Scale–Stable Statistics of Renormalized Prime Gaps

Allen Proxmire

December 14, 2025

Abstract

Prime gaps grow on average with the size of the primes, obscuring potential regularities in their local behavior. A standard normalization by the logarithm of the prime restores approximate scale invariance in the marginal gap distribution, but empirical verification of this stability is often treated implicitly. In this paper we present a systematic empirical study of the renormalized prime gap sequence $g_n/\log p_n$. We demonstrate that the normalized gaps exhibit scale–stable statistics across wide ranges of primes, maintain local stationarity in sliding windows, and differ subtly from randomized null models that preserve marginal gap statistics. These results establish a clean empirical baseline for subsequent investigations of correlation structure in prime gaps.

1 Introduction

Let p_n denote the nth prime and define the prime gaps

$$g_n = p_{n+1} - p_n.$$

It is well known that the average size of g_n grows approximately like $\log p_n$, reflecting the decreasing density of primes. As a result, raw prime gaps are not stationary in scale, complicating statistical comparisons across different regions of the number line.

A standard approach is to consider the renormalized gaps

$$\tilde{g}_n = \frac{g_n}{\log p_n},$$

which are expected, under probabilistic models of the primes, to exhibit approximately scale–invariant behavior. While this normalization is widely used in analytic and probabilistic studies, direct empirical demonstrations of its effectiveness and limitations are often informal.

The purpose of this paper is to provide a careful empirical characterization of the renormalized prime gap sequence. We focus exclusively on first–order statistics: marginal distributions, scale dependence, local stationarity, and comparison to randomized null models. No assumptions about dynamical laws or correlations are made. These baseline results serve as a foundation for the companion study [1], which examines higher–order correlation structure.

2 Data and Computational Setup

All computations in this paper use the sequence of primes up to 10^7, corresponding to approximately 6.6×10^5 consecutive prime gaps. Prime lists were generated using standard sieving methods.

For each prime p_n we compute the gap g_n and the normalized gap $\tilde{g}_n = g_n/\log p_n$. Statistical comparisons are performed across disjoint ranges of p_n as well as in sliding windows along the sequence.

To assess the role of ordering and dependence, we also construct a permutation null model in which the sequence of prime gaps $\{g_n\}$ is randomly permuted. This preserves the marginal gap distribution while destroying all sequential structure.

3 Marginal Distributions Across Scales

Figure 1 shows histograms of the renormalized gaps $\tilde{g}_n$ across two disjoint ranges of primes. The strong overlap of the distributions indicates that normalization by $\log p_n$ effectively removes the dominant scale dependence present in the raw gaps.

We begin by examining the empirical distribution of $\tilde{g}_n$ across disjoint ranges of primes. Specifically, we compare normalized gaps for primes in the intervals $[10^5, 10^6)$ and $[10^6, 10^7)$.

Histograms of $\tilde{g}_n$ in these ranges show strong overlap, with nearly identical central mass and tail behavior. This confirms that normalization by $\log p_n$ effectively removes the dominant scale dependence present in the raw gaps.

Quantitative comparison using Kolmogorov–Smirnov tests reveals no statistically significant difference between the distributions at these scales, supporting approximate scale invariance of the renormalized gaps.

4 Local Stationarity

To probe local stability, we examine the mean of $\tilde{g}_n$ in sliding windows along the prime sequence. As shown in Figure 3, the windowed mean fluctuates narrowly around a constant value with no observable drift, confirming local stationarity of the renormalized gaps.

The resulting time series fluctuates narrowly around a constant value with no observable drift. This demonstrates that the renormalized gap sequence is not only globally scale–stable but also locally stationary over long stretches of primes.

Such local stationarity is essential for meaningful analysis of higher–order structure, as it ensures that any detected correlations are not artifacts of slow global trends.

5 Comparison with Permutation Null Model

To assess whether the observed behavior arises purely from marginal statistics, we compare the renormalized gap distribution to that obtained from a permutation of the prime gaps. Figure 2 shows that the distributions are extremely similar, as expected, though subtle deviations are detectable.

While the renormalized gaps are scale–stable, it is important to assess whether their behavior differs from that of a purely randomized sequence with the same marginal statistics.

We therefore compare the empirical distribution of $\tilde{g}_n$ to that obtained from a permutation of the original gap sequence. By construction, the null model preserves the distribution of gap sizes but destroys all sequential dependence.

The distributions are extremely similar, as expected for a first–order observable. However, subtle differences are detectable in both the body and tails of the distribution, indicating that the ordering of gaps contributes weak but measurable effects beyond marginal statistics alone.

These differences are small in magnitude and do not, by themselves, imply correlation. Rather, they motivate further investigation of higher–order structure, undertaken in the companion paper.

6 Discussion

The results presented here establish several baseline facts about the renormalized prime gap sequence:

- Normalization by $\log p_n$ yields strong scale invariance across multiple decades of primes.
- The normalized gaps exhibit local stationarity in sliding windows.
- Marginal statistics are well approximated by randomized null models, with only subtle deviations.

These observations justify the use of $\tilde{g}_n$ as a stable observable for studying more refined features of the prime gap sequence. Importantly, none of the results in this paper rely on assumptions about correlations or dynamics; they are purely descriptive.

7 Conclusion

We have provided a systematic empirical study of the renormalized prime gap sequence. The results confirm that normalization by $\log p_n$ effectively removes scale dependence and yields a locally stationary sequence with stable marginal statistics.

This paper establishes a clean statistical baseline. In the companion preprint [1], we build upon this foundation to examine correlations between consecutive renormalized gaps and to quantify their local structure.

References

[1] A. Proxmire, *Local Correlation Structure in Renormalized Prime Gaps*, companion preprint, 2025.

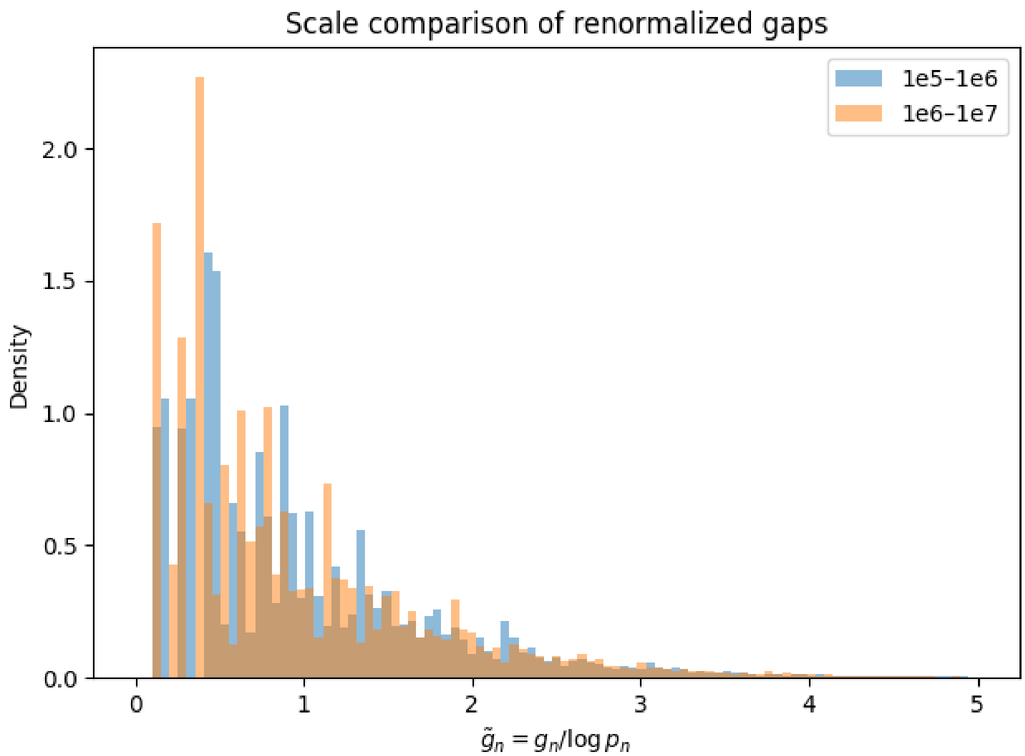

Figure 1: Histogram of renormalized prime gaps $\tilde{g}_n = g_n/\log p_n$ across two disjoint prime ranges. The near coincidence of the distributions demonstrates approximate scale invariance of the normalized gaps.

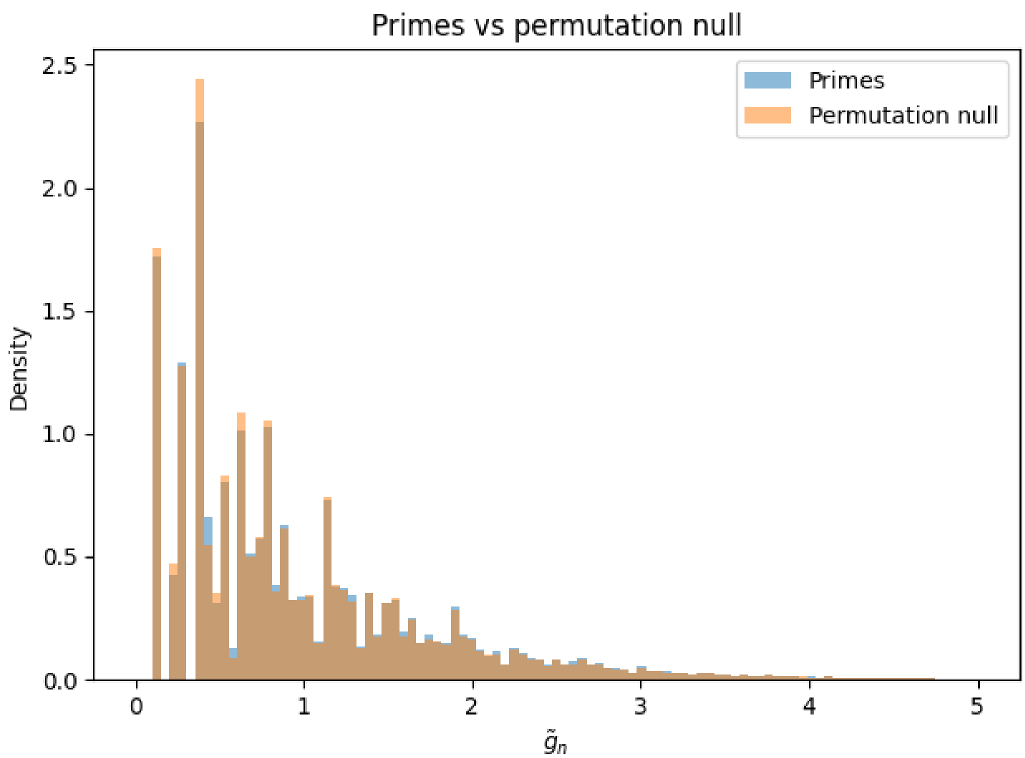

Figure 2: Comparison of renormalized gap distributions for primes and a permutation null model. The null preserves the marginal gap statistics but destroys all sequential structure.

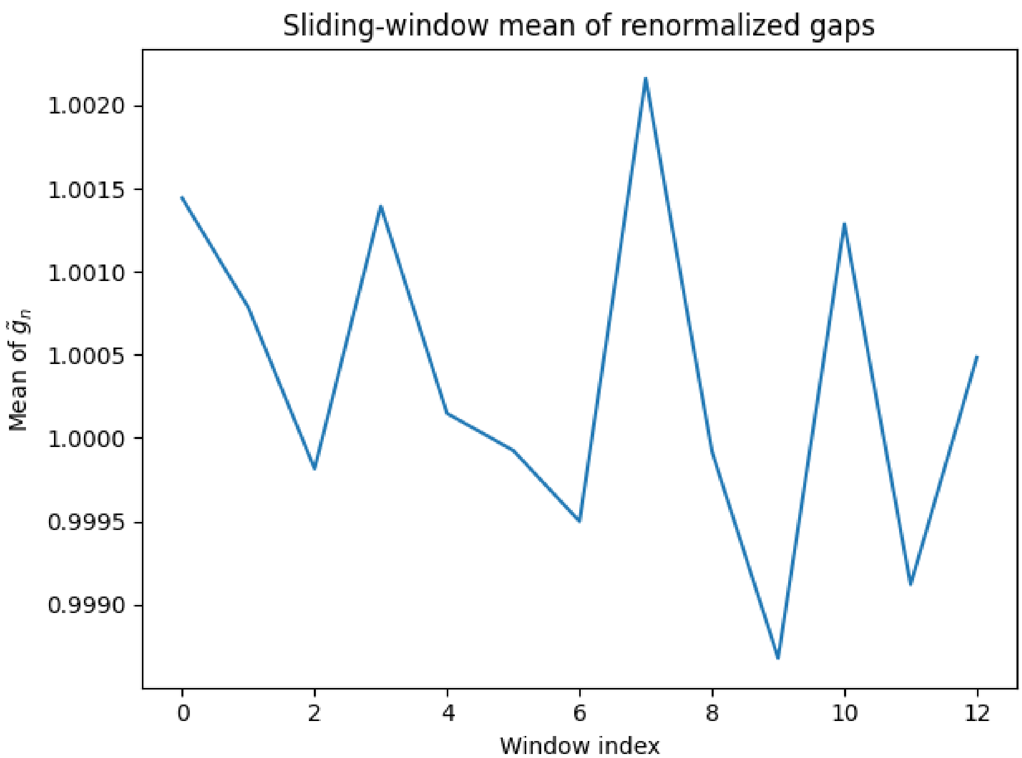

Figure 3: Sliding-window mean of the renormalized prime gaps $\tilde{g}_n$. The absence of systematic drift indicates local stationarity over long stretches of the prime sequence.