Israel, Russia, Serbia, Turkey and Ukraine.

Source: Author calculations.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

TABLE 2 — EFFECT OF PLAYING IN THE SAME GROUP OF THE CHAMPIONS LEAGUE DURING THE FOLLOWING MONTHS.

	β	s.e.	Obs.	Route FE	Year FE & country trends	Month FE & month trends
Month of the match [a]	0.067***	(0.016)	6136	Yes	Yes	Yes
January	0.084**	(0.035)	1344	Yes	Yes	No
February	0.064**	(0.032)	1346	Yes	Yes	No
March [b]	0.062**	(0.028)	1340	Yes	Yes	No
April [b]	0.045	(0.030)	1358	Yes	Yes	No
May [b]	0.042	(0.030)	1354	Yes	Yes	No
June	0.030	(0.023)	1584	Yes	Yes	No
July	0.033	(0.024)	1584	Yes	Yes	No
August	0.019	(0.025)	1584	Yes	Yes	No

Notes: The table reports estimates of β in equation (3). The sample includes all routes across cities that had at least 1 team taking part in the Champions League group phase during the current edition. I exclude all routes that: (i) could not have been treated given the seeding structure of the random draw; (ii) connect cities in Israel, Russia, Serbia, Turkey or Ukraine and (iii) had their teams met in the later stages of the competition either in the current or in the previous edition of the Champions League. The dependent variable (log $P_{ij,m}$) has the top and bottom 0.5 percent of observations winsorized. Standard errors in parentheses are clustered at the route-month level. See text for details.

Source: Author calculations.

[a] September through December. [b] Knock-out phase.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

TABLE 3 — ROBUSTNESS FOR THE EFFECT OF PLAYING IN THE SAME GROUP OF THE CHAMPIONS LEAGUE.

PANEL A – Data collapsed at route-month level						
	β	s.e.	Obs.	Route FE	Year FE & country trends	Month FE & month trends
Month of the match[a]	0.063***	(0.016)	3068	Yes	Yes	Yes
January	0.084**	(0.041)	672	Yes	Yes	No
February	0.064*	(0.038)	673	Yes	Yes	No
March[b]	0.063*	(0.034)	670	Yes	Yes	No
April[b]	0.045	(0.035)	679	Yes	Yes	No
May[b]	0.042	(0.036)	677	Yes	Yes	No
June	0.03	(0.027)	792	Yes	Yes	No
July	0.033	(0.028)	792	Yes	Yes	No
August	0.019	(0.029)	792	Yes	Yes	No

PANEL B – Dyadic standard errors						
	β	s.e.	Obs.	Route FE	Year FE & country trends	Month FE & month trends
Month of the match[a]	0.067***	(0.016)	6136	Yes	Yes	Yes
January	0.084***	(0.031)	1344	Yes	Yes	No
February	0.064**	(0.029)	1346	Yes	Yes	No
March[b]	0.062**	(0.025)	1340	Yes	Yes	No
April[b]	0.045	(0.031)	1358	Yes	Yes	No
May[b]	0.042	(0.030)	1354	Yes	Yes	No
June	0.03	(0.022)	1584	Yes	Yes	No
July	0.033	(0.023)	1584	Yes	Yes	No
August	0.019	(0.024)	1584	Yes	Yes	No

Notes: The table reports estimates of β in equation (3). The sample includes all routes across cities that had at least 1 team taking part in the Champions League group phase during the current edition. I exclude all routes that: (i) could not have been treated given the seeding structure of the random draw; (ii) connect cities in Israel, Russia, Serbia, Turkey or Ukraine and (iii) had their teams met in the later stages of the competition either in the current or in the previous edition of the Champions League. Panel A reports estimates when observations are collapsed at the route-month level: standard error here are corrected for heteroschedasticity. Panel B reports the same estimates of table 2, but standard errors are dyadic. In both panels the dependent variable has the top and bottom 0.5 percent of observations winsorized. See text for details.

Source: Author calculations.

[a] September through December. [b] Knock-out phase.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

TABLE 4 — RESULTS FROM 1000 PLACEBO SIMULATIONS OF THE TREATMENT.

	Simulations with p-value smaller than 0.05	*95th percentile of estimates in simulations*	β [a]	*Simulations with estimates larger than* β [a]
	(1)	(2)	(3)	(4)
Month of the match [b]	8.6 %	0.038	0.067	0.1%
January	5.2 %	0.068	0.084	1.5%
February	6.2 %	0.067	0.064	5.7%
March [c]	6.3 %	0.061	0.062	4.6%
April [c]	6.2 %	0.051	0.045	6.9%
May [c]	5.5 %	0.048	0.042	8.4%
June	4.9 %	0.039	0.030	10.3%
July	4.1 %	0.037	0.033	6.7%
August	4.4 %	0.039	0.019	22.5%

Notes: Column (1) reports the percentage of simulations in which the effect of a placebo treatment was estimated to be different from 0 at the 5 percent confidence level. The placebo treatment is assigned according to the same rules used to form the Champions League groups; in each simulation I ran a regression identical to (3) substituting the true treatment G_{ij} with this placebo treatment randomly assigned. The number of simulations on which these statistics are computed is 1000. Column (2) reports the 95th percentile of the β estimated on the 1000 simulations of the placebo treatment. Column (3) reports the estimates of β from table 2. Column (4) reports the percentage of simulation with an estimated β larger than the one reported in column 2. See text for details.

Source: Author calculations.

[a] From table 2. [b] September through December. [c] Knock-out phase.