MATHEMATICS OF COMPUTATION
VOLUME 61, NUMBER 203
JULY 1993, PAGES 215-224

ESTIMATES OF THE LEAST PRIME FACTOR OF A BINOMIAL COEFFICIENT

P. ERDŐS, C. B. LACAMPAGNE, AND J. L. SELFRIDGE

Dedicated to the memory of D. H. Lehmer

ABSTRACT. We estimate the least prime factor p of the binomial coefficient $\binom{N}{k}$ for $k \geq 2$. The conjecture that $p \leq \max(N/k, 29)$ is supported by considerable numerical evidence. Call a binomial coefficient **good** if $p > k$. For $1 \leq i \leq k$ write $N - k + i = a_i b_i$, where b_i contains just those prime factors $> k$, and define the **deficiency** of a good binomial coefficient as the number of i for which $b_i = 1$. Let $g(k)$ be the least integer $N > k + 1$ such that $\binom{N}{k}$ is good. The bound $g(k) > ck^2/\ln k$ is proved. We conjecture that our list of 17 binomial coefficients with deficiency > 1 is complete, and it seems that the number with deficiency 1 is finite. All $\binom{N}{k}$ with positive deficiency and $k \leq 101$ are listed.

1. GOOD BINOMIAL COEFFICIENTS

Consider a sequence of $k \geq 2$ positive integers $\{n + i\} = n + 1, \ldots, n + k$, with $n + i = a_i b_i$, where $p | a_i$ implies $p \leq k$, and $p | b_i$ implies $p > k$ for any prime p; i.e., the a_i have the prime factors up to k, and the b_i have all the larger prime factors.

Since the binomial coefficient $\binom{n+k}{k}$ is an integer, $k! \mid \prod a_i$. We are concerned with the least prime factor of any number in the sequence, except for those composing $k!$, that is, the least prime factor of $\binom{n+k}{k}$.

Definition 1A. A sequence of consecutive integers $a_1 b_1, \ldots, a_k b_k$, where $p | a_i$ implies $p \leq k$, and $p | b_i$ implies $p > k$, is said to be **good** if $\prod a_i = k!$.

We denote the least prime factor of m by $p(m)$. We frequently find it convenient to write $N = n + k$.

Definition 1B. The binomial coefficient $\binom{N}{k}$ is said to be **good** if $p(\binom{N}{k}) > k$.

Note that Definition 1B is equivalent to stating that the k-sequence $\{N - k + i\}$ is good or that $\prod a_i = k!$ or that $\gcd(\binom{N}{k}, k!) = 1$. Further, when k is fixed and $\binom{N}{k}$ is good, we say that N is good with respect to k.

Ecklund, Erdős, and Selfridge [1] studied the function $g(k)$, the least integer $N > k + 1$ such that $p(\binom{N}{k}) > k$, so $g(k)$ is the least N which is good with respect to k. They showed that $g(k) > 2k$ for $k > 4$ and established weak upper and lower bounds on $g(k)$. They determined $g(k)$

Received by the editor July 27, 1992 and, in revised form, December 11, 1992..
1991 *Mathematics Subject Classification*. Primary 11B65, 11N37.

TABLE 1. Relative minima and maxima of $g(k)$, $k \leq 149$

Relative minima			Maxima		Maxima	
k	$d > 0$	$g(k)$	k	$g(k)$	k	$g(k)$
2		6	2	6	61	2237874623
4	2	7	3	7*	71	3184709471
5	1	23	5	23*	72	4179979724
8	2	44	6	62*	73	1 5780276223
10	3	46	7	143*	74	1 9942847999
11	4	47	9	159	75	4 8899668971
12	2	174	12	174**	83	79 7012560343
14	2	239	13	2239	89	352 4996442239
16	3	241	17	5849	103	509 2910127863
28	9	284	20	43196	104	600 3175578749
33	3	6459	24	193049	107	626 0627365739
35		37619	29	240479	108	974 6385386989
40		85741	31	341087	109	7324 5091349869
42	2	96622	32	371942*	110	9479 4806842238
52		366847	38	487343	111	22226 1611307119
58		4703099	39	767919	113	51796 8108138869
99		5675499	41	3017321*	114	59819 9028602614
100		3935600486	43	24041599	115	1271435 6616655615
102		17 5209712494	44	45043199	139	2597202 7636644319
106		48 8898352367	47	232906799	141	6333152 3816662671
135		315 7756005623				
136		413 8898693368				
148		1180 8400809148				

$d = d(g(k), k)$ $\quad$ $^* d = 1$ $\quad$ $^{**} d = 2$

for $2 \leq k \leq 40$ and $k = 42, 46, 52$, and showed by direct search that $g(k) > 2.5 \times 10^6$ for all other $k \leq 100$. After discussion with us, Scheidler and Williams [5] found all values of $g(k)$ for $k \leq 140$, using the new open architecture sieve at the University of Manitoba. The list of these values appears in [5], and the sieving continues. In Table 1 we present $g(k)$ where $g(k) \leq g(t)$ for $k < t < 150$ and where $g(k) \geq g(t)$ for $t < k$. No doubt, $g(k)$ increases faster than polynomially and surely $g(k) < (1+c)^{\pi(k)}$, but we have no proof. When k is large, it is clear to every right-thinking person that $\binom{N}{k}$ has a prime factor in $(k/2, k)$ for every $N < \exp(ck/\ln k)$. It seems that $g(k)$ increases very irregularly, and no doubt,

$$\limsup g(k+1)/g(k) = \infty, \qquad \liminf g(k+1)/g(k) = 0.$$

Note that $g(29)/g(28) > 846$, and $g(99)/g(98) < 1/1872$.

Up to $k = 148$, $g(k+1) - g(k) = 0$ or 1 only at $k = 3, 10, 18$ and 36.

It was shown in [1] that there is an absolute constant $c > 0$ for which $g(k) > k^{1+c}$. Since c is small, the following is an improved lower bound.

Theorem 1. *There holds* $g(k) > c_1 k^2/\ln k$, *for some absolute constant* $c_1 > 0$.

Proof. We first show that if $k^{7/4} \leq N < c_1 k^2/\ln k$, where $k > k_0(c_1)$, then $\binom{N}{k}$ has a prime factor p satisfying

$$k/2 < p < k/2 + k^{3/4}. \tag{1}$$

Begin by noting that Ingham [3] proved that the number of primes p satisfying

(1) is greater than $k^{3/4}/(c_2 \ln k)$ for some absolute constant c_2. By an averaging argument, there are two primes p_1 and p_2 satisfying

$$k/2 < p_1 < p_2 < p_1 + c_2 \ln k < k/2 + k^{3/4} \tag{2}$$

and $p_1 < k/2 + k^{3/4} - 2$. Now we show that for $k^{7/4} \le N < c_1 k^2/\ln k$, either p_1 or p_2 divides $\binom{n+k}{k}$. Let tp_1 be the largest multiple of p_1 which is less than or equal to n, so that $n < (t+1)p_1$. It follows immediately from (2) that if $tp_1 \le n - 2k^{3/4} + 4$, then $(t+2)p_1 \le n + k$. Thus $p_1 | \binom{n+k}{k}$. Therefore, we can assume

$$tp_1 > n - 2k^{3/4} + 4. \tag{3}$$

We wish to show that $tp_2 > n$ and $(t+1)p_2 \le n + k$, which imply that $p_2 | \binom{n+k}{k}$.

Observe that $p_1 \le k$ and $k(k^{3/4} - 1) = k^{7/4} - k \le n < (t+1)p_1$ imply that $t > k^{3/4} - 2$. Thus, (2) and (3) yield $tp_2 \ge tp_1 + 2t > n - 2k^{3/4} + 4 + 2k^{3/4} - 4 = n$ and

$$(t+1)p_2 < (t+1)p_1 + (t+1)c_2 \ln k < n + k/2 + k^{3/4} + tc_2 \ln k. \tag{4}$$

But $t \le n/p_1 < 2n/k < 2c_1 k/\ln k$. Put $\delta = 1/2k_0^{1/4}$. By (4) we have

$$(t+1)p_2 < n + k/2 + k^{3/4} + 2c_1c_2k < n + k \qquad \text{if} \quad c_1 < (\tfrac{1}{4} - \delta)/c_2 .$$

Now only the case $k^{1+c} < n + k < k^{7/4}$ remains to be considered. For such n we prove that $\binom{n+k}{k}$ has a prime factor p satisfying $k/2 < p < k/2 + k^{7/8}$. To see this, let p be any prime in the interval $(k/2, k/2 + k^{7/8})$ and let $t_p p$ be the largest multiple of p less than or equal to n. If $t_p p \le n - 2k^{7/8}$, then $(t_p + 2)p \le n + k$, and hence $p | \binom{n+k}{k}$. Thus, a prime p is unusable only if $n - 2k^{7/8} < t_p p \le n$, that is, $(n - 2k^{7/8})/t_p < p \le n/t_p$. Thus, for fixed t_p there are at most $2k^{7/8}/t_p$ unusable primes with the same multiplier t_p. Next we estimate the number of possible values of t_p. Clearly, $t_p < n/(k/2)$ and

$$t_p > \frac{n - 2k^{7/8}}{k/2 + k^{7/8}} > \frac{2n}{k + 3k^{7/8}} \quad \text{provided that} \quad n > 2k + 6k^{7/8}.$$

Thus, the number of possible values of t_p is at most

$$\frac{2n}{k} - \frac{2n}{k + 3k^{7/8}} < \frac{6n}{k^{9/8}},$$

so the number of unusable primes is at most

$$\frac{2k^{7/8}}{t_p} \cdot \frac{6n}{k^{9/8}} < 6k^{3/4}\left(1 + \frac{3}{k^{1/8}}\right).$$

But in $(k/2, k/2 + k^{7/8})$ there are, by Ingham's result, more than $k^{7/8}/(c \ln k)$ primes, which is of higher order than $k^{3/4}$. Thus there are many usable primes.

This completes our proof that $g(k) > c_1 k^2/\ln k$. In fact, our proof can be easily modified to show that if $n < c_1 k^2/\ln k$, then there are more than

$c_3k/\ln k$ primes in $(k/2, k)$ which divide $\binom{n+k}{k}$. No doubt this still holds for much larger values of N, but we cannot prove this. □

Theorem 2. *If neither $k+1$ nor $k+2$ is prime, then $g(k) \geq (k+1)q - 1$, where q is the largest prime power divisor of $k+2$.*

Proof. Suppose $\binom{N}{k}$ is good. Lemma 1 below implies that $N \equiv -1 \bmod (k+1)$ and $N \equiv -1$ or $-2 \bmod q$. Then $N \equiv -1$ or $k \bmod (k+1)q$. But $N > k$, so $N \geq (k+1)q - 1$. □

Corollary 1. *If $k+1$ is composite and $k+2 = p^a$ $(a > 1)$, then $g(k) \geq k^2 + 3k + 1$.*

Examples. $g(7) \geq 71$, $g(14) \geq 239$ (=!), $g(23) \geq 599$, $g(25) \geq 701$.

Corollary 2. *If $k+1$ is composite and $k+2 = 2p^a$ $(a \geq 1)$, then $g(k) \geq (k^2+3k)/2$.*

Examples. $g(8) \geq 44$ (=!), $g(20) \geq 230$, $g(24) \geq 324$, $g(32) \geq 560$, $g(44) \geq 1034$.

We confidently conjecture that $g(k) > k^2$ for $k > 16$, with the notable exception $g(28) = 284$. We conjecture that $g(k) > k^3$ for $k > 35$, with a charming near exception $g(99) = 5675499$. It would not surprise us if $g(k) > k^5$ for $k > 100$.

Lemma 1. *The prime p does not divide $\binom{N}{k}$ if and only if each of the digits of N base p is greater than or equal to the corresponding digit of k base p, or, equivalently, $N \bmod p^a \geq k \bmod p^a$ for all $a \leq b$ such that $p^{b-1} \leq k < p^b$, where $n \bmod m := n - m\lfloor \frac{n}{m} \rfloor$.*

The proof of Lemma 1 is easy and well known, but we include it for completeness.

Proof. If the digit in the p^a column of N base p is greater than or equal to the corresponding digit of k base p, then there are the same number of multiples of p^a among $n+1, \ldots, n+k$ as there are among $1, \ldots, k$. If this holds for each a, then p does not divide $\binom{N}{k}$. Otherwise, there are more multiples of p^a among $n+1, \ldots, n+k$ than among $1, \ldots, k$, for some a, and p divides $\binom{N}{k}$. For the alternative version, if $N \bmod p^a \geq k \bmod p^a$, then $n+1, \ldots, n+k$ have the same number of multiples of p^a as $1, \ldots, k$ have, for each a, and thus p does not divide the binomial coefficient $\binom{N}{k}$. □

2. Least prime factors of binomial coefficients

The main problem: Estimate $p(\binom{N}{k})$, the least prime factor p of $\binom{N}{k}$.

Case 1. $N > k^2$.

Conjecture. We conjecture [6] that in this case, $p(\binom{N}{k}) \leq N/k$ except for $\binom{62}{6}$.

If any b_i is composite, then $p(\binom{N}{k}) \leq \sqrt{b_i} \leq \sqrt{N} < N/k$. If $\binom{N}{k}$ is not good, then $p(\binom{N}{k}) \leq k < N/k$, satisfying our conjecture. For the remainder of Case 1, we will only consider good binomial coefficients.

To see why we must allow for $p(\binom{N}{k}) = N/k$, suppose

$$(5) \qquad a_i < k, \quad \text{for all} \quad i < k \quad \text{and} \quad a_k = k .$$

Then $p(\binom{N}{k}) = N/k$ if and only if the b_i are all prime. Sequences satisfying (5) were characterized in [2]. For a given k the a_i are always a permutation of $1, \dots, k$ and the number of such permutations is always a power of 2. For example, when $k = 2459$, there are five operations which we call 2^{11} swap, 3^7 swap, 7^4 swap, 2459 swap (followed by symmetric flip), and 1229 + 2 twin prime double swap, yielding 32 permutations which can be written immediately. A more mundane example is provided by $k = 5$, which admits a 2^2 swap, yielding just two permutations. When $n = 210$ for instance, the a_i are 1, 4, 3, 2, 5 and the b_i are 211, 53, 71, 107, 43.

If none of the $b_i = 1$ and $k|n+i$, then $p((n+i)/k) \le (n+i)/k \le N/k$. So our conjecture remains to be verified only for those good binomial coefficients where at least one of the $b_i = 1$. If N is very large compared to k (for example, if $N > k!$), then $b_i > 1$ for all i.

Schinzel has conjectured that for every k there is an $N \ge 2k$ such that the b_i are all prime. We note that $\binom{N}{k}$ can be the product of fewer than k primes all greater than k. For example, when $k = 25$ and $n = 2080$, then $b_{10} = b_{13} = b_{20} = 1$, and the other b_i are all prime. Since $\binom{2105}{25}$ is good, it is the product of these 22 primes, each greater than 25.

Definition 2. If $\binom{N}{k}$ is good, we call the number of i such that $b_i = 1$ the **deficiency** of $\binom{N}{k}$ and use the notation $d(N, k)$ or d in context.

Thus $d(2105, 25) = 3$ and $d(215, 5) = 0$.

Theorem 3. *If* $\binom{N}{k}$ *is good, and* $N > c_4 2^k \sqrt{k}$ (*where* $c_4 < 0.4$ *when* $k \ge 94$), *then* $d(N, k) = 0$. (For $k < 94$, see Tables 2 and 3.)

The proof of Theorem 3 depends on the following lemma.

Lemma 2. *If* $\binom{N}{k}$ *is good, then* $a_i | i\binom{k}{i}$.

Proof of Lemma 2. We use the alternative form $a_i = a_{k-j}$, $0 \le j < k$. If $p^a | a_{k-j}$, then $p^a | n+k-j$, so $N \bmod p^a = j \bmod p^a$. Since $\binom{N}{k}$ is good, $j \bmod p^a \ge k \bmod p^a$ by Lemma 1. Put $qp^a \le j < (q+1)p^a$; there are q multiples of p^a among $1, \dots, j$. But since $k \bmod p^a \le j \bmod p^a$, there are $q+1$ multiples of p^a among the $j+1$ numbers $k, k-1, \dots, k-j$. The same argument shows that there is one more multiple of p^d among $k, \dots, k-j$ than among $1, \dots, j$ for any $d \le a$. Each $d \ge 1$ contributes a count of one to the power of p dividing

$$\frac{k(k-1)\cdots(k-j+1)(k-j)}{j!} = (k-j)\binom{k}{j},$$

which is therefore a multiple of p^a. Thus, $a_{k-j} | (k-j)\binom{k}{k-j}$ for $0 \le j < k$. □

Proof of Theorem 3. By Lemma 2, $a_i \le i\binom{k}{i}$. The maximum for this bound occurs when $i = \lfloor (k+1)/2 \rfloor$. If k is even, $a_i \le (k/2)\binom{k}{k/2} < 2^k\sqrt{k/2\pi}$, using Stirling's formula. For k odd, $k \ge 95$, $a_i < 0.4\, 2^k\sqrt{k}$. So if $N - k$ is larger than this maximum, then none of the b_i equals 1. □

TABLE 2. Binomial coefficients $\binom{N}{k}$ with deficiency $d > 1$

d	k	N	d	k	N
9	28	284*	2	4	7*
				8	44*
4	11	47*		10	74
				12	174*
3	10	46*+1		14	239*
	16	241*		27	5179
	25	2105*		28	8413+1
	27	1119*		42	96622*
	33	6459*			

* $N = g(k)$ +1: N and $N+1$ are both solutions

We ran a program (based on a more precise bound on a_i) to find binomial coefficients with positive deficiency for $k \le 101$ and all possible N. Tables 2 and 3 give the results of our computer search.

Notice that in Table 3 below there is a positive deficiency at $k = 100$ near the top of our search range. But for $k \ge 20$ our published remark [3, p. 523] that for each k there seems to be an N such that $d(N, k) = 1$ is way off base. In fact, the first k for which $d(N, k) = 0$ for all good $\binom{N}{k}$ is 13.

We conjecture that $k = 42$ is the last k with $d(N, k) > 1$, and from looking at the tables, one gets the idea that there are only finitely many binomial coefficients with $d(N, k) > 0$. The only values of $k \le 42$ for which $d = 0$ for all good N are 13, 20, 21, 22, 24, 29, 31, 37. However, for $42 < k \le 101$ there are only 13 values of k with $d(N, k) = 1$ for some N, and none with $d > 1$ for any N.

Notice also that whenever we found $d(N, k) > 1$, then $d(g(k), k) > 1$. There are 11 values of $k \le 46$ with $d(g(k), k) = 1$. We conjecture that $d(g(k), k) = 0$ for $k > 46$; this has been checked up to $k = 149$.

If $N > k^2$, then $\binom{62}{6}$ is the only known exception to $p(\binom{N}{k}) \le N/k$. If there were to be any further exceptional binomial coefficients $\binom{N}{k}$ in Case 1, the following four conditions would all have to be satisfied: 1. $d(N, k) > 0$, so $k > 101$. 2. $N \ge g(k)$. 3. If $b_i \ne 1$, then b_i is prime. 4. If $b_i \ne 1$, then $a_i \le k$.

Remark 1. The probability that one of the k consecutive integers of $\{n + i\}$ is divisible by a prime just larger than k is close to 1. Thus, if $N > k^3$, there will almost always be a prime factor less than or equal to N/k.

Case 2. $2k \le N \le k^2$.

Lemma 3. *We have* $p(r) | \binom{rk}{k}$.

Proof. Let $p = p(r)$ and $p^a \| rk$. Then $rk \bmod p^a = 0 < k \bmod p^a$. Thus, by Lemma 1, $p(r) | \binom{rk}{k}$. So at $N = k^2$, $\binom{N}{k}$ has a prime factor $p(k) \le k = N/k$. □

Remark 2. If $N = k^2 - 1$, then $p(\binom{N}{k}) < N/k$.

Proof. Let $p = p(k-1)$. Then $N \bmod p = 0$ and $k \bmod p = 1$. By Lemma 1, $p | \binom{N}{k}$, so $p(\binom{N}{k}) \le k - 1 < N/k$. □

TABLE 3. Binomial coefficients with deficiency $d(N, k) = 1$

k	Values of N
3	7*
4	13+1
5	23*
6	62*
7	143*
8	89, 143
9	319, 509
10	94+1, 122, 187, 286, 319, 362, 367, 635
11	1391
12	188, 237, 797, 3967, 5549
14	719
15	719* 799, 2319, 3967, 5471, 10015
16	566, 1241, 1591+1, 2293+1, 8017+1, 11447, 20599, 25748, 102967+1
17	74267
18	5718, 20599, 36474, 350074
19	2099*
23	35423*
26	76922, 177659
27	2239, 49279, 3834683, 70204063
28	2239, 20479, 22813, 49279, 150718, 153404, 218974, 225244, 281533, 434719, 469214, 1285213+1, 1352093, 2713213+1, 16046653, 22465309, 70204063, 187210813+1
30	3834687, 4750206, 13572799, 17235294, 26613311, 40820414+1, 775587614
32	371942* 1828859
34	69614* 657719
35	236663, 1869047
36	239797, 336621, 1828863, 1869047, 4352423, 69537661+1, 97582118, 261813614, 300402671, 1296447917+1, 2634716718, 7425110718
38	40465463
39	5776999, 13161839, 151479719, 228986799, 1 1732392319
40	96620, 4171067, 37396798, 117929965, 228986798, 652046569, 698703290, 1 8379537195+1
41	3017321*
46	692222*
48	2668 7672013624
52	17692343, 23836084, 364728823, 208 3691314934, 2508 1469531324
57	561133817, 618031933, 49 8565957819
58	16794619, 28676734, 1589319934, 2052428219, 2385269114, 4398350459, 1 2678949498+1, 4 2659680319, 49 8565957819, 825 1483160059
59	1 2678949499
65	11 3642398319, 21 5662310621, 174 8704870373
66	13814 3173371874, 34549607 6602971874
70	691 0212567374, 15361 9118501974, 170014 0546689622, 31407132 6474420095
78	4722 9486938863
95	2 1619814065 5426106847
96	709762 7778251372, 8 0351139737 6448532847
100	12155716 2475124854

* $N = g(k)$ +1: N and $N + 1$ are both solutions

When N/k is small, it is unacceptably small as a bound for $p(\binom{N}{k})$. In fact, there are infinitely many binomial coefficients with $p(\binom{N}{k}) = 3 > N/k$ when $N < 3k$.

Theorem 4. *For each* $k > 2$ *there is an* N, $2k \le N < 4k$, *such that* $p(\binom{N}{k}) > N/k$.

Proof. Let $2^{r-1} \le k < 2^r$ and let $N = 2^r + k$. Then by Lemma 1, $\binom{N}{k}$ is odd, since $N \equiv k \bmod 2^a$ for $a = 1, \dots, r$. But if $\binom{N}{k}$ is odd and $2k < N < 3k$, then $p(\binom{N}{k}) \ge 3 > N/k$. Now there will always be an $N = 2^r + k$, $2k < 2^r + k < 3k$, except when $2^r = 2k$. But when $2^{r-1} = k$, then $\binom{N}{k}$ is odd for all values of N such that $3k \le N < 4k$. Now take $N = 3^s + k$, $2 \cdot 3^s - 1$ or $2 \cdot 3^s + k$, whichever is in the region. By Lemma 1, 3 does not divide $\binom{N}{k}$. So $p(\binom{N}{k}) \ge 5 > N/k$. □

We now combine Cases 1 and 2.

Definition 3. If $p(\binom{N}{k}) > N/k$, then $\binom{N}{k}$ is said to be **exceptional**.

It would be very interesting if someone could prove our conjecture that the number of exceptional $\binom{N}{k}$ with $p > 17$ is finite. We wrote a program to find all exceptional $\binom{N}{k}$ where $p > 5$ and $k \le 12000$. The most unusual exception is for $p = 29$: $p(\binom{284}{28}) = 29$. We have found one exception for $p = 23$: $p(\binom{474}{66}) = 23$. There appear to be exactly two exceptions for $p = 19$: $\binom{62}{6}$ and $\binom{959}{56}$.

We have two further conjectures:

1. The only exceptional $\binom{N}{k}$ with $p > 17$ are these four.

2. $p(\binom{N}{k}) \le N/k$ if $N > 17\frac{1}{8}k$.

Notice that 1. is a bit stronger and implies 2.

We have found eight exceptional $\binom{N}{k}$ with $p = 17$: $\binom{241}{16}$, $\binom{439}{33}$, $\binom{317}{56}$, $\binom{482}{130}$, $\binom{998}{256}$, $\binom{998}{260}$, $\binom{14273}{896}$, $\binom{13277}{900}$, and two near misses: $\binom{239}{14}$ and $\binom{956}{56}$.

Our program gave only one output for $331 < k < 625$, namely, $p(\binom{3574}{406}) = 13$, and no $p > 13$ for any k other than those already listed. Thus, at this point in time, it is still possible that $p(\binom{N}{k}) \le \max(N/k, 13)$, with the twelve exceptions listed above.

3. A RELATED PROBLEM

We now turn our attention to a related problem. We study the number of indices i where $b_i = 1$ without requiring that $\binom{N}{k}$ be good.

Theorem 5. *Assume* $N \ge 2k$ *and denote by* $f(N, k)$ *the number of indices* i *for which* $b_i > 1$. *Then* $f(N, k) \ge (1-\epsilon)\pi(k)$ *for* $k > k_0(\epsilon)$.

Proof. We use the strong form of the prime number theorem: the number of primes p such that $N - k < p \le N$ is greater than $(1-\epsilon)\pi(k)$ if $N < k^{1+c}$ for any $c \le 1/3$. This immediately gives the theorem for $2k \le N < k^{4/3}$.

In fact, the contribution to $b_i > 1$ from primes $k < p \le N/2$ becomes significant even when $N = (2+\delta)k$. When $N = 3k$, use the quoted theorem for $k < p \le N/2$ and $2k < p \le N$ to get $f(N, k) \ge (\frac{3}{2} - \epsilon)\pi(k)$ for large k, and similarly one can get $f(N, k) > (1+\mu)\pi(k)$ when $N \ge (2+\delta)k$.

Now if $N \geq k^{4/3}$, we will show that $f(N, k) > k/8$ for k large enough. First observe that $\binom{N}{k} \geq N^k/k^k$. The well-known result that $p^\alpha \leq N$ when $p^\alpha \| \binom{N}{k}$ follows easily from Lemma 1. Thus we have

$$N^k/k^k \leq \binom{N}{k} < N^{\pi(k)} N^{f(N,k)}. \tag{6}$$

Now $N^{3/4} \geq k$ and, for $k > 10^4$, $\pi(k) < k/8$. Thus, from (6), using $N/k \geq N^{1/4}$, we get $N^{k/4} \leq N^k/k^k < N^{k/8} N^{f(N,k)}$ and thus $f(N, k) > k/8 > \pi(k)$, which finishes the proof of the theorem. □

It would be quite difficult to give good explicit inequalities for $f(N, k)$.

Corollary 3. *If $N \geq 2k$, there are at least $(1 - \epsilon)\pi(k)$ primes greater than k dividing $\binom{N}{k}$. In fact, when $N \leq k^2$, the count of primes greater than k is $f(N, k)$, since all b_i are prime.*

Notice that $f(21, 10) = 3 < \pi(20) - \pi(10) = 4$. Is it true that for every t there are integers N and k for which $f(N, k) \leq \pi(2k) - \pi(k) - t$? For $f(213, 100)$ we can take $t = 3$.

We conjecture that there are examples with arbitrarily large t. Suppose there is a large gap in the primes between p_r and p_{r+1}. Take $N = p_{r+1} - 1$ and $2k = p_r + 1$. Then $f(N, k)$ is the number of primes between $N - k$ and N, plus the number of primes between k and $N/2$. Since $\pi(N) = \pi(2k)$, we get

$$f(N, k) = \pi(2k) - \pi(k) - (\pi(N - k) - \pi(N/2)).$$

As an example, the first gap of 320 between consecutive primes is listed by Lander and Parkin [4]. Put $N = 2300942868$ and $k = 1150471275$. The primes between $N/2$ and $N - k$ are 1150471000 + 297, 307, 319, 369, 373, 393, 417, giving an example with $t = 7$.

From the strong form of the prime number theorem, we deduce that only the interval $2k < N < 2k + k^c$ has to be searched. No doubt the only possible examples with large t lie in the interval $2k < N < 2k + c \ln k$, where $c = c(t) > 0$.

Acknowledgment

We thank the referee for his careful reading.

Bibliography

1. E. F. Ecklund, P. Erdős, and J. L. Selfridge, *A new function associated with the prime factors of* $\binom{n}{k}$, Math. Comp. **28** (1974), 647–649.
2. P. Erdős, C. B. Lacampagne, and J. L. Selfridge, *Prime factors of binomial coefficients and related problems*, Acta Arith. **49** (1988), 507–523.
3. A. E. Ingham, *On the differences between consecutive primes*, Quart. J. Math. Oxford **8** (1937), 255–266.
4. L. J. Lander and T. R. Parkin, *On first appearance of prime differences*, Math. Comp. **21** (1967), 483–488.
5. R. Scheidler and H. C. Williams, *A method of tabulating the number-theoretic function* $g(k)$, Math. Comp. **59** (1992), 251–257.

6. J. L. Selfridge, *Some problems on the prime factors of consecutive integers*, Abstract 747-10-9, Notices Amer. Math. Soc. **24** (1977), A456–A457.

(P. Erdős) c/o Dr. R. L. Graham, AT&T Bell Laboratories, Murray Hill, New Jersey 07974

Office of Research, U. S. Department of Education, Washington, DC 20008
E-mail address: clacampa@nsf.gov

Department of Mathematics, Northern Illinois University, DeKalb, Illinois 60115