Using Soccer Games as an Instrument to Forecast the Spread of COVID-19 in Europe[*]

Juan-Pedro Gómez[¤]

Maxim Mironov[§]

February 2021

Forthcoming at the *Finance Research Letters*

Abstract

We provide strong empirical support for the contribution of soccer games held in Europe to the spread of the COVID-19 virus in March 2020. We analyze more than 1,000 games across 194 regions from 10 European countries. Daily cases of COVID-19 grow significantly faster in regions where at least one soccer game took place two weeks earlier, consistent with the existence of an incubation period. These results weaken as we include stadiums with smaller capacity. We discuss the relevance of these variables as instruments for the identification of the causal effect of COVID-19 on firms, the economy, and financial markets.

JEL Codes: I18, C36, G01

Keywords: COVID-19, soccer, super-spreaders, instrumental variables, identification strategy

Declarations of interest: none

---

[*] © <2021>. This manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/ We thank Antonio De Vito, Rüdiger Fahlenbrach, Garen Markarian, Kevin Rageth, and Pablo Ruiz-Verdú, and an anonymous referee for their help and comments. We also thank seminar participants in the 2020 MadBar Conference and the Solbridge International Business School seminar series. The usual caveat applies. Research Reported in this paper was partially funded by the Spanish Ministry of Economy and Competitiveness (MCIU), State Research Agency (AEI) and European Regional Development Fund (ERDF) Grant No. PGC2018-101745-A-I00.

[¤] IE Business School, IE University. María de Molina 12, 28006 Madrid, Spain. Email: juanp.gomez@ie.edu

[§] IE Business School, IE University. María de Molina 12, 28006 Madrid, Spain. Email: Maxim.Mironov@ie.edu

1. Introduction

There is anecdotal evidence that soccer games have contributed to the spread of the COVID-19 pandemic in Europe.[1] In this paper, we provide strong empirical support for this conjecture and discuss the implications of our findings for the identification of the causal impact of COVID-19 on firms, the economy, and financial markets.

Although it makes sense to assume that the original outbreak of the pandemic in China at the end of 2019 is exogenous, this becomes a more questionable assumption for the propagation of cases across countries and regions in Europe during the first quarter of 2020. For instance, the *uninstrumented* number of cases, especially at the beginning of pandemic, is likely to overestimate the incidence of COVID-19 in well-connected versus remote cities.[2] Similarly, cities and regions with more inhabitants and higher population density are likely to experience faster virus spread (Rocklöv and Sjödin (2020)). On the one side, these regions tend to accumulate a higher percentage of firms and human capital, hence making any correlation between the number of cases and firm variables (like productivity, growth, solvency, or liquidity) potentially spurious. On the other side, these regions are likely to concentrate more economic and medical resources to detect and counterattack the pandemic. Thus, the *raw* number of COVID-19 cases might capture the inverse quality of the regional health system, which is likely correlated with firm performance and regional growth.

To overcome these endogeneity issues, we propose four variables related to soccer games played across European regions from 10 countries during the first quarter of 2020. These variables constitute a novel and valuable instrument to explore the causal effect of COVID-19 infections on firm performance, management decisions, and the economy. Methodologically, the *exclusion restriction* is well founded. National leagues and pan-European tournaments, like the UEFA Champions and Europa leagues, were scheduled well before the original outbreaks of COVID-19 in China. Although there is evidence of the behavioral impact of victories and losses of soccer matches on stock returns (e.g., Edmans, García, and Norli (2007)), our soccer-related instruments are independent from the game's output. As far as we know, there is neither theory nor evidence that links *directly* the number of attendants to a soccer match or the capacity of the venue where it is played with, for instance, stock returns, cash holdings, or dividends of firms headquartered in the region, or, alternatively, growth in regional product or unemployment. Theoretically, the physical interaction among spectators in large venues as well as their arrival and departure from stadiums increase the likelihood of being infected with the virus, ultimately working as "super-spreader" events. The evidence in this paper is consistent with this conjecture and offers solid support for the *relevance* of these instruments to predict the spread of COVID-19 cases across European regions.

We collect data from soccer games from all competitions (domestic and international) played in 194 regions across Belgium, France, Italy, Germany, the Netherlands, Poland, Spain, Sweden, Switzerland, and the UK, between January 1 and until the end of March 2020 (most games in Europe were canceled after March 10). In our main analysis, we include games played in venues with a minimum capacity of 25,000 people. In total, there are 1,051 qualifying games during this period.[3] We also collect the

---

[1] "The first three S viruses identified in Spain are from samples taken on February 26 and 27 in Valencia. A week before, 2,500 soccer fans from the region had traveled to Milan to see Atalanta play Valencia, an event that was described as a `biological bomb' by the mayor of Bergamo, Giorgio Gori." El País, April 23, 2020.

[2] There is evidence that regions with international airports and hubs are more likely to be affected first and more severely by the virus (Paraskevas and Dimitriou (2020)).

[3] For robustness, we also collet data from games that took place in stadiums with a minimum capacity of 10,000 spectators, increasing the sample up to 2,314 matches.

confirmed cases of COVID-19 in these regions until the end of March, plus three economic and demographic variables: gross regional product, population, and density.

We construct four variables related to the soccer matches. Namely: a dummy variable that takes a value of one if there was a soccer game in the region, zero otherwise; a variable that accumulates the number of games played in the region; a variable that accumulates all the spectators who attended those games; and a variable that accumulates the capacity (maximum number of spectators) of the venues where the matches took place.

We document the following findings. First, for any single country and day from March 1 through 14, the rate of change in the number of COVID-19 cases relative to the previous day is, on average, 5.5 percentage points higher in regions where there was at least one soccer game two weeks earlier relative to regions with no games in the same period (as reference, the average rate of change is 23% per day during this period). Additionally, the daily increment of cases is, on average, about 6 basis points higher for every 1% increase in the attendance and venue capacity of games played two weeks earlier. These results are significant at the 1% level, and robust to the inclusion of regional demographic and economic control variables known to affect the virus spread (e.g., Rocklöv and Sjödin (2020)). Second, games celebrated, either the previous week or earlier than 2 weeks before, have no significant effect in the increment of daily cases. This is consistent with the incubation period and the lack of massive testing in the early stages of the pandemic.[4] Third, as we expand the sample to include games celebrated in venues with smaller capacity, the statistical significance of the coefficient on the three soccer-related variables decreases, turning non-significantly different from zero when we include stadiums with a minimum capacity above 10,000 spectators. This evidence is consistent with the effect of "super-spreaders" of the virus documented in other large events (e.g., Dave et al (2020) and Felbermayr, Hinz, and Chowdhry (2020)). Fourth, the games played by a (local) team of a given region in another region have no significant effect on the number of cases in the local region, regardless of the game attendance or the venue capacity. Thus, there is no evidence that soccer fans moving to other regions or people gathering in bars in the local region to watch the game have contributed significantly to the spread of the virus.

The rest of the paper is organized as follows. Section 2 describes the data. Results are presented in Section 3. We discuss the limitations of the analysis in Section 4, before concluding with Section 5. Our variables and their sources are described in the Appendix.

2. Data

Our sample consists of 2,162 region-day observations.[5] We collect the accumulated number of diagnosed cases of COVID-19 per day and region from day 1 through 14 of March 2020, in 194 regions from Belgium, France, Italy, Germany, the Netherlands, Poland, Spain, Sweden, Switzerland, and the UK.[6] We call this variable *Cases*.[7] Panel A of Table 1 shows that, on average, there are 96 accumulated cases per day and region with an average of 35 accumulated cases per million regional inhabitants and day (variable *Cases/Population*).

Then, we collect data from soccer games from all competitions (domestic and international) played in the 194 regions between January 1 and until the end of March 2020 (most games in Europe were canceled after March 10). Originally, we only include games played in venues with a minimum capacity

---

[4] "Coronavirus disease 2019 (COVID-19) Situation Report – 73," WHO, April 2, 2020.

[5] Data on COVID-19 cases from Poland start on March 4, from Switzerland on March 6, and from England on March 9.

[6] We are unable to obtain regional data of COVID-19 cases from Northern Ireland, Scotland, or Wales. Hence, only English regions are considered.

[7] Table A in the Appendix shows the exact definition and source for each variable.

of 25,000 people. In total, there are 1,051 qualifying games during the sample period. From each game, we collect date, playing teams, attendance (when available), venue capacity, and the region and country where it is located. Finally, we also collect the following demographic variables from each region: *Population*, *Density*, and *Gross Regional Product (GRP)* per capita.

First, we want to explore if there is a pattern in the relation between the attendance to these events and the propagation of the virus. Every day, from March 1 through 14, we calculate the number of matches (*# Games*), *Attendance* and venue *Capacity* that took place in each region 1, 2,…, and up 30 days before. Figure 1 plots the average value of each variable across the 14 days and 194 regions for each day lag. Notice that game attendance and venue capacity are highly correlated across lags (correlation coefficient 0.98). The average match attendance is about 60% of venue capacity and this percentage is very stable across lags. The figure shows periodic spikes around 7, 21, and 28-day lags for the 3 variables. Considering that the first day of our sample is Sunday, March 1, these spikes reflect the higher concentration of soccer matches on weekends (70% of soccer matches take place on weekends). Figure 2 confirms this by plotting the number of soccer games across all regions in our sample, from January 14 through March 14. In the horizontal axis, we include Saturdays. We can see that a disproportionate number of games fall on Saturday or Sunday.

[Insert Figure 1 about here]

[Insert Figure 2 about here]

Thus, in order to smooth out the effect of weekends, we accumulate games, attendance and venue capacity over weekly windows. For every region in our sample and for every day from March 1 through 14, we estimate the number of soccer matches, the accumulated attendance, and the accumulated venue capacity 1, 2…, and up to 6 weeks earlier. We also calculate the variable *I_Games* that takes a value of 1 if there was at least one soccer match in the region during a given week, zero otherwise.

Table 1, Panel A reports the statistics accumulated over the 6 weeks window. From March 1 through 14, on average, there were games in 44% percent of the regions over the previous 6 weeks. Additionally, for every day and region, there were on average 3.29 games accumulated over the previous 6 weeks, attended by an average of 78,953 (accumulated) people and played in venues with an average (accumulated) capacity of 136,092 spectators. Table B in the Appendix includes a list of all regions, with the accumulated number of cases, games, attendance and venue capacity in our sample.[8]

[Insert Table 1 about here]

Panel B of Table 1 presents the average of each variable across the 14 sample days and 194 regions for each week lag. Except for the first week,[9] the estimates are very similar across weeks. On average, across weeks 2 through 6, 33% of the regions celebrated at least one soccer match per week. There

---

[8] There are 112 regions with no qualifying games (i.e., played in venues with minimum capacity of 25,000 spectators) during the sample period. Thus, the median value of the three variables in Table 1 is zero.

[9] Games were canceled throughout Europe around March 10. Thus, the variable estimates from March 11 through 14 over the first week-lag are smaller than the corresponding estimates for weeks 2 through 6.

were 0.55 games per week and region, attended by 13,192 people and played in venues with average capacity for about 22,590 spectators.[10]

3. Results

We proceed now to analyze the relation between, on the one side, the number, attendance, and venue capacity of the soccer games celebrated until all competitions were interrupted, and, on the other, the propagation of COVID-19 cases across days and regions during the first two weeks in March 2020.

There is evidence that the incubation period of COVID-19 (that is, the "pre-symptomatic" period between becoming infected and developing symptoms of the disease) can be as long as two weeks. Thus, there is likely a lag between the time when the match spectators become infected and the time they are tested after developing symptoms compatible with the disease. This is especially relevant in the first two weeks of March 2020 when mass testing (in particular, across asymptomatic people) had not been yet implemented in any country. Figure 3 shows that by March 15, all countries in our sample, except Switzerland and (marginally) Germany, had a ratio of COVID-19 tests per thousand people below 0.2. Most likely, at the onset of the pandemic, only people with symptoms were tested and, eventually, diagnosed as new cases of COVID-19 infections. Therefore, considering the incubation window and that only symptomatic people were tested at that point, we expect the predictive power of our instruments to become significant with a lag after the game.

[Insert Figure 3 about here]

To test this prediction, we run the following panel regression in region *r* and day *t* from March 1 through 14, 2020:[11]

$$\begin{aligned}\Delta Log\left(1 + Cases_{r,t}\right) \\ &= a + b_1 \Delta Log\left(1 + Cases_{r,t-1}\right) + b_2 Log(Population_r) \\ &+ b_3 Log(Density_r) + b_4 Log(GRP_r) + \sum_{w=1}^{6} c_w WX_{r,t-w} \\ &+ FE_{c\times t} + \epsilon_{r,t}.\end{aligned} \tag{1}$$

$\Delta Log\left(1 + Cases_{r,t}\right)$ represents the (log) difference between 1 plus the number of cases in region *r* and day *t* and day *t-1*. Likewise, $\Delta Log\left(1 + Cases_{r,t-1}\right)$ is the same variable lagged 1 day. For every lagged week $w = \{1,2,\dots,6\}$ and region *r*, the variable $WX_{r,t-w}$ represents, alternatively, the dummy variable, I_Games_{t-w}, that takes a value of one if there was a soccer match in the region any day $t \in (t-(1+7\times(w-1)), t-7\times w)$; the natural logarithm of 1 plus the accumulated number of match attendants over the week, $Log(1 + Attendance_{t-(1+7\times(w-1))} - Attendance_{t-7\times w})$; and the natural logarithm of 1 plus the accumulated venue capacity of the games played over the week, $Log(1 + Capacity_{t-(1+7\times(w-1))} - Capacity_{t-7\times w})$. We control for each region's population,

---

[10] If the region did not have any games, the capacity is zero. Thus, the average capacity is below 25,000, the minimum required stadium capacity to be included in the sample.

[11] In other to keep all observations, we add 1 to *Cases* since otherwise the logarithm of zero is not defined.

density and gross regional product per capita (GRP). Our object of interest is the series of coefficients on the weekly lagged predictors, $c_{w=\{1,2,\ldots,6\}}$. $FE_{c\times t}$ represents country times day fixed effects. All variables are defined in Appendix A. Standard errors are clustered at the region level.

Table 2 presents the results from regression (1) for the three soccer variables. The rate at which the daily number of cases of COVID-19 increases is positive and significantly related to the increase of cases the previous day. It is also higher in more populated and wealthier (higher Log(*GRP*)) areas. With respect to the lagged soccer variables, only the coefficient c_2 corresponding to *I_Games*, Log(*Attendance*), or Log(*Capacity*) two weeks earlier is significant. The other lags are non-significant for any of the three variables. In specification (1), for any single country and day from March 1 through 14, the rate of change in the number of COVID-19 cases relative to the previous day is, on average, 5.5 percentage points higher in regions where there was a soccer game two weeks earlier relative to regions with no games in the same period. This result is statistically significant at the 1% level as well as economically significant (the average growth rate of cases was 23% per day during this period). Specifications (2) and (3) show that the rate of change is, on average, about 6 basis points higher for every 1% increase in attendance and venue capacity, respectively. Both results are significant at the 1% level. These results are consistent with the documented incubation period of the virus and the lack of massive testing during the sample period.

[Insert Table 2 about here]

Finally, we test if our results change when we include venues with smaller minimum capacity. There is evidence of the role played by large gatherings of people in the dissemination of the virus. These are known as “super-spreader” events (e.g., Dave et al (2020) and Felbermayr, Hinz, and Chowdhry (2020)). To test the importance of the minimum venue capacity, we expand the sample to include games that took place in venues with a minimum capacity of 10,000 spectators. The extended sample includes 2,314 games.

Table 3 presents the results of regression (1) when we consider games held in venues with a minimum capacity of 20, 15 and 10 thousand spectators, respectively, for each of the three soccer variables. Like in Table 2, the daily increment in the number of cases of COVID-19 is positive and significantly related to the increase of cases the previous day. It is also higher in more populated and wealthier (higher Log(*GRP*)) areas. When we include stadiums with a minimum capacity of 20,000 spectators, the rate of change in the number of COVID-19 cases relative to the previous day is, on average, higher by 4.2 percentage points in regions where there was a soccer game two weeks earlier relative to regions with no games in the same period. This is lower than the 5.5% difference in Table 2. The coefficient is significant at the 5% level (down from 1% in Table 2). The *Attendance* and *Capacity* variables show a similar qualitative pattern. However, when we expand the minimum capacity to 15,000 spectators, the coefficient is not statistically different from zero for any of the three variables (only marginally at the 10% for *Attendance*). These results are confirmed when the minimum capacity is lowered to 10,000 spectators.

[Insert Table 3 about here]

We interpret these results as consistent with the evidence of other super-spread events. A minimum agglomeration is needed for the spread of the virus to be statistically detectable.

4. Limitations of the analysis

In this section, we discuss some limitations of our analysis. In the first place, our regressions only explain, on average, 18% of the change in daily cases. Thus, the coefficients on the soccer variables should be interpreted in a cross-sectional way: they help explain differences in the incidence of COVID-19 across regions in the early stages of the pandemic, rather than the absolute numbers of contagions within each region. Furthermore, relative to our sample period, people's awareness has increased and governments around the world have taken measures to promote public hygiene and social distancing. Currently, we would expect any public gathering or mass event to result in much lower COVID-19 spreading. For this reason, using soccer games as an instrument variable is only applicable during the outbreak of the pandemic across Europe in March. This limitation is shared by other studies based on large gatherings, like motorcycle rallies and ski resorts, mentioned in the Introduction. Unlike these events, however, soccer competitions have two advantages as an instrument. First, they take place across several countries, hence expanding the sample size considerably. Second, the games are staggered through the first quarter of 2020, in contrast with other mass events like Carnival celebrations, which take place rather simultaneously across Europe in the same period.

Finally, another limitation is that people might have also caught the corona virus in bars where soccer matches were broadcasted, without being physically present in the match venue. To assess the impact of this *indirect* via of contagion, we perform the following exercise. For every game in our sample, we replicate Table 2 but considering the spread of cases in the region when a local team plays outside the region. In this case, we might expect an increase of bar attendance in the region of the local team but not mass gathering of people as we predict in the region where the game is actually played.[12] That is, in regression (1), for every lagged week w={1,2,…,6} and region *r*, the variable $WX_{r,t-w}$ now represents, alternatively, the dummy variable, I_Games_{t-w}, that takes a value of one if there was a soccer match in which a team from region *r* played outside that region any day $t \in (t-(1+7\times(w-1)), t-7\times w)$; the natural logarithm of 1 plus the accumulated number of match attendants to those games, $Log(1+Attendance_{t-(1+7\times(w-1))}-Attendance_{t-7\times w})$, or the natural logarithm of 1 plus the accumulated venue capacity of those games, $Log(1+Capacity_{t-(1+7\times(w-1))}-Capacity_{t-7\times w})$. We include the same set of controls as in equation (1). Standard errors are clustered at the region level.

Results are reported in Table 4. Even accounting for the impact of cross-border movements of fans, the celebration of any game where a local team plays outside the region has no significant effect on the virus spread in the region, regardless of the venue attendance or capacity.

[Insert Table 4 about here]

5. Conclusions and implications

---

[12] Arguably, this is not a perfect experiment since fans of a local team from a given region might have travelled to attend the game when the team plays in another region, later spreading the virus at home (see footnote 1). The number of local fans travelling to another region is likely to increase with the game attendance and the venue capacity. We cannot disentangle this effect from the virus spread from bar attendants in the local region.

The evidence about the soccer variables introduced in this paper may help overcome potential endogeneity issues in the analysis of how the spread of COVID-19 has affected the economy and firm decisions. Despite the limited time span (March 2020) of these variables, the impact of the COVID-19 pandemic is so deep and unprecedented, that we believe this analysis is relevant. Gómez and Mironov (2020), for instance, show that, only after instrumenting the number of COVID-19 cases with the soccer variables, there is evidence of a causal relation between the propagation of the virus and the cross-section of stock returns from firms headquartered in these regions. The accumulated drop in stock performance during March and April 2020 is significantly higher for firms in regions with higher incidence of (instrumented) COVID-19 cases only for companies whose CEOs older than 60 years. The evidence shows that older people are more likely to suffer from severe illness or even death in case of contagion. Thus, the market is discounting the likelihood of the company's CEO possibly dying of COVID-19. These instruments could also be used to analyze the causal effect of the virus on the drop in regional gross product or employment, or corporate variables like revenue, cash holdings, dividends, investments, inventories, and accounts payable, as more data becomes available.

References

Dave, Dhaval M., Andrew I. Friedson, Drew McNichols, and Joseph J. Sabia, 2020, The Contagion Externality of a Superspreading Event: The Sturgis Motorcycle Rally and COVID-19, NBER Working Paper No. 27813.

Edmans, A., D. García, and Ø. Norli, 2007. Sports Sentiment and Stock Returns. The Journal of Finance 62, 1967-1998.

Felbermayr, Gabriel , Julian Hinz, and Sonali Chowdhry, 2020, Après-ski: The Spread of Coronavirus from Ischgl through Germany, Working Paper.

Gomez, Juan-Pedro and Mironov, Maxim, COVID-19 and the Value of CEOs: The Unintended Effect of Soccer Games across European Stocks (July 7, 2020). Available at SSRN: https://ssrn.com/abstract=3645401

Paraskevas Nikolaou, and Loukas Dimitriou, 2020. Identification of critical airports for controlling global infectious disease outbreaks: Stress-tests focusing in Europe, Journal of Air Transport Management 85.

Rocklöv, Joacim, and Henrik Sjödin, 2020. High population densities catalyze the spread of COVID-19, Journal of Travel Medicine 27.

Table 1
Summary Statistics for the Sample of Region-Days

In Panel A, each observation is a duple region-day. Every day from March 1 through March 14, 2020, *Cases* is the accumulated number of diagnosed cases of COVID-19 in the region during that period. *Cases/Population* is the number of cases per million inhabitants. We consider all regions in Belgium, France, Italy, Germany, the Netherlands, Poland, Spain, Sweden, Switzerland, and the UK. The distribution of observations across regions is in Table B.1 of Appendix B. Every day from March 1 through March 14, *# Games, Attendance*, and *Capacity* is the accumulated number of soccer matches played in the region, their attendance, and the venue capacity, respectively, over the previous 6 weeks. *I_Games* is a dummy variable that takes a value of 1 if there was at least one soccer match in the region where the firm is located during the previous 6 weeks, zero otherwise. *Population* is thousands of inhabitants per region; *Density* is number of inhabitants per square-Km; *GRP* is the Gross Regional Product per capita in USD. Log (*x*) denotes the natural logarithm of *x*. Δ Log(1+x_t)=Log((1+x_t)/(1+x_{t-1})). In Panel B, we report the average across regions of the weekly accumulated number of games, attendance and venue capacity for up to 6 weekly lags. Table A in the Appendix includes the definition and source of each variable.

Panel A. Accumulated variables per day and region

	Mean (1)	Median (2)	St. dev. (3)	# Regions (4)	# Obs. (5)
Cases	96	8	507	194	2,162
Cases/Population	35	7	87	194	2,162
Log(1+*Cases*)	2.434	2.197	1.902	194	2,162
Δ Log(1+*Cases*)	0.228	0.152	0.286	194	2,073
Log(1+*Cases/Population*)	-11.486	-11.554	1.643	194	2,162
# Games	3.296	0	5.162	194	2,162
I_Games	0.444	0	0.497	194	2,162
Attendance	78,953	0	162,921	194	2,162
Capacity	136,092	0	244,261	194	2,162
Log(1+*Attendance*)	5.115	0	5.826	194	2,162
Log(1+*Capacity*)	5.481	0	6.149	194	2,162
Population, 000	2,287	1,199	2,782	194	2,162
Density	451	160	1,046	194	2,162
GRP	37,428	35,240	14,728	194	2,162
Log*(Population*)	13.920	13.997	1.344	194	2,162
Log*(Density*)	5.091	5.081	1.327	194	2,162
Log*(GRP*)	10.464	10.470	0.359	194	2,162

Panel B. Statistics by Weekly Lags

Weeks ago	*# Games* (1)	*I_Games* (2)	*Attendance* (3)	Log (1+*Attendance*) (4)	*Capacity* (5)	Log (1+*Capacity*) (6)
1	0.476	0.299	11,320	2.661	19,374	3.251
2	0.547	0.327	13,192	3.101	22,594	3.560
3	0.587	0.339	13,919	3.277	24,276	3.703
4	0.551	0.328	13,143	3.206	22,436	3.583
5	0.558	0.342	13,574	3.306	23,328	3.738
6	0.576	0.321	13,805	3.138	24,083	3.525

Table 2
Regression of Change in *Cases* on Weekly Lagged *Games*, *Attendance* and *Capacity*

This table reports the coefficients from the following regression:

$$\Delta Log(\ 1 + Cases_{r,t}\ ) = a + b_1 \Delta Log(\ 1 + Cases_{r,t-1}\ ) + b_2 Log(Population_r) + b_3 Log(Density_r) + b_4 Log(GRP_r) + \sum_{w=1}^{6} c_w WX_{r,t-w} + FE_{c\times t} + \epsilon_{r,t}.$$

$\Delta Log(\ 1 + Cases_{r,t}\ )$ represents (log) difference between 1 plus the number of cases in region *r* and day *t* with respect to day *t-1*. Likewise, $\Delta Log(\ 1 + Cases_{r,t-1}\ )$ is the same variable lagged 1 day. For every lagged week w={1,2,…,6} and region *r*, the variable $WX_{r,t-w}$ represents, alternatively, the dummy variable, I_Games_{t-w}, that takes a value of one if there was a soccer match in the region any day $t \in (t-(1+7\times(w-1)), t-7\times w)$; the natural logarithm of 1 plus the accumulated number of match attendants over the week, $Log(1 + Attendance_{t-(1+7\times(w-1))} - Attendance_{t-7\times w})$, or the natural logarithm of 1 plus the accumulated venue capacity over the week, $Log(1 + Capacity_{t-(1+7\times(w-1))} - Capacity_{t-7\times w})$. We control for each region's *Population*, *Density* and Gross Regional Product per capita (*GRP*). $FE_{c\times t}$ Represents country times day fixed effects. Appendix A includes the definition and source of each variable. Standard errors (in parenthesis) are clustered at the region level. ***, **, * represent statistical significance at the 1, 5, and 10% level, respectively.

	I_Games (1)	Log(1+*Attendance*) (2)	Log(1+*Capacity*) (3)
$\Delta Log(\ 1 + Cases_{t-1}\ )$	0.056	0.056	0.055
	(0.028)**	(0.028)**	(0.028)**
Log(*Population*)	0.027	0.027	0.028
	(0.007)***	(0.007)***	(0.007)***
Log(*Density*)	-0.002	-0.002	-0.002
	(0.006)	(0.006)	(0.006)
Log(*GRP*)	0.049	0.049	0.048
	(0.025)**	(0.025)*	(0.025)**
Lagged week 1 (c_1)	-0.028	0.000	-0.003
	(0.021)	(0.002)	(0.002)
Lagged week 2 (c_2)	0.055	0.006	0.005
	(0.02)***	(0.002)***	(0.002)***
Lagged week 3 (c_3)	-0.016	-0.003	-0.001
	(0.025)	(0.002)	(0.002)
Lagged week 4 (c_4)	-0.015	-0.001	-0.001
	(0.02)	(0.002)	(0.002)
Lagged week 5 (c_5)	-0.004	-0.001	0.000
	(0.022)	(0.002)	(0.002)
Lagged week 6 (c_6)	-0.012	-0.002	-0.002
	(0.022)	(0.002)	(0.002)
Country × Day FE	Y	Y	Y
R-sq	0.180	0.180	0.181
Number of Obs.	2,073	2,073	2,073
Number of Regions	194	194	194

Table 3
Regression of Change in *Cases* on Weekly Lagged *Games*, *Attendance* and *Capacity*
Sorted by minimum venue Capacity (below 25K spectators)

This table reports the coefficients from the following regression:

$$\Delta Log(\,1 + Cases_{r,t}\,) = a + b_1 \Delta Log(\,1 + Cases_{r,t-1}\,) + b_2 Log(Population_r) + b_3 Log(Density_r) + b_4 Log(GRP_r) + \sum_{w=1}^{6} c_w WX_{r,t-w} + FE_{c\times t} + \epsilon_{r,t}.$$

$\Delta Log(\,1 + Cases_{r,t}\,)$ represents (log) difference between 1 plus the number of cases in region *r* and day *t* with respect to day *t-1*. Likewise, $\Delta Log(\,1 + Cases_{r,t-1}\,)$ is the same variable lagged 1 day. For every lagged week w={1,2,…,6} and region *r*, the variable $WX_{r,t-w}$ represents, alternatively, the dummy variable, I_Games_{t-w}, that takes a value of one if there was a soccer match in the region any day $t \in (t-(1+7\times(w-1), t-7\times w)$; the natural logarithm of 1 plus the accumulated number of match attendants over the week, $Log(1 + Attendance_{t-(1+7\times(w-1))} - Attendance_{t-7\times w})$, or the natural logarithm of 1 plus the accumulated venue capacity over the week, $Log(1 + Capacity_{t-(1+7\times(w-1))} - Capacity_{t-7\times w})$. We control for each region's *Population*, *Density* and Gross Regional Product per capita (*GRP*). $FE_{c\times t}$ Represents country times day fixed effects. Appendix A includes the definition and source of each variable. >20K, >15K, and >10K represent the minimum capacity of venues included in the sample. Standard errors (in parenthesis) are clustered at the region level. ***, **, * represent statistical significance at the 1, 5, and 10% level, respectively.

	I_Games			Log*(1+Attendance)*			Log*(1+Capacity)*		
	>20K	>15K	>10K	>20K	>15K	>10K	>20K	>15K	>10K
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\Delta Log(\,1 + Cases_{t-1}\,)$	0.059	0.059	0.058	0.059	0.059	0.060	0.059	0.059	0.059
	(0.028)**	(0.029)**	(0.029)**	(0.028)**	(0.029)**	(0.029)**	(0.028)**	(0.029)**	(0.029)**
Log(*Population*)	0.025	0.022	0.017	0.025	0.022	0.017	0.026	0.023	0.019
	(0.007)***	(0.007)***	(0.007)**	(0.007)***	(0.007)***	(0.008)**	(0.007)***	(0.008)***	(0.008)**
Log(*Density*)	-0.002	-0.003	-0.005	-0.003	-0.003	-0.004	-0.001	-0.002	-0.004
	(0.007)	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)
Log(*GRP*)	0.048	0.052	0.057	0.050	0.050	0.053	0.049	0.051	0.054
	(0.025)*	(0.025)**	(0.025)**	(0.025)**	(0.025)**	(0.025)**	(0.025)**	(0.025)**	(0.025)**
Lagged week 1 (c_1)	0.003	0.014	0.007	0.002	0.002	0.003	0.000	0.001	0.000
	(0.021)	(0.021)	(0.02)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Lagged week 2 (c_2)	0.042	0.008	-0.012	0.005	0.004	0.001	0.004	0.001	-0.001
	(0.02)**	(0.022)	(0.021)	(0.002)**	(0.002)*	(0.002)	(0.002)**	(0.002)	(0.002)
Lagged week 3 (c_3)	-0.021	0.003	0.019	-0.003	-0.001	0.000	-0.002	0.000	0.001
	(0.029)	(0.027)	(0.025)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)

Lagged week 4 (c_4)	-0.030	-0.042	-0.031	-0.002	-0.004	-0.004	-0.003	-0.004	-0.003
	(0.019)	(0.02)**	(0.024)	(0.002)	(0.002)*	(0.002)*	(0.002)	(0.002)*	(0.002)
Lagged week 5 (c_5)	-0.020	-0.022	-0.011	-0.002	-0.002	-0.001	-0.002	-0.002	-0.001
	(0.024)	(0.024)	(0.023)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Lagged week 6 (c_6)	0.014	0.045	0.060	0.000	0.001	0.005	0.001	0.003	0.005
	(0.024)	(0.027)*	(0.028)**	(0.002)	(0.003)	(0.003)*	(0.002)	(0.003)	(0.003)*
Country × Day FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
R-sq	0.178	0.178	0.179	0.179	0.178	0.178	0.178	0.177	0.177
Number of Obs.	2,073	2,073	2,073	2,073	2,073	2,073	2,073	2,073	2,073
Nr. of Regions	194	194	194	194	194	194	194	194	194

Table 4
Regression of Change in *Cases* on Weekly Lagged *Games*, *Attendance* and *Capacity* when a Regional Local Team Plays in a Different Region

This table reports the coefficients from the following regression:

$$\Delta Log(\,1 + Cases_{r,t}\,) = a + b_1 \Delta Log(\,1 + Cases_{r,t-1}\,) + b_2 Log(Population_r) + b_3 Log(Density_r) + b_4 Log(GRP_r) + \sum_{w=1}^{6} c_w WX_{r,t-w} + FE_{c\times t} + \epsilon_{r,t}.$$

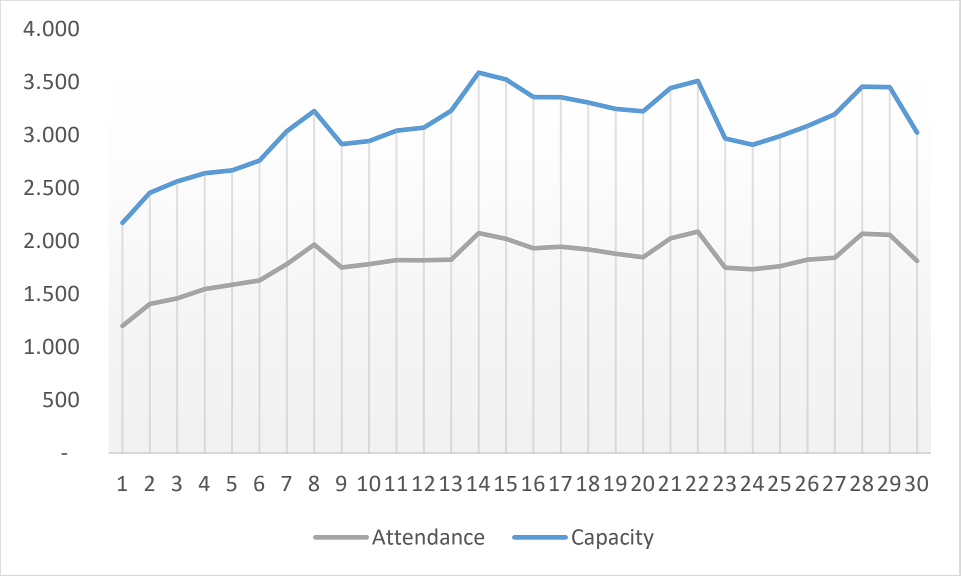
$\Delta Log(\,1 + Cases_{r,t}\,)$ represents (log) difference between 1 plus the number of cases in region *r* and day *t* with respect to day *t-1*. Likewise, $\Delta Log(\,1 + Cases_{r,t-1}\,)$ is the same variable lagged 1 day. For every lagged week w={1,2,…,6} and region *r*, the variable $WX_{r,t-w}$ represents, alternatively, the dummy variable, I_Games_{t-w}, that takes a value of one if there was a soccer match where a local team from region *r* played outside that region any day $t \in (t-(1+7\times(w-1)), t-7\times w)$; the natural logarithm of 1 plus the accumulated number of match attendants to those games, $Log(1 + Attendance_{t-(1+7\times(w-1))} - Attendance_{t-7\times w})$, or the natural logarithm of 1 plus the accumulated venue capacity of those games, $Log(1 + Capacity_{t-(1+7\times(w-1))} - Capacity_{t-7\times w})$. We control for each local region's *Population*, *Density* and Gross Regional Product per capita (*GRP*). $FE_{c\times t}$ Represents country times day fixed effects. Appendix A includes the definition and source of each variable. Standard errors (in parenthesis) are clustered at the region level. ***, **, * represent statistical significance at the 1, 5, and 10% level, respectively.

	I_Games (1)	*Log(1+Attendance)* (2)	*Log(1+Capacity)* (3)
$\Delta Log(\,1 + Cases_{t-1}\,)$	0.058	0.058	0.057
	(0.029)**	(0.029)**	(0.029)**
Log(*Population*)	0.031	0.029	0.031
	(0.007)***	(0.007)***	(0.007)***
Log(*Density*)	0.000	-0.001	0.000
	(0.006)	(0.006)	(0.006)
Log(*GRP*)	0.049	0.050	0.049
	(0.024)**	(0.024)**	(0.024)**
Lagged week 1 (*c1*)	-0.022	-0.002	-0.002
	(0.016)	(0.002)	(0.001)
Lagged week 2 (*c2*)	-0.013	0.000	-0.001
	(0.016)	(0.002)	(0.002)
Lagged week 3 (*c3*)	-0.002	-0.002	0.000
	(0.016)	(0.002)	(0.001)
Lagged week 4 (*c4*)	0.021	0.002	0.002
	(0.015)	(0.001)	(0.001)
Lagged week 5 (*c5*)	-0.014	-0.001	-0.001
	(0.016)	(0.002)	(0.001)
Lagged week 6 (*c6*)	-0.016	-0.001	-0.002
	(0.015)	(0.001)	(0.001)
Country × Day FE	Y	Y	Y
R-sq	0.178	0.178	0.178
Number of Obs.	2,073	2,073	2,073
Number of Regions	194	194	194

Figure 1
Instrument variables estimated with lags from 1 through 30 days

For every region in our sample and for every day from day 1 through 15 of March 2020, we estimate *# Games*, *Attendance*, and venue *Capacity x* days earlier, where *x* takes the value of 1 through 30. Panel A (B) presents the average *Attendance* and *Capacity* (*# Games*) over the 2,162 observations for every lag from 1 through 30 days. Variables are defined in Table 1.

Panel A
Average Attendance and Capacity



Panel B
Average # Games

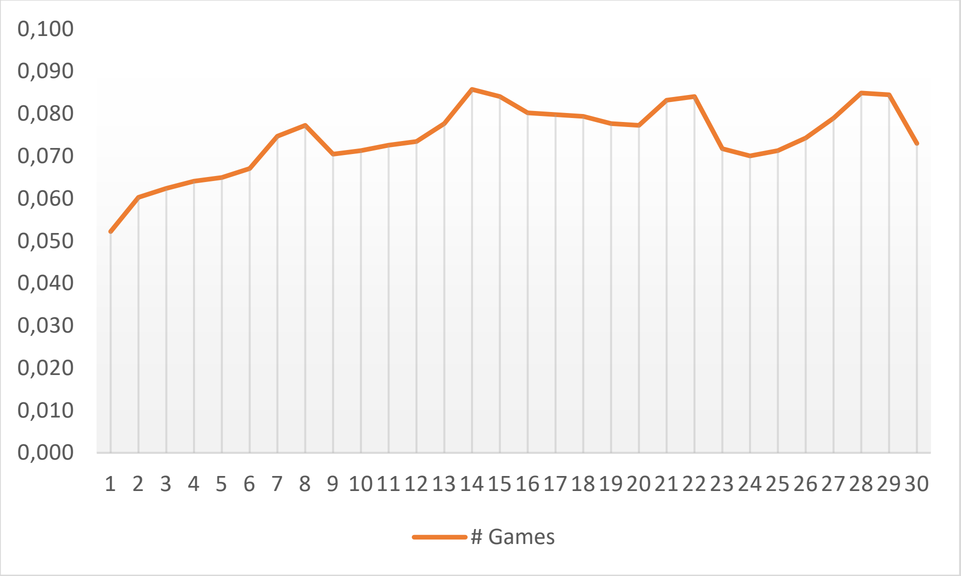


Figure 2

Total number of soccer games per day in our sample

The figure represents the total numbe of games each day from january 14 through March 14 across all regions in our sample. In the horizontal axis, we include all Satudays.

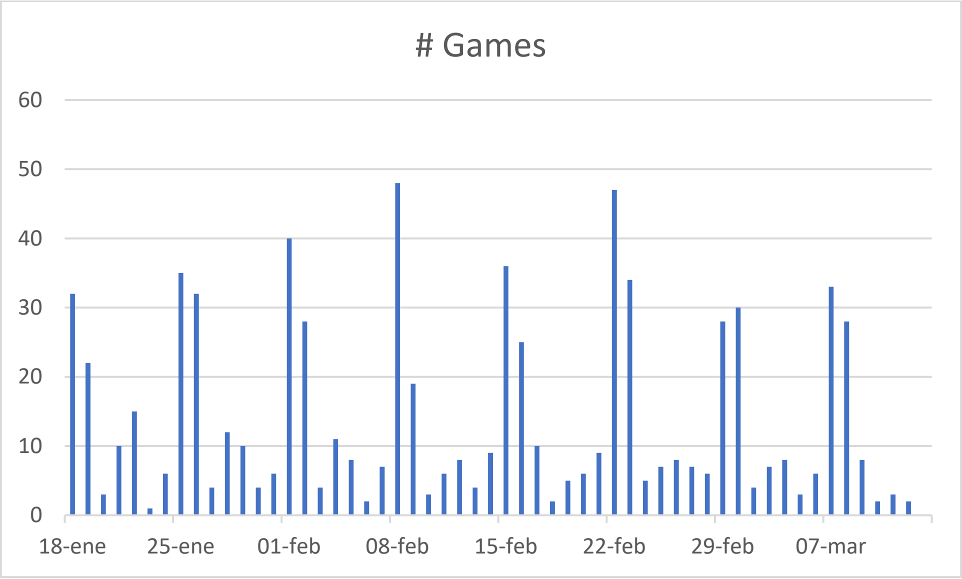
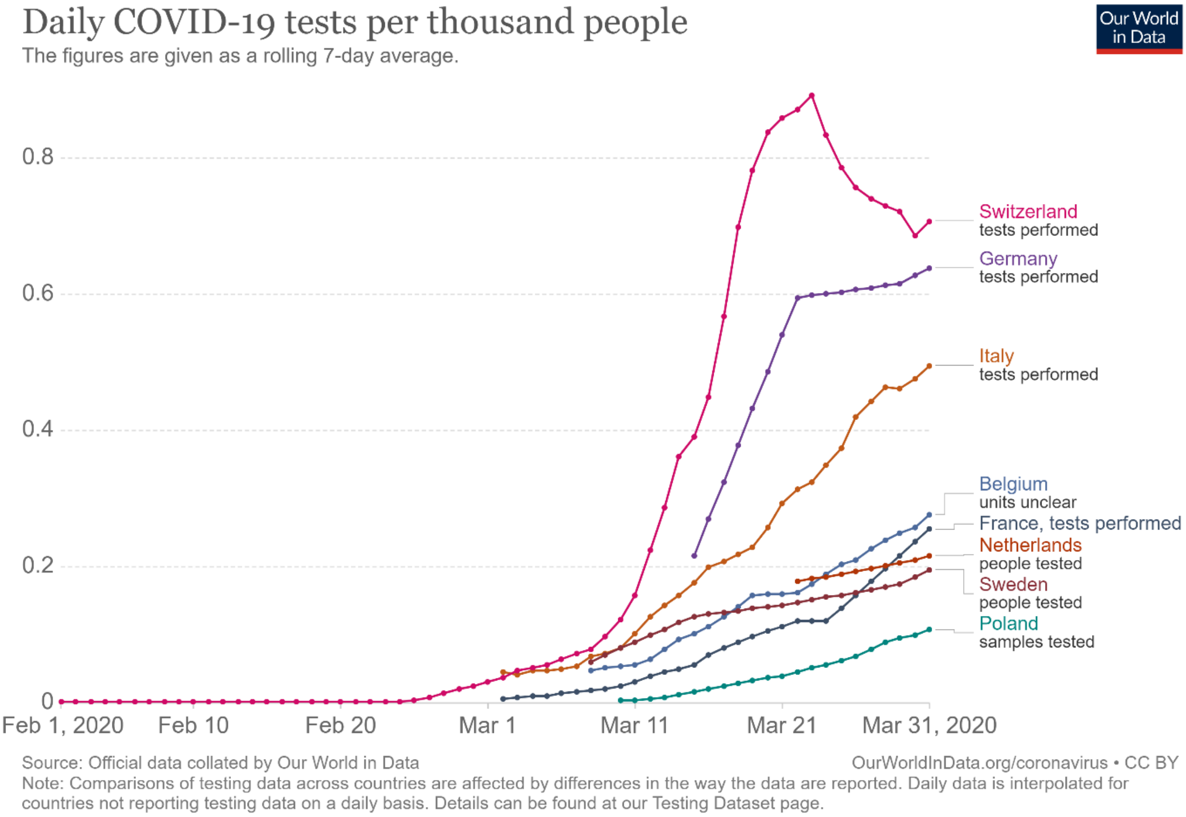


Figure 3
Daily COVID-19 test per thousand people

The figure shows the number of daily test of COVID-19 per thousand people from February 1 through March 31, 2020, for the countries in our sample for which there is data available. The graph is retrieved from https://ourworldindata.org/coronavirus-testing. Data is collected by Our World in Data by Oxford Martin School at the University of Oxford. Data description and sources per coutry can be found at https://ourworldindata.org/coronavirus-testing#source-information-country-by-country



Daily COVID-19 tests per thousand people

The figures are given as a rolling 7-day average.

Source: Official data collated by Our World in Data

OurWorldInData.org/coronavirus • CC BY

Note: Comparisons of testing data across countries are affected by differences in the way the data are reported. Daily data is interpolated for countries not reporting testing data on a daily basis. Details can be found at our Testing Dataset page.

Appendix

Table A

Variables definition and source

Main variables

Cases	Accumulated number of COVID-19 diagnosed cases per region from the following sources:

Country	Agency/Website	Country	Agency/Website
Belgium	Epistat	Poland	Serwis Rzeczypospolitej Polskiej
France	Santé Publique France	Spain	Instituto de Salud Carlos III
Italy	Dipartimento della Protezione Civile	Sweden	Folkhalsomyndigheten
Germany	Robert Koch Institute	Switzerland	FOPH
The Netherlands	RIVM	UK	GOV.UK

Cases/Population	Accumulated number of COVID-19 diagnosed cases per million inhabitant per region.
# Games	Accumulated number of soccer matches per region. Collected from the website https://www.thesportsman.com/football
I_Games	A dummy variable that takes a value of 1 if there was a soccer match in the region where the firm is located, zero otherwise.
Attendance	Accumulated number of attendants to all soccer matches in each region. Various websites, including www.footlive.com, www.azscore.com, www.soccerway.com, www.fbref.com, and www.sofascore.com.
Capacity	Accumulated maximum capacity in all venues with a minimum capacity of 25,000 spectators that hosted soccer matches per region. Retrieved from the website: https://en.wikipedia.org/wiki/List_of_European_stadiums_by_capacity.

Demographic variables

Population	Thousands of inhabitants in the region in 2018
Density	Thousands of inhabitants per square-Km in the region in 2018
GRP	Gross Regional Product: USD per capita in 2018

Country	Agency/Website	Country	Agency/Website
Belgium	NBB.Stat	Poland	Statistics Poland
France	INED	Spain	INE
Italy	ISTAT	Sweden	SCB
Germany	DESTATIS	Switzerland	FSO
The Netherlands	CBS	UK	ONS

Table B
Statistics per Region and Day

Each day is one observation. Every day from March 1 through March 14, 2020, *Cases* is the accumulated number of diagnosed cases of COVID-19 in the region until that day. *# Games, Attendance,* and *Capacity* are the accumulated number of soccer matches played in venues with capacity of at least 25,000 spectators in the region, their attendance, and the venue capacity, respectively, over the previous 6 weeks. *Population* is thousands of inhabitants per region; *Density* is number of inhabitants per square-Km, both as of 2018. The table reports the average value of each variable and region from March 1 through 14. Appendix A describes all variables and their source.

Country Region	Cases	# Games	Attendance	Capacity	Population	Density	# Obs.
Belgium Brussels	70	-	-	-	1,199	7,381	14
Belgium Flanders	322	9.21	140,116	276,198	6,553	481	14
Belgium Wallonia	165	9.79	78,607	293,571	3,624	214	14
France Auvergne-Rhône-Alpes	171	12.29	362,220	665,764	7,917	113	14
France Bourgogne-Franche-Comté	117	-	-	-	2,818	59	14
France Brittany	66	3.36	93,004	99,969	3,307	121	14
France Centre-Val de Loire	16	-	-	-	2,578	66	14

France Corsica	31	-	-	-	330	38	14
France Grand Est	346	6.50	132,825	180,914	5,555	97	14
France Hauts-de-France	187	7.86	251,519	348,176	6,007	189	14
France Normandy	33	4.14	35,226	104,321	3,336	111	14
France Nouvelle-Aquitaine	41	2.93	64,568	123,337	5,936	70	14
France Occitanie	62	3.00	42,800	99,450	5,808	80	14
France Pays de la Loire	25	6.43	107,287	204,370	3,738	116	14
France Provence-Alpes-Côte d'Azur	78	8.64	274,276	431,566	5,022	160	14
France Île-de-France	293	4.21	190,014	201,987	12,117	1,009	14
Germany Badendeath Württemberg	208	12.14	319,334	443,882	10,880	304	14
Germany Bavaria	228	9.50	442,626	515,194	12,844	182	14
Germany Berlin	57	3.43	154,714	255,939	3,520	3,946	14
Germany Brandenburg	13	-	-	-	2,485	84	14
Germany Bremen	13	3.57	148,673	150,357	671	1,598	14
Germany Hamburg	35	6.29	247,474	277,885	1,787	2,367	14
Germany Hesse	47	5.36	252,136	275,893	6,176	292	14
Germany Lower Saxony	60	7.00	177,349	264,286	7,927	167	14
Germany Mecklenburgdeath Vorpommern	12	2.79	34,076	80,786	1,612	69	14
Germany North Rhinedeath Westphalia	448	36.29	1,307,019	1,701,549	17,865	524	14
Germany Rhinelanddeath Palatinate	28	7.57	173,961	322,803	4,053	204	14
Germany Saarland	9	-	-	-	996	388	14
Germany Saxony	21	6.43	221,229	239,863	4,085	221	14
Germany Saxonydeath Anhalt	10	3.50	58,349	95,375	2,245	110	14
Germany Schleswigdeath Holstein	16	-	-	-	2,859	181	14
Germany Thuringia	8	-	-	-	2,171	134	14
Italy Abruzzo	33	-	-	-	1,312	121	14
Italy Aosta Valley	13	-	-	-	126	39	14
Italy Apulia	50	10.86	117,088	411,101	4,029	206	14
Italy Basilicata	4	-	-	-	563	56	14
Italy Bolzano	39	-	-	-	521	79	14
Italy Calabria	14	3.79	43,359	104,270	1,947	128	14
Italy Campania	102	14.64	162,686	618,583	5,802	424	14
Italy Emilia-Romagna	1,204	10.07	92,181	320,486	4,459	199	14
Italy Friuli-Venezia Giulia	90	8.14	57,341	204,646	1,215	153	14
Italy Lazio	109	13.79	313,579	973,740	5,879	341	14
Italy Liguria	129	10.36	98,136	379,061	1,551	286	14
Italy Lombardy	4,773	20.79	512,609	1,195,928	10,061	422	14
Italy Marche	313	-	-	-	1,525	162	14
Italy Molise	11	-	-	-	306	69	14
Italy Piedemont	332	11.00	124,778	415,073	4,356	172	14
Italy Sardinia	17	-	-	-	1,640	68	14
Italy Sicily	55	11.29	76,326	357,356	5,000	194	14
Italy Trentino-South Tyrol	50	-	-	-	1,072	79	14
Italy Tuscany	197	6.86	91,255	324,343	3,730	162	14
Italy Umbria	32	-	-	-	882	104	14

Italy Veneto	775	4.93	46,283	192,436	4,906	267	14
Netherlands Drenthe	7	-	-	-	493	188	14
Netherlands Flevoland	3	-	-	-	422	299	14
Netherlands Friesland	2	7.14	74,363	186,429	650	196	14
Netherlands Gelderland	24	4.07	62,696	101,786	2,084	420	14
Netherlands Groningen	1	-	-	-	586	252	14
Netherlands Limburg	26	-	-	-	1,118	521	14
Netherlands North Brabant	129	2.50	86,400	87,500	2,563	523	14
Netherlands North Holland	26	4.29	225,453	235,671	2,878	1,082	14
Netherlands Overijssel	7	6.00	80,600	181,230	1,162	350	14
Netherlands South Holland	36	3.07	143,993	157,187	3,706	1,317	14
Netherlands Utrecht	42	-	-	-	1,354	981	14
Netherlands Zeeland	3	-	-	-	384	216	14
Poland Greater Poland	2	3.00	31,614	137,490	3,398	114	11
Poland Holy Cross	0	-	-	-	1,273	109	11
Poland Kuyavia-Pomerania	-	-	-	-	2,068	115	11
Poland Lesser Poland	1	3.55	58,265	118,773	3,287	217	11
Poland Lower Silesia	4	2.91	23,987	124,425	2,887	145	11
Poland Lublin	3	-	-	-	2,162	86	11
Poland Lubusz	1	-	-	-	1,009	72	11
Poland Masovia	4	3.55	82,805	110,274	5,204	146	11
Poland Opole	1	-	-	-	1,033	110	11
Poland Podlaskie	-	-	-	-	1,191	59	11
Poland Pomerania	0	1.91	19,007	80,151	2,220	121	11
Poland Silesia	5	-	-	-	4,646	377	11
Poland Subcarpathian	2	-	-	-	2,099	118	11
Poland Warmia–Masuria	2	-	-	-	1,427	59	11
Poland West Pomerania	2	-	-	-	1,693	74	11
Poland Łódź	2	-	-	-	2,549	140	11
Spain Andalucia	99	10.14	339,071	453,185	8,450	96	14
Spain Aragon	38	3.50	90,463	117,628	1,349	28	14
Spain Asturias	31	5.93	104,055	179,357	1,077	102	14
Spain Canarias	34	2.50	30,219	81,000	2,118	284	14
Spain Cantabria	17	-	-	-	594	112	14
Spain Castilla y Leon	55	3.00	61,847	83,538	2,546	27	14
Spain Castilla-La Mancha	87	-	-	-	2,122	27	14
Spain Cataluña	166	7.86	373,219	576,342	7,571	236	14
Spain Ceuta	0	-	-	-	84	4,422	14
Spain Extremadura	20	-	-	-	1,108	27	14
Spain Galicia	36	5.57	126,220	185,571	2,781	94	14
Spain Islas Baleares	14	-	-	-	1,119	224	14
Spain La Rioja	113	-	-	-	324	64	14
Spain Madrid	916	9.00	588,469	676,052	6,499	809	14
Spain Melilla	1	-	-	-	81	6,216	14
Spain Murcia	15	3.00	-	93,537	1,474	130	14

Spain Navarra	44	-	-	-	645	62	14
Spain Pais Vasco	193	10.64	400,259	447,445	2,193	303	14
Spain Valencia	73	11.71	227,139	448,804	5,129	221	14
Sweden Blekinge	3	-	-	-	160	54	14
Sweden Dalarna	1	-	-	-	287	10	14
Sweden Gotland	1	-	-	-	59	19	14
Sweden Gävleborg	2	-	-	-	287	16	14
Sweden Halland	10	-	-	-	329	60	14
Sweden Jämtland	3	-	-	-	130	3	14
Sweden Jönköping	12	-	-	-	361	34	14
Sweden Kalmar	2	-	-	-	245	22	14
Sweden Kronoberg	3	-	-	-	200	24	14
Sweden Norrbotten	2	-	-	-	250	3	14
Sweden Skåne	54	-	-	-	1,362	123	14
Sweden Stockholm	156	4.29	37,971	175,786	2,344	360	14
Sweden Södermanland	4	-	-	-	295	48	14
Sweden Uppsala	12	-	-	-	376	46	14
Sweden Värmland	13	-	-	-	281	16	14
Sweden Västerbotten	3	-	-	-	270	5	14
Sweden Västernorrland	3	-	-	-	245	11	14
Sweden Västmanland	1	-	-	-	274	53	14
Sweden Västra Götaland	56	-	-	-	1,710	71	14
Sweden Örebro	3	-	-	-	302	35	14
Sweden Östergötland	2	-	-	-	462	44	14
Switzerland Aargau	18	-	-	-	678	388	9
Switzerland Appenzell Ausserrhoden	2	-	-	-	55	220	9
Switzerland Appenzell Innerrhoden	0	-	-	-	16	87	9
Switzerland Basel-Landschaft	25	-	-	-	290	502	9
Switzerland Basel-Stadt	55	6.00	75,895	227,964	200	5,072	9
Switzerland Bern	42	1.33	34,498	42,385	1,035	158	9
Switzerland Fribourg	17	-	-	-	319	141	9
Switzerland Geneva	92	5.00	11,914	150,420	499	1,442	9
Switzerland Glarus	1	-	-	-	40	51	9
Switzerland Graubünden; Grisons	24	-	-	-	198	26	9
Switzerland Jura	5	-	-	-	73	82	9
Switzerland Luzern	8	-	-	-	410	233	9
Switzerland Neuchâtel	24	-	-	-	177	206	9
Switzerland Nidwalden	2	-	-	-	43	138	9
Switzerland Obwalden	2	-	-	-	38	66	9
Switzerland Schaffhausen	0	-	-	-	82	246	9
Switzerland Schwyz	8	-	-	-	159	143	9
Switzerland Solothurn	4	-	-	-	273	308	9
Switzerland St. Gallen	9	-	-	-	508	222	9
Switzerland Thurgau	3	-	-	-	276	229	9
Switzerland Ticino	120	-	-	-	353	110	9

Switzerland Uri	0	-	-	-	36	33	9
Switzerland Valais	17	-	-	-	344	53	9
Switzerland Vaud	109	-	-	-	799	188	9
Switzerland Zug	7	-	-	-	127	416	9
Switzerland Zürich	67	5.11	25,964	133,420	1,521	701	9
UK Bedfordshire	3	-	-	-	669	542	6
UK Berkshire	12	-	-	-	911	722	6
UK Bristol	3	-	-	-	463	4,224	6
UK Buckinghamshire	7	3.33	28,249	101,667	809	432	6
UK Cambridgeshire	2	-	-	-	853	252	6
UK Cheshire	2	-	-	-	1,059	452	6
UK Cornwall	5	-	-	-	568	160	6
UK Cumbria	7	-	-	-	499	74	6
UK Derbyshire	6	5.83	150,093	195,983	1,053	401	6
UK Devon	21	-	-	-	1,194	178	6
UK Dorset	3	-	-	-	772	274	6
UK Durham	3	-	-	-	867	324	6
UK East Riding of Yorkshire	2	4.33	49,732	110,067	600	242	6
UK East Sussex	9	5.00	63,266	153,750	845	472	6
UK Essex	8	-	-	-	1,833	499	6
UK Gloucestershire	5	-	-	-	916	291	6
UK Greater London	145	31.50	1,211,548	1,447,249	8,899	5,671	6
UK Greater Manchester	27	13.17	415,219	563,642	2,813	2,204	6
UK Hampshire	18	3.00	87,876	97,515	1,844	489	6
UK Herefordshire	1	-	-	-	192	88	6
UK Hertfordshire	18	-	-	-	1,184	721	6
UK Isle of Wight	1	-	-	-	142	372	6
UK Kent	10	-	-	-	1,846	494	6
UK Lancashire	6	4.33	54,252	135,924	1,498	487	6
UK Leicestershire	4	3.83	118,206	123,863	1,053	489	6
UK Lincolnshire	2	-	-	-	1,088	156	6
UK Merseyside	10	6.50	318,321	322,475	1,423	2,200	6
UK Norfolk	-	2.00	54,120	54,488	904	168	6
UK North Yorkshire	5	4.00	83,202	139,952	1,159	134	6
UK Northamptonshire	6	-	-	-	748	316	6
UK Northumberland	-	-	-	-	320	64	6
UK Nottinghamshire	9	4.00	113,541	122,412	1,154	535	6
UK Oxfordshire	14	-	-	-	688	264	6
UK Rutland	-	-	-	-	40	104	6
UK Shropshire	2	-	-	-	498	143	6
UK Somerset	2	-	-	-	965	232	6
UK South Yorkshire	7	8.00	206,392	297,166	1,403	904	6
UK Staffordshire	4	4.00	92,488	120,356	1,131	417	6
UK Suffolk	1	5.00	95,139	151,555	759	200	6
UK Surrey	11	-	-	-	1,190	716	6

UK Tyne and Wear	8	7.00	254,218	348,211	1,136	2,105	6
UK Warwickshire	4	-	-	-	571	289	6
UK West Midlands	12	19.33	425,726	592,587	2,916	3,235	6
UK West Sussex	4	-	-	-	859	431	6
UK West Yorkshire	11	7.67	203,289	254,780	2,320	1,143	6
UK Wiltshire	6	-	-	-	720	207	6