The Prize Winners Dataset: Extending Jones and Weinberg to 2026, and Beyond the Nobel

Enrico D. Turri*

July 26, 2026

Link to latest version

Abstract

This note accompanies the dataset on scientific prize winners available on my website. The dataset extends the data underlying the burden-of-knowledge facts documented by Jones (2009, 2010) and Jones and Weinberg (2011) in two directions: temporally, to 2026, and across prizes, by adding the Turing Award, the Abel Prize, the Wolf Prize in Mathematics, and the Fields and Clark Medals — for a total of 1,046 awards. The note describes the dataset, whose main variables are, for each awardee, the date of birth, the date of the award and the date of the great achievement that warranted the prize, and it extends selected figures and tables in the papers cited above. The main findings are that (i) the original results can be replicated, though not perfectly (parts of the papers' datasets are not available online), (ii) the upward trend in age at great achievement persisted into the new millennium, and (iii) if anything, it has accelerated.

Keywords: Burden of knowledge; Age and invention; Scientific creativity; Nobel Prize; Research productivity.

JEL Codes: O31, O33, J24, I23, N30.

*London School of Economics. Email: e.turri@lse.ac.uk. I built this dataset mostly for fun, starting from Benjamin Jones's work because I always found his papers on the topic stimulating and based on cool data. This is an accompanying note (which, for example, doesn't properly survey the literature) which I wrote as a first point of access to the data. Feel free to use and extend the dataset, and please reach out to me if you find any errors or you have suggestions.

1 Dataset

The Prize Winners Dataset, available on my website at edturri.github.io, covers

- the Nobel Prize in Physics, Chemistry, Medicine and Economics;[1]
- the Fields Medal;
- the John Bates Clark Medal;
- the A.M. Turing Award;
- the Wolf Prize in Mathematics;
- the Abel Prize,

from the year each prize was first awarded through July 2026. For every winner it records the date of birth and three key ages:

1. the age at which the winner earned their highest degree;
2. the age at the time of the work that eventually led to the award;
3. the age at which the winner received the prize.

The first and the third follow directly from standard biographical information, easily sourced online. To date the prize-winning work, I follow the measurement strategy of Jones (2010) and Jones and Weinberg (2011), who exploit the text of the prize citation to back out the year of the "great achievement". In practice, citations often describe a research window rather than a single year; in those cases I take the midpoint of the window and subtract the year of birth.

Two of the prizes, the Fields and Clark Medals, are awarded only to recipients under the age of 40, and the corresponding records carry an age-restriction flag. For the two lifetime-achievement prizes, the Abel and Wolf Prizes, the citation spans a career rather than a single discovery, so the midpoint is better read as a career midpoint than as a date of discovery.

The data come from the awarding bodies' own documents, supplemented for mathematics by the MacTutor History of Mathematics Archive and the Mathematics Genealogy Project. Every newly collected Nobel observation carries a source URL. Coverage is near-complete: the age at great achievement is available for 1,044 of the 1,046 records, and the degree year for all of them. Table 1 summarises the panel by prize and Nobel category.

[1] I write "Nobel Prize in Economics" throughout, although the prize in Economics was not among those established by Alfred Nobel's will: it is formally the *Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel*.

2 Figures and Tables

Summary statistics.

Prize	N	Award years	Age at degree	Age at great achievement	Age at prize	Recognition lag
Nobel Prize (all)	756	1901–2025	26.7 (3.6)	39.4 (8.5)	59.7 (12.8)	20.3 (11.6)
Physics	228	1901–2025	26.6 (3.4)	38.1 (9.7)	57.7 (15.0)	19.6 (13.3)
Chemistry	199	1901–2025	26.1 (3.8)	40.5 (8.1)	59.2 (12.1)	18.7 (10.7)
Medicine	230	1901–2025	26.8 (3.5)	40.3 (7.7)	58.8 (11.7)	18.5 (9.8)
Economics	99	1969–2025	27.7 (4.0)	37.8 (7.7)	67.0 (8.3)	29.2 (8.9)
Fields Medal‡	64	1936–2022	25.6 (2.5)	30.5 (3.6)	35.8 (3.3)	5.2 (2.0)
John Bates Clark Medal‡	48	1947–2026	26.8 (2.2)	31.6 (2.5)	37.7 (1.6)	6.1 (2.3)
Turing Award	81	1966–2025	27.1 (3.7)	35.1 (6.1)	59.0 (11.3)	24.0 (10.8)
Wolf Prize in Mathematics	68	1978–2024	25.0 (2.7)	37.4 (5.9)	67.8 (8.2)	30.4 (8.4)
Abel Prize	29	2003–2026	25.0 (2.2)	36.5 (5.5)	76.3 (6.5)	39.8 (8.6)
All prizes	1,046	1901–2026	26.5 (3.5)	37.9 (8.2)	58.1 (14.2)	20.2 (12.4)

‡ Age-capped prize (award restricted to age <40).

Table 1: Values are means with standard deviations in parentheses. N counts laureate–prize records.

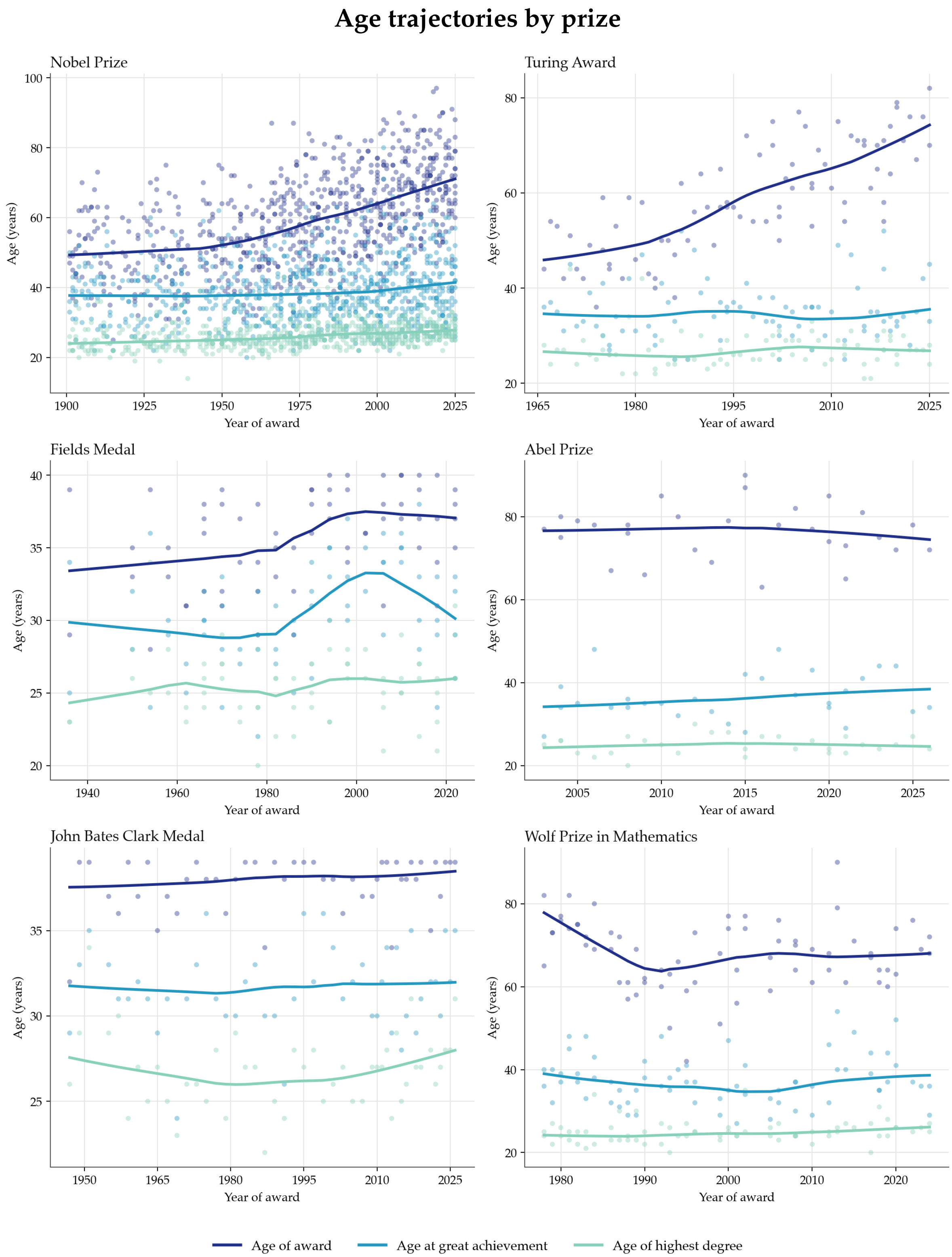

Figure 1: Age at the award, age at the great achievement, and age at the highest degree, plotted against the year of the award, for the six prizes in the dataset. Points are individual winners; lines are LOWESS smooth fits. The Fields and Clark Medals are capped at 40 years of age. The Abel Prize and the Wolf Prize honour a career rather than a single discovery.

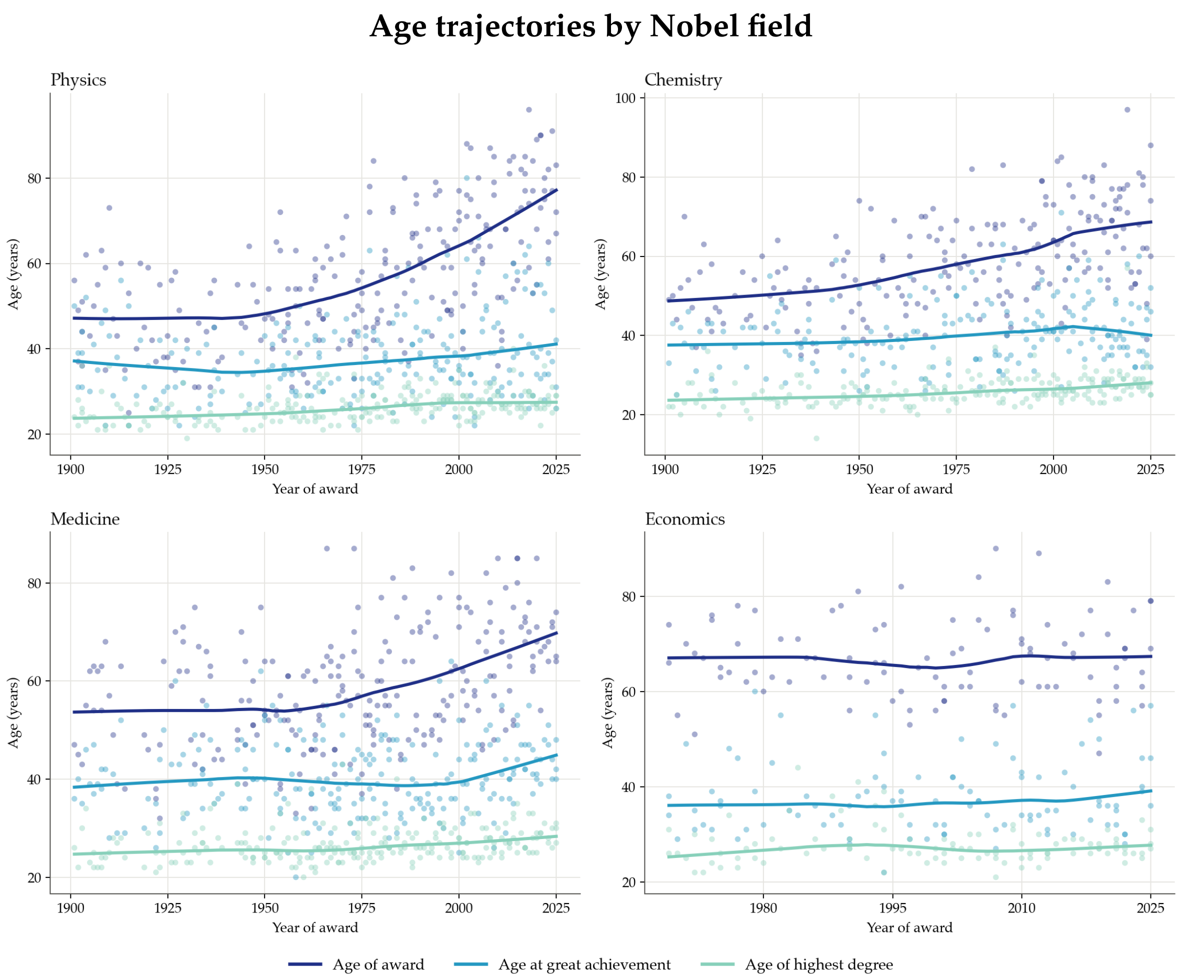

Figure 2: Age at the award, age at the great achievement, and age at the highest degree, plotted against the year of the award for Nobel prize winners by field. Points are individual laureates; lines are LOWESS smooth fits.

Age trends among prize winners.

Panel A. Nobel prizes

	(1)	(2)	(3)	Award years	N	Mean age at achievement
Jones (2010), published	5.83^{***} (1.37)	6.34^{***} (1.36)	7.79^{***} (1.54)	1901–2003	544	38.6
Jones (2010), replicated	5.50^{***} (1.33)	5.91^{***} (1.33)	6.87^{***} (1.52)	1901–2003	542	38.6
Full sample	6.95^{***} (1.00)	7.47^{***} (1.00)	8.35^{***} (1.12)	1901–2025	756	39.4
Field fixed effects	No	Yes	Yes			
Country of birth fixed effects	No	No	Yes			

Panel B. All prizes

Prize	(1)	(2)	Award years	N	Mean age at achievement
Nobel Prize	6.95^{***} (1.00)	8.01^{***} (1.12)	1901–2025	756	39.4
Fields Medal‡	5.54^{***} (1.89)	5.76^{**} (2.65)	1936–2022	64	30.5
John Bates Clark Medal‡	1.91 (1.51)	0.80 (2.11)	1947–2026	47	31.6
Turing Award	14.09^{**} (6.08)	12.51^{*} (7.33)	1966–2025	80	35.1
Wolf Prize in Mathematics	7.77^{*} (4.11)	12.10^{**} (5.42)	1978–2024	68	37.4
Abel Prize	26.86^{***} (7.13)	30.50^{***} (7.41)	2003–2026	29	36.5
Country of birth fixed effects	No	Yes			

Panel C. Testing for acceleration

	Achievement year	(Achievement year)^{2†}	Implied trend/century		N
			1920	2010	
Nobel Prize — all fields	7.03^{***} (0.99)	12.90^{***} (2.85)	−0.7	22.5	756
Nobel Prize — sciences	7.55^{***} (1.04)	12.36^{***} (3.03)	0.1	22.4	657
All uncapped prizes	5.91^{***} (0.91)	15.02^{***} (2.62)	−3.1	23.9	933

Significance: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.
‡ Age-capped prize (award restricted to age <40).
† Quadratic in the year of achievement, centered at 1950.

Table 2: OLS regressions of the age at great achievement on the year of achievement, divided by 100, so that each coefficient is the number of years of age added per century; standard errors in parentheses. Panel A compares the coefficients published by Jones (2010) with the same specifications re-estimated on our dataset using the Jones's time window and the full panel. Panel B estimates the trend prize by prize. Panel C adds a quadratic in the achievement year and reports the slope it implies in 1920 and in 2010.

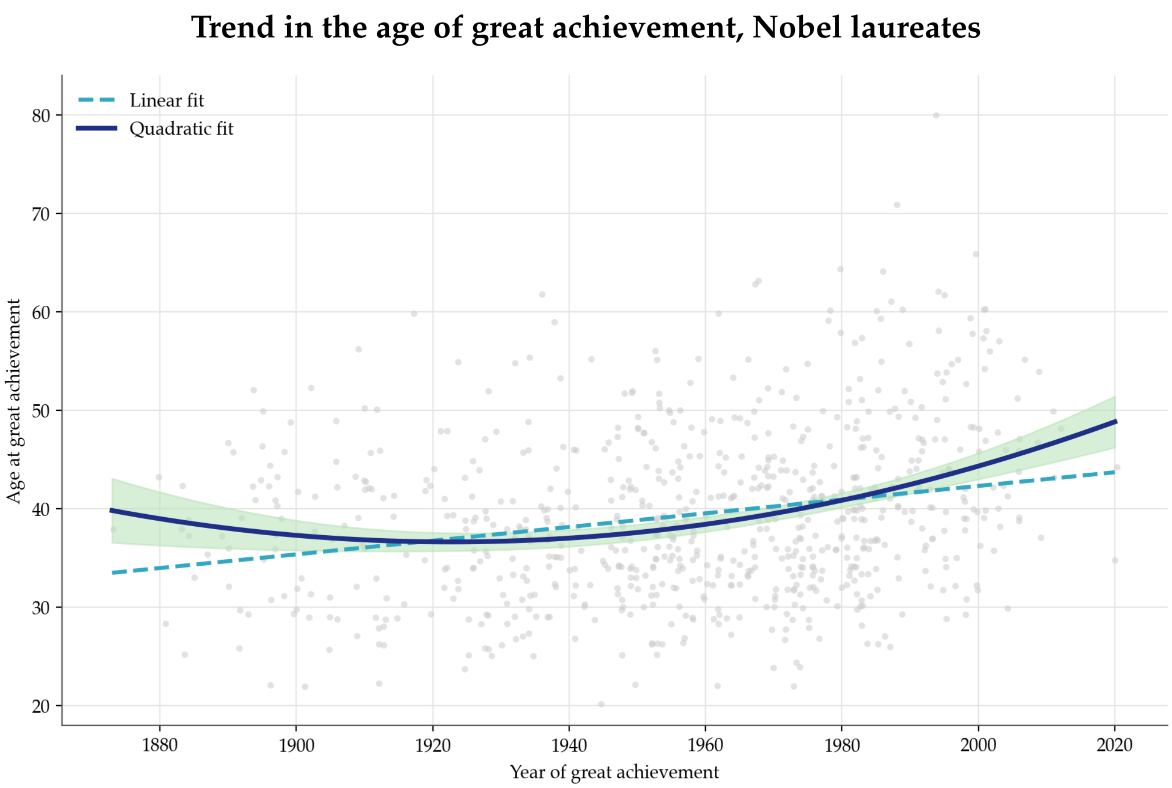

Figure 3: Each grey point is one laureate, plotted at the year of great achievement and at the age reached in that year. The dashed line is the fitted linear trend of Table 2, Panel A, column (1); the solid line adds a quadratic term, with its shaded 95% confidence band. The quadratic fit is the graphical counterpart of Panel C: the trend is essentially flat until the interwar decades and steepens thereafter, reaching about 22 years per century by 2010.

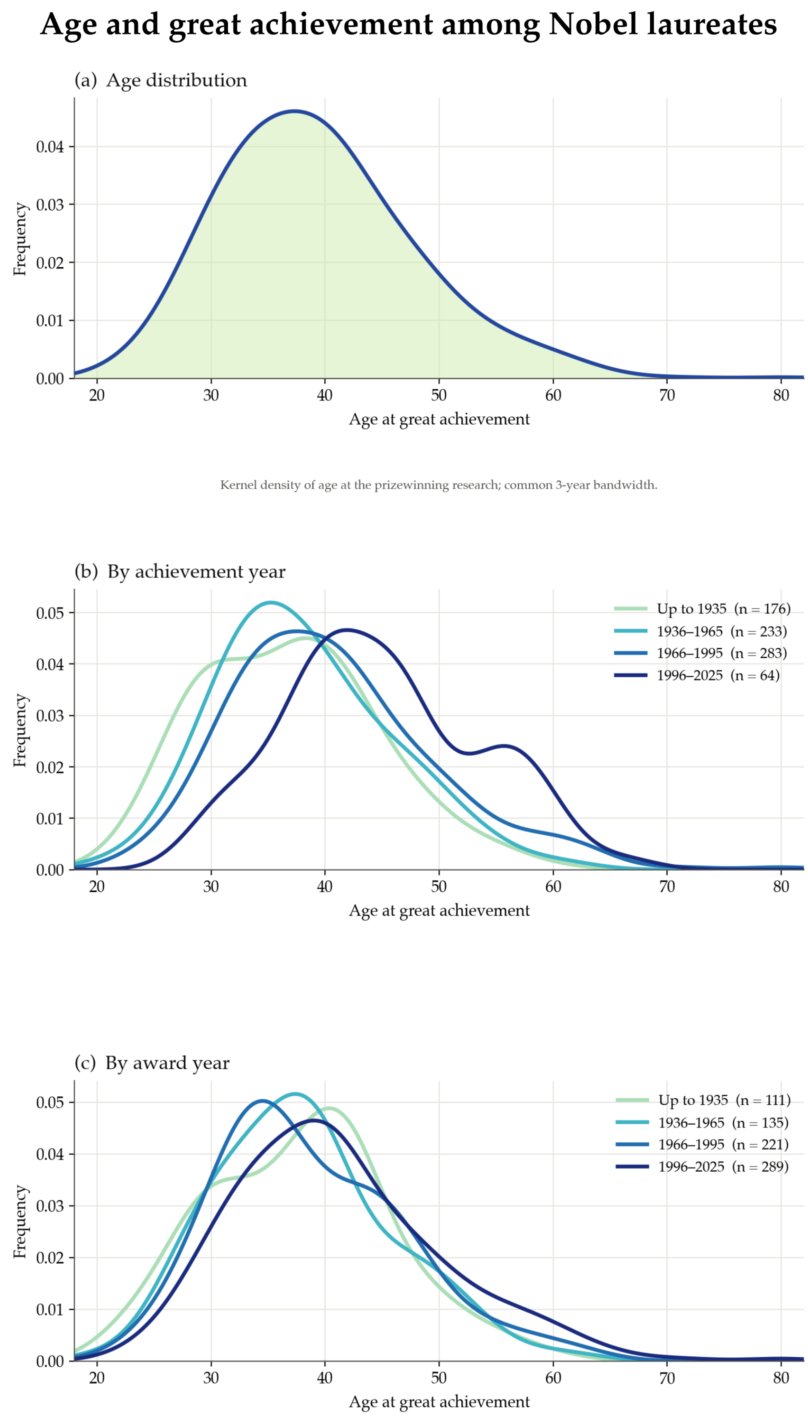

Figure 4: Kernel densities of the age at which the prizewinning research was done, pooling the four Nobel fields. Panel (a) is the whole sample; panel (b) splits laureates into thirty-year bins by the year of achievement; panel (c) splits them by the year of the award. All panels use a three-year bandwidth.

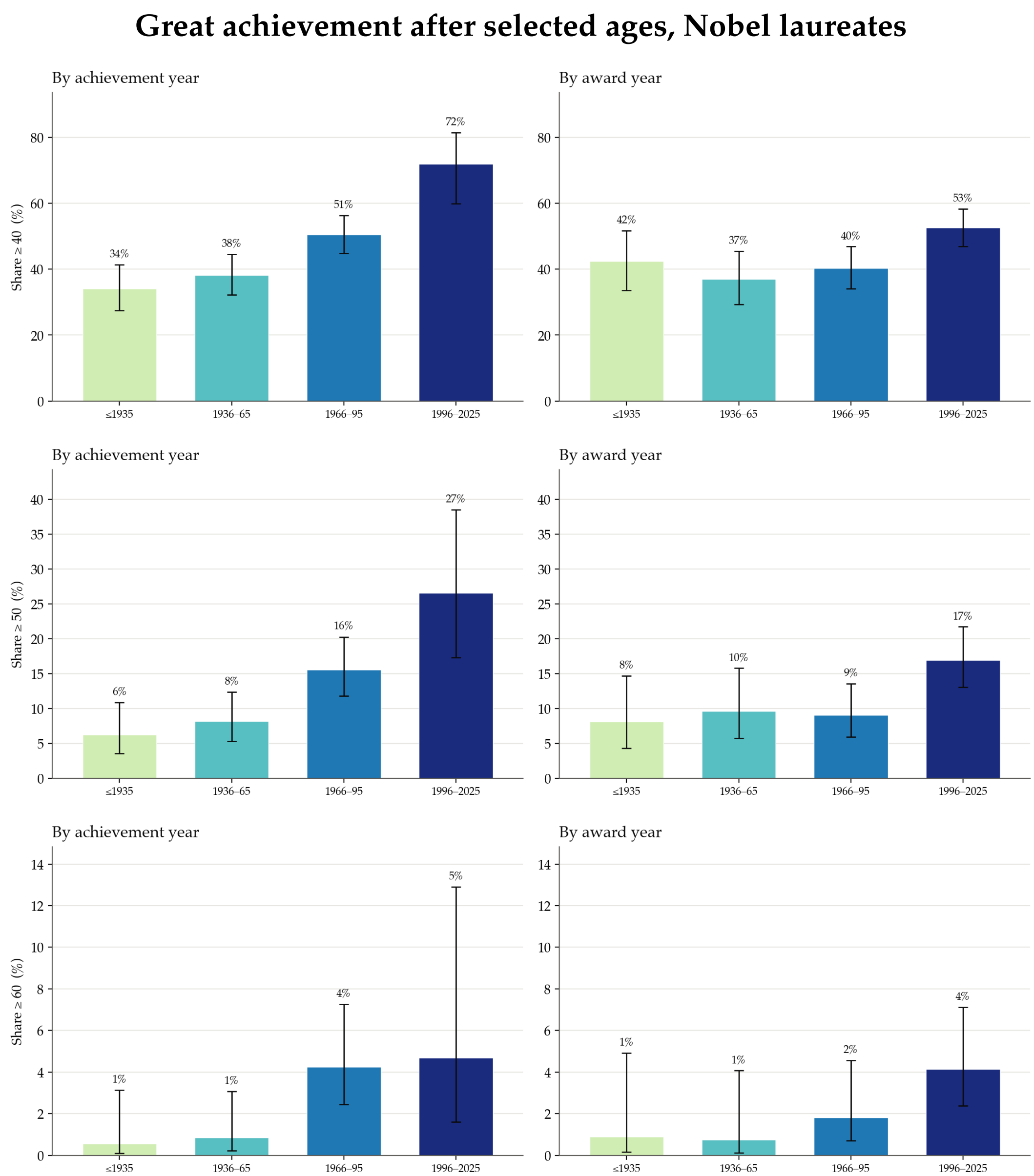

Figure 5: Share of laureates whose prizewinning work was done at or after age 40 (top row), 50 (middle) and 60 (bottom), by thirty-year period. The left column dates laureates by the year of the achievement, the right column by the year of the award. Bars are sample shares with 95% confidence intervals; note the differing vertical scales across rows.

The life cycle of innovation potential.

	(1) Jones (2010), published	(2) Jones (2010), replicated	(3) Full sample
Early life cycle logistic curve			
μ_0	22.8 (2.84)	23.78 (4.94)	21.58 (3.31)
μ_1	10.29 (4.56)	9.98 (6.95)	13.82 (5.76)
ω	2.30 (0.46)	2.14 (0.47)	2.67 (0.44)
Later life cycle logistic curve			
θ_0	43.7 (3.54)	43.29 (2.86)	43.75 (1.70)
θ_1	0.00 (0.00)	0.00 (3.89)	0.00 (5.25)
ρ	6.24 (0.99)	6.05 (0.74)	6.26 (0.67)
Observations	181	178	269
Log likelihood	-633.6	-618.5	-945.4

Table 3: Maximum-likelihood estimates of the two-logistic life-cycle model of Jones (2010), fitted to US-born Nobel laureates with the US population by age as the population at risk. The early curve has location $\mu = \mu_0 + \mu_1 s$ and scale ω, the later curve location $\theta = \theta_0 + \theta_1 s$ and scale ρ, where s is the achievement year measured in centuries since 1900; μ_1 is the rightward shift per century in the age at which innovation potential begins. Column (1) reports the published estimates, column (2) re-estimates them on Jones's estimation window, and column (3) uses the full 1901–2025 panel. Bootstrap standard errors (300 replications) in parentheses.

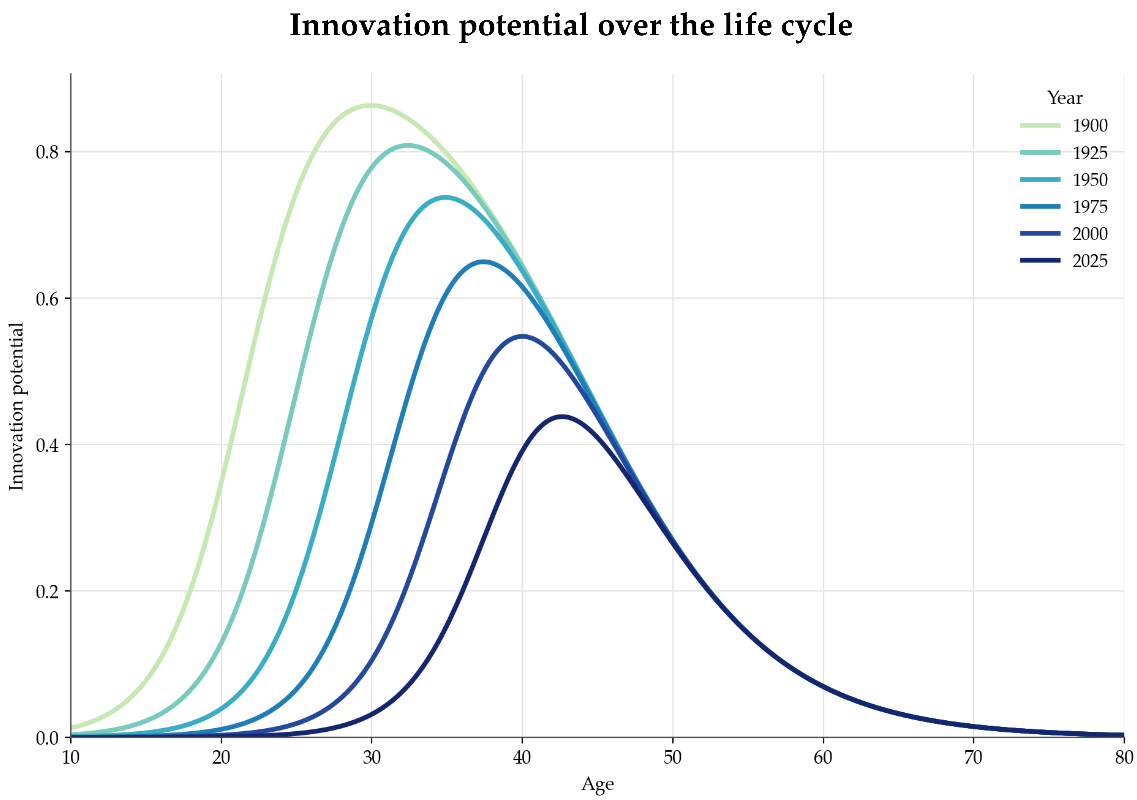

Figure 6: Fitted life-cycle curves implied by the estimates in Table 3, column (3), evaluated every twenty-five years. Innovation potential is the product of an ascending logistic in age, which governs the arrival at the research frontier, and a descending logistic, which governs the decline thereafter. The curves shift right and flatten across the century: the modal age of prizewinning work moves from about 30 in 1900 to about 43 in 2025, and the potential attained at that peak roughly halves.

3 Results

This is a non-exhaustive list of results and intuitions that can be drawn from the figures and tables above.

- **The original results replicate.** Restricting the analysis to the time window of Jones (2010) (Nobel prizes until 2003), the age of achievement time trend reproduces at 5.5 years per century (close to the published 5.83) and the life-cycle innovation model reproduces within one standard error of every published parameter (Table 2, Panel A; Table 3). Data for Jones (2010) isn't available online. While this dataset coincides with the data of Jones and Weinberg (2011), the latter doesn't include winners of the Nobel Prize in Economics which I independently reconstructed.

- **The age at achievement trend persists in the third millennium.** On the full 1901–2025 Nobel panel the achievement-age trend is again positive and significant, and rises to 6.95 years per century (Table 2, Panel A).

- **Great achievement after 50 years old has become more frequent in the last 30 years.** While until 1996 most of the trend was due to the decreasing number of early discoverers (those below 30 years old especially), and the frequency of great achievement after 50 years of age rose only gradually, in the last 30 years the mass of the right tail has jumped, especially between 50 and 60 years of age (Figures 4 and 5).

- **The trend might be accelerating.** Introducing a quadratic term (centred at 1950) in the regression, I find it is strongly significant and positive. This means that the rise in great achievement age might be accelerating over time. The implied slope goes from roughly zero around 1920 to about 22 years per century by 2010 (Table 2, Panels A and C; Figure 3).

- **Innovation potential peaks later and lower over the century.** The estimated life cycle shifts right and its peak flattens: the modal age of prizewinning work moves from about 30 in 1900 to about 43 in 2025, and the potential attained at that peak roughly halves (Figure 6; Table 3). The shift is driven by the early life cycle logistic curve; the late life cycle one is not affected. From the figure, it is also apparent that innovation potential at higher ages is not improving, which paints a bleak picture for the future. However, some additional exploratory analysis I conducted seems to suggest that the logistic model under-explains the rise in the number of prizewinning discoveries at later ages (after 50 years old) seen in the data in the last quarter of the century.

- **Heterogeneity by award.** Turing laureates do their prizewinning work youngest among uncapped prizes, at about 35. The two age-capped prizes — the Fields and Clark Medals — mechanically compress the recognition lag and report an age of great discovery lower by about six years relative to their uncapped counterparts (Figure 1; Table 2, Panel B). However, the Fields medallists still exhibit a positive trend in the age of prize-worthy achievement, which might put a question mark on the long-term viability of such an age-capped prize. The Clark Medal is the only prize not to report any significant trend. The Wolf and Abel prizes exhibit strong trends. However, since they mostly recognise a career rather than a single discovery, and the year of achievement is a midpoint, this result might be a by-product of longer careers. The strong trend in the Turing prize might be due to informatics being a relatively younger field.

4 Will AI Help Us Overcome the Burden of Knowledge?

In 2024, half of the Nobel Prize in Chemistry went to Demis Hassabis and John M. Jumper, the DeepMind researchers behind AlphaFold, for the prediction of protein structures.[2] This was arguably the first Nobel Prize awarded for a scientific discovery made possible by artificial intelligence.[3] Beyond its symbolic value as a milestone for AI, the award stands out for the youth of one of its recipients: Jumper was 39 at the time of the award and 35 when AlphaFold2 won the CASP14 assessment in 2020, against 48 and 44 for Hassabis. At 39, Jumper is the youngest Chemistry laureate since Richard Synge in 1952, who was 38. This raises the question: will AI reverse the burden of knowledge?

AI might lower the age at great achievement, both by compressing training time and by making it cheaper to search over the space of recombinations. On the other hand, it might instead raise that age, if it expands the frontier faster than it lowers the cost of reaching it. The dataset can't answer the question today, but it may be able to in a few years' time. Given the recognition lag that major scientific prizes involve, the age-capped prizes — the Fields and Clark Medals — are likely to register any change in trend first.

[2]The other half was awarded to David Baker, for computational protein design.

[3]The distinction is between AI as a *tool* and AI as the *subject*: the 2024 Physics prize to Hopfield and Hinton, announced a day earlier, honoured the machine-learning methods themselves rather than a discovery enabled by them.

References

Jones, B. F. (2009). The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? *Review of Economic Studies 76*(1), 283–317.

Jones, B. F. (2010). Age and great invention. *Review of Economics and Statistics 92*(1), 1–14.

Jones, B. F. and B. A. Weinberg (2011). Age dynamics in scientific creativity. *Proceedings of the National Academy of Sciences 108*(47), 18910–18914.