

**Supplementary Information for Bridging the short-term and long-term dynamics
of economic structural change**

James McNerney^{1,2,3,*}, Yang Li^{1,4}, Andres Gomez-Lievano^{1,5}, and Frank Neffke^{1,2}

¹Center for International Development, Kennedy School of Government, Harvard University,
Cambridge MA 02139, USA

²Complexity Science Hub, Vienna, Austria

³Lincoln Laboratory, Massachusetts Institute of Technology, Lexington MA, USA

⁴School of Resources and Environment, Nanchang University, Nanchang 330031, China

⁵Analysis Group, Inc., Boston MA 02199, USA

*james.mcnerney@ll.mit.edu

Contents

Supplementary Note 1: Cancellation of imaginary components in complex conjugate pairs	3
Supplementary Note 2: Export basket variation with HS data	4
Supplementary Note 3: Individual country trajectories and 2018 locations	5
Supplementary Note 4: Higher-order eigenvectors and eigenvalues	7
Supplementary Note 5: Correlations of coordinates with complexity metrics	9
Supplementary Note 6: Numerically comparing PCI*/PCI and ECI*/ECI across proximity measures	11
Supplementary Note 7: ECI* values of countries and regions	13
Supplementary Note 8: Comparing our dynamical modeling framework with co-determined complexity metrics	17
Supplementary Note 9: Complexity of oil extraction and artistic performance	18
Supplementary Note 10: Dimensionality reduction	19
Supplementary Note 11: Predicting country income and growth	21
Supplementary Note 12: Benefits of seeing complexity metrics as coordinates of structural change	24
Supplementary Note 13: Relationship between left and right eigenvectors	25
Supplementary Note 14: On the ECI* weights when $\Phi = \Phi^P$	27

Supplementary Note 1: Cancellation of imaginary components in complex conjugate pairs

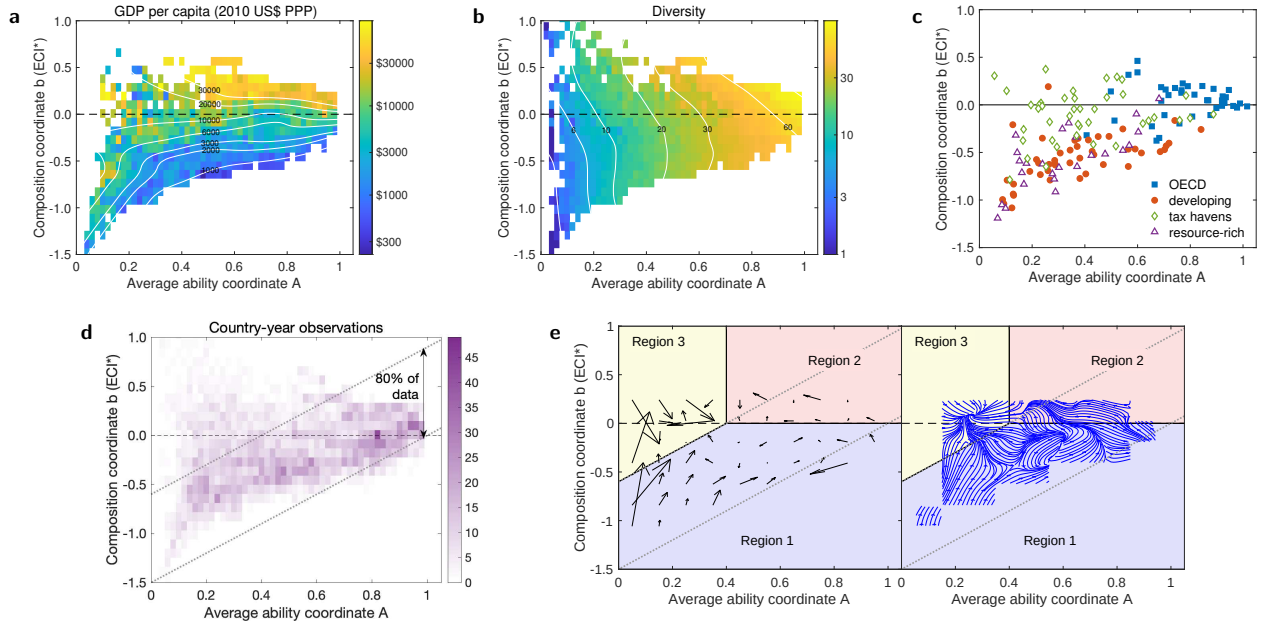
Typically the matrix Φ that maps the relatedness of different economic activities is constructed to be symmetric, which results in eigenvalues and eigenvectors of L_Φ that are real-valued. However, a symmetric Φ is not a requirement of our dynamical theory; here we discuss how complex eigenvalues and eigenvectors can be handled without any special difficulties in interpreting the eigenmodes. The eigenvalues and eigenvectors of real matrices occur in complex conjugate pairs and real-valued modes can be constructed with a standard trick by combining both components of the pair. As noted in the main text, the projection of region c onto the μ th eigenmode is given by $c_\mu(t) = \mathbf{w}_\mu^\dagger \mathbf{R}(t)$. If the elements of $\mathbf{w}_\mu = \mathbf{a}_\mu + i\mathbf{b}_\mu$ contain imaginary parts, then it has a paired eigenvector $\mathbf{w}_{\mu'} = \mathbf{w}_\mu^* = \mathbf{a}_\mu - i\mathbf{b}_\mu$, and wherever this is the case one may combine the eigenvectors and consider the coefficients in pairs, i.e.

$$c_\mu(t) + c_{\mu'}(t) = \mathbf{w}_\mu^\dagger \mathbf{R}(t) + \mathbf{w}_{\mu'}^\dagger \mathbf{R}(t) = [\mathbf{w}_\mu^\dagger + (\mathbf{w}_\mu^\dagger)^*] \mathbf{R}(t) \quad (1)$$

for which the summed vector $\mathbf{w}_\mu^\dagger + (\mathbf{w}_\mu^\dagger)^*$ (and therefore the combined projection onto both modes, $c_\mu(t) + c_{\mu'}(t)$) is manifestly real.

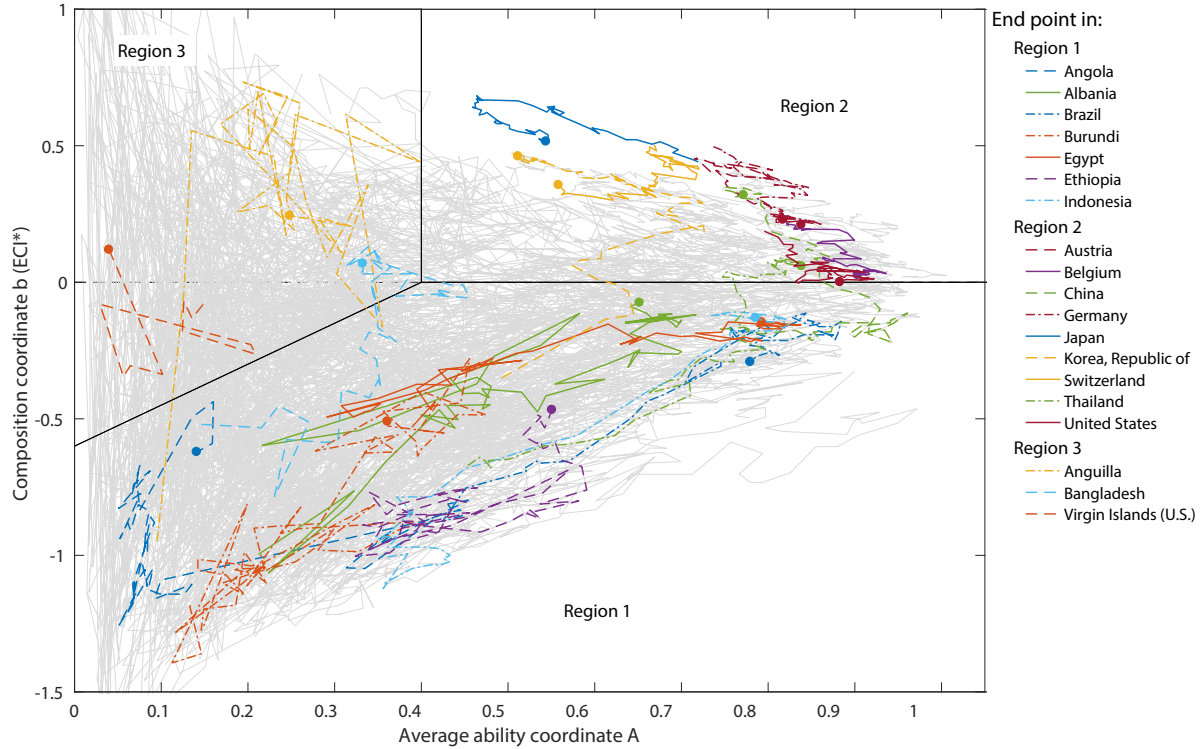
Supplementary Note 2: Export basket variation with HS data

Here we perform a robustness check to probe how country patterns in Fig. 3 of the main text are impacted by the choice of product classification scheme. Supplementary Fig. 1 depicts the same panels using HS classifications [242 countries over 1995 - 2018 (5494 observations) at the HS 3-digit level (173 product categories)]. HS classifications give very similar results to SITC. Income increases primarily as a function of the b coordinate when holding A fixed, while diversity increases primarily as a function of the A coordinate when holding b fixed (Supplementary Figs. 1a,b). We see a very similar arrangement of types of economies (Supplementary Fig. 1c) and concentration of data (Supplementary Fig. 1d), with 80% of export baskets occupying a diagonal band stretching from low A / low b to high A / above-average b . The movement of countries (Supplementary Figs. 1e) is also very similar, though significantly noisier, with slightly less than half the number of observations as in the SITC-based data. As in the SITC-based data, countries in Region 3 with positive values of b tend to move quickly toward lower levels, converging near zero. Countries in Region 2 tend to sustain a high b , and often move toward lower values of A . Countries in lower parts of Region 1 move in a lower-left-to-upper-right direction. The main difference is among countries in middle parts of Region 1, where in the 1995 - 2018 period a number of countries with A values between about 0.3 and 0.6 progressed toward lower values of A . This is due to the difference in time period (SITC data covers 1962 - 2018 while HS covers 1995 - 2018) rather than classification system, and movement patterns under the two classifications are very similar when confining SITC data to 1995 - 2018. (We note that while the A and b coordinates summarize a canonical development pathway toward higher income and product diversity, not every country has followed this path or done so in every time period.)



Supplementary Figure 1: Country export baskets based on HS classifications. We show baskets of 242 countries over 1995 - 2018 (5494 observations) at the HS 3-digit level (173 product categories). Compare with Fig. 3 from the main text based on SITC data. **a-b** GDP per capita and average diversity d within each bin. White lines are contours of a LOWESS-smoothed surface fit to bin averages. **c** Positions of export baskets in 2018 of four categories of economies. **d** Number of country-year observations in each bin. **e** Quiver and streamline plots for the average directional change over 20 years for countries beginning at different starting points in the plane. We use the same visualization parameters as in the main text and divide the plane into 10 equal-size bins horizontally and 12 equal-size bins vertically (120 bins total). For each bin we obtain the horizontal and vertical components of movement by computing the average horizontal change ΔA and average vertical change Δb over the next 20 years using all observations starting in the bin.

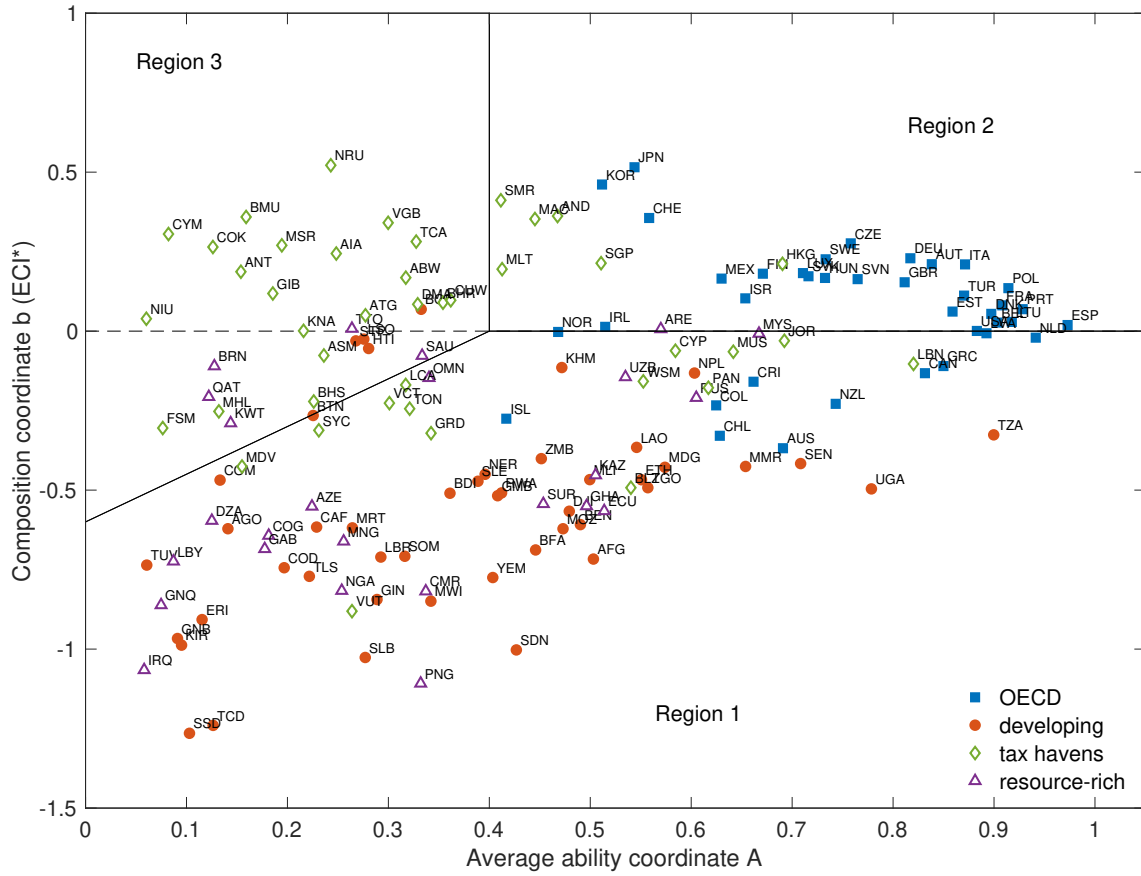
Supplementary Note 3: Individual country trajectories and 2018 locations



Supplementary Figure 2: Trajectories of selected countries in the A - b phase plane. The labels Region 1, Region 2, and Region 3 refer to the same parts of the plane as in Figure 4e of the main paper. Round markers identify the end point of each trajectory.

Supplementary Fig. 2 shows the trajectories of individual countries in our data from which the average directional movement depicted in Fig. 3e of the main text was derived. In some parts of this plane only a few countries have been observed, and in these sparse areas individual countries can dominate the behavior of directional lines in Fig. 3e. An instance of this small-numbers effect is seen in Japan's movements in Region 2.

Countries individually follow a wide variety of paths. For example, Indonesia follows the canonical trajectory of countries in Region 1, moving right and upward over the period of our data. Egypt for a long time moves back and forth in the plane, eventually moving upward toward the end of the period. Other countries in Region 1 such as Albania and Ethiopia similarly start with erratic paths, but then embark on steady development. China is a clear example of a canonical development trajectory in later stages, starting from $(A, b) = (0.744, -0.172)$ and moving right and upward until around 1981, when its A coordinate begins to fall while its b coordinate continues to rise, whereupon it looks more like European countries. Korea's trajectory is similar, moving right and upward until turning earlier in time (in 1975) and at a lower A value, after which its A coordinate decreases while its b coordinate increases. In Region 3 the erratic paths of tax havens (which are highly exposed to external factors, and may be less driven by internal development processes) is apparent as well. Supplementary Fig. 3 shows the ending positions of countries in the plane in 2018.

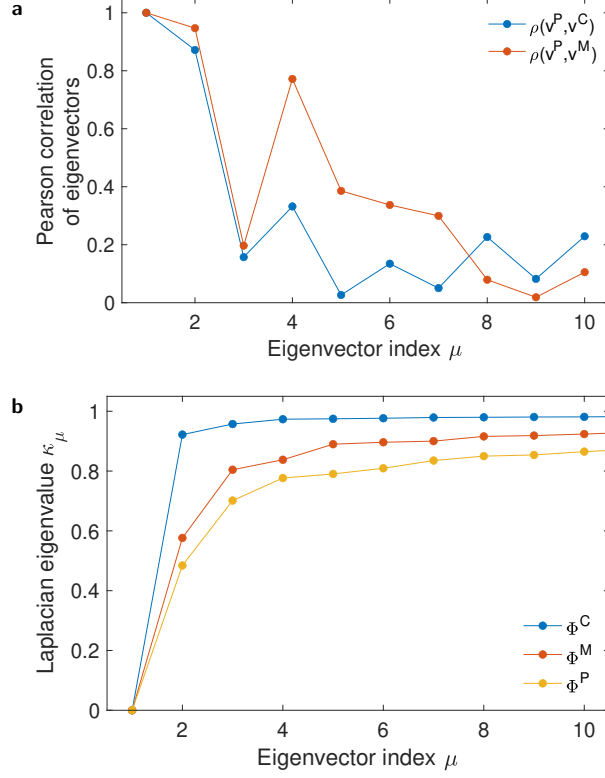


Supplementary Figure 3: Locations of countries in the A - b plane in 2018. Country codes are given in Table 1. Positions of export baskets in 2018 of four categories of economies. ‘Developing’ refers to economies on the UN’s least developed economies list [1], ‘tax havens’ to non-OECD and non-developing countries appearing in Hines (2010) [2], and ‘resource-rich’ to economies not elsewhere categorized whose natural resource rent [3] exceeded 5% of GDP in 2018.

Supplementary Note 4: Higher-order eigenvectors and eigenvalues

Here we examine the empirical structure of the right eigenvectors and eigenvalues of $\tilde{\Phi}$ for the three choices of proximity matrix Φ describe in the Methods (Φ^P , Φ^M , and Φ^C). Higher-order eigenvectors (i.e. those beyond the second eigenvalue) are of interest because in principle they could describe other modes of economic transformation besides the particular diversity and compositional changes that we study in the main text. However, we find that the proximity matrices in the literature only have strong agreement in the structure of the first two eigenvectors, and do not straightforwardly resolve how many eigenvectors matter. We row-normalize each proximity matrix to obtain the matrices $\tilde{\Phi}^P$, $\tilde{\Phi}^M$, and $\tilde{\Phi}^C$ and examine the eigenvectors of the resulting graph Laplacian L_{Φ} (i.e. $L_{\Phi} = I - \tilde{\Phi}^P$, $L_{\Phi} = I - \tilde{\Phi}^M$, or $L_{\Phi} = I - \tilde{\Phi}^C$). We put their eigenvectors in ascending order by eigenvalue κ_{μ} , and use the eigenvectors of $\tilde{\Phi}^P$ as a reference point, computing the correlation of the μ th eigenvector of $I - \tilde{\Phi}^P$ with the μ th eigenvector of $L_{\Phi} = I - \tilde{\Phi}^M$ or $L_{\Phi} = I - \tilde{\Phi}^C$ (Supplementary Fig. 4a). The first eigenvectors have a correlation $\rho = 1$ since $\tilde{\Phi}^P$, $\tilde{\Phi}^M$, and $\tilde{\Phi}^C$ share the same first right eigenvector, $\mathbf{v}_1 = \mathbf{1}$. The first two eigenvectors have very high correlations, but the correlations fall off beyond this (with a notable spike for the 4th eigenvector). We see related issues when we consider the eigenvalues (Supplementary Fig. 4b). The spectra derived from different proximity matrices differ not only in their values but also in the number of eigenvalues that appear significantly different from others.

Resolving the structure of higher-order eigenvectors and estimating how many eigenvectors appear relevant for long-term structural transformation means addressing the question of what the right proximity matrix is. Future work could, for example, investigate which proximity matrices show the highest predictive ability [4]. Alternatively, one could avoid directly constructing the proximity matrix Φ and adopt an agnostic approach such as Dynamical Mode Decomposition, which involves estimating the best-fitting linear operator that advances a system forward in time.



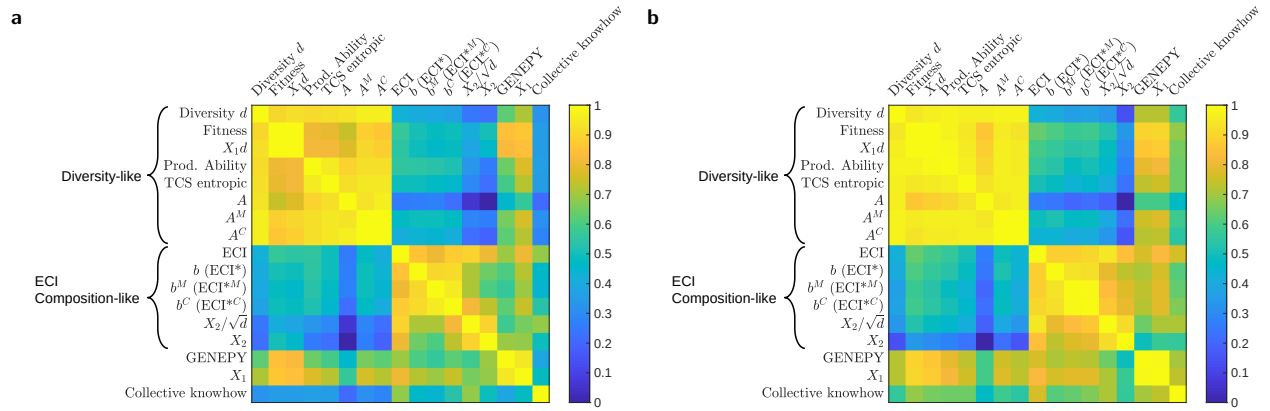
Supplementary Figure 4: Higher order eigenvectors of normalized proximity matrices. **a** Pearson correlations showing how eigenvectors of different proximity matrices become less similar to one another as one moves to higher-order eigenvectors. The series labeled $\rho(\mathbf{v}^P, \mathbf{v}^C)$ gives correlations between eigenvectors of $\tilde{\Phi}^P$ and $\tilde{\Phi}^C$, and the series labeled $\rho(\mathbf{v}^P, \mathbf{v}^M)$ gives correlations between eigenvector of $\tilde{\Phi}^P$ and $\tilde{\Phi}^M$. **b** Variation of the first 10 eigenvalues of the graph Laplacian L_Φ for the proximity metrics Φ^P , Φ^M , and Φ^C .

Supplementary Note 5: Correlations of coordinates with complexity metrics

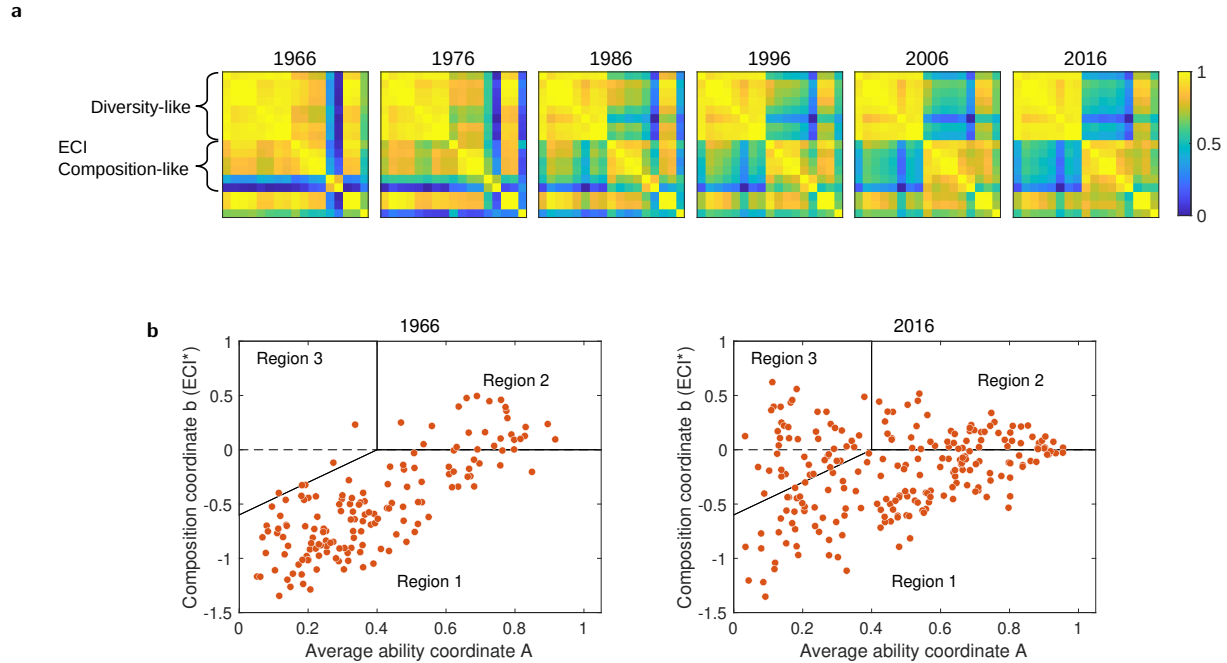
Here we further discuss connections of the A and b coordinates with complexity metrics. As noted in the main text, nearly all complexity metrics fall into one of two groups (Supplementary Fig. 5) that correspond to these two coordinates. The main text discusses theoretical and empirical similarities of the A coordinate with country Fitness [5] and of the b coordinate with the ECI [6]. Two other metrics, Production Ability [7] and the entropic measure of Teza, Caraglio, and Stella [8, 9], are also strongly correlated with the A coordinate. A recently introduced hybrid metric known as GENEPLY [10] contains aspects of both the ECI and country Fitness. GENEPLY is computed by first obtaining the two eigenvectors with the largest eigenvalues of a country-to-country similarity matrix. Letting $X_{c,1}$ and $X_{c,2}$ denote the elements of these eigenvectors, and letting λ_1 and λ_2 denote the eigenvalues, the GENEPLY of country c is $GENEPLY_c = \left(\sum_{i=1}^2 \lambda_i X_{c,i}^2 \right)^2 + 2 \sum_{i=1}^2 \lambda_i^2 X_{c,i}^2$. By design, $X_{c,1}$ is closely related to Fitness and $X_{c,2}$ is closely related to the ECI. Sciarra et al. (2020) show that, for example, when the country-country proximity matrix is $N_{cc'} \equiv \sum_p \frac{M_{cp} M_{c'p}}{d_c d_{c'} (k_p')^2}$ where $k_p' = \sum_c M_{cp} / k_c$, then $X_{c,1} d_c = F_c$, and $X_{c,2} / \sqrt{d_c}$ is strongly correlated with ECI_c. We find empirically that the $X_{c,1}$ and $X_{c,2}$ components underlying GENEPLY also have similar relationships to the A and b coordinates, i.e. $X_{c,1} d_c$ is highly correlated to A , and $X_{c,2} / \sqrt{d_c}$ is highly correlated to ECI*. Examining the bare $X_{c,1}$ and $X_{c,2}$ components by themselves, we find that $X_{c,2}$ resembles composition-like quantities (slightly less than $X_{c,2} / \sqrt{d}$), while $X_{c,1}$ does not fall in either group. GENEPLY strongly resembles its $X_{c,1}$ component (Spearman correlation 0.987), more so than $X_{c,2}$ (Spearman correlation 0.493). Of the metrics we examine, only collective knowhow [11] shows no similarly strong empirical relationship to the A and b coordinates. The existence of these two groups emphasizes that, when viewed through the analytical lens of our dynamical modeling framework, most existing complexity metrics emphasize different aspects of the same development process: either changes in economic diversity, or a particular pattern of compositional shifts in activities. However, as Fig. 2 in the main paper illustrated, both aspects of change are important characteristics of economic development.

Further comments are in order regarding the GENEPLY metric. Sciarra et al. highlight that GENEPLY merges information about countries from coordinates on two axes. Interestingly, we also arrived at a representation of country development on a 2D plane, using a very different starting point for our calculations. GENEPLY is motivated by the goal of reconciling different methods for inferring complexity from data within a consistent linear algebra and networks framework. In contrast, our 2D plane arose from our analysis of a dynamical model based on the Principle of Relatedness. We note three differences in the resulting planes that define each of our approaches. First, as noted already above, our A and b coordinates correspond most closely to $X_{c,1} d_c$ and $X_{c,2} / \sqrt{d_c}$ rather than Sciarra et al.'s $X_{c,1}$ and $X_{c,2}$ themselves. Second, a fundamental input in the method of Sciarra et al. is a country-country proximity matrix $N_{cc'}$. The elements of the eigenvectors of this matrix then directly give countries' coordinates in the plane. In contrast, the fundamental input in the framework here is a product-product proximity matrix $\Phi_{pp'}$. Our approach comes with a two-step process: Eigenvectors derived from Φ first define directions of movement, and then country coordinates are obtained from projections onto these axes. Third, Sciarra et al. interpret their coordinates in terms of their relationship to Fitness and the ECI. Here, we exploit the fact that, by analyzing the changes in the world economy captured by movement in the directions $\mathbf{v}_1$ and $\mathbf{v}_2$, we can give economic meaning to the coordinates A and b (ECI*) that define our 2D plane, as capturing changes in average ability or activity composition.

Correlations within the two groups of diversity-like and composition-like metrics were high across time (Supplementary Fig. 6a). In addition, we note that correlations across these two groups are high in early years of our data, where regions overwhelmingly lie in a diagonal band (Supplementary Fig. 6b) such that a region's A coordinate is strongly related to its b coordinate. We are mainly interested in correlations within each group of metrics, and do not determine the extent to which these changes in cross-group correlations derive from aspects of our data (e.g. new countries becoming available in our dataset over time) or from changes in the nature of development. An exception to the high within-group correlations is X_2 (and $X_2 / \sqrt{d}$), which in early years is weakly correlated with other compositional quantities. GENEPLY was originally studied on data in the years 1995 - 2017, and thus we are examining this quantity in a period well outside the one in which it was tested.



Supplementary Figure 5: Correlations of A and b (ECI*) coordinates and complexity measures. The figure shows several proposed measures of complexity (ECI, Fitness, Production Ability, the entropic measure of Teza, Caraglio, and Stella, and GENEPI), and related quantities. **a** Pearson correlations. **b** Spearman rank correlations (identical to Fig. 5 in the main text). The coordinate A is our average ability coordinate used throughout the text, defined using the proximity matrix Φ^P , while A^M and A^C refer to the same coordinate obtained from using the proximity matrices Φ^M and Φ^C respectively. All quantities were computed for the year 2016 using UN Comtrade data.

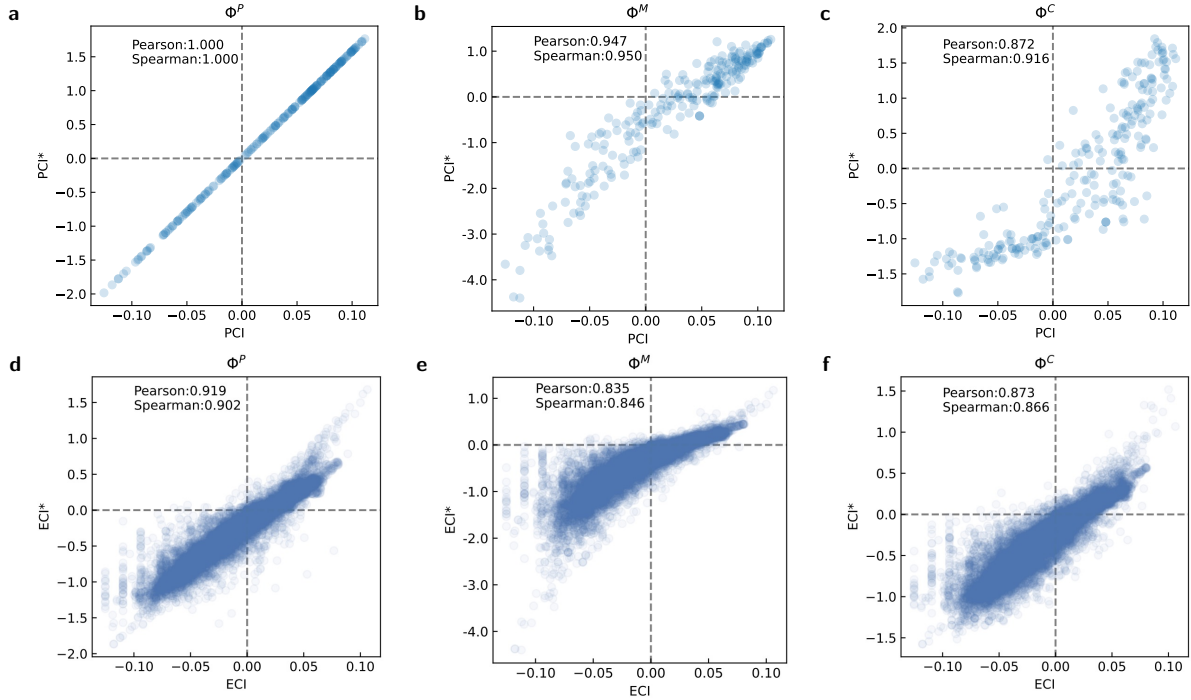


Supplementary Figure 6: Correlations of coordinates and complexity measures over time. **a** Correlations over time among our A and b (ECI*) coordinates, complexity metrics, and related quantities. **b** Scatter of regions in the A - b plane in 1966 and 2016.

Supplementary Note 6: Numerically comparing PCI*/PCI and ECI*/ECI across proximity measures

The PCI* is defined by the second eigenvector of L_Φ , which means that it will vary depending on how the proximity matrix Φ between activities is constructed. This matrix has been implemented in a variety of ways that all find empirical support. To compare the PCI* and ECI* with the PCI and ECI numerically we therefore compute the former using several choices of proximity matrix found in the literature. (The conventional PCI is always computed from the usual Φ^P matrix.) We consider three proximity matrices. The first is Φ^P , with $\Phi_{pp'}^P \equiv \sum_c \frac{M_{cp}M_{cp'}}{D_c}$. The second is the minimum conditional probability-based [20] proximity measure Φ^M noted in the main text, $\Phi_{pp'}^M \equiv \min\{(M^T M U^{-1})_{pp'}, (U^{-1} M^T M)_{pp'}\}$, and the third is a correlation-based proximity measure [21] Φ^C , $\Phi_{pp'}^C \equiv \frac{1}{2}(1 + \rho_{pp'})$, where $\rho_{pp'}$ is the Pearson correlation of the RCAs for products p and p' across locations c . Supplementary Fig. 7 compares the PCI* to the PCI and the ECI* to the ECI for each proximity matrix. The general structure of the second eigenvector is not sensitive to the choice of proximity matrix and consistently resembles the vector of conventional PCIs. Similarly, the ECI* numerically resembles the conventional ECI, which is not surprising because the ECI* depends on the PCI*. Pearson and Spearman correlations are shown for each comparison, all with high values between 0.83 and 1. Panels (a) and (d) show the comparison in the special case $\Phi = \Phi^P$ discussed above. As noted already, when proximities are constructed using Φ^P , the PCI* and PCI are equal up to a constant overall factor. Supplementary Fig. 7d is the comparison of ECI* with ECI in this case, and displays the same results as Fig. 4. While the **PCI*** and **PCI** coincide in one special case ($\Phi = \Phi^P$), the same is never true of the ECI* and ECI. Differences between these quantities will always remain because of the different averaging weights used in equations 8 and 9: $\text{ECI}_c = \sum_p \left(\frac{M_{cp}}{\sum_{p'} M_{cp'}} \right) \text{PCI}_p$ and $\text{ECI}_c^* = \sum_p \left(\frac{R_{cp} \pi_p}{\sum_{p'} R_{cp'} \pi_{p'}} \right) \text{PCI}_p^*$. In the closest approach, where $\Phi = \Phi^P$, the Pearson correlation between the ECI* and the ECI is 0.919, and the Spearman rank correlation is 0.902 (Supplementary Fig. 7d). In Supplementary Note 14 we discuss why the weights used to calculate ECI* may have certain desirable properties.

Although the closest approach of the $b(t)$ coordinate to the conventional ECI occurs when the proximity matrix is taken to be Φ^P , there is no inherent reason to assume this particular matrix. Proximities between activities could be quantified with other existing matrices, or entirely new ones to be developed, and could be selected based on the quality of forecasts of future transitions in activities, rather than on arguments about how best to infer complexity.



Supplementary Figure 7: Comparison of PCI* with PCI (a-c) and ECI* with ECI (d-f) for the proximity matrices Φ^P , Φ^M , and Φ^C . The PCI* is defined by the second eigenvector of L_Φ , which means that it will vary depending on how the proximity matrix Φ between activities is constructed. This matrix has been implemented in a variety of ways that all find empirical support in regression models of country diversification. Therefore, to compare the PCI* and ECI* with the PCI and ECI numerically, we compute the former using several choices of proximity matrix found in the literature. The general structure of the second eigenvector is not sensitive to the choice of proximity matrix and consistently resembles the vector of conventional PCIs. Similarly, the ECI* numerically resembles the conventional ECI.

Supplementary Note 7: ECI* values of countries and regions

Below we show the b coordinate / ECI* values for countries and regions in our data with populations above 10,000 (Table 1). As noted in the main text, an economy's ECI* indicates how far along the compositional b axis it is, or equivalently, how much its activity basket is weighted toward products that load positively on the $\mathbf{v}_2$ eigenvector of $\tilde{\Phi}$. (As noted in the main text, positive elements of $\mathbf{v}_2$ tend to be manufacturing products, while negative elements tend to be agriculture products.) The ranking of countries by complexity metrics often attracts attention from interest in knowing what countries have the most sophisticated economies. As we emphasized though, our derivation of ECI* gives no reason to interpret it as a complexity metric, but as an indication of how much a country weights its activity basket towards particular goods. Beyond this, it is not immediately clear what additional significance the ranking below has. As noted in the main text, the ECI* strongly correlates with the conventional ECI and ranks countries similarly.

Supplementary Table 1: ECI* of countries and regions in 2018.

Rank	Code	Region	ECI*	ECI	ECI rank
1	JPN	Japan	0.4589	0.0694	1
2	KOR	Korea, Republic of	0.4005	0.0571	3
3	CHE	Switzerland	0.3380	0.0494	5
4	CZE	Czechia	0.2683	0.0501	4
5	SGP	Singapore	0.2600	0.0456	6
6	CHN	China	0.2433	0.0430	7
7	DEU	Germany	0.2278	0.0578	2
8	HUN	Hungary	0.2252	0.0348	12
9	AUT	Austria	0.2011	0.0401	9
10	ITA	Italy	0.1961	0.0332	16
11	SVK	Slovakia	0.1949	0.0352	11
12	SVN	Slovenia	0.1754	0.0369	10
13	SWE	Sweden	0.1637	0.0331	17
14	FIN	Finland	0.1627	0.0234	23
15	MAC	Macao	0.1616	0.0178	29
16	HKG	Hong Kong	0.1602	0.0341	14
17	MEX	Mexico	0.1570	0.0412	8
18	LUX	Luxembourg	0.1505	0.0230	25
19	MLT	Malta	0.1440	0.0233	24
20	ROU	Romania	0.1426	0.0283	19
21	STP	Sao Tome and Principe	0.1421	0.0273	22
22	GBR	United Kingdom of Great Britain and Northern Ireland	0.1364	0.0347	13
23	POL	Poland	0.1056	0.0216	26
24	IRL	Ireland	0.1050	0.0281	20
25	FRA	France	0.1036	0.0281	20
26	PRK	Korea (Democratic People's Republic of)	0.1026	0.0160	31
27	THA	Thailand	0.0920	0.0333	15
28	ISR	Israel	0.0899	0.0182	28
29	DNK	Denmark	0.0669	0.0169	30
30	CUW	Curaçao	0.0643	0.0124	36
31	USA	United States of America	0.0582	0.0288	18
32	ABW	Aruba	0.0436	-0.0059	61
33	BLR	Belarus	0.0388	0.0108	39
34	EST	Estonia	0.0375	0.0103	42
35	PRT	Portugal	0.0333	0.0129	34
36	TUR	Turkey	0.0319	0.0147	32
37	BEL	Belgium	0.0253	0.0129	34
38	LTU	Lithuania	0.0241	0.0064	48

39	PHL	Philippines	0.0138	0.0207	27
40	SRB	Serbia	0.0131	0.0069	46
41	ESP	Spain	0.0111	0.0129	34
42	BRB	Barbados	0.0103	0.0064	48
43	HRV	Croatia	0.0100	0.0106	41
44	MYS	Malaysia	0.0089	0.0121	38
45	NLD	Netherlands	0.0084	0.0124	36
46	BIH	Bosnia and Herzegovina	0.0083	0.0086	43
47	GUM	Guam	0.0030	0.0034	52
48	BGR	Bulgaria	-0.0067	0.0081	44
49	TUN	Tunisia	-0.0191	0.0065	47
50	LVA	Latvia	-0.0194	0.0023	53
51	MKD	North Macedonia	-0.0336	0.0074	45
52	IND	India	-0.0418	0.0012	54
53	VNM	Viet Nam	-0.0568	0.0107	40
54	BHR	Bahrain	-0.0673	-0.0167	86
55	PSE	Palestine, State of	-0.0783	0.0035	51
56	SLV	El Salvador	-0.0792	-0.0011	55
57	ARE	United Arab Emirates	-0.0804	-0.0211	103
58	LBN	Lebanon	-0.0920	-0.0050	59
59	SWZ	Eswatini	-0.0983	-0.0054	60
60	MNE	Montenegro	-0.1190	-0.0015	56
61	CAN	Canada	-0.1197	-0.0196	92
62	JOR	Jordan	-0.1219	-0.0113	74
63	CRI	Costa Rica	-0.1248	-0.0039	57
64	CYP	Cyprus	-0.1264	-0.0076	67
65	GRC	Greece	-0.1313	-0.0069	65
66	WSM	Samoa	-0.1346	-0.0072	66
67	UKR	Ukraine	-0.1425	-0.0110	72
68	NPL	Nepal	-0.1450	-0.0048	58
69	MDA	Moldova, Republic of	-0.1453	-0.0139	82
70	IDN	Indonesia	-0.1583	-0.0115	76
71	NOR	Norway	-0.1588	-0.0321	126
72	ZAF	South Africa	-0.1593	-0.0165	84
73	BWA	Botswana	-0.1647	-0.0233	110
74	DOM	Dominican Republic	-0.1733	-0.0091	68
75	BRA	Brazil	-0.1789	-0.0211	103
76	FJI	Fiji	-0.1837	-0.0115	76
77	MUS	Mauritius	-0.1839	-0.0093	69
78	KGZ	Kyrgyzstan	-0.1905	-0.0120	79
79	LKA	Sri Lanka	-0.1945	-0.0060	62
80	GEO	Georgia	-0.2032	-0.0165	84
81	NZL	New Zealand	-0.2082	-0.0234	111
82	ARM	Armenia	-0.2092	-0.0206	102
83	PAN	Panama	-0.2100	-0.0213	105
84	SAU	Saudi Arabia	-0.2200	-0.0068	64
85	LCA	Saint Lucia	-0.2253	-0.0104	71
86	TON	Tonga	-0.2269	0.0051	50
87	COL	Colombia	-0.2332	-0.0320	125
88	MAR	Morocco	-0.2461	-0.0216	107
89	EGY	Egypt	-0.2474	-0.0129	80
90	ALB	Albania	-0.2543	-0.0198	96
91	BGD	Bangladesh	-0.2639	-0.0103	70
92	TTO	Trinidad and Tobago	-0.2654	-0.0424	148

93	BTN	Bhutan	-0.2727	-0.0170	87
94	CPV	Cabo Verde	-0.2729	-0.0198	96
95	ISL	Iceland	-0.2750	-0.0201	99
96	RUS	Russian Federation	-0.2779	-0.0310	123
97	LSO	Lesotho	-0.2796	-0.0174	88
98	KEN	Kenya	-0.2834	-0.0227	109
99	PAK	Pakistan	-0.2841	-0.0204	101
100	GRD	Grenada	-0.2923	-0.0116	78
101	KHM	Cambodia	-0.2939	-0.0110	72
102	TZA	Tanzania, United Republic of	-0.2975	-0.0131	81
103	URY	Uruguay	-0.3025	-0.0180	90
104	UZB	Uzbekistan	-0.3167	-0.0202	100
105	ARG	Argentina	-0.3270	-0.0269	115
106	BHS	Bahamas	-0.3365	-0.0143	83
107	COM	Comoros	-0.3374	-0.0197	94
108	NER	Niger	-0.3419	-0.0386	137
109	AUS	Australia	-0.3464	-0.0386	137
110	PYF	French Polynesia	-0.3511	-0.0215	106
111	SLE	Sierra Leone	-0.3535	-0.0472	153
112	CHL	Chile	-0.3556	-0.0317	124
113	GTM	Guatemala	-0.3571	-0.0194	91
114	SYR	Syrian Arab Republic	-0.3582	-0.0197	94
115	VCT	Saint Vincent and the Grenadines	-0.3606	-0.0200	98
116	HND	Honduras	-0.3631	-0.0269	115
117	BRN	Brunei Darussalam	-0.3683	-0.0918	184
118	OMN	Oman	-0.3711	-0.0503	156
119	HTI	Haiti	-0.3732	-0.0305	121
120	ETH	Ethiopia	-0.3797	-0.0284	118
121	DJI	Djibouti	-0.3803	-0.0196	92
122	ZMB	Zambia	-0.3856	-0.0264	114
123	NIC	Nicaragua	-0.3945	-0.0357	131
124	UGA	Uganda	-0.3992	-0.0254	113
125	BDI	Burundi	-0.4034	-0.0174	88
126	BLZ	Belize	-0.4161	-0.0226	108
127	LAO	Lao People's Democratic Republic	-0.4207	-0.0292	119
128	NAM	Namibia	-0.4240	-0.0276	117
129	JAM	Jamaica	-0.4333	-0.0377	136
130	MLI	Mali	-0.4339	-0.0239	112
131	NCL	New Caledonia	-0.4410	-0.0115	76
132	SEN	Senegal	-0.4478	-0.0337	128
133	PRY	Paraguay	-0.4492	-0.0301	120
134	BFA	Burkina Faso	-0.4578	-0.0415	144
135	PER	Peru	-0.4605	-0.0418	146
136	RWA	Rwanda	-0.4705	-0.0354	130
137	TGO	Togo	-0.4886	-0.0333	127
138	FSM	Micronesia (Federated States of)	-0.4959	-0.0063	63
139	CIV	Côte d'Ivoire	-0.4989	-0.0344	129
140	BEN	Benin	-0.5061	-0.0400	140
141	SUR	Suriname	-0.5084	-0.0361	132
142	MDG	Madagascar	-0.5087	-0.0369	134
143	KWT	Kuwait	-0.5104	-0.0757	177
144	MMR	Myanmar	-0.5172	-0.0364	133
145	KAZ	Kazakhstan	-0.5267	-0.0526	158
146	TJK	Tajikistan	-0.5369	-0.0306	122

147	MDV	Maldives	-0.5456	-0.0528	159
148	TLS	Timor-Leste	-0.5540	-0.0561	162
149	IRN	Iran (Islamic Republic of)	-0.5614	-0.0465	151
150	GHA	Ghana	-0.5620	-0.0549	161
151	ZWE	Zimbabwe	-0.5625	-0.0455	150
152	GUY	Guyana	-0.5647	-0.0405	142
153	ECU	Ecuador	-0.5865	-0.0561	162
154	CAF	Central African Republic	-0.5976	-0.0400	140
155	CUB	Cuba	-0.6024	-0.0397	139
156	GMB	Gambia	-0.6047	-0.0375	135
157	BOL	Bolivia (Plurinational State of)	-0.6095	-0.0485	154
158	AFG	Afghanistan	-0.6204	-0.0419	147
159	QAT	Qatar	-0.6373	-0.0412	143
160	AZE	Azerbaijan	-0.6391	-0.0775	178
161	MNG	Mongolia	-0.6781	-0.0465	151
162	MWI	Malawi	-0.6794	-0.0582	166
163	KIR	Kiribati	-0.6813	-0.0417	145
164	MOZ	Mozambique	-0.6885	-0.0569	164
165	SOM	Somalia	-0.6918	-0.0433	149
166	LBR	Liberia	-0.6919	-0.0639	171
167	TKM	Turkmenistan	-0.7012	-0.0508	157
168	COG	Congo	-0.7116	-0.0603	167
169	VUT	Vanuatu	-0.7418	-0.0573	165
170	VEN	Venezuela (Bolivarian Republic of)	-0.7744	-0.0668	174
171	MRT	Mauritania	-0.7812	-0.0653	173
172	CMR	Cameroon	-0.7945	-0.0734	176
173	GAB	Gabon	-0.8041	-0.0816	180
174	COD	Congo, Democratic Republic of the	-0.8174	-0.0502	155
175	YEM	Yemen	-0.8281	-0.0535	160
176	GIN	Guinea	-0.8386	-0.0725	175
177	AGO	Angola	-0.8878	-0.0835	182
178	SLB	Solomon Islands	-0.8900	-0.0647	172
179	ERI	Eritrea	-0.9048	-0.0628	169
180	SDN	Sudan	-0.9184	-0.0636	170
181	NGA	Nigeria	-0.9403	-0.0817	181
182	DZA	Algeria	-0.9439	-0.0815	179
183	GNB	Guinea-Bissau	-0.9880	-0.0620	168
184	PNG	Papua New Guinea	-1.0346	-0.0870	183
185	LBY	Libya	-1.1201	-0.1116	186
186	IRQ	Iraq	-1.2961	-0.2211	189
187	TCD	Chad	-1.3215	-0.1132	187
188	GNQ	Equatorial Guinea	-1.3545	-0.1137	188
189	SSD	South Sudan	-1.4068	-0.1038	185

Supplementary Note 8: Comparing our dynamical modeling framework with co-determined complexity metrics

Several complexity metrics emerge from taking country and product metrics to be co-determined by a system of equations [6, 5, 10], i.e.

$$\text{country complexity} = F(\text{product complexities}) \quad (2)$$

$$\text{product complexity} = G(\text{country complexities}). \quad (3)$$

For example, one obtains the ECI by taking F as an arithmetic average complexity of the products that a place exports competitively (i.e. with $\text{RCA} > 1$), and taking G as an arithmetic average complexity of the countries that competitively export a given product. Alternatively, one reaches country Fitness by replacing G with a harmonic average. Here, we elaborate on how the mathematical structure of our framework differs from the approach above.

First, as noted in the main text, the calculations of the A and b coordinates are functions of product variables, and in this way our framework contains an equation that very closely resembles supplementary equation (2). (Although we interpret these variables differently and not as complexity measures.)

However, our framework does not in general have an analog to supplementary equation (3). To see this, consider the calculation of the ECI and PCI. Let M , D , and U be the matrices of binarized RCAs, diversities, and ubiquities respectively (see Methods, Review of conventional PCI and ECI). The calculations of the ECI and PCI satisfy

$$\mathbf{ECI} \propto \bar{F} \mathbf{PCI} \quad (4)$$

$$\mathbf{PCI} \propto \bar{G} \mathbf{ECI} \quad (5)$$

where $\bar{F} \equiv D^{-1}M$ and $\bar{G} \equiv U^{-1}M$ are matrices. Plugging supplementary equation (4) into supplementary equation (5) gives

$$\mathbf{PCI} \propto U^{-1}M^T D^{-1}M \mathbf{PCI}. \quad (6)$$

This is an eigenvector equation, with the PCI vector taken by definition to be the second eigenvector of $\bar{G}\bar{F} = U^{-1}M^T D^{-1}M$.

The corresponding relationship in our dynamical framework is

$$\mathbf{PCI}^* \propto \tilde{\Phi} \mathbf{PCI}^*, \quad (7)$$

where $\tilde{\Phi}$ is the row-normalized proximity matrix between products used in our dynamical model (main text equation (2)), and $\mathbf{PCI}^*$ is its second eigenvector. The key point is that $\tilde{\Phi}$ can be constructed in many ways, which in general differ from $U^{-1}M^T D^{-1}M$. If $\tilde{\Phi}$ is constructed as $U^{-1}M^T D^{-1}M$ (the case $\Phi = \Phi^P$) then the vectors $\mathbf{PCI}^*$ and $\mathbf{PCI}$ will be in exact proportion to one another (Fig. 7a of the main text). However, many other constructions of Φ are possible (such as the cases $\Phi = \Phi^C$ and $\Phi = \Phi^M$, Figs. 7b and 7c). A given construction of the proximity matrix can be assessed by how accurately it predicts product transitions. However selected, this matrix is then the starting point for downstream calculations of other product and country quantities in our framework (main text Fig. 5a), and in general will differ from the results of combining supplementary equations (4)-(5) (such as the matrix $U^{-1}M^T D^{-1}M$ above in the case of the conventional PCI and ECI). In principle Φ could even be constructed with a different dataset from the one that is used to project countries onto the eigenvectors. For example, one could conceivably use patent data to estimate relatedness between products by examining correspondences between products and the technology classes they depend on, while continuing to use trade data to determine product weights in order to project countries onto the eigenvectors of $\tilde{\Phi}$.

Supplementary Note 9: Complexity of oil extraction and artistic performance

Product complexity metrics sometimes give complexity values that arguably run counter to expectations. In the main text, we noted (as examples) oil production and pottery and works-of-art. Oil production ranks near the bottom in the complexity value assigned to it by both the PCI and product Quality. However, one might anticipate oil production to have high complexity, given that it requires knowledge of multiple fields of engineering, geology, chemistry, physics, transportation and logistics, business operations, and other areas. At the other end of the spectrum, pottery and works-of-art are ranked high in complexity by these metrics, even though this product class might be expected to combine relatively fewer productive capabilities. To probe these expectations, we look at the occupational diversity and wages of the two NAICS industries “Oil and gas extraction” and “Independent artists, writers, and performers” using data from the US Bureau of Labor Statistics [12].

We estimate the number of occupations engaged in each by several measures:

- The number of occupations with RCA greater than 1. (RCA here is defined as θ_{oi}/θ_o , where θ_{oi} is the share of occupation o in the total employment of industry i , and θ_o is the average such share of occupation o across industries.)
- The number of occupations in the industry with an employment share θ_{oi} greater than 1%.
- The Hill number of occupations in the industry, $\exp(-\sum_o \theta_{oi} \log \theta_{oi})$.

The number of occupations engaged in by these industries is consistent with the broad expectations set up above. Oil extraction ranks near the top in occupational diversity (top 13% of industries) while artistic performance ranks relatively low (bottom 39%). In addition we examine the median wages of occupations engaged in each industry, since it has been argued that wages should be higher in complex industries. We find similar results, with oil and gas extraction ranking near the top (4th), and artistic performance ranking relatively low (173rd) out of 247 industries.

Supplementary Table 2: Alternate correlates of complexity for “Oil & gas extraction” and “Independent artists, writers, and performers” Data comes from the US Bureau of Labor Statistics [12]. Industry ranks (out of 247 industries) are given in parentheses.

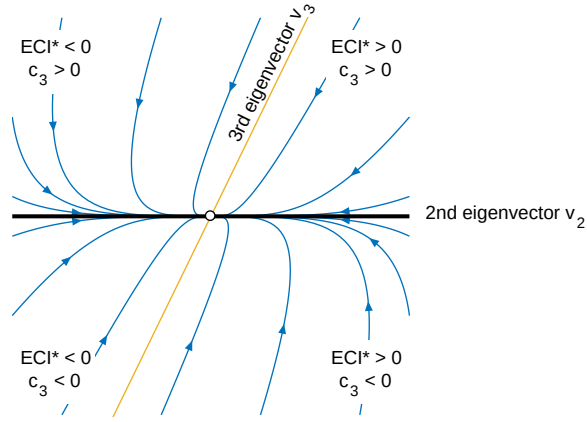
NAICS	Name	Num. occ. RCA > 1	Num. occ. emp. share > 1%	Hill num.	Median Wage (2023 USD)
211100	Oil and Gas Extraction	23 (3)	21 (33)	22.21 (11)	97,430 (4)
711500	Independent Artists, Writers, and Performers	9 (174)	12 (204)	11.07 (153)	49,850 (173)

Supplementary Note 10: Dimensionality reduction

We can relate the theory here to recent works that study the geography of economic activity and complexity metrics through the lens of dimensionality reduction. For instance, Mealy et al. (2019) [13] point out that the conventional ECI can be seen as a dimensionality reduction tool, equivalent to a spectral clustering algorithm that orders countries along an axis in a way that tries to preserve the proximities between baskets of activities. Similarly, Sciarra et al. (2020) [10] put forward a complexity index that reduces the dimensionality of economic activity data by combining information from two eigenvectors of a country-country similarity graph. Brummitt et al. (2020) [14], in turn, apply machine learning methods to extract dimensions that characterize variation in historical export baskets across countries and time, finding axes that strongly correlate with the PCIs. In the works above, dimensionality reduction involves strategies to summarize complex data. In contrast, in exploring our dynamical model, we used the tool of eigenmode decomposition. Eigenmode decomposition is not focused on discovering structure in data per se, but on describing a dynamical process. Nevertheless, it is well understood that these approaches are related, and here we briefly summarize the relationship.

First, the outcome of eigenmode decomposition is very similar to dimensionality reduction techniques: It returns dimensions (the eigenvectors $\mathbf{v}_\mu$) in which one can describe high-dimensional data (the matrix of activity baskets $\mathbf{R}(t)$), along with the coordinate projections of each data point onto these axes (the coefficients $c_\mu(t)$). The two approaches differ in how directions in the vector space are determined and interpreted. Typically, dimensionality reduction techniques aim to identify directions that capture large amounts of variation in data, while eigenmodes characterize coherent dynamical patterns. This difference is closely related to the distinction between proper orthogonal decomposition (POD) and dynamical mode decomposition (DMD) in applications of machine learning to dynamical systems (e.g. [15]): Whereas POD modes optimally reconstruct a data set of snapshots of a system over time, DMD modes identify coherent structures in the dynamics.

The eigenmode decomposition not only generates a representation of data in a vector space, but shows that the model predicts that countries' activity baskets will converge to a lower-dimensional subspace that corresponds to the modes with the largest time scales. This is a well-known outcome of the interaction of structure and dynamics in networks [16, 17]. Heterogeneity in the time scales of different eigenmodes leads countries' activity baskets to collapse onto a low-dimensional subspace of possible activity baskets as differences between countries shrink first among quickly-decaying dimensions, leaving activity baskets scattered primarily along slowly-decaying dimensions. A simple example is given in Supplementary Fig. 8, which depicts the phase portrait of the model if we assume that the characteristic time scale of the 2nd eigenmode is four times that of the 3rd, $\tau_2/\tau_3 = 4$. Phase lines show the path of a country that obeys the model, and show that variation along the 3rd eigenvector shrinks much more quickly than variation along the 2nd eigenvector. This illustrates how the theory here can be viewed through a dimensionality reduction lens, and indeed encourages such a perspective.



Supplementary Figure 8: Illustration of the model dynamics. We display the phase portrait in the 2D plane corresponding to the 2nd and 3rd eigenmodes in a case where $\tau_2/\tau_3 = 4$. Different points in the plane correspond to different coordinate projections for an activity basket $\mathbf{R}(t)$ onto the 2nd and 3rd eigenvectors $\mathbf{v}_2$ and $\mathbf{v}_3$. We present a skewed coordinate system as a reminder that these vectors in general are non-orthogonal, $\mathbf{v}_2 \cdot \mathbf{v}_3 \neq 0$. The four quadrants are labelled by whether the projection onto each eigenvector is positive or negative. Blue lines show the paths of a country that obeys the model depending on its starting position. The model has a stable fixed point where the projections onto both eigenvectors are zero. Because the speed of convergence is much faster along the 3rd eigenvector, however, variation along this direction quickly shrinks, while variation along the 2nd eigenvector is sustained for a longer time.

Supplementary Note 11: Predicting country income and growth

The literature on economic complexity indices often evaluates different indices by how well they predict countries’ income (GDP per capita) and income growth. Here we examine this predictive ability for the A and b (ECI*) coordinates as well, comparing them with the ECI, Fitness, and GENEPIY. Our sample uses all countries not classified as resource-dependent or tax havens by the criteria in Fig. 3 of the main text. (We take ‘developing’ countries to refer to economies on the UN’s least developed economies list [1], ‘tax havens’ to non-OECD and non-developing countries appearing in Hines (2010) [2], and ‘resource-rich’ to economies not elsewhere categorized whose natural resource rent [3] exceeded 5% of GDP in 2018.) For each country we calculate annual GDP per capita levels and growth rates over 5-, 10-, and 25-year non-overlapping periods between 1962 and 2018. We fit the following equation using Ordinary Least Squares (OLS) estimation:

$$y_{ct} = \alpha + X_{ct}\beta + \varepsilon_{ct}. \quad (8)$$

Here α is a constant, ε_{ct} a disturbance term, and y_{ct} is either country c ’s log-transformed GDP per capita in year t , or the growth thereof. X_{ct} contains one or more country metrics. In the growth regressions, it also contains the log-transformed GDP per capita at the start of the period. Since the A and b coordinates emerged jointly from the dynamical model in the main text, we also estimate a regression model where A and b are included simultaneously. In all models we cluster standard errors at the country level.

Income. All the complexity metrics and the A and b coordinates are strongly predictive of a country’s income (Table 3). The ECI explains more variance than Fitness, while GENEPIY has the weakest explanatory force. We observe the strongest predictive power for the b coordinate (ECI*), which outperforms all complexity indices.

Supplementary Table 3: GDP per capita levels.

	Base	ECI	Fitness	GENEPIY	A	b	A and b
ECI		17.787*** (1.353)					
Fitness			0.417*** (0.048)				
GENEPIY				3.552*** (0.621)			
A					3.502*** (0.439)		0.606 (0.531)
b						2.745*** (0.211)	2.468*** (0.263)
Constant	8.116*** (0.140)	8.585*** (0.096)	7.535*** (0.134)	7.968*** (0.134)	6.177*** (0.264)	8.991*** (0.110)	8.566*** (0.369)
# obs.	6,200	62,00	6,200	6,200	6,200	6,200	6,200
R^2	0.019	0.440	0.343	0.113	0.300	0.461	0.465
Adj. R^2	0.010	0.435	0.337	0.105	0.293	0.456	0.460

Models predict countries’ log(GDP/cap), pooling across all years and countries.

Standard errors, clustered at the country level, in parentheses.

Two-sided t-test of slope equal to 0: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Growth. Tables 4 - 6 give the results of growth regressions over 5-, 10-, and 25-year horizons. Variables ranked similarly in their predictiveness of growth as they did for income. Across time horizons, the b coordinate explains the largest variance. Among complexity indices, the ECI explains more variance than Fitness, which in turn outperforms GENEPIY. Notably, the A coordinate has better relative predictive power in this setting than when predicting income levels, and ranks second behind the b coordinate over all horizons.

These results are interesting in that, despite our framework making no assumptions about how structural transformation relates to wealth, the continuous weights used in the calculation of the b coordinate (ECI*)

Supplementary Table 4: 5-year GDP per capita growth.

	Base	ECI	Fitness	GENEPY	A	b	A and b
log(gdp/cap)	0.003 (0.004)	-0.015*** (0.005)	-0.005 (0.005)	0 (0.004)	-0.012** (0.005)	-0.020*** (0.005)	-0.021*** (0.006)
ECI		0.754*** (0.137)					
Fitness			0.009*** (0.003)				
GENEPY				0.069** (0.029)			
A					0.180*** (0.034)		0.095** (0.037)
b						0.140*** (0.022)	0.101*** (0.023)
Constant	0.073** (0.033)	0.238*** (0.047)	0.123*** (0.040)	0.088** (0.035)	0.094*** (0.031)	0.302*** (0.052)	0.249*** (0.050)
# obs.	1,155	1,155	1,155	1,155	1,155	1,155	1,155
R^2	0.096	0.144	0.110	0.101	0.153	0.166	0.177
Adj. R^2	0.088	0.135	0.100	0.091	0.144	0.158	0.167

Models predict countries' log(GDP/cap) growth, pooling across all countries and 5-year periods.

Standard errors, clustered at the country level, in parentheses.

Two-sided t-test of slope equal to 0: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Supplementary Table 5: 10-year GDP per capita growth.

	Base	ECI	Fitness	GENEPY	A	b	A and b
log(gdp/cap)	0.003 (0.008)	-0.034*** (0.012)	-0.014 (0.010)	-0.003 (0.008)	-0.027** (0.011)	-0.044*** (0.013)	-0.046*** (0.013)
ECI		1.374*** (0.313)					
Fitness			0.019*** (0.005)				
GENEPY				0.159*** (0.057)			
A					0.377*** (0.078)		0.214** (0.088)
b						0.268*** (0.050)	0.177*** (0.056)
Constant	0.157** (0.071)	0.493*** (0.113)	0.269*** (0.086)	0.199*** (0.074)	0.204*** (0.067)	0.638*** (0.123)	0.503*** (0.119)
# obs.	496	496	496	496	496	496	496
R^2	0.111	0.172	0.136	0.122	0.201	0.210	0.227
Adj. R^2	0.102	0.161	0.125	0.112	0.191	0.200	0.216

Models predict countries' log(GDP/cap) growth, pooling across all countries and 10-year periods.

Standard errors, clustered at the country level, in parentheses.

Two-sided t-test of slope equal to 0: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Supplementary Table 6: 25-year GDP per capita growth.

	Base	ECI	Fitness	GENEPY	A	b	A and b
$\log(\text{gdp}/\text{cap})$	0.054 (0.034)	0.031 (0.033)	-0.006 (0.047)	0.001 (0.046)	-0.066 (0.043)	-0.009 (0.036)	-0.067 (0.043)
ECI		4.507*** (0.867)					
Fitness			0.055*** (0.018)				
GENEPY				1.418*** (0.501)			
A					1.204*** (0.250)		1.039*** (0.326)
b						0.623*** (0.133)	0.173 (0.135)
Constant	-0.181 (0.305)	0.069 (0.310)	0.220 (0.382)	0.166 (0.376)	0.206 (0.277)	0.462 (0.343)	0.331 (0.246)
# obs.	85	85	85	85	85	85	85
R^2	0.031	0.144	0.091	0.074	0.324	0.203	0.332
Adj. R^2	0.019	0.123	0.069	0.052	0.308	0.184	0.307

Models predict countries' $\log(\text{GDP}/\text{cap})$ growth, pooling across all countries and 25-year periods.

Standard errors, clustered at the country level, in parentheses.

Two-sided t-test of slope equal to 0: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

lead to better alignment with per capita income and growth than the dichotomous weights used in the ECI. We leave the question of why open for future research.

Supplementary Note 12: Benefits of seeing complexity metrics as coordinates of structural change

In discussing our dynamical modeling framework, we emphasized that quantities very similar to complexity metrics (the A and b coordinates) can be motivated using standard arguments of dynamical systems theory. These arguments lead to measures that have little to do with ‘complexity’ per se, calling into question the usual interpretation of complexity metrics. However, we could also ask what benefits or opportunities are suggested by the framework here, which invites us to see complexity metrics as quantities that track the path of structural change by summarizing an economy’s basket of activities in interpretable, low-order terms. We describe four possible benefits below.

First, a theory like the one here can potentially give interpretations to the signs and magnitudes of some complexity metrics. For example, the conventional PCI and ECI are usually standardized to have zero mean and unit variance, implying that only the way they rank products and countries matters. Yet our theory assigns clear meaning to signs and magnitudes: Elements of $\mathbf{v}_2$ (**PCI***) convey which products are vacated or reached, and how large the change in these products is as structural change occurs. Similarly, $b(t)$ (**ECI***) conveys more than just rank information – its magnitude conveys how far a country has progressed in a particular direction of change. (As a caveat, we note that there are still numerical and sign differences among these metrics and our coordinates, due to differences in how each one is constructed, including choices about the product proximity matrix selected; the use of RCAs as measures of ability; and how RCAs are transformed.)

Second, interpreting complexity metrics as coordinates raises the possibility of tracking countries’ structural change in absolute terms. A longstanding concern of research on economic growth is whether countries are converging in income. A related question is whether countries are converging in industrial composition. It is clearly not possible to track such a convergence if complexity metrics can only provide rank information about countries. But it is possible to do this using the A and b coordinates, or any complexity metric, if one can justify that the cardinal values of these metrics are meaningful, so that different countries can potentially reach similar coordinate values over time.

Third, by deriving these coordinate measures from a dynamical model of the process of development, improving the description of this process – e.g. developing better proximity matrices Φ , or improving on the dynamical model presented in the main text – can feed into better summary measures of structural change, offering a path for refinement of these measures.

Finally, because our coordinates depend on the network of products Φ , our results raise the question of what happens as this network changes. While it is conventional to assume that Φ is fixed for the purposes of predicting product transitions, proximities between different products can evolve [18] as the technologies underpinning these products change, which in turn could shift the directional tendency of structural change. While in the past countries have undergone well-documented patterns of development (e.g. the shift out of agricultural products toward manufactured goods [19, 14]) this canonical trajectory could change. A framework like the one here can account for such shifts in a similar way that index numbers do in other settings. Change in a region’s $b(t)$ coordinate (for example) would be split into two components – one due to changes in a country’s weights on the elements of $\mathbf{v}_2$, and one due to changes in the elements of $\mathbf{v}_2$ themselves. This would allow one to decompose movements of a country into sum of a long-term movement along the direction of development, and shifts in the direction of development itself, whereby the relative sequence of different products can rise and fall over time. This would also be one way for studies to evaluate the quality of the approximation that Φ is fixed in time, as this would be reflected in the rate of change of $\mathbf{v}_2$.

Supplementary Note 13: Relationship between left and right eigenvectors

Here we derive a relationship between the left and right eigenvectors of L_Φ . This relationship is used to help determine our normalization convention for these vectors, as well as to derive the ECI-like expression for the b coordinate (equation 28, main text).

First, note that $L_\Phi = I - \tilde{\Phi}$ and $\tilde{\Phi}$ share the same eigenvectors. The matrix $\tilde{\Phi} = D^{-1}\Phi$ with $D = \text{Diag}(\Phi\mathbf{1})$ is a row-normalized stochastic matrix, and so its principal right eigenvector may be taken to be a vector of 1s, $\mathbf{v}_1 = \mathbf{1}$. Its principal left eigenvector can be understood as a vector of stationary probabilities $\boldsymbol{\pi}$ for the Markov chain described by $\tilde{\Phi}$, $\mathbf{w}_1 = \boldsymbol{\pi}$.

Define $\Pi = \text{Diag}(\boldsymbol{\pi})$. If Φ is symmetric, then the left and right eigenvectors of $\tilde{\Phi}$ are related by

$$\Pi \mathbf{v}_\mu = \alpha \mathbf{w}_\mu \quad (9)$$

where α is a constant that depends on the normalization of the eigenvectors.

Proof: By definition the right eigenvectors of $\tilde{\Phi}$ satisfy $\tilde{\Phi} \mathbf{v}_\mu = \lambda_\mu \mathbf{v}_\mu$, which can be rearranged as

$$(D^{-1/2} \Phi D^{-1/2})(D^{1/2} \mathbf{v}_\mu) = \lambda_\mu (D^{1/2} \mathbf{v}_\mu).$$

The vector $D^{1/2} \mathbf{v}_\mu$ is thus a right eigenvector of the symmetric matrix $Q \equiv D^{-1/2} \Phi D^{-1/2} = D^{1/2} \tilde{\Phi} D^{-1/2}$. Similarly, the left eigenvectors of $\tilde{\Phi}$ satisfy $\mathbf{w}_\mu^\dagger \tilde{\Phi} = \lambda_\mu \mathbf{w}_\mu^\dagger$, which can be rearranged as

$$(\mathbf{w}_\mu^\dagger D^{-1/2})(D^{-1/2} \Phi D^{-1/2}) = \lambda_\mu (\mathbf{w}_\mu^\dagger D^{-1/2}),$$

showing that $D^{-1/2} \mathbf{w}_\mu$ is a left eigenvector of the same matrix Q . Finally, since Q is symmetric, any right eigenvector is also a left eigenvector, and thus

$$D^{1/2} \mathbf{v}_\mu = \alpha D^{-1/2} \mathbf{w}_\mu$$

for some constant α . We therefore have $D \mathbf{v}_\mu = \alpha \mathbf{w}_\mu$. In particular, the first left and right eigenvectors of $\tilde{\Phi}$ are $\mathbf{v}_1 \propto \mathbf{1}$ and $\mathbf{w}_1 \propto \boldsymbol{\pi}$, and so we have $D \mathbf{1} \propto \boldsymbol{\pi}$. Since D is diagonal, in index form this reads $D_{pp} \propto \pi_p$, i.e., D is a matrix that is proportional to Π , giving us equation (9).

Normalization of eigenvectors. Our convention for normalizing eigenvectors helps determine the numerical scales of the A and b coordinates (and in principle other coordinate projections). In general left and right eigenvectors are biorthogonal, obeying $W^T V = \Psi$ for some diagonal matrix Ψ . It simplifies many calculations to require that $\Psi = I$ so that left and right eigenvectors obey

$$\mathbf{w}_\nu^\dagger \mathbf{v}_\mu = \delta_{\nu\mu}. \quad (10)$$

This condition does not determine a normalization for the $\mathbf{w}_\mu$ and $\mathbf{v}_\mu$ eigenvectors, since these can be rescaled as $\mathbf{v}_\mu \rightarrow c \mathbf{v}_\mu$ and $\mathbf{w}_\mu \rightarrow \frac{1}{c} \mathbf{w}_\mu$ while preserving equation (10). However, we can pin down a normalization for the eigenvectors by requiring that they also satisfy equation (9) with proportionality constant $\alpha = 1$. One way to accomplish this is to normalize left eigenvectors such that the weighted 2-norm

$$\|\mathbf{w}\| \equiv \sqrt{\mathbf{w}^T \Pi^{-1} \mathbf{w}} \quad (11)$$

is 1 for each left eigenvector, and to normalize right eigenvectors such that the weighted 2-norm

$$\|\mathbf{v}\| \equiv \sqrt{\mathbf{v}^T \Pi \mathbf{v}} \quad (12)$$

is 1 for each right eigenvector. In addition to satisfying equation (9) with $\alpha = 1$, this normalization convention has the nice side effect that the first right eigenvector $\mathbf{v}_1 = \mathbf{1}$ is a vector of 1s, and the elements of the first left eigenvector $\mathbf{w}_1 = \boldsymbol{\pi}$ sum to 1. (We note in passing that equations (11) and (12) are dual norms. If $\|\mathbf{v}\|$ is a norm for vector $\mathbf{v}$, the dual norm of $\mathbf{w}$ is the least upper bound of $\mathbf{w}^T \mathbf{v}$ for all $\mathbf{v}$ such that $\|\mathbf{v}\| \leq 1$. To see

that these are dual norms, recall the Cauchy-Schwarz inequality $|\mathbf{a}^T \mathbf{b}| \leq \|\mathbf{a}\|_2 \|\mathbf{b}\|_2$. Making the replacements $\mathbf{a} = \Pi^{-1/2} \mathbf{w}$ and $\mathbf{b} = \Pi^{1/2} \mathbf{v}$ lets us write this in terms of the norms equations (11)-(12):

$$|\mathbf{w}^T \mathbf{v}| \leq \|\Pi^{-1/2} \mathbf{w}\|_2 \|\Pi^{1/2} \mathbf{v}\|_2 = \sqrt{\mathbf{w}^T \Pi^{-1} \mathbf{w}} \sqrt{\mathbf{v}^T \Pi \mathbf{v}} = \|\mathbf{w}\| \|\mathbf{v}\|. \quad (13)$$

As $|\mathbf{w}^T \mathbf{v}|$ reaches its largest value when $\|\mathbf{v}\| = 1$, the least upper bound of $|\mathbf{w}^T \mathbf{v}|$ is $\sqrt{\mathbf{w}^T \Pi^{-1} \mathbf{w}}$, which is equation (11).)

Supplementary Note 14: On the ECI* weights when $\Phi = \Phi^P$

As discussed in the Methods, the special case $\Phi = \Phi^P$ is noteworthy because **PCI*** coincides with **PCI**. There are also reasons why this special case is interesting for the ECI*. First, note that in this case L_Φ 's first left eigenvector $\mathbf{w}_1 = \boldsymbol{\pi}$ is proportional to the vector of product ubiquities:

$$\boldsymbol{\pi} \propto \mathbf{u}. \quad (14)$$

To see this, observe that $L_\Phi = I - \tilde{\Phi}$ and $\tilde{\Phi}$ have the same eigenvectors, and that $\mathbf{u}$ is a left eigenvector of $\tilde{\Phi}^P = U^{-1}M^TD^{-1}M$ with eigenvalue 1:

$$\mathbf{u}^T \mathcal{P} = \mathbf{u}^T U^{-1}(M^T D^{-1} M) = \mathbf{1}^T M^T D^{-1} M = \mathbf{d}^T D^{-1} M = \mathbf{1}^T M = \mathbf{u}^T. \quad (15)$$

Taking $\boldsymbol{\pi} \propto \mathbf{u}$, the ECI* (equation (9) of the main text) becomes

$$\text{ECI}_c^*(t) = \sum_p \left(\frac{R_{cp}(t)u_p}{\sum_{p'} R_{cp'}(t)u_{p'}} \right) \text{PCI}_p. \quad (16)$$

Thus, in this special case, the weights used by the ECI* are ubiquity-adjusted RCAs. These weights emerge as a by-product of using the proximity matrix $\Phi = \Phi^P$, and do not represent an intentional choice about how different products in the export basket should be weighted. However, such a ubiquity adjustment could nevertheless be beneficial from a measurement standpoint, because it puts more weight on products for which high RCAs are harder to achieve and are therefore more exceptional. For example, many countries in the world produce textiles, but only a few produce ostrich eggs. As a result, almost all ostrich egg-producers have high RCAs in this product. In contrast, to achieve a high RCA in a high-ubiquity product such as textiles, a country must outcompete many others. In general, achieving high RCAs will be more difficult in ubiquitous products than in nonubiquitous ones, and it would be reasonable to regard high values of R_{cp} as more impressive the more ubiquitous product p is. The conventional ECI not only thresholds RCAs in weighting products in a country's activity basket, but it also makes no adjustment for how competitive each product's export market is. In contrast, the ECI* uses unthresholded RCAs, and in the special case $\Phi = \Phi^P$, it emphasizes the RCAs of ubiquitous products. We do not offer a more general interpretation of the ergodic weights π_p that applies outside the special case $\Phi = \Phi^P$, but exploring such issues could reveal interesting aspects of measurement in those contexts as well.

Supplementary References

- [1] United Nations. Un list of least developed countries. URL <https://unctad.org/topic/least-developed-countries/list>.
- [2] Hines Jr, J. R. Treasure islands. *Journal of Economic Perspectives* **24**, 103–26 (2010).
- [3] The World Bank. World Bank Development Indicators (2018). URL <http://data.worldbank.org/>.
- [4] Li, Y. & Neffke, F. M. Evaluating the principle of relatedness: Estimation, drivers and implications for policy. *Research Policy* 104952 (2024).
- [5] Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A. & Pietronero, L. A new metrics for countries' fitness and products' complexity. *Scientific Reports* **2**, 1–7 (2012).
- [6] Hidalgo, C. A. & Hausmann, R. The building blocks of economic complexity. *Proceedings of the National Academy of Sciences* **106**, 10570 – 10575 (2009).
- [7] Bustos, S. & Yildirim, M. A. Production ability and economic growth. *Research Policy* 104153 (2020).
- [8] Teza, G., Caraglio, M. & Stella, A. L. Growth dynamics and complexity of national economies in the global trade network. *Scientific reports* **8**, 1–8 (2018).
- [9] Teza, G., Caraglio, M. & Stella, A. L. Entropic measure unveils country competitiveness and product specialization in the world trade web. *Scientific reports* **11**, 1–11 (2021).
- [10] Sciarra, C., Chiarotti, G., Ridolfi, L. & Laio, F. Reconciling contrasting views on economic complexity. *Nature Communications* **11**, 1–10 (2020).
- [11] Gomez-Lievano, A. & Patterson-Lomba, O. Estimating the drivers of urban economic complexity and their connection to economic performance. *Royal Society open science* **8**, 210670 (2021).
- [12] US Bureau of Economic Analysis. Occupational employment and wage statistics. URL <https://www.bls.gov/oes/data-overview.htm>.
- [13] Mealy, P., Farmer, J. D. & Teytelboym, A. Interpreting economic complexity. *Science Advances* **5**, eaau1705 (2019).
- [14] Brummitt, C. D., Gómez-Liévano, A., Hausmann, R. & Bonds, M. H. Machine-learned patterns suggest that diversification drives economic development. *Journal of the Royal Society Interface* **17**, 20190283 (2020).
- [15] Tu, J. H., Rowley, C. W., Luchtenburg, D. M., Brunton, S. L. & Kutz, J. N. On dynamic mode decomposition: Theory and applications. *Journal of Computational Dynamics* **1**, 391 (2014).
- [16] Simon, H. A. & Ando, A. Aggregation of variables in dynamic systems. *Econometrica: Journal of the Econometric Society* 111–138 (1961).
- [17] Schaub, M. T., Delvenne, J.-C., Lambiotte, R. & Barahona, M. Structured networks and coarse-grained descriptions: A dynamical perspective. *Advances in Network Clustering and Blockmodeling* 333–361 (2019).
- [18] Hidalgo, C. A. The dynamics of economic complexity and the product space over a 42 year period. *CID Working Paper Series* (2009).
- [19] Imbs, J. & Wacziarg, R. Stages of diversification. *American Economic Review* **93**, 63–86 (2003).
- [20] Hidalgo, C. A., Klinger, B., Barabási, A.-L. & Hausmann, R. The product space conditions the development of nations. *Science* **317** (2007).
- [21] Hausmann, R., Stock, D. P. & Yildirim, M. A. Implied comparative advantage. *Research Policy* 104143 (2021).