Canad. Math. Bull. Vol. **66** (1), 2023, pp. 185–195
http://dx.doi.org/10.4153/S0008439522000212

On the average value of $\pi(t) - \mathrm{li}(t)$

Daniel R. Johnston

Abstract. We prove that the Riemann hypothesis is equivalent to the condition $\int_2^x (\pi(t) - \mathrm{li}(t))\,\mathrm{d}t < 0$ for all $x > 2$. Here, $\pi(t)$ is the prime-counting function and $\mathrm{li}(t)$ is the logarithmic integral. This makes explicit a claim of Pintz. Moreover, we prove an analogous result for the Chebyshev function $\theta(t)$ and discuss the extent to which one can make related claims unconditionally.

1 Introduction

Let $\pi(x)$ denote the number of primes less than or equal to x and

$$\mathrm{li}(x) = \int_0^x \frac{1}{\log t}\mathrm{d}t.$$

In his celebrated 1859 article, Riemann [21] remarked on the apparent truth of the inequality

$$\pi(x) < \mathrm{li}(x),$$

for all $x \geq 2$. In 1903, Schmidt [25, p. 204] showed that such a result would imply the Riemann hypothesis. However, in 1914, Littlewood [16] managed to prove that $\pi(x) - \mathrm{li}(x)$ changes sign infinitely often. More precisely, he showed that for some positive constant c, there are arbitrarily large values of x such that

$$\pi(x) - \mathrm{li}(x) > \frac{c\sqrt{x}\log\log\log x}{\log x} \quad \text{and} \quad \pi(x) - \mathrm{li}(x) < -\frac{c\sqrt{x}\log\log\log x}{\log x}.$$

It is an open problem to determine the smallest value of x such that $\pi(x) > \mathrm{li}(x)$. Large computations have shown that $\pi(x) < \mathrm{li}(x)$ for all $x \leq 10^{19}$ [6, Theorem 2] and that the first sign change occurs before $x = 1.4 \times 10^{316}$ [1, 24].

Although $\pi(x) < \mathrm{li}(x)$ is not true in general, one can ask whether, in a precise sense, $\pi(x) < \mathrm{li}(x)$ is true on average. Namely, several authors [12–14, 19, 26] assert that the Riemann hypothesis implies

$$\int_2^x (\pi(t) - \mathrm{li}(t))\,\mathrm{d}t < 0, \quad x > x_0, \tag{1.1}$$

Received by the editors January 16, 2022; revised March 2, 2022; accepted March 5, 2022.
Published online on Cambridge Core March 14, 2022.
AMS subject classification: **11M26**, 11N05.
Keywords: Distribution of primes, prime-counting function, logarithmic integral, Riemann hypothesis, Riemann zeta function.

for some sufficiently large x_0. Pintz [19] claims that (1.1) is in fact equivalent to the Riemann hypothesis and is likely to hold for all $x > 2$ under such assumptions. Using explicit bounds on prime-counting functions, we are able to prove this claim.

Theorem 1.1 *The Riemann hypothesis is equivalent to the condition*

$$\int_2^x (\pi(t) - \mathrm{li}(t))\,\mathrm{d}t < 0, \quad \textit{for all } x > 2. \tag{1.2}$$

We also prove an analogous result for the Chebyshev function

$$\theta(x) = \sum_{p \le x} \log p.$$

Theorem 1.2 *The Riemann hypothesis is equivalent to the condition*

$$\int_2^x (\theta(t) - t)\,\mathrm{d}t < 0, \quad \textit{for all } x > 2. \tag{1.3}$$

It is natural to ask whether a modification of (1.2) or (1.3) is true unconditionally. In this direction, we consider the weighted integrals

$$I_1(x, f) = \int_2^x (\pi(t) - \mathrm{li}(t)) f(t)\,\mathrm{d}t \quad \text{and} \quad I_2(x, f) = \int_2^x (\theta(t) - t) f(t)\,\mathrm{d}t \tag{1.4}$$

for some choice of function $f(t)$. Given that one has to go quite far to find a value of t such that[1] $\pi(t) - \mathrm{li}(t) > 0$, one should intuitively take $f(t)$ to be positive and decreasing as to give more weight to the negative bias for small values of t. In [19], Pintz considers $f(t) = \exp(-\log^2 t/y)$ for sufficiently large y. However, using an explicit form of Mertens' theorems, we show that $I_1(x, f) < 0$ and $I_2(x, f) < 0$ for the simpler and asymptotically larger function $f(t) = 1/t^2$. Analogous results also hold for the prime-counting functions

$$\psi(x) = \sum_{p^m \le x} \log p \quad \text{and} \quad \Pi(x) = \sum_{p^m \le x} \frac{1}{m}.$$

Theorem 1.3 *Unconditionally, we have*

$$\int_2^x \frac{\pi(t) - \mathrm{li}(t)}{t^2}\,\mathrm{d}t < 0, \qquad \int_2^x \frac{\theta(t) - t}{t^2}\,\mathrm{d}t < 0,$$
$$\int_2^x \frac{\Pi(t) - \mathrm{li}(t)}{t^2}\,\mathrm{d}t < 0, \qquad \int_2^x \frac{\psi(t) - t}{t^2}\,\mathrm{d}t < 0,$$

for all $x > 2$.

Note that since $\pi(t) < \Pi(t)$ and $\theta(t) < \psi(t)$, the apparent negative bias of $\Pi(t) - \mathrm{li}(t)$ and $\psi(t) - t$ is less pronounced than that of $\pi(t) - \mathrm{li}(t)$ or $\theta(t) - t$. In fact, $\int_2^x \Pi(t) - \mathrm{li}(t)\,\mathrm{d}t$ and $\int_2^x \psi(t) - t\,\mathrm{d}t$ change sign infinitely often (see Lemma 2.8).

[1] The first value of t with $\theta(t) - t > 0$ is also expected to be around 10^{316} (see [20]).

https://doi.org/10.4153/S0008439522000212 Published online by Cambridge University Press

Thus, unlike as in Theorem 1.3, Theorems 1.1 and 1.2 do not hold if $\pi(t)$ and $\theta(t)$ are replaced with $\Pi(t)$ and $\psi(t)$, respectively.

Finally, we show that one cannot do much better than Theorem 1.3 without further knowledge of the location of the zeros of $\zeta(s)$.

Theorem 1.4 *Let $\omega = \sup\{\Re(s) : \zeta(s) = 0\}$. If $1/2 < \omega \le 1$ and $c < 1 + \omega \le 2$, then*

$$\int_2^x \frac{\pi(t) - \mathrm{li}(t)}{t^c}\mathrm{d}t = \Omega_+(1), \qquad \int_2^x \frac{\theta(t) - t}{t^c}\mathrm{d}t = \Omega_+(1),$$
$$\int_2^x \frac{\Pi(t) - \mathrm{li}(t)}{t^c}\mathrm{d}t = \Omega_+(1), \qquad \int_2^x \frac{\psi(t) - t}{t^c}\mathrm{d}t = \Omega_+(1).$$

Here, as per usual, the notation $g(x) = \Omega_+(1)$ means that there exist arbitrarily large values of x such that $g(x) > 0$.

Remark 1.5 Despite the restrictions in Theorem 1.4, it is conceivable that one may be able to use a slightly (asymptotically) larger weight than $f(t) = 1/t^2$ in Theorem 1.3. For instance, $f(t) = \log t/t^2$. Such a result would most likely require the use of an explicit zero-free region, e.g., [18, Theorem 1] or [8, Theorem 5]. We do not pursue this here.

2 Useful lemmas

In this section, we list a series of useful lemmas. Most of the following results are explicit bounds on prime-counting functions which follow directly from existing results in the literature.

Lemma 2.1 [6, Theorem 2] *For all $2 \le x \le 10^{19}$, $\pi(x) - li(x) < 0$ and $\theta(x) - x < 0$.*

Lemma 2.2 [22, Equations (3.5), (3.6), (3.15), and (3.16)] *We have*

$$\frac{x}{\log x} < \pi(x) < \frac{1.3x}{\log x}, \qquad x \ge 17, \tag{2.1}$$

$$x - \frac{x}{\log x} < \theta(x) < x + \frac{0.5x}{\log x}, \qquad x \ge 41. \tag{2.2}$$

Lemma 2.3 *We have*

$$\Pi(x) - 1.9x^{1/2} < \pi(x) < \Pi(x) - x^{1/2}/\log x, \qquad x \ge 17, \tag{2.3}$$

$$\psi(x) - 1.5x^{1/2} < \theta(x) < \psi(x) - 0.98x^{1/2}, \qquad x \ge 121. \tag{2.4}$$

Proof Let $M = \lfloor \frac{\log x}{\log 2} \rfloor$. Then,

$$\Pi(x) = \sum_{m=1}^{M} \frac{1}{m}\pi(x^{1/m}) \le \pi(x) + \frac{M}{2}\pi(x^{1/2}) < \pi(x) + \frac{1.3}{\log 2}x^{1/2} < \pi(x) + 1.9x^{1/2}$$

https://doi.org/10.4153/S0008439522000212 Published online by Cambridge University Press

by Lemma 2.2. On the other hand,

$$\Pi(x) = \sum_{m=1}^{M} \frac{1}{m}\pi(x^{1/m}) > \pi(x) + \frac{1}{2}\pi(x^{1/2}) > \pi(x) + \frac{x^{1/2}}{\log x},$$

as required. The inequalities in (2.4) then follow from equations (3.36) and (3.37) in [22]. ∎

Lemma 2.4 *For all $x > 1$, we have*

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

Proof For $x \geq 1{,}865$, the result follows from [2, Lemma 5.9]. For smaller values of x, the result follows via simple computations. ∎

Lemma 2.5 *We have*

$$-30{,}000 < \int_2^{3{,}000} (\pi(t) - \mathrm{li}(t))\,\mathrm{d}t < -29{,}000,$$
$$-140{,}000 < \int_2^{3{,}000} (\theta(t) - t)\,\mathrm{d}t < -130{,}000,$$
$$-2{,}900 < \int_2^{3{,}000} (\psi(t) - t)\,\mathrm{d}t < -2{,}800.$$

Proof Follows by directly computing each integral on Mathematica. ∎

Lemma 2.6 *Assuming the Riemann hypothesis, $\left|\int_2^x (\psi(t) - t)\,dt\right| \leq 0.08x^{3/2}$ for all $x \geq 3{,}000$.*

Proof By [12, Theorem 27],

$$\psi_1(x) := \int_2^x \psi(t)\,\mathrm{d}t = \frac{x^2}{2} - \sum_\rho \frac{x^{\rho+1}}{\rho(\rho+1)} - x\frac{\zeta'}{\zeta}(0) + \frac{\zeta'}{\zeta}(-1) - \sum_{r=1}^{\infty} \frac{x^{1-2r}}{2r(2r-1)},$$

where the sum is taken over all nontrivial zeros $\rho = \beta + i\gamma$ of the Riemann zeta function. First, we note that

$$\left|\sum_\rho \frac{x^{\rho+1}}{\rho(\rho+1)}\right| \leq x^{1+\omega}\sum_\rho \frac{1}{\gamma(\gamma+1)} \leq 0.04621x^{3/2},$$

because $\sum_\rho 1/\gamma^2 = 0.046209\ldots$ [3, Corollary 1]. Next, $(\zeta'/\zeta)(0) = \log 2\pi$ [7, Section 3.8] and $(\zeta'/\zeta)(-1) = 1.985\ldots$ as computed on Mathematica. Finally,

$$0 \leq \sum_{r=1}^{\infty} \frac{x^{1-2r}}{2r(2r-1)} \leq \frac{x}{2}\sum_{r=1}^{\infty}\frac{1}{x^{2r}} = \frac{x}{2(x^2-1)}.$$

https://doi.org/10.4153/S0008439522000212 Published online by Cambridge University Press

Thus, noting that $\frac{x^2}{2} = \int_2^x t\,\mathrm{d}t + 2$,

$$\left|\int_2^x (\psi(t) - t)\,\mathrm{d}t\right| \le x^{3/2}\left(0.04621 + \frac{\log(2\pi)}{x^{1/2}} + \frac{2 + 1.986}{x^{3/2}} + \frac{1}{2x^{1/2}(x^2 - 1)}\right)$$
$$\le 0.08x^{3/2},$$

for all $x \ge 3{,}000$. ■

Lemma 2.7 (Cf. [12, pp. 103–104]) *Let $Q(x) = \Pi(x) - li(x)$, $R(x) = \psi(x) - x$, and $R_1(x) = \int_2^x R(t)\,dt$. Then,*

$$Q(x) = \frac{R(x)}{\log x} + \frac{R_1(x)}{x\log^2 x} + \int_{3{,}000}^x \left(\frac{R_1(t)}{t^2\log^2 t} + \frac{2R_1(t)}{t^2\log^3 t}\right)\mathrm{d}t + C, \tag{2.5}$$

where

$$C = Q(3{,}000) - \frac{R(3{,}000)}{\log(3{,}000)} - \frac{R_1(3{,}000)}{3{,}000\log^2(3{,}000)} = -0.4351\ldots < 0.$$

Proof Using integration by parts,

$$\mathrm{li}(x) - \mathrm{li}(3{,}000) = \frac{x}{\log x} + \int_{3{,}000}^x \frac{\mathrm{d}t}{\log^2 t} - \frac{3{,}000}{\log(3{,}000)},$$

so that by partial summation,

$$\Pi(x) - \Pi(3{,}000) = \mathrm{li}(x) - \mathrm{li}(3{,}000) + \frac{R(x)}{\log x} - \frac{R(3{,}000)}{\log(3{,}000)} + \int_{3{,}000}^x \frac{R(t)}{t\log^2 t}\mathrm{d}t.$$

Hence,

$$Q(x) = \frac{R(x)}{\log x} + \int_{3{,}000}^x \frac{R(t)}{t\log^2 t}\mathrm{d}t + Q(3{,}000) - \frac{R(3{,}000)}{\log(3{,}000)}.$$

A further application of integration by parts then gives the desired result. ■

Lemma 2.8 *Let $c \in \mathbb{R}$ and define*

$$\Pi_{1,c}(x) := \int_2^x \frac{\Pi(t)}{t^c}\mathrm{d}t \quad \textit{and} \quad \psi_{1,c}(x) := \int_2^x \frac{\psi(t)}{t^c}\mathrm{d}t.$$

If $\omega = \sup\{\Re(s) : \zeta(s) = 0\}$ and $c < 1 + \omega \le 2$, then for all $\delta > 0$,

$$\Pi_{1,c}(x) - \int_2^x \frac{\mathrm{li}(t)}{t^c} = \Omega_\pm(x^{1+\omega-c-\delta}) \quad \textit{and} \quad \psi_{1,c}(x) - \int_2^x \frac{t}{t^c}\mathrm{d}t = \Omega_\pm(x^{1+\omega-c-\delta}).$$

Proof We begin with the integral expression [12, Equation (18), p. 18]

$$\log\zeta(s) = s\int_1^\infty \frac{\Pi(x)}{x^{s+1}}\mathrm{d}x, \quad \Re(s) > 1.$$

https://doi.org/10.4153/S0008439522000212 Published online by Cambridge University Press

Using integration by parts,

$$\begin{aligned}\int_1^\infty \frac{\Pi(x)}{x^{s+1}}\mathrm{d}x &= \int_1^\infty \frac{d\Pi_{1,c}(x)}{x^{s+1-c}}\\ &= \left[\frac{\Pi_{1,c}(x)}{x^{s+1-c}}\right]_1^\infty + (s+1-c)\int_1^\infty \frac{\Pi_{1,c}(x)}{x^{s+2-c}}\mathrm{d}x\\ &= (s+1-c)\int_1^\infty \frac{\Pi_{1,c}(x)}{x^{s+2-c}}\mathrm{d}x,\end{aligned}$$

noting that since $\Pi(x) = O(x)$,

$$\left|\frac{\Pi_{1,c}(x)}{x^{s+1-c}}\right| = O\left(x^{c-1-\Re(s)}\int_2^x t^{1-c}\mathrm{d}t\right) = O(x^{1-\Re(s)}) = o(1).$$

Hence,

$$\log\zeta(s) = s(s+1-c)\int_1^\infty \frac{\Pi_{1,c}(x)}{x^{s+2-c}}\mathrm{d}x, \quad \Re(s) > 1. \tag{2.6}$$

Using an analogous argument with the integral expression [12, Equation (17), p. 18]

$$-\frac{\zeta'(s)}{\zeta(s)} = s\int_1^\infty \frac{\psi(x)}{x^{s+1}}\mathrm{d}x, \quad \Re(s) > 1,$$

one obtains

$$-\frac{\zeta'(s)}{\zeta(s)} = s(s+1-c)\int_1^\infty \frac{\psi_{1,c}(x)}{x^{s+2-c}}\mathrm{d}x, \quad \Re(s) > 1. \tag{2.7}$$

Equipped with (2.6) and (2.7), one can then follow a standard argument (e.g., [4, p. 80], [12, pp. 90–91]) *mutatis mutandis* to obtain the desired result. ■

3 Proofs of Theorems 1.1 and 1.2

We begin with the case where the Riemann hypothesis is false. By Lemma 2.8 with $c = 0$, there are arbitrarily large values of x such that $\int_2^x \Pi(t) - \mathrm{li}(t)\,\mathrm{d}t > Kx^\kappa$ for some positive constants $K > 0$ and $\kappa > 3/2$. For such values of x, we then have by Lemma 2.3 that

$$\begin{aligned}\int_2^x (\pi(t) - \mathrm{li}(t))\,\mathrm{d}t &= \int_2^{17} (\pi(t) - \mathrm{li}(t))\,\mathrm{d}t + \int_{17}^x (\pi(t) - \mathrm{li}(t))\,\mathrm{d}t\\ &> -\int_{17}^x 1.9t^{1/2}\,\mathrm{d}t + \int_{17}^x (\Pi(t) - \mathrm{li}(t))\,\mathrm{d}t + O(1)\\ &= Kx^\kappa + O(x^{3/2}).\end{aligned}$$

Thus, there are arbitrarily large values of x such that

$$\int_2^x (\pi(t) - \mathrm{li}(t))\,\mathrm{d}t > 0,$$

as required. The same reasoning holds for the integral over $\theta(t) - t$ using the corresponding bounds for $\theta(t)$ in Lemmas 2.3 and 2.8.

https://doi.org/10.4153/S0008439522000212 Published online by Cambridge University Press

Now, suppose the Riemann hypothesis is true. To show (1.2) and (1.3), it suffices to consider $x > 10^{19}$ in light of Lemma 2.1. We begin with the integral over $\theta(t) - t$. By Lemmas 2.3, 2.5, and 2.6, we have

$$\begin{aligned}
\int_2^x (\theta(t) - t)\,dt &= \int_2^{3{,}000} (\theta(t) - t)\,dt + \int_{3{,}000}^x (\theta(t) - t)\,dt \\
&< -130{,}000 - \int_{3{,}000}^x 0.98 t^{1/2}\,dt + \int_{3{,}000}^x (\psi(t) - t)\,dt \\
&< -130{,}000 - \frac{1.96}{3}\left(x^{3/2} - (3{,}000)^{3/2}\right) + 0.08x^{3/2} - \int_2^{3{,}000} (\psi(t) - t)\,dt \\
&< -19{,}000 - 0.57x^{3/2} < 0,
\end{aligned}$$

as required.

The integral over $\pi(t) - \mathrm{li}(t)$ requires more work. First, we apply Lemmas 2.3, 2.5, and 2.7 to obtain

$$\begin{aligned}
\int_2^x (\pi(t) - \mathrm{li}(t))\,dt &< \int_2^{3{,}000} (\pi(t) - \mathrm{li}(t))\,dt + \int_{3{,}000}^x (\pi(t) - \mathrm{li}(t))\,dt \\
&< -29{,}000 - \int_{3{,}000}^x \frac{t^{1/2}}{\log t}\,dt + \int_{3{,}000}^x (\Pi(t) - \mathrm{li}(t))\,dt \\
&< -29{,}000 - \int_{3{,}000}^x \frac{t^{1/2}}{\log t}\,dt + \int_{3{,}000}^x \left(\frac{R(t)}{\log t} + \frac{R_1(t)}{t\log^2 t}\right. \\
&\qquad \left. + \int_{3{,}000}^t \left(\frac{R_1(u)}{u^2\log^2 u} + \frac{2R_1(u)}{u^2\log^3 u}\,du\right)\right)dt, \qquad (3.1)
\end{aligned}$$

using the notation from Lemma 2.7. Now, by integration by parts and Lemmas 2.5 and 2.6,

$$\begin{aligned}
\int_{3{,}000}^x \left(\frac{R(t)}{\log t} + \frac{R_1(t)}{t\log^2 t}\right)dt &= \frac{R_1(x)}{\log x} - \frac{R_1(3{,}000)}{\log(3{,}000)} + \int_{3{,}000}^x \frac{2R_1(t)}{t\log^2 t}\,dt \\
&\le 0.08\frac{x^{3/2}}{\log x} + 370 + 0.16\int_{3{,}000}^x \frac{t^{1/2}}{\log^2 t}\,dt. \qquad (3.2)
\end{aligned}$$

Next,

$$\begin{aligned}
\int_{3{,}000}^t \left(\frac{R_1(u)}{u^2\log^2 u} + \frac{2R_1(u)}{u^2\log^3 u}\right)du &\le 0.08\int_{3{,}000}^t \left(\frac{1}{u^{1/2}\log^2 u} + \frac{2}{u^{1/2}\log^3 u}\right)du \\
&\le 0.08\int_{3{,}000}^t \frac{3}{u^{1/2}\log^2 u}\,du.
\end{aligned}$$

Since $u^{1/4}/\log^2 u$ is increasing for $u \ge 3{,}000$, we then have

$$0.08\int_{3{,}000}^t \frac{3}{u^{1/2}\log^2 u}\,du \le 0.24\frac{t^{1/4}}{\log^2 t}\int_{3{,}000}^t \frac{1}{u^{3/4}}\,du \le 0.96\frac{t^{1/2}}{\log^2 t}.$$

https://doi.org/10.4153/S0008439522000212 Published online by Cambridge University Press

Thus,

$$\int_{3,000}^{t}\left(\frac{R_1(u)}{u^2\log^2 u}+\frac{2R_1(u)}{u^2\log^3 u}\right)\mathrm{d}u \le 0.96\frac{t^{1/2}}{\log^2 t}. \tag{3.3}$$

Substituting (3.2) and (3.3) into (3.1) then gives

$$\int_2^x (\pi(t)-\mathrm{li}(t))\,\mathrm{d}t < -28,630-\int_{3,000}^{x}\frac{t^{1/2}}{\log t}\mathrm{d}t+0.08\frac{x^{3/2}}{\log x}+1.12\int_{3,000}^{x}\frac{t^{1/2}}{\log^2 t}\mathrm{d}t. \tag{3.4}$$

Now, using integration by parts,

$$\int_{3,000}^{x}\frac{t^{1/2}}{\log t}\mathrm{d}t=\frac{2}{3}\frac{x^{3/2}}{\log x}-\frac{2}{3}\frac{(3,000)^{3/2}}{\log(3,000)}+\frac{2}{3}\int_{3,000}^{x}\frac{t^{1/2}}{\log^2 t}\mathrm{d}t \ge \frac{2}{3}\frac{x^{3/2}}{\log x}-14,000. \tag{3.5}$$

Moreover,

$$1.12\int_{3,000}^{x}\frac{t^{1/2}}{\log^2 t}\mathrm{d}t \le 1.12\frac{x^{1/4}}{\log^2 x}\int_{3,000}^{x}t^{1/4}\,\mathrm{d}t \le 0.9\frac{x^{3/2}}{\log^2 x}. \tag{3.6}$$

Applying (3.5) and (3.6), we have that (3.4) reduces to

$$\int_2^x (\pi(t)-\mathrm{li}(t))\,\mathrm{d}t < -14,000-0.58\frac{x^{3/2}}{\log x}+0.9\frac{x^{3/2}}{\log^2 x}<0,$$

as required.

4 Proof of Theorem 1.3

First, we note that, via a simple computation, all four integrals in question are negative for $2<x\le 200$. Thus, we may assume throughout that $x>200$.

First, we deal with the integrals involving $\pi(t)$ and $\Pi(t)$. By an explicit form of Mertens' theorem [22, Equation (3.20)], we have

$$\sum_{p\le x}\frac{1}{p}<\log\log x+B+\frac{1}{\log^2 x}, \tag{4.1}$$

where $B=0.2614\ldots$. Now, by partial summation and Lemma 2.2,

$$\sum_{p\le x}\frac{1}{p}=\frac{\pi(x)}{x}+\int_2^x\frac{\pi(t)}{t^2}\mathrm{d}t>\frac{1}{\log x}+\int_2^x\frac{\pi(t)}{t^2}\mathrm{d}t. \tag{4.2}$$

Moreover, by integration by parts and Lemma 2.4,

$$\begin{aligned}\log\log x&=\int_e^x\frac{1}{t\log t}\mathrm{d}t=\frac{\mathrm{li}(x)}{x}-\frac{\mathrm{li}(e)}{e}+\int_e^x\frac{\mathrm{li}(t)}{t^2}\mathrm{d}t\\ &\le\frac{1}{\log x}+\frac{2}{\log^2 x}-\frac{\mathrm{li}(e)}{e}+\int_2^x\frac{\mathrm{li}(t)}{t^2}\mathrm{d}t-\int_2^e\frac{\mathrm{li}(t)}{t^2}\mathrm{d}t.\end{aligned} \tag{4.3}$$

https://doi.org/10.4153/S0008439522000212 Published online by Cambridge University Press

Substituting (4.2) and (4.3) into (4.1) gives

(4.4)
$$\int_2^x \frac{\pi(t) - \mathrm{li}(t)}{t^2} \mathrm{d}t < \frac{2.5}{\log^2 x} + B - \frac{\mathrm{li}(e)}{e} - \int_2^e \frac{\mathrm{li}(t)}{t^2} \mathrm{d}t < \frac{3}{\log^2 x} - 0.62 < 0,$$

as desired. For the integral involving $\Pi(t)$, we then note that by Lemma 2.3,

$$\begin{aligned}
\int_2^x \frac{\pi(t) - \mathrm{li}(t)}{t^2} \mathrm{d}t &= \int_2^{200} \frac{\pi(t) - \mathrm{li}(t)}{t^2} \mathrm{d}t + \int_{200}^x \frac{\pi(t) - \mathrm{li}(t)}{t^2} \mathrm{d}t \\
&> -0.59 + \int_{200}^x \frac{\Pi(t) - 1.9t^{1/2} - \mathrm{li}(t)}{t^2} \mathrm{d}t \\
&= -0.59 + \left(\frac{3.8}{\sqrt{x}} - \frac{3.8}{\sqrt{200}} \right) + \int_2^x \frac{\Pi(t) - \mathrm{li}(t)}{t^2} \mathrm{d}t \\
&\qquad - \int_2^{200} \frac{\Pi(t) - \mathrm{li}(t)}{t^2} \mathrm{d}t \\
&> \int_2^x \frac{\Pi(t) - \mathrm{li}(t)}{t^2} \mathrm{d}t - 0.56 + \frac{3.8}{\sqrt{x}}.
\end{aligned}$$

Substituting this into (4.4) then gives

$$\int_2^x \frac{\Pi(t) - \mathrm{li}(t)}{t^2} \mathrm{d}t < -0.06 - \frac{3.8}{\sqrt{x}} + \frac{3}{\log^2 x} < 0.$$

We argue similarly for the integrals involving $\theta(t)$ and $\psi(t)$. In particular, we have [22, Equation (3.23)]

$$\sum_{p \le x} \frac{\log p}{p} < \log x + E + \frac{1}{\log x},$$

where $E = -1.332\ldots$. Then, by Lemma 2.2,

$$\sum_{p \le x} \frac{\log p}{p} = \frac{\theta(x)}{x} + \int_2^x \frac{\theta(t)}{t^2} \mathrm{d}t > 1 - \frac{1}{\log x} + \int_2^x \frac{\theta(t)}{t^2} \mathrm{d}t$$

and

$$\log(x) = \int_1^x \frac{1}{t} \mathrm{d}t = \int_2^x \frac{t}{t^2} \mathrm{d}t + \log 2.$$

Therefore,

$$\int_2^x \frac{\theta(t) - t}{t^2} \mathrm{d}t < E - 1 + \log 2 + \frac{1.5}{\log x} < -1.63 + \frac{2}{\log x} < 0, \tag{4.5}$$

as required. The corresponding result for $\psi(t)$ then follows similar to before by applying Lemma 2.3 to (4.5). In particular, we have that

$$\int_2^x \frac{\psi(t) - t}{t^2} \mathrm{d}t < -0.83 - \frac{3}{\sqrt{x}} + \frac{2}{\log x} < 0.$$

5 Proof of Theorem 1.4

The result for the integrals involving $\Pi(t)$ and $\psi(t)$ follows immediately from Lemma 2.8. For the integral involving $\pi(t)$, we first note that by Lemma 2.8, for any choice of $\delta > 0$, there exist arbitrarily large values of x such that

$$\int_2^x \frac{\Pi(t) - \mathrm{li}(t)}{t^c} \, \mathrm{d}t > Kx^{1+\omega-c-\delta},$$

for some positive constant $K > 0$. For such values of x, we then have by Lemma 2.3 that

$$\begin{aligned} \int_2^x \frac{\pi(t) - \mathrm{li}(t)}{t^c} \, \mathrm{d}t &= \int_2^{17} \frac{\pi(t) - \mathrm{li}(t)}{t^c} \, \mathrm{d}t + \int_{17}^x \frac{\pi(t) - \mathrm{li}(t)}{t^c} \, \mathrm{d}t \\ &> -\int_{17}^x 1.9t^{1/2-c} \, \mathrm{d}t + \int_{17}^x \frac{\Pi(t) - \mathrm{li}(t)}{t^c} \, \mathrm{d}t + O(1) \\ &> -\int_{17}^x 1.9t^{1/2-c} \, \mathrm{d}t + Kx^{1+\omega-c-\delta} + O(1). \end{aligned} \tag{5.1}$$

The integral in (5.1) satisfies

$$\int_{17}^x 1.9t^{1/2-c} \, \mathrm{d}t = \begin{cases} O(x^{3/2-c}), & c < 3/2, \\ O(\log x), & c = 3/2, \\ O(1), & 3/2 < c < 1 + \omega. \end{cases}$$

Thus, if we take any choice of $\delta < \omega - 1/2$ when $c \leq 3/2$, and $\delta < 1 + \omega - c$ when $3/2 < c < 1 + \omega$, we have

$$\int_2^x \frac{\pi(t) - \mathrm{li}(t)}{t^c} \mathrm{d}t > 0$$

for arbitrarily large values of x as required. The same reasoning holds for the integral over $(\theta(t) - t)/t^c$ using the corresponding bounds for $\theta(t)$ in Lemmas 2.3 and 2.8.

6 Discussion and further work

The general idea in this paper was to consider averaged versions of arithmetic functions in order to gain insight into biases occurring in number theory. The functions $\pi(t) - \mathrm{li}(t)$ and $\theta(t) - t$, in particular, exhibit an apparent negative bias, and our results reflect this.

There are many other biases occurring in number theory, and it would be interesting to consider averaged versions of these. For example, we have:

(a) The bias in Mertens' theorems [5, 15].
(b) The Chebyshev bias for primes in arithmetic progressions [10, 23].
(c) The bias of $\lambda(n)$ and related functions [11, 17].

One could also attempt to extend our results to more general number fields. In this direction, it is worth noting that Garcia and Lee [9] recently proved explicit versions of Mertens' theorems for number fields. Using Garcia and Lee's results could thus allow one to generalize Theorem 1.3, whose proof followed directly from an explicit version of Mertens' theorems in the standard setting.

Acknowledgment Thanks to my supervisor Tim Trudgian and colleagues Ethan, Michaela, and Shehzad for the fun and fruitful discussions on this project.

References

[1] C. Bays and R. H. Hudson, *A new bound for the smallest x with $\pi(x) > li(x)$*. Math. Comp. **69**(2000), 1285–1296.
[2] M. Bennett, G. Martin, K. O'Bryant, and A. Rechnitzer, *Explicit bounds for primes in arithmetic progressions*. Illinois J. Math. **62**(2018), nos. 1–4, 427–532.
[3] R. Brent, D. Platt, and T. Trudgian, *Accurate estimation of sums over zeros of the Riemann zeta-function*. Math. Comp. **90**(2021), 2923–2935.
[4] K. Broughan, *Equivalents of the Riemann hypothesis. Volume 1: arithmetic equivalents*, Cambridge University Press, Cambridge, 2017.
[5] J. Büthe, *On the first sign change in Mertens' theorem*. Acta Arith. **171**(2015), no. 2, 183–195.
[6] J. Büthe, *An analytic method for bounding $\psi(x)$*. Math. Comp. **87**(2018), no. 312, 1991–2009.
[7] H. M. Edwards, *Riemann's zeta function*, Dover Publications, New York, 1974.
[8] K. Ford, *Zero-free regions for the Riemann zeta function*. In: *Number theory for the millennium, II (Urbana, IL, 2000)*, A K Peters, Natick, MA, 2002, pp. 25–56.
[9] S. R. Garcia and E. S. Lee, *Unconditional explicit Mertens' theorems for number fields and Dedekind zeta residue bounds*. Ramanujan J. **57**(2022), no. 3, 1169–1191.
[10] A. Granville and G. Martin, *Prime number races*. Amer. Math. Monthly **113**(2006), no. 1, 1–33.
[11] P. Humphries, *The distribution of weighted sums of the Liouville function and Pólya's conjecture*. J. Number Theory **133**(2013), no. 2, 545–582.
[12] A. E. Ingham, *The distribution of prime numbers*, Cambridge University Press, New York, 1964.
[13] J. Kaczorowski, *On sign-changes in the remainder-term of the prime number formula, I*. Acta Arith. **44**(1984), 365–377.
[14] A. A. Karatsuba, *On the Riemann asymptotic formula for $\pi(x)$*. Dokl. Math. **69**(2004), no. 3, 423–424.
[15] Y. Lamzouri, *A bias in Mertens' product formula*. Int. J. Number Theory **12**(2016), no. 1, 97–109.
[16] J. E. Littlewood, *Sur la distribution des nombres premiers*. CR Acad. Sci. Paris **158**(1914), no. 1914, 1869–1872.
[17] G. Martin, M. J. Mossinghoff, and T. S. Trudgian, *Fake mu's*. Preprint, 2021. arXiv:2112.05227
[18] M. J. Mossinghoff and T. S. Trudgian, *Nonnegative trigonometric polynomials and a zero-free region for the Riemann zeta-function*. J. Number Theory **157**(2015), 329–349.
[19] J. Pintz, *On an assertion of Riemann concerning the distribution of prime numbers*. Acta Math. Hungar. **58**(1991), nos. 3–4, 383–387.
[20] D. J. Platt and T. S. Trudgian, *On the first sign change of $\theta(x) - x$*. Math. Comp. **85**(2016), no. 299, 1539–1547.
[21] B. Riemann, *Ueber die Anzahl der Primzahlen unter einer gegebenen Grosse*. Ges. Math. Werke und Wissenschaftlicher Nachlaß. **2**(1859), 145–155.
[22] J. B. Rosser and L. Schoenfeld, *Approximate formulas for some functions of prime numbers*. Illinois J. Math. **6**(1962), no. 1, 64–94.
[23] M. Rubinstein and P. Sarnak, *Chebyshev's bias*. Exp. Math. **3**(1994), no. 3, 173–197.
[24] Y. Saouter, T. Trudgian, and P. Demichel, *A still sharper region where $\pi(x) - li(x)$ is positive*. Math. Comp. **84**(2015), no. 295, 2433–2446.
[25] E. Schmidt, *Über die Anzahl der Primzahlen unter gegebener Grenze*. Math. Ann. **57**(1903), no. 2, 195–204.
[26] S. B. Stechkin and A. Y. Popov, *The asymptotic distribution of prime numbers on the average*. Russian Math. Surveys **51**(1996), 6.

School of Science, The University of New South Wales Canberra, Northcott Drive, Campbell, ACT 2612, Australia
e-mail: daniel.johnston@adfa.edu.au