<u>ShinyGate: New data shows shiny World Cup 2018 stickers really are rarer</u>

Dr Tim Paulden (Innovation & Development Manager, ATASS Sports)

A new data set logging nearly 11,000 World Cup 2018 stickers collected over the last three months indicates a substantial bias in the sticker frequencies – with "shinies" being systematically rarer than other stickers. Your #GotGotNeed habit might end up costing you a packet – unless you embrace the power of teamwork…

★ ★ ★

12 minute read

No football World Cup would be quite the same without the official Panini sticker collection – now a much-loved national tradition. In schools up and down the country, pupils have been spending break times poring over stacks of "swaps" in search of that elusive Eliseu (#118) or troublesome Troisi (#223) – and the craze has also spread to social media, where avid collectors are exchanging their duplicates under the nostalgic hashtag #GotGotNeed.

But completing the World Cup 2018 sticker album isn't exactly a cheap proposition. This year, a packet of five stickers costs £0.80, which means that the album's 682 slots will cost at least £109 to fill – and as any seasoned collector will tell you, the occurrence of duplicates can increase the cost substantially. Recently, several national newspapers published calculations by Professor Paul Harper [1] indicating that the average cost to complete the 2018 album was a mammoth £773 (assuming that you don't do any sticker swaps, and don't send off for the last few stickers when you're nearly finished). Prof Harper also noted that swapping could cut the cost substantially – for instance, a group of ten collectors who work together by swapping duplicates could reduce the average completion cost to £250 per person.

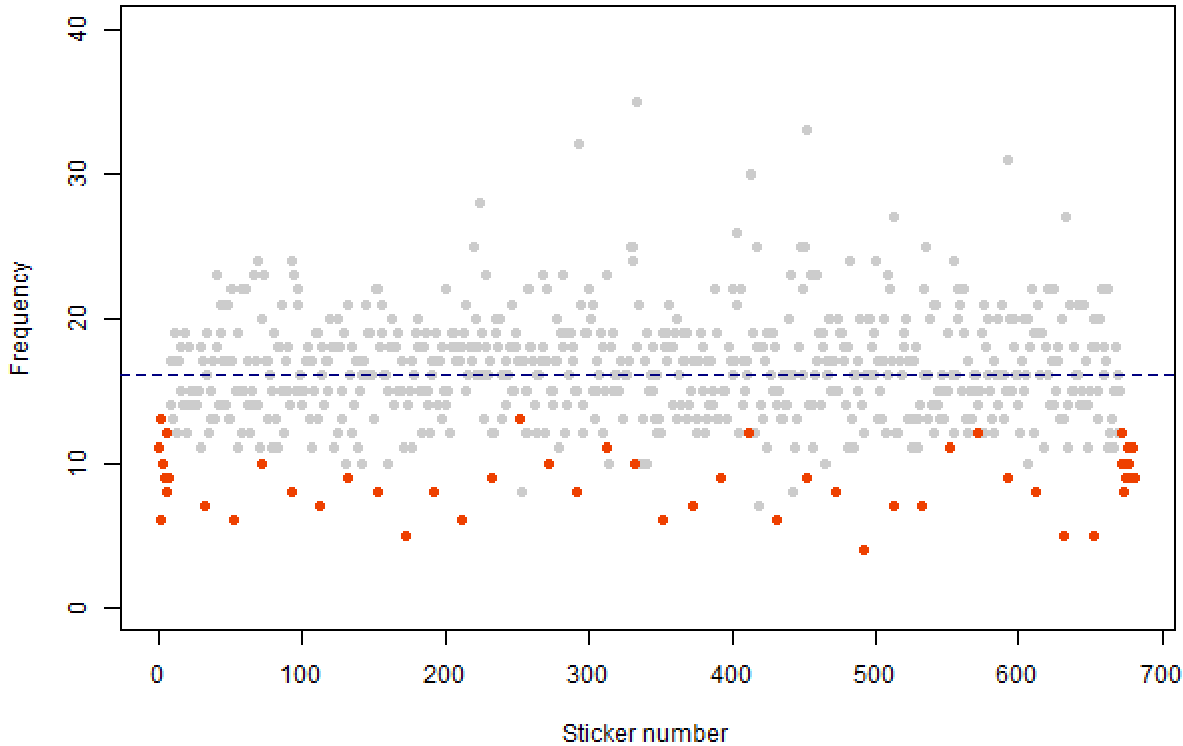
These calculations – which are instances of the well-known "coupon collector's problem" [2] – are all based on the assumption that stickers in the collection are uniformly distributed (in other words, each sticker is equally likely to arise as any other). This assumption of uniformity among the stickers appears to be accepted wisdom: indeed, according to a May 2014 article by The Economist – "The Economics of Panini Football Stickers" [3] – Panini "insists that each sticker is printed in the same volume and randomly distributed". The same Economist article also cites work by Sardy & Velenik [4] who collected 6,000 stickers for the World Cup 2010 album in Switzerland, and concluded that the overall distribution of frequencies did not depart significant from uniformity.

However, the statements and analysis appearing in the Economist article are now four years old, and were all made with reference to World Cup 2014 – so it is natural to wonder to what extent they still hold true in 2018. In other words, is the uniformity assumption still valid for the World Cup 2018 album, or might there now be certain stickers that are rarer than others? We decided to find out.

Here at ATASS Sports, we have eleven avid collectors of World Cup 2018 stickers, all working together on their respective collections (one of which is displayed proudly as double-page spreads around the walls of the research area). To streamline the process of determining swaps, each collector has been systematically logging their sticker numbers in a data file – recording both the stickers successfully filled in, and those available to swap. With the aid of these files, any collector can run a quick piece of R code to find out if there's anyone with whom they can make fruitful exchanges. Progress has been good – at the time of writing, two of the eleven collections (including the one on the wall) are 100% complete, and the remaining nine are substantially complete.

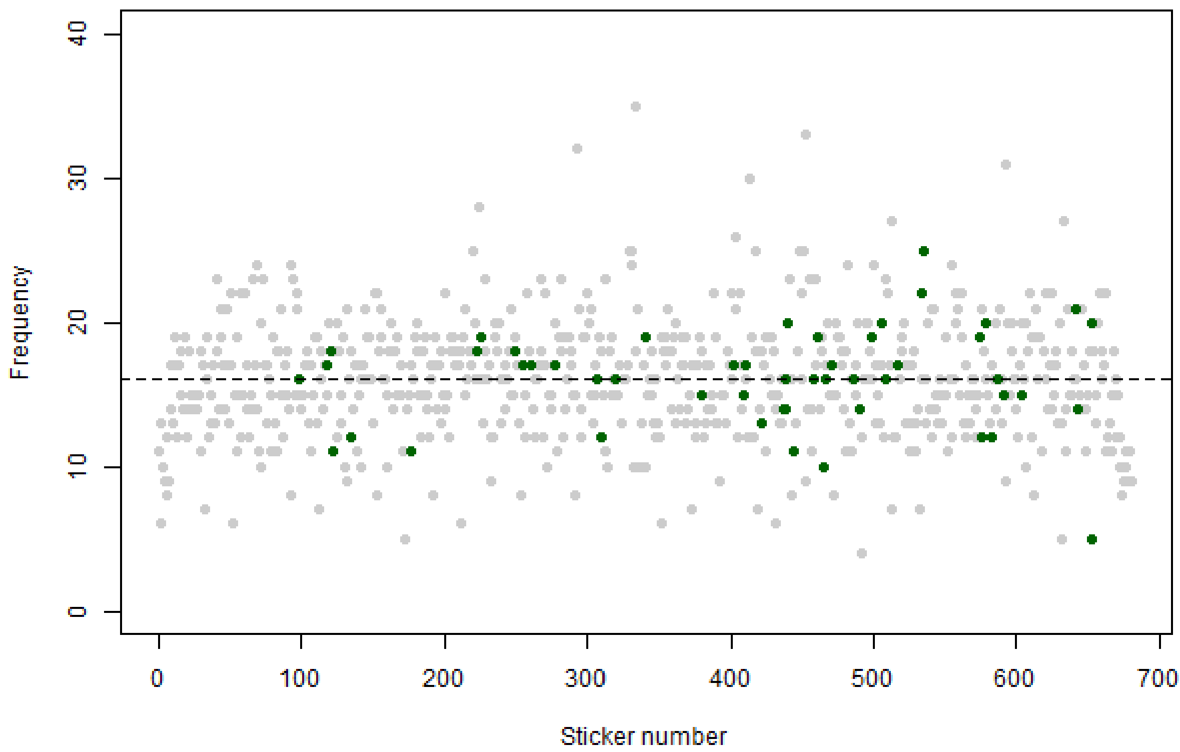
Early in July 2018, we collated the data files from all eleven collectors into a single "master" data set, and discovered that in total, just shy of 11,000 stickers had been collected. This data set gave us an accurate record of how many times each sticker number (from #0 to #681) had arisen. Moreover, we also knew the numbers of the 50 special "shiny" stickers within the collection (#0 to #7, #32, #52, #72, #92, …, #612, #632, #652, and #672 to #681).

When we plotted the data, we were pretty taken aback. The figure below shows the number of times each sticker arose, with the points corresponding to the "shiny" stickers in orange, and the "normal" stickers appearing in grey. The dashed blue line represents the overall average for the number of appearances – just a touch over 16.



Every single one of the 50 shiny stickers had arisen noticeably fewer times than the global average indicated by the blue dashed line. If all the stickers in the collection really were of equal rarity, we'd certainly expect – simply through random chance – to see some shiny stickers more than average, and others less than average. Seeing all 50 points fall beneath the average line, and by such a considerable margin, would be vanishingly unlikely – akin to flipping a coin 50 times and seeing heads come up every single time.

Here's a quick illustration of what a genuinely unbiased random pattern should look like. Rather than highlighting the 50 shiny stickers, suppose we instead highlight the 50 stickers whose official names start with "J" – that is, "José Giménez" (#98), "José Fonte" (#117), "João Moutinho" (#121), and so on, through to the "Japan team" sticker (#653). (The letter "J" was chosen because there just happen to be exactly 50 such stickers.) We obtain the following plot:



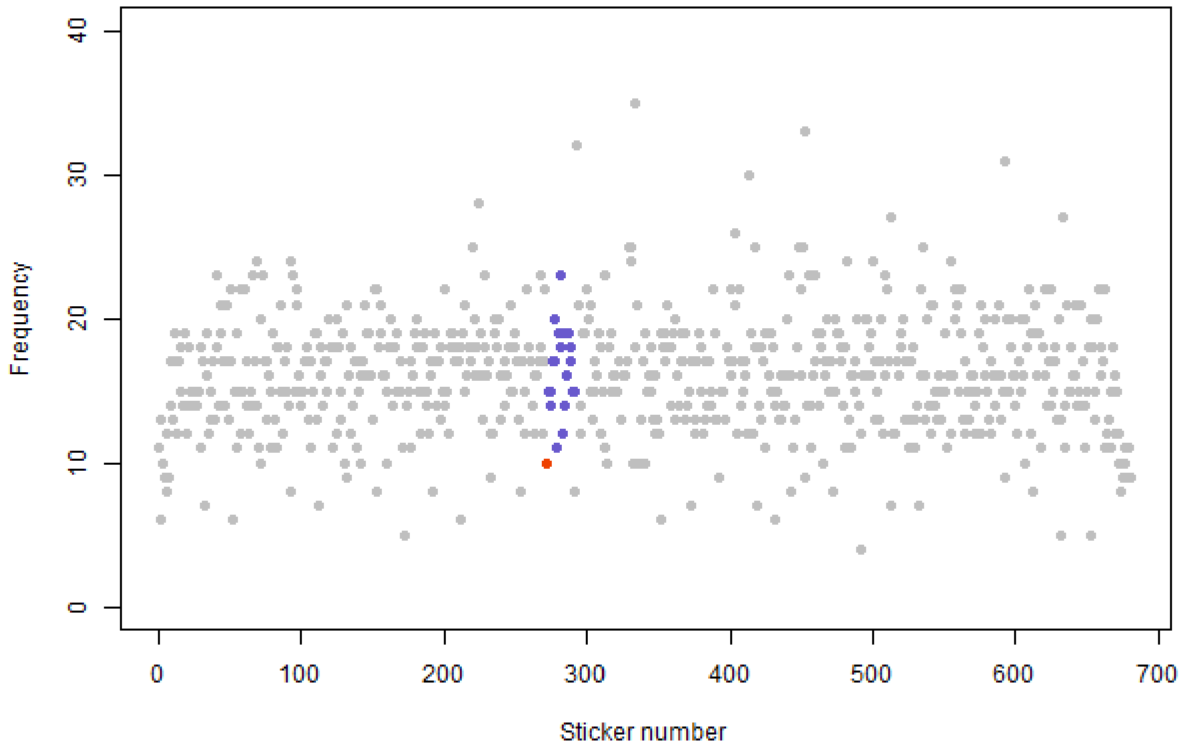
We see that the stickers whose names start with "J" show no particular bias – some occur more than average, while others occur less than average. We would expect a similar kind of pattern to arise for any arbitrary subset of 50 stickers – unless there is a reason why that subset's counts should be systematically different.

Significant others

The strength of the anomalous pattern observed among the shiny stickers becomes even more striking if we slice up the data on a team-by-team basis. In the sticker album, each of the 32 competing teams receives a double-page spread with 20 stickers to be filled in, including one “shiny” sticker – the team's logo. These 32 spreads account for 640 of the 682 stickers in the album – around 94% of the total.

If we focus attention on just these 640 stickers, we can ask the following question: for each of the 32 teams, which of the 20 associated stickers arose the fewest times? Following this line of questioning will help us quantify just how unlikely the observed patterns in the shiny sticker counts really are.

As an illustration, consider the team page for Argentina, on which the 20 stickers to be collected are numbered from #272 to #291. The figure below highlights the observed counts for each of these stickers – sticker #272, the shiny team logo, is highlighted in orange, while the other 19 stickers (#273 to #291) appear in purple. We see clearly that Argentina's shiny team logo (#272) appeared fewer times than any of the other 19 stickers.



If we repeat this process for the other teams competing, does the same pattern arise? Amazingly, we found that for 27 of the 32 teams, the shiny team logo occurred strictly fewer times than any of the other 19 stickers for that team. (There were also two further countries for which the shiny tied with another sticker for the fewest number of occurrences – but we will disregard these for simplicity.)

If all the stickers in the album really were equally rare, what would be the probability of us observing such an extreme pattern? Well, for each individual team, the probability that we observe the shiny sticker strictly fewer times than any of the other 19 team stickers is slightly less than 0.05. (If ties were impossible, the probability would be exactly 0.05, by symmetry.) For simplicity, we will take the probability to be equal to

0.05, on the understanding that our final answer will be an overestimate. When we then look across all 32 teams, the number of teams for which the shiny sticker occurs strictly fewer times than any other is described statistically by a binomial distribution with n=32 and p=0.05. Using R's statistical functions for the binomial distribution, we find the probability of an outcome as extreme as the one we saw (that is, 27 or more out of 32) is around 10^{-30} – that is, 0.000000000000000000000000000001. And remember, this value is an *overestimate*.

In other words, if all stickers were indeed equally rare, the probability of seeing the kind of anomalies we observe in the shiny sticker counts is less likely than winning the National Lottery four weeks on the trot.

★ ★ ★

Shiny unhappy people

Our discovery that shiny stickers are systematically rarer than others (by a factor of around two) contradicts the conventional wisdom that all stickers are equally likely to arise. But this bias also has an impact on the average cost required to complete the album – and we can quantify this effect using a computer simulation in R.

As an initial test, we first reran the calculations described by Prof Harper for the equal-probability case, and obtained the same average cost of £773 widely quoted in the newspapers. We then investigated the effect of modifying the sticker probabilities slightly so that they were no longer all equal. Motivated by the observed frequencies in our data set, we made the assumption that the 682 stickers actually come in two types: 50 "shiny" stickers, each arising with a probability of 1/1250 (= 0.0008), and 632 "normal" stickers, each arising with a probability of 3/1975 (≈ 0.001519). The probability that a random sticker is a "shiny" then equals 0.04, which matches our observed data. Under this model, we found the average cost to complete the album rises to around £943, an increase of 22%. In other words, if certain stickers (in this case, the 50 shiny ones) are indeed slightly rarer than others, the average cost to complete the album is higher than in the equal-probability case.

The effect is even more pronounced in Harper's collaborative scenario, where a group of ten friends work together by swapping duplicate stickers with each other. In the equal-probability case, Harper found that the average cost under this scenario was around £250 per person, and our code produced the same value. However, if we once again modify the model by introducing two different types of stickers as described above (50 shiny stickers, and 632 normal stickers), we find that the average cost rises substantially to around £370 per person – an increase of nearly 50%. Even in the swapping scenario, then, the increased rarity of shiny stickers has the effect of raising the expected cost by over £100 per person.

But while these inflated costs may make rather bleak reading for the aspiring sticker collector, there is a ray of hope…

Clinical finishing

As briefly mentioned when we first introduced Prof Harper's model, one detail that is absent from his calculations is the fact that Panini permits collectors to make a one-off purchase of up to 50 stickers to finish their collection, at a cost of £0.22 per sticker. In other words, once you have successfully found 632 of the 682 stickers, you have the option of completing the album for an additional outlay of £11, plus a small delivery charge of £1. While some sticker aficionados may be reluctant to resort to this option, it can bring about an enormous reduction in the total cost, since the great majority of the cost of completing an album is spent on acquiring the final few stickers.

In case this idea is unfamiliar, consider the situation in which you have already collected 681 of the 682 stickers and are just looking for the final sticker to complete the album. Under Prof Harper's equal-probability model with no swapping, you will go on buying more and more packets until you succeed in finding it, and on average you will end up spending £109 (!) just to acquire this single sticker. If we extend this concept to the situation in which you have found 632 of the 682 stickers (i.e. 50 more remain), we find that acquiring these final 50 stickers will cost you an extra £491 on average – much more than the average amount you needed to spend to collect the first 632 stickers, which is around £284. (One would like to hope, of course, that not many collectors would be quite this frivolous with their money.)

What happens if we add the “buy last 50 stickers” option into the equal-probability model proposed by Prof Harper? We find the average cost to complete the whole album drops to around £296, rather than Prof Harper’s original value of £773. Similarly, if we add the same option into our modified “50 shiny stickers / 632 normal stickers” probability model, the average cost drops down to £306, rather than £943.

In other words, once a collector reaches the latter stages of completing the album, attempting to finish it simply through buying further packets of stickers quickly becomes a very foolish idea indeed – and the cost will start to approach the exorbitant values calculated by Prof Harper.

In passing, it's also interesting that we see the effect of the shiny bias almost vanish in the above numbers (£306 versus £296) – this makes sense, as the final 50 stickers we send off for can include all the shiny stickers that we have failed to collect. In this scenario, we see their relative rarity has only a small impact on our average spend.

Dream team

Finally, what happens when we add the “buy last 50 stickers” option into the scenario with ten friends working together? This case is a bit more complicated, as we need to consider how the group members swap stickers between themselves. However, the most interesting case to consider is the “complete collaboration” scenario, where the group agree to collect stickers communally until they have collectively found 6320 of the 6820 stickers they need, and then they send off for the last 500 stickers – i.e. 50 per person. (The group might even agree not to commit the stickers to a particular album until they reach this stage, and then share out the pile so that each person receives 632 different stickers, and can send off for the remaining 50.)

If we assume equal probability for each sticker, the total cost of completing the album under this collaborative strategy averages out at around £136 per person. For our modified "50 shiny stickers / 632 normal stickers" model, it comes out as £141. These costs are surprisingly close to the "optimal" cost of around £109 to complete the album.

If the same strategy is extended to a larger group, the cost per person will continue to fall – though only slightly. For instance, for a group of 20 people collaborating as described above, the average cost per person comes out as £125 or £129 under the two respective models. (In practice, you'd need to weigh up whether such a small improvement is worth the additional organisational cost.)

★ ★ ★

Final whistle

In conclusion, we were surprised to discover an unexpected pattern hidden within our data set of 11,000 World Cup 2018 stickers, indicating that the shiny stickers in the collection were rarer than others by roughly a factor of two. (We could speculate that the factor might be exactly two – and that this may be related to the manner in which the different types of stickers are being printed and packaged.)

We also discovered that the implications of this bias in the sticker distribution are highly dependent on the behaviour being followed by collectors. While our initial calculations showed that this bias would tend to increase the average cost of completing the album, when we looked a little deeper, we saw that the brunt of the impact actually occurred when the sticker album was already close to completion. By this stage, the savvy collector should have already stopped purchasing further stickers in favour of pursuing opportunities to swap their duplicates, and then – eventually – sending off for any remaining stickers.

In particular, even when including the shiny sticker bias, we found that a group of ten collectors working together – and taking advantage of the "buy last 50 stickers" option – would spend an average of just £141 each to finish their collections, even without swapping duplicates outside the group. Amazingly, this average cost is more than six times smaller than the value of £943 we calculated for the no-swapping, no-sending-off scenario – a powerful demonstration that even in the domain of collecting football stickers, developing a realistic mathematical model can pay big dividends.

Note: We'd be delighted to hear from anyone who has independently collected similar data on the distribution of stickers, and is interested in comparing the features of the data sets.

References

[1] https://www.theguardian.com/football/2018/mar/29/cost–to–fill–panini–world–cup–sticker–book–is–734–says–maths–prof

[2] https://en.wikipedia.org/wiki/Coupon_collector%27s_problem

[3] https://www.economist.com/the–economist–explains/2014/05/27/the–economics–of–panini–football–stickers

[4] https://www.unige.ch/math/folks/velenik/Vulg/Paninimania.pdf