REVERSE CHEBYSHEV BIAS IN THE DISTRIBUTION OF SUPERPRIMES

WALDEMAR PUSZKARZ

Abstract. We study the distribution of superprimes, a subsequence of prime numbers with prime indices, mod 4. Rather unexpectedly, this subsequence exhibits a reverse Chebyshev bias: terms of the form $4k+1$ are more common than those of the form $4k+3$, whereas the opposite is the case in the sequence of all primes. The effect, while initially weak and easy to overlook, tends to be several times larger than the Chebyshev bias for all primes for samples of comparable size, at least, by one simple measure. By two other measures, it can be seen as fairly strong; by the same measures the ordinary Chebyshev effect is very strong. Both of these measures also imply that the reverse Chebyshev bias for superprimes is more volatile than the ordinary Chebyshev bias.

1. Introduction

Superprimes are prime numbers whose indices are prime too. For this reason, they are also sometimes referred to as the prime indexed primes (or PIPs).

Even if perhaps somewhat esoteric, this subsequence of primes has attracted a good deal of research attention spanning decades as evidenced by this sample of the literature dedicated to this subject and its variations, including higher order analogues [1, 2, 3, 4, 5, 6, 7, 8]. For instance, one of the earliest papers on it to be found in the literature [8], has proved that every integer greater than 96 may be represented as a sum of distinct superprimes.

In particular, the sequence has been shown to form a small set [6], a distinction it shares with, for instance, the sequence of twin primes, the latter being the primes that differ only by 2, the smallest such difference systematically possible for primes. However, as can easily be seen from the list of the first ten superprimes - 3, 5, 11, 17, 31, 41, 59, 67, 83, 109 - no two consecutive superprimes, except 3 and 5, corresponding to indices 2 and 3, form a pair of primes for they are always separated by more than 2.

The issue of the small set is worth emphasizing for if this clearly infinite subsequence of primes forms such a set - which, by the set very definition means the sum of reciprocals of its terms is finite - then this suggest that the subsequence of twin primes can also be infinite. This has been conjectured for quite some time now - it is known as the twin prime conjecture and dates back at least to the mid 19th century [9] - but, despite considerable advances towards the proof that have been made recently [10], it remains an open problem, one of the most important unsolved problems of contemporary number theory.

Date: September 5, 2023.

2010 *Mathematics Subject Classification.* 11N05.

Key words and phrases. primes, superprimes, prime index primes, statistical bias.

In this paper, we present experimental evidence that superprimes are more likely to be of the form $4k+1$ than of the alternative form $4k+3$, which is quite unexpected for the generic primes follow the opposite pattern known as the Chebyshev bias (or effect).

This phenomenon, named after Pafnuty Lvovich Chebyshev, a founding father of Russian mathematics, who discovered it in the mid 19th century, has been researched rather extensively ever since (see, e.g., [11, 12, 13]) both with the methods of experimental mathematics and of analytical number theory.

A branch of the latter, known as comparative prime number theory [14, 15], defined as the study of the discrepancies in distributions of primes in different residue classes, is, in large measure, devoted to the study of this effect in its various manifestations, which is to say, that apart from the standard Chebyshev effect, in which one observes primes of the form $4k+3$ more often than of the other admissible form, there is also the $6k + 5$ versus $6k + 1$ effect, perhaps less familiar, but also considerably investigated, as well as effects of this kind in other congruence classes, also studied in the literature, even if to a lesser degree.

What started as a seemingly inconsequential observation by Chebyshev communicated by him in a letter to Fuss in 1853 has become a pretty mainstream part of modern number theory connected to the Generalized Riemann Hypothesis [16] first also proposed in the 19th century.

To the best of our knowledge, the effect we report on here has not been previously discussed in the literature. This may, in no small part, be due to the fact that it cannot be noticed for the smallest of samples that make sense statistically, say, containing at least 100 data points, as the data presented next shows. But it does reassert itself for larger samples quite decisively by some simple measures we used to detect it, the measures that can also detect the Chebyshev bias for all primes.

2. Data and Analysis

This section presents and explains the data from the computer experiments we conducted that show that superprimes exhibit a clear reverse Chebyshev bias. To obtain this data we employed Wolfram Mathematica [21] and verified it with the help of PARI/GP [22].

However, for the sake of completeness and comparison, what follows first is a table demonstrating the ordinary Chebyshev bias, in which prime numbers of the form $4k + 3$ are slightly more common than those of the form $4k + 1$.

Exp	**#P(4k+3)**	**#P(4k+1)**	**Diff**	**Pct Diff**
2	52	47	5	5.05051
3	504	495	9	0.90090
4	5015	4984	31	0.31003
5	50050	49949	101	0.10100
6	500201	499798	403	0.04030
7	5000547	4999452	1095	0.01095
8	50001251	49998748	2503	0.00250
9	500003168	499996831	6337	0.00063
10	5000014372	4999985627	28745	0.00029

Table 1. Data to illustrate the Chebyshev bias.

The table does this for the ranges containing from 10^2 up to 10^{10} first primes, the range exponent placed in its leftmost column, labeled **Exp**. The next two columns, labeled **#P(4k+3)** and **#P(4k+1)**, contain the numbers of primes of the form $4k+3$ and $4k+1$, respectively, in the ranges, samples, whose size is specified in the first column.

Then we have columns labeled **Diff** and **Pct Diff**. The first of them informs about the differences between the numbers in the same rows of columns 2 and 3. This can be considered an absolute measure of the strength of the Chebyshev bias. The second one expresses these differences in relative (percentage) terms, i.e., it is the ratio (expressed as a percent) of the number in the next-to-last column to the total number of primes in both residue classes, which is $10^k - 1$, k being the number in the first column; -1 is due to 2, the first prime, not belonging to either of the congruence classes under consideration.

As the 4th column in the table above shows, the primes equal 3 mod 4 are indeed more common than the primes in the complementary residue class and this happens over a number of expanding ranges. It is thus systematic. Had the numbers in this column fluctuated between positive and negative values, one would not be able to ascertain a bias. But that's not the case, hence the Chebyshev observation, drawn from a much smaller sample of primes, appears to be correct.

Let us also note, as evidenced by the last two columns of Table 1, that the effect in question gradually, yet rather rapidly, decreases in strength when measured in relative terms, even if at the same time the strength absolute measure keeps increasing.

The next table contains the data for the superprimes (SP) juxtaposed against the data for ordinary primes (OP).

Exp	**#OP(4k+3)**	**#SP(4k+1)**	**OP Dev**	**SP Dev**	**Rel Dev**
2	52	49	2	-1	0.5
3	504	506	4	6	1.5
4	5015	5077	15	77	5.13
5	50050	50197	50	197	3.94
6	500201	500957	201	957	4.76
7	5000547	5000791	547	791	1.45
8	50001251	50008450	1251	8450	6.75
9	500003168	500018392	3168	18392	5.81

TABLE 2. Data for the Chebyshev bias in primes and superprimes.

Its first column refers both to the regular Chebyshev effect (OP) and the Chebyshev effect for superprimes (SP). In the former case, it means the same as the corresponding column in Table 1: the sample size exponent k of the first $10^k - 1$ primes. In the latter, it means the sample size exponent k of the first 10^k superprimes. Clearly, these sizes are very comparable, but the numbers they contain are not necessarily the same.

The second column, labeled **#OP(4k+3)**, contains the number of ordinary primes of the form $4k+3$, while the third column, labeled **#SP(4k+1)**, contains the number of superprimes of the form $4k+1$, for the sample sizes specified in column 1. The next two columns, labelled **OP Dev** and **SP Dev**, show the deviation from the expected values assumed to be half the sample size. This is strictly correct only for the superprimes, yet for the primes, this assumption improves with the sample size and for most of the samples studied here, it is more than good enough.

Finally, the column labeled **Rel Dev** shows the absolute ratio of **SP Dev** to **OP Dev**. This last column indicates that the reverse Chebyshev bias in the sequence of superprimes is stronger than the Chebyshev bias among all primes, for some ranges (4 out of 8) even about 5 times stronger. Of course, that does not mean that it will not start dissipating at some point, perhaps even at the very next range containing 10^{10} superprimes that we could not explore here due to computational limitations[1]. It may even eventually reverse for larger prime indices, becoming similar to the ordinary Chebyshev bias, though perhaps differing in strength.

It is, however, pretty surprising that it is both so strong and quite persistent as well. Even if it does not register at all in the smallest of ranges we investigated: the deviation of -1 means practically a wash and can be expected due to fluctuations. Such fluctuations can be stronger than in the Chebyshev effect for all primes as the effect under consideration is stronger too. Moreover, the smaller the sample, the greater the impact of fluctuations can be.

That the fluctuations here are indeed more pronounced than in the ordinary Chebyshev bias can be seen from the charts we show below after first introducing another simple way to examine the strength of both effects.

Let us then present one more, very elementary, measure to demonstrate that the effect under consideration is genuine and also fairly strong. We essentially use it to explicate certain differences between these two effects; the cumulative deviation measure that we used to discover the reverse effect (Table 2) is more fundamental.

It is a frequency measure. Consider 100 different cells each containing 10^6 superprimes. For simplicity, let us assume they all come from the range containing the first 10^8 superprimes, which also means that they are contiguous, and not totally (pseudo)random, as they might be too. Randomizing them would most likely produce a different outcome due to lost correlations.

How many of them contain superprimes equal 1 mod 4 in excess of 500000? Answering this question quantitatively gives us some measure of the effect strength, for the more of them do so, the stronger the effect.

We can do the same exercise for the Chebyshev effect for all primes for the first 10^8 primes with cells of the same size, but this time considering primes equal 3 mod 4. We can then compare these numbers to see how these effects stack up against one another.

It turns out that for the Chebyshev effect as many as 60 out of 100 cells have an excess of primes of the form $4k + 3$ compared to the expected value of 500000. That's 60%. What is the number for the effect we are really after here? It is fairly close: 55%.

For practical purposes, it seems reasonable to assume that an effect genuinely exists if at least 52% of the cells in the procedure described above favor it. That's a 4 point spread between those in favor and those against it. If this spread is 3 times greater, it makes sense to call the effect strong, and when it is 5 times this (rather high) threshold value, it seems warranted to call it very strong.

[1] On a typical home computer, it may take about 20 times longer to compute the number of superprimes of the form $4k + 1$ among the first 10^6 such primes that follow right after the first 10^9 superprimes than it does so for the very first 10^6 superprimes.

By these standards, the Chebyshev effect for all primes, with the spread in question at 20 is very strong[2], while the reverse Chebyshev bias for superprimes, with the spread at 10, can be considered fairly strong.

This simple approach is illustrated by histograms and graphs for the Chebyshev bias and its reverse version under discussion that follow next (figures 1 through 4).

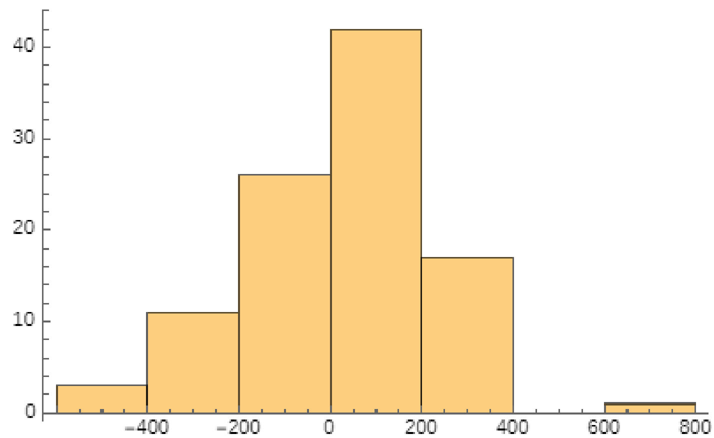

Figure 1. Chebyshev bias for primes.

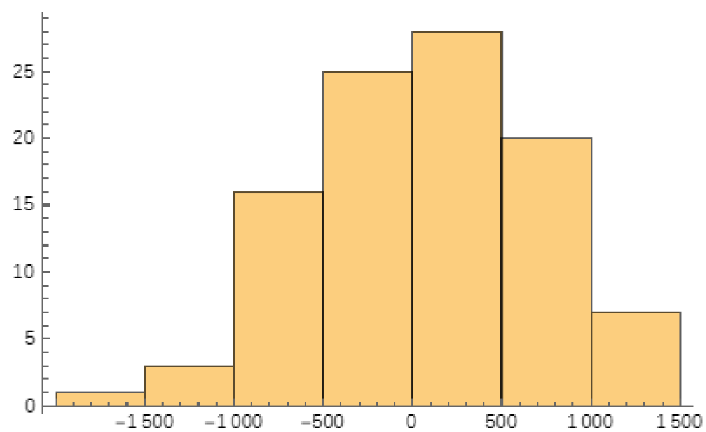

Figure 2. Reverse Chebyshev bias for superprimes.

These histograms clearly demonstrate that the Chebyshev bias for primes is distinctly sharper while at the same time much less volatile. Out of the 100 cells only one has an excess over 500, while such deviations from the expected value are much more common in the case of superprimes. They occur in both directions, with several exceeding even 1000 in absolute terms. The picture we get here is that the reverse Chebyshev bias for superprimes features more, and even quite extreme, swings. In comparison, the ordinary Chebyshev bias is much quieter.

[2]In the chosen range of the first 10^8 primes, at least. In other ranges, with other cell sizes, and especially with randomization, things may look different.

This notable difference in the characters of these two effects - one quiet, the other volatile - can perhaps be even better appreciated through the graphs below.

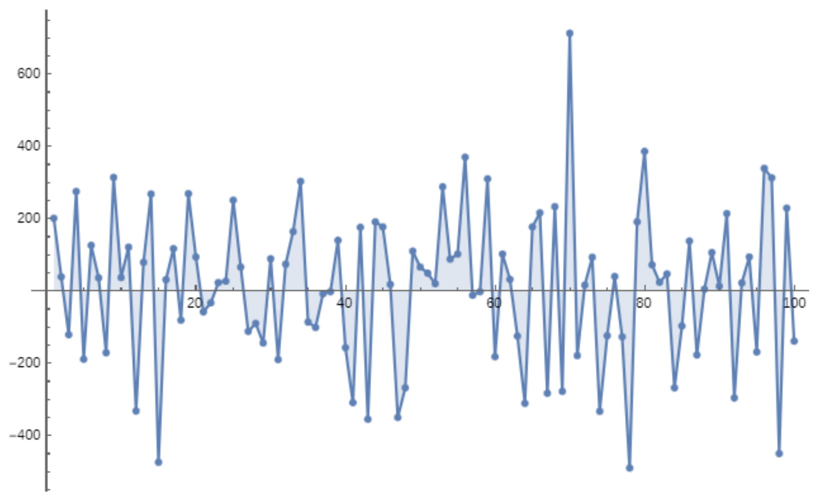

Figure 3. Chebyshev bias for primes.

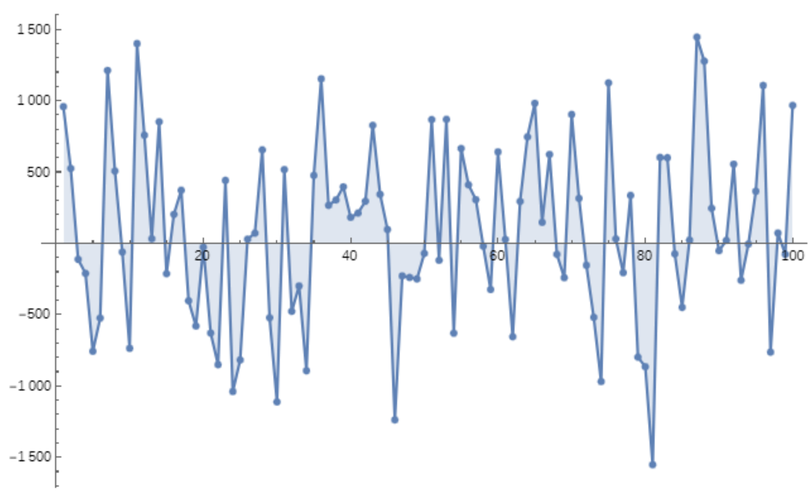

Figure 4. Reverse Chebyshev bias for superprimes.

The information that the figures above provide allows us to better understand two discrepancies in Table 2. One we already mentioned above: the negative deviation in the smallest sample of superprimes (column 5 of the said table). We already indicated that this could be due to fluctuations, and these figures, that reveal how wild such fluctuations can be, confirm our hypothesis. But there is another discrepancy in the very same column: it is between samples with exponents 6 and 7. In the latter case the deviation is smaller than in the former, unlike in the case of the regular Chebyshev bias where such deviations steadily grow. This too can be attributed to the wild fluctuations under discussion.

To get still better an idea about the effect studied here, another way of looking at it, from the point of view of a race between two groups of numbers (competing residue classes) [13], is entertained below, culminating in Table 3. It provides yet another visualization of this effect, and, in particular, the fluctuations in question.

This is accomplished with the help of three graphs, each showing the difference - let us denote it by Δ_{SP} - between the number of superprimes equal 1 mod 4, $\pi(n, 1, 4)$, and equal 3 mod 4, $\pi(n, 3, 4)$, as a function of the current prime index n - the prime with this index, $p(n)$, serves as an index for the n-th superprime, $p(p(n))$.

Hence, $\Delta_{SP}(n) = \pi(n, 1, 4) - \pi(n, 3, 4)$, and on the charts below it is measured along the vertical axis, with n running along the horizontal one from 1 to 10^k, where $k = 3, 4, 5$.

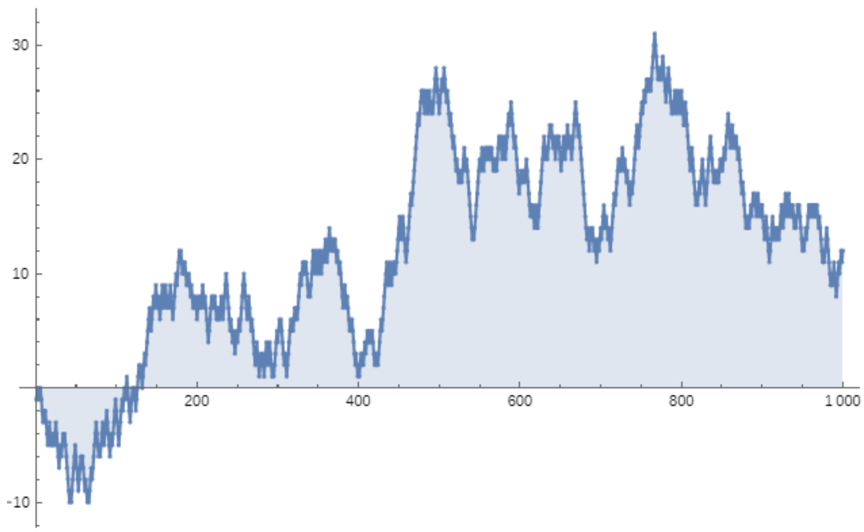

Figure 5. $\Delta_{SP}(n)$ **with** n **from 1 to 1000.**

Note that, as already alluded to before, initially, up to about $n = 100$, $\Delta(n)$ has no intention to get into the positive territory, but when it does so, it does so with aplomb, its maximum in this territory reaching 30, while it never got below 10 in the negative territory.

The next two pictures show that once it gets into the positive territory, it tends to stay there, the effect that is quite noticeable. One measure of that is how often it touches the horizontal axis: only 105 times up to $n = 10^5$, the last time being for $n = 6138$. There are no further touchings of the horizontal axis up to, at least, $n = 10^6$.

For comparison, there are 284 such touchings for the Chebyshev bias for the first 10^6 primes, although most of them initially are not followed by negative values of its $\Delta_P(n)$, understood here as $\pi(n, 3, 4) - \pi(n, 1, 4)$, that stays non-negative until $n = 2946$.

However, in the index range discussed, there are 1196 instances of negative values of $\Delta_{SP}(n)$ for the reverse effect of superprimes, which should be contrasted with 1940 such instances for the $\Delta_P(n)$ in the regular Chebyshev effect.

Thus, even though the Chebyshev effect for primes is initially positive in $\Delta_P(n)$ for nearly its first three thousands values, it experiences the negative values more often compared to the reverse Chebyshev effect for superprimes in this particular sample (range). In the other ranges (see Table 3 below), the ordinary Chebyshev effect shows a decisive strength, but the reverse effect for superprimes is fairly strong too, save for the smallest of the samples investigated.

Figure 6 shows that the height of positive peaks easily exceeds the depth of negative valleys. In fact, there are no new such valleys in Figure 7, that is, beyond those already in the previous two.

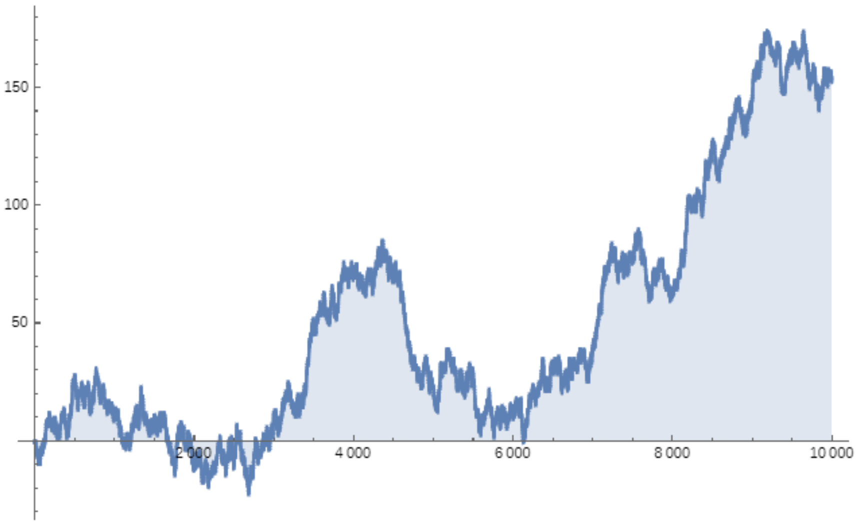

Figure 6. Δ_{SP} **with** n **from 1 to 10000.**

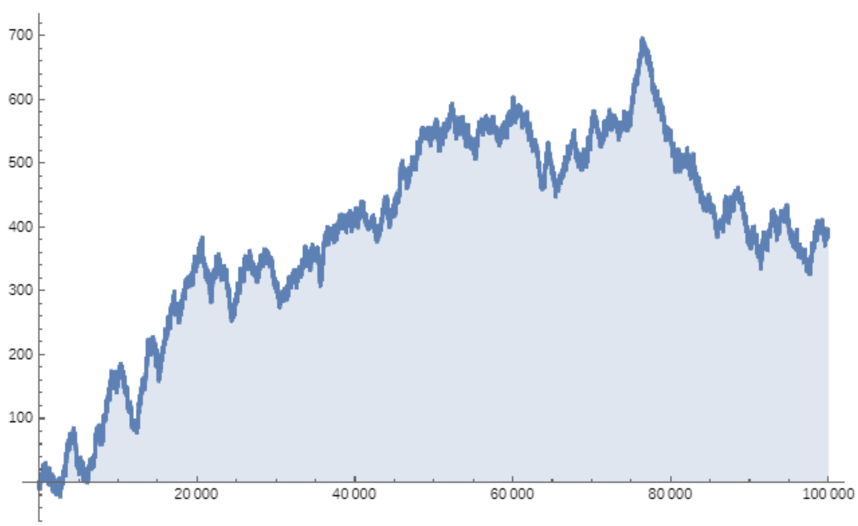

Figure 7. Δ_{SP} **with** n **from 1 to 100000.**

Note also that $\Delta_{SP}(10^k)$ $(k = 3, 4, 5)$ - the values of $\Delta_{SP}(n)$ at the right hand edges of these three charts - is twice the value in the next-to-last column of Table 2 for range exponents equal k.

It seems quite in order to compare the above charts illustrating the reverse Chebyshev effect for superprimes with at least one analogous for the regular Chebyshev effect for primes. We will now do this for the primes $p(n)$ up to $n = 10^5$, which in the index range corresponds to the last chart for the superprimes.

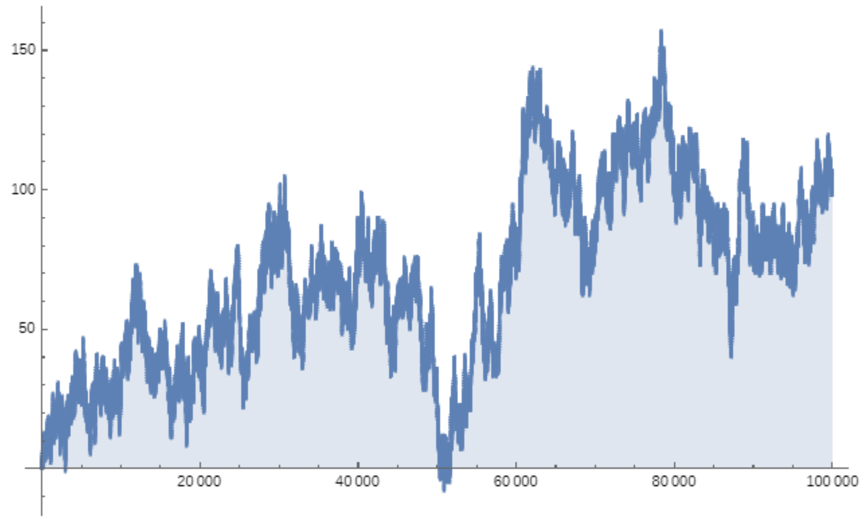

Figure 8. $\Delta_P(n)$ with n from 1 to 100000.

To continue our analysis in greater detail, let us introduce new symbols: $\#\Delta_a^+$ and $\#\Delta_a^-$, where $a \in \{P, SP\}$. The plus/minus signs refer to the positive/negative counts of respective Δ's. For instance, $\#\Delta_{SP}^+$ is the number of instances for which Δ_{SP} attains positive values in a given range.

Table 3 below collects in columns 2-4 the data for $\#\Delta_{SP}^+$, $\#\Delta_{SP}^-$, and the ratio (**Ratio SP**) of the former to Total. In columns 5-7, it collects the data for $\#\Delta_P^+$, $\#\Delta_P^-$, and the ratio (**Ratio P**) of the former to Total. Total is the sample size, i.e., 10^k or $10^k - 1$, the first value applicable to the superprimes, the other to primes, where k is the range exponent (**Exp**). The ratios are expressed as percentages and rounded off to four digits.

Exp	$\#\Delta_{SP}^+$	$\#\Delta_{SP}^-$	**Ratio SP**	$\#\Delta_P^+$	$\#\Delta_P^-$	**Ratio P**
2	0	97	0.00	95	0	95.96
3	874	116	87.40	995	0	99.60
4	8699	1196	86.99	9988	1	99.89
5	98699	1196	98.70	99649	239	99.65
6	998699	1196	99.87	997775	1940	99.78
7	9821043	177669	98.21	9997775	1940	99.98
8	93711972	6284497	93.71	99991536	7804	99.99

Table 3. Races between residue classes: superprimes and primes.

It is these ratios that are crucial here for they demonstrate both the greater smoothness of the Chebyshev effect for the primes (lesser volatility - staying steady about 99% in most ranges investigated) as well as its greater strength (only in the case of $k = 6$, the superprime Chebyshev effect is stronger) compared to the reverse effect for the superprimes, confirming the previous observations made using other statistical measures.

For the sake of completeness, let us also include here a sample PARI/GP code that can be used to obtain the data for the tables shown above.

What follows first is the code for the core data presented in Table 2. The code will print out a range exponent followed by a comma, and then followed by the number of superprimes equal 1 mod 4 in the range.

```
for(k=2, 9, c=1; a=2; c1=0; while(c<prime(10^k), a=nextprime(a+1);
c++; a%4==1&&isprime(c)&&c1++); print(); print1(k, ", ", c1, ", "))
```

With this code, using a typical home computing power, one should expect to invest several tens of hours to get the data for the biggest range of superprimes we have investigated here.

The following code will produce all the columns of Table 1 separated by commas.

```
for(k=2, 10, c=0; forprime(n=3, prime(10^k), n%4==3&&c++); print();
print1(k, ", ", c, ", ", 10^k-c-1, ", ", 2*c+1-10^k, ", ");
printf("%.5f", 100*((2*c+1-10^k)/(10^k-1))))
```

While this code is incomparably faster, on a relatively modest home computer it may still take several hours to get all the nine rows of this table.

The two pieces of code that follow next will produce the data for Table 3. The first piece computes the data for columns 1-4. The output consists of four columns, separated by commas, the first containing the values of the range exponent.

```
for(k=2, 8, c=1; a=2; c1=0; c3=0; cp=0; cn=0;
while(c<prime(10^k), a=nextprime(a+1); c++; if(isprime(c),
if(a%4==1, c1++, c3++); c1>c3&&cp++; c3>c1&&cn++)); print();
print1(k, ", ", cp, ", " cn, ", "); printf("%5.2f", 100*cp/10^k))
```

The second piece computes the data for columns 5-7 of Table 3. The output consists of four columns, separated by commas. The first of them also contains the values of the range exponent, but it is the remaining columns that we actually need for the table.

```
for(k=2, 8, c=1; a=2; c1=0; c3=0; cp=0; cn=0;
forprime(n=3, prime(10^k), if(n%4==3, c3++, c1++); c3>c1&&cp++;
c1>c3&&cn++); print(); print1(k, ", ", cp, ", " cn, ", ");
printf("%5.2f", 100*cp/(10^k-1)))
```

3. Conclusion

The investigation of the biases in the distribution of primes has a fairly long tradition dating back to the mid 19th century when the observation by Chebyshev brought to the realization of mathematicians that prime numbers are not really as random as they may appear or are assumed to be for various theoretical ends [15]: those of the form $4k + 3$ are slightly more common (or occur more frequently) and, what's important, systematically so, than those of the complementary form, $4k + 1$, indicating a deviation from a naively understood random distribution predicated on the absence of such systematic discrepancies.

Nowadays, if only in part, because of this and related bias phenomena, one thinks of the distribution of primes as pseudorandom rather than random, as emphasized, among others, by T. Tao [18]. Pseudorandom sequences are sequences that are not actually random but behave as if they were, exhibiting many properties of randomness, even though they were produced by deterministic algorithms.

Thanks to the modern computer technology, research on the Chebyshev bias and related phenomena has entered a new and exciting era [11, 13]. Without this technology and specialized mathematical software that comes with it, like Wolfram Mathematica [21] or PARI/GP [22], this paper, that revealed a reverse Chebyshev bias in the distribution of superprimes, would not be possible.

Using these software tools we were able to uncover a new effect in the distribution of some fairly popular subsequence of primes, prime indexed primes (superprimes), and to demonstrate its existence through data and visualization, while employing rather simple statistical methods. We contrasted this effect with the Chebyshev bias for generic primes, finding it fairly strong, and also more volatile than the latter, even if initially hard to register for its starts in earnest for prime indices exceeding 100.

An interesting example of the research related to ours, and also pretty recent, is a work by Lemke Oliver and Soundararajan [17], who discovered an effect similar to the Chebyshev bias, in that it also deals with distributions of prime numbers viewed through the prism of a congruence class.

More specifically, this effect concerns the last digits of primes that appear to exhibit an anti-sameness bias of sorts. Namely, it turns out that prime numbers with the same last digit repel one another, i.e., they are less likely to appear next to one another in the sequence of primes than one would expect it.

The Chebyshev and related effects in the distribution of prime numbers are fairly weak in the sense that whereas the deviation from the expected values keeps increasing in absolute terms, it decreases in relative terms, as pointed out in the previous section.

Not all bias phenomena in primes are like that, though. In particular, as discovered by us recently [19, 20] some can be both quite noticeable and free of the weakness mentioned. Preliminary research on this novel kind of effect in the sequence of superprimes indicates, rather unsurprisingly, that this is also the case there.

When it comes to the effect this paper is about, more research, with more powerful computing resources than we had at our disposal, is needed to get a more thorough understanding of this rather unexpected phenomenon, especially regarding its further behavior. Does the effect continue beyond the ranges of superprimes we have investigated here, or does it begin to falter?

This paper is only a starting point in investigating of this phenomenon, and since it is purely experimental in its approach, an analytical approach is certainly very needed and welcome here.

Acknowledgements. The author is grateful to the developers of Wolfram Mathematica [21] and PARI/GP [22] whose software was indispensable to this research.

References

[1] May, Michael P. *Properties of Higher-Order Prime Number Sequences*. Missouri Journal of Mathematical Sciences 32, no. 2: 158-170 (2020).

[2] Miska, Piotr, János T. Tóth, and Błażej Żmija. *On distribution of subsequences of primes having prime indices with respect to the* (R)*-denseness and convergence exponent*. arXiv:1908.10421 [math.NT] (2019).

[3] Guariglia, Emanuel. *Primality, fractality, and image analysis*. Entropy 21, no. 3: 304 (2019).

[4] Cattani, Carlo, and Armando Ciancio. *On the fractal distribution of primes and prime-indexed primes by the binary image analysis*. Physica A: Statistical Mechanics and its Applications 460: 222-229 (2016).

[5] Batchko, Robert G. *A prime fractal and global quasi-self-similar structure in the distribution of prime-indexed primes*. arXiv:1405.2900 [math.GM] (2014).

[6] Broughan, Kevin A., and A. Ross Barnett. *On the subsequence of primes having prime subscripts*. Journal of Integer Sequences, Vol. 12, Article 09.2.3. (2009).

[7] Powell, Barry J. *Primitive densities of certain sets of primes*. Journal of Number Theory 12, no. 2: 210-217 (1980).

[8] Dressler, Robert E., and S. Thomas Parker. *Primes with a prime subscript*. Journal of the ACM (JACM) 22, no. 3: 380-381 (1975).

[9] Dunham, William. *A note on the origin of the twin prime conjecture*. Notices of the International Consortium of Chinese Mathematicians 1, no. 1: 63-65 (2013).

[10] Maynard, James. *The twin prime conjecture*. Japanese Journal of Mathematics 14: 175-206 (2019).

[11] Rubinstein, Michael, and Peter Sarnak. *Chebyshev's bias*, Experimental Mathematics 3, no. 3: 173-197 (1994).

[12] Kaczorowski, Jerzy. *On the distribution of primes (mod4)*. Analysis 15, no. 2: 159-172 (1995).

[13] Granville, Andrew, and Greg Martin. *Prime number races*. The American Mathematical Monthly 113, no. 1: 1-33 (2006).

[14] Knapowski, Stanislaw, and Pál Turán. *Comparative prime-number theory. I*. Acta Mathematica Hungarica 13, no. 3: 299-314 (1962).

[15] Martin, Greg, and Justin Scarfy. *Comparative prime number theory: a survey*. arXiv:1202.3408 [math.NT] (2012).

[16] Broughan, Kevin. *Equivalents of the Riemann Hypothesis: Volume 2, Analytic Equivalents*. Vol. 165. Cambridge University Press (2017).

[17] Lemke Oliver, Robert J., and Kannan Soundararajan. *Unexpected biases in the distribution of consecutive primes*. Proc. Natl. Acad. Sci. USA 113, no. 31: E4446–E4454, (2016).

[18] Terence Tao: Structure and Randomness in the Prime Numbers, UCLA slides.

[19] Puszkarz, Waldemar. *Statistical Bias in the Distribution of Prime Pairs and Isolated Primes*. arXiv:1807.00406 [math.NT] (2018).

[20] Puszkarz, Waldemar. *Singly and doubly even multiples of 6 and statistical biases in the distribution of primes*. ResearchGate publication (2021).

[21] Mathematica, Wolfram Research Inc., `http://www.wolfram.com`.

[22] The PARI Group, PARI/GP, Univ. Bordeaux, `http://pari.math.u-bordeaux.fr/`.

Email address: `psi_bar@yahoo.com`