How production networks amplify economic growth

James McNerney[1], Charles Savoie[2], Francesco Caravelli[3], J. Doyne Farmer[2,4]

Technological improvement is the most important cause of long-term economic growth,[1] but the factors that drive it are still not fully understood. In standard growth models technology is treated in the aggregate, and a main goal has been to understand how growth depends on factors such as knowledge production.[2] But an economy can also be viewed as a network, in which producers purchase goods, convert them to new goods, and sell them to households or other producers.[3] Here we develop a simple theory that shows how the network properties of an economy can amplify the effects of technological improvements as they propagate along chains of production. A key property of an industry is its output multiplier, which can be understood as the average number of production steps required to make a good. The model predicts that the output multiplier of an industry predicts future changes in prices, and that the average output multiplier of a country predicts future economic growth. We test these predictions using data from the World Input Output Database and find results in good agreement with the model. The results show how purely structural properties of an economy, that have nothing to do with innovation or human creativity, can exert an important influence on long-term growth.

Economic output is made by a complex network of industries that buy goods from one another, convert them to new goods, and sell them to households or other industries. Studies have examined a number of characteristics of production networks that hold across diverse economies, including their link weight and industry size distributions,[4–8] community structure,[7] and path-length properties.[9] Economies typically have a heterogeneous network structure with a few highly central industries that are strong suppliers to the rest of the network,[6,8,10] a feature that has been incorporated into models where short-term fluctuations of economic output are generated by shocks to individual industries.[8,11–15] Here, we focus on how the structure of production networks affects long-term economic growth. Over the long term, changing industry productivities significantly alter prices and production flows in the network. Multiple economic processes contribute to productivity change,[16] but long-term improvements are thought to result primarily from improvements to technology,[16] which is widely understood to be the principle driver of growth.[1] However, the processes of technological change and growth remain imperfectly understood.

As technology evolves, the network structure of production plays a key role in amplifying changes to prices and production flows in an economy. A key property of an industry i is its output multiplier[17] $\mathcal{L}_i$, which can be defined recursively as

$$\mathcal{L}_i = \sum_j \mathcal{L}_j a_{ji} + 1, \tag{1}$$

where a_{ji} is the fraction of good j in producer i's expenditures.

To build intuition for this quantity, we make an ecosystem analogy. A species in a food web can be represented as a node in a network, with links to the species it eats. A species' place in the food web is often characterized by its trophic level – informally, its position along a food chain.[18] Photosynthesizers, which use sunlight as a resource, have trophic level 1 by convention, while species that consume only photosynthesizers have trophic level 2, and so on. Food webs typically have complex structures where each node obtains inputs from multiple trophic levels, so that trophic levels are usually not integers. Letting a_{ji} represent the energy fraction of prey j in species i's diet, Eq. (1) states that the trophic level $\mathcal{L}_i$ of species i is one greater than the average trophic levels of the species it consumes.[19]

Similarly, an economy can be regarded as a network in which an industry producing a good is a node, with links to the input goods it uses for production. (For simplicity, we lump together goods and services, calling them both "goods" for brevity, and assume each industry produces only one good.) Household factors of production such as labor are the base resource, so that producers that pay only households occupy trophic position 1. With a_{ji} as the fraction of good j in producer

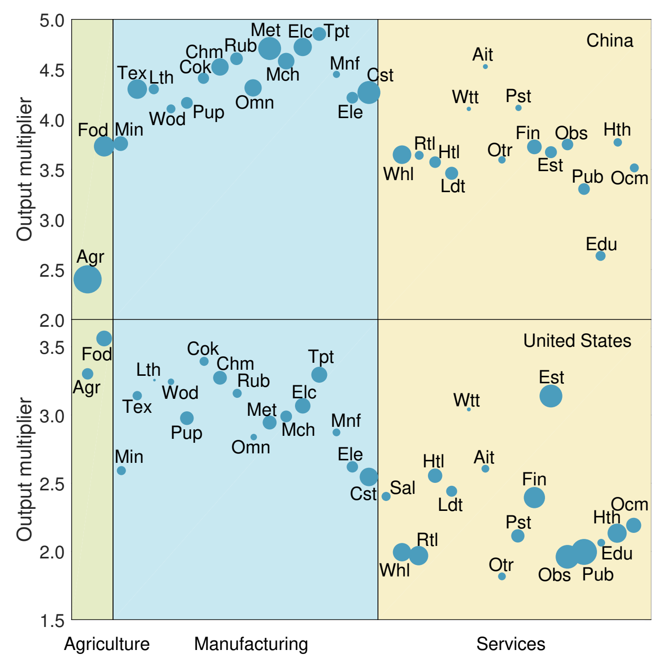

Figure 1: **Output multipliers in the Chinese and U.S. economies.** The output multiplier of each industry is plotted against a standard industry classification. Node size corresponds to an industry's gross output. Three-letter industry codes are given in Extended Data Table 1.

i's expenditures, Eq. (1) gives a measure of the trophic level of an industry. In economics, $\mathcal{L}_i$ is also called i's total backward linkage[17] or downstreamness.[20] Letting A denote the matrix with elements a_{ij} and $\boldsymbol{\mathcal{L}}$ the vector with elements $\mathcal{L}_i$, rearranging Eq. (1) gives the vector of output multipliers as $\boldsymbol{\mathcal{L}} = (I - A^T)^{-1}\mathbf{1}$, where $\mathbf{1}$ is a vector of 1s.

The output multiplier can also be understood in terms of network path lengths.[20–22] Regarding the elements a_{ji} as transition probabilities in a Markov chain,[23] $\mathcal{L}_i$ gives the average length of all production chains ending at industry i, following each path backward through inputs until it reaches households. (See Supplementary Information.) As a result, two factors influence the output multiplier of a producer: the fraction of its expenditures that go directly to purchasing labor, and the output multipliers of the goods that it buys. Higher labor expenditures make it more likely that a dollar spent will go directly to the household node, realizing the shortest possible path length of 1, and lowering the output multiplier. Similarly, dollars spent on goods from producers with high output multipliers will take more steps to reach the household node than dollars spent on goods with low output multipliers.

The output multipliers of an economy collectively help characterize the economy's network structure. Two examples are shown in Fig. 1. Using data from the World Input Output Database (WIOD),[24] we plot the output multipliers of China and the United States. (The WIOD provides money flows between producers, aggregated into 35 industries in 40 countries, representing about 86% of global GDP.) The output multipliers of industries in China are higher than those of the U.S. for two reasons. First, China is heavily concentrated in manufacturing industries such as Electrical and Optical Equipment (Elc) or Basic Metals and Fabricated Metals (Met), which tend to have high output multipliers because they have many steps of production.[25] In contrast, the U.S. is heavily concentrated in industries such as Public Administration and Defense and Compulsory Social Security (Pub) or Renting of Machinery & Equipment and Other Business Activities (Obs), which tend to have shallow production chains. The second reason is that China's labor share of gross expenditures is lower. The difference in the output multiplier

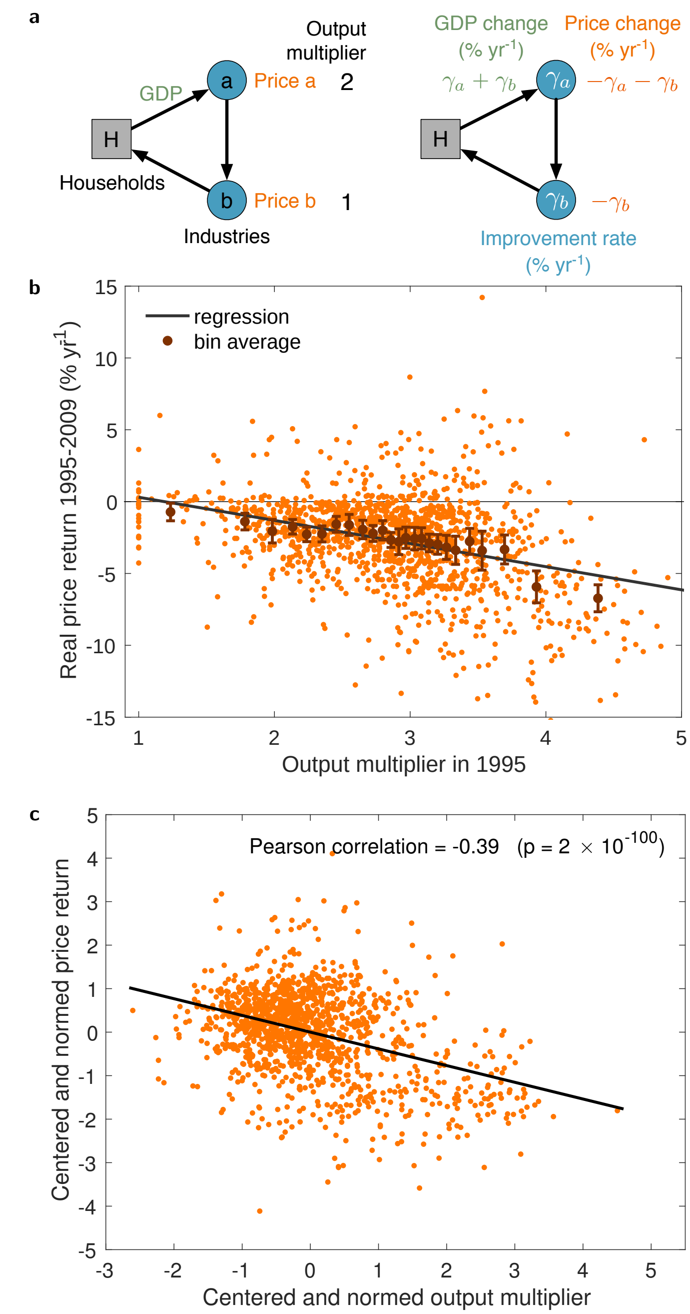

Figure 2: **Industry improvements and price effects.** **a**, A simple production network. Arrows show direction of payments. Second diagram shows effects of productivity shocks on prices and GDP. Improvement rates are shown inside nodes, and resulting rates of change to prices and GDP are shown outside. **b**, Rates of change of real prices for the 1400 country-industry pairs (orange dots) over the period 1995 - 2009 versus industries' output multipliers in 1995. Returns with similar output multipliers were put into bins of about 45 points each, and the average return was computed in each bin (brown dots). Vertical lines give error bars of two standard deviations around the bin mean. The black line is an OLS regression fit. **c**, Variation in price returns versus variation in output multipliers within each of the 35 industry categories in the WIOD. For each industry, we center and normalize its 40 observations across countries by the industry's mean and standard deviation across countries. Similarly, we center and normalize the industry's price returns over 1995 - 2009 by its mean and standard deviation across countries. We plot the centered and normalized variables against one another for all industries (orange dots). The black line is a regression fit.

of agriculture (Agr) in the two countries is illustrative. In the U.S., agricultural industries have high output multipliers similar to manufacturing industries, reflecting a high degree of mechanization. In China, agriculture is more labor-intensive, giving it a lower output multiplier relative to other industries.

Output multipliers have long been used for predicting the impacts of a change in final demand, such as a government stimulus.[26] Additional final demand for a good requires the industry producing it to buy more inputs, increasing its production and setting off a chain reaction that increases the gross output of the economy. Intuitively, this amplification factor is greater when production chains are longer.

Here we go further and propose that production chains play a key role in long-term, technology-driven growth. Let ϕ_{ij} denote the amount of good j needed by producer i per unit of i's output. Neglecting markups, so that prices and costs are the same, the price p_i of each good i is equal to its total cost of production, $p_i = \sum_j \phi_{ij} p_j$. This equation determines prices, so as the matrix of input needs $\phi_{ij}(t)$ evolves, prices change accordingly. Modern work in economics usually adopts a framework that considers the optimization decisions of producers, and the model here can be understood in these terms, though such stronger assumptions are not necessary. See Supplementary Information.

We consider a stylized model of technological improvement in which producers become more efficient in their use of inputs over time. Each producer i reduces its use of good j at rate $\gamma_{ij} = -\dot{\phi}_{ij}/\phi_{ij}$. An industry's improvement is captured by its rate of productivity growth γ_i, which can be expressed as the cost-weighted average of the rates of change of its input uses, $\gamma_i = \sum_j \gamma_{ij} a_{ji}$. (See Supplementary Information.) We plug this assumption into the Leontief framework and then convert from the flow of goods to the flow of money, which is described by the matrix A defined earlier. A simple example of the network dynamics represented by this model is shown in Fig. 2a. There are three nodes, households H and two industries a and b. Households buy good a from industry a, which buys good b from industry b, which buys labor from households.[1] When an industry's productivity rises, it requires less input per unit of good produced, causing its price to fall due to the lowered cost of production. The lower price is also passed to downstream industries, helping their prices fall as well. A basic prediction of the model is that the real growth rate g of an economy is equal to the negative of the average real price return r of the final goods it produces, i.e. that economies grow at the rate at which real prices decrease. As a result, the fall in prices corresponds to economic growth. In the simple network in Fig. 2a, all of GDP is spent on good a, and thus the growth rate equals the rate of decrease of good a's price.

In the Supplementary Information we derive a number of predictions of the model for price evolution in a production network with arbitrary structure. Let $r_i = \dot{p}_i/p_i$ denote the real (i.e. inflation-adjusted) price return of an industry. The price returns of industries $\mathbf{r}$ are related to industries' rates of productivity improvement $\boldsymbol{\gamma}$ through their network interactions by $\mathbf{r} = -H^T\boldsymbol{\gamma}$, where the matrix $H = (I - A)^{-1}$ is known as the Leontief inverse. A consequence of this result is the prediction that industries with larger output multipliers will experience faster price reduction on average. Treating γ_i over a given period as a random number, we write it as a sum of its average value across industries $\bar{\gamma}$ and a deviation $\Delta\gamma_i$. Then the expected value of r_i conditioned on the output multiplier is $E[r_i|\mathcal{L}_i] = -\bar{\gamma}\mathcal{L}_i - \sum_j E\left[\Delta\gamma_j H_{ji}|\mathcal{L}_i\right]$. Under the assumption that the deviations $\Delta\gamma_j$ are uncorrelated with the matrix elements H_{ji}, this reduces to

$$E[r_i|\mathcal{L}_i] = -\bar{\gamma}\mathcal{L}_i, \tag{2}$$

i.e. an industry's expected real price return is proportional to its output multiplier. This formula captures the intuitive idea that industries with longer production chains will tend to realize faster price reduction. The model also predicts co-movement among prices that is shaped by the network structure. We show in the Supplementary Information that the covariances of the price returns $R_{ij} = E[r_i r_j] - E[r_i]E[r_j]$ depend on the variances of the productivity improvement rates $D_m = E[\gamma_m^2] - E[\gamma_m]^2$ and the network structure as $R_{ij} = \sum_m H_{mi} D_m H_{mj}$.

We test these predictions with the price returns of the 1400 industries (40 countries × 35 industry categories) in the WIOD data. We compute price returns of industries over the period 1995 - 2009 and take industries' output multipliers from the year 1995. Comparing these quantities (Fig. 2b) shows the clear deviation of the mean industry behavior with the output multiplier, such that industries with larger output multipliers tend to realize faster price reduction. Regressing the returns against the output multipliers gives a slope of -1.6% per year, with a p-value of 6×10^{-42} and $R^2 = 0.13$. The downward tendency can also be seen by

[1] An unrealistic feature adopted for clarity is that industry a purchases no labor, though this does not affect the intuition that follows.

binning price returns by industries' output multipliers and computing the average return within each bin. Productivity improvement rates tend to be larger for industries with higher output multipliers (Pearson correlation 0.11, $p = 3 \times 10^{-5}$), a correlation that increases the magnitude of the slope in Fig. 2b. To see whether this drives the relationship between price changes and output multipliers, we shuffle improvement rates across industries to remove the correlation with the output multipliers (Extended Data Fig. 1), finding that the output multipliers retain a highly significant correlation with price returns even with this effect removed.

We can improve the predictions of price returns by exploiting the persistent network structure of industries. We estimate productivity growth rates from price returns by a dual method,[16] which means that productivity growth rates in a given year are not independent of the price returns from that year, and cannot be used for the test of the model prediction here. Instead, we split our data into periods, and use independent productivity data from earlier periods to predict price returns in a later period. We split the years 1995 - 2009 into three periods of nearly equal length (5 years, 5 years, 4 years), using data from the first two periods to predict productivity changes $\hat{\boldsymbol{\gamma}}$ in the third under an AR(1) time-series model. We then use $\hat{\boldsymbol{\gamma}}$ to generate predicted returns for the last period as $\mathbf{r} = -H^T\hat{\boldsymbol{\gamma}}$. The predicted and actual price returns are significantly correlated (Extended Data Fig. 2) with a regression slope close to 1 after accounting for the correlation noted earlier between the productivity improvement rates and output multipliers. We also find a good agreement with the model's predictions for the covariance of price returns. Considering all pairs of industries leads to about 1 million covariances, for which we compare the actual values to the predictions given by $\sum_m H_{mi} D_m H_{mj}$, finding a slope of 1.14 and a p-value that is smaller than our machine's precision.

It has been noted that manufacturing industries tend to have larger output multipliers.[25] As a result, the model predicts that manufacturing industries will tend to realize faster rates of price reduction. This is compatible with a well-known observation that manufacturing industries tend to experience faster rates of productivity growth than service industries.[27] However, the model makes the very specific prediction that variation in the output multipliers should predict variations in price returns. Thus, even the same manufacturing industry (e.g. Chemicals and Chemical Products) in different countries should realize different rates of price reduction depending on the value of its output multiplier in these countries. We test this for the 35 industry categories in the WIOD (which includes both manufacturing and other non-manufacturing industries), obtaining the 40 values of the industry's output multiplier observed across countries in the year 1995. Against these we regress the average price returns over 1995 - 2009 (Extended Data Table 3). Remarkably, 34 out of 35 industries have a negative slope as predicted by the model, which in most cases is statistically significant. To assess this behavior in the data as a whole, we pool industries in the following way. Let $\mathcal{L}_{ic}$ denote the output multiplier of industry i in country c, and let r_{ic} be the price return of industry i in country c. To capture the cross-country variation within a given industry, we center and normalize its output multipliers by its mean $\bar{\mathcal{L}}_i$ and standard deviation $\sigma_{\mathcal{L}_i}$ across countries, $(\mathcal{L}_{ic} - \bar{\mathcal{L}}_i)/\sigma_{\mathcal{L}_i}$. Similarly we compute the centered and normalized price returns $(r_{ic} - \bar{r}_i)/\sigma_{r_i}$. The two quantities have a negative correlation of -0.39 (Fig. 2c), indicating that a higher relative output multiplier results in faster price reduction, with a p-value equal to 2×10^{-100}. We also directly compare the predictive ability of industry labels to that of output multipliers, finding that the latter are much better predictors of price change. (See Supplementary Information.)

To better understand how the effects of industry improvement are propagated through the network, we examine how much price reduction is inherited from others versus being generated by local improvements. An industry's price return can be decomposed as $r_i = -\gamma_i + \sum_j r_j a_{ji}$, where $-\gamma_i$ accounts for the direct benefits of i's own improvement and $\sum_j r_j a_{ji}$ accounts for price changes passed to i through input goods (see Extended Data Fig. 3). Industries' price returns are highly correlated with both components, with a correlation of 0.91 to the direct component and 0.71 to the inherited component. (See Supplementary Information.) Inherited price reductions tend to contribute more to price reduction (mean value -1.65 % yr^{-1}) than the direct improvements (-1.06 % yr^{-1}), while the direct component has a wider distribution, and thus explains more of the variation in price returns.

As noted earlier, a basic prediction is that decreases in prices correspond to GDP growth. The model predicts that the rate of real GDP growth g for a closed economy is proportional to the average output multiplier $\bar{\mathcal{L}}$,

$$g = \tilde{\gamma}\bar{\mathcal{L}}. \tag{3}$$

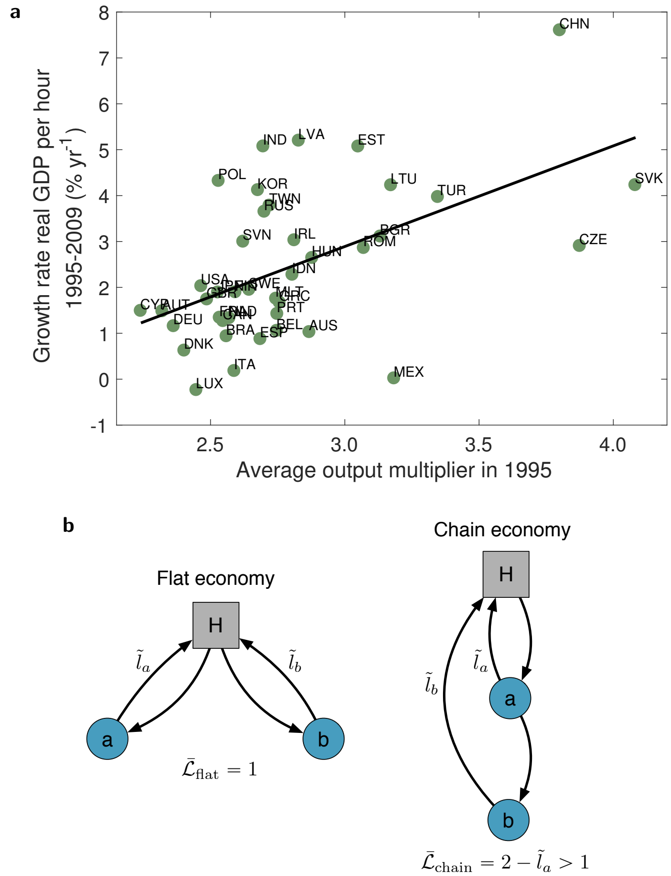

Figure 3: **Country growth rates versus average output multiplier.** **a**, Growth rates of real GDP per hour over 1995 - 2009 for 40 WIOD countries versus $\bar{\mathcal{L}}$ in 1995. Solid line is an OLS regression fit. Three-letter country codes are given in Extended Data Table 2. **b**, Diagram of flat and chain economies. In the flat economy households purchase two final goods that each require only labor as an input. $\tilde{l}_a$ and $\tilde{l}_b$ are the fractions of each industry's expenditures devoted to labor. In the chain economy households purchase one final good, which has a two-step production process. If industries in both economies realize productivity improvements at the same rate, the chain economy is expected to realize faster growth.

Here $\tilde{\gamma}$ is the average rate of productivity improvement $\tilde{\gamma} = \sum_i \eta_i \gamma_i$ of a country's industries, with weights η_i giving the share of producer i in the country's gross output. The factor $\bar{\mathcal{L}} = \sum_i \theta_i \mathcal{L}_i$ is a weighted average of the output multipliers of the country's industries, where θ_i is the GDP share of producer i. $\bar{\mathcal{L}}$ measures the average length of production chains in an economy. Eq. (3) indicates that, all else equal, longer production chains are expected to yield faster growth.

The average output multiplier is a key variable characterizing an economy's production network structure, with predictive value for future growth. The average output multiplier varies slowly, in contrast with the average improvement rate $\tilde{\gamma}$, which fluctuates considerably from year to year (Extended Data Fig. 4). In Fig. 3a, we plot the growth rates of real GDP per hour for the WIOD countries over 1995 - 2009 against their average output multiplier in 1995. The two quantities have a Pearson correlation $\rho = 0.53$, a high value for a single economic variable, with a p-value of 4×10^{-4}.

To intuitively understand why the average output multiplier predicts growth, consider the two economies in Fig. 3b. In the flat economy, with average output multiplier $\bar{\mathcal{L}}_{\text{flat}} = 1$, households buy two final goods, each of which pays only households for inputs. In the chain economy, with $\bar{\mathcal{L}}_{\text{chain}} > 1$, households buy one final good, which has a two-industry production chain. If industries in both economies realize productivity improvements at the same rate, the chain economy grows more quickly because improvements accumulate along the chain connecting industries a and b.

Data on production networks varies in its level of aggregation, ranging from a few industries to hundreds of industries. This raises the concern that the average output multiplier will vary with the granularity of the data. However, the average output multiplier of a closed economy has been shown to be independent of the level of aggregation and equal to $\bar{\mathcal{L}} = \mathcal{O}/Y$, where $\mathcal{O}$ is gross output and Y is net output (GDP).[22] In the practical context of an open economy, computing the average output multiplier for the U.S. at different levels of network resolution shows that it changes little over a wide range of levels of aggregation.[22] (See also Supplementary Information.) Note that for this to be true it is essential that node self-payments are properly accounted for.[2]

Is the average output multiplier capturing something new in growth economics or is it a proxy for something known? To get insight we compare the average output multiplier to 14 variables that commonly appear in growth models. We regress these variables one at a time against country GDP growth rates, using average values over the period 1995 - 2009 (Extended Data Table 4). The model prediction $\tilde{\gamma}\bar{\mathcal{L}}$ (Fig. 2d) has the highest R^2 of any variable, with the average improvement rate $\tilde{\gamma}$ second. The next best is gross capital formation, followed by the average output multiplier $\bar{\mathcal{L}}$, with $R^2 = 0.37$. Several of these variables have significant correlations with $\bar{\mathcal{L}}$, the highest being to gross capital formation. (We also perform multivariate regressions of growth rates against these variables with and without the average output multiplier as a regressor, see Supplementary Information.) The average output multiplier also has low correlations with measures of economic complexity,[29,30] and potentially could be combined with such measures to make better forecasts. In Extended Data Table 4 we also see that the average improvement rate has a correlation of 0.45 with the average output multiplier. This suggests that the relation between growth rates and output multipliers shown in Fig. 3a has two sources: (i) the theoretical prediction that, all else equal, countries with longer production chains should grow faster, and (ii) the empirical observation that countries with longer production chains tend to have higher average improvement rates $\tilde{\gamma}$. Our model says nothing about the second observation, though it is plausible that factors such as investment could simultaneously increase the length of production chains and the rate of technological improvement.

The results here point to an important role played by production in amplifying economic growth. Structural properties of an economy, computed only from its network of production, are seen to influence rates of price reduction and output growth. The model and observations suggest that the growth of a country over long periods is influenced by changes in the lengths of its production chains, as characterized by its average output multiplier. One expects an undeveloped economy to have short chains of production. As manufacturing becomes more prominent and more sophisticated, the average output multiplier increases. Finally, as service industries become more prominent, the average output multiplier decreases. Our model suggests that, all else equal, an economy will accelerate its growth during the manufacturing stage and relax back to a slower growth rate once it becomes more developed. In Fig. 3a, developed economies have low average output multipliers and low growth rates while economies that are developing a strong manufacturing sector, such as China or Slovakia, tend to have high average output multipliers and high growth. The WIOD does not contain data for undeveloped countries, so we cannot confirm that their average output multipliers are low, though it would be very surprising if it were otherwise. In the future, improved models and an increased understanding of how this network evolves can shed further light on how economies develop and on the processes of technological change and growth.

METHODS SUMMARY

We computed output multipliers using data from the World Input-Output Database (WIOD).[24] We treated the world as one large economy and computed a 1400×1400 matrix A of input coefficients a_{ij} corresponding to 35 industries in 40 countries. We took its Leontief inverse and computed the 1400-dimensional vector $\boldsymbol{\mathcal{L}} = (I - A^T)^{-1}\mathbf{1}$ giving the output multipliers of each industry in each country. We computed the average output multiplier of a country by taking the GDP-weighted average of the industry output multipliers for that country. These calculations were done for each year, and where specified averages were taken over the 14-year period from 1995 to 2009. WIOD includes industry production price indices from which we computed the vector $\mathbf{r}$ of rates of price change for each industry in each country. Local improvement rates for each industry were estimated from $\mathbf{r}$ using $\hat{\boldsymbol{\gamma}} = (A^T - I)\mathbf{r}$. For our regressions we used data from the World Bank[31] and Penn World Tables.[32]

[2] Interestingly, an equivalent result was obtained in ecology, showing that the average path length of the average energy input to an ecosystem equals the ratio of the ecosystem's total energy throughput to total energy input.[28]

References

[1] Solow, R. M. A contribution to the theory of economic growth. *The Quarterly Journal of Economics* **70**, 65–94 (1956).

[2] Romer, P. M. Endogenous technological change. *Journal of Political Economy* **98**, 65–94 (1990).

[3] Leontief, W. *Input-Output Economics* (Oxford University Press, 1986).

[4] Xu, M., Allenby, B. R. & Crittenden, J. Interconnectedness and resilience of the U.S. economy. *Advances in Complex Systems* (2011).

[5] Atalay, E., Hortaçsu, A., Roberts, J. & Syverson, C. Network structure of production. *Proceedings of the National Academy of Sciences* **108**, 5199–5202 (2011).

[6] Jones, C. I. Misallocation, economic growth, and input-output economics. NBER Working Papers 16742, National Bureau of Economic Research (2011).

[7] McNerney, J., Fath, B. D. & Silverberg, G. Network structure of inter-industry flows. *Physica A* **392**, 6427–6441 (2013).

[8] Carvalho, V. M. *Aggregate Fluctuations and the Network Structure of Intersectoral Trade*. Ph.D. thesis, University of Chicago (2008).

[9] Antràs, P., Chor, D., Fally, T. & Hillberry, R. Measuring the upstreamness of production and trade flows. *American Economic Review* **102**, 412–416 (2012).

[10] Blöchl, F., Theis, F. J., Vega-Redondo, F. & Fisher, E. O. Vertex centralities in input-output networks reveal the structure of modern economies. *Physical Review E* **83** (2011).

[11] Horvath, M. Sectoral shocks and aggregate fluctuations. *Journal of Monetary Economics* **45**, 69–106 (2000).

[12] Foerster, A. T., Sarte, P.-D. G. & Watson, M. W. Sectoral versus aggregate shocks: A structural factor analysis of industrial production. *Journal of Political Economy* **119**, 1–38 (2011).

[13] Gabaix, X. The granular origins of aggregate fluctuations. *Econometrica* **79**, 733–772 (2011).

[14] Acemoglu, D., Carvalho, V. M., Asuman, O. & Alireza, T.-S. The network origins of aggregate fluctuations. *Econometrica* **80**, 1977–2016 (2012).

[15] Contreras, M. G. A. & Fagiolo, G. Propagation of economic shocks in input-output networks: A cross country analysis. *Physical Review E* **90**, 062812 (2014).

[16] Organization for Economic Cooperation and Development. Measuring Productivity - OECD Manual: Measurement of Aggregate and Industry-level Productivity Growth. Tech. Rep. (2001).

[17] Rasmussen, P. N. *Studies in Inter-sectoral Relations* (Amsterdam: North-Holland, 1957).

[18] Post, D. M. The long and short of food-chain length. *Trends in Ecology & Evolution* **17**, 269 – 277 (2002).

[19] Levine, S. Several measures of trophic structure applicable to complex food webs. *Journal of Theoretical Biology* **83**, 195–207 (1980).

[20] Miller, R. E. & Temurshoev, U. Output upstreamness and input downstreamness of industries/countries in world production. *International Regional Science Review* (2015).

[21] Kemeny, J. G. & Snell, J. L. *Finite Markov Chains* (Van Nostrand, 1960).

[22] Fally, T. On the fragmentation of production in the U.S. (2011). Unpublished.

[23] Leontief, W. & Brody, A. Money-flow computations. *Economic Systems Research* **5**, 225–233 (1993).

[24] Timmer, M. P., Dietzenbacher, E., Los, B., Stehrer, R. & de Vries, G. J. An illustrated user guide to the World Input–Output Database: The case of global automotive production. *Review of International Economics* (2015).

[25] Park, S.-H. & Chan, K. S. A cross-country input-output analysis of intersectoral relationships between manufacturing and services and their employment implications. *World Development* **17**, 199–212 (1989).

[26] Miller, R. E. & Blair, P. D. *Input-output analysis: Foundations and extensions* (Cambridge University Press, 2009).

[27] Baumol, W. J. Macroeconomics of unbalanced growth: The anatomy of urban crisis. *The American Economic Review* **57**, 415–426 (1967).

[28] Finn, J. T. Measures of ecosystem structure and function derived from analysis of flows. *Journal of Theoretical Biology* **56**, 363–380 (1976).

[29] Hidalgo, C. A. & Hausmann, R. The building blocks of economic complexity. *Proceedings of the National Academy of Sciences* **106**, 10570 – 10575 (2009).

[30] Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A. & Pietronero, L. A new metrics for countries' fitness and products' complexity. *Scientific reports* **2**, 723 (2012).

[31] The World Bank. World Bank Development Indicators. URL `http://data.worldbank.org/`.

[32] Feenstra, R. C., Inklaar, R. & Timmer, M. P. The next generation of the Penn World Table. *American Economic Review* **105**, 3150–3182 (2015).

Acknowledgements We thank Robert Axtell, Jerry Silverberg, David Pugh, Vasco Carvalho, Ricardo Hausmann, Sergey Paltsev, and Francois Lafond for valuable discussions. This work was enabled by the National Science Foundation under grant SBE0738187 and in part by the European Commission project FP7-ICT-39 2013-611272 (GROWTHCOM) and a Leading Technology Policy Fellowship from the MIT Institute for Data, Systems, and Society.

METHODS

Description of data We used data from the World Input-Output Database (WIOD).[24] This consisted of a worldwide input-output table for 35 industries in 40 countries covering the period 1995 - 2009, for a total of 14 years. We excluded 2010 and 2011 from the analysis because many countries lacked data on labor compensation needed to compute the output multipliers. The 40 countries together accounted for about 88% of world GDP. The data included industry production price indices from which we computed the vector $\mathbf{r}$ of rates of price change in each industry and each country. For our regressions we used data from the World Bank[31] and Penn World Tables.[32]

Calculation of industry output multipliers We treated the world as one large economy and constructed the 1400×1400 matrix A of input coefficients corresponding to all industries in all countries. We took the Leontief inverse and computed the 1400-dimensional vector $\boldsymbol{\mathcal{L}} = (I - A^T)^{-1}\mathbf{1}$ whose elements give the output multiplier of each industry in each country. Industries and their output multipliers are listed in Extended Data Table 1. We interpreted the labor coefficient $\tilde{\boldsymbol{\ell}}$ (see Supplementary Information) in two ways, either accounting for all payments to households (using value added, row code r64) or accounting for labor income only (using the labour compensation field and WIOD exchange rates to convert to U.S. dollars). All results here used the former unless otherwise noted. We found the results were qualitatively similar either way. The main difference was that output multipliers were smaller when including all payments to households, since this increases the flow of money to the household sector and thus shortens path lengths. WIOD did not contain data for labor and capital income separately for the Rest Of the World (ROW) region. We compared the results of excluding ROW altogether with including it using an assumed fraction of value added to represent labor income, finding qualitatively similar results either way. Results shown are based on including ROW with an assumed labor fraction 0.5, similar to the global average (0.57 in 2009) computed across the WIOD countries.

Calculation of average output multipliers We computed the average output multiplier of each country as a weighted sum of the output multipliers of its industries. The weight of industry i in country c was given by the share of i in c's contribution to world final demand, i.e. $Y_{i,c}/\sum_i Y_{i,c}$ where $Y_{i,c}$ is the total final demand of industry i in country c. The final demand $Y_{i,c}$ accounts for consumption and investment payments by all countries (i.e. column codes c37-c42, summed over countries) and excludes net exports, since in WIOD the latter are accounted for within the input-output table. The average output multiplier was computed in each year and for the regressions shown in Extended Data Table 4 it was averaged over the 14-year period 1995 - 2009. Countries and their average output multipliers are listed in Extended Data Table 2.

Calculation of industry returns The nominal industry return $r'_{i,c}$ of industry i in country c was computed as the log return of (i, c)'s gross output price index. These returns were computed for each year and for the whole period 1995 - 2009. The wage rate in a country was computed as the ratio of the total labor income earned to total hours worked by industries in the country, and the log return of this was computed to give ρ_c. The real price return $r_{i,c}$ was then computed as $r'_{i,c} - \rho_c$.

Calculation of productivity growth rates We estimated productivity growth rates as $\hat{\boldsymbol{\gamma}} = (A^T - I)\mathbf{r}$. This estimation method represents a dual approach to estimating productivity changes,[16] computing the average growth rate of an industry's input prices and subtracting the growth rate of its output price, ascribing the difference to improvements by the industry.

Calculation of average improvement rates The average local improvement rate $\hat{\gamma}_c$ for country c was estimated as $\hat{\gamma}_c = \sum_i \eta_{i,c}\,\gamma_{i,c}$, where $\eta_{i,c}$ is the share of industry (i, c)'s gross output in the total gross output of country c.

Test of $\mathbf{r} = -H^T\boldsymbol{\gamma}$ We split the period 1995 - 2009 into three periods of nearly equal length (5 years, 5 years, 4 years), labelled periods I, II, and III, and use data from periods I and II to predict price returns in period III. In each period, we computed the period average price returns $\mathbf{r}_I$, $\mathbf{r}_{II}$, $\mathbf{r}_{III}$, and productivity growth rates $\boldsymbol{\gamma}_I$, $\boldsymbol{\gamma}_{II}$, $\boldsymbol{\gamma}_{III}$. We treated the productivity growth rates as observations from an AR(1) time-series model and fit $\boldsymbol{\gamma}_{II} = a\mathbf{1} + b\boldsymbol{\gamma}_I + \boldsymbol{\varepsilon}$, obtaining the fitted coefficients $\hat{a}$ and $\hat{b}$. We then computed predicted growth rates for the third period as $\hat{\boldsymbol{\gamma}}_{III} = \hat{a}\mathbf{1} + \hat{b}\boldsymbol{\gamma}_{II}$. Since the predicted productivity growth rates use data from only the first two periods they are fully independent of price returns in the third period. We then computed the prediction $\hat{\mathbf{r}}_{III} = -H^T\hat{\boldsymbol{\gamma}}_{III}$ using the Leontief inverse H from the final year of period II. There were thus two effects that could limit the predictive performance of $\hat{\mathbf{r}}_{III} = -H^T\hat{\boldsymbol{\gamma}}_{III}$, the exclusion of Leontief inverse data from period III and the exclusion of productivity estimates from period III.

The predicted and actual price returns have a highly significant correlation (Extended Data Fig. 2a) with p-value $\sim 3 \times 10^{-41}$. This prediction performed significantly better than a prediction based only on the output multipliers, which is expected since the full prediction of the model $\mathbf{r} = -H^T\boldsymbol{\gamma}$ includes the additional information of productivity growth rates across industries and the full network structure as captured by the Leontief inverse. The slope is notably larger than 1, showing that actual returns are larger than predicted ones in this case. Productivity growth rates have been seen to be correlated with output multipliers (Extended Data Fig. 1). Taking this correlation into account (Fig. 2b) results in a slope that is nearly 1. Here we fit productivity growth rates in periods I and II as $\boldsymbol{\gamma}_{II} = a\mathbf{1} + b\boldsymbol{\gamma}_I + c\boldsymbol{\mathcal{L}} + \boldsymbol{\varepsilon}$, where $\boldsymbol{\mathcal{L}}$ are the output multipliers from the final year of period II. We then estimated productivity growth rates in period III as $\hat{\boldsymbol{\gamma}}_{III} = \hat{a}\mathbf{1} + \hat{b}\boldsymbol{\gamma}_{II} + \hat{c}\boldsymbol{\mathcal{L}}$, and computed predicted price returns as before.

Test of prediction for covariances of price returns We compute the covariances of price returns between every pair of industries in the WIOD, leading to about 9.5×10^5 unique covariances after removing industries with zero expenditures. For each pair of industries i and j, we compute $R_{ij} = E[r_i r_j] - E[r_i]E[r_j]$ using the year-to-year price returns r_i and r_j over 1995 - 2009. We compute the variances of each industry m's productivity improvement rates as $D_m = E[\gamma_m^2] - E[\gamma_m]^2$ from the year-to-year productivity changes over 1995 - 2009. We then compute predicted price return covariances as $R_{ij} = \sum_m H_{mi} D_m H_{mj}$, using the Leontief inverse H from the initial year 1995.

Extended Data

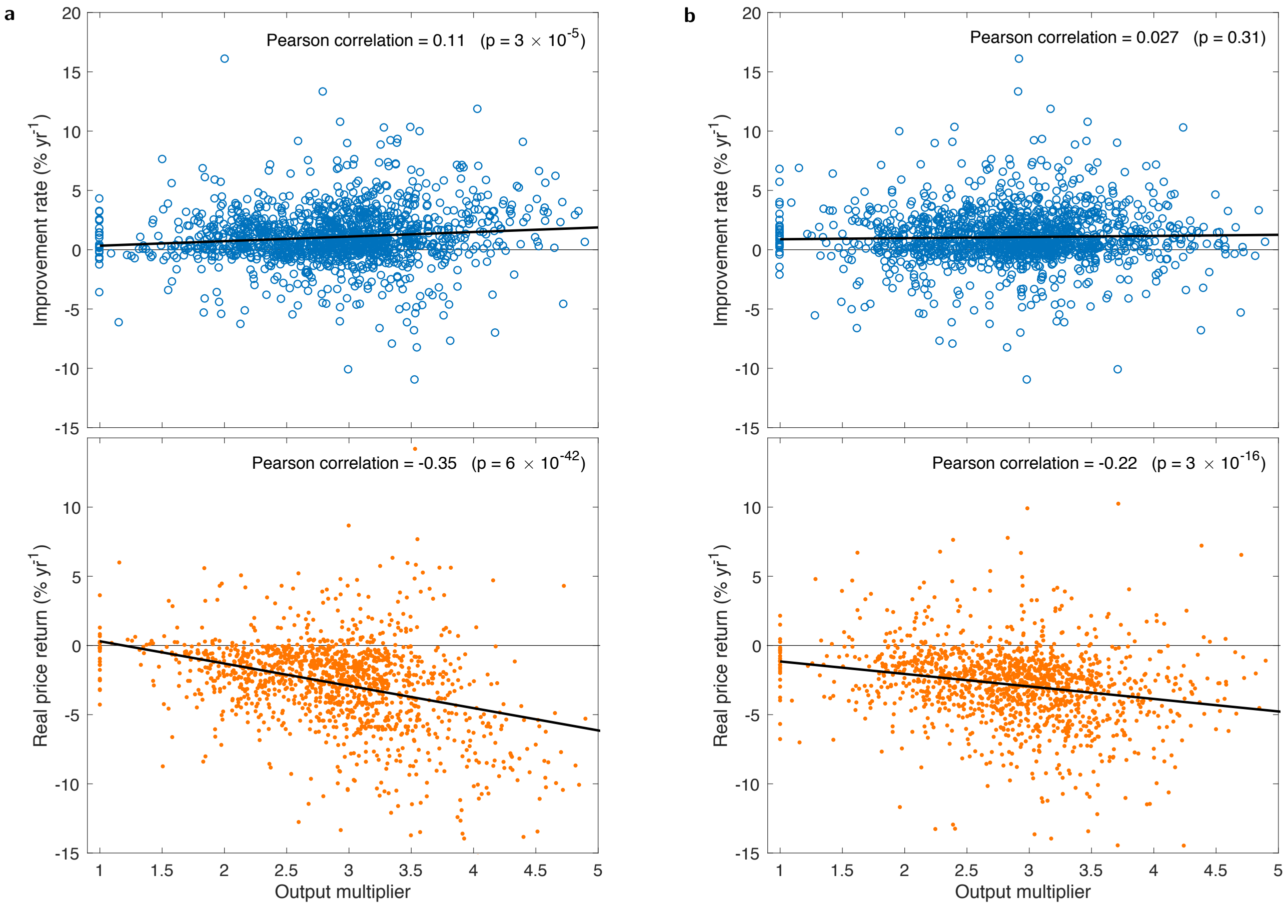

Figure 1: (Extended Data) **Correlation between returns and output multipliers before and after shuffling improvement rates.** **a**, Observed improvement rates and price returns versus output multipliers. Improvement rates have a small positive correlation with output multipliers. **b**, Improvement rates were shuffled across industries to remove the correlation with the output multipliers. Resulting industry returns were then computed with these improvement rates using the model. Results vary from one shuffle to another, with those shown here being typical.

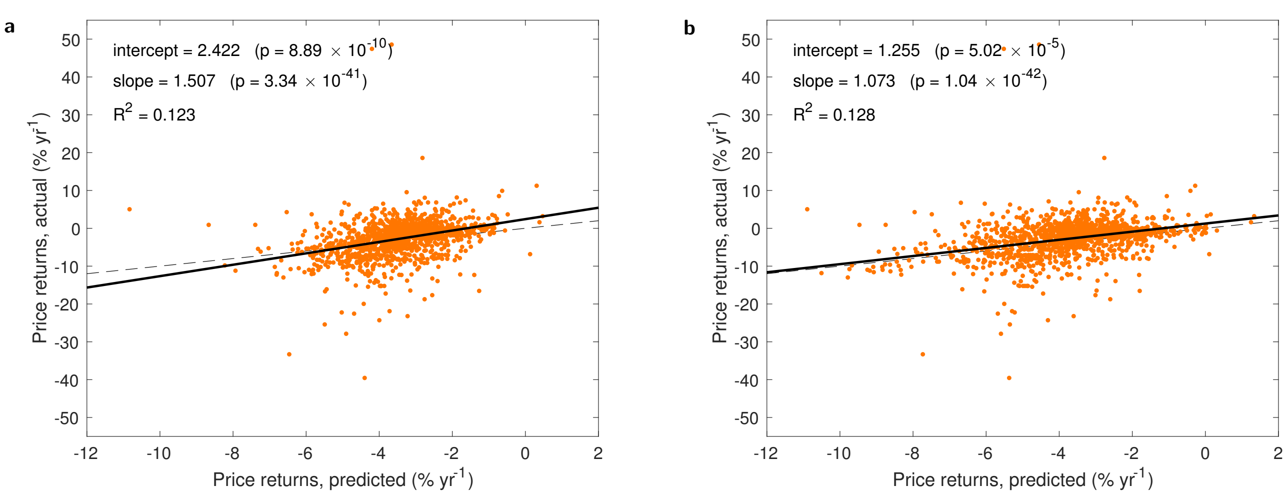

Figure 2: (Extended Data) **Comparison of actual price returns with predicted values based on $\mathbf{r} = -H^T\boldsymbol{\gamma}$.** See text for description of the test procedure. Dashed lines are the identity lines. **a** Results from using the time-series model $\hat{\boldsymbol{\gamma}}_{III} = \hat{a}\mathbf{1} + \hat{b}\boldsymbol{\gamma}_{II}$. **b** Results from using the time-series model $\hat{\boldsymbol{\gamma}}_{III} = \hat{a}\mathbf{1} + \hat{b}\boldsymbol{\gamma}_{II} + \hat{c}\boldsymbol{\mathcal{L}}$.

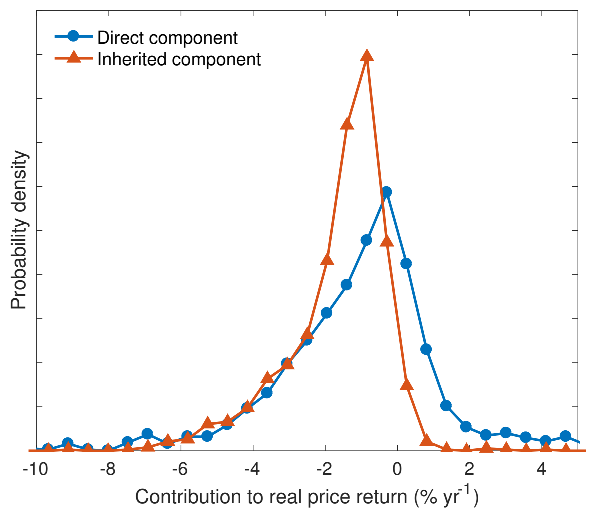

Figure 3: (Extended Data) **Probability densities of direct and inherited components of price returns.** The price return r_i of industry i can be decomposed as $r_i = -\gamma_i + \sum_j r_j a_{ji}$, where $-\gamma_i$ is i's direct contribution to its price reduction and $\sum_j r_j a_{ji}$ is the contribution from inherited price changes passed to i through input goods. The distribution of each component is shown using a histogram with 30 evenly-spaced bins.

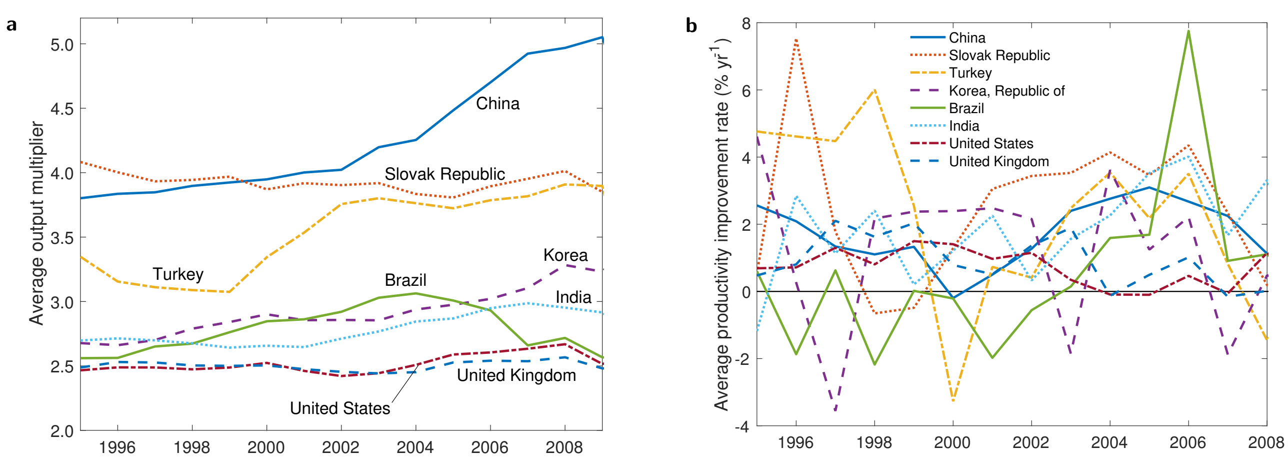

Figure 4: (Extended Data) **Average output multipliers and average productivity improvement rates over time.** Average output multipliers (a) and improvement rates (b) for a selection of countries over the period 1995 - 2009.

Code	Industry	Average $\mathcal{L}_i$ †	Average γ_i † (% yr^{-1})	Average r_i † (% yr^{-1})
Cok	Coke, Refined Petroleum and Nuclear Fuel	3.66(±0.30)	−1.05(±4.79)	0.32(±4.95)
Tpt	Transport Equipment	3.63(±0.53)	1.93(±2.31)	−4.12(±3.14)
Chm	Chemicals and Chemical Products	3.60(±0.39)	1.86(±2.21)	−3.77(±2.95)
Elc	Electrical and Optical Equipment	3.55(±0.49)	2.66(±2.30)	−5.66(±3.39)
Met	Basic Metals and Fabricated Metal	3.54(±0.44)	0.78(±3.01)	−2.35(±4.04)
Rub	Rubber and Plastics	3.48(±0.44)	2.02(±2.18)	−4.19(±3.06)
Ele	Electricity, Gas and Water Supply	3.47(±0.48)	0.22(±2.28)	−1.51(±2.87)
Fod	Food, Beverages and Tobacco	3.44(±0.29)	0.27(±1.85)	−3.05(±2.90)
Mch	Machinery, Nec	3.37(±0.41)	1.87(±2.84)	−3.84(±3.41)
Ait	Air Transport	3.36(±0.43)	1.69(±3.32)	−3.59(±4.02)
Wtt	Water Transport	3.33(±0.33)	0.81(±2.32)	−2.56(±3.11)
Lth	Leather, Leather and Footwear	3.31(±0.41)	1.77(±1.89)	−3.94(±2.81)
Tex	Textiles and Textile Products	3.30(±0.40)	2.07(±1.73)	−4.43(±2.62)
Mnf	Manufacturing, Nec; Recycling	3.30(±0.44)	1.76(±2.19)	−3.77(±3.24)
Pup	Pulp, Paper, Paper , Printing and Publishing	3.30(±0.40)	1.90(±1.49)	−4.23(±2.71)
Est	Real Estate Activities	3.30(±0.60)	0.04(±1.63)	−1.59(±1.92)
Wod	Wood and Products of Wood and Cork	3.29(±0.38)	1.05(±2.31)	−3.38(±3.59)
Omn	Other Non-Metallic Mineral	3.29(±0.37)	1.77(±1.51)	−3.48(±2.42)
Cst	Construction	3.19(±0.46)	−0.48(±1.42)	−1.23(±2.25)
Otr	Other Supporting and Auxiliary Transport Activities; Activities of Travel Agencies	3.02(±0.58)	0.80(±1.96)	−2.36(±2.35)
Min	Mining and Quarrying	3.01(±0.36)	−0.93(±2.94)	−0.38(±3.24)
Pst	Post and Telecommunications	2.85(±0.44)	2.69(±2.96)	−4.77(±3.17)
Ldt	Inland Transport	2.83(±0.45)	0.91(±1.12)	−2.33(±1.85)
Htl	Hotels and Restaurants	2.83(±0.39)	0.27(±1.48)	−1.94(±2.10)
Sal	Sale, Maintenance and Repair of Motor Vehicles and Motorcycles; Retail Sale of Fuel	2.68(±0.47)	1.23(±2.08)	−2.62(±2.58)
Agr	Agriculture, Hunting, Forestry and Fishing	2.66(±0.53)	2.72(±1.77)	−4.49(±2.73)
Whl	Wholesale Trade and Commission Trade, Except of Motor Vehicles and Motorcycles	2.66(±0.44)	1.38(±1.32)	−2.91(±2.01)
Ocm	Other Community, Social and Personal Services	2.63(±0.40)	−0.11(±1.41)	−1.15(±1.75)
Obs	Renting of M&Eq and Other Business Activities	2.61(±0.45)	−0.03(±1.23)	−1.40(±1.61)
Fin	Financial Intermediation	2.57(±0.39)	1.87(±2.43)	−3.38(±3.23)
Rtl	Retail Trade, Except of Motor Vehicles and Motorcycles; Repair of Household Goods	2.43(±0.43)	1.54(±1.30)	−2.78(±1.95)
Hth	Health and Social Work	2.29(±0.46)	−0.59(±1.71)	−0.51(±1.83)
Pub	Public Admin and Defence; Compulsory Social Security	2.17(±0.38)	0.17(±1.63)	−1.18(±1.90)
Edu	Education	1.75(±0.38)	−0.25(±2.09)	−0.45(±2.13)
Pvt	Private Households with Employed Persons	1.07(±0.27)	0.57(±1.57)	−0.65(±1.66)

† Averages are over countries and the period 1995 - 2009. Numbers in parentheses give standard deviations across countries.

Table 1: (Extended Data) **Cross-country average properties of industries from the WIOD dataset.**

Code	**Country**	**GDP per cap. in 1995** (2011 PPP$)	**Ave. growth per cap. (1995 - 2009)** (% yr^{-1})	**Ave. improvement rate $\tilde{\gamma}_c$ (1995 - 2009)** (% yr^{-1})	**Ave. output multiplier $\bar{\mathcal{L}}_c$ (1995 - 2009)**
AUS	Australia	30,347	2.18	0.12	2.89
AUT	Austria	33,544	1.65	0.25	2.58
BEL	Belgium	32,361	1.51	0.16	2.90
BGR	Bulgaria	8,434	4.19	0.30	3.40
BRA	Brazil	11,012	1.47	0.55	2.79
CAN	Canada	32,100	1.55	0.71	2.67
CHN	China	2,550	8.65	1.74	4.26
CYP	Cyprus	26,444	1.64	0.74	2.33
CZE	Czech Republic	19,093	2.62	2.30	3.49
DEU	Germany	33,849	1.01	0.31	2.53
DNK	Denmark	36,670	1.05	0.48	2.49
ESP	Spain	25,630	1.83	-0.10	2.79
EST	Estonia	11,068	4.78	2.81	3.12
FIN	Finland	27,303	2.45	0.65	2.78
FRA	France	30,822	1.15	0.62	2.60
GBR	United Kingdom	28,513	1.64	0.92	2.50
GRC	Greece	21,641	2.57	1.48	2.62
HUN	Hungary	15,136	2.65	0.92	3.05
IDN	Indonesia	6,022	2.09	0.40	2.95
IND	India	2,058	4.91	1.83	2.78
IRL	Ireland	26,002	3.88	1.04	2.95
ITA	Italy	32,730	0.53	-0.06	2.76
JPN	Japan	31,224	0.37	0.35	2.66
KOR	Korea, Republic of	16,798	3.83	1.19	2.91
LTU	Lithuania	9,229	5.53	2.62	2.84
LUX	Luxembourg	64,018	2.23	-0.32	2.77
LVA	Latvia	8,145	5.78	1.90	3.07
MEX	Mexico	12,609	1.16	-0.41	3.19
MLT	Malta	20,720	1.86	0.78	2.90
NLD	Netherlands	35,005	1.86	0.42	2.67
POL	Poland	11,149	4.40	0.17	2.89
PRT	Portugal	21,974	1.44	0.30	2.77
ROM	Romania	10,271	3.76	1.85	3.10
RUS	Russia	12,012	3.90	1.74	2.73
SVK	Slovak Republic	12,876	4.25	2.46	3.93
SVN	Slovenia	18,244	3.10	1.04	2.79
SWE	Sweden	31,044	1.96	0.68	2.69
TUR	Turkey	11,530	2.10	2.24	3.54
TWN	Taiwan	no data	3.22	1.32	2.77
USA	United States	39,476	1.48	0.73	2.52
RoW	Rest of World	9,139	N.A.	N.A.	2.87

The first column gives the ISO code of each country. GDP per capita data is from the World Bank.[31]

Table 2: (Extended Data) **Summary statistics for countries in the WIOD dataset for 1995 - 2009.**

Industry code	Industry name	Slope	p-value	
Agr	Agriculture, Hunting, Forestry and Fishing	-2.21	1.08×10^{-3}	**
Min	Mining and Quarrying	-0.28	0.83	
Fod	Food, Beverages and Tobacco	-3.66	5.51×10^{-4}	***
Tex	Textiles and Textile Products	-2.94	1.82×10^{-3}	**
Lth	Leather, Leather and Footwear	-2.85	2.77×10^{-3}	**
Wod	Wood and Products of Wood and Cork	-4.61	1.44×10^{-4}	***
Pup	Pulp, Paper, Paper , Printing and Publishing	-3.77	1.21×10^{-5}	***
Cok	Coke, Refined Petroleum and Nuclear Fuel	-0.08	0.98	
Chm	Chemicals and Chemical Products	-3.36	1.70×10^{-3}	**
Rub	Rubber and Plastics	-3.83	1.84×10^{-4}	***
Omn	Other Non-Metallic Mineral	-4.41	2.61×10^{-8}	***
Met	Basic Metals and Fabricated Metal	-2.29	0.11	
Mch	Machinery, Nec	-4.36	3.08×10^{-4}	***
Elc	Electrical and Optical Equipment	-3.88	6.68×10^{-4}	***
Tpt	Transport Equipment	-2.25	9.43×10^{-3}	**
Mnf	Manufacturing, Nec; Recycling	-4.53	4.78×10^{-7}	***
Ele	Electricity, Gas and Water Supply	-0.69	0.41	
Cst	Construction	-2.00	8.19×10^{-3}	**
Sal	Sale, Maintenance and Repair of Motor Vehicles and Motorcycles; Retail Sale of Fuel	-1.78	0.02	*
Whl	Wholesale Trade and Commission Trade, Except of Motor Vehicles and Motorcycles	-1.78	4.02×10^{-3}	**
Rtl	Retail Trade, Except of Motor Vehicles and Motorcycles; Repair of Household Goods	-2.08	1.26×10^{-4}	***
Htl	Hotels and Restaurants	-2.27	3.39×10^{-3}	**
Ldt	Inland Transport	-1.87	1.02×10^{-3}	**
Wtt	Water Transport	-0.68	0.57	
Ait	Air Transport	-4.32	6.18×10^{-4}	***
Otr	Other Supporting and Auxiliary Transport Activities; Activities of Travel Agencies	-0.59	0.38	
Pst	Post and Telecommunications	0.74	0.42	
Fin	Financial Intermediation	-0.65	0.55	
Est	Real Estate Activities	-1.04	0.03	*
Obs	Renting of M&Eq and Other Business Activities	-1.02	0.09	
Pub	Public Admin and Defence; Compulsory Social Security	-1.54	0.01	*
Edu	Education	-2.30	6.30×10^{-3}	**
Hth	Health and Social Work	-1.02	0.08	
Ocm	Other Community, Social and Personal Services	-1.34	0.05	
Pvt	Private Households with Employed Persons	-1.69	0.07	

Significance levels: * = 0.05, ** = 0.01, *** = 0.001.

Table 3: (Extended Data) **Regression slopes and p-values for yearly price returns by industry.**

Explanatory variable	R^2	p-value	Correlation with $\mathcal{L}$	p-value
Model prediction $\tilde{\gamma}\mathcal{L}$	0.602	4.00×10^{-9}	0.603	3.84×10^{-5}
Improvement rate $\tilde{\gamma}$	0.568	1.99×10^{-8}	0.454	3.23×10^{-3}
Gross capital formation	0.478	1.09×10^{-6}	0.614	3.17×10^{-5}
Average output multiplier $\bar{\mathcal{L}}$	0.366	3.51×10^{-5}	1.000	0
Urban population	0.360	5.36×10^{-5}	-0.390	1.41×10^{-2}
Health expenditure	0.360	5.48×10^{-5}	-0.553	2.61×10^{-4}
TFP level	0.264	6.88×10^{-4}	-0.430	5.56×10^{-3}
Labor share of gross output	0.239	1.38×10^{-3}	-0.883	4.89×10^{-14}
School enrollment	0.220	2.94×10^{-3}	-0.397	1.35×10^{-2}
Tax revenue	0.186	6.07×10^{-3}	-0.378	1.77×10^{-2}
Population growth rate	0.128	2.55×10^{-2}	-0.146	3.76×10^{-1}
Savings rate	0.115	3.47×10^{-2}	0.321	4.66×10^{-2}
Researchers in R&D	0.098	5.28×10^{-2}	-0.403	1.09×10^{-2}
Labor share of GDP	0.075	8.82×10^{-2}	-0.486	1.48×10^{-3}
log(GDP per capita in 1995)	0.068	1.08×10^{-1}	-0.218	1.83×10^{-1}
Inflation rate	0.043	2.07×10^{-1}	0.345	3.15×10^{-2}
Depreciation rate	0.022	3.58×10^{-1}	-0.019	9.07×10^{-1}
Index human capital	0.012	5.08×10^{-1}	-0.156	3.36×10^{-1}

Table 4: (Extended Data) **Correlations between growth rates and variables associated with growth.** Variables are averages over the period 1995 - 2009.