MATHEMATICS OF COMPUTATION
VOLUME 53, NUMBER 188
OCTOBER 1989, PAGES 721-741

The Probability that a Random Probable Prime is Composite*

By Su Hee Kim and Carl Pomerance**

Abstract. Consider a procedure which (1) chooses a random odd number $n \leq x$, (2) chooses a random number b, $1 < b < n-1$, and (3) accepts n if $b^{n-1} \equiv 1 \pmod{n}$. Let $P(x)$ denote the probability that this procedure accepts a composite number. It is known from work of Erdös and the second author that $P(x) \to 0$ as $x \to \infty$. In this paper, explicit inequalities are established for $P(x)$. For example, it is shown that $P(10^{100}) < 2.77 \times 10^{-8}$ and that $P(x) \leq (\log x)^{-197}$ for $x \geq 10^{10^5}$.

Introduction. Suppose one wants to produce a random prime $p \leq x$, drawn with the uniform distribution. One possible solution is to choose a random number n, $1 < n \leq x$, and apply a test to n that can tell if it is prime or composite. This procedure is repeated independently until a prime is found. By the prime number theorem, the expected number of trials until a prime is drawn is about $\log x$. If one wishes to choose an odd prime, the trials n may be restricted to odd numbers. The expected number of trials is then about $\frac{1}{2}\log x$.

There are many algorithms which can be used to decide if n is prime or composite. However, using the Fermat congruence is a very cheap test that is usually recommended as a preliminary procedure before a more time-consuming test is attempted. Namely, one chooses a random number b, $1 < b < n-1$, and checks if $b^{n-1} \equiv 1 \pmod{n}$. If n is prime, then this congruence will hold. If this congruence holds, then n is called a *probable prime to the base b*. This procedure can prove an input n is composite, but cannot establish primality.

How good is this test at producing random primes? Specifically, let $P(x)$ denote the probability that n is composite given that

(i) n is chosen at random with $1 < n \leq x$, n odd,

(ii) b is chosen at random with $1 < b < n-1$, and

(iii) n is a probable prime to the base b.

It is well known that there are some composite numbers n, namely the Carmichael numbers, such that (iii) holds for every b coprime to n. However, Carmichael numbers are rare, so presumably the odds of choosing one in (i) is small. In fact, extensive numerical evidence suggests that $P(x)$ is quite small when x is large.

In practice, if a large random number n passes a random probable prime test, then one strongly conjectures that n is prime. As Henri Cohen has colorfully put it, such an n can be considered an "industrial grade prime." That is, although n

Received November 8, 1988.

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 11Y11; Secondary 11A51, 11N56.

*This paper is based on the first author's master's thesis at the University of Georgia.

**Supported in part by an NSF grant.

0025-5718/89 $1.00 + $.25 per page

has not been proved prime, the probability it is composite is so small that n might be used as a prime for industrial (cryptographic) purposes.

We do know theoretically that if x is sufficiently large, then $P(x)$ is small. Indeed, from Theorem 2.2 in Erdös and Pomerance [2], we have that

$$P(x) \leq \exp(-(1+o(1))\log x \log\log\log x/\log\log x) \tag{1.1}$$

as $x \to \infty$. In particular, $\lim P(x) = 0$.

Although we have the strong inequality (1.1) and the practical experience of many people to draw on, we still do not have any good estimate for $P(x)$ for various finite values of x. The problem is the "$o(1)$" in (1.1) which renders the inequality computationally useless.

In this paper we replace the asymptotic inequality (1.1) with a weaker, but explicit inequality. The argument is loosely based on the proof in [2] of (1.1) above, but a number of difficulties are encountered. For delicate estimates involving prime numbers, we use the results of Rosser and Schoenfeld [5]. However, the rest of our work is elementary and involves only moderate computation.

We prove that

$$P(x) \leq (\log x)^{-197} \quad \text{for } x \geq 10^{10^5}. \tag{1.2}$$

For smaller values of x, our results are summarized in Table 1. To find an upper estimate for $P(x)$ for some x not in the table with $10^{60} < x < 10^{10^5}$, one can find the largest x_0 in the table with $x_0 < x$ and multiply the estimate at x_0 by $\log x/\log x_0$.

It is highly likely that our upper bounds can be improved upon. To some extent, it is a matter of how hard one is willing to work. Sometimes we make trivial estimates for simplicity, but a more careful estimation would give a better result.

One possible way to gain an improvement is to replace the Fermat congruence with the strong probable prime test of Selfridge. This test is just as easy to perform and it "lies" less frequently about composite numbers. To describe this test, let $n > 1$ be an odd number. First one computes s, t with $n-1 = 2^s t$ and t odd. Next, one chooses a number b, $1 < b < n-1$. The number n passes the test (and is called a *strong probable prime to the base* b) if either

$$b^t \equiv 1 \pmod{n} \quad \text{or} \quad b^{2^i t} \equiv -1 \pmod{n} \quad \text{for some } i < s. \tag{1.3}$$

Every odd prime must pass this test. Moreover, Monier [3] and Rabin [4] have shown that if $n > 1$ is an odd composite, then the probability that it is a strong probable prime to a random base b, $1 < b < n-1$, is less than $\frac{1}{4}$.

Let $P_1(x)$ denote the same probability as $P(x)$, except that (iii) is changed to

(iii)′ n is a strong probable prime to the base b.

Based on the Monier-Rabin theorem, one is tempted to say that $P_1(x) \leq \frac{1}{4}$, but as pointed out in [1], this reasoning is fallacious. In fact, if α is the probability that a random odd number up to x is prime and β is an upper bound for the probability that an odd composite number up to x passes a random strong probable prime test, then

$$P_1(x) \leq \frac{(1-\alpha)\beta}{\alpha + (1-\alpha)\beta}. \tag{1.4}$$

TABLE 1***

x	Upper bound for $P(x)$	x	Upper bound for $P(x)$
1.0E + 60	7.16E − 2	1.0E + 0300	5.8E − 0029
1.0E + 70	2.87E − 3	1.0E + 0400	5.7E − 0042
1.0E + 80	8.46E − 5	1.0E + 0500	2.3E − 0055
1.0E + 90	1.70E − 6	1.0E + 0600	1.7E − 0068
1.0E + 100	2.77E − 8	1.0E + 0700	1.8E − 0082
1.0E + 110	4.03E − 10	1.0E + 0800	5.4E − 0096
1.0E + 120	5.28E − 12	1.0E + 0900	1.0E − 0109
1.0E + 130	7.54E − 14	1.0E + 1000	1.2E − 0123
1.0E + 140	1.08E − 15	1.0E + 2000	8.6E − 0262
1.0E + 150	1.49E − 17	1.0E + 3000	3.8E − 0397
1.0E + 160	1.81E − 19	1.0E + 4000	7.8E − 0537
1.0E + 170	2.27E − 21	1.0E + 5000	7.6E − 0680
1.0E + 180	2.76E − 23	1.0E + 6000	3.9E − 0820
1.0E + 190	3.26E − 25	1.0E + 7000	1.1E − 0951
1.0E + 200	3.85E − 27	1.0E + 8000	7.3E − 1081
		1.0E + 9000	1.7E − 1207
		1.0E + 10000	*1.6E − 1331*
		1.0E + 100000	1.3E − 10584

From Monier-Rabin, we have that $\beta \leq \frac{1}{4}$. Thus all we get from this theorem is that

$$P_1(x) \leq \frac{1-\alpha}{1+3\alpha}. \tag{1.5}$$

If x is very large, then α is very small and so (1.5) is a quite weak result.

However, presumably much is lost using the worst case upper bound β. This is attained only for very special composites which, like Carmichael numbers, are rare.

The results of this paper also apply to $P_1(x)$, since we trivially have $P_1(x) \leq P(x)$. If one were to concentrate solely on $P_1(x)$, it is possible that considerably stronger estimates could be obtained. We remark that by using the formulas of Monier [3] for the number of b for which n is a probable prime, respectively strong probable prime, our estimates for $P(x)$ can be multiplied by $\frac{1}{2}$ when applied to $P_1(x)$.

Consider finally a procedure which chooses a random odd number $n \leq x$ and then performs k strong probable prime tests on n with k independently drawn random numbers b, $1 < b < n-1$. Let $P_k(x)$ denote the probability that this

***The notation aEn means $a \times 10^n$.

procedure accepts a composite number. Combining our results with the Monier-Rabin theorem, we have

$$P_k(x) \le 4^{-(k-1)} P_1(x)/(1-P_1(x)) \le 4^{-(k-1)} P(x)/(1-P(x)). \tag{1.6}$$

The popularly believed inequality is that $P_k(x) \le 4^{-k}$, but as we have seen, the reasoning for this is fallacious. However, if we have $P(x) \le \frac{1}{5}$, then (1.6) does imply that $P_k(x) \le 4^{-k}$ for every k. In particular, from the results of this paper, this inequality holds for all $x \ge 10^{60}$.

2. The Basic Method. Let

$$F(n) = \#\{b \in (\mathbf{Z}/n)^* : b^{n-1} \equiv 1 \pmod n\}.$$

If $n > 1$ is odd, then $b = \pm 1$ both satisfy $b^{n-1} \equiv 1 \pmod n$. Thus for these n, $F(n) - 2$ is the number of b, $1 < b < n-1$, with $b^{n-1} \equiv 1 \pmod n$. Also note that by Fermat's theorem, if p is a prime, then $F(p) = p - 1$. We thus have for $x \ge 5$,

$$\begin{aligned} P(x) &= \frac{\sum_{n \le x, n \text{ odd, composite}} (F(n)-2)}{\sum_{1 < n \le x, n \text{ odd}} (F(n)-2)} \\ &\le \frac{\sum_{n \le x, n \text{ odd, composite}} F(n)}{\sum_{2 < p \le x} (p-3)}, \end{aligned} \tag{2.1}$$

where here and throughout the paper, p denotes a prime.

Hence, to get an upper bound for $P(x)$, we shall be interested in obtaining a lower bound for $\sum_{2<p\le x}(p-3)$ and an upper bound for $\sum_{n \le x,\ n \text{ odd, composite}} F(n)$.

For this purpose we shall prove two theorems.

THEOREM 2.1. *For $x \ge 37$, we have*

$$\sum_{2<p\le x} (p-3) \ge \frac{x^2}{2(2+\log x)}.$$

THEOREM 2.2. *Suppose c, L_1 and L are arbitrary real numbers satisfying $\frac{1}{2} < c < 1$, $1 < L_1 < L$. Then for any $x > L^2$, we have*

$$\begin{aligned} &\sum_{\substack{n \le x, n \text{ odd,} \\ \text{composite}}} F(n) \\ &\le \frac{x^2}{4L_1} + \frac{x^2}{L}(1+\log L_1) + \frac{x^2}{2(L-1)} \cdot \left(\frac{L_1}{L-1} + 1\right)(2+\log L_1)^2 \\ &\quad + \frac{K_c}{1-c} x^{1+c} L^{2(1-c)} (1+\log L_1) \sum_{m \le L^2 L_1} \frac{\tau_{L_1}(m)}{m} \exp(2^{-c} f_c(m)), \end{aligned}$$

where

$$K_c = \exp\left(\sum_{p>2} \sum_{k=2}^{\infty} k^{-1} p^{-kc}\right),$$

$$f_c(m) = \prod_{p \mid m} (1-p^{-c})^{-1},$$

and $\tau_{L_1}(m)$ is the number of divisors of m up to L_1.

Before we prove Theorems 2.1 and 2.2, we state a theorem that is an immediate consequence of them and (2.1). Say that $g(x, c, L, L_1)$ is the right member of the inequality in Theorem 2.2.

THEOREM 2.3. *For all real numbers c, L and L_1 with $\frac{1}{2} < c < 1$, $1 < L_1 < L$ and for all $x > L^2 > 37$, we have*

$$P(x) \leq 2(2 + \log x)g(x, c, L, L_1)/x^2.$$

Thus, our upper bound for $P(x)$ depends on the choices of the variables x, c, L and L_1.

Now let us prove Theorems 2.1 and 2.2.

Proof of Theorem 2.1. Let $\pi(x)$ denote the number of primes not exceeding x. We have

$$\pi(x) > x/(\log x - \tfrac{1}{2}) \quad \text{for } x \geq 67 \tag{2.2}$$

and

$$\pi(x) < x/(\log x - \tfrac{3}{2}) \quad \text{for } x > e^{3/2} \ (x \geq 4.48169) \tag{2.3}$$

by (3.3) and (3.4) in [5, p. 69].

Using partial summation, (2.2) and (2.3), we have

$$\begin{aligned} \sum_{2<p\leq x} (p-3) &= (x-3)\pi(x) - \int_3^x \pi(t)\,dt \\ &> \frac{x(x-3)}{\log x - \frac{1}{2}} - \int_3^x \pi(t)\,dt. \end{aligned} \tag{2.4}$$

Now for $x \geq 245$,

$$\begin{aligned} \int_3^x \pi(t)\,dt &\leq \int_{245}^x \frac{t\,dt}{\log t - \frac{3}{2}} + \int_3^{245} \pi(t)\,dt \\ &= \frac{x^2}{2(\log x - \frac{3}{2})} - \frac{245^2}{2(\log 245 - \frac{3}{2})} \\ &\quad + \frac{1}{2}\int_{245}^x t\left(\log t - \frac{3}{2}\right)^{-2} dt + 7154 \\ &\leq \frac{x^2}{2(\log x - \frac{3}{2})} + \frac{1}{2}\int_{245}^x t\left(\log t - \frac{3}{2}\right)^{-2} dt. \end{aligned}$$

Let $S = \frac{1}{2}\int_{245}^x t(\log t - \frac{3}{2})^{-2}\,dt$. We have

$$\begin{aligned} S &< \frac{1}{4}x^2\left(\log x - \frac{3}{2}\right)^{-2} + \frac{1}{2}\int_{245}^x t\left(\log t - \frac{3}{2}\right)^{-3} dt \\ &< \frac{1}{4}x^2\left(\log x - \frac{3}{2}\right)^{-2} + \frac{1}{4}S \end{aligned}$$

because $\log t - \frac{3}{2} > 4$ for $t \geq 245$. Therefore,

$$S < \tfrac{1}{3}x^2(\log x - \tfrac{3}{2})^{-2},$$

so that

$$\int_3^x \pi(t)\,dt < \frac{x^2}{2(\log x - \frac{3}{2})} + \frac{x^2}{3(\log x - \frac{3}{2})^2}.$$

Putting this estimate in (2.4), we have for $x \geq 245$

$$\sum_{2<p\leq x} (p-3) \geq \frac{x(x-3)}{\log x - \frac{1}{2}} - \frac{x^2}{2(\log x - \frac{3}{2})} - \frac{x^2}{3(\log x - \frac{3}{2})^2}. \tag{2.5}$$

We replace the right side of (2.5) with the simpler expression $x^2/2(2+\log x)$ which is smaller for all $x > 20,000$. Moreover, we have checked numerically that

$$\sum_{2<p\le n-1} (p-3) > \frac{n^2}{2(2+\log n)}$$

for every integer n with $38 \le n \le 20,000$. Thus we have Theorem 2.1.

Proof of Theorem 2.2. Since $F(n)$ is the cardinality of a subgroup of $(\mathbf{Z}/n)^*$, we have that for any n, $F(n)|\phi(n)$, where ϕ is Euler's function. That is, $F(n) = \phi(n)/k$ for some integer $k > 0$.

Let $\mathbf{C}_k(x)$ denote the set of odd, composite $n \le x$ such that $F(n) = \phi(n)/k$, and let $C_k(x) = \#\mathbf{C}_k(x)$.

For any $x > L^2$ where $L > L_1 > 1$,

$$\begin{aligned}\sum_{\substack{n\le x, n \text{ odd},\\ \text{composite}}} F(n) &= \sum_{k=1}^{\infty} \sum_{n\in\mathbf{C}_k(x)} F(n) = \sum_{k=1}^{\infty} \sum_{n\in\mathbf{C}_k(x)} \frac{\phi(n)}{k} \\ &= \sum_{k\le L_1} \frac{1}{k} \sum_{n\in\mathbf{C}_k(x)} \phi(n) + \sum_{k>L_1} \sum_{n\in\mathbf{C}_k(x)} \frac{\phi(n)}{k} \\ &\le x \sum_{k\le L_1} \frac{C_k(x)}{k} + \frac{1}{L_1} \sum_{\substack{1<n\le x\\ n \text{ odd}}} (n-2) \\ &\le x \sum_{k\le L_1} \frac{C_k(x)}{k} + \frac{x^2}{4L_1}.\end{aligned} \tag{2.6}$$

It will thus be desirable to obtain an upper bound for $C_k(x)$. Three classes are considered to estimate $C_k(x)$ for $k \le L_1$:

(i) $n \le x/L$,

(ii) n is divisible by some prime $p > L$,

(iii) $n > x/L$ and every prime p in n is at most L.

Let $C_{k,1}(x)$, $C_{k,2}(x)$ and $C_{k,3}(x)$ denote the number of $n \le x$ counted by $C_k(x)$ for each class respectively. Thus,

$$C_k(x) \le C_{k,1}(x) + C_{k,2}(x) + C_{k,3}(x). \tag{2.7}$$

Obviously,

$$C_{k,1}(x) \le x/2L. \tag{2.8}$$

We now state a result that will be useful for classes (ii) and (iii). This result is (2.11) in [2].

LEMMA 2.4. *If $F(n) = \phi(n)/k$, then $\lambda(n)|k(n-1)$, where $\lambda(n)$ is the Carmichael universal exponent function; that is, $\lambda(n)$ is the least positive integer with $b^{\lambda(n)} \equiv 1 \pmod n$ for all integers b with $(b,n) = 1$.*

Let d be a natural number. We consider those n counted by $C_k(x)$ with $d|n$. If $d|n$, then $\lambda(d)|\lambda(n)$, so that the condition $\lambda(n)|k(n-1)$ from Lemma 2.4 implies $\lambda(d)|k(n-1)$. Thus, the number of n counted by $C_k(x)$ with $d|n$ is at most the number of composite numbers $n \le x$ with

$$n \equiv 0 \pmod d, \qquad k(n-1) \equiv 0 \pmod{\lambda(d)}. \tag{2.9}$$

The latter congruence is equivalent to

$$n-1 \equiv 0 \left(\mod \frac{\lambda(d)}{(k,\lambda(d))} \right).$$

If there is any such n that satisfies (2.9), it is necessary that

$$\left(d, \frac{\lambda(d)}{(k,\lambda(d))} \right) = 1.$$

Thus, by the Chinese remainder theorem, the number of n counted by $C_k(x)$ with $d|n$ is at most

$$1 + \left[\frac{x(k,\lambda(d))}{d\lambda(d)} \right]. \tag{2.10}$$

Further, if $d = p$ is prime, then the solution $n = p$ of (2.9) should not be counted since it is not composite. Thus for p prime, the number of n counted by $C_k(x)$ with $p|n$ is at most

$$\left[\frac{x(k,p-1)}{p(p-1)} \right]. \tag{2.11}$$

We now estimate $C_{k,2}(x)$ by using (2.11). For any $k \le L_1 < L$,

$$\begin{aligned} C_{k,2}(x) &\le x \sum_{p>L} \frac{(k,p-1)}{p(p-1)} = x \sum_{d|k} \sum_{\substack{p>L \\ (k,p-1)=d}} \frac{d}{p(p-1)} \\ &\le x \sum_{d|k} \sum_{\substack{p>L \\ p-1=md \\ \text{for some } m}} \frac{d}{(md+1)md} < x \sum_{d|k} \sum_{m>(L-1)/d} \frac{1}{m^2 d} \\ &\le x \sum_{d|k} \frac{1}{d} \left(\frac{d^2}{(L-1)^2} + \int_{(L-1)/d}^{\infty} \frac{1}{t^2}\, dt \right) \\ &= x \sum_{d|k} \left(\frac{d}{(L-1)^2} + \frac{1}{(L-1)} \right) \le \frac{x}{L-1} \left(\frac{L_1}{L-1} + 1 \right) \tau(k), \end{aligned} \tag{2.12}$$

where $\tau(k)$ is the number of divisors of k.

Suppose n is in class (iii). Let d_0 be the least divisor of n with $d_0 > x/L^2$. If p is any prime factor of d_0 and $d_0 > x/L$, then $d_0/p \ge d_0/L > x/L^2$, which gives a contradiction. Hence, n must have a divisor d with

$$\frac{x}{L^2} < d \le \frac{x}{L}. \tag{2.13}$$

Thus by (2.10),

$$C_{k,3}(x) \le \sum{}' \left(1 + \left[\frac{x(k,\lambda(d))}{d\lambda(d)} \right] \right), \tag{2.14}$$

where $\sum'$ denotes a sum over odd d satisfying (2.13).

We thus have

$$
\begin{aligned}
C_{k,3}(x) &\le \frac{x}{2L} + x \sum_{d\lambda(d)\le x(k,\lambda(d))}{}' \frac{(k,\lambda(d))}{d\lambda(d)} \\
&= \frac{x}{2L} + x \sum_{m\le L^2} \frac{1}{m} \sum_{\lambda(d)/(k,\lambda(d))=m}{}' \frac{1}{d} \\
&= \frac{x}{2L} + x \sum_{u\mid k} \sum_{m\le L^2} \frac{1}{m} \sum_{\substack{\lambda(d)=mu \\ (k,\lambda(d))=u}}{}' \frac{1}{d} \\
&\le \frac{x}{2L} + x \sum_{u\mid k} \sum_{m\le L^2} \frac{1}{m} \sum_{\lambda(d)=mu}{}' \frac{1}{d}.
\end{aligned} \tag{2.15}
$$

Using partial summation for the inner sum in (2.15), we have

$$
\sum_{\lambda(d)=mu}{}' \frac{1}{d} = \frac{1}{x/L} \sum_{\lambda(d)=mu}{}' 1 + \int_{x/L^2}^{x/L} \frac{1}{t^2} \sum_{\substack{d\le t \\ \lambda(d)=mu}}{}' 1\, dt. \tag{2.16}
$$

We thus shall be interested in obtaining an upper bound for $\Lambda(t, mu)$, the number of odd $d \le t$ with $\lambda(d) = mu$.

LEMMA 2.5. *Let $\lambda(m)$ be the Carmichael universal exponent function. Then,*

$$
\begin{aligned}
\Lambda(x,n) &:= \#\{m \le x\colon m \text{ odd},\ \lambda(m) = n\} \\
&\le K_c x^c \exp(2^{-c} f_c(n))
\end{aligned}
$$

for any $x > 1$, $\frac{1}{2} < c < 1$, where K_c and $f_c(n)$ are defined in Theorem 2.2.

The proof of Lemma 2.5 will be given later. Using it now in (2.16), we have

$$
\begin{aligned}
\sum_{\lambda(d)=mu}{}' \frac{1}{d} &\le \frac{L}{x}\left(\frac{x}{L}\right)^c K_c \exp(2^{-c} f_c(mu)) \\
&\quad + \int_{x/L^2}^{x/L} \frac{1}{t^2} t^c K_c \exp(2^{-c} f_c(mu))\, dt \\
&= \left(\frac{L}{x}\right)^{1-c} K_c \exp(2^{-c} f_c(mu)) \\
&\quad + K_c \exp(2^{-c} f_c(mu)) \cdot \frac{1}{c-1} \cdot \left[\left(\frac{x}{L}\right)^{c-1} - \left(\frac{x}{L^2}\right)^{c-1}\right] \\
&\le \frac{K_c}{1-c}\left(\frac{L^2}{x}\right)^{1-c} \exp(2^{-c} f_c(mu)).
\end{aligned} \tag{2.17}
$$

Putting this estimate in (2.15), we have

$$
\begin{aligned}
C_{k,3}(x) &\le \frac{x}{2L} + x \sum_{u\mid k} \sum_{m\le L^2} \frac{1}{m} \cdot \frac{K_c}{1-c}\left(\frac{L^2}{x}\right)^{1-c} \exp(2^{-c} f_c(mu)) \\
&\le \frac{x}{2L} + \frac{K_c}{1-c} x^c L^{2(1-c)} \sum_{u\mid k} \sum_{m\le L^2} \frac{1}{m} \exp(2^{-c} f_c(mu)).
\end{aligned} \tag{2.18}
$$

Using estimates (2.8), (2.12) and (2.18) in (2.7), we have

$$C_k(x) \le \frac{x}{L} + \frac{x}{L-1}\left(\frac{L_1}{L-1}+1\right)\tau(k) + \frac{K_c}{1-c}x^c L^{2(1-c)} \sum_{u|k} \sum_{m \le L^2} \frac{1}{m} \exp(2^{-c} f_c(mu)).$$

Using this estimate in (2.6), we get

$$\sum_{\substack{n \le x,\ n \text{ odd,} \\ \text{composite}}} F(n)$$

$$\le \frac{x^2}{4L_1} + \frac{x^2}{L}\sum_{k \le L_1}\frac{1}{k} + \frac{x^2}{L-1}\left(\frac{L_1}{L-1}+1\right)\sum_{k \le L_1}\frac{\tau(k)}{k} + \frac{K_c}{1-c}x^{1+c}L^{2(1-c)}\sum_{k \le L_1}\frac{1}{k}\sum_{u|k}\sum_{m \le L^2}\frac{1}{m}\exp(2^{-c}f_c(mu)). \tag{2.19}$$

The single sums on the right of (2.19) are dealt with in the following lemma.

LEMMA 2.6. *For any $x \ge 1$, we have*

$$\sum_{k \le x}\frac{1}{k} \le 1 + \log x, \qquad \sum_{k \le x}\frac{\tau(k)}{k} < \frac{1}{2}(2+\log x)^2,$$

where $\tau(k)$ is the number of divisors of k.

We defer the proof of Lemma 2.6 until later.

We deal with the final triple sum on the right of (2.19) as follows. We have, using Lemma 2.6,

$$\sum_{k \le L_1}\frac{1}{k}\sum_{u|k}\sum_{m \le L^2}\frac{1}{m}\exp(2^{-c}f_c(mu))$$

$$= \sum_{u \le L_1}\sum_{m \le L^2}\sum_{v \le L_1/u}\frac{1}{muv}\exp(2^{-c}f_c(mu))$$

$$\le (1+\log L_1)\sum_{u \le L_1}\sum_{m \le L^2}\frac{1}{mu}\exp(2^{-c}f_c(mu))$$

$$\le (1+\log L_1)\sum_{\mu \le L^2 L_1}\frac{\tau_{L_1}(\mu)}{\mu}\exp(2^{-c}f_c(\mu)),$$

where $\tau_{L_1}(\mu)$ is the number of divisors of μ up to L_1. Using this estimate and Lemma 2.6 on the right of (2.19) immediately gives the theorem.

We now prove Lemmas 2.5 and 2.6.

Proof of Lemma 2.5. If $c > 0$,

$$\Lambda(x,n) = \sum_{\substack{m \le x \\ m \text{ odd} \\ \lambda(m)=n}} 1 \le x^c \sum_{\substack{\lambda(m)=n \\ m \text{ odd}}} m^{-c} \le x^c \sum_{\substack{p|m \\ \Rightarrow (p-1)|n, \\ p \text{ odd}}} m^{-c}$$

$$= x^c \prod_{\substack{(p-1)|n \\ p \text{ odd}}} (1-p^{-c})^{-1}.$$

Hence,

$$\Lambda(x,n) \le x^c \exp\left(-\sum_{\substack{(p-1)|n \\ p \text{ odd}}} \log(1-p^{-c})\right)$$

$$= x^c \exp\left(\sum_{\substack{(p-1)|n \\ p \text{ odd}}} (p^{-c} + \tfrac{1}{2}p^{-2c} + \tfrac{1}{3}p^{-3c} + \cdots)\right).$$

We have

$$\sum_{\substack{(p-1)|n \\ p \text{ odd}}} p^{-c} \le \sum_{\substack{d|n \\ d \text{ even}}} (d+1)^{-c} < \sum_{\substack{d|n \\ d \text{ even}}} d^{-c} \le 2^{-c} \sum_{d|n} d^{-c}$$

$$\le 2^{-c} \prod_{p|n} (1-p^{-c})^{-1} = 2^{-c} f_c(n).$$

We recall that $K_c = \exp(\sum_{p>2} \sum_{k=2}^{\infty} \frac{1}{k} p^{-kc})$, which is finite for $c > \frac{1}{2}$. Thus we have Lemma 2.5.

Proof of Lemma 2.6. From Euler's summation formula,

$$\sum_{k \le x} \frac{1}{k} \le 1 + \int_1^x \frac{1}{t}\, dt = 1 + \log x.$$

Using partial summation,

$$\sum_{k \le x} \frac{\tau(k)}{k} = \frac{1}{x} \sum_{k \le x} \tau(k) + \int_1^x \frac{1}{t^2} \sum_{k \le t} \tau(k)\, dt.$$

We have by the first part of the lemma

$$\sum_{k \le t} \tau(k) = \sum_{k \le t} \sum_{d|k} 1 = \sum_{d \le t} \left[\frac{t}{d}\right] \le \sum_{d \le t} \frac{t}{d} \le t(1 + \log t).$$

Thus,

$$\sum_{k \le x} \frac{\tau(k)}{k} \le \frac{1}{x} \cdot x(1 + \log x) + \int_1^x \frac{1 + \log t}{t}\, dt$$

$$= 1 + \log x + \frac{1}{2} \log^2 x + \log x$$

$$< \frac{1}{2}(\log x + 2)^2.$$

3. The Range $x \ge 10^{300}$. In this section we shall use Theorem 2.3 to prove (1.2) and establish the estimates in Table 1 for $x \ge 10^{300}$. For the record, we make the following formal statement.

THEOREM 3.1. *If $x \ge 10^{10^5}$, then $P(x) \le 1/(\log x)^{197}$.*

We shall prove Theorem 3.1 by choosing

$$L = (\log x)^{200}, \quad L_1 = \frac{L}{(\log L)^2}, \quad c = 0.75 \tag{3.1}$$

in Theorem 2.3. However, there is some substantial work to do since the last term in Theorem 2.2 is not in closed form. The last term in Theorem 2.2 is

$$\frac{K_c}{1-c}x^{1+c}L^{2(1-c)}(1+\log L_1)\sum_{m\le L^2L_1}\frac{\tau_{L_1}(m)}{m}\exp(2^{-c}f_c(m)), \tag{3.2}$$

where

$$K_c=\exp\left(\sum_{p\text{ odd}}\sum_{k=2}^{\infty}k^{-1}p^{-kc}\right),$$
$$f_c(m)=\prod_{p\mid m}(1-p^{-c})^{-1},$$

and $\tau_{L_1}(m)$ is the number of divisors of m up to L_1.

To get an upper bound for (3.2), we first get an upper bound for K_c. Let p_i denote the ith prime. We have

$$\begin{aligned}\sum_{p>2}\sum_{k=2}^{\infty}\frac{1}{k}p^{-kc}&\le\frac{1}{2}\sum_{p>2}p^{-2c}+\frac{1}{3}\sum_{p>2}p^{-3c}(1+p^{-c}+p^{-2c}+\cdots)\\&=\frac{1}{2}\sum_{p>2}p^{-2c}+\frac{1}{3}\sum_{p>2}\frac{p^{-3c}}{1-p^{-c}}=\sum_{p>2}p^{-2c}\left(\frac{1}{2}+\frac{1}{3(p^c-1)}\right)\\&\le\sum_{i=2}^{11}p_i^{-2c}\left(\frac{1}{2}+\frac{1}{3(p_i^c-1)}\right)+\left(\frac{1}{2}+\frac{1}{3(37^c-1)}\right)\sum_{p\ge37}p^{-2c}.\end{aligned}$$

Now,

$$\begin{aligned}\sum_{p\ge37}p^{-2c}&<\sum_{k=18}^{\infty}(2k+1)^{-2c}<\sum_{k=18}^{\infty}\int_{2k}^{2k+1}t^{-2c}\,dt\\&<\int_{36}^{37}t^{-2c}\,dt+\frac{1}{2}\sum_{k=19}^{\infty}\left(\int_{2k-1}^{2k}+\int_{2k}^{2k+1}\right)t^{-2c}\,dt\\&=\int_{36}^{37}t^{-2c}\,dt+\frac{1}{2}\int_{37}^{\infty}t^{-2c}\,dt\\&=\frac{1}{2c-1}\left(36^{1-2c}-\frac{1}{2}\cdot37^{1-2c}\right).\end{aligned}$$

Thus, if

$$\begin{aligned}K_c'=&\sum_{i=2}^{11}p_i^{-2c}\left(\frac{1}{2}+\frac{1}{3(p_i^c-1)}\right)\\&+\frac{1}{(2c-1)}\left(\frac{1}{2}+\frac{1}{3(37^c-1)}\right)\left(36^{1-2c}-\frac{1}{2}\cdot37^{1-2c}\right),\end{aligned}$$

then

$$K_c\le\exp(K_c'). \tag{3.3}$$

We now obtain an upper bound for $\sum_{m\le L^2L_1}\tau_{L_1}(m)/m$.

LEMMA 3.2. *If* $1 < L_1 < L$, *then*

$$\sum_{m \le L^2 L_1} \frac{\tau_{L_1}(m)}{m} \le (1 + \log L_1)(1 + \log L^2 L_1).$$

Proof. We have

$$\sum_{m \le L^2 L_1} \frac{\tau_{L_1}(m)}{m} = \frac{1}{L^2 L_1} \sum_{m \le L^2 L_1} \tau_{L_1}(m) + \int_1^{L^2 L_1} \frac{1}{t^2} \sum_{m \le t} \tau_{L_1}(m)\, dt.$$

Note that by Lemma 2.6

$$\sum_{m \le t} \tau_{L_1}(m) = \sum_{m \le t} \sum_{\substack{d \mid m \\ d \le L_1}} 1 = \sum_{d \le L_1} \left[\frac{t}{d}\right]$$

$$\le t \sum_{d \le L_1} \frac{1}{d} \le t(1 + \log L_1).$$

Thus,

$$\sum_{m \le L^2 L_1} \frac{\tau_{L_1}(m)}{m} \le (1 + \log L_1) + \int_1^{L^2 L_1} \frac{1}{t}(1 + \log L_1)\, dt$$

$$= (1 + \log L_1)(1 + \log L^2 L_1).$$

This completes the proof of Lemma 3.2.

Using Lemma 3.2, we can get an upper bound for

$$\sum_{m \le L^2 L_1} \frac{\tau_{L_1}(m)}{m} \exp(2^{-c} f_c(m)).$$

We define

$$m_i = p_1 p_2 \cdots p_i, \tag{3.4}$$

the product of the first i primes, and let j be an integer such that

$$m_j \le L^2 L_1 < m_{j+1}. \tag{3.5}$$

PROPOSITION 3.3. *If* $1 < L_1 < L$, *then*

$$\sum_{m \le L^2 L_1} \frac{\tau_{L_1}(m)}{m} \exp(2^{-c} f_c(m))$$

$$\le (1 + \log L_1)(1 + \log L^2 L_1) \exp(2^{-c} f_c(m_j)).$$

Proof. Suppose that m has k distinct prime factors $q_1, q_2, \ldots, q_k$ and that $p_1, p_2, \ldots, p_k$ are the first k primes. Then

$$f_c(m) = \prod_{1 \le i \le k} (1 - q_i^{-c})^{-1} \le \prod_{1 \le i \le k} (1 - p_i^{-c})^{-1} = f_c(m_k).$$

Because j is chosen as $p_1 p_2 \cdots p_j \le L^2 L_1 < p_1 p_2 \cdots p_{j+1}$, clearly j is the largest possible value for the number of primes in $m \le L^2 L_1$. Thus $f_c(m_j)$ is a universal upper bound for $f_c(m)$ for any $m \le L^2 L_1$. Thus,

$$\sum_{m \le L^2 L_1} \frac{\tau_{L_1}(m)}{m} \exp(2^{-c} f_c(m)) \le \exp(2^{-c} f_c(m_j)) \sum_{m \le L^2 L_1} \frac{\tau_{L_1}(m)}{m}$$

$$\le \exp(2^{-c} f_c(m_j))(1 + \log L_1)(1 + \log L^2 L_1)$$

by Lemma 3.2.

Now we prove Theorem 3.1. Assume $x \geq 10^{10^5}$ and L, L_1, c are as given in (3.1). From Theorem 2.3 and Proposition 3.3, we have

$$\begin{aligned} P(x) &\leq \frac{\sum_{n\leq x,\ n \text{ odd, composite}} F(n)}{\sum_{2<p\leq x}(p-3)} \\ &\leq 2(2+\log x)\left(\frac{(\log L)^2}{4L} + \frac{\log L}{L} + \frac{(\log L)^2}{L}\right. \\ &\qquad \left. + \frac{3K_c}{1-c}\cdot\frac{L^{2(1-c)}}{x^{1-c}}(\log L)^3 \exp(2^{-c} f_c(m_j))\right) \qquad (3.6) \\ &= 2(2+\log x)\left(\frac{5(\log L)^2 + 4\log L}{4L}\right. \\ &\qquad \left. + \frac{3K_c}{1-c}\cdot\frac{L^{2(1-c)}}{x^{1-c}}(\log L)^3 \exp(2^{-c} f_c(m_j))\right). \end{aligned}$$

We shall be interested in getting an upper bound for $2^{-c} f_c(m_j)$. We have

$$2^{-c} f_c(m_j) = 2^{-c} \prod_{p\mid m_j} (1-p^{-c})^{-1} = \frac{1}{2^c - 1} \prod_{i=2}^{j} (1-p_i^{-c})^{-1}. \qquad (3.7)$$

Now,

$$\log \prod_{i=2}^{j} (1-p_i^{-c})^{-1} = -\sum_{i=2}^{j} \log(1-p_i^{-c}) = \sum_{i=2}^{j} p_i^{-c} + \sum_{i=2}^{j}\sum_{k=2}^{\infty} \frac{1}{k} p_i^{-kc}.$$

Hence,

$$\prod_{i=2}^{j} (1-p_i^{-c})^{-1} \leq K_c \exp\left(\sum_{i=2}^{j} p_i^{-c}\right). \qquad (3.8)$$

Putting (3.8) in (3.7), we have

$$2^{-c} f_c(m_j) \leq \frac{K_c}{2^c - 1} \exp\left(\sum_{i=2}^{j} p_i^{-c}\right). \qquad (3.9)$$

Now for $j > 16$,

$$\sum_{i=2}^{j} p_i^{-c} = \sum_{i=2}^{16} p_i^{-c} + \sum_{i=17}^{j} p_i^{-c}. \qquad (3.10)$$

Using partial summation for $\sum_{i=17}^{j} p_i^{-c}$, we have

$$\sum_{i=17}^{j} p_i^{-c} = \sum_{59\leq p\leq p_j} p^{-c} = j p_j^{-c} - 16\cdot 59^{-c} + c\int_{59}^{p_j} t^{-c-1}\pi(t)\,dt. \qquad (3.11)$$

We use the following upper bound for $\pi(t)$:

$$\pi(t) \leq \frac{t}{\log t}\left(1 + \frac{3}{2\log t}\right) \quad \text{for } t > 1, \qquad (3.12)$$

which is (3.2) in [5, p. 69]. Thus, for $S = \int_{59}^{p_j} t^{-c-1}\pi(t)\,dt$, we have

$$\begin{aligned} S &\le \int_{59}^{p_j} \frac{t^{-c}}{\log t}\left(1+\frac{3}{2\log t}\right)dt \\ &= \frac{1}{1-c}\cdot\frac{t^{1-c}}{\log t}\Big]_{59}^{p_j} + \left(\frac{3}{2}+\frac{1}{1-c}\right)\int_{59}^{p_j}\frac{t^{-c}}{\log^2 t}\,dt. \end{aligned} \tag{3.13}$$

Let $I = \int_{59}^{p_j}(t^{-c}/\log^2 t)\,dt$. Assuming $p_j > 3481 = 59^2$, we have

$$\begin{aligned} I &= \int_{59}^{\sqrt{p_j}} \frac{t^{-c}}{\log^2 t}\,dt + \int_{\sqrt{p_j}}^{p_j}\frac{t^{-c}}{\log^2 t}\,dt \\ &\le \frac{1}{\log^2 59}\cdot\frac{1}{1-c}\cdot t^{1-c}\Big]_{59}^{\sqrt{p_j}} \\ &\quad + \frac{1}{\log^2\sqrt{p_j}}\cdot\frac{1}{1-c}\cdot t^{1-c}\Big]_{\sqrt{p_j}}^{p_j}. \end{aligned}$$

Thus,

$$\begin{aligned} S &\le \frac{1}{1-c}\frac{p_j^{1-c}}{\log p_j} - \frac{1}{1-c}\frac{59^{1-c}}{\log 59} \\ &\quad + \left(\frac{3}{2}+\frac{1}{1-c}\right)\left(\frac{1}{(1-c)\log^2 59}p_j^{(1-c)/2} - \frac{1}{(1-c)\log^2 59}59^{1-c}\right. \\ &\quad \left. + \frac{4}{1-c}\cdot\frac{p_j^{1-c}}{\log^2 p_j} - \frac{4}{1-c}\frac{p_j^{(1-c)/2}}{\log^2 p_j}\right). \end{aligned} \tag{3.14}$$

Now by (3.12),

$$j = \pi(p_j) < \frac{p_j}{\log p_j}\left(1+\frac{3}{2\log p_j}\right). \tag{3.15}$$

Assembling (3.10), (3.11), (3.14), (3.15), and simplifying with $c = 0.75$, we have

$$\sum_{i=2}^{j} p_i^{-0.75} \le 4\cdot\frac{p_j^{1/4}}{\log p_j} + 67.5\frac{p_j^{1/4}}{\log^2 p_j} + p_j^{1/8} - 66\cdot\frac{p_j^{1/8}}{\log^2 p_j} - 3.7. \tag{3.16}$$

From (3.5) and Theorems 9 and 10 in [5, p. 71], we have

$$p_j \le 1.04\log L^2 L_1 \le 3.12\log L \quad \text{for } \log L^2 L_1 > 2703.$$

Putting this in (3.16) and taking $L = (\log x)^{200}$, we get

$$\sum_{i=2}^{j} p_i^{-0.75} \le 4.8(\log\log x)^{1/4} \quad \text{for } x \ge 10^{10^5}.$$

From (3.3) with $c = 0.75$, we have $K_c < e^{0.4}$. Thus, from the above and (3.9), we have

$$\begin{aligned} \exp(2^{-c} f_c(m_j)) &\le \exp\left(\frac{K_c}{2^c-1}\exp\left(\sum_{i=2}^{j} p_i^{-0.75}\right)\right) \\ &< \exp(2.2\exp(4.8(\log\log x)^{1/4})) < x^{0.1} \end{aligned} \tag{3.17}$$

for $x \ge 10^{10^5}$.

We put (3.17) in (3.6), getting

$$P(x) \le 2(2+\log x) \cdot \left(\frac{5(\log L)^2 + 4\log L}{4L} + 18.0\frac{L^{1/2}}{x^{0.15}}(\log L)^3\right) \tag{3.18}$$

for $x \ge 10^{10^5}$, where $L = (\log x)^{200}$. Since $x^{0.15} > (\log x)^{1200}$ for all $x \ge 10^{10^5}$, it is easy to check that (3.18) implies $P(x) \le 1/(\log x)^{197}$ for all $x \ge 10^{10^5}$, which was to be proved.

To establish the estimates in Table 1 for $x \ge 10^{300}$, we use the same values of the parameters L_1, c in Theorem 2.3 as given in (3.1) and then we choose L optimally. We also use Proposition 3.3 for the sum in the last term in Theorem 2.2, to obtain (3.6). The principal difference between the range $x \ge 10^{10^5}$ and $10^{300} \le x \le 10^{10^5}$ is that instead of using (3.7)–(3.17) to estimate $\exp(2^{-c} f_c(m_j))$, we directly compute this quantity, which is not too hard to do when given a finite value of x that is not too large.

4. A Refinement of the Basic Method. To get good results for smaller values of x, we shall use a more elaborate version of Theorem 2.2 and Proposition 3.3.

THEOREM 4.1. *Suppose c, L_1, L, L_2 and M are arbitrary real numbers satisfying $\frac{1}{2} < c < 1$, $10 < L_1 < L < L_2 < M/2$, $L^{3/2} \le 10M$. Then for any $x > L^2$, we have*

$$\begin{aligned}\sum_{\substack{n\le x,\ n \text{ odd},\\ \text{composite}}} F(n) \le{}& \frac{x^2}{4L_1} + \frac{50}{99}\frac{x^2}{L_2-1}\left(\frac{L_1}{L_2-1}+1\right)(2+\log L_1)^2 \\ &+ xL_2^2\left(2+\frac{\log x}{\log 10}\right)(1+\log L_1) + \frac{100}{99}\frac{x^2}{M}(1+\log L_1)^2 \\ &+ \frac{125}{3564}\frac{x^2(1+\log L_2)^2}{M-2L_2}(4+\log L_1)^4 + \frac{50}{99}\frac{x^2}{L}(1+\log L_1) \\ &+ \frac{100K_c}{(1-c)(10^{1+c}-1)}x^{1+c}M^{1-c}(1+\log L_1) \\ &\cdot \sum_{m\le L^2L_1}\frac{\tau_{L_1}(m)}{m}\exp(2^{-c}f_c(m)),\end{aligned}$$

where K_c, $f_c(m)$ and $\tau_{L_1}(m)$ are defined in Theorem 2.2.

Proof. Although the assertion appears to be considerably more complicated, the proof of Theorem 4.1 follows fairly directly from the same methods used to prove Theorem 2.2. By the same argument that establishes (2.6), we have

$$\sum_{\substack{x/10<n\le x\\ n \text{ odd, composite}}} F(n) \le x\sum_{k\le L_1}\frac{C_k(x)-C_k(x/10)}{k} + \frac{1}{L_1}\sum_{\substack{x/10<n\le x\\ n \text{ odd}}}(n-2). \tag{4.1}$$

To obtain the theorem, we add (4.1) at $x, x/10, \dots, x/10^u$, where $x/10^u > L^2 > x/10^{u+1}$. Thus it shall be sufficient to prove that

$$\begin{aligned}\sum_{k\le L_1} &\frac{C_k(x)-C_k(x/10)}{k}\\ &\le \frac{1}{2}\cdot\frac{x}{L_2-1}\left(\frac{L_1}{L_2-1}\right)(2+\log L_1)^2+\frac{x}{M}(1+\log L_1)^2\\ &\quad+\frac{5}{144}\frac{x(1+\log L_2)^2}{M-2L_2}(4+\log L_1)^4+\frac{1}{2}\frac{x}{L}(1+\log L_1)\\ &\quad+L_2^2(1+\log L_1)\\ &\quad+\frac{10^{1-c}K_c}{1-c}x^cM^{1-c}(1+\log L_1)\sum_{m\le L^2L_1}\frac{\tau_{L_1}(m)}{m}\exp(2^{-c}f_c(m)).\end{aligned}\tag{4.2}$$

The expression $C_k(x)-C_k(x/10)$ is the cardinality of the set of odd, composite integers n with $x/10 < n \le x$ and $F(n)=\phi(n)/k$. We let $B_{k,1}(x), B_{k,2}(x), B_{k,3}(x)$ denote the number of such n that satisfy, respectively,

(1) n is divisible by some prime $p > L_2$,
(2) n has a divisor $pq > M$, where $q \le p$ are prime, but n is not counted by (1),
(3) n is not counted by (1) or (2).

Thus, $C_k(x)-C_k(x/10)=B_{k,1}(x)+B_{k,2}(x)+B_{k,3}(x)$.

From the argument which gives (2.12) we have immediately that

$$B_{k,1}(x)\le\frac{x}{L_2-1}\left(\frac{L_1}{L_2-1}\right)\tau(k),$$

so that from Lemma 2.6, we have

$$\sum_{k\le L_1}\frac{B_{k,1}(x)}{k}\le\frac{1}{2}\cdot\frac{x}{L_2-1}\left(\frac{L_1}{L_2-1}\right)(2+\log L_1)^2.\tag{4.3}$$

To analyze $B_{k,2}(x)$, we consider separately the case $q<p$ and $q=p$. The contribution to $B_{k,2}(x)$ from the case $q<p$ is, by (2.10), at most

$$\begin{aligned}&\sum_{\substack{q<p\le L_2\\ pq>M}}1+\frac{x(k,[p-1,q-1])}{pq[p-1,q-1]}\\ &\qquad\le\frac{1}{2}L_2^2+\frac{1}{2}x\sum_{\substack{q,p\le L_2\\ pq>M\\ p\ne q}}\frac{(k,[p-1,q-1])}{pq[p-1,q-1]}.\end{aligned}\tag{4.4}$$

For this last sum, we write $p-1=md$, $q-1=nd$, where $(m,n)=1$. Since $pq>M$ and $p,q\le L_2$ imply $(p-1)(q-1)>M-2L_2:=M'$, we have

$$\sum_{\substack{p,q\le L_2\\ pq>M\\ p\ne q}}\frac{(k,[p-1,q-1])}{pq[p-1,q-1]}\le\sum_{d\le L_2}\sum_{\substack{m,n\le L_2/d\\ mn>M'/d^2\\ (m,n)=1}}\frac{(k,mnd)}{m^2n^2d^3}.\tag{4.5}$$

If we let $u_1 = (k, m)$, $u_2 = (k, n)$, then the condition $(m, n) = 1$ implies $u_1 u_2 | k$. Let $u_3 = (k/u_1 u_2, d)$. Thus, $(k, mnd) = u_1 u_2 u_3$. Let u_4 be such that $k = u_1 u_2 u_3 u_4$, and let μ, ν, δ be such that $m = u_1 \mu$, $n = u_2 \nu$, $d = u_3 \delta$. Thus, from (4.5),

$$\begin{aligned}
&\sum_{\substack{p,q \le L_2 \\ pq > M \\ p \ne q}} \frac{(k, [p-1, q-1])}{pq[p-1, q-1]} \\
&\quad \le \sum_{\substack{u_1 u_2 u_3 u_4 = k \\ (u_1, u_2) = 1}} \sum_{\substack{\delta \le L_2/u_3 \\ \nu \le L_2/u_2}} \sum_{\mu > M'/(\nu \delta^2 u_1 u_2 u_3^2)} \frac{1}{\mu^2 \nu^2 \delta^3 u_1 u_2 u_3^2} \qquad (4.6) \\
&\quad < \frac{5}{3M'} \sum_{u_1 u_2 u_3 u_4 = k} \sum_{\substack{\delta \le L_2/u_3 \\ \nu \le L_2/u_2}} \frac{1}{\nu \delta} \\
&\quad \le \frac{5}{3M'} (1 + \log L_2)^2 \sum_{u_1 u_2 u_3 u_4 = k} 1,
\end{aligned}$$

where we used the inequality

$$\sum_{\mu > y} \frac{1}{\mu^2} < \frac{5}{3y} \quad \text{for } y > 0. \qquad (4.7)$$

The proof of (4.7) follows from the facts that

$$\begin{aligned}
&\sum_{\mu > y} \frac{1}{\mu^2} = \sum_{\mu=1}^{\infty} \frac{1}{\mu^2} = \frac{\pi^2}{6} < \frac{5}{3} < \frac{5}{3y} \quad \text{for } 0 < y < 1, \\
&\sum_{\mu > y} \frac{1}{\mu^2} \le \sum_{\mu=2}^{\infty} \frac{1}{\mu^2} < \frac{2}{3} \le \frac{5}{3y} \quad \text{for } 1 \le y \le 2.5, \\
&\sum_{\mu > y} \frac{1}{\mu^2} < \int_{y-1}^{\infty} \frac{1}{t^2}\, dt = \frac{1}{y-1} < \frac{5}{3y} \quad \text{for } y > 2.5.
\end{aligned}$$

Putting (4.6) into (4.4), we have

$$\begin{aligned}
&\sum_{k \le L_1} \frac{1}{k} \sum_{\substack{q < p \le L_2 \\ pq > M}} 1 + \frac{x(k, [p-1, q-1])}{pq[p-1, q-1]} \qquad (4.8) \\
&\quad \le \frac{1}{2} L_2^2 (1 + \log L_1) + \frac{5x}{6M'} (1 + \log L_2)^2 \sum_{k \le L_1} \frac{\tau_{(4)}(k)}{k},
\end{aligned}$$

where $\tau_{(i)}(k)$ is the number of ordered factorizations of k into i positive factors.

It is not hard to prove by induction on i that

$$\sum_{k \le y} \frac{\tau_{(i)}(k)}{k} \le \frac{1}{i!} (i + \log y)^i \qquad (4.9)$$

for any natural number i and for any $y \ge 1$; in fact, Lemma 2.6 gives the cases $i = 1, 2$. We shall use (4.9) with $i = 4$ in a moment.

We now consider the contribution to $B_{k,2}(x)$ when $q = p$. If $p^2 | n$ and $F(n) = \phi(n)/k$, we have from (2.9) that

$$n \equiv 0 \pmod{p^2}, \qquad k(n-1) \equiv 0 \pmod{p(p-1)}.$$

Thus, $p|k$, and the number of such $n \le x$ is at most

$$1 + \left[\frac{x(k,p-1)}{p^2(p-1)}\right] \le 1 + \frac{x}{p^2} < 1 + \frac{x}{M}$$

if $p^2 > M$. Since $k \le L_1$, the number of primes $p|k$ with $p > M^{1/2}$ is at most $(\log L_1)/\log(M^{1/2}) < \log L_1$. Thus, the contribution to $B_{k,2}(x)$ when $q = p$ is at most $(1 + x/M)\log L_1$.

Using this result together with (4.8) and (4.9) gives

$$\sum_{k \le L_1} \frac{B_{k,2}(x)}{k} \le L_2^2(1 + \log L_1) + \frac{x}{M}(1 + \log L_1)^2 + \frac{5x}{144M'}(1 + \log L_2)^2(4 + \log L_1)^4. \tag{4.10}$$

We now turn our attention to $B_{k,3}(x)$, the number of odd, composite n with $x/10 < n \le x$, $p \le L_2$ for every prime $p|n$, and $pq \le M$ for all primes p, q with $pq|n$. Factor such a number n as $q_1 q_2 \cdots q_t$, where $q_1 \ge q_2 \ge \cdots \ge q_t$ are primes. Then, $q_1 \le L_2$ and $q_1 q_2 \le M$. Note that

$$\frac{n}{q_1 q_2} \ge \frac{n}{M} > \frac{x}{10M}.$$

Suppose $n/q_1 q_2 > x/L$. Then, $q_1 q_2 \le L$, so that $q_2 \le L^{1/2}$. Thus, n has a divisor d with $x/L^{3/2} < d \le x/L$. But $x/L^{3/2} \ge x/(10M)$, so that in either case, n has a divisor d satisfying

$$\frac{x}{10M} < d \le \frac{x}{L}. \tag{4.11}$$

We next repeat the calculations (2.15)–(2.18), but with $\sum'$ now representing a sum over odd d satisfying (4.11). Thus,

$$B_{k,3}(x) \le \frac{x}{2L} + \frac{K_c}{1-c} x^c (10M)^{1-c} \sum_{u|k} \sum_{m \le L^2} \frac{1}{m} \exp(2^{-c} f_c(mu)),$$

and so, as in Section 2, we get

$$\sum_{k \le L_1} \frac{B_{k,3}(x)}{k} \le \frac{x}{2L}(1 + \log L_1) + \frac{K_c}{1-c} x^c (10M)^{1-c}(1 + \log L_1) \cdot \sum_{m \le L^2 L_1} \frac{\tau_{L_1}(m)}{m} \exp(2^{-c} f_c(m)). \tag{4.12}$$

Our proof is now complete, since adding (4.3), (4.10) and (4.12) gives (4.2), which, as we have seen, is sufficient for the proof of the theorem.

We shall also wish to use a sharper result than Proposition 3.3. Let

$$2 \le l < L_1, \qquad \alpha_l = (1 - l^{-c})^{-\log(L^2 L_1)/\log l}, \tag{4.13}$$

where α_l depends on the choice of c, L, L_1, l.

PROPOSITION 4.2. *If* $\frac{1}{2} < c < 1$, $1 < L_1 < L$ *and* l, α_l *are given by* (4.13), *we have*

$$\sum_{m \le L^2 L_1} \frac{\tau_{L_1}(m)}{m} \exp(2^{-c} f_c(m)) \le (1 + \log L_1) \sum_{i=0}^{j} \frac{1}{i!} \left(2 \sum_{p \le l} \frac{1}{p} \right)^i \exp(2^{-c} \alpha_l f_c(m_i)) \left(1 + \log \frac{L^2 L_1}{m_i} \right),$$

where m_i *and* j *are defined in* (3.4) *and* (3.5).

To prove Proposition 4.2, we state a result that will be useful. We first make a definition as follows. Let $S(l,k) = \sum \frac{1}{u}$, where u runs over the squarefree integers that are the product of k distinct primes up to l.

LEMMA 4.3. *For any nonnegative integer* k *and any* $l \ge 2$, *we have*

$$S(l,k) \le \frac{1}{k!} \left(\sum_{p \le l} \frac{1}{p} \right)^k.$$

Proof. This elementary result follows by expanding the right side of this inequality with the multinomial theorem.

We now prove Proposition 4.2.

Proof of Proposition 4.2. To estimate $f_c(m)$, we ask how many primes $p > l$ can divide m. The number of such primes is at most $\log(L^2 L_1)/\log l$. Thus,

$$f_c(m) = \prod_{p \mid n} (1 - p^{-c})^{-1} = \prod_{\substack{p \mid m \\ p > l}} (1 - p^{-c})^{-1} \cdot \prod_{\substack{p \mid m \\ p \le l}} (1 - p^{-c})^{-1}$$
$$\le (1 - l^{-c})^{-\log(L^2 L_1)/\log l} \cdot \prod_{\substack{p \mid m \\ p \le l}} (1 - p^{-c})^{-1} \le \alpha_l f_c(m_i)$$

if m has exactly i distinct primes up to l. In general, let $\omega_l(m)$ be the number of distinct prime factors of m at most l and $\omega(m)$ be the number of distinct primes in m.

Now,

$$\sum_{m \le L^2 L_1} \frac{\tau_{L_1}(m)}{m} \exp(2^{-c} f_c(m)) \le \sum_{i=0}^{j} \sum_{\substack{\omega_l(m) = i \\ m \le L^2 L_1}} \frac{\tau_{L_1}(m)}{m} \exp(2^{-c} f_c(m)).$$

For the inner sum, we have

$$\sum_{\substack{\omega_l(m) = i \\ m \le L^2 L_1}} \frac{\tau_{L_1}(m)}{m} \exp(2^{-c} f_c(m)) \le \exp(2^{-c} \alpha_l f_c(m_i)) \sum_{\substack{\omega_l(m) = i \\ m \le L^2 L_1}} \frac{\tau_{L_1}(m)}{m}$$
$$\le \exp(2^{-c} \alpha_l f_c(m_i)) \sum_{\substack{\omega(u) = \omega_l(u) = i \\ u \text{ squarefree}}} \sum_{t \le L^2 L_1 / m_i} \frac{\tau_{L_1}(ut)}{ut}$$

since any $m \le L^2 L_1$ with $\omega_l(m) = i$ may be factored as ut, where u is the product of i distinct primes up to l and t is an integer at most $L^2 L_1/u \le L^2 L_1/m_i$. Using

$$\tau_{L_1}(ut) \le \tau_{L_1}(u)\tau_{L_1}(t) \le \tau(u)\tau_{L_1}(t) = 2^i \tau_{L_1}(t),$$

we have

$$\begin{aligned}
&\sum_{\substack{\omega_l(m)=i \\ m \le L^2 L_1}} \frac{\tau_{L_1}(m)}{m} \exp(2^{-c} f_c(m)) \\
&\le \exp(2^{-c}\alpha_l f_c(m_i)) \sum_{\substack{\omega(u)=\omega_l(u)=i \\ u \text{ squarefree}}} \frac{2^i}{u} \sum_{t \le L^2 L_1/m_i} \frac{\tau_{L_1}(t)}{t} \qquad (4.14) \\
&\le \exp(2^{-c}\alpha_l f_c(m_i)) \frac{1}{i!} \left(2 \sum_{p \le l} \frac{1}{p}\right)^i (1 + \log L_1) \left(1 + \log \frac{L^2 L_1}{m_i}\right)
\end{aligned}$$

by Lemmas 3.2 and 4.3. Therefore, taking the summation of (4.14) from $i = 0$ to j completes the proof.

Putting together Theorems 2.1 and 4.1 and Proposition 4.2, we have the following result.

THEOREM 4.4. *If* $\frac{1}{2} < c < 1$, $10 < L_1 < L < L_2 < M/2$, $L^{3/2} \le 10M$, *and* $x > L^2$, *then*

$$\begin{aligned}
P(x) \le 2(2 + \log x) \Bigg\{ &\frac{1}{4L_1} + \frac{50}{99}\left(\frac{L_1}{L_2 - 1} + 1\right) \frac{(2 + \log L_1)^2}{L_2 - 1} \\
&+ \frac{1}{x} L_2^2 \left(2 + \frac{\log x}{\log 10}\right)(1 + \log L_1) + \frac{100}{99} \cdot \frac{(1 + \log L_1)^2}{M} \\
&+ \frac{125}{3564} \frac{(1 + \log L_2)^2}{M - 2L_2} (4 + \log L_1)^4 + \frac{50}{99} \cdot \frac{1 + \log L_1}{L} \\
&+ \frac{K_c}{(1 - c)(10^{1+c} - 1)} \left(\frac{M}{x}\right)^{1-c} (1 + \log L_1)^2 \\
&\cdot \sum_{i=0}^{j} \frac{1}{i!} \left(2 \sum_{p \le l} \frac{1}{p}\right)^i \left(1 + \log \frac{L^2 L_1}{m_i}\right) \exp(2^{-c}\alpha_l f_c(m_i)) \Bigg\},
\end{aligned}$$

where l, α_l *are given by* (4.13), m_i *is defined in* (3.4) *and* j *is defined in* (3.5).

While admittedly looking complicated, Theorem 4.4 can be readily used to get explicit upper bounds for $P(x)$ for various values of x. The art is to choose the many free parameters optimally. Of the seven terms in the brackets, it is clear that some dominate others. For example, the fourth term is small compared to the fifth term. We choose the parameters so that the first four terms are the least important and the seventh is the most important. We feel these choices are close to the optimal ones. Our results are recorded in Table 2 and summarized in Table 1 in the Introduction. An asterisk in the M column signifies that M was chosen as $L^{3/2}/10$.

TABLE 2

x	L	L_1	L_2	M	l	c	Upper Bound for $P(x)$
1.0E + 60	3.6E + 5	5.4E + 3	2.0E + 6	6.2E + 9	350	0.7125	7.16E − 2
1.0E + 70	1.1E + 7	1.7E + 5	1.2E + 8	5.9E + 11	600	0.7125	2.87E − 3
1.0E + 80	7.2E + 8	1.1E + 7	1.4E + 10	9.9E + 13	850	0.7125	8.46E − 5
1.0E + 90	4.5E + 10	7.0E + 8	1.4E + 12	1.4E + 16	1400	0.7100	1.70E − 6
1.0E + 100	3.3E + 12	5.5E + 10	1.7E + 14	2.5E + 18	1850	0.7100	2.77E − 8
1.0E + 110	2.6E + 14	5.3E + 12	2.3E + 16	4.3E + 20	1850	0.7100	4.03E − 10
1.0E + 120	1.2E + 16	4.4E + 14	2.5E + 18	*1.3E + 23	2350	0.7075	5.28E − 12
1.0E + 130	6.7E + 17	5.6E + 16	4.2E + 20	*5.5E + 25	2590	0.7075	7.54E − 14
1.0E + 140	6.8E + 19	5.7E + 18	5.3E + 22	*5.6E + 28	2800	0.7075	1.08E − 15
1.0E + 150	2.7E + 21	2.2E + 20	2.4E + 24	*1.4E + 31	3250	0.7075	1.49E − 17
1.0E + 160	4.2E + 23	3.8E + 22	5.1E + 26	*2.7E + 34	4900	0.7050	1.81E − 19
1.0E + 170	3.9E + 25	3.5E + 24	5.6E + 28	*2.4E + 37	6300	0.7025	2.27E − 21
1.0E + 180	3.7E + 27	3.3E + 26	6.2E + 30	*2.3E + 40	8000	0.7025	2.76E − 23
1.0E + 190	2.9E + 29	2.4E + 28	5.4E + 32	*1.6E + 43	9300	0.7025	3.26E − 25
1.0E + 200	2.7E + 31	3.3E + 30	8.2E + 34	*1.4E + 46	12000	0.7025	3.85E − 27

Department of Computer Science
University of South Carolina
Columbia, South Carolina 29208

Department of Mathematics
University of Georgia
Athens, Georgia 30602

1. P. BEAUCHEMIN, G. BRASSARD, C. CRÉPEAU, C. GOUTIER & C. POMERANCE, "The generation of random numbers that are probably prime," *J. Cryptology*, v. 1, 1988, pp. 53–64.

2. P. ERDÖS & C. POMERANCE, "On the number of false witnesses for a composite number," *Math. Comp.*, v. 46, 1986, pp. 259–279.

3. L. MONIER, "Evaluation and comparison of two efficient probabilistic primality testing algorithms," *Theoret. Comput. Sci.*, v. 12, 1980, pp. 97–108.

4. M. O. RABIN, "Probabilistic algorithm for testing primality," *J. Number Theory*, v. 12, 1980, pp. 128–138.

5. J. B. ROSSER & L. SCHOENFELD, "Approximate formulas for some functions of prime numbers," *Illinois J. Math.*, v. 6, 1962, pp. 64–94.