An evolutionary model of investment strategies in the Premier League

Harry Foster

March 7, 2025

1 Introduction

"How much money should I put into the club?" - a question key to the owner of every club in the premier league. In modern times, football players are being bought and sold for record transfer fees, and owners are often reluctant to invest into their club beyond the initial buyout due to the costs involved. This paper aims to model the evolution of investment strategies over time, showing how the decisions of other clubs will influence the decisions of each other club in the league.

2 Initial Formulation

We model our fictional league after the English Premier League, which has 20 teams, all of whom play each other twice per season. These teams earn 3 points per win, 1 per draw, and 0 per loss. There is the opportunity to spend money on transfers at specific times during the season, and in our model we will model this as a budget set for each season, which is to be spent on the team itself.

We assume that costs not related to the transfer market are covered by merchandise sales, TV revenue, and other business ventures undertaken by the club. In this way, we have 3 main sources of revenue for each club:

Premier League prize payments: The higher a team finishes in the Premier League, the more the F.A pays them as an incentive for competitive strategies.

Sponsorships: Teams that perform better will obtain more lucrative sponsorships, which pay them more money.

Special Payments: The extremities of the final table will have special revenues due to their result. These are the European places, where teams will receive more money from TV rights during the next year, and the relegation places, where teams will lose fans and prime TV spots, meaning a large loss in revenue.

These sources of revenue can be calculated alongside the average number of points needed to finish in each position in the Premier League over the last 10 years, and we obtain a quartic regression model for the number of points obtained against the profit obtained by a club spending £i million on transfers in the Premier with population $\mathbf{v} = (v_1, v_2, v_3, v_4, v_5)$, where v_is are investment strategies "Low Spend", "Medium-low Spend", "Medium Spend", "Medium-High Spend", and "High Spend" respectively:

$P(i, \mathbf{v}) = -(1059.558838 + i) + 51.712382X_i(\mathbf{v}) - 0.953450X_i(\mathbf{v})^2 + 0.009183X_i(\mathbf{v})^3 - 0.000034X_i(\mathbf{v})^4$

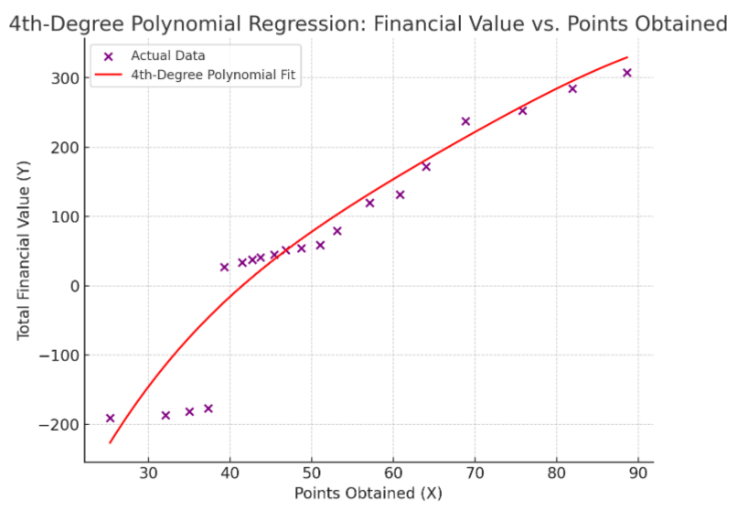


Figure 1

Using this function, we can model the changing investment strategies in the Premier league as a Moran process, where each generation is a season, the duplication of player x and removal of player y represent a club who changes strategy from y to x. This regression can serve as the basis of our fitness function in such a model, where we define X as the number of points a team expects to obtain in a season when investing £i.

$X_i(\mathbf{v}) = 2\sum_{j=1}^{5}(A_{i,j}v_j) - 2A_{i,i}v_i$

Where A is the matrix of expected points in a game between a club spending £i and a club spending £j, defined as:

$$\begin{pmatrix} 1.6 & 1.3 & 1.0 & 0.8 & 0.5 \\ 1.9 & 1.6 & 1.3 & 1.0 & 0.8 \\ 2.1 & 1.9 & 1.6 & 1.3 & 1.0 \\ 2.3 & 2.1 & 1.9 & 1.6 & 1.3 \\ 2.5 & 2.3 & 2.1 & 1.8 & 1.5 \end{pmatrix}$$

So we obtain $P_i(\mathbf{v}) = -(1059.558838 + i) + 51.712382(2\sum_{j=1}^{5}(A_{i,j}v_j) - 2A_{i,i}v_i) - 0.953450(2\sum_{j=1}^{5}(A_{i,j}v_j) - 2A_{i,i}v_i)^2 + 0.009183(2\sum_{j=1}^{5}(A_{i,j}v_j) - 2A_{i,i}v_i)^3 - 0.000034(2\sum_{j=1}^{5}(A_{i,j}v_j) - 2A_{i,i}v_i)^4$ and so our fitness function is $F_i(\mathbf{v}) = e^{0.04P_i(\mathbf{v})}$, in order to ensure all strategies admit a positive fitness with $P_i(\mathbf{v}) > P_j(\mathbf{v}) \implies F_i(\mathbf{v}) > F_j(\mathbf{v})$. We can also notice that $P_i(\mathbf{v}) > 0 \implies F_i(\mathbf{v}) > 1$, and $P_i(\mathbf{v}) < 0 \implies F_i(\mathbf{v}) \in (0, 1)$, which is a useful quality for analysing the system later on.

We will consider the current Premier League population, **v**=(2,4,6,6,2) as our starting population.

3 Analysis

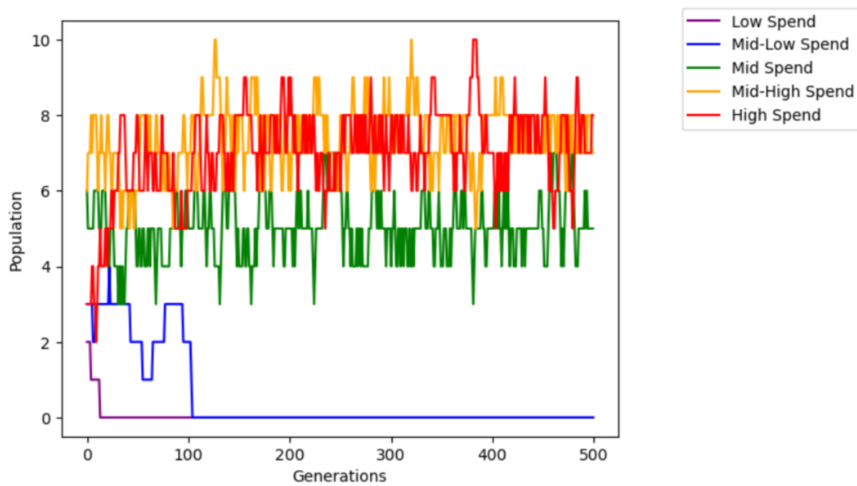
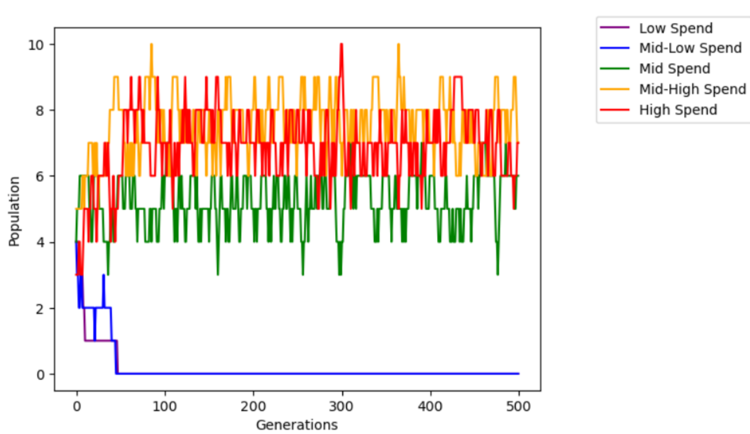
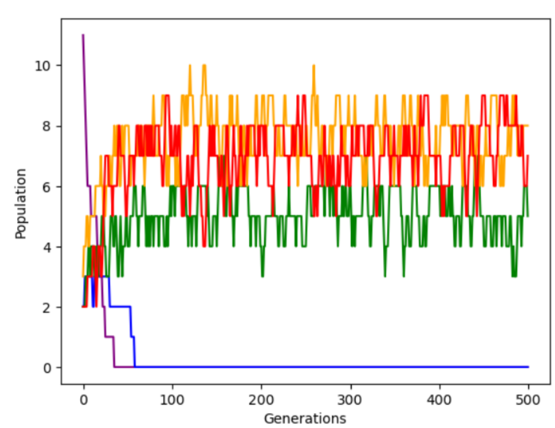


Figure 2: $\mathbf{v}$ = (2,4,6,6,2)

Within our model, we see that the "Low" and "Mid-low" strategies become obsolete fairly quickly, with no budget clubs remaining after generation 110. We will, in fact, see that in any model with at least 1 team of each strategy, the budgeting strategies become obsolete within the first 200 generations. However, the rest of the strategies remain present for the entire simulation (if we simulate 100 million generations, we still see that it is near certain that these strategies remain for every generation). The higher cost strategies enter a state where they oscillate around the populations $v_a = (0, 0, 5, 8, 7)$ and $v_b = (0, 0, 5, 7, 8)$. Specifically, the average population once the budget strategies become obsolete is (0, 0, 5.1, 7.6, 7.3). We denote this oscillating system as $v*$. Now, let us look at the following starting populations to see how the model enters a similar process each time.



(a) $\mathbf{v}$ = (4,4,4,4,4)



(b) $\mathbf{v}$ = (12,2,2,2,2)

Figure 3

In all starting populations with representation for each strategy, we quickly (within about 100 generations) enter $v*$. Let us observe the duplication and removal probabilities at the average population:

Duplication: (0, 0, 0.245, 0.360, 0.394)

Removal: (0, 0, 0.253, 0.381, 0.366)

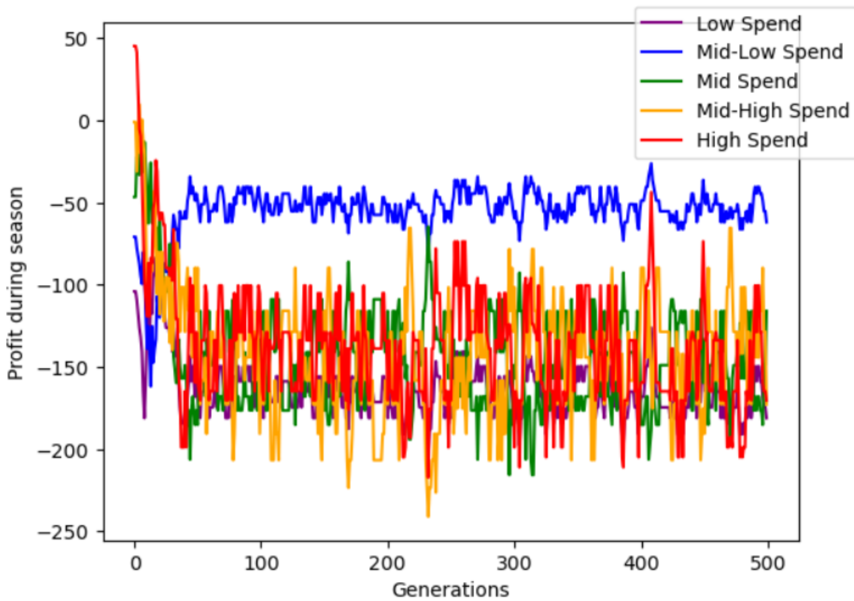
In most cases, of the mid-high or high spend teams will be duplicated, and one of them will be removed. In this scenario, either the population remains constant, or one club will change from a high spend to a mid high spend (or visa versa). Let us look at the two cases v_a and v_b

In the first case, we have our new duplication probabilities at approximately (0, 0, 0.27, 0.22, 0.51), and our new removal probabilities at (0, 0, 0.25, 0.4, 0.35). Thus, the most likely scenario is that this system either remains the same, or removes the additional Mid-High spending club and duplicates a High-Spend team, and so the next generation will have population v_b

The second case is much the same. We get duplication probabilities (0, 0, 0.2, 0.631, 0.169) and removal probabilities (0, 0, 0.25, 0.35, 0.4). This exhibits the same behaviour as population v_a with the roles of the "Mid-High" and "High" spenders reversed.

So we see that this relatively stable system is achieved due to two different populations having their most likely outcomes being to alternate between each other. It can also be shown that any other less likely result will also lead to a population which favours returning to v_a or v_b during the next few generations.

One interesting fact to notice is that in populations without the budget strategies, all teams in the league will be making a net loss. Our model's profits can be plotted to get a better idea of this:

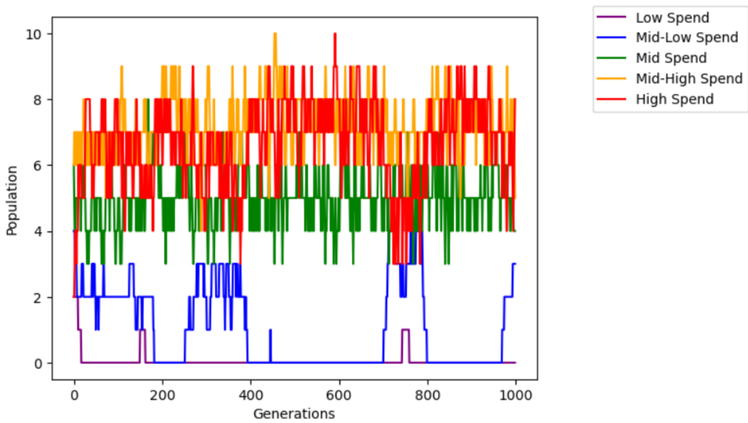


Immediately we see that our initial population favours the higher spending strategies by a large amount. These strategies will dominate the early stages of the model by punishing the budget strategies and gaining additional points, forcing such investors to pivot to other investment plans (in our model, having higher fitness so that more budget clubs will not be duplicated, and eventually will be removed from the model through stochastic processes while the higher spending clubs are duplicated and resist any reduction in their population). Once there are no remaining budget teams, all clubs begin expecting to make a significant net loss, around £100-200 million. This successfully mirrors the real English Football League, where 85% of clubs make a net loss. However, we can notice the theoretical relative success of the "Mid-Low Spend" strategy in $v*$, despite the fact that it is not present in these populations. If we could introduce a single team of this type to this system, we might expect to see it inspire other teams to also spend mid-low.

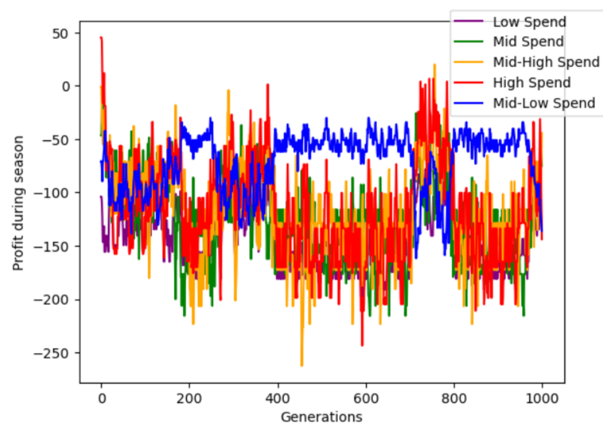
4 Updated Model with Mutation

We now update the model to include mutation, representing clubs who receive harsh financial penalties for overspending and clubs who are purchased by new wealthy owners. This is defined as a matrix of probabilities for each mutation, and one mutation can occur per season. The probability of a team being bought is proportional to the fitness of higher value strategies compared to low cost strategies, specifically by the probability $p_{i,j} = 0.005(1 - e^{\frac{F_j(\mathbf{v})}{F_j(\mathbf{v})}})$ for (i,j) such that $F_i(\mathbf{v}) < F_j(\mathbf{v})$. We assume that new owners will always invest more money than previous owners and will spend at least a medium amount, in order to gain the approval of the fans. "High Spend" teams may also spend too much and be punished under financial fair play rules. We thus obtain the mutation matrix M:

$$\begin{pmatrix} 0 & 0 & p_{1,3} & p_{1,4} & p_{1,5} \\ 0 & 0 & p_{2,3} & p_{2,4} & p_{2,5} \\ 0 & 0 & 0 & p_{3,4} & p_{3,5} \\ 0.001 & 0 & 0 & 0 & 0 \\ 0.001 & 0.01 & 0.05 & 0 & 0 \end{pmatrix}$$



(a) Population over time in the Mutation model



(b) Profit over time in the Mutation Model

Figure 4

The simulation of the mutation model over 1000 generations gives us insight into how the standard $v*$ oscillation can be potentially disturbed by the introduction of previously obsolete strategies. We immediately see by comparing the graphs that during periods where no team has taken on the budget investment strategies, teams in the league makes a monumental loss per season on average.However as we saw previously the theoretical loss of a "Mid-Low Spend" club is much less than that of any other strategy in these populations. Once a mutation occurs forcing a high-paying team to take on a budget investment strategy, one of two systems are created. If the club is forced to spend "Low", they will quickly be corrected and return to spending more money in the following generations (or rather, will not be duplicated and eventually removed through stochastic processes). However, if the club is forced to spend "Mid-Low", the strategy actually remains as part of the system for around 100-150 generations. We first notice that the presence of budget strategies both reduces the loss of higher cost strategies, but also increase the loss of budget strategies. This appears to even out so that the chance of each strategy being duplicated is approximately even. The "Mid-Low" spending team will take up around 1-3 spaces in the population before it's fitness drops to significantly less than that of teams who spend more money, reducing the number of "High" spending teams present. Eventually, this is corrected through stochastic processes to return to our oscillating $v*$ population, as the low number of "Mid-Low" teams present allows them to be removed through random selection, and their fitness does not reach a high enough value to resist this process effectively. We can also notice that if the "Mid-Low Spend" teams ever reach 3-4 members of the population, the loss of "(Mid-) High Spend" clubs becomes quite low, for example around generation 700-800 in this simulation, and actually sometimes breaks through to make a profit due to the increase in the number of clubs who provide a large number of points per game. Thus, budget strategies never make up a high proportion of the overall population, as the presence of lower spending clubs will promote an increase in the spending of teams in the league by increasing their expected point total.

The upwards mutation, representing the purchasing of budget teams by wealthy owners, does not seem to have much of an effect on the model as the "(Mid-) Low spend" strategies already go obsolete within approximately 200 generations of entering the simulation (either at the beginning or when re-entering the model through mutation).

5 Final Remarks

To conclude, we can see that the Premier League system is not a profitable business for club owners. In order to make a profit, one needs to invest a large amount of money while the league population leans towards budget strategies. Once the average investment into the clubs in the league increases, it becomes an exercise in minimising your loss rather than increasing profit.

6 Appendix

(1) "How Much Every Premier League Club Earns From Kit Sponsorship" - Give Me Sport, 14^{th} Sep 2024
https://www.givemesport.com/how-much-every-premier-league-club-earns-from-kit-sponsorship/
(2) "Premier League prize money: How much clubs make per position" - 90min.com, 17^{th} Aug 2024
https://www.90min.com/premier-league-prize-money
(3) "UEFA Champions League prize money 2024/25: Total purse breakdown for winners, knockout rounds, matches in league phase" Sporting News, 4^{th} Mar, 2025
https://www.sportingnews.com/uk/football/news/champions-league-prize-money-breakdown-ucl-winners-uefa/6e9cd9ee671ddd07fc6507ff
(4) "UEFA Champions League, Europa League, Conference League prize money: How much do winners receive? How much for a win?" - TNT Sports, 16^{th} Sep, 2024
https://www.tntsports.co.uk/football/champions-league/2024-2025/champions-league-europa-league-prize-money-conference-league-winners-final-win-draw_sto20037553/story.shtml
(5) "Conference League prize money 24/25: Here's how much Hearts will earn from playing in the Conference League" - The Scottsman, 3^{rd} Oct, 2024
https://www.scotsman.com/sport/football/hearts/conference-league-prize-money-2425-heres-how-much-hearts-will-earn-from-playing-in-the-conference-league-4808098 (6) Transfer Activity of the last 10 years - Trasfermarkt (accessed 7^{th} Mar, 2025)
https://www.transfermarkt.co.uk/premier-league/fuenfjahresvergleich/wettbewerb/GB1

(7) "Financial sustainability of men's football clubs: Bigger and better – but also riskier" - LCP.com, 31^{st} Jul, 2024
https://www.lcp.com/en/insights/in-brief/financial-sustainability-of-mens-football-clubs-bigger-and-better-but-also-riskier