RESEARCH

Supplementary Information
Success and luck in creative careers

Milán Janosov[1], Federico Battiston[1] and Roberta Sinatra[2,3,4*]

*Correspondence: rsin@itu.dk
[2]Department of Computer Science, IT University of Copenhagen, 2300 Copenhagen, Denmark
Full list of author information is available at the end of the article

S1 Data

S1.1 Data sets

Our research was based on four different data that were collected during the period of June-August 2017.

IMDb dataset. We collected information on individuals active in the movie industry based on the Internet Movie Database (IMDb) [1]. To this we first used the Advanced Title Search[1] function and sent multiple queries to obtain the list of all movie identifiers which received a vote from at least users. Using the list of unique movie identifiers (∼1.3 million) we downloaded the HTML source code of each movie's site. After processing all the raw HTML files about the movies we ∼0.8 million distinct names being as director, producer, scriptwriter, composer, and art-director, and created the career by associating each movie to the corresponding individuals, for each profession separately (e.g. directors, producers). We attached the six different success measures present in the database: average rating, rating count, metascore [2], gross revenue, the number of user and critic reviews to each career and constructed the individuals' career trajectories as time series of these quantities.

Discogs and LastFM dataset. To cover individuals active in the music industry we relied on Discogs[2] [3], a crowd-sourced music discography website. Via its search functionality, we listed all the master releases from the genres of rock, pop, electronica, folk, funk, hip-hop, classical, and jazz music to obtain a comprehensive list of ∼0.4 million artists combined. After crawling their discographies based on their unique identifiers from Discogs and parsing them into tracklists, we used the API of LastFM[3] [4], a music-providing service, to extract the play counts used as impact measures. For each , of the artists and kept only those which had been played at least once. This way, we obtained a dataset consisting of ∼31 million songs. Then we combined the timestamped discography and the song – play count datasets to reconstruct the musicians' careers for each genre.

Goodreads dataset. We gathered data about book authors using Goodreads[4] [5], a social network site for readers, by crawling the HTML website of the profile of ∼2.1 million individuals authored ∼6.6 million books. By extracting information from the authors' biography profiles we built their career trajectories. Goodreads

[1]www.imdb.com/search/title
[2]www.discogs.com/search/
[3]www.lastfm.com
[4]www.goodreads.com

provides three different ways for measuring impact: the average users' ratings of a book, the total number of ratings, and the number of editions a book has, from which we used the rating count for further analysis.

Web of Science dataset. We used the Web of Science [6] database to reconstruct the careers of scientists from 15 scientific disciplines: Agronomy, Applied Physics, Biology, Chemistry, Engineering, Environmental Science, Geology, Health Science, Mathematics, Physics, Political Science, Space Science Or Astronomy, Theoretical Computer Science, Zoology. In total, we analyze the careers of 1.2 million scientists, who authored 87.4 million papers. Each paper has been associated with the number of citations received. The career of a scientist consists of her publication record and the citation impact of each paper.

After collecting these data sets, to limit the analysis to careers with sustained productivity, we set filtering thresholds to 10 movies and papers for individuals (except art-directors, for whom it was 20), 50 books for authors,

S1.2 Measuring success in artistic domains

Our research premise is that success is a social phenomenon and as such we aim to capture "a community's reactions to the performance of the individuals" [7, 8]. For this reason, the movies, songs, books, and scientific papers in our database are associated with measures of success of different nature based on their social context. On one hand, there are success measures that are based on the evaluation of experts of the field, who have supposedly more insights on the underlying performance associated with the artistic product. On the other hand, success measures based on the opinion of the general public have larger statistics. However, they are also more likely to be biased by external factors, such as the rich-gets-richer phenomenon or the peer-effect [9]. From a statistical perspective, success measures can be either obtained as an average of responses over time or as the result of cumulative activities through time (Figure S1). We based our analysis on the cumulative measures since these are the only ones present in all different available data sets. This also allowed us to adapt existing techniques and methodologies previously used for the study of paper and scientific careers.

		Measure name	Example value	Source of opinion: Public	Source of opinion: Expert	Statistical nature: Averaging	Statistical nature: Cumulative
	Film	Average rating	8.1 (10)	X		X	
		Rating count	1040234	X			X
		Metascore	90 (100)		X	X	
		Critic reviews	298		X		X
	Music	Play count	65489	X			X
	Books	Average rating	4.4 (5)	X		X	
		Rating count	135468	X			X

Figure S1 Taxonomy of success measures. Success measures associated with movies, songs and books differ based on their definition (public opinion or experts' evaluation), and their statistical nature (averaged or cumulated over time), defining four distinct categories of success measures. Measures associated with all possible four combinations are present for movies (IMDb [1]). Songs (Discogs and Lastfm [3, 4]) are only associated with a public opinion-based cumulative measure. Data about books (Goodreads [5]) provide only public opinion based measures.

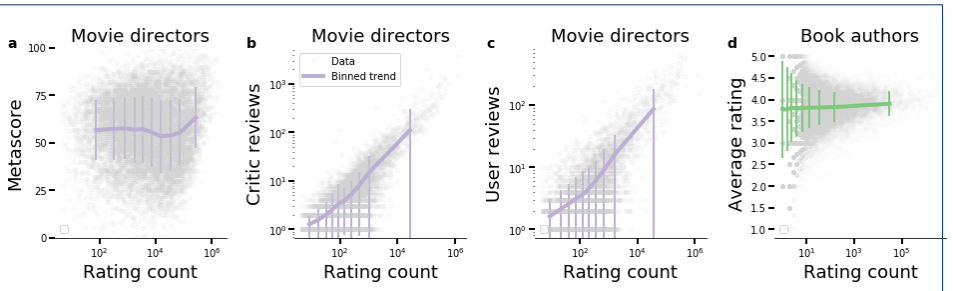

Figure S2 Correlations between different success measures. The correlations between movies' rating counts and their **a**, Metascores (Spearman's rank correlation coefficient $r_S \approx 0.151$); **b**, number of critic reviews ($r_S \approx 0.839$); and **c**, number of user reviews ($r_S \approx 0.641$) for the case of movie directors. Subfigure **d**, shows the low correlation between the number of ratings and the average rating a book receives on Goodreads ($r_S \approx 0.022$). In the scatter plots, each dot represents a product (i.e. a movie or a book), while the continuous lines show the percentile-binned trends using 10 bins.

S1.3 Correlations between different success measures

Two of our data sets, covering movies and books, contain more than one type of success measures. Here we compare them by computing the correlations between pairs of measures of a different kind. We find that different cumulative measures show high correlations with each other (see Fig. S2b-c), indicating that results are robust to the choice of the specific cumulative measure. Averaged measures, like Metascore (Fig. S2a) or average rating Fig. S2d), do not correlate well with cumulative measures, indicating a different process generating these measures. Since these averaged measures have a broad distribution, and previous literature offers methods and finding mainly about cumulative measures, we opted to use cumulative measures .

S2 Q-model

S2.1 Testing the random impact rule

The random impact rule states that the chronological rank of the best product (N^*) over a career with a length of N (measured as the number of creative products throughout ones career) is uniformly randomly distributed across a large sample of careers, meaning that the probability distribution $P(N^*/N)$ is well approximated by a uniform $U(0,1)$ distribution, as prior work has already shown for scientific fields [10] and other creative domains [11]. To test this hypothesis in our creative domains, we compared the observed success cumulative distribution function (CDF) $P(> N^*/N)$ with both the CDF of the theoretical $U(0,1)$ distribution, and the CDF in a set of synthetic careers. In synthetic careers, we randomly reshuffled the products, making sure that N^* takes a uniformly randomly assigned position over the career. To obtain statistically reliable results, we repeated this randomization 100 times. We quantified the goodness of the fit by computing the R^2 deviation original and the randomized data from the theoretical null model, i.e. the cumulative distribution function of the $U(0,1)$ uniform distribution. The goodness of the fit for all the studied professions is measured by the R^2 value comparing the data to the $U(0,1)$. These results are summarized in Table S1.

Field	R^2_{random}	R^2_S	R^2_N	R^2_p	R^2_Q
Agronomy	0.99955	0.99955	0.988	0.9808	0.9419
Applied Physics	0.99979	0.99979	0.982	0.9572	0.9654
Biology	0.99967	0.99967	0.984	0.9884	0.9822
Book authors	0.97773	0.97773	0.894	0.9796	0.9796
Chemistry	0.99963	0.99963	0.989	0.9945	0.996
Classical musicians	0.96694	0.96694	0.963	0.9444	0.9933
Electronic music artists	0.99553	0.99553	0.947	0.9857	0.9813
Engineering	0.99973	0.99973	0.986	0.9776	0.9793
Environmental science	0.99969	0.99969	0.968	0.9932	0.9933
Folk musicians	0.96508	0.96508	0.923	0.981	0.973
Funk musicians	0.95236	0.95236	0.934	0.9916	0.988
Geology	0.99967	0.99967	0.982	0.9885	0.9824
Health Science	0.99962	0.99962	0.991	0.9689	0.9811
Hip-hop artists	0.94326	0.95512	0.882	0.9915	0.9875
Jazz musicians	0.96488	0.96488	0.869	0.989	0.9844
Mathematics	0.99969	0.99969	0.983	0.9856	0.9778
Movie art directors	0.97207	0.97207	0.922	0.9989	0.9994
Movie directors	0.9982	0.9982	0.96	0.9914	0.9875
Movie producer	0.99941	0.99941	0.916	0.9817	0.974
Physics	0.99972	0.99972	0.989	0.9961	0.9972
Political Science	0.99973	0.99973	0.984	0.9913	0.9844
Pop musicians	0.99407	0.99407	0.903	0.9742	0.9844
Rock musicians	0.95565	0.95565	0.948	0.9934	0.9952
Script writers	0.99957	0.99957	0.877	0.9841	0.9751
Soundtrack composers	0.99923	0.99923	0.974	0.9801	0.9822
Space Science or Astronomy	0.99959	0.99959	0.965	0.9928	0.9922
Theoretical Computer Science	0.99954	0.99954	0.982	0.9918	0.9886
Zoology	0.9996	0.9996	0.976	0.9807	0.9206

Table S1

S2.2 Q-model

The Q-model, proposed in [10], assumes that when the distribution of the impact of scientific papers can be described by log-normal functions, then the impact can be expressed as the trivariate log-normal distribution of three variables: *(i)* the productivity of the individuals (e.g. the number of papers they publish, N) *(ii)* an individual based parameter only depending on the individual's prior works' success and *(iii)* a random parameter representing outer factors (p). By transforming these variables to the logarithmic space ($\hat{N} = \log N$, $\hat{Q} = \log Q$, $\hat{p} = \log p$), the impact $P(\hat{S})$ distribution reads:

$$P(\hat{S}) \quad = \quad P(\hat{p}, \hat{Q}, \hat{N}) = \frac{1}{\sqrt{(2\pi)^3}} \exp\left(-\frac{1}{2}(\mathbf{X} - \mu)^T \Sigma^{-1} (\mathbf{X} - \mu) \right), \tag{1}$$

where $\mathbf{X} = (\hat{p}, \hat{Q}, \hat{N})$, $\mu = (\mu_N, \mu_p, \mu_Q)$ is the average vector, and Σ the covariance matrix: $\Sigma = \begin{pmatrix} \sigma_p^2 & \sigma_{p,Q} & \sigma_{p,N} \\ \sigma_{p,Q} & \sigma_Q^2 & \sigma_{Q,N} \\ \sigma_{p,N} & \sigma_{Q,N} & \sigma_N^2 \end{pmatrix}$. If the cross-terms $\sigma_{p,Q}$ and $\sigma_{p,N}$ are close to zero, then the distribution of p does not depend on variables capturing individual. In this case, a number of simplifications can be made, and the impact rescaled by the individual parameter Q collapses on the same distribution for all individuals. To obtain the covariance matrix of the trivariate log-normal distribution of Eq. 1, we fit the theoretical distribution to the data by using CMA-ES [12, 13] (Covariance Matrix Adaptation Evolution Strategy Evolution trategies), from which we obtained the parameters in Table S2. The shown results are consistent with the reported findings of scientific careers in [10].

S2.3 Requirements of the Q-model

To apply the Q-model to a creative domain, the data has to fulfill requirements.

Field	μ_N	μ_p	μ_Q	σ_N	σ_Q	σ_p	σ_{pQ}	σ_{pN}	σ_{QN}
Agronomy	4.467	2.148	4.147	1.711	0.452	0.514	-0.069	0.025	-0.005
Applied Physics	4.521	1.224	2.938	1.823	0.399	0.452	-0.046	-0.015	0.012
Movie art directors	4.214	4.557	3.877	1.606	0.437	0.465	-0.047	0.030	0.037
Biology	4.305	2.334	4.078	1.867	0.434	0.519	-0.072	-0.074	-0.027
Books authors	3.514	2.868	5.571	1.682	0.396	0.476	-0.054	-0.014	-0.062
Chemistry	4.366	2.462	4.586	1.629	0.440	0.485	-0.059	-0.132	0.083
Classical musicians	5.313	4.762	5.896	1.873	0.488	0.502	-0.072	0.024	0.003
Soundtrack composer	3.687	3.935	5.312	1.634	0.358	0.416	-0.030	0.005	-0.059
Movie directors	4.174	4.673	4.992	1.783	0.426	0.469	-0.053	0.097	-0.017
Electronic music artists	5.144	4.531	3.761	1.833	0.345	0.415	-0.035	0.026	0.038
Engineering	4.976	2.170	3.768	1.601	0.457	0.500	-0.063	0.017	0.046
Environmental Science	3.489	2.112	4.194	1.517	0.445	0.512	-0.063	0.034	-0.025
Folk musicians	5.246	4.071	6.260	1.744	0.434	0.475	-0.054	-0.026	-0.060
Funk musicians	5.608	3.612	6.081	1.656	0.447	0.477	-0.058	-0.076	0.011
Geology	4.592	4.749	4.441	1.802	0.374	0.435	-0.037	0.071	-0.056
Health Science	4.114	3.986	3.513	1.702	0.448	0.494	-0.061	0.048	-0.023
Hip-hop artists	5.354	4.703	5.433	1.472	0.481	0.495	-0.068	-0.072	-0.078
Jazz musicians	4.380	3.340	5.337	1.695	0.379	0.413	-0.030	-0.022	-0.099
Mathematics	4.557	4.662	4.212	1.647	0.378	0.434	-0.040	-0.030	-0.102
Physics	4.552	1.760	3.396	1.571	0.383	0.454	-0.046	0.055	0.037
Political Science	4.723	2.889	4.041	1.807	0.384	0.461	-0.055	-0.002	0.010
Pop musicians	6.071	2.553	4.421	1.911	0.417	0.454	-0.046	-0.024	0.018
Movie producers	3.823	4.424	4.010	1.499	0.397	0.476	-0.052	-0.081	-0.019
Psychology	4.565	4.696	4.514	1.000	0.737	2.013	0.220	0.009	-0.012
Rock musicians	5.518	3.858	3.643	1.572	0.451	0.509	-0.069	-0.017	0.046
Space Science Astronomy	5.231	2.164	2.523	1.672	0.371	0.462	-0.046	-0.056	0.094
Theoretical Computer Science	4.058	4.781	4.216	1.799	0.426	0.456	-0.045	-0.006	0.042
Script writers	3.024	2.944	4.128	1.620	0.461	0.516	-0.073	-0.005	-0.073
Zoology	4.058	3.655	3.523	1.723	0.377	0.429	-0.042	-0.028	-0.048

Table S2 Optimization results. The table reports the parameters of the $P(N)$, $P(Q)$, and $P(p)$ distributions obtained evolutionary optimization for all the studied fields.

S2.3.1 Fitting the impact distributions

To model the distribution of the success measure on the different fields – rating count for movies and books, play count for songs, and citations for scientific papers – we assumed a log-normal shape and fitted the cumulative distribution function of the data (examples from each data set are on Figure S3). We quantified the goodness of the fit by computing R^2 values, whose values are reported for all the fields in Table

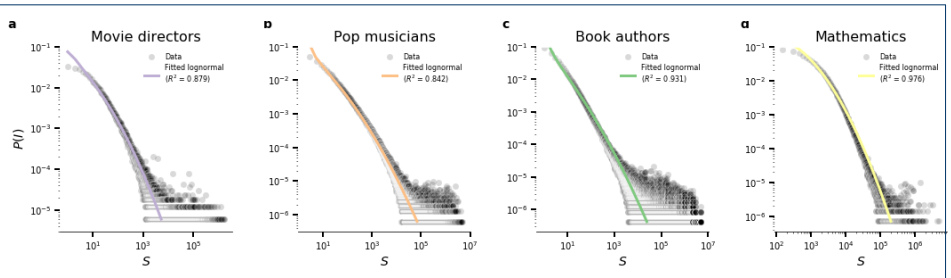

Figure S3 Log-normal fitting of the impact distributions. We report the impact distribution , namely movie directors, pop musicians, book authors, and mathematicians. The distribution is in light grey, and the fitted curve is a colored continuous line. We measure the goodness of the fit by the coefficient of determination

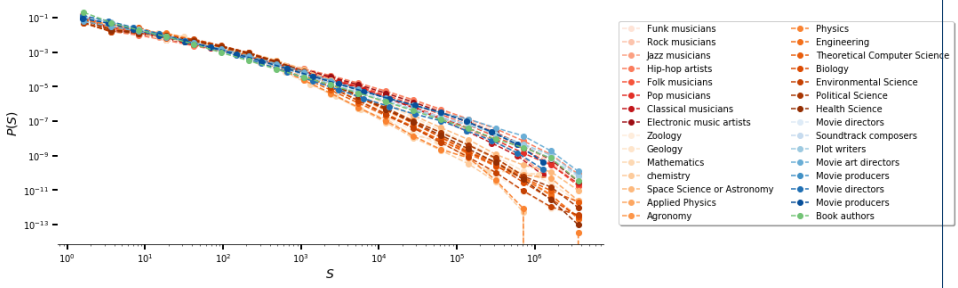

Figure S4 Re-scaled impact distributions. The min-max re-scaled impact distributions of the different fields.

Since different fields reach different audience, the impact of creative products across domains spans different ranges . In order to compare the decompositions of impacts across the Q and p components for several fields, a min-max scaling to the measured impacts. This transforms $P(S_a)$, the impact distribution of field a, in the following way:

$$P(S_a) \quad \rightarrow \quad \frac{P(S_a) - \min(P(S_a)}{\max(P(S_a)) - \min(P(S_a))} \cdot \max(P(S_c)), \tag{2}$$

where $P(S_c)$ denotes the distribution of all the fields combined. The re-scaled impact distributions of the different fields are visualized in Figure S4.

S2.3.2 Career length distributions

Figure S5 shows the log-normal distributions fitted on the career distributions for four selected, representative fields. The goodness of the fit is summarized in Table .

S2.3.3 $P(Q)$ and $P(p)$

The distributions $P(Q)$ and $P(p)$ described by log-normal functions, as illustrated by the fitted graphs on four representative fields on Figure S6, and the results being summarized in Table S1.

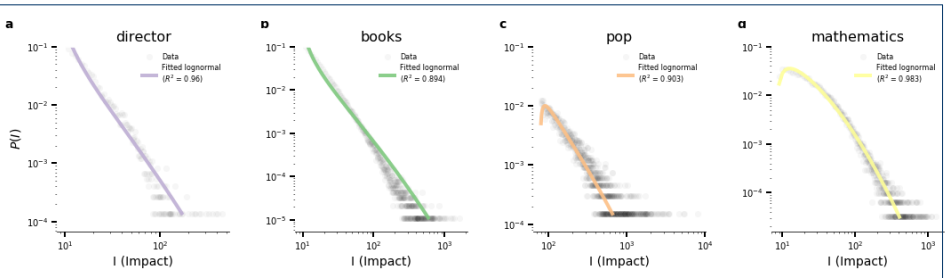

Figure S5 The career length distributions fitted by log-normal curves of four selected fields. We fit the productivity distribution (number of creative products each individual produced) by log-normal curves and characterized the goodness of fit by computing their R^2 values.

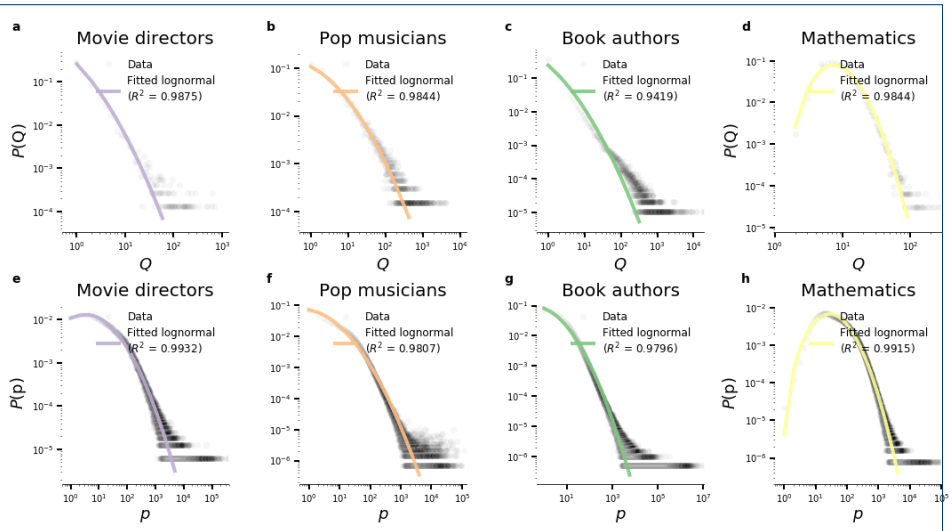

Figure S6 The distributions of Q and p in different creative domains. Based on the validation of Q-model in the studied fields, we are able to decompose the success of the works in the product of two parameters: the individual-based Q parameter, which only depends on the career trajectory of the individual and is a unique constant for everyone, and p, a probabilistic parameter which is a random number drawn from the same distribution for all the products on the same field. Figures illustrate how the values of the Q and p parameters (grey scatter plot and colored binned trend) are distributed in the different fields and how well they compare to the log-normal model.

S2.4 Comparison to the data

As the random-impact-rule holds, this means that the best product within a creative career occurs at random. However, can we say the same about the magnitude of the success of an individual's best hit? If each artistic product has the same probability to be the most successful, and success does not depend on any intrinsic ability of an individual, success will only be affected by its productivity. this hypothesis (black lines, Figure S8), known as the *R-model* [10], does not capture the observed patterns of impact in artistic domains (colored lines, Figure S8). This finding was first observed in Ref. [10] for scientific careers.

We tested model for the success of creative products in individual careers, based on the random impact rule, by generating sets of careers on each field based on the random impact rule. We then compared the highest impact in synthetic careers

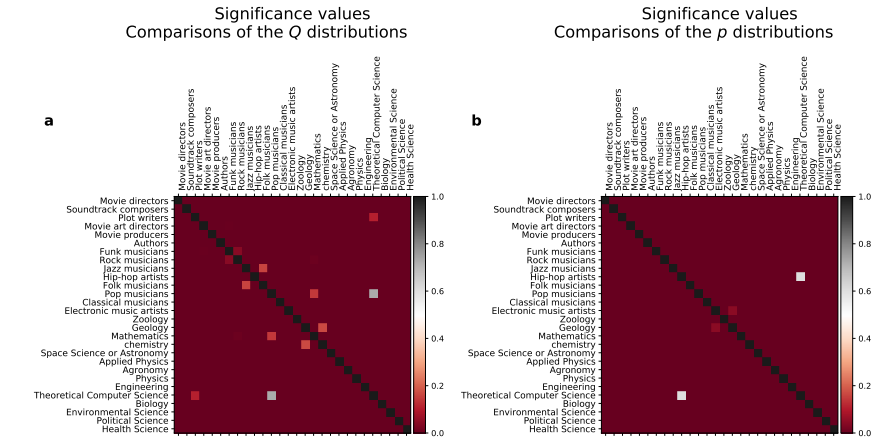

Figure S7

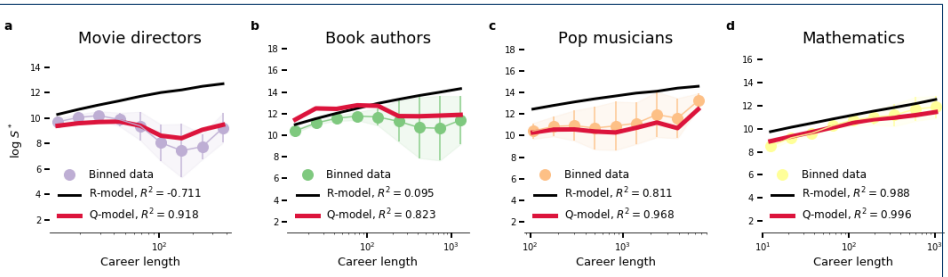

Figure S8 Validation of the Q-model. We compare the prediction of the Q-model in terms of the highest impact works of the individuals as the function of their productivity to the random model and to the original data.

to that of the observed data. To ensure that the set of synthetic careers is directly comparable to data, we constructed them by randomly reshuffling the time events of the careers found in the data, then repeated this random shuffling 100 times and averaged them to minimize the level of noise.

We also compared the expected highest impact of the individuals as a function of their productivity based on the Q-model. In order to do so we generated synthetic careers by combining the given career length N_i and measured Q_i parameter of the individual i, and randomly re-distributed the possible p_j parameters (picking exactly N_i p_j values for individual i) among them to compute the impacts of the synthetic careers by using the equation proposing the Q-model ($S_{i,\alpha} = Q_i p_{i,\alpha}$). After repeating this 100 times to minimize the noise level we arrived at a set of synthetic careers following the Q-model. We conducted this comparison on all the studied fields, for which the results are summarized in Table S3 .

Field	R^2
Jazz musicians	0.588
Funk musicians	0.636
Electronic music artists	0.7
Movie art directors	0.704266055671
Script writers	0.808
Soundtrack composer	0.812
Hip-hop artists	0.812
Book authors	0.823
Classical musicians	0.868
Rock musicians	0.871
Movie directors	0.918
Movie producers	0.925
Pop musicians	0.968
Agronomy	0.985
Environmental Science	0.988
Biology	0.991
Space Science or Astronomy	0.991
Zoology	0.992
Geology	0.993
Applied Physics	0.994
Engineering	0.994
Theoretical Computer Science	0.994
Chemistry	0.995
Mathematics	0.996
Physics	0.998
Political Science	0.999
Health Science	1.0

Table S3 Validation of the Q-model. Goodness of the fit of the Q-model for the different studied data sets expressed by the measured R^2 values.

Field	Correlation
Soundtrack composers	0.106
Plot writers	-0.005
Movie art directors	-0.144
Movie producers	0.063
Book authors	0.047
Funk musicians	0.022
Rock musicians	-0.029
Jazz musicians	0.101
Hip-hop artists	0.037
Folk musicians	0.067
Pop musicians	0.141
Classical musicians	0.184
Electronic music artists	0.063

Table S4

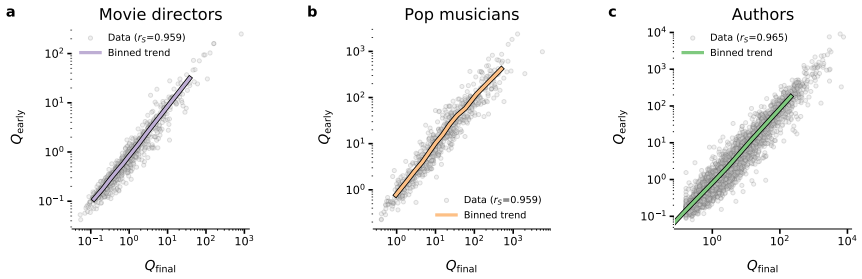

Figure S9

Field	Correlation
Classical musicians	0.956
Electronic music artists	0.947
Folk musicians	0.96
Funk musicians	0.953
Hip-hop artists	0.948
Jazz musicians	0.951
Movie art directors	0.96
Movie directors	0.972
Movie producers	0.981
Plot writers	0.974
Pop musicians	0.959
Rock musicians	0.968
Soundtrack composers	0.955

Table S5

S3 Randomness in networking

We tested the relationship between the collaboration network of an individual and her success for several creative fields (movie directors, pop musicians, mathematicians). Results show two different types of networking behavior. For one type of individual, their impact peaks first, and an increase in network centrality follows. For the others, the opposite is observed. Figure S11-S12 shows the distribution of the Q parameter and the impact S for these two groups of individuals, their network relevance measured by . Results show that there is no significant difference between the success patterns of these two groups, in Table S6-S7). the value of τ, the shifting parameters determined from the data associated with each individuals' career to the τ values we obtain in a randomized null-model data which we generate by reshuffling the original time series.

	Pop music		Mathematicians		Film directors	
	d	p	d	p	d	p
PageRank	0.048	<0.1	0.022	<0.001	0.02202	<0.1
Degree	0.032	<0.1	0.076	<0.001	0.05878	<0.001
Clustering	0.073	<0.01	0.054	<0.001	0.00662	<0.001
Strength	0.131	<0.001	0.203	<0.001	0.078	<0.001
Betweenness	0.240	<0.001	0.075	<0.001	0.113	<0.001
Closeness	0.099	<0.001	0.061	<0.001	0.094	<0.001
Constraint	0.088	<0.001	0.171	<0.001	0.032	<0.001
Coreness	0.065	<0.001	0.083	<0.001	0.039	<0.001

Table S6

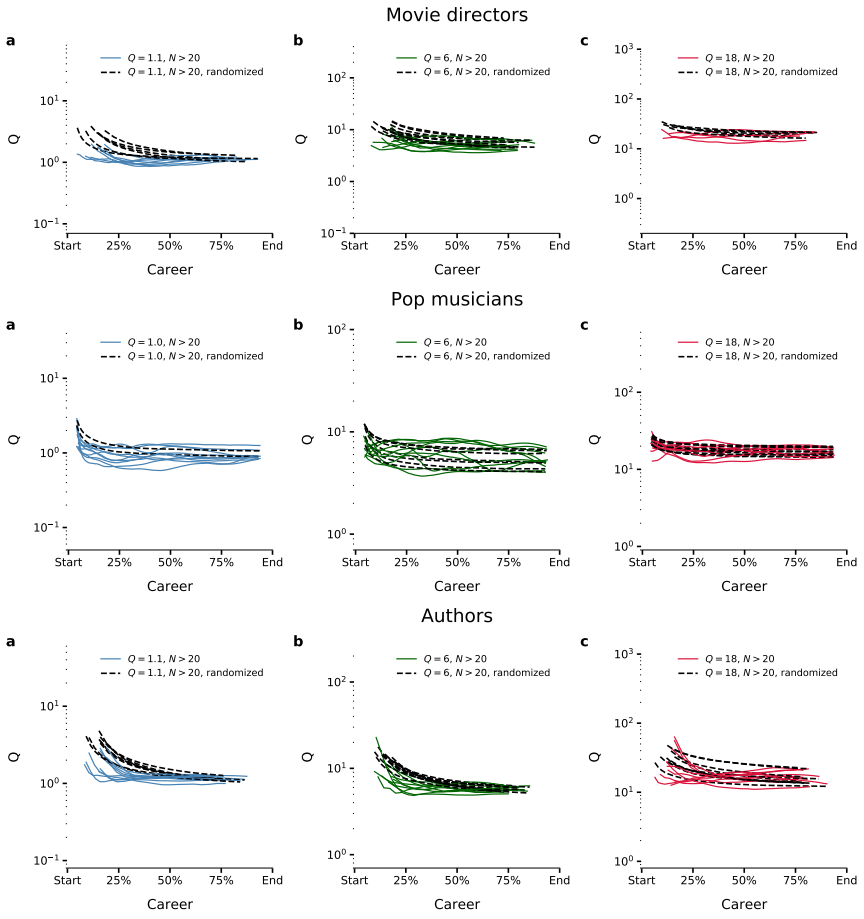

Figure S10 The evolution of Q over individual careers. The figure shows how the values of the Q parameter change over the course of a set of randomly chosen individual careers. We report a different field on each row, and within the field we look at individuals' with different average Q values, going from low (left) to high (right). For averaging Q over time, we use a moving window of $\Delta N = 10$ for movies, tracks, and books, respectively. The measured data is shown by coloured lines, while the results corresponding to the randomized case are shown by black. The figures show that the average Q-parameters stabilize no later than the first quarter of the careers.

	Pop music		Mathematicians		Film directors	
	d	p	d	p	d	p
PageRank	0.074	<0.001	0.039	<0.001	0.041	<0.001
Degree	0.029	<0.1	0.085	<0.001	0.068	<0.001
Clustering	0.073	<0.001	0.029	<0.01	0.057	<0.001
Strength	0.131	<0.001	0.203	<0.001	0.078	<0.001
Betweenness	0.24	<0.001	0.075	<0.001	0.113	<0.001
Closeness	0.099	<0.001	0.061	<0.001	0.094	<0.001
Constraint	0.088	<0.001	0.071	<0.001	0.032	<0.001
Coreness	0.065	<0.001	0.083	<0.001	0.039	<0.001

Table S7

	Pop music		Mathematicians		Film directors	
	d	p	d	p	d	p
PageRank	0.129	0.557	0.273	1.00E-05	0.142	0.60065
Degree	0.129	0.515	0.269	1.00E-05	0.121	0.79516
Clustering	0.114	0.698	0.292	0	0.158	0.47926
Strength	0.177	0.582	0.191	0.558	0.078	0.999
Betweenness	0.154	0.752	0.172	0.692	0.095	0.981
Closeness	0.143	0.839	0.156	0.772	0.068	1.000
Constraint	0.121	0.940	0.164	0.726	0.109	0.945
Coreness	0.146	0.805	0.179	0.613	0.059	1.000

Table S8 .

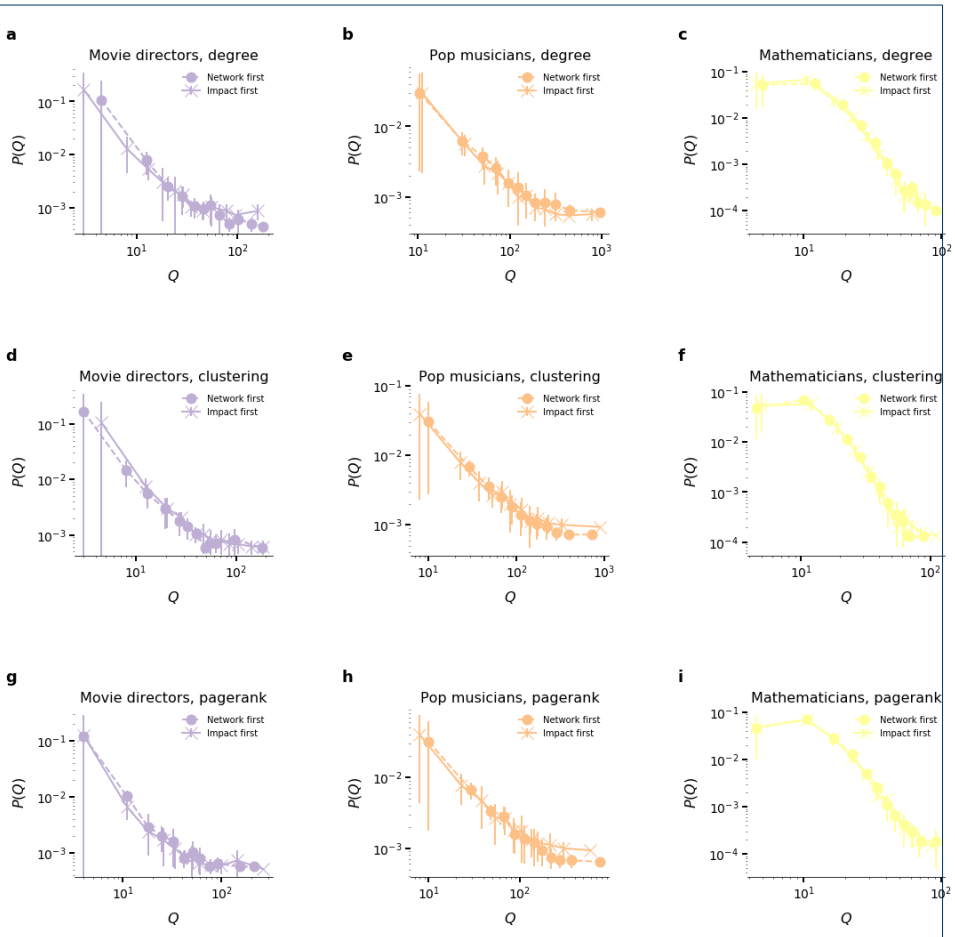

Figure S11 Q **distributions, where individuals are split based on their professions (director, pop musician, mathematician), while their network centralities are captured by their degree, PageRank, and clustering.**

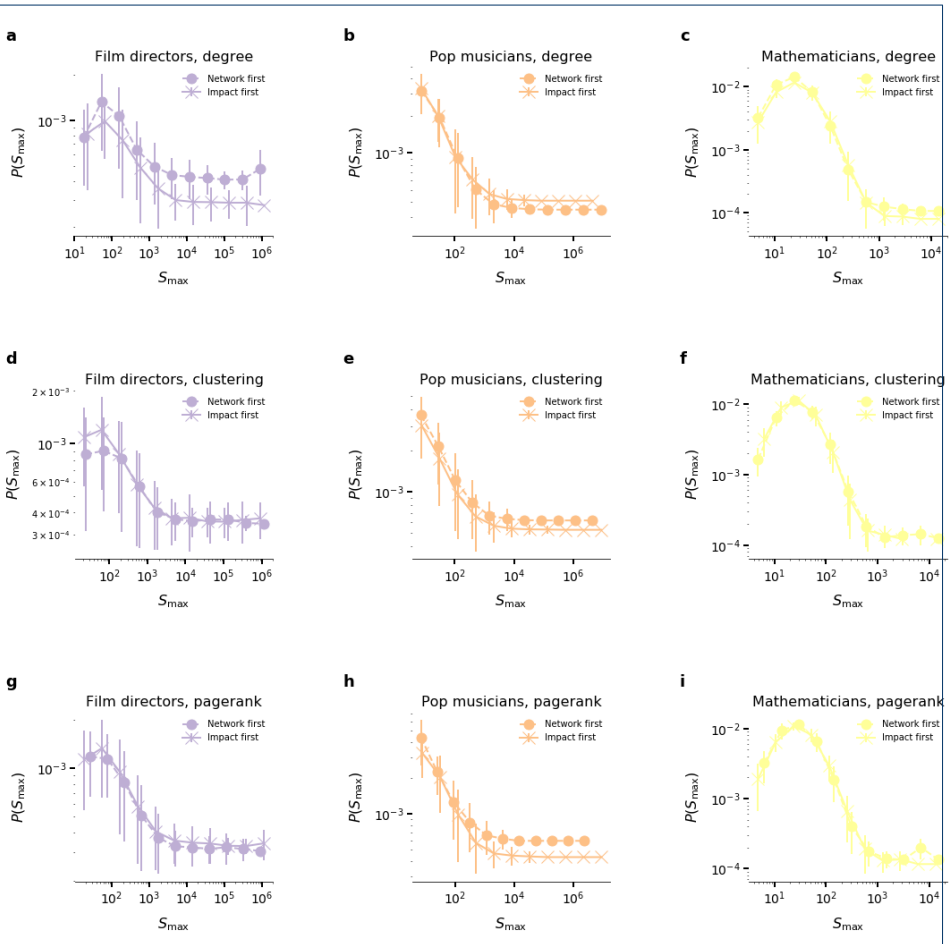

Figure S12 S **distributions, where individuals are split based on their professions (director, pop musician, mathematician), while their network centralities are captured by their degree, PageRank, and clustering.**

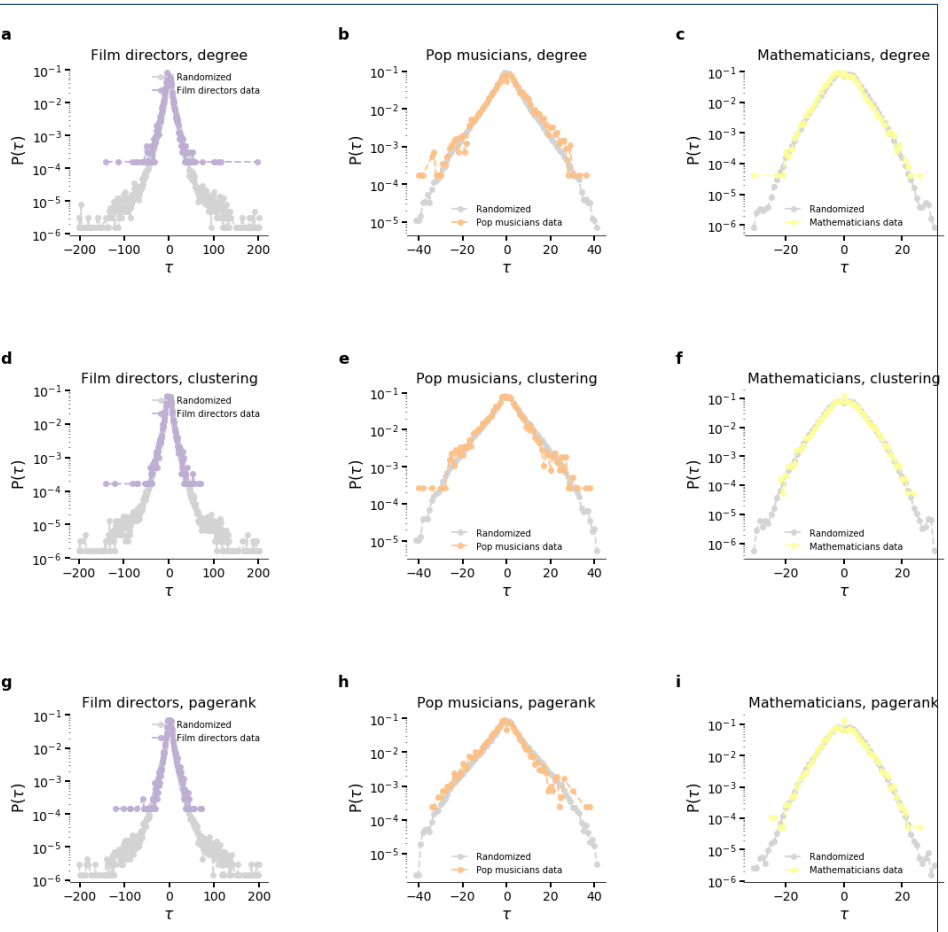

Figure S13 **Comparing the τ distributions between the randomized and the measured network data for different careers and network measures.**

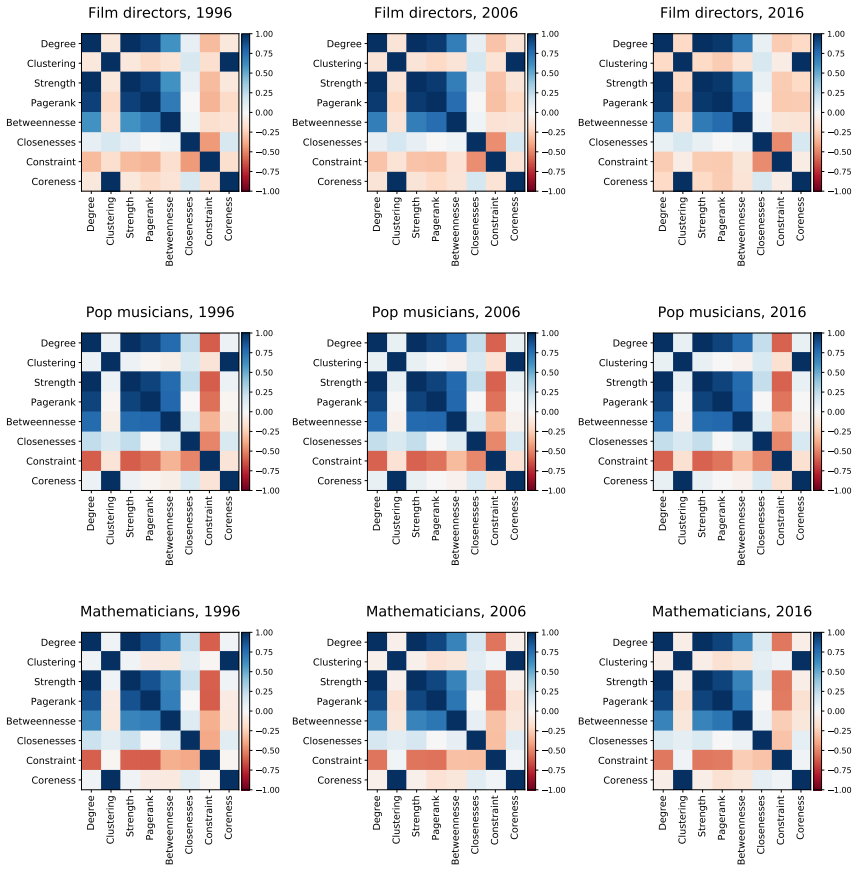
Film directors, 1996
Film directors, 2006
Film directors, 2016
Pop musicians, 1996
Pop musicians, 2006
Pop musicians, 2016
Mathematicians, 1996
Mathematicians, 2006
Mathematicians, 2016
Degree
Clustering
Strength
Pagerank
Betweennesse
Closenesses
Constraint
Coreness
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00

Figure S14

Author details
[1]Department of Network and Data Science, Central European University, 1051 Budapest, Hungary. [2]Department of Computer Science, IT University of Copenhagen, 2300 Copenhagen, Denmark. [3]ISI Foundation, 10126 Torino, Italy. [4]Complexity Science Hub Vienna, 1080 Vienna, Austria.

References

1. www.imdb.com: Internet Movie Database. Date accessed: 2017.02.04
2. http://www.metacritic.com: Metacritic database using expert's evaluations. Date accessed: 2017.02.04
3. www.discogs.com: Discogs music release database. Date accessed: 2017.02.04
4. www.last.fm: LastFM. Date accessed: 2017.02.06
5. www.goodreads.com: Goodreads book database. Date accessed: 2017.02.04
6. https://webofknowledge.com: Web of Science. Date accessed: 2018.11.06
7. Yucesoy, B., Barabási, A.-L.: Untangling performance from success. EPJ Data Science **5**(1), 17 (2016)
8. Barabási, A.-L.: The Formula: The Universal Laws of Success. Little, Brown and Company, Hachette UK (2018)
9. Muchnik, L., Aral, S., Taylor, S.J.: Social influence bias: A randomized experiment. Science **341**(6146), 647–651 (2013)
10. Sinatra, R., Wang, D., Deville, P., Song, C., Barabási, A.-L.: Quantifying the evolution of individual scientific impact. Science **354**(6312), 5239 (2016)
11. Liu, L., Wang, Y., Sinatra, R., Giles, C.L., Song, C., Wang, D.: Hot streaks in artistic, cultural, and scientific careers. Nature **559**(7714), 396 (2018)
12. Hansen, N., Ostermeier, A.: Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation. In: Evolutionary Computation, 1996., Proceedings of IEEE International Conference On, pp. 312–317 (1996). IEEE
13. Vásárhelyi, G., Virágh, C., Somorjai, G., Nepusz, T., Eiben, A.E., Vicsek, T.: Optimized flocking of autonomous drones in confined environments. Science Robotics **3**(20), 3536 (2018)