What can we learn from the football-industry workers' behavior?

CEU/MKE 2021

Gábor Békés (CEU, KRTK, CEPR)

2021-09-01

What can we learn from the football-industry workers' behavior?

Quite a bit.

What can we learn from the football-industry workers' behavior?

Quite a bit.

Thank you for your attention.

What can we learn from the football-industry workers' behavior?

When are we interested to learn from football?

- **When** to learn from the football-industry workers' behavior?

- Is there external validity?
  - If relevant outside of football
- Can it offer relevant setting?
  - If offers a setup that suits answering interesting questions
- Is there unique data opportunity?
  - If it has data that allows more adequate methods, estimations

What can we learn from the football-industry workers' behavior?

My research questions

- How globalized/diverse organizations operate?
  - We know: how ties are formed, diverse and performance
  - We don't know: how collaboration happens
- Is there homophily bias in collaboration?

- Favoritism under social pressure – what are the mechanisms?
- How careers affected by peers when starting out?
- Is changing CEOs have short run effect on performance?
- Is team stability and coherence help adaptation to adversity?

What can we learn from the football-industry workers' behavior?

My research questions

Data collection and preparation ongoing for many years with Endre Borza, Bence Szabo

Several research projects with Gianmarco Ottaviano, Marton Fleck, Bence Szabo, Gabor Kezdi, Balazs Lengyel, Endre Borza

What can we learn from the football-industry workers' behavior?

This talk

- The football industry and what we can learn from it about human behavior
  - Is there external validity?
  - Can it offer relevant setting?
  - Is there unique data opportunity?
- Collaboration and Homophily in Global Teams
- Other projects

Az előadást Kézdi Gábor emlékének ajánlom.

I. The football industry and what we can learn from it

About football 1: Is there external validity?

THE FOOTBALL INDUSTRY

- Organization/company (club, squad, team)
- Worker/employee(player)
- CEO/manager (manager/coach)
- Owner/shareholder(owner/shareholder)
- Judge/arbitrator (referee)
- Product/service (a game)
- Alliance/business group (a league)
- Rule-setter (football federation, UEFA)

About football 1: Is there external validity?

THE FOOTBALL INDUSTRY

Some key features

- Service industry
- Knowledge intensive
  - Not a Math PhD, but talent, skills, practice
- Medium sized organizations dominate
- Globalized product
- Globalized workforce

About football 1: Is there external validity?

THE FOOTBALL INDUSTRY

Some key features

- Service industry
- Knowledge intensive
  - Not a Math PhD, but talent, skills, practice
- Medium sized organizations dominate
- Globalized product
- Globalized workforce

Industries that share features

- Ad agency, marketing
- Consulting, law firm
- Design, fashion
- Software, technology
- Art: movies, film, music, theaters

About football 1: Is there external validity?

What looks special to football but is not

- Not clear what is the product: game, player, or the league?
  - Movies: film, actor, studio
  - Gaming: game, tournament, e-sport
  - Gmail: email, you
- There is competition and collaboration between businesses
  - Film industry and Oscars, standards
  - Accountancy, consulting
- It seems all business and games round around a few big stars
  - Also in media, consulting, even healthcare
- Many can consume a game at once
  - Non-rival goods, similar to all media, games

Football industry

WHAT IS *REALLY* SPECIAL ABOUT FOOTBALL AND SPORT?

Football industry

WHAT IS *REALLY* SPECIAL ABOUT FOOTBALL AND SPORT?

- Many people are genuinely interested in it
  - Unlike cement

About Football 2: Can it offer relevant setting?

WHAT IS GREAT ABOUT THE FOOTBALL INDUSTRY

Football has important features that makes it great to study human behavior

- Large variations in organizations and workers
- Clear incentives, rules, motivation and ”production” technology
  - Same used globally
- Very open and competitive
  - Both at team, player level

About Football 2: Can it offer relevant setting?

What is great about football: openly observable

A very special feature that makes it useful to study behavior is that production of games is openly observable.

- Open production: games are played in front of spectators but are not rehearsed
- Because of rules and open play – easy to partial what is special
- Data is very special: super detailed as we observe many things because game is played openly.

About Football 2: Can it offer relevant setting?

AND ANYWAY, HOW DO WE KNOW ABOUT OTHER INDUSTRIES?

Advertising[a]

[a] "Mad Men" (2007-2014)

Software development[a]

[a] "Social Network" (2010)

About Football 2: Can it offer relevant setting?

WE ARE NOT ALONE USING FOOTBALL (SPORT) DATA

- Game theory (penalty shootouts)
  - Palacios-Huerta (2003) (ReStud)
- Migration and taxation (football stars),
  - Kleven, Landais, and Saez (2013) (AER)
- Diversity/homophily in teams (in hockey)
  - Kahane, Longley, and Simmons (2013) (ReStat)
- Discrimination (race and wages in football)
  - Szymanski (2000) (JPE)
- Exposure to celebrity and Islamophobia (Salah, Liverpool)
  - Alrababa'h et al. (2021) (Am PolSci Rev)
- Peer effects (spillovers from good players in basketball)
  - Arcidiacono, Kinsler, and Price (2017) (JoLE)

About football 3: the unique data opportunity

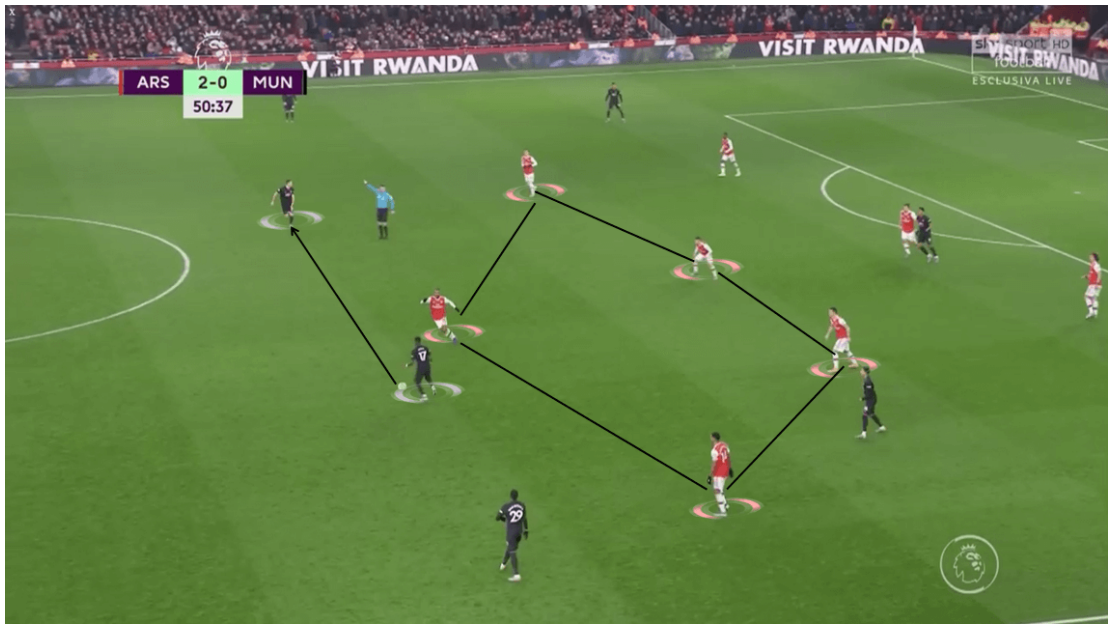
OPENLY OBSERVED ACTION.

- Production process is observed openly
  - Capital goods
  - Labor
  - Production process
- Data collected

About football 3: the unique data opportunity

OPENLY OBSERVED AND ANALYZED ACTION.

- Production process is observed openly
  - Capital goods
  - Labor
  - Production process
- Data collected
- Data analyzed by pros, reviews written: domain knowledge



About football 3: the unique data opportunity

MATCH EVENTS: DATA ON THE PRODUCTION PROCESS

- Event data - each event separately recorded with a timestamp
- One game: about 1000 events / team
- Pass events filtered: origin, destination, location, time

- Created by OPTA (cameras + algorithms + humans)
- Scraped from a sports website (`whoscored.com`) describes various aspects of industry

About football 3: the unique data opportunity

Players: data on the worker careers

- Player characteristics: nationalities, position
- Careers
- Player valuation (Not wages, but market values)

- Background, career information available widely
- Player valuations - scraped from `transfermarkt.com/`, a player information database

About football 3: the unique data opportunity

Data - technical detour: entity resolution

- This is a relational dataset, many data tables
  - Events and match features linked w match ID
  - Match info and player history linked w player ID

- Multiple data table, player names
- Entity resolution tasks is massive
  - Player names matching from resources

About football 3: the unique data opportunity

Data used: core dataset scope

- 5 top leagues (France, Germany, Spain, Italy, England),
- 8-10 seasons (2011/12-2020/21) every teams play every other 2x
- 20 (18) teams per league, 14,608 games in total

- 10.7 million passes in total (800 passes/game/team)

- 154 team, with 30 strong squad, regular churning (twice a year)
- 7000 players in sample from 132 countries

II. Collaboration and Homophily in Global Teams (with Gianmarco Ottaviano)

Globalization and homophily in teams

Motivation

- Globalization - mixing the best of global expertise in multinational teams
  - Diversity benefits: learning, innovation
  - Hurdles: communication, trust
- Interaction between people of different cultural background key to understand function of teams
- Homophily = association of similar people (shared cultural background)
- Here: Same nationality
  - Also: language, shared history

Globalization and homophily in teams

Motivation

- What we know most: how teams are formed, tie-formation, friendship networks
- Our focus is collaboration (work for a common purpose)
- How do barriers related to cultural background (nationality, language) affect collaboration in multinational teams?

Globalization and homophily in teams

Related literature

- Same ethnicity workers collaborate better, learn less: Lazear, 1999
- 'Diversity spillovers' : Ottaviano and Peri, 2005 Buchholz, 2021
- Diversity in teams: team performance and composition
  - Hockey: Kahane, Longley, and Simmons, 2013, Football: Tovar, 2020.
  - Ethnic conflict: Hjort, 2014, Laurentsyeva, 2019,
  - Team form process: Calder-Wang, Gompers, and K. Huang, 2021
- Partners in science: Freeman and W. Huang, 2015, AlShebli, Rahwan, and Woon, 2018
- Homophily in network formation in friendships: Currarini, Jackson, and Pin, 2009; Currarini, Jackson, and Pin, 2010
- More in psychology, management (Lawrence and Shah, 2020; Ertug et al., 2021).

Globalization and homophily in teams

Motivation: Measuring homophily

- Homophily = Opportunity (**induced**) + Preference (**choice**)
  - Opportunity: mechanically induced - distributions across categories define the probability they choose similar others
- Challenge: partial out induced homophily to measure choice homophily in a setup with external validity to modern workplaces
  - Option A: With experiment / find a case with random team formation
  - Option B: With observational data / model baseline.

Globalization and homophily in teams

Football is great setup 1: Observe collaboration

- Ideal setting (rules, objectives, workplace, globalized)
- Collaboration = pass rate between a pair of players
  - Collaboration is individual choice.
  - Team formation is decided by manager
- Observe collaboration and human characteristics, repeatedly in great detail
- Model baseline – a discrete choice model of players' passing behavior.
  - Pass rate between players is pinned down by their characteristics and opportunities during the matches
  - Estimate directly in model.

Globalization and homophily in teams

Football is great setup 2: Abundant, useful data

- Exhaustive dataset recording passing events from professional football
- 8 seasons, 10.7 million passes (origin and destination player ID, location)
- all 7000 players' characteristics, and career paths.

Collaboration and Homophily in Global Teams

MODEL: SETUP

- Football team $N = 11$ players, two players indexed o, d.
- The passer's decision = problem of passing the ball to the receiver who generates the highest expected benefit for the *team*.
- Game = series of short units of time (t) up to T ('periods').
  - Players o and d are together in the football pitch for $T^{o,d}$ periods.
  - In any t, a player is identified by his ID and position.
  - Two periods: t ('current period') and $t + 1$ ('future period').
- A 'pass' (o, d, t) = player o ('passer') to teammate d ('receiver'). Started by o in t, received by d in $t + 1$
- Passer takes into account the current and future implications for the team's payoff.

Collaboration and Homophily in Global Teams

Model: The player's decision

The passer's decision = pass the ball to receiver who generates the highest expected benefit for the team. Benefit to have + option value.

$$U_t^o = \ln u_t^o + \max_{\{d\}_{d=1}^{N=11}} \left\{ \beta \varphi^d E\left[U_{t+1}^d\right] - \widetilde{c}^{o,d} + z_t^d \right\}. \tag{1}$$

- $\ln u_t^d$ = benefit due to player d's characteristics
- $\widetilde{c}^{o,d}$ = challenges – 'passing cost'
- φ^d = probability of successful pass to receiver d
  - z_t^d = random part ('shock')
  - β = relative value of pass vs hold. Style (team, p1).

Collaboration and Homophily in Global Teams

Model: Passing cost

Model passing cost

$$\widetilde{c}^{o,d} = \left(g^{o,d}\right)^{\gamma} \left(l^{o,d}\right)^{\lambda} \tag{2}$$

- $g^{o,d}$ captures all distance-related frictions
- $l^{o,d}$ captures all non-distance-related frictions.
  - This is where we may see homophily - same nationality indicator
  - May be high if o and d find it hard to collaborate
  - In model, assume separability (true empirically)

Collaboration and Homophily in Global Teams

Model: Passing cost

- Solve for passing probability.

- Aggregate to pass frequency from passer to receiver in season-half
- Relative to total passes by passer when both players are on the pitch
  - approximately = time spent together * pass intensity of p1

Collaboration and Homophily in Global Teams

DATA: CREATING THE WORK DATASET- AGGREGATION

- From a choice model, aggregate to relative frequencies
- Aggregate to half-seasons (16-20 games), compromise between games and seasons
  - Noise is high / randomness of games
  - Squads are large, only 11 players at field at once, lot of variation across games, selection major issue for a single game.
  - Assume player quality is stable within a half-season (4 months) but may vary across half -seasons.

- Data: Origin player * Destination player * Season-half
- Count of passes for player pairs during a season-half
  - compared to total passes by origin player when both are playing

Collaboration and Homophily in Global Teams

DATA: CREATING THE WORK DATASET - NATIONALITY

- Same nationality definition = two players have **a** common nationality
  - 26% of players have two or three nationalities
  - Born in a country and moved to another as minor and got nationality
  - Parents have multiple nationalities
  - Example: French-Algerian dual citizenship player will have common nationality with *both* a French and an Algerian player.

Man City players[a]

[a] Ryad Mahrez(Alg, Fr) and Aymeric Laporte(Fr, Es) in 2020

Collaboration and Homophily in Global Teams

Estimated model: Poisson with player fixed effects

- Poisson regression for pass count
- SameNat=Same Nationality Indicator

$$E(Pass|.) = exp(\gamma SameNat_{p1,p2} + \lambda PassFeat_{p1,p2,t} + \\ +1 \ln TotalPass_{p1} + \gamma^1_{p1,t} + \gamma^2_{p1,t}) \quad (3)$$

  - PassFeat is features of passes, e.g. avg distance
  - $\gamma^1_{p1,t}$ is $player_1 \times half\_season$ FE
  - $\ln TotalPass_{p1}$ is the exposure

Globalization and homophily in teams

Review of results 1

- Baseline difference is 7.3% (some induced and choice)
- Once partialling out induced homophily, we find that pairs with a same nationality will pass 2.5% more (choice homophily)

- Language plays some role
  - Same nationality: 2.8%, same language: 1.3%
- Turns out, it's more about shared history, culture (colonial past)
  - Same nationality: 3%, colonial past: 2.2%, language: 0.6%

Globalization and homophily in teams

Review of results 2

- In model, estimation: Pass rate
- Who plays together is not exogeneous
  - Playing together may not be a confounder, but a mechanism
  - Managers (coaches) observe players and assemble teams
- Same nationality pairs play 1.4% more
- Adding up intensive and extensive margin: 3.9% same nationality premium

Globalization and homophily in teams

Review of results 3

- Findings stronger for deep collaboration (intensive passing plays) - twice as large coefficient
  - Partly: preferences adding up
- Bias is higher for younger players
- No difference wrt to quality

Globalization and homophily in teams

What did we learn from football workers' behavior?

1. Focus on everyday workplace collaboration high skilled, lowly charged context in a developed area with no real conflicts.
2. Well defined measure of collaboration at individual level through time (not rare pair formation)
3. Careful model of baseline via rules of the game and empirics (go beyond randomization)
4. Large data, rich and precise measures of individual characteristics, large data allows FE

III. Some other research answers

Favoritism and Social Pressure revisited

Investigating sources of social pressure

- Favoritism is the practice of giving unfair preferential treatment to one organization (or person) at the expense of another.
- Source of favoritism: social pressure - influence exerted indirectly/informally
- Favoritism by judge via social pressure. Mechanism?

- Revisit Garicano, Palacios-Huerta, and Prendergast, 2005
  - Referees favor home team - more extra time when losing.
  - Social pressure driven by home crowds that push referees.

- We confirm bias towards home team (smaller in magnitude)
  - Despite all games being televised and data publicly shared.

Favoritism and Social Pressure revisited

Investigating sources of social pressure

- Compare games when home is losing 0:1 vs when away is losing 0:1. Close, one goal matters
- 14 sec difference, even with controls

- Bias is not social pressure from the home crowds.
- Empty stadiums owing to Covid-19 as source of external variation in crowd size, we see zero change in this bias when there is no attendance.

- Offer alternative story
- Instead, we find that the bias mostly comes from favoring big teams of these leagues.

Stardust:

Stardust

The impact of replacing football team managers

Event study, English Premier League (Ch24)

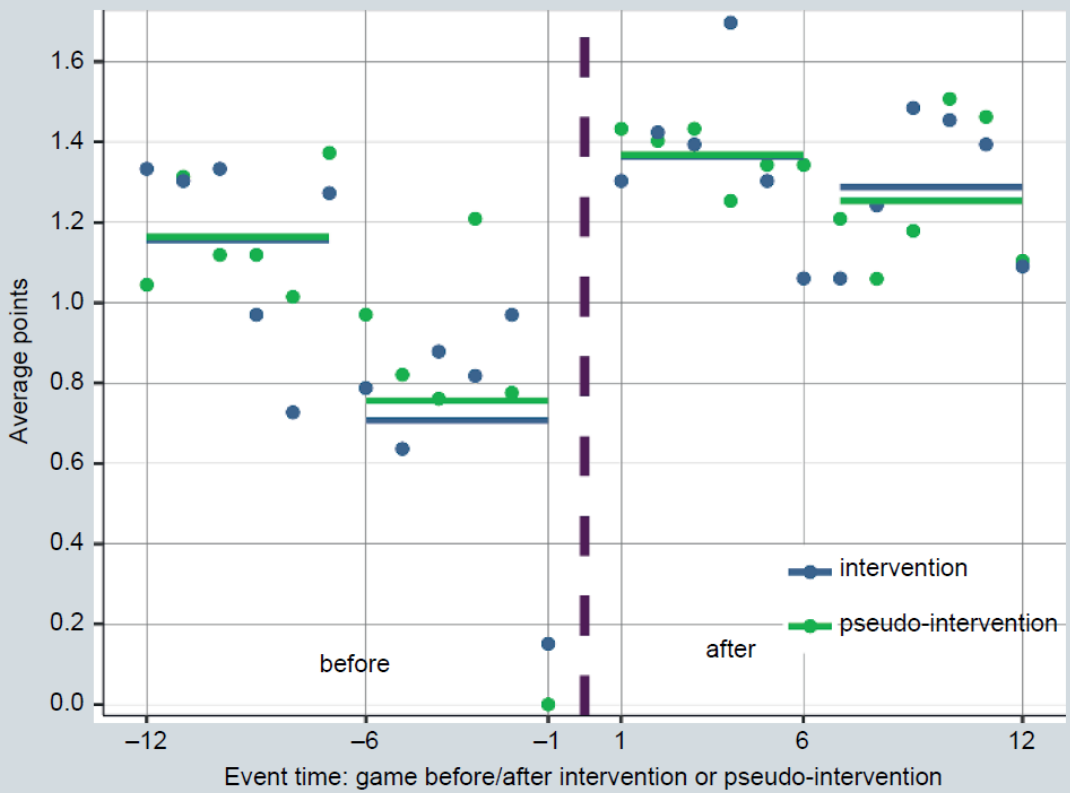


Figure 24.5 Average points before and after manager change and pseudo-intervention

Note: Average points (wins 3 pts, draws 1 pt, losses 0 pt) by event time (games before/after the management change). Six-week average lines added. Interventions: blue dots and lines. Pseudo-interventions: green dots and lines.

Source: `football` dataset. English Premier League, 11 seasons, balanced panel of 12 games before and after 33 manager changes and 67 pseudo-interventions. N=2400 team-games.

What can we learn from football workers' behavior?

What did we learn

[ezt meg ki kell találni]

- Football industry is not specific
- Data is great because of open play
- Human behavior can be observed from many angles

Thanks for the attention!

Barcelona, 1959 with Kocsis, Czibor, Kubala

References I

Alrababa'h, Ala' et al. (2021). "Can Exposure to Celebrities Reduce Prejudice? The Effect of Mohamed Salah on Islamophobic Behaviors and Attitudes". *American Political Science Review* 115.4, pp. 1111–1128.

AlShebli, Bedoor K., Talal Rahwan, and Wei Lee Woon (2018). "The preeminence of ethnic diversity in scientific collaboration". *Nature communications* 9, pp. 1–10.

Arcidiacono, Peter, Josh Kinsler, and Joseph Price (2017). "Productivity Spillovers in Team Production: Evidence from Professional Basketball". *Journal of Labor Econonomics* 35.1, pp. 191–225.

Buchholz, Maximilian (Apr. 2021). "Immigrant diversity, integration and worker productivity: uncovering the mechanisms behind 'diversity spillover' effects". *Journal of Economic Geography* 21.2, pp. 261–285.

Calder-Wang, Sophie, Paul A Gompers, and Kevin Huang (Apr. 2021). *Diversity and Performance in Entrepreneurial Teams*. Working Paper 28684. National Bureau of Economic Research.

Currarini, Sergio, Matthew O Jackson, and Paolo Pin (2009). "An economic model of friendship: Homophily, minorities, and segregation". en. *Econometrica* 77.4, pp. 1003–1045.

— (Mar. 2010). "Identifying the roles of race-based choice and chance in high school friendship network formation". *PNAS* 107.11, pp. 4857–4861.

References II

Ertug, Gokhan et al. (2021). "What Does Homophily Do? A Review of the Consequences of Homophily". *Academy of Management Annals*.

Freeman, Richard B. and Wei Huang (2015). "Collaborating with People Like Me: Ethnic Coauthorship within the United States". *Journal of Labor Economics* 33.S1, pp. 289–318.

Garicano, Luis, Ignacio Palacios-Huerta, and Canice Prendergast (May 2005). "Favoritism Under Social Pressure". *The Review of Economics and Statistics* 87.2, pp. 208–216.

Hjort, Jonas (2014). "Ethnic divisions and production in firms". *The Quarterly Journal of Economics* 129.4, pp. 1899–1946.

Kahane, Leo, Neil Longley, and Robert Simmons (Mar. 2013). "The Effects of Coworker Heterogeneity on Firm-Level Output: Assessing the Impacts of Cultural and Language Diversity in the National Hockey League". *The Review of Economics and Statistics* 95.1, pp. 302–314.

Kleven, Henrik Jacobsen, Camille Landais, and Emmanuel Saez (Aug. 2013). "Taxation and International Migration of Superstars: Evidence from the European Football Market". *American Economic Review* 103.5, pp. 1892–1924.

Laurentsyeva, Nadzeya (Dec. 2019). *From friends to foes: National identity and collaboration in diverse teams*. Working Paper 226.

References III

Lawrence, Barbara S. and Neha Parikh Shah (2020). “Homophily: Measures and Meaning”. *Academy of Management Annals* 14.2, pp. 513–597.

Lazear, Edward (1999). “Language and culture”. *Journal of Political Economy* 107.6, S95–S126.

Ottaviano, Gianmarco I.P. and Giovanni Peri (2005). “Cities and cultures”. *Journal of Urban Economics* 58.2, pp. 304–337.

Palacios-Huerta, Ignacio (Apr. 2003). “Professionals Play Minimax”. *The Review of Economic Studies* 70.2, pp. 395–415.

Szymanski, Stefan (2000). “A Market Test for Discrimination in the English Professional Soccer Leagues”. *Journal of Political Economy* 108.3, pp. 590–603.

Tovar, Jorge (Apr. 2020). “Performance, Diversity And National Identity Evidence From Association Football”. *Economic Inquiry* 58.2, pp. 897–916.

Appendix Slide

▸ Back to Main