The Geography of Sport in Finland

Seppo Suominen
Haaga-Helia University of Applied Sciences
Helsinki, Finland
The Geography of Sport in Finland

- Introduction
- Relevant literature
- A model
- Data and prescriptive statistics
- Empirical results
- Conclusions
Introduction

- The sports geography shows that in Finland the most of the men’s ice hockey or floorball top league teams are located in large cities while a more traditional baseball in Finland is more rural.
- During a rather long period from 1990 to 2015 men’s ice hockey has been played in 15 different cities.
- Other popular team sports leagues have not been closed during the sample period but still the geography shows that the locations have been rather stable.
Introduction

|                          | Ice hockey, # 340 | Football, # 327 | Baseball, # 337 | Floorball, # 328 | Volleyball, # 282 | Basketball, # 323 |
|--------------------------|-------------------|-----------------|-----------------|------------------|-------------------|-------------------|
| Regular number of teams  | 12 - 15           | 10 – 14         | 11 – 15         | 10 - 14          | 8 – 12            | 10 – 16           |
| in highest league        |                   |                 |                 |                  |                   |                   |
| Different teams          | 18                | 33              | 28              | 46               | 33                | 27                |
| Different towns          | 15                | 23              | 27              | 23               | 24                | 21                |
| HHI (towns)              | 935               | 719             | 519             | 1124             | 640               | 610               |
| Pop 2005, min            | 31190             | 10716           | 3414            | 7413             | 3834              | 7844              |
| Pop 2005, 25 %           | 59017             | 22233           | 9886            | 57085            | 14035             | 18083             |
| Pop 2005, median         | 122720            | 76191           | 21885           | 174984           | 24243             | 54802             |
| Pop 2005, 75 %           | 174984            | 127337          | 37374           | 203029           | 57617             | 174984            |
| Pop 2005, max            | 560905            | 560905          | 560905          | 560905           | 560905            | 560905            |
Introduction

- The Herfindahl-Hirschman index \( H = \sum_{i=1990/91}^{2015/26} s_i^2 \) measuring the concentration of ice hockey top teams is 0.0935 or 935. Floorball has been more concentrated since the Herfindahl-Hirschman index is 1124.
Introduction

- The stability of team locations raises some questions:
  - Why teams survive in some locations,
  - How many top teams a particular town can sustain,
  - How differentiated these towns are in terms of different sport types.
Relevant literature

- Using NHL data Jones and Ferguson (1988) show that the major attributes that have an impact on the chances of franchise survival are population and location in Canada.
- The locational quality is the key element in teams’ revenue determination.
- Coates and Humphreys (1997) show that sports environment and real income growth are negatively interrelated.
- Chapin (2000) and Newsome and Comer (2000) emphasise that since the Second World War sport facilities or venue were built in suburban locations but not in city centres however, since 1980’s most of the new professional sport venues have been located in central city areas although these locations are rather expensive to acquire.
- Oberhofer, Philippovich and Winner (2015) show using German football data that financial resources have a positive impact on survival in the highest league (Bundesliga) while the local market size measured by population has a low but negative effect on survival.
A model

- The monopolistic competition assumption is suitable for analysing the equilibrium number of different top league sports teams (brands) in a town.
- Following Shy (1995) a simplified version of Dixit and Stiglitz (1977) model is used.
- The utility function of the sport spectator is given by a CES (constant elasticity of substitution) utility function:
  \[ u(q_1, q_2, q_3, \ldots) = \sum_{i=1}^{N} \sqrt{q_i} \]
A model

- The representative consumer’s income is made up of total wages paid by the firms producing these brands and the sum of their profits. The wage rate is normalised to equal 1, hence all monetary values are all denominated in units of labour. The budget constraint is then.

\[ (3) \sum_{i=1}^{N} p_i q_i \leq I = L + \sum_{i=1}^{N} p_i q_i \]

- Where L denotes labour supply.
A model

- The Lagrangian ($\mathcal{L}$) is the following.

\[
\mathcal{L}(q_i, p_i, \lambda) = \sum_{i=1}^{N} \sqrt{q_i} - \lambda \left[ I - \sum_{i=1}^{N} p_i q_i \right]
\]

- The demand and price elasticity ($\varepsilon_i$) for each brand are given by

\[
q_i(p_i) = \frac{1}{4\lambda^2(p_i)^2}, \quad \varepsilon_i = \frac{\partial q_i}{\partial p_i} \frac{p_i}{q_i} = -2
\]
**A model**

- Each brand is produced by a single sport club. All clubs have identical cost structure with increasing returns to scale. Formally, the cost function \( C_i \) of a sport club producing \( q_i \) units of brand \( i \) is given by

\[
C_i(q_i) = F + cq_i \quad \text{if} \quad q_i > 0, \text{or} \quad C_i(q_i) = 0 \quad \text{if} \quad q_i = 0
\]

- Each sport club behaves as a monopoly over its brand and maximises its profit (8)

\[
\max_{q_i} \pi_i = p_i(q_i)q_i - (F + cq_i)
\]
A model

- In the monopolistic competition model free entry of clubs will result in each club making zero profits in the long run and each club has excess capacity.
- Since each club that is making some production and maximising its profit the marginal costs must equal marginal revenue

\[(9) \quad MC(q_i) = MR_i(q_i) = p_i \left(1 + \frac{1}{\varepsilon_i}\right) = p_i \left(1 + \frac{1}{-2}\right) = \frac{p_i}{2} = c\]
A model

- The zero profit condition denotes that \( q_i = \frac{F}{c} \). The labour market equilibrium presumes that labour supply (L) equals labour demanded for production: \( \sum_{i=1}^{N} (F + cq_i) = L \) which implies that \( N \left[ F + c \left( \frac{F}{c} \right) \right] = L \).

- The monopolistic competition equilibrium is therefore given by

\[
(10) \quad p_i = 2c \quad \text{and} \quad q_i = \frac{F}{c} \quad \text{and} \quad N = \frac{L}{2F}
\]
A model

- H1: If the town is small in terms of population \((L)\), the variety of sports offered in a town is small \((N)\).
- H2: When the fixed costs \((F)\) due to nature of the sports are high, the variety of sports offered in a town is low \((N)\).
- H3: The number of spectators \((q_i)\) is more correlated with fixed costs \((F)\) than with population \((L)\).
Data and prescriptive statistics

- Data on the number of top sport teams in a town is typically count data. We have in the data some towns that have had during the 1990 – 2015 period only once a top team and the corresponding figure in Helsinki is 215.

- Men: Ice hockey, Football, Baseball (not American style!), Floorball, Volleyball, Basketball.
Ice hockey
Football
Baseball
Basketball
Volleyball
Floorball
Sports geography
blue
black
orange
red
yellow
white
Empirical results

- There are two commonly used estimation methods for count data: Poisson regression and Negative Binomial regression.
- The $y_i$ variable is the aggregate number of teams in the highest league of these six sports from 1990 to 2015. Bigger towns naturally have the highest score: Helsinki has 214 (population in 2005 was 560905), Espoo 83 (231704), Tampere 176 (204337), Vantaa 41 (187281), Turku 92 (174868), Oulu 72 (173436) and Jyväskylä 97 (124205). Espoo and Vantaa are the neighbouring cities of Helsinki and it seems that Helsinki is cannibalising their score.
- The Dixit-Stiglitz model equilibrium proposes that the score ($N$) is related to labour (incomes so that wage is equalised to one), hence a relevant $x_i$ variable takes into account both (the logarithm of) the population (2005) and personal incomes (2007).
### Empirical results

| $y_i = \text{Score}$ | Poisson | Negative Binomial | Poisson | Negative Binomial |
|-----------------------|---------|-------------------|---------|-------------------|
| Log Population        | 0.924  | 0.749             | 0.924  | 0.749             |
|                       | (0.023)** | (0.092)**          |         |                   |
| Dummy: Population     |         |                   |         |                   |
| $< 15000$             | -0.318 | -0.848            | -0.318 | -0.848            |
|                       | (0.135)* | (0.256)**          |         |                   |
| Dummy: 15000 $<$     |         |                   |         |                   |
| Population $< 30000$  | 0.128  | -0.504            | 0.128  | -0.504            |
|                       | (0.128) | (0.256)*           |         |                   |
| Dummy: 30000 $<$     |         |                   |         |                   |
| Population $< 50000$  | ref    | ref               | ref    | ref               |
| Dummy: 50000 $<$     |         |                   |         |                   |
| Population $< 100000$| 1.207  | 1.068             | 1.207  | 1.068             |
|                       | (0.121)** | (0.355)**          |         |                   |
| Dummy: 100000 $<$    |         |                   |         |                   |
| Population $< 200000$| 1.809  | 1.226             | 1.809  | 1.226             |
|                       | (0.122)** | (0.530)**          |         |                   |
| Dummy: 200000 $<$    |         |                   |         |                   |
| Population $< 500000$| 2.883  | 2.270             | 2.883  | 2.270             |
|                       | (0.120)** | (0.603)**          |         |                   |
| Log Incomes           | -2.417 | -2.118            | -2.417 | -2.118            |
|                       | (0.245)** | (0.851)*          |         |                   |
| Constant              | 17.982 | 16.814            | 25.796 | 25.956            |
|                       | (2.500)** | (8.435)**          | (2.655)** | (7.288)**          |
| $\alpha$              | 0.503  | 0.324             | 0.503  | 0.324             |
|                       | (0.116)** | (0.074)**          |         |                   |
| McFadden Pseudo $R^2$| 0.668  | 0.719             | 0.668  | 0.719             |
|                       | 0.370  |                   | 0.370  |                   |
| $\chi^2$              | 1682.164*** | 309.444***          | 1810.004*** | 190.838***          |
| Overdispersion tests: $g = \mu$ | 3.669 | 5.804             | 3.669  | 5.804             |
| Overdispersion tests: $g = \mu^2$ | 1.856 | 2.984             | 1.856  | 2.984             |
Empirical results

- The Poisson and Negative Binomial regression results show that the score is positively related to town size and negatively to incomes. It seems that sports is not favoured by high income consumers.

- H3: The number of spectators \( (q_i) \) is more correlated with fixed costs \( (F) \) than with population \( (L) \).

|               | Budget     | Population | Spectators |
|---------------|------------|------------|------------|
| Budget        | 1          | 0.331      | 0.958      |
| Population    |            | 1          | 0.389      |
| Spectators    |            |            | 1          |
Conclusions

- A larger town in terms of population is able to sustain a bigger number of teams and also a larger variation of sports.
- Somewhat surprisingly towns with lower incomes are on average able to sustain a bigger number of different teams.
- The location decisions of top teams: They seem to have less survival possibilities in high income towns than in low income towns and regions.
- The Dixit-Stiglitz model also proposes that the attendance or the number of spectators is more related and correlated with the cost structure of the team than with the population statistics. The results seem to verify this hypothesis.
Thank you for your attention