Supplementary Information: Background growth, count noise and the first accepted year of a patent take-off rule

Andreas Peters

Abstract

Complete analytical proofs, empirical source definitions, sensitivity results, controlled-background experiments, criterion margins and the count-noise simulation accompany the main article. Official observations, constructed paths, simulated counts and conditional mathematical benchmarks are distinguished throughout.

Local author-review draft. Not cleared for submission; author information and declarations require personal confirmation.

S1 Filter identities and response to a growth-rate change

S1.1 Calendar definition and linearity

Let $N_{it} \geq 0$, $\ell_{it} = \log(1 + N_{it})$, and $z_{it} = \ell_{it} - \ell_{i,t-1}$. The identities below apply to any nonnegative level sequence, including the real-valued benchmark used for the dating proof. They do not depend on interpreting its levels as patent families, applications or grants. Missing calendar observations remain missing: adjacent stored rows cannot be substituted for consecutive years.

For positive integers p and q, define the post-period and pre-period endpoint growth rates by

$$r_{ik}^{\mathrm{post},p} = \frac{\ell_{i,k+p-1} - \ell_{i,k-1}}{p}, \qquad r_{ik}^{\mathrm{pre},q} = \frac{\ell_{i,k-1} - \ell_{i,k-q-1}}{q}, \qquad D_{ik}^{(p,q)} = r_{ik}^{\mathrm{post},p} - r_{ik}^{\mathrm{pre},q}. \tag{S1}$$

Evaluation requires the complete level block $k - q - 1, \ldots, k + p - 1$. There are $q + 1$ pre-period levels, p post-period levels, and $p + q$ growth intervals. In the worked rule, $(p, q) = (3, 4)$, so the block runs from $k - 5$ through $k + 2$.

Proposition S1.1 (Difference of adjacent growth means) *For every complete candidate block,*

$$D_{ik}^{(p,q)} = \frac{1}{p}\sum_{h=0}^{p-1} z_{i,k+h} - \frac{1}{q}\sum_{h=1}^{q} z_{i,k-h}. \tag{S2}$$

The operator is linear in the growth sequence and annihilates a constant growth rate.

Proof The first sum telescopes to $\ell_{i,k+p-1} - \ell_{i,k-1}$, and the second to $\ell_{i,k-1} - \ell_{i,k-q-1}$. Division by the respective numbers of intervals gives Eq. (S2). Its p positive weights sum to one and its q negative weights to minus one. Their total is zero. A constant input $z_{it} = c$ therefore gives $c - c = 0$; linearity follows from the fixed weighted sum. □

For the worked rule the weights, in calendar order on changes ending in $k-4, \ldots, k+2$, are

$$\left(-\tfrac{1}{4}, -\tfrac{1}{4}, -\tfrac{1}{4}, -\tfrac{1}{4}, \tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}\right). \tag{S3}$$

The contrast is a comparison of mean growth, not a direct measure of the count level. The familiar moving-sum weights are retained to make the subsequent nonlinear dating argument reproducible, rather than as a new filter construction.

An optional additive representation makes the cancellation explicit. Suppose, as an algebraic specification rather than an estimated model, that

$$z_{it} = \alpha_i + \lambda_i \theta_t + u_{it}, \qquad D_{ik}^{(p,q)} = \lambda_i D^{(p,q)}(\theta)_k + D^{(p,q)}(u_i)_k. \tag{S4}$$

Here α_i and λ_i are constant over the candidate block. The second identity follows by linearity and the removal of α_i. It says nothing about whether a common component can be identified or fitted to a particular panel. Time-varying loadings cannot be taken outside the operator in this way.

No distributional assumption is needed for this decomposition. In particular, the filtered residual need not vanish or be independent of the filtered common input. An observed contrast cannot be attributed to one of those terms merely because the decomposition can be written down. Estimation and identification would require additional restrictions that the identity itself does not supply.

S1.2 Exact step response

Consider a permanent increase in a growth input, $\theta_t = a + \Delta \mathbf{1}\{t \geq \tau\}$, with $\Delta > 0$, and put $d = k - \tau$. Write $[x]_0^m = \min\{m, \max\{x, 0\}\}$.

Proposition S1.2 (Response of an asymmetric contrast) *The filtered contribution of this input to Eq.* (S4) *is*

$$\lambda_i D^{(p,q)}(\theta)_k = \lambda_i \Delta \psi_{p,q}(d), \qquad \psi_{p,q}(d) = \frac{[d+p]_0^p}{p} - \frac{[d]_0^q}{q}. \tag{S5}$$

For integer displacements, $\psi_{p,q}$ is zero for $d \leq -p$ and $d \geq q$, positive for $-p < d < q$, and reaches its maximum of one at $d = 0$.

Table S1 Changed-interval counts for the worked $(p, q) = (3, 4)$ contrast.

d	Post-period count	Pre-period count	$\psi_{3,4}(d)$
$d \le -3$	0	0	0
-2	1	0	1/3
-1	2	0	2/3
0	3	0	1
1	3	1	3/4
2	3	2	1/2
3	3	3	1/4
$d \ge 4$	3	4	0

Proof Among post-period changes indexed by $k+h$, $h = 0, \ldots, p-1$, exactly $[d+p]_0^p$ fall on or after τ. Among pre-period changes indexed by $k-h$, $h = 1, \ldots, q$, the corresponding number is $[d]_0^q$. Each changed post-period observation adds Δ/p to the contrast; each changed pre-period observation subtracts Δ/q. The constant component a cancels, which proves the formula.

If $d \le -p$, neither window includes the changed rate. For $-p < d < 0$, the pre-period count is zero and the post-period count is strictly between zero and p. At $d = 0$, all post-period and no pre-period changes use the new rate, giving one. For $0 < d < q$, the post-period count is p and the pre-period count lies strictly between zero and q. When $d \ge q$, both averages contain the new rate throughout and their difference is zero. □

The interval counts in Table S1 show how the response before τ arises: an earlier candidate's post-period already contains observations from τ onward. It is retrospective window overlap, not foresight. More generally, adding this growth increment to any background adds precisely $\Delta\psi_{p,q}(d)$ to its contrast. The incremental identity does not determine the maximum of the total contrast on an irregular background or the first date accepted by the full rule. Those are different quantities.

S2 Complete proof of the first-acceptance ladder

S2.1 Rule and benchmark

For $(p, q) = (3, 4)$, retain the stored growth threshold $c_1 = 0.1133286853$, the gap threshold $c_2 = 0.1$, the median level ratio 1.5, at least two strictly positive anchored post-period changes, and a post-period median of at least 20. The three anchored changes are $N_k - N_{k-1}$, $N_{k+1} - N_k$ and $N_{k+2} - N_{k+1}$. Both growth thresholds and the level/count thresholds use weak inequalities; the change signs are strict. A chronological acceptance at k suppresses $k+1, \ldots, k+5$ and permits $k+6$ again.

The proof concerns a flat pre-change level $N_t = N_0 \ge 20$ for $t < \tau$ and a permanent increase from zero to $\Delta > 0$ in $\log(1 + N_t) - \log(1 + N_{t-1})$. Define

$$M_j = (1 + N_0)e^{j\Delta} - 1 \quad (j \ge 0), \qquad N_t = M_{t-\tau+1} \quad (t \ge \tau). \tag{S6}$$

Table S2 Full-rule components under the pure growth step. Medians are count levels; positive changes are counted among the three anchored differences. The count floor holds in every row.

d	r^{post}/Δ	D/Δ	Pre median	Post median	Positive changes
-2	1/3	1/3	N_0	N_0	1
-1	2/3	2/3	N_0	M_1	2
0	1	1	N_0	M_2	3
1	1	3/4	N_0	M_3	3
2	1	1/2	N_0	M_4	3
3	1	1/4	M_1	M_5	3

Thus $M_0 = N_0$ and M_j increases strictly with j. The path is a real-valued construction, not a claim about integer-valued observations. Rounding it may alter the rule and is outside this exact proof.

Every candidate from $\tau - 2$ through $\tau + 3$ is assumed to have its complete calendar block. Evaluation is chronological, with no inherited cooldown excluding a candidate before the first acceptance. Later cooldown exclusions do not affect which candidate was accepted first. No disturbance term or non-flat pre-change trend is included.

Put

$$\begin{aligned} Q(N_0) &= \log\left(\frac{1.5N_0 + 1}{N_0 + 1}\right), \\ T_{-1}(N_0) &= Q(N_0), \qquad T_0(N_0) = \frac{Q(N_0)}{2}, \\ T_1(N_0) &= \max\left\{c_1, \frac{4c_2}{3}, \frac{Q(N_0)}{3}\right\}. \end{aligned} \tag{S7}$$

Proposition S2.1 (First acceptance of the five-criterion rule) *Under the stated benchmark and eligibility assumptions,*

$$k^*(\Delta) = \begin{cases} \tau - 1, & \Delta \geq T_{-1}(N_0), \\ \tau, & T_0(N_0) \leq \Delta < T_{-1}(N_0), \\ \tau + 1, & T_1(N_0) \leq \Delta < T_0(N_0), \\ \textit{no acceptance}, & 0 < \Delta < T_1(N_0). \end{cases} \tag{S8}$$

No first acceptance occurs at $\tau - 2$, $\tau + 2$ or a later year.

Proof Table S2 evaluates the ingredients at every offset where the contrast is positive. Since every level is at least $N_0 \geq 20$, the count floor holds throughout. The pre-period median also exceeds one, so the denominator of the level ratio is that median itself, without activation of its lower bound.

For $d \leq -3$, post-period growth and the gap are zero, so both positive growth floors fail. At $d = -2$, the post-median equals the pre-median and only one change is positive. Both the ratio and positivity requirements fail, irrespective of Δ.

Table S3 Complete-rule thresholds, rounded to six decimals. The final row is a large-level limit, not a finite admissible count.

N_0	T_{-1}	T_0	T_1
20	0.389465	0.194732	0.133333
50	0.398908	0.199454	0.133333
100	0.402159	0.201080	0.134053
500	0.404800	0.202400	0.134933
$N_0 \to \infty$	0.405465	0.202733	0.135155

At $d = -1$, the positivity and count conditions hold. The two growth conditions require $\Delta \geq \max\{3c_1/2, 3c_2/2\}$. The level-ratio condition is

$$\frac{M_1}{N_0} \geq 1.5 \quad \Longleftrightarrow \quad \Delta \geq Q(N_0). \tag{S9}$$

For $N_0 \geq 20$, $Q(N_0) \geq \log(31/21) > 3c_1/2 > 3c_2/2$. Consequently, all five conditions hold exactly when $\Delta \geq T_{-1}$.

At $d = 0$, three changes are positive and the growth conditions require $\Delta \geq \max\{c_1, c_2\}$. The post-median is M_2, so the level ratio requires $\Delta \geq Q(N_0)/2$. Since $\log(31/21)/2 > c_1 > c_2$, that ratio is binding and the qualification threshold is T_0.

At $d = 1$, positivity and the count floor again hold. Post growth requires $\Delta \geq c_1$, the gap requires $\Delta \geq 4c_2/3$, and $M_3/N_0 \geq 1.5$ requires $\Delta \geq Q(N_0)/3$. Their conjunction gives T_1. The thresholds satisfy

$$T_1 \leq \frac{\log(1.5)}{3} < \frac{\log(31/21)}{2} \leq T_0 < T_{-1}. \tag{S10}$$

The first inequality holds because each term defining T_1 is at most $\log(1.5)/3$. The ordering gives the first three branches of Eq. (S8) under chronological evaluation.

At $d = 2$, a necessary growth-gap condition is $\Delta \geq 2c_2 = 0.2$. But $T_1 \leq \log(1.5)/3 < 0.2$, so a candidate meeting this necessary condition has already qualified at $d = 1$, whatever additional restriction its own median ratio imposes. At $d = 3$, the necessary gap condition is still stronger, $\Delta \geq 4c_2 = 0.4$, and the same reasoning applies. No later first qualification can arise from these offsets. For $d \geq 4$, the contrast is zero and fails the positive gap threshold. Thus there is no acceptance anywhere when $0 < \Delta < T_1$, completing the proof. □

S2.2 Binding thresholds and the growth-only comparison

The growth floor c_1 never binds T_1 because $4c_2/3 > c_1$. Their remaining branches meet when $Q(N_0) = 4c_2$, at

$$N_0^\star = \frac{e^{4c_2} - 1}{1.5 - e^{4c_2}} \simeq 60.1598. \tag{S11}$$

For $20 \leq N_0 < N_0^\star$, the gap floor determines T_1; above that value the level ratio does, with equality at the boundary. Here the function $Q(N_0)$ increases to $\log(1.5)$, but does not equal that limit at finite N_0. Keeping the pseudocount is therefore necessary for exact boundary calculations. Table S3 reports selected finite-level values and their limit.

The two growth criteria alone have a different first-qualification ladder. On the same zero-growth baseline, their thresholds at $d = -2, -1, 0$ are $3c_1$, $3c_1/2$ and c_1, since $c_1 > c_2$. Thus the first candidate is $\tau - 2$ for $\Delta \geq 3c_1$, $\tau - 1$ for $3c_1/2 \leq \Delta < 3c_1$,

and τ for $c_1 \leq \Delta < 3c_1/2$. Below c_1, even post-period growth after the change is insufficient. If the constant pre-change growth rate is instead at least c_1, its post-growth condition always holds and the same three regions are governed by c_2 in place of c_1. Both statements concern the two growth conditions, not the full rule.

The stored c_1 is about 7.0×10^{-12} below $\log(1.12)$. The difference changes none of the complete-rule boundaries above because c_1 is dominated wherever it enters. It should nevertheless not be silently replaced in the implementation. The proof establishes exact real-arithmetic boundaries; testing a floating-point implementation at an equality boundary requires attention to numerical representation.

Calendar eligibility and chronological processing are part of the result. If an otherwise qualifying early candidate is unavailable or excluded by an inherited cooldown, a later first acceptance need not follow the displayed ladder. The pointwise filter identity remains valid wherever its input block exists. Conversely, the no-acceptance branch of the exact permanent-step proof concerns the whole specified path, whereas no observed or perturbed label in a bounded search window means only non-acceptance within that window.

S3 Preserved aggregation and rank results

S3.1 A pooled between-year share is not a path identifier

Let E be a finite set of observed series-year growth cells. For each represented year, let $n_t > 0$ be its number of cells and $\bar{z}_t$ its mean; let $\bar{z}_E$ be the mean across all cells, weighted equally. Define

$$S_{\mathrm{yr}} = \frac{\sum_t n_t(\bar{z}_t - \bar{z}_E)^2}{\sum_{(i,t)\in E}(z_{it} - \bar{z}_E)^2}, \tag{S12}$$

only when the denominator is positive. Otherwise the statistic is undefined. The numerator is a row-weighted between-year sum of squares. Decomposing $z_{it} - \bar{z}_E = (z_{it} - \bar{z}_t) + (\bar{z}_t - \bar{z}_E)$ and summing eliminates the cross term because $\sum_{i:(i,t)\in E}(z_{it} - \bar{z}_t) = 0$. Consequently the total sum of squares equals its within-year and between-year parts, and $0 \leq S_{\mathrm{yr}} \leq 1$ when defined. The accounting identity is not a factor-identification result.

Proposition S3.1 (An explicit scalar collision) *The map from a common time path to S_{yr} is not injective, even for a balanced panel with two series and four years.*

Proof Take $z_{it} = \theta_t + u_{it}$, with $u_{1t} = 1$ and $u_{2t} = -1$. Compare

$$\theta^A = (1, -1, 0, 0), \qquad \theta^B = (1/\sqrt{2}, 1/\sqrt{2}, -1/\sqrt{2}, -1/\sqrt{2}). \tag{S13}$$

Both have zero mean and squared norm two. Their population variances are therefore $1/2$. Each annual cross-sectional mean equals θ_t, so the numerator in Eq. (S12) is $2 \times 2 = 4$. The within-year sum of squares is $4(1^2 + (-1)^2) = 8$. For both panels, $S_{\mathrm{yr}} = 4/(4 + 8) = 1/3$, although their temporal paths differ. □

The collision concerns the specified variance share, not every possible encoding into a scalar. The statistic discards temporal ordering and does not locate an episode. No

estimated common-factor share is produced by the example. Its four growth observations are also insufficient for the seven growth intervals required by the worked contrast at one candidate.

An explicit longer embedding shows why the scalar cannot replace the filtered path. Index the four entries in Eq. (S13) by $t = 1, \ldots, 4$ and set $\theta_t = 0$ outside that range. At candidate $k = 3$, the pre-period growth inputs are indexed by $-1, 0, 1, 2$ and the post-period inputs by $3, 4, 5$. Substitution gives

$$D^{(3,4)}(\theta^A)_3 = 0, \qquad D^{(3,4)}(\theta^B)_3 = -\frac{7}{6\sqrt{2}}. \tag{S14}$$

The constructed extension is an evaluation of the two paths, not an additional patent dataset. The four-year collision alone proves the stated non-injectivity. Changing the scalar or restricting the admissible paths would define another identification question.

S3.2 Within-year ranks under homogeneous loading

Retain the conditional representation $z_{it} = \alpha_i + \lambda_i \theta_t + u_{it}$ from Eq. (S4). No fitted decomposition is assumed.

Proposition S3.2 (Translation invariance of ranks) *If $\lambda_i \equiv \lambda$ for all compared series, the within-year ordering of z_{it} equals that of $\alpha_i + u_{it}$, including ties. For a fixed candidate year and identical complete windows, the ordering of $D_{ik}^{(p,q)}$ equals that of $D^{(p,q)}(u_i)_k$.*

Proof For fixed t, the common contribution $\lambda\theta_t$ is the same constant for every series. Pairwise differences and ties are unchanged by adding it. For fixed k, Eq. (S4) adds the same value $\lambda D^{(p,q)}(\theta)_k$ to every filtered residual, and the same argument applies. The series-specific constant α_i has already been removed by the filter. □

With heterogeneous loadings, a pairwise difference instead contains $(\lambda_i - \lambda_j)\theta_t$, which can change the ordering. Nor does a within-year result assert that pooled ranks across different years are unchanged: their common additions can differ. Translation invariance also does not imply invariance of a rule with fixed numerical thresholds, medians and a count floor. Ranking and binary acceptance remain separate operations.

S3.3 The additive-Gaussian population benchmark

A different calculation concerns a randomly sampled cell from a specified population, rather than a ranking conditional on one year. Assume unit loading and mutually independent Gaussian variables α, u and θ, with variances σ_α^2, σ_u^2 and σ_θ^2. Let

$$\begin{aligned} g &= \alpha + u, \qquad z = g + \theta, \\ \sigma_g^2 &= \sigma_\alpha^2 + \sigma_u^2 > 0, \qquad \rho = \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_g^2}. \end{aligned} \tag{S15}$$

Means can be removed without changing correlations. The assumptions specify a Gaussian population; they do not describe an estimated distribution of EPO observations.

Proposition S3.3 (Pooled rank correlation in the Gaussian benchmark) *Under Eq.* (S15),

$$\mathrm{Corr}(z, g) = \sqrt{1-\rho}, \qquad \rho_S(z, g) = \frac{6}{\pi} \arcsin\left(\frac{\sqrt{1-\rho}}{2}\right). \tag{S16}$$

The population Spearman correlation decreases strictly from one as ρ increases on $[0, 1)$ and tends to zero as $\rho \to 1$.

Proof Independence gives $\mathrm{Cov}(z, g) = \sigma_g^2$ and $\mathrm{Var}(z) = \sigma_g^2 + \sigma_\theta^2$. Division by the two standard deviations yields the Pearson correlation $r = \sqrt{1-\rho}$. Let (Z, G) be the standardised pair, with marginal distribution functions F_Z and F_G. For continuous margins the two probability transforms are uniform on $[0, 1]$, each with variance $1/12$, and hence

$$\rho_S = 12\,\mathbb{E}[F_Z(Z)F_G(G)] - 3. \tag{S17}$$

Take Z' and G'' independently from their respective margins and independently of (Z, G). Conditional independence then gives

$$\mathbb{E}[F_Z(Z)F_G(G)] = \Pr(Z' < Z,\ G'' < G). \tag{S18}$$

The normal pair $(Z - Z', G - G'')$ has variances two and covariance r, so its correlation is $r/2$.

For a standard bivariate-normal pair (X, Y) with correlation v, write $X = W_1$ and $Y = vW_1 + \sqrt{1-v^2}W_2$, where W_1, W_2 are independent standard normals. Their polar angle is uniform. The angular regions satisfying $X > 0$ and $Y > 0$ overlap over an angle $\pi/2 + \arcsin(v)$, and therefore

$$\Pr(X > 0, Y > 0) = \frac{1}{4} + \frac{\arcsin(v)}{2\pi}. \tag{S19}$$

Using $v = r/2$ in Eqs. (S18) and (S17) gives $\rho_S = (6/\pi)\arcsin(r/2)$, which proves the stated expression. Finally, arcsin is strictly increasing and $\sqrt{1-\rho}$ strictly decreasing on the specified domain. At $\rho = 0$ the expression is one; its limit as $\rho \to 1$ is zero. □

There is no conflict with within-year rank invariance: that proposition conditions on a fixed year, whereas this correlation pools over a population in which the common term varies. Heterogeneous loadings generally yield a mixture across series rather than this single bivariate-normal population. The expression must not be transferred to that setting without a separate argument. Neither ρ nor the correlation is automatically equal to a finite-panel between-year share, and none supplies an accuracy threshold for the take-off classifier. The results preserve the original manuscript's aggregation qualifications without making them premises of the empirical dating analysis.

S4 Data provenance and analysis populations

The empirical application uses annual European patent-application activity, reported by applicant geography and technology field in the EPO Data Desk release of 23 March 2026. Its extraction date is 2 February 2026 and its calendar covers 2011–2025 (European Patent Office 2026). All 35 named technology fields were acquired through

the public report's ordinary export function. No underlying application-record export was available. The resulting data therefore support an application-count analysis, not deduplication of patent families or identification of the dates at which inventions were made.

Direct European applications and PCT applications entering the European phase are included. For the latter, fulfilment of the regional-phase conditions determines the reporting year; it need not be the priority year or the year in which the international application was originally filed (European Patent Office n.d.a). Geographic origin follows the residence of the first-named applicant, rather than inventor residence or a fractional allocation among co-applicants (European Patent Office n.d.b). The 2025 column is a complete reporting-calendar column at the stated extraction date, not a year-to-date observation. Completeness of that column does not imply that subsequent EPO releases cannot revise its values.

Figure S1 shows three accepted series chosen by the rule stated in its caption. Each field export contains 181 source labels and 15 annual values per label. Three report aggregates, All countries, EPO member states and Rest of the World, were excluded in the export filter. Within the exported labels, EU27 countries is an overlapping aggregate; Unknown and Not Applicable are retained as separate provenance categories, not treated as countries. Netherlands Antilles, Serbia And Montenegro and U.S.S.R. are preserved as historical geographic labels outside the primary population. Adding these three labels gives the 178-geography comparison below. No historical count is allocated to a successor state. Other territory labels and the original name Swaziland remain in the primary roster. The roster defines 175 reporting geographies, not 175 sovereign states with fully harmonised borders.

Table S4 Source and analysis populations. A series is a geography–technology pair; a candidate is one complete eight-level window in that series.

Population or record type	Number
Technology fields	35
Exported source labels per field	181
Original field–label–year count cells	95,025
Primary geography labels	175
Primary geography–technology series	6,125
Primary annual count cells, 2011–2025	91,875
Primary series with zero counts in every year	2,971
Calendar-eligible candidates, 2016–2023	49,000
Candidates with pre-period median at least 20	5,298
Candidates satisfying all five conditions before cooldown	268
Accepted candidates after cooldown	201
Accepted candidates within the pre-median-at-least-20 subset	112

Table S4 distinguishes source cells, series and candidate windows. All 91,875 primary cells have explicit nonnegative integer counts, unique geography–technology–year keys and complete calendars. Zeros are preserved; no absent cell was converted to zero. The 201 accepted candidates belong to 201 different series, 37 geographies and 34

technology fields. Of the 268 otherwise qualifying candidates, 67 are excluded by cooldown. Calendar eligibility is not conditional on any growth or scale threshold, so the 49,000-window denominator includes all-zero and small-count series. The alternative denominator of 5,298 uses only the pre-period median. Neither denominator is a count of independent statistical trials.

The normalized panel carries source-file SHA-256 identifiers, and the supplied preparation code describes the source roster and transformations. Original EPO field workbooks, ingestion and reconciliation manifests, the dated preparation protocol, and the alternative input panels are absent from this replication distribution. Consequently, the original export-to-panel chain cannot be reconstructed from these files alone. The archived primary panel does support independent recalculation of its candidate and event outputs. Historical-label inputs remain recorded preparation results, distinct from independently reproduced source-ingestion checks. No public preregistration is claimed.

An independent check against the current official field–country workbook (European Patent Office n.d.b) reproduces 16,270 matches and exactly the five recorded discrepancies among 16,275 overlapping cells for 2016–2025. The retrieved workbook's SHA-256 digest is recorded in the replication manifest. This corroborates the overlapping cells and their recorded differences; it does not recover the original 35 field exports or verify the 2011–2015 cells against those exports. Independent reconstruction of the five-cell substitution likewise leaves all 268 qualifications and 201 accepted events unchanged, while changing numerical components in seven windows.

S5 Specification sensitivity and the local segmentation comparator

The baseline uses five pre-period levels, $k-5, \ldots, k-1$, and three post-period levels, $k, \ldots, k+2$. Its conditions are post-growth at least the stored constant 0.1133286853, growth acceleration at least 0.1, post-to-pre median ratio at least 1.5 with the pre denominator floored at one, at least two strictly positive anchored count changes, and post-period median at least 20. Acceptance prevents another acceptance for five subsequent candidate years. Every scan starts in 2016 with no earlier accepted state; pre-2016 acceptance histories are unobserved.

Alternative definitions change one setting at a time. Write p for the number of post-growth intervals and q for the number of pre-growth intervals. A changed window uses $q+1$ pre-period levels and p post-period levels, requiring $\lceil 2p/3 \rceil$ positive anchored changes. The scale floor remains 20. All definitions in Table S5 are compared on 2016–2022, from the same initial state. The common calendar contains 42,875 candidate windows and 192 baseline events. The complete baseline calendar extends to 2023 and contains another nine accepted events; comparing its total of 201 directly with the common-calendar alternatives would mix different coverage.

Jaccard agreement is $|E \cap B|/|E \cup B|$; it is undefined if both sets are empty. None of the displayed comparisons has an empty union. Changes in event counts alone do not identify the extent of redating: for example, the ratio-1.25 specification has 280 events but retains only 166 baseline event keys. Cooldown can replace a later baseline

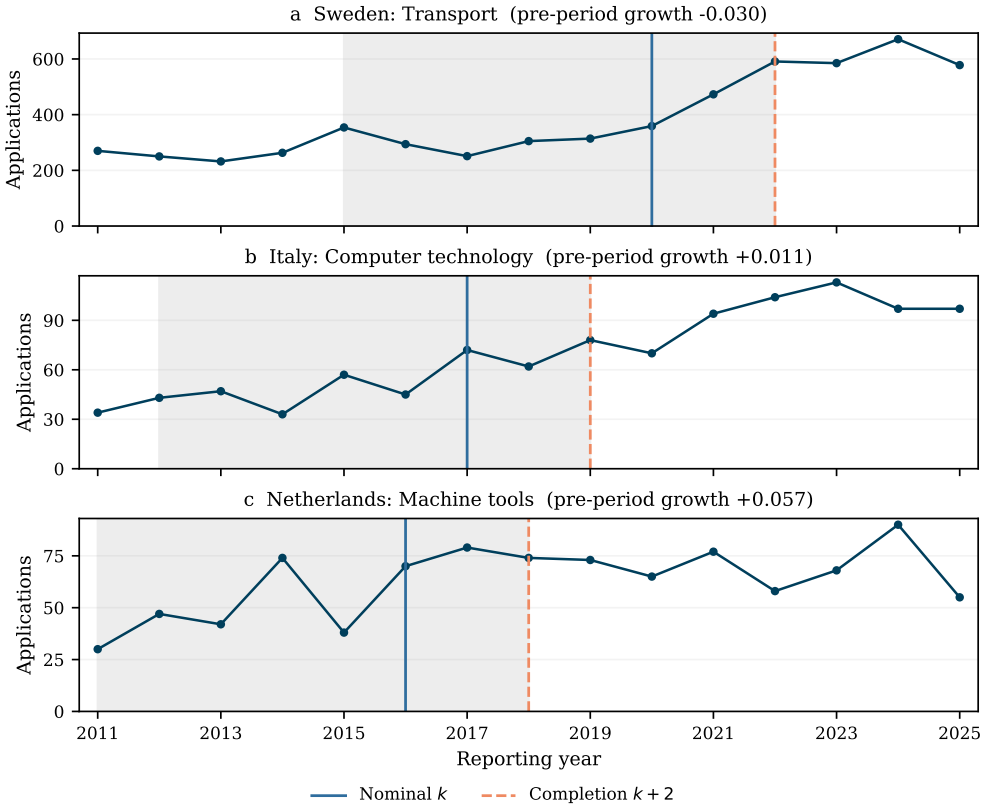


Fig. S1 Three accepted events at the lower quartile, median and upper quartile of pre-period growth. Selection rule, fixed before the plots were seen: among the 112 accepted events with a pre-period median of at least 20, order by pre-period growth with ties broken by field name and year, take relative positions 0.25, 0.50 and 0.75, and move to the next position if a geography would appear twice; the rule returns positions 27, 55 and 83. The shaded interval is the complete candidate window, not a confidence band. Vertical lines distinguish the nominal year from window completion. Source: current EPO Data Desk export; country names and technology categories follow the source.

acceptance with an earlier acceptance after a threshold is relaxed. Such changes describe dependence on the operational definition, not superiority of one setting.

For the local segmentation comparator, the seven annual log-growth rates within each accepted event's eight-level window are split after two, three, four or five observations. Each split receives two fitted constant means; unweighted residual sum of squares is minimised and exact ties select the earliest split. Its reference year s is the first year of the second segment, hence $s \in \{k-2, k-1, k, k+1\}$. The recorded differences $k-s$ are -1 for 33 events, zero for 87, one for 29 and two for 52, accounting for all 201 events. Sixteen fitted mean shifts are negative, despite the complete rule's positive-growth requirements. Selection of the window by the rule and the restricted split range prevent interpreting this comparison as an independent test against technological onset.

The observed-path illustration uses the restricted 112-event selection specified in Figure S1. It does not use quartiles of all 201 events. The complete event records

Table S5 One-setting specification sensitivity, common candidate calendar 2016–2022. E is the alternative event set and B the 192-event baseline set. Event identity requires the same geography, technology and year.

Definition	$\lvert E\rvert$	$\lvert E \cap B\rvert$	$\lvert E \cup B\rvert$	Jaccard
Baseline	192	192	192	1.000
Acceleration threshold 0.075	205	185	212	0.873
Acceleration threshold 0.125	171	168	195	0.862
Median ratio 1.25	280	166	306	0.542
Median ratio 1.75	129	109	212	0.514
Post-growth threshold log(1.08)	192	192	192	1.000
Post-growth threshold log(1.16)	177	177	192	0.922
No cooldown	253	192	253	0.759
Three-year cooldown	194	192	194	0.990
Growth windows $(p, q) = (3, 3)$	187	141	238	0.592
Growth windows $(p, q) = (4, 4)$	138	73	257	0.284

accompany the selected examples; the plotting selection does not define the analysis population.

S6 Controlled growth changes on observed backgrounds

The experiment begins with every complete geography–technology block $\tau-7, \ldots, \tau+5$. The 2011–2025 calendar permits three anchors, $\tau \in \{2018, 2019, 2020\}$, giving 18,375 potential pairs. Eligibility requires both the median over $\tau-5, \ldots, \tau-1$ and the count at $\tau-1$ to be at least 20. A pair is excluded if the observed path already satisfies all five conditions anywhere in $\tau-2, \ldots, \tau+3$, or has a previously accepted event in $\tau-7, \ldots, \tau-3$. The historical check is restricted to the available scan beginning in 2016.

Table S6 Sequential selection of perturbation pairs. Exclusions are applied from left to right and are mutually exclusive in this accounting. Each anchor begins with 6,125 complete blocks.

Anchor	Complete	Low scale	Prior local qualification	Earlier accepted event	Retained
2018	6,125	5,494	107	0	524
2019	6,125	5,479	95	29	522
2020	6,125	5,476	86	50	513
Total	18,375	16,449	288	79	1,559

The 1,559 retained pairs cover 547 distinct series, 32 geographies and all 35 technology fields. Unconditional exclusion flags overlap; their totals must not be added

as though they partitioned the population. The sequential accounting in Table S6 does partition it. The retained counts for 2018, 2019 and 2020 cover 31, 31 and 30 geographies, respectively, and all 35 fields in each year.

Each retained block supplies three backgrounds: the observed log-count path, a constant path at $\ell_{i,\tau-1}$, and a linear log-count path anchored there with slope equal to mean growth over the four preceding intervals. Fitted negative log levels would be floored at zero, but no flooring occurs in the retained blocks. For background b_t, the construction is

$$\widetilde{\ell}_t = b_t + \delta \max\{t - \tau + 1, 0\}, \qquad \widetilde{N}_t = \exp(\widetilde{\ell}_t) - 1, \tag{S20}$$

with $\delta \in \{0, 0.10, 0.15, 0.20, 0.30, 0.40\}$. No random noise is generated. What is added is a permanent log-growth increment on a specified background, not another observed patent dataset. Inverse-transformed levels can be fractional; they are not rounded into fictitious observed application counts. The design yields $1,559 \times 3 \times 6 = 28,062$ constructed paths.

Local scans begin without an active cooldown and examine candidate offsets $d = k - \tau = -2, \ldots, +3$. The outcome is the first accepted offset or no local label by $+3$. Negative offsets describe retrospective assignment to an earlier nominal year, not information available before the imposed change. For example, the window of candidate $\tau - 2$ closes at τ. No label by $+3$ does not establish the absence of a later label.

The complete distributions in Table S7 retain every candidate offset. Constant-background dates satisfy the first-acceptance ladder in Section S2. Dependence on the anchor count remains: the median-ratio threshold is $Q(N_0) = \log\{(1.5N_0+1)/(N_0+1)\}$, so equal growth increments need not give identical dates across different constant levels. Observed backgrounds yield a wider distribution of offsets. At $\delta = 0.20$, for example, 297 pairs receive an earlier nominal date, 581 the imposed year, 568 a later date, and 113 no local label. All of them are conditional finite-panel frequencies, not estimates of an error distribution against an observed technological onset.

Figure S2 compares the three backgrounds at each increment. Within exactly the same population, a component comparison retains only post-growth and acceleration conditions, with the original thresholds and candidate grid. The first complete-rule acceptance cannot precede the growth-only acceptance: at each candidate the complete qualification set is a subset of the two-condition set, and both scans begin without an active cooldown. All 28,062 recorded comparisons obey this ordering. The statement concerns first local acceptance, not an ordering of subsequent cooldown-constrained event sequences.

Table S8 gives the paired component counts. At zero increment, selection already excludes complete-rule labels on the observed path, whereas the two growth conditions alone label 416 of the 1,559 pairs. The difference is a consequence of the selection and qualification definitions; it is not a false-positive comparison. The additional level, positivity and scale conditions can delay or suppress a local label, but technological accuracy is not observed. Overlapping anchors within the same series are not treated as independent trials.

Threshold equality is evaluated according to the rule's weak inequalities, not by widening a threshold with an empirical tolerance. For the constant and unfloored

Table S7 First accepted offsets under the complete rule. Every row has denominator 1,559 pairs. Zero cells denote observed zero frequencies; "None" means no local label by +3, not no eventual label.

Background	δ	-2	-1	0	$+1$	$+2$	$+3$	None
Constant	0.00	0	0	0	0	0	0	1,559
Constant	0.10	0	0	0	0	0	0	1,559
Constant	0.15	0	0	0	1,559	0	0	0
Constant	0.20	0	0	463	1,096	0	0	0
Constant	0.30	0	0	1,559	0	0	0	0
Constant	0.40	0	463	1,096	0	0	0	0
Observed	0.00	0	0	0	0	0	0	1,559
Observed	0.10	23	67	163	151	143	112	900
Observed	0.15	35	138	343	305	203	125	410
Observed	0.20	48	249	581	331	166	71	113
Observed	0.30	81	552	698	160	42	20	6
Observed	0.40	112	885	490	62	6	3	1
Trend	0.00	0	0	0	0	0	0	1,559
Trend	0.10	0	0	618	0	0	0	941
Trend	0.15	0	525	319	305	0	0	410
Trend	0.20	0	619	454	275	78	0	133
Trend	0.30	260	584	535	118	44	0	18
Trend	0.40	260	813	423	56	3	1	3

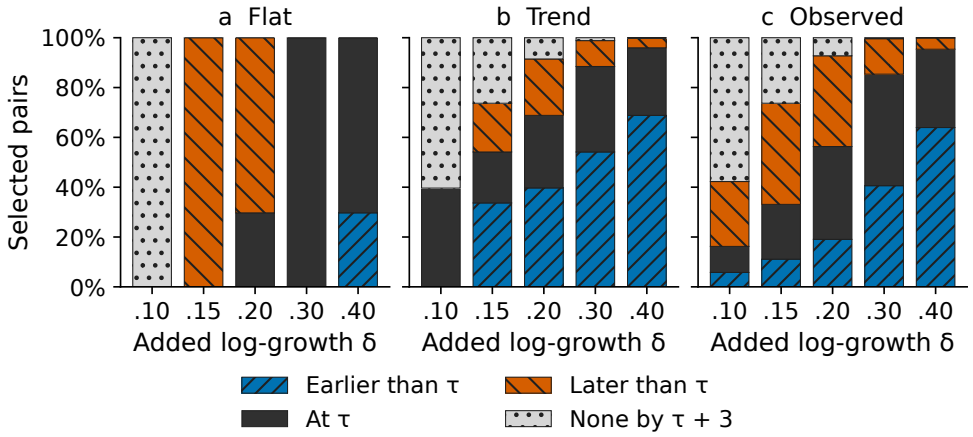


Fig. S2 First full-rule dates after constructed permanent growth increments. Each bar contains the same 1,559 selected series–anchor pairs. Earlier and later refer to the prescribed change year τ, not to an observed onset. The linear-trend and flat backgrounds are constructed comparators. The omitted zero-increment condition has no full-rule labels in any background; for the observed background that outcome follows from sample selection and is not a specificity estimate.

Table S8 Complete rule compared with its two growth conditions. Each row contains 1,559 paired comparisons. "Later" requires both dates to be finite. "Suppressed" means the growth-only rule labels locally and the complete rule does not. "Neither" refers only to the grid through +3.

Background	δ	Same date	Later	Suppressed	Neither
Constant	0.00	0	0	0	1,559
Constant	0.10	0	0	0	1,559
Constant	0.15	0	1,559	0	0
Constant	0.20	0	1,559	0	0
Constant	0.30	0	1,559	0	0
Constant	0.40	0	1,559	0	0
Observed	0.00	0	0	416	1,143
Observed	0.10	238	421	533	367
Observed	0.15	384	765	273	137
Observed	0.20	474	972	80	33
Observed	0.30	513	1,040	6	0
Observed	0.40	515	1,043	1	0
Trend	0.00	0	0	0	1,559
Trend	0.10	618	0	332	609
Trend	0.15	525	624	181	229
Trend	0.20	619	807	45	88
Trend	0.30	260	1,281	10	8
Trend	0.40	260	1,296	0	3

linear backgrounds, the background contrast is exactly zero. The prescribed decimal increments and the rational coefficients of ψ are therefore represented as fractions, so the comparison $\delta\psi(d) \geq 0.1$ is exact. Unmodified observed counts are retained as integers rather than recovered through an exponential round trip. The fitted trend passes through both observed support points at $\tau - 5$ and $\tau - 1$; their unmodified counts are likewise preserved exactly, which prevents an exact median-ratio or scale equality from being changed by numerical reconstruction.

For observed backgrounds, a binary-arithmetic distance within 10^{-12} of the acceleration threshold selects a higher-precision calculation but does not determine its pass state. With

$$R = \frac{(1 + N_{i,k+2})^4 (1 + N_{i,k-5})^3}{(1 + N_{i,k-1})^7}, \qquad D_{ik} = \frac{\log R}{12}, \tag{S21}$$

the required background contrast is compared with the exact rational remainder $0.1 - \delta\psi(d)$. If that remainder is zero, integer comparison of the numerator and denominator resolves the inequality exactly. Otherwise, decimal logarithms at 80 and 120 digits must give the same decision. A separate reference calculation evaluates all constructed outcomes at both precisions while preserving the rational contrast and exact trend support points. No unresolved precision disagreement remains between those reference calculations. The procedures implement the stated mathematical inequalities;

they do not create another acceptance criterion or permit an arbitrary choice between conflicting labels.

S7 Comparison of source versions and geographic definitions

The separate official 2016–2025 field and field-by-country workbooks provide a corroborating comparison, not an extension of the primary calendar (European Patent Office n.d.b). Their year sheets do not all share one extraction date. The field-by-country sheets for 2016–2018 state 27 January 2020; those for 2019, 2020, 2021 and 2022 state 1 February 2021, 24 January 2022, 30 January 2023 and 29 January 2024, respectively. Both 2023 and 2024 sheets state 3 February 2025, while the 2025 sheet states 2 February 2026. Consequently, a disagreement cannot be presumed to be a failed extraction from the current report.

Across 16,275 directly comparable individual-country cells for the 35 named fields in 2016–2025, five values differ (Table S9). The workbook's Other countries category is not an individually comparable country and is not included in that denominator. Neither these comparisons nor the corroborating workbooks independently verify 2011–2015.

Table S9 All differing individually comparable source cells. All refer to 2024. The workbook retains its own status date of 3 February 2025; Data Desk is the primary source.

Technology	Geography	Workbook	Data Desk
Civil engineering	Australia	6	46
Civil engineering	Canada	10	37
Civil engineering	Israel	13	25
Civil engineering	Taiwan, Province of China	24	30
Other consumer goods	Taiwan, Province of China	25	27

A separate diagnostic copy replaces exactly these five primary cells with the workbook values. It changes numerical components in seven of the 49,000 common candidate windows: 2022 Civil engineering for Australia and Canada; 2022 and 2023 Civil engineering for Israel and Taiwan, Province of China; and 2022 Other consumer goods for Taiwan, Province of China. Individual condition flags change in three of those windows. In all seven, complete qualification and acceptance remain false in both sources. The accepted event set is unchanged, although its invariance does not imply invariance of the component statistics.

Table S10 separates the two source diagnostics. Including the three historical labels adds 105 series and 840 complete candidate windows, giving 49,840 candidates in total. No accepted event is added or removed. The result does not resolve the historical allocation of their source counts; it only describes the effect of their inclusion under the fixed rule.

Table S10 Sensitivity to source-cell replacement and historical geographic labels, baseline calendar 2016–2023. Changes are relative to the primary definition. Metric changes concern the common primary candidate windows; newly added series are not counted as changed common windows.

Definition	Series	Events	Changed metrics	Changed complete decisions
Primary 175 geographies	6,125	201	n.a.	n.a.
Including three historical labels	6,230	201	0	0
Five-cell replacement diagnostic	6,125	201	7	0

Aggregate comparisons require a different denominator. Among 350 field–year comparisons in 2016–2025, the sum over all 180 exported labels other than EU27 exceeds the official world field total in 39 cases, by one to three applications. Removing Not Applicable alone would arithmetically reconcile 38 of those cases; a one-application difference remains for Electrical machinery, apparatus, energy in 2017. Such arithmetic does not establish the EPO's production definition of Not Applicable. The primary 175-geography sum is below the official world total in 79 field–years, by one to eight applications, because its population excludes non-geographic and historical labels and the sources also differ. Aggregate residuals are preserved, not assigned to a chosen country or interpreted as missing primary rows.

The five-cell replacement is deliberately a hybrid diagnostic. It neither reconstructs a complete historical vintage nor establishes which conflicting value is correct. The supplied primary outputs and recorded alternative-output diagnostics are kept distinct from the newly reconstructed five-cell comparison; the original complete source reconciliation is not included. Agreement of their event sets is therefore a bounded finding about this rule on these recorded versions, not a general assurance against future source revisions.

S8 Proof of Proposition 1 and of Corollary 1

Throughout, $A = e^L$ with $L = \ell_{\tau-1}$, $g \geq 0$, $\delta > 0$, $A-1 \geq 20$, with non-negative counts throughout all required finite windows, and the path is $\ell_t = L + g\, s_t + \delta \max\{s_t, 0\}$ with $s_t = t - \tau + 1$, so that $N_t = Ae^{g s_t + \delta \max\{s_t,0\}} - 1$. Because $g \geq 0$ and $\delta > 0$, ℓ_t is non-decreasing in t, and it is strictly increasing for $t \geq \tau$. Fix an offset $d \in \{-2, \ldots, 3\}$ and the candidate $k = \tau + d$.

Medians. The pre-window $\{k-5, \ldots, k-1\}$ and the post-window $\{k, k+1, k+2\}$ each contain an odd number of consecutive observations of a monotone path, so their medians are the middle-time observations, $m^{\text{pre}} = N_{k-3}$ and $m^{\text{post}} = N_{k+1}$, with $s_{k-3} = d-2$ and $s_{k+1} = d+2$. Hence $m^{\text{post}} = Ae^{g(d+2)+\delta \max\{d+2,0\}} - 1$ and $m^{\text{pre}} = Ae^{g(d-2)+\delta \max\{d-2,0\}} - 1$. For $d \in \{-2, -1, 0, 1\}$ the pre-median carries no step and the post-median carries $(d+2)$ years of it, which gives the level conditions in Eqs. (5)–(8) of the main text.

Growth conditions. By the contrast identity, $D_k = \delta\psi(d)$ exactly, since the constant slope g contributes $g - g = 0$; the gap conditions are $\delta\psi(d) \geq c_2$, that is $\delta \geq 3c_2$, $\frac{3}{2}c_2$, c_2, $\frac{4}{3}c_2$ for $d = -2, -1, 0, 1$. The post-period rate is the mean of the three growth increments ending at $k, k+1, k+2$; of these, $\max\{0, \min\{d+3, 3\}\}$ carry the step, so $r^{\text{post}} = g + \delta[d+3]_0^3/3$.

Positivity. The three increments ending at k, $k+1$, $k+2$ equal g for intervals before τ and $g + \delta$ from τ onward. For $d \geq -1$ at least two of them contain τ or later, so at least two are $g + \delta > 0$. For $d = -2$ the increments are g, g, $g + \delta$; at least two are strictly positive if and only if $g > 0$.

Floor. Since the path is non-decreasing and $N_{\tau-1} = A - 1 \geq 20$, every post-median from $d = -2$ onward is at least $A - 1 \geq 20$.

Collecting the four items yields Eqs. (5)–(8). For $d \geq 2$ the gap condition requires $\delta \geq 2c_2$ ($d = 2$) or $\delta \geq 4c_2$ ($d = 3$). Either implies E_1: the gap at $d = 1$ needs $\delta \geq \frac{4}{3}c_2 < 2c_2$; the growth condition $g + \delta \geq c_1$ holds because $2c_2 = 0.2 > c_1$; the level condition at $d = 1$ holds because $Ae^{3g+3\delta} - 1 \geq Ae^{0.6} - 1 > 1.5A - 1.5 \geq 1.5\max\{Ae^{-g} - 1, 1\}$ for $A \geq 21$; the floor holds by monotonicity. Hence no first acceptance occurs at $\tau + 2$ or later, and chronological evaluation without inherited cooldown gives $k^* = \tau + \min\{d : E_d\}$. For declining backgrounds ($g < 0$) the medians at $d \in \{-1, 0\}$ remain the middle-time observations, because the pre-window then lies entirely before τ and is monotone, and the post-window lies entirely from $\tau - 1$ onward with constant increment $g + \delta$; the two conditions therefore hold as stated whenever $g + \delta > 0$, with the floor conditions now binding. For $d = -2$ and $d \geq 1$ on a declining background the windows straddle τ and the median index depends on δ; those cases are evaluated numerically (Fig. 2 of the main text) and are not part of the proposition. □

Corollary. Letting $A \to \infty$ in the level conditions gives $e^{g(d+2)+\delta(d+2)} \geq 1.5e^{g(d-2)}$, that is $4g + (d+2)\delta \geq \log 1.5$, for $d \in \{-1, 0, 1\}$, and $e^{4g} \geq 1.5$ for $d = -2$. Combining with the gap and growth conditions yields the thresholds in Eq. (9) and the crossover values $(\log 1.5 - 2c_2)/4$, $(\log 1.5 - 1.5c_2)/4$ and $\log 1.5/4$. □

Verification. In the controlled experiment, each trend is anchored to the observed count at $\tau - 1$, with slope derived from the observed endpoints at $\tau - 5$ and $\tau - 1$. Those integer endpoints are mathematically consistent with the defined exponential path; their restoration in numerical evaluation avoids roundoff at exact boundaries. Finite-count predictions agree in all 6,504 non-negative-slope cases, including the separate zero-increment checks. Within the positive-increment domain, the large-count corollary agrees in 5,316 of 5,420 cases (98.1 per cent). Declining-background cases are excluded from this comparison because the corollary does not cover their complete first-acceptance rule.

For illustration grids with arbitrary A and g, no integer endpoint restoration is justified. Figure 2 of the main text therefore evaluates the continuous path without rounding arbitrary support values. Independent direct-median calculations also agree with the finite-count predicates on 30,000 randomly sampled admissible non-negative-growth paths and with the two declining-background predicates on 10,000 paths. These numerical checks supplement the proof; they do not replace it.

Table S11 lists qualification thresholds at $A-1=100$ and their large-count limits. A qualification threshold at a later offset need not define a first-acceptance region when an earlier offset already qualifies.

Table S11 Qualification thresholds on constant-growth backgrounds at $A-1=100$: finite-count (with pseudocount, monotone case $g \geq 0$ only) and large-count form. A dash marks an offset at which acceptance is impossible or not covered by the exact statement.

g	$\tau-2$		$\tau-1$		τ		$\tau+1$	
	exact	large	exact	large	exact	large	exact	large
0.0000	–	–	0.4022	0.4055	0.2011	0.2027	0.1341	0.1352
0.0200	–	–	0.3220	0.3255	0.1610	0.1627	0.1333	0.1333
0.0500	–	–	0.2016	0.2055	0.1009	0.1027	0.1333	0.1333
0.0639	–	–	0.1500	0.1500	0.1000	0.1000	0.1333	0.1333
0.0800	–	–	0.1500	0.1500	0.1000	0.1000	0.1333	0.1333
0.1000	–	–	0.1500	0.1500	0.1000	0.1000	0.1333	0.1333
0.1014	0.3000	0.3000	0.1500	0.1500	0.1000	0.1000	0.1333	0.1333
0.1200	0.3000	0.3000	0.1500	0.1500	0.1000	0.1000	0.1333	0.1333

S9 Proof of Proposition 2 and an acceleration illustration

Let $z_t = a + c(t - t_0)$. By Eq. (2) of the main text, $D_k = \frac{1}{3}\sum_{h=0}^{2} z_{k+h} - \frac{1}{4}\sum_{h=1}^{4} z_{k-h}$. The first mean is $a + c(k + 1 - t_0)$ and the second is $a + c(k - 2.5 - t_0)$, so $D_k = c\,(k + 1 - k + 2.5) = 3.5c$ for every k. The gap condition $D_k \geq c_2$ therefore holds at every candidate when $c \geq c_2/3.5 = 0.02857\ldots$ and at none otherwise. The remaining conditions depend on the level path $\ell_t = \ell_{t_0} + a(t - t_0) + \frac{1}{2}c(t - t_0)(t - t_0 + 1)$ and on the first eligible candidate, not on any event. □

Illustration on a flat start ($a = 0$ at the first observation of a 2011–2025 calendar): for $c \in \{0.0286, 0.03, 0.05\}$ and levels 20, 100 and 500 the contrast is constant at $3.5c$ to machine precision and the first accepted nominal year is 2016, the first eligible candidate, in all nine cases; for $c = 0.02$ the contrast is $0.07 < c_2$ and no candidate is accepted at any level.

S10 Fluctuation bounds, variances and the dispersion check

Writing $\ell_t = b_t + \varepsilon_t$ and using $D_k = \frac{1}{3}(\ell_{k+2} - \ell_{k-1}) - \frac{1}{4}(\ell_{k-1} - \ell_{k-5})$, the deviation enters as $D_k(\varepsilon) = \frac{1}{3}\varepsilon_{k+2} - \frac{7}{12}\varepsilon_{k-1} + \frac{1}{4}\varepsilon_{k-5}$; the absolute coefficients sum to $\frac{1}{3} + \frac{7}{12} + \frac{1}{4} = \frac{7}{6}$, which is the bound in Eq. (10). The post rate $r_k^{\mathrm{post}} = \frac{1}{3}(\ell_{k+2} - \ell_{k-1})$ moves by at most $\frac{2}{3}\eta$. Under independent level deviations with variance σ_ℓ^2, $\operatorname{Var} D_k = \sigma_\ell^2(\frac{1}{9} + \frac{49}{144} + \frac{1}{16}) = \frac{37}{72}\sigma_\ell^2$. Under independent growth deviations u_t with variance σ^2, Eq. (2) gives $\operatorname{Var} D_k = \sigma^2(3 \cdot \frac{1}{9} + 4 \cdot \frac{1}{16}) = \frac{7}{12}\sigma^2$. The two are not interchangeable: differencing independent level deviations produces growth deviations with lag-one correlation $-\frac{1}{2}$.

The dispersion check used in Section 4.4 of the main text fits, for each of the 670 observed series with mean count of at least 20 over 2011–2025, a least-squares line to $\log(1 + N_t)$ against time, sets $\hat{\mu}_t = \exp(\hat{\ell}_t) - 1$, and computes the Pearson dispersion $\hat{\phi} = \sum_t (N_t - \hat{\mu}_t)^2 / \hat{\mu}_t / 13$ and the moment estimate $\hat{\alpha} = \sum_t \{(N_t - \hat{\mu}_t)^2 - \hat{\mu}_t\} / \sum_t \hat{\mu}_t^2$. The quantiles of $\hat{\phi}$ at 0.10, 0.25, 0.50, 0.75, 0.90 are 1.15, 1.62, 2.85, 5.74, 12.69, with $\hat{\phi} > 1$ in 93.7 per cent of series; the quantiles of $\hat{\alpha}$ are −0.0001, 0.0039, 0.0125, 0.0305, 0.0674, with $\hat{\alpha} \geq 0.10$ in 6.4 per cent. By mean count the median $\hat{\alpha}$ is 0.020 for means of 20–49 (208 series), 0.015 for 50–149 (223) and 0.009 for 150 and above (239). The fit absorbs any trend, so the estimate describes scatter around a log-linear path and not a residual process; it is used only to place the simulation parameters.

S11 Criterion margins, calendar composition and stratified outcomes

Relative slack is $v_j / h_j - 1$ for criterion value v_j and threshold h_j, among the two growth criteria, the median ratio and the count floor. The minimum identifies the tightest of these four continuous criteria. Positivity is discrete and reported separately. The term binding margin in the archived numerical outputs refers to this minimum; the normalisation is descriptive and does not show which condition determines the first accepted date. The main-text condition-removal table reports that distinct counterfactual comparison.

For each accepted event the distance from each threshold is $r^{\text{post}} - c_1$, $D - c_2$, ratio−1.5 and $m^{\text{post}} - 20$; the binding criterion is the one with the smallest distance divided by its threshold. Table 2 of the main text reports the quantiles. Further counts: 19 events have a gap within 0.02 of c_2; 78 events have negative pre-period growth; 48 events lie within 5 per cent and 87 within 10 per cent of their binding threshold. Among the 5,298 scale-eligible candidates, 157 satisfy all five conditions (112 accepted, 45 excluded by cooldown), 518 fail exactly one, 615 two, 1,868 three, 2,001 four and 139 all five. Of the 518 single failures, 301 fail the ratio (44 with a ratio of at least 1.4, 96 with at least 1.3), 206 the gap (20 with a gap of at least 0.08), 4 the post-growth rate, 7 positivity and none the floor.

Events by nominal year, 2016 to 2023: 48, 29, 29, 20, 25, 26, 15 and 9. Events by geography (top ten): China 20, Republic of Korea 14, Singapore 12, Denmark 11, Ireland 10, Belgium 9, Hong Kong SAR 9, Taiwan 8, and Australia, Canada, India, Norway and Türkiye 7 each; 37 geographies and 34 fields carry at least one event. Table S12 lists the fields in which the earliest nominal year or the set of earliest-labelled geographies changes under the $(4, 4)$ growth windows on the common 2016–2022 calendar; among the 33 fields with events under both definitions, the set changes in 24 and the year alone in two. In 23 of the 34 baseline fields the baseline's earliest event carries the nominal year 2016.

Table S13 gives the first-date outcomes on observed backgrounds by pre-period slope for every step size, Table S14 the outcomes by scale and by anchor year, and Table S15 the agreement of the observed-background date with the trend-background and constant-background dates of the same pair.

Table S12 Earliest-labelled geographies per field, baseline against $(4, 4)$ growth windows, common calendar 2016–2022. Names follow the source.

Field	Baseline	$(4, 4)$ windows
Audio-visual technology	Austria / Belgium (2016)	Switzerland (2018)
Basic communication processes	Austria / China, People's Republic of (2017)	Austria (2016)
Basic materials chemistry	Saudi Arabia (2018)	India (2018)
Chemical engineering	Australia (2017)	Australia (2018)
Civil engineering	China, People's Republic of / Israel (2016)	Türkiye (2016)
Computer technology	Barbados / Cayman Islands / Singapore (2016)	Austria / Barbados / Cayman Islands / Ireland / Singapore (2016)
Control	Canada / Taiwan, Province of China (2016)	Belgium / Canada (2016)
Digital communication	Cayman Islands / Italy / Türkiye (2016)	Cayman Islands / Singapore (2016)
Engines, pumps, turbines	Belgium / China, People's Republic of (2016)	Denmark (2016)
Food chemistry	Belgium / Israel (2016)	Israel (2016)
Handling	Australia / Norway (2018)	Norway (2018)
IT methods for management	Australia / Belgium / Ireland / Singapore (2016)	Cayman Islands / Germany / Ireland / Singapore (2016)
Machine tools	Greece / Netherlands (2016)	Greece (2016)
Materials, metallurgy	Australia / Denmark / Korea, Republic of / Liechtenstein (2017)	Australia (2016)
Mechanical elements	Belgium (2017)	Türkiye (2018)
Medical technology	India / Puerto Rico (2016)	Ireland / Portugal (2018)
Optics	Israel (2016)	Taiwan, Province of China (2017)
Organic fine chemistry	China, People's Republic of / Czech Republic / Korea, Republic of (2016)	China, People's Republic of (2017)
Other consumer goods	China, People's Republic of (2016)	United Kingdom (2016)
Other special machines	Taiwan, Province of China (2016)	Korea, Republic of (2016)
Pharmaceuticals	Singapore (2017)	Singapore (2016)
Semiconductors	China, People's Republic of (2016)	Austria (2016)
Surface technology, coating	Netherlands (2017)	Spain (2016)
Textile and paper machines	Austria / Sweden (2017)	Sweden (2017)
Thermal processes and apparatus	China, People's Republic of / Poland (2016)	China, People's Republic of (2016)
Transport	China, People's Republic of / Liechtenstein / Türkiye (2016)	China, People's Republic of / Israel / Liechtenstein (2016)

S12 Count simulation: all cells

Figure S3 plots the headline shares of this section.

Design as in Section 4.4 of the main text. Seeds: 20260908 for the 54 Poisson and $\alpha = 0.10$ cells and 20260909 for the 27 cells at $\alpha = 0.03$ added after the dispersion check, with separate generators so that the original cells keep their random streams; 2,000 sequences per cell; negative-binomial counts drawn as Poisson with gamma-distributed mean (shape $1/\alpha$, scale $\alpha\mu_t$). Table S16 reports the cells without a step and Table S17 the cells with a step, and Table S18 the latent AR(1) addition. In every cell, the full rule never accepts earlier than the growth-only rule on the same sequence, as set inclusion requires; the assertion is part of the code.

Table S13 First full-rule dates on observed backgrounds by pre-period slope class (fitted mean growth over $\tau - 5$ to $\tau - 1$; class boundaries fixed before the reading) and step size δ. Earlier and later refer to the prescribed change year τ; none means no local label by $\tau + 3$.

δ	Slope class	Pairs	Earlier	At τ	Later	None
0.10	below −0.05	178	2	19	64	93
0.10	−0.05 to 0	297	2	30	85	180
0.10	0 to 0.05	461	13	67	101	280
0.10	0.05 to 0.10	363	34	34	84	211
0.10	above 0.10	260	39	13	72	136
0.15	below −0.05	178	7	29	90	52
0.15	−0.05 to 0	297	4	71	137	85
0.15	0 to 0.05	461	29	149	173	110
0.15	0.05 to 0.10	363	66	70	131	96
0.15	above 0.10	260	67	24	102	67
0.20	below −0.05	178	7	54	94	23
0.20	−0.05 to 0	297	9	133	131	24
0.20	0 to 0.05	461	75	225	140	21
0.20	0.05 to 0.10	363	110	122	111	20
0.20	above 0.10	260	96	47	92	25
0.30	below −0.05	178	20	91	65	2
0.30	−0.05 to 0	297	59	192	45	1
0.30	0 to 0.05	461	189	233	38	1
0.30	0.05 to 0.10	363	202	127	33	1
0.30	above 0.10	260	163	55	41	1
0.40	below −0.05	178	42	110	26	0
0.40	−0.05 to 0	297	129	152	16	0
0.40	0 to 0.05	461	332	120	8	1
0.40	0.05 to 0.10	363	286	71	6	0
0.40	above 0.10	260	208	37	15	0

Table S14 First full-rule dates on observed backgrounds by pre-period median scale and by anchor year, for $\delta = 0.20$ and $\delta = 0.30$.

δ	Stratum	Pairs	Earlier	At τ	Later	None
0.20	20-29	180	39	55	69	17
0.20	30-59	313	62	113	119	19
0.20	60-149	451	89	155	177	30
0.20	150+	615	107	258	203	47
0.20	anchor 2018	524	122	211	153	38
0.20	anchor 2019	522	104	176	203	39
0.20	anchor 2020	513	71	194	212	36
0.30	20-29	180	67	74	39	0
0.30	30-59	313	137	116	58	2
0.30	60-149	451	173	210	67	1
0.30	150+	615	256	298	58	3
0.30	anchor 2018	524	253	213	57	1
0.30	anchor 2019	522	209	230	80	3
0.30	anchor 2020	513	171	255	85	2

Table S15 Pairs (of 1,559) whose first offset on the observed background equals the first offset on the same pair's trend background and constant background, and pairs whose trend and constant offsets coincide.

δ	Observed = trend	Observed = constant	Trend = constant
0.00	1559	1559	1559
0.10	603	900	941
0.15	362	305	305
0.20	487	377	301
0.30	718	698	535
0.40	899	533	471

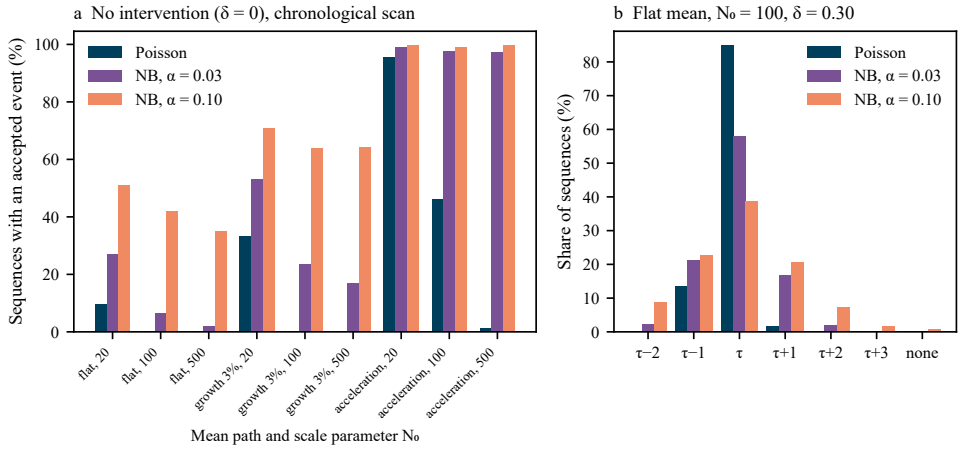


Fig. S3 Count noise around known mean paths. Panel a: share of sequences with at least one accepted event under the chronological scan when the mean contains no step, by mean path, scale parameter N_0 and count law. For the acceleration path the mean at $\tau - 1$ is $N_0 e^{0.405}$, and the deterministic mean path fails the growth-gap condition. Panel b: distribution of the first accepted offset in the local scan for a step of $\delta = 0.30$ on a flat mean at $N_0 = 100$. Simulated counts, 2,000 sequences per cell.

S13 No-change nulls for the observed panel

Each scenario imposes a constant log-linear mean growth rate while retaining a fitted level and stipulated count variation. The fitted path need not represent the history that would have occurred without technological change. For series i a log-linear Poisson mean $\mu_{it} = \exp\{a_i + b_i(t - 2018)\}$ is fitted to the 2011–2025 counts by damped Newton iterations on the two parameters, using every non-empty series (3,154 of the 6,125). Fitted slopes have median 0.026 with quartiles −0.038 and 0.097. Dispersion enters in three forms. Poisson counts set the variance equal to the mean. Pooled dispersion assigns $\alpha = 0.03$ to series with a mean count below 20 and the class medians 0.0202, 0.0146 and 0.0088 of Section S10 to the classes 20–49, 50–149 and 150 and above. Series-wise dispersion uses the moment estimator

Table S16 Cells with $\delta = 0$. Any: share of sequences with at least one accepted event under the chronological scan $\tau - 5, \ldots, \tau + 8$; Events: mean number of accepted events per sequence; Growth only: the same share for the two growth conditions; Local: share with a full-rule label in the local scan $-2, \ldots, +3$. Per cent of 2,000 sequences unless stated.

Mean path	N_0	Law	Any	Events	Growth only	Local
Flat	20	Poisson	9.6	0.097	88.6	4.1
Flat	20	NB 0.03	27.1	0.282	96.5	12.9
Flat	20	NB 0.10	50.8	0.566	99.4	24.0
Flat	100	Poisson	0.0	0.000	11.1	0.0
Flat	100	NB 0.03	6.4	0.065	84.1	2.7
Flat	100	NB 0.10	42.0	0.451	99.1	20.6
Flat	500	Poisson	0.0	0.000	0.0	0.0
Flat	500	NB 0.03	2.0	0.020	75.4	1.0
Flat	500	NB 0.10	35.0	0.368	98.7	17.9
Growth 3%	20	Poisson	33.3	0.353	95.9	15.7
Growth 3%	20	NB 0.03	53.0	0.583	99.0	27.5
Growth 3%	20	NB 0.10	71.0	0.847	99.8	40.6
Growth 3%	100	Poisson	0.2	0.002	34.7	0.1
Growth 3%	100	NB 0.03	23.4	0.240	94.8	11.9
Growth 3%	100	NB 0.10	64.0	0.758	99.7	33.6
Growth 3%	500	Poisson	0.0	0.000	0.1	0.0
Growth 3%	500	NB 0.03	17.0	0.173	90.6	8.0
Growth 3%	500	NB 0.10	64.3	0.755	99.7	35.1
Acceleration	20	Poisson	95.6	1.354	100.0	66.9
Acceleration	20	NB 0.03	99.1	1.557	100.0	73.2
Acceleration	20	NB 0.10	99.6	1.637	100.0	77.0
Acceleration	100	Poisson	46.0	0.493	78.3	27.1
Acceleration	100	NB 0.03	97.7	1.435	100.0	70.6
Acceleration	100	NB 0.10	98.9	1.629	100.0	76.0
Acceleration	500	Poisson	1.2	0.012	5.2	1.0
Acceleration	500	NB 0.03	97.2	1.390	100.0	70.0
Acceleration	500	NB 0.10	99.5	1.634	100.0	78.4

$\hat{\alpha}_i = \max\{0, \sum_t [(N_{it} - \mu_{it})^2 - \mu_{it}] / \sum_t \mu_{it}^2\}$ around the fitted trend, which has median 0.033 on the 201 event series and 0 on the remaining series. That difference may reflect changes, noise, selection and misspecification; it does not isolate acceleration. Counts are drawn as a gamma–Poisson mixture, and 500 replicate panels per null pass through the unchanged chronological scan with cooldown.

A fourth null treats the event series differently. For an event at nominal year k the trend is fitted on the pre-period only, the years 2011 to $k - 1$, with the slope taken from that window and the intercept refitted so that the fitted counts reproduce the pre-period total; the resulting mean path is extrapolated over the whole panel. Series without an accepted event keep their full-sample trend. Pre-period continuation uses the observed event to select its fitting window. It supplies an exploratory continuation scenario rather than an independently specified counterfactual. Calendar dispersion classes use the full-period mean, unlike the pre-period mean classes used for the event-level draws below. For 69 of the 201 event series those choices assign different dispersion

Table S17 Cells with a step. Local-scan first offsets in per cent of 2,000 sequences; Before: share first accepted in the chronological scan at $\tau - 3$ or earlier, that is with a window closing before the step; paired comparison with the growth-only rule on the local grid (same finite date, later, suppressed) in sequences.

Mean path	N_0	Law	δ	-2	-1	0	$+1$	$+2$	$+3$	none	Before	Same	Later	Suppr.
Flat	20	Poisson	0.15	2.0	6.4	17.8	23.0	19.6	12.7	18.6	2.6	357	1271	340
Flat	20	Poisson	0.30	3.3	24.5	55.8	14.8	1.5	0.1	0.1	2.5	460	1539	1
Flat	20	NB 0.03	0.15	4.5	8.8	18.2	20.2	17.9	12.7	17.8	6.8	402	1243	331
Flat	20	NB 0.03	0.30	5.6	23.6	45.7	18.4	6.0	0.6	0.3	6.2	525	1470	4
Flat	20	NB 0.10	0.15	7.3	10.7	18.1	17.6	15.5	12.5	18.4	13.3	505	1128	339
Flat	20	NB 0.10	0.30	10.1	21.3	34.5	22.4	8.7	2.1	1.0	12.7	648	1332	19
Flat	100	Poisson	0.15	0.0	0.3	8.8	27.7	21.0	8.8	33.6	0.0	161	1168	622
Flat	100	Poisson	0.30	0.0	13.4	84.9	1.8	0.0	0.0	0.0	0.0	159	1841	0
Flat	100	NB 0.03	0.15	1.6	4.1	18.8	22.2	18.9	14.2	20.4	0.9	302	1291	389
Flat	100	NB 0.03	0.30	2.4	21.1	57.9	16.8	2.0	0.1	0.0	1.0	340	1660	0
Flat	100	NB 0.10	0.15	6.6	10.6	17.6	18.4	16.4	11.6	19.0	10.5	427	1194	353
Flat	100	NB 0.10	0.30	8.8	22.7	38.6	20.6	7.3	1.6	0.6	9.9	600	1388	11
Flat	500	Poisson	0.15	0.0	0.0	0.6	49.0	9.6	0.6	40.2	0.0	20	1176	778
Flat	500	Poisson	0.30	0.0	1.2	98.9	0.0	0.0	0.0	0.0	0.0	3	1997	0
Flat	500	NB 0.03	0.15	0.6	4.2	15.9	24.0	20.9	12.8	21.7	0.6	244	1322	409
Flat	500	NB 0.03	0.30	1.3	22.5	58.9	15.7	1.5	0.1	0.1	0.8	300	1698	2
Flat	500	NB 0.10	0.15	5.0	9.6	17.9	17.1	17.2	13.7	19.7	8.0	406	1201	377
Flat	500	NB 0.10	0.30	7.9	22.5	40.7	20.1	6.9	1.4	0.7	8.9	584	1402	12
Growth 3%	20	Poisson	0.15	6.9	13.1	27.8	23.9	14.5	7.5	6.5	8.8	651	1219	120
Growth 3%	20	Poisson	0.30	9.5	36.2	46.8	7.5	0.2	0.0	0.0	9.3	759	1241	0
Growth 3%	20	NB 0.03	0.15	8.6	15.2	25.8	23.0	13.6	7.2	6.8	13.3	712	1153	119
Growth 3%	20	NB 0.03	0.30	13.1	32.0	41.0	12.6	1.2	0.2	0.1	14.5	878	1121	1
Growth 3%	20	NB 0.10	0.15	10.6	15.4	21.9	18.6	15.1	9.0	9.6	20.6	715	1094	171
Growth 3%	20	NB 0.10	0.30	15.4	27.3	33.7	17.0	4.9	1.2	0.7	20.1	891	1096	13
Growth 3%	100	Poisson	0.15	0.1	5.2	41.6	29.7	11.5	3.8	8.3	0.2	439	1395	143
Growth 3%	100	Poisson	0.30	0.6	54.2	45.2	0.1	0.0	0.0	0.0	0.0	477	1523	0
Growth 3%	100	NB 0.03	0.15	4.2	13.3	30.0	24.1	14.4	7.0	7.3	5.9	564	1291	126
Growth 3%	100	NB 0.03	0.30	7.3	36.7	47.9	7.6	0.6	0.1	0.0	6.4	617	1383	0
Growth 3%	100	NB 0.10	0.15	10.6	15.3	20.1	21.4	14.4	9.3	9.1	19.2	690	1128	165
Growth 3%	100	NB 0.10	0.30	15.4	28.9	36.0	14.9	4.3	0.4	0.2	18.4	846	1150	4
Growth 3%	500	Poisson	0.15	0.0	0.2	61.0	26.7	5.8	0.3	6.1	0.0	483	1396	111
Growth 3%	500	Poisson	0.30	0.0	62.7	37.4	0.0	0.0	0.0	0.0	0.0	453	1547	0
Growth 3%	500	NB 0.03	0.15	3.5	12.6	30.9	26.0	14.7	5.7	6.8	3.4	496	1368	122
Growth 3%	500	NB 0.03	0.30	6.2	39.3	48.1	6.3	0.2	0.0	0.0	3.8	634	1366	0
Growth 3%	500	NB 0.10	0.15	10.5	14.6	22.3	20.2	14.8	8.6	9.2	18.9	662	1154	169
Growth 3%	500	NB 0.10	0.30	14.6	27.3	38.0	16.2	3.0	0.7	0.3	16.9	821	1173	6
Acceleration	20	Poisson	0.15	14.8	31.8	38.3	12.4	2.1	0.5	0.2	16.1	1088	908	2
Acceleration	20	Poisson	0.30	21.3	61.4	16.8	0.5	0.0	0.0	0.0	15.3	1109	891	0
Acceleration	20	NB 0.03	0.15	17.3	29.7	31.9	14.8	4.8	1.2	0.6	24.1	1096	893	5
Acceleration	20	NB 0.03	0.30	26.1	47.0	24.3	2.5	0.2	0.0	0.0	23.2	1188	812	0
Acceleration	20	NB 0.10	0.15	20.5	24.8	25.4	16.9	7.8	3.4	1.4	30.4	1044	929	20
Acceleration	20	NB 0.10	0.30	24.9	37.0	29.3	7.3	1.3	0.3	0.2	31.2	1174	823	3
Acceleration	100	Poisson	0.15	3.8	53.2	41.0	2.1	0.1	0.1	0.0	0.6	952	1048	0
Acceleration	100	Poisson	0.30	6.7	91.5	1.9	0.0	0.0	0.0	0.0	0.8	679	1321	0
Acceleration	100	NB 0.03	0.15	16.1	32.8	34.5	11.9	3.5	0.9	0.4	15.0	1021	971	3
Acceleration	100	NB 0.03	0.30	23.2	53.0	23.2	0.7	0.1	0.0	0.0	15.5	1070	930	0
Acceleration	100	NB 0.10	0.15	19.3	25.8	27.9	16.8	6.6	2.7	1.0	29.5	1047	934	17
Acceleration	100	NB 0.10	0.30	25.8	38.3	28.5	6.6	0.8	0.1	0.0	28.4	1133	867	0
Acceleration	500	Poisson	0.15	0.2	83.8	16.0	0.0	0.0	0.0	0.0	0.0	1256	744	0
Acceleration	500	Poisson	0.30	0.3	99.7	0.0	0.0	0.0	0.0	0.0	0.0	204	1796	0
Acceleration	500	NB 0.03	0.15	14.1	32.8	36.5	12.6	3.4	0.7	0.1	12.0	991	1007	2
Acceleration	500	NB 0.03	0.30	20.9	57.8	20.9	0.4	0.1	0.0	0.0	12.4	1015	985	0
Acceleration	500	NB 0.10	0.15	19.5	27.6	26.9	15.5	7.5	2.1	1.1	28.7	1041	937	18
Acceleration	500	NB 0.10	0.30	25.5	40.7	26.1	6.9	0.7	0.3	0.0	27.3	1152	848	0

Table S18 Latent AR(1) deviation on the growing mean ($g = 0.03$, $N_0 = 100$), Poisson counts conditional on the mean, $\sigma_v = 0.15$. Local first offsets in per cent; Any and Before as in the previous tables.

ρ	δ	−2	−1	0	+1	+2	+3	none	Any	Before
0.0	0.00	0.8	1.5	1.4	1.1	1.2	1.4	92.9	16.8	4.6
0.0	0.15	3.1	11.0	30.5	26.1	15.4	7.5	6.5	99.0	4.8
0.0	0.30	6.8	38.0	49.3	5.7	0.2	0.1	0.0	100.0	4.0
0.6	0.00	1.8	1.5	1.7	1.6	1.4	1.7	90.6	19.1	4.6
0.6	0.15	3.9	12.6	28.6	22.6	15.0	7.9	9.6	98.9	4.8
0.6	0.30	5.6	40.6	48.0	5.5	0.4	0.1	0.0	100.0	4.8

values. Sixty-eight endpoint-only series have no finite log-linear Poisson slope maximum-likelihood estimate. The numerical fitting routine approaches a boundary solution; small changes in those fitted parameters under an independent rerun did not change the reproduced calendar outcomes.

Table S19 reports the calendars by nominal year. Simulated profiles are flatter than the observed 48 events in 2016 and nine in 2023. Observation-window boundaries limit interpretation, but the comparison does not establish their contribution to either feature. Nor is any fitted scenario an upper or lower bound on the number of genuine changes.

Table S19 Accepted events by nominal year, observed and under four no-change nulls, means of 500 replicate panels each. Nulls: Poisson counts around the fitted trend; pooled dispersion by level class; dispersion estimated for each series around its own trend; pre-period continuation for event series.

Year	Observed	Poisson trend	Pooled dispersion	Series dispersion	Continuation
2016	48	13.7	23.3	27.4	18.8
2017	29	12.7	21.1	25.0	17.9
2018	29	12.4	20.5	23.0	17.2
2019	20	12.3	20.5	23.5	16.8
2020	25	12.4	20.4	22.7	16.5
2021	26	12.1	20.1	21.7	15.7
2022	15	13.3	21.8	23.8	17.3
2023	9	14.4	24.0	27.0	19.8

Under the pre-period continuation scenario, approximately 29 per cent of the 201 observed event series receive at least one simulated label, compared with 2.5 per cent of the 2,953 other non-empty series. This contrast combines different selected populations and fitting windows; it cannot be attributed uniquely to scale or growth.

A separate comparison uses the full-period pooled-dispersion scenario for both groups. Qualification frequencies were estimated from 5,000 panels, independently of 500 further panels used to select simulated events. The median frequency is 0.0642 at observed events and 0.0922 at simulated events. In paired comparisons, the observed

event has the strictly lower frequency with probability 0.59345; ties account for 0.00127. Giving ties half weight yields 0.59408. Separate simulation batches avoid scoring a simulated event using a frequency estimate that includes its own draw. The weak separation remains conditional on the fitted scenario and is not a classifier accuracy or a significance test.

S14 Event-level continuation probabilities and date stability

Two quantities are computed for each accepted event and reported in Table S20. Continuation probability asks how often the five conditions hold at the nominal year when the series continues its own pre-period trend. A Poisson trend on 2011 to $k-1$ is extrapolated to $k+2$, counts for the whole stretch are drawn as negative binomial with the pooled dispersion of the class determined by the five local pre-window counts, and 2,000 replicates give the share in which all five conditions hold at k. Values are pointwise. No correction is made for the 25,232 candidate evaluations in the non-empty panel. The calendar simulation does not replace a multiplicity correction, and these selected, fitted qualification frequencies are not calibrated test p values. With zero successes in 2,000 independent draws, the Monte Carlo estimate is zero and the one-sided 95 per cent binomial upper bound is $1 - 0.05^{1/2000} \simeq 0.001497$.

Date stability asks how often the printed year survives count variation around the observed path. Replicates $N^*_{it} \sim \mathrm{NB}(N_{it}, \alpha)$ with the same locally assigned pooled dispersion are drawn for 2011–2025, and the full chronological scan is run on each. Recorded are the share of replicates whose first accepted year equals k, the share within one year of k and the share with no accepted event in that series. Poisson replicates provide a lower-variance comparison, with mean same-year shares of 0.52 against 0.45 under the negative binomial. Neither number is sampling error of a population: the observed path is held fixed and only the count law varies.

Table S20 Event-level results by binding margin and by post-period scale. Columns: number of events; mean share of bootstrap replicates reproducing the nominal year, reproducing a year within one of it, and producing no event in that series; median continuation probability at the nominal year; number of events whose continuation probability is below 0.05. Bootstrap and continuation replicates: 2,000 per event.

Class	Events	Same year	Within one	No event	Median p	$p < 0.05$
< 0.05	48	0.33	0.50	0.34	0.013	37
0.05 to 0.10	39	0.40	0.58	0.28	0.013	30
≥ 0.10	114	0.53	0.70	0.18	0.027	66
Post-period median < 50	111	0.42	0.59	0.27	0.013	73
Post-period median ≥ 50	90	0.49	0.68	0.19	0.026	60
All events	201	0.45	0.63	0.24	0.020	133

Fourteen events meet their binding condition exactly, eight at the floor with a post-period median of exactly 20 and six at the level ratio; their mean same-year share is 0.30. Among the conditions, the level ratio fails most often in the continuation replicates, in 99 of the 201 events, followed by the growth increase in 55 and the floor in 45, which repeats on stochastic paths the ordering that Table 2 of the main text finds on the observed ones.

S15 A whole-series comparator

Least squares fits the continuous broken stick $\ell_{it} = a + b_1(t - t_0) + b_2 \max\{0, t - t_0\}$ for every knot t_0 in 2013–2023, and the knot with the smallest residual sum of squares is retained. First year of the new regime is $t_0 + 1$, matching the meaning of the nominal year k as the first year of the post-window. An F statistic against the straight line with one and twelve degrees of freedom accompanies each fit; the knot search makes its nominal 5 per cent point anti-conservative, and no correction is applied because the statistic is used descriptively.

Offsets $k - (t_0 + 1)$ for the 201 event series appear in Table S21. Both tails are heavy, 33 series with the rule at least five years earlier and 29 at least five years later. In 113 series the fitted slope change is non-positive, so the selected one-knot fit has a non-positive slope change even though a take-off was accepted somewhere inside it; 70 series have a slope increase of at least 0.10 and 49 combine that with an F statistic above the nominal point. Among the 670 series with mean count of at least 20, which contain 150 of the event series, 105 have a slope increase of at least 0.10 and an F statistic above the uncorrected fixed-knot reference, and 42 of those carry an accepted event.

Table S21 Offset between the nominal year of the rule and the first year of the new regime under a continuous broken stick fitted to the whole series, $k - (t_0 + 1)$, for the 201 accepted events. Negative values place the rule earlier.

Offset in years	Event series
≤ -5	33
-4 to -2	62
-1	7
0	19
$+1$	12
$+2$ to $+4$	39
$\geq +5$	29

S16 Left censoring and holding-address geographies

Two features of the accepted calendar invite the objection that the headline results rest on them. Nominal year 2016 is the first the scan can accept and carries 48 of the 201

events. Seven reporting geographies where first-applicant attribution plausibly records holding or licensing addresses, Singapore, Ireland, Hong Kong SAR, Liechtenstein, Luxembourg, Barbados and the Cayman Islands, together carry 48 events. Table S22 recomputes the headline statistics on the common calendar without the censored year, without those geographies and without both. Neither omission changes the qualitative result. Level ratio remains the binding criterion in 66 to 74 per cent of the remaining events in every subset, the share of events within 10 per cent of their binding threshold stays between 43 and 50 per cent, and the earliest-labelled geography of a field still changes in 22 to 25 of the fields that retain events under the $(4, 4)$ windows. Removal is reported, not adopted: the main results keep every event, because a rule applied to an official report is being studied as it would be used.

Table S22 Headline statistics on the common 2016–2022 calendar under omission of the censored first year and of holding-address geographies. Fields first: fields whose set of earliest-labelled geographies changes, over fields with events under both definitions. Order preserved: pairs of series labelled under both definitions whose within-field order is unchanged under the $(4, 4)$ windows.

Subset	Events	Ratio binds	Floor binds	Within 10%	Fields $(4, 4)$	Fields 1.25	Order kept
Baseline, 2016–2022	192	127	48	83	24/33	19/34	211/260
Without the censored year 2016	144	95	37	67	25/31	20/34	85/120
Without holding-address geographies	148	108	26	67	25/33	19/34	90/110
Without both	111	82	19	56	22/30	19/34	50/65

S17 Change shapes other than a permanent growth step

Blocks of thirteen years $\tau - 7, \ldots, \tau + 5$ carry a background $N_0 \exp\{g(t - \tau)\}$ and one of three alternative shapes. A level jump multiplies the mean by J from τ onwards. A growth phase adds δ to each of the three log increments $\tau - 1 \to \tau$, $\tau \to \tau + 1$ and $\tau+1 \to \tau+2$ and then holds the resulting multiplier constant; the background continues at its original rate. A logistic transition multiplies the mean by $1 + (M - 1)\sigma((t - \tau)/s)$ with $s = 1.37$, which places about eighty per cent of the transition inside six years. Deterministic evaluation uses real-valued counts and the weak inequalities of the production rule; the stochastic cells draw negative binomial counts at $\alpha = 0.03$ with 2,000 sequences each. Table S23 gives every cell.

Deterministic first offsets happen to agree across the three tested scales. Scale nevertheless enters the log-growth criteria through $\log(1 + N)$, as well as through the count floor. A pure level jump on a flat background fails positivity because only one of the anchored differences is positive. With the tested background rates $g = 0.03$ and 0.06, the jump first qualifies at $\tau - 1$; no universal offset is claimed for other rates. The logistic doubling tested here never meets the gap threshold, whereas tripling over the same span is dated two years before its midpoint. For a three-year growth phase at

$\delta = 0.20$ on a flat background, no label is accepted, although the permanent step is accepted at $\tau + 1$. At $\delta = 0.30$ their first dates agree in the tested deterministic cells. Finite shared prefixes can prevent a windowed rule from distinguishing lasting and temporary changes, but the reported designs are not equivalent at every step size.

Table S23 Alternative change shapes, all cells. Deterministic first accepted offset relative to τ for three constant-growth backgrounds, and shares of 2,000 negative binomial sequences ($\alpha = 0.03$) whose first local offset is absent, before τ or at τ, in per cent, for flat and growing backgrounds. Permanent steps are shown for comparison.

Shape	N_0	Deterministic			Noisy, $g = 0$			Noisy, $g = 0.03$		
		$g = 0$	.03	.06	None	Earl.	At τ	None	Earl.	At τ
Level jump ×1.5	20	none	$\tau - 1$	$\tau - 1$	45	26	20	24	41	24
	100	none	$\tau - 1$	$\tau - 1$	53	21	20	27	40	26
	500	none	$\tau - 1$	$\tau - 1$	52	21	23	27	42	25
Level jump ×2	20	none	$\tau - 1$	$\tau - 1$	15	55	26	8	69	21
	100	none	$\tau - 1$	$\tau - 1$	12	61	25	6	73	20
	500	none	$\tau - 1$	$\tau - 1$	11	62	25	6	74	20
Growth phase 0.20	20	none	τ	$\tau - 1$	27	17	27	12	31	32
	100	none	τ	$\tau - 1$	30	11	33	12	25	41
	500	none	τ	$\tau - 1$	29	8	35	12	22	44
Growth phase 0.30	20	τ	$\tau - 1$	$\tau - 1$	5	30	45	2	45	41
	100	τ	$\tau - 1$	$\tau - 1$	4	24	56	1	44	47
	500	τ	$\tau - 1$	$\tau - 1$	3	24	59	1	45	49
Permanent step 0.20	20	$\tau + 1$	τ	$\tau - 1$	5	16	28	1	28	32
	100	$\tau + 1$	τ	$\tau - 1$	3	9	34	1	23	43
	500	$\tau + 1$	τ	$\tau - 1$	3	9	35	1	24	44
Permanent step 0.30	20	τ	$\tau - 1$	$\tau - 1$	0	29	45	0	45	42
	100	τ	$\tau - 1$	$\tau - 1$	0	24	57	0	41	49
	500	τ	$\tau - 1$	$\tau - 1$	0	22	61	0	45	49
Logistic ×2	20	none	none	none	37	30	14	20	43	16
	100	none	none	none	48	26	13	23	43	16
	500	none	none	none	51	22	13	25	42	16
Logistic ×3	20	$\tau - 2$	$\tau - 2$	$\tau - 2$	16	56	15	10	66	13
	100	$\tau - 2$	$\tau - 2$	$\tau - 2$	15	60	14	10	72	11
	500	$\tau - 2$	$\tau - 2$	$\tau - 2$	17	62	14	10	73	11

S18 The priority-year populations in detail

Source and selection. OECD SDMX, agency OECD.STI.PIE, dataflow DSD_PATENTS@DF_PATENTS_WIPO version 1.0, retrieved on 9 September 2026 in SDMX-CSV. Five keys were retrieved, all with frequency annual, date type PRIORITY, both agent roles, all WIPO fields except the total and all reference areas except the three reported aggregates OECD, EU27_2020 and W: 6F0.A.AP.PATN (European Patent Office applications), US50.A.AP.PATN (United States Patent and Trademark Office applications), 9P50_1.A.AP.PATN (applications under the Patent Cooperation Treaty), 9P50_2.A.PF.PATN_FM (triadic patent families) and 9P50_3.A.PF.PATN_FM (IP5 patent families). Priority years 2008–2022 form a 15-level block of the same length as the reporting-year panel, so the calendar of complete candidate windows is 2013–2020 in every population. Counts are fractional, because a document classified

in several fields contributes a share to each, and they are used as retrieved. The five retrieved files and their SHA-256 digests travel with the analysis archive; the leading twelve hexadecimal digits appear in Table S24.

The rule is unchanged throughout, including the log-transform pseudocount and the lower bound on the ratio denominator and the count floor of 20, both of which apply to fractional counts without modification. No threshold, window or cooldown setting was adjusted for any population, and no result from any of them was used to choose any setting. Table 6 of the main text gives the binding criteria and margins; Table S24 compares the same office or family population definition under different attribution roles. The aggregate exports do not permit document-level matching or isolation of attribution-coverage differences.

Table S24 Priority-year populations, 2008–2022, and the effect of the attribution rule within each. Events are counted under inventor and under applicant attribution; "both" counts field–geography series labelled under each; "same" and "±1" compare their first accepted nominal years on those series; "fields moved" counts fields, among those labelled under both, whose set of earliest-labelled geographies changes. The last column gives the leading digits of the SHA-256 digest of the retrieved export.

Population	Key	Meas.	Inv.	App.	Both	Same	±1	Fields moved	Digest
EPO	6F0	AP	139	164	94	70	84	21/33	b8c82f7a0e77
USPTO	US50	AP	114	130	56	38	50	28/32	77e2b607165f
PCT	9P50_1	AP	156	185	113	90	107	25/34	1ff5de78c1ef
Triadic	9P50_2	PF	77	90	57	45	53	15/32	39e9ef1a1af5
IP5	9P50_3	PF	107	122	58	38	45	24/34	eef2527db8af

Applicant attribution yields more labelled series in each population. Where series have two events, comparison uses the minimum accepted year, not the last row encountered. For EPO counts the event/series totals are 139/138 under inventor attribution and 164/161 under applicant attribution. Among the 94 shared labelled series, 70 first years match and 84 are within one year. Corresponding exact matches are 38 of 56 for USPTO, 90 of 113 for PCT, 45 of 57 for triadic families and 38 of 58 for IP5 families. These differences establish sensitivity to attribution; they do not identify which applications were filed through holding entities.

What the comparison does not establish. The populations differ from the reporting-year panel in year definition, geography roster, count type and calendar window at once, so event-level correspondence between them is not computed and no accuracy claim is made in any direction. What is compared is the shape of the outcome: which criterion binds, how close events sit to their thresholds, how much of the calendar falls on the first year the scan can accept and how far a lead statement survives a change of attribution. Families are counted by the OECD definition of a triadic and of an IP5 family; other family definitions exist and would give other counts (Martínez 2011).

S19 Reporting against the GLOBAL guidance list

GLOBAL was developed as reporting guidance for bibliometric analyses of publications and citations (Ng et al. 2025). It is adapted here as an organising checklist, without claiming validation as a patent-study standard. The abstract states scope, data and principal results. Introduction and rule sections define the questions, counting unit and three dates. Data and design sections give populations, eligibility, transformations, computational conventions and the distinct roles of constructed paths and simulations. Results distinguish criterion margins, condition-removal effects, specification sensitivity, fitted scenarios and attribution comparisons. The supplement supplies full proofs, detailed grids and source keys.

Uncertainty is reported as scenario distributions and count-perturbation frequencies, with fitted-model, selection and multiplicity limitations explicit. Discussion separates operational sensitivity from untested historical validity. The manuscript records the current replication availability and identifies the personal declarations that remain open in this local review draft. The original EPO export chain is not complete in the supplied replication distribution, which limits source-level reconstruction despite reproducibility from the normalized panel. The official aggregate datasets are not a literature-search sample; export filters and SDMX keys replace a bibliographic search strategy. No public preregistration or calibrated inferential test of technological onset is claimed.

References

European Patent Office (2026) European patent applications and grants. EPO Data Desk, URL https://datadesk.epo.org/dashboards/european-patent-applications-and-grants-2026-03, published 23 March 2026. Accessed 8 September 2026

European Patent Office (n.d.a) Clarification on data sources. URL https://www.epo.org/en/about-us/statistics/clarification-data-sources, accessed 8 September 2026

European Patent Office (n.d.b) European patent applications per field of technology and per country of residence of the applicant for each individual year 2016–2025. EPO Technology Dashboard 2025, statistical workbook, URL https://link.epo.org/web/about-us/statistics/european-patent-applications-per-field-of-technology-and-per-country-of-residence-of-the-applicant-for-each-individual-year-2016-2026-epo-technology-dashboard-2025.xlsx, sheets 2016–2025. Accessed 8 September 2026

Martínez C (2011) Patent families: When do different definitions really matter? Scientometrics 86(1):39–63. https://doi.org/10.1007/s11192-010-0251-3

Ng JY, Liu H, Masood M, et al (2025) Guidance list for reporting bibliometric analyses (GLOBAL): A two-round modified delphi study. In: Proceedings of the 20th International Conference on Scientometrics and Informetrics (ISSI 2025), pp 837–861, https://doi.org/10.51408/issi2025_069