2021

What You Exported Matters: Persistence in Productive Capabilities across Two Eras of Globalization

Isabella M. Weber

Gregor Semieniuk,

Tom Westland

Junshang Liang

What You Export*ed* Matters: Persistence in Productive Capabilities across Two Eras of Globalization

Isabella M. Weber,[1] Gregor Semieniuk,[2] Tom Westland,[3] Junshang Liang[4]

27 January 2021

Abstract
Does what you exported matter? We build a new global commodity-level export database for the previous era of globalization and find persistence in productive capabilities proxied by economic complexity, export diversification, and sophistication across a century. We also show that productive capabilities at the turn of the 20th century are a powerful predictor of today's income levels. We demonstrate that our results are not driven by persistence in geography or institutions. The persistence mechanism is the complementarity between past and future productive capabilities with one important qualification, the persistent negative effect of European overseas colonization. We also study shocks that undermined persistence, confirm the resource curse hypothesis for the long run and find a positive but slow effect of democratization.
JEL codes: F14, F63, N10, O10, O50

I. INTRODUCTION

Why are some nations rich and others poor? This is one of the cardinal questions in economics. In recent years, a growing number of persistence studies locates the answer in the distant past (Voth, 2021). Most focus on non-economic factors as drivers of long-term growth such as culture, institutions, geography and genetics. In this paper, we analyze persistence in a key economic aspect: countries' productive capabilities. Complexity economics demonstrates that export data can reveal countries' productive capabilities (Hidalgo & Hausmann, 2009) and that what you export matters for how fast you grow (Hausmann, Hwang, & Rodrik, 2007; S. Lall, 2000). We construct a new global export database on the commodity level for the previous era of globalization. We show that countries' productive capability rankings are persistent across the two eras of globalization (1897-1906 to 1998-2007) and that productive capabilities a century ago can predict today's GDP per capita.

The importance of productive capabilities for economic growth is a long-standing insight (Dahlman, Ross-Larson, & Westphal, 1987; Fagerberg & Srholec, 2008). Recently Weitzman (1998), Hausmann and Hidalgo (2011) and van Dam & Frenken (2020) have argued that initial productive capabilities matter: The more capabilities a country has, the more it can develop. If new products require a recombination of certain tasks, then having the capability to carry out

[1] Department of Economics and Political Economy Research Institute, University of Massachusetts Amherst
[2] Political Economy Research Institute and Department of Economics, University of Massachusetts Amherst
[3] Department of History, University of Cambridge
[4] Department of Economics, University of Massachusetts Amherst

these tasks improve the likelihood to develop the new product. The capabilities to produce batteries and internal combustion engine cars certainly help with developing electric vehicles. This combinatorial feature makes productive capabilities a natural candidate for persistence.

Export competitiveness can proxy capabilities. Export diversification reveals insights about the variety of capabilities while complexity and sophistication of exports add information on their quality (Hidalgo & Hausmann, 2009; Lectard & Rougier, 2018). Until now, the long-run persistence of productive capabilities could not be analyzed due to a lack of historical data. We employ colonial statistics to create a new database that enables us to measure export diversification, complexity and sophistication in a consistent fashion across a century.

We use the previous globalization as reference period since this was both the first time that the whole world was integrated into international trade which is necessary to use exports as meaningful proxy. It also coincided with the high point of colonial statistics which enables us to achieve almost global coverage and to disaggregate our data to the level of specific commodities. To the best of our knowledge ours is the first database that goes beyond Lamartine Yates's (1959) 'census' of world trade for the year 1913 that records broad commodity types.[5] To make our historical data compatible with UN COMTRADE we translate commodity classifications in primary sources into the Standard International Trade Classification. As a result, we can measure the same variables for both the historical reference period (1897-1906) and the contemporary one (1998-2007) – unlike most other persistence studies (Bisin & Moro, 2021).

Figures 1a and 1b map countries' export diversification by decile in 1897-1906 to 1998-2007 measured as the number of different commodities exported by a country.[6] This illustrates our data coverage and the trend of persistence. Europe and the United States occupy the top rank then and now while Asian countries are largely in the upper medium ranks in both periods. Some notable cases of falling behind are Russia and the Middle East as well as several African countries, while Brazil and Argentine are hopeful cases of catching up.

Our regression analysis shows that all our proxies of productive capabilities (diversification, complexity and sophistication) for the previous era of globalization are highly statistically significant and quantitatively important predictors of their respective contemporary values. Put simply, what you exported matters: Countries that lacked productive capabilities then, tend to be trapped and catching up is the exception rather than the rule. This is robust to varying the revealed comparative advantage threshold and SITC digit level as well as to adding standard control variables in the trade diversification and growth literature including resource rent, population, human capital, trade openness, FDI and polity. Based on all three proxies, persistence across the two waves of globalization is quantitatively vastly more important than the whole range of controls combined. This finding is confirmed by principal component analysis for which we add the manufacturing share as a more traditional measure of industrial development. We also find persistence in relation to the intermediate periods 1962-1971 and 1979-1988 and can therefore rule out re-emergence.

[5] For want of more comprehensive data, Yates's (1959) work is still used in exercises that decompose trade in the previous era of globalization to the sectoral level (Findlay & O'Rourke, 2007; Galor, 2005; O'Rourke, 2018).
[6] Revealed comparative advantage by country x in commodity y is defined as the share of x in the value of y's exports divided by the share of x in the value of global exports (Balassa, 1965).

A robust stylized fact in development economics is that the ways in which countries are integrated into global trade matters for their prosperity.[7] We test this relation for the first time for the long run of a century and find the more diversified, complex and sophisticated a country's exports at the turn of the 20th century, the higher the wealth of the nation today. 56 per cent of the variance in today's per capita GDP can be explained by countries' economic complexity in 1897-1906 (see Table 6 Column 3). All our measures of productive capabilities in the previous era of globalization are highly significant at the 1 per cent significance level in relation to per capita GDP in the present period. This finding is robust to the aforementioned variations in data configuration and controls.

We adopt Comin, Easterly and Gong's (2010) model for persistence in technology to devise a simple framework to analyze drivers of persistence in productive capabilities. We test two alternative persistence mechanisms: complementarities between past and future capabilities or a persistent third variable, such as geography or institutions. A large and influential literature places geography center stage to explain growth, development and trade (Frankel & Romer, 1999; Gallup, Sachs, & Mellinger, 1998; Sachs, 2001, 2003). Similarly, institutions have been foregrounded as primary drivers of a country's economic performance (Acemoglu, Egorov, & Sonin, 2021; Rodrik, Subramanian, & Trebbi, 2004). We show that the correlates of our proxies measured a century ago do not explain productive capabilities today. This is subject to an important exception: The productive capabilities of European overseas colonies were significantly lower than those of other countries across both periods. Meanwhile, complementarities are a good explanation, even after controlling for persistent covariates.

We follow Voigtländer and Voth (2012) and examine cases of path-defiance. We analyze policy and institutional shocks as well as changes to natural endowments. We find that broad brush proxies for development policy paradigms do not alter persistence, while the path-defying effect of democratization appears to be slow. Countries that discovered oil showed downward path-defiance, significantly reducing their export diversification, complexity and sophistication. This confirms the resource curse hypothesis for the long run (Frankel, 2010).

Our results contribute to the persistence literature that examines historical sources of contemporary development (Nunn, 2020). A set of recent studies uncovers the importance of a variety of institutional, geographic and cultural factors that persist over long periods and can be linked to contemporary economic outcomes (Voth, 2021). A prominent finding is the importance of colonial heritage in the long run (Dell, 2010; Nunn & Wantchekon, 2011) which is confirmed in our study. Unlike most of these studies, we measure economic indicators directly in the historical period. Moreover, we use the same measures in the historical and in the contemporary period (Bisin & Moro, 2021). This is thanks to our rich data on commodity-level exports. Along with tax data that has recently been harnessed for long-term economic analysis (Piketty & Saez, 2003), trade statistics are some of the earliest detailed systematic records that can be aggregated to the national level and they are available on an almost global scale.

[7] See for example Hausmann et al. (2007); Hausmann & Hidalgo (2011); Hesse (2009); Hirschman (1958); Hidalgo, Winger, Barabási, & Hausmann (2007); Lederman & Maloney (2003); Prebisch (1950); Rosenstein-Rodan (1943); Zhu & Li (2017).

The wide coverage of countries across a century relates our study to the convergence literature. Few topics in economics are as consequential both for theory and policy as the convergence hypothesis, which states that "initial conditions have no implications for the long-run distribution of per capita income" (Johnson & Papageorgiou, 2020, 130). A recent literature review shows that the convergence hypothesis is not backed by empirical evidence, yet the mechanisms preventing convergence are poorly understood (ibid.). The data on GDP per capita over time is not universally robust, even in the post-war period (Jerven, 2013). This creates a need to consider other metrics of convergence and divergence across a longer time span. Our paper suggests that initial conditions in productive capabilities matter in the long-run and affect GDP outcomes even today.

Our study also speaks to development economics seeking to unlock conditions for successful growth. A large body of literature investigates how export diversification can serve growth (Lederman & Maloney, 2012). Our paper highlights the importance of long-term persistence in export patterns which points to important constraints. We find that many of the usual policy variables (e.g. trade openness, FDI, human capital, institutions) did not diminish the importance of past productive capabilities in explaining today's development outcomes. Path-defiant cases therefore merit further study to identify policy strategies that succeeded in undermining persistence.

The paper proceeds as follows. Section II provides historical background on the previous globalization and introduces our data and measures of productive capabilities. In section III, we describe our empirical strategy and present our main results. Section IV analyzes the persistence mechanism and cases of path-defiance. Section V concludes.

II. HISTORICAL CONTEXT, DATA AND MEASURES

Our data covers a decade of merchandise exports for the previous and current waves of globalization (1897-1906 and 1998-2007). We construct a new global commodity-level database. We exploit colonial trade data, among the richest of colonial records, and record commodity level export data for 100 countries. This is about double the number of countries for which the Maddison Project reports GDP in 1900 and 1913 (Bolt & van Zanden, 2020). We use three export-based measures devised by Hausman and Hidalgo (2009) and others (diversification, economic complexity and export sophistication) as proxies for productive capabilities for both periods. This allows us to study the persistence of countries' position on the global development ladder across a century.

A Two Eras of Globalization and Colonial Trade Statistics

To test the long-term global persistence of productive capabilities using export data it is desirable to compare two periods of roughly similar degrees of global market integration. Previous studies on the evolution of economic complexity, diversification and underlying productive capabilities have been limited to the post-1962 period covered by the UN COMTRADE database (Cadot, Carrère, & Strauss-Kahn, 2013; Hausmann et al., 2007), which spans a phase of high levels of protectionism followed by a phase with increasingly free markets. For this paper we have

constructed a new dataset for the previous era of globalization which enables us to test persistence across two periods of high levels of market integration and across a considerably longer time span.

The previous era of globalization (1870-1913) was the first time that nearly all territories and peoples of the world were comprehensively integrated in a global division of labor. As such this was a formative period for global export patterns. Long-distance trade had gained importance in earlier periods, but only in the late 19th century globalization fundamentally shaped the internal organization of economies around the world (O'Rourke & Williamson, 2002; Williamson, 2011; Pascali, 2017). World trade boomed from 1870 to 1913 accounting for the first time for more than 20 per cent of world GDP (Estevadeordal, Frantz, & Taylor, 2003).[8] Measures such as global price convergence, savings-investment correlations and current account imbalances demonstrate global market integration (Jones & Obstfeld, 1997; Obstfeld & Taylor, 2004; Schularick, 2006; Taylor et al., 2002). The trade boom was facilitated by a transport and communication revolution (steamships, railways and telegraphs); backed by the gold standard and internationalizing financial markets; and fueled by a "growth miracle" in the capitalist core countries (Federico & Persson, 2007; Jacks, Meissner, & Novy, 2009; Lewis, 1969; Pascali, 2017; Williamson, 2011).

With the first World War the previous globalization started to disintegrate. The post-World War II period was marked by decades of protectionism, the spread of state socialism and state-led development. The world only began to re-globalize in the 1980s under the renewed rise of free markets. According to the Sachs-Warner index, the share of countries with free trade policies was as low as 15.6 per cent in 1960, but increased from 26 per cent in 1985 to 73 per cent in 2000 (Wacziarg & Welch, 2003).

We compare two decades of data across a century. This allows us to trace the persistence in productive capabilities in relation to the first period of a comprehensive and truly global division of labor. For our comparison across a century, our main concern in choosing our contemporary reference period is to capture globalization in its maturity. It is contested whether the ongoing wave of globalization is ebbing (O'Rourke, 2018) but the pre-2008 period is certainly part of its high point before a return of protectionism (Fajgelbaum, Goldberg, Kennedy, & Khandelwal, 2020). Trade openness measured in terms of the share of global merchandise trade in world GDP peaked in 2008 at 24.6 per cent (UNCTAD, 2020). We use the decade before the Great Recession (1998-2007) and after the Asian financial as a period comparable to the last era of globalization in terms of the relative absence of a major global crisis. For this contemporary period we rely on UN COMTRADE data available at the Observatory of Economic Complexity.[9]

We construct our database for the decade 1897-1906 for which we achieve close to global coverage (see Table 1 and Figure 1a). Peak openness in the last era of globalization varies across countries with the United States and Canada peaking as early as 1900 and Japan as late as 1929 (Bairoch & Kozul-Wright, 1996, 6). The end of the last era of globalization is usually dated as

[8] Estimates of the rate of growth of GDP and of trade in the period 1870-1913 vary but trade certainly outpaced output (Bairoch & Kozul-Wright, 1998, 5). Mitchie & Kitson (1995) suggest that output growth was 2.7 per cent and trade growth 3.5 per cent; Maddison (1989) estimates 2.5 per cent and 3.9 per cent respectively.
[9] The data is available here: https://legacy.oec.world/en/resources/data/.

1913. 1897-1906 is sufficiently close to this end point to capture mature globalization and marks a decade without major global crises.[10]

By the turn of the 20th century the global division of labor brought about by the Great Divergence (Pomeranz, 2000) was firmly established. At the onset of the industrial revolution in 1750, the share in world manufacturing output of Asia, Africa and Latin America combined accounted for 73 per cent and in 1860 still for 36.6 per cent. By 1913 this share had fallen to 7.5 per cent and that of Europe, North America and Japan conversely increased to 92.5 per cent (Bairoch, 1982; Galor, 2011; Nayyar, 2013a). As a result of this increasing asymmetry, living standards diverged drastically in the 19th century with the new industrial core pulling ahead (Findlay & O'Rourke, 2007, 414; O'Rourke & Williamson, 1999; Pritchett, 1997; Williamson, 2011, 6). This paper explores whether the asymmetry in global productive capabilities brought about by the Great Divergence persists until the present era of globalization.

In constructing our database, we exploit the fact that the turn of the twentieth century represents the high point in terms of colonial trade statistics. This was the Age of High Imperialism and the peak of colonial administrative capacity (Ferguson, 2002; Mitchener & Weidenmier, 2008). By then Europe's imperial powers had opened up the rest of the world through direct colonization as in the case of most of Africa or through gunboat diplomacy towards the nominally independent countries as in several Asian countries (Bairoch & Kozul-Wright, 1998; Findlay & O'Rourke, 2007b; Nayyar, 2006, 2013a; Williamson, 2002). When globalization came to a close in 1914, Europe controlled 84 per cent of the world's surface which had increased from 67 per cent in 1878 (Headrick, 1981, 3).

Export data are some of the richest, and most accurate, of the colonial statistics. Trade data are more reliable than most colonial statistics. Censuses were only rarely undertaken, and were often little more than guesstimates, especially in African colonies (Frankema & Jerven, 2014). Detailed data on the overall structures of colonial economies was rarely collected in any systematic fashion in this period. In contrast, colonial governments relied on trade taxes for revenue and therefore kept careful records of the international flow of goods (Gardner, 2013). This was the case both for colonies and the colonial home countries of Europe which at the time experienced a protectionist backlash that ended the preceding period of openness. The self-governing white settler colonies and some Latin American countries also raised tariffs and to this end kept careful trade statistics (Bairoch & Burke, 1989; Findlay & O'Rourke, 2003, 2007; O'Rourke & Williamson, 1999, 93-118; Williamson, 2006, 2011). The United Kingdom stands as major exception of continued economic openness at home but nonetheless collected detailed statistics on its own trade, and, as the ruler of the most expansive colonial empire, ensured that statistical processes were implanted in the territories it ruled (Touchelay, 2019). After the First World War, French data becomes much less available in compiled form, and though the British statistical compendia were still published, export statistics begin to become generally less detailed at the commodity level.

[10] Of course, no single decade will be perfectly free from political or economic change: our period includes, for example, the Russo-Japanese War, the Constitutional Revolution in Persia, and the colonization of several parts of the African continent by European powers. However, for the most part major economic crises were absent: the Long Depression in Europe and the United States had ended by the middle of the 1890s, as had the aftershocks of the Barings Crisis of 1890 and the Encilhamento in Brazil.

We collect our data from various sources in five languages. Exports from British colonies were taken from the annual *Blue Books*, annual *Colonial Reports*, and the *Statistical Abstract of the British Empire*. These were all official statistical documents, with data on exports compiled from colony-level customs records. Data for French colonies were compiled from the annual *Statistiques Coloniales: Commerce*, a comprehensive statistical compendium drawn up by the Ministry of Colonies in Paris from colony-level customs records. The *Statistiques Coloniales* were compiled rigorously prior to the First World War, but cease thereafter, replaced by a piecemeal set of colony-specific documents. Trade figures for the German overseas empire were published annually in the official *Statistisches Jahrbuch für das Deutsche Reich*. For some countries where systematic official data is lacking, we used contemporaneous non-government sources in Portuguese and Spanish, and drew on several country-specific sources.

Data on other countries were mainly gathered from diplomatic reports on trade, principally those drawn up by British consular officials in the countries in question; though potentially less reliable than official country-level statistics, the consular reports are available for countries which did not publish statistical publications of their own; British consuls often obtained figures directly from customs offices. Data for most European countries, and a portion of independent countries and colonies, were taken from the US compilation *Statistical Abstract of Foreign Countries* (1909). This was a document drawn up by the US Department of Commerce in order to provide comprehensive statistical information on world trade for the benefit of American businesses.

B Data Overview

We have constructed a new database of global merchandise exports that allows us to create a unique snapshot of the economic development around the world at the turn to the 20th century. Digitizing comprehensive primary sources on trade flows, we have recorded export data for up to 100 countries on the commodity-level. This compares with only 45 and 69 countries for which the Maddison Project reports GDP in 1900 and 1913 respectively, and of which only one country is in Sub-Saharan Africa (Bolt & van Zanden, 2020).

To our knowledge this is the first such global database recording commodity-level exports and the first effort to go beyond Lamartine Yates's (1959) 'census' of global trade for the year 1913. Yates hand-coded exports from primary sources. However, he did not publish the underlying data, and his sources are not rigorously documented. Furthermore, his coding was—while extremely detailed for historical work in the 1950s—not highly disaggregated when compared to our dataset: he listed 57 separate commodities, while we document nearly ten times as many, with 545 unique SITC codes recorded in our historical dataset. Other recent datasets only record aggregate trade flows (Federico & Tena-Junguito, 2019) or are limited to regional or country-level coverage (Frankema, Williamson, & Woltjer, 2018; Meissner & Tang, 2018).

Our dataset provides near-global coverage for ten years 1897-1906 of commodity level exports. In order to stay true to our aim to compile detailed commodity-level data, some countries had to be excluded that could enter the aggregate datasets, like Afghanistan or Mongolia. Due to difficulties in collecting data on trade that did not flow via ports, trade data tends to be sparse in

our period for landlocked developing countries, such as those in central Asia or central Africa (in the latter case, statistics would become more detailed as colonial rule was entrenched). In one case, that of Nepal, we could reconstruct the commodity composition of exports by relying on Indian trade statistics, since this was overwhelmingly where Nepalese exports were directed. Nevertheless, the resulting dataset is close to other estimates of global total exports, as Table 1 shows. For the decade 1897-1906 we cover no less than 80% of global trade as measured by Federico and Tena-Junguito (2019) in every year except 1897, and 84% on average.

Our country-level aggregates might be more accurate than those recorded in aggregate trade databases. Recording commodity level data, we avoid certain problems that beset aggregate datasets from that period: for example, our dataset explicitly excludes monetary bullion. This is in contrast to other datasets in which bullion-inclusive aggregate export figures have been taken from primary sources with either little or incomplete correction, (this is often the case, for example, in the trade aggregates presented in Mitchell's *Historical Statistics* compilations). Because we worked from the commodity level, we could explicitly exclude such flows.

We constructed our dataset as follows. Historical records of trade statistics were scanned and converted to digital format using optimal character recognition technology. The original source documents were compared with the digitized versions and mistakes corrected manually by at least two checkers. The quality of our data has been ensured by performing a number of robustness checks. This has included checking for unusually large changes in exports between years and benchmarking total exports against existing datasets for aggregate trade data. All export flows in local currencies were converted to US dollars at prevailing exchange rates, using the Federico and Tena-Junguito (2019) market exchange rate dataset.

We compiled the data on the basis of *special exports*: that is, all domestic exports of goods produced wholly or partly within each country, excluding transit trade, re-exports, and purely monetary flows of bullion and specie. We therefore get as close as possible to a dataset listing only exports that were produced within the country that exported them. Most original source material list transit trade, re-exports and bullion and specie separately at the commodity level. In two major cases—the Netherlands and Belgium—so-called 'disguised transit' is a well-known problem whereby transit trade was incorrectly listed as special trade. We relied on existing studies of those countries' exports to correct for this problem (Horlings, 2002; van Zanden & Lindblad, 1989). Additionally, we use international prices (Federico & Tena-Junguito, 2019) to correct for inaccurate 'administrative' pricing in several countries, especially the Netherlands, Belgium and Spain.

Country borders are not fixed over time. The goal of this project is to help explain patterns of structural change and diversification over time, and to this end, we wished as far as possible to record exports from geographical units that most closely approximate current nation states. To this end, we have used subnational sources for some countries or colonies where these were available to help. For example, in reconstructing the trade of countries in the former Ottoman Empire, we drew on port-level records from British and French consuls to estimate exports from areas that are today independent countries, like Turkey, Lebanon and Israel. Similarly, we deducted all exports at the commodity level from British Burma, for which independent statistics exist, from the aggregate figures for British India. Inevitably, this procedure was not always

possible, and we have sometimes been forced to assign all exports from a larger geographical unit, like French Indochina, to its most economically important component country (in this case Vietnam).

In total, the dataset reports detailed exports of up to 545 separate commodities per country-year between 1987-1906, for 100 polities. This represents 2194 unique product descriptions in our primary sources in their original languages. For example, unprocessed rubber, 'raw rubber', 'caoutchouc' and 'Rohkautschuk' are 4 different descriptions but all translate to the same SITC code because they represent the same commodity. We have manually translated the historical commodity classifications into the contemporary SITC (rev.4) system at the four-digit level. Each product description was investigated and matched to the closest product description in the full SITC nomenclature by a coder. In case of doubt, the full team of three coders considered the description and assigned the most appropriate SITC code. A dictionary to translate from original product descriptions to SITC codes was thus created. After assembling the dictionary, we then performed several manual checks, looking at items that had been assigned the same SITC code to ensure all of the product descriptions were appropriate. This translation exercise ensures that our data is directly comparable with contemporary trade statistics such as UN COMTRADE, which are also recorded and presented using the SITC nomenclature.

Throughout the paper we use data at the SITC-3 level. SITC-3 is also used as disaggregation level for diversification analysis in other recent studies (e.g. Lall, Weiss, & Zhang, 2005; Zhu & Fu, 2013) as well as in Michaely's (1984) pioneering work. Our data is most complete and reliable at the 3-digit level for two reasons. First, the product descriptions in the primary sources were in some cases too vague to attribute the more specific four-digit category. Second, the SITC classification system is set up to record some kinds of detail and not others: for example, in classifying wood and timber, the SITC system pays close attention to the treatment and transformation of wood, without caring much about the *kind* of wood (i.e., from what kind of trees). In primary sources from the early twentieth century, by contrast, we often have reasonably good information about the kind of wood but only at times about its treatment. So, going down to the SITC-4 level would introduce a bias where it would seem that a country is exporting more distinct commodities simply because its trade statistics include the treatment of wood instead of its kind. For the more general 3-digit classification this problem is not generally an issue in our data.

C Measures of Productive Capabilities

Exports can reveal the productive capabilities required to gain competitiveness in certain sectors (Sutton & Trefler, 2016). On the most basic level productive capabilities include technology and inputs. For example, a country without furnaces or iron will not be able to export steel. But productive capabilities required to gain competitiveness also involve more intangible inputs such as knowledge, management techniques, a disciplined labor force, a conducive institutional environment, access to finance and much more (Lall, 1992). For export data to capture these underlying conditions of competitiveness requires a certain degree of market integration. After all, we can only approximate countries' productive capabilities through their exports if they participate in international trade. In this paper we therefore compare two periods of deep market integration during two eras of globalization.

In order to use exports as a reliable indicator for countries' underlying production capabilities we want to filter out accidental exports. A country might be exporting a certain good in small quantities due to a glut on its market which might lead to the good being sold below its cost. We would not want to infer from such an export that a country is competitive on the world market in producing this good. We therefore follow the literature that uses export data to proxy for productive capabilities (Hausmann & Hidalgo, 2011; Sutton & Trefler, 2016) and adopt a revealed comparative advantage (RCA) as defined by Balassa (1965) of 0.5 to prevent our results from being driven by very small export lines while at the same time avoiding being too restrictive. That is, we calculate each commodity's share in total exports of a country and the world, and only include those export lines for which the share in a country's export is at least half of that in world trade. Due to missing data in some years for some countries in both periods, we take the arithmetic average of annual export flows of those years for which data is available, rather than the sum over all years. We then base all our measures on the average annual flows of each period.[11]

The most straightforward proxy for productive capabilities based on non-bilateral export data is diversity, a simple count of the number of distinct commodities above an RCA threshold of 0.5 (Hausmann et al., 2014, pp. 24–25; Hidalgo & Hausmann, 2009).[12] The idea is that the more different kinds of commodities a country is exporting the more productive capabilities it has. Export diversification has long been used as an indicator for levels of industrialization in the development literature (Felipe, Kumar, Abdon, & Bacate, 2012). Absolute levels of diversification have increased markedly across the two globalizations. In the last era of globalization, the United Kingdom was by far the most diversified country with 71 distinct export lines at the SITC 3 level for the years 1897-1906, whereas the mean diversification level was only 17 (Table 2). In contrast, the most diversified countries in the present era of globalization, France and Spain, both had 201 distinct export lines in the years 1998-2007 while the mean of all countries of 84 surpassed the UK's diversification in the last era of globalization. This reflects that the novelty of the present era of globalization compared to the previous one lies in the quantitative extent to which manufacturing processes have been broken up and distributed across borders (Findlay & O'Rourke, 2007, p. 511).

Since we are here interested in countries' productive capabilities in relation to all others in the cross section, we normalize all diversification levels by dividing through the maximum exports lines of the most diversified country in the respective period (Great Britain and France). This rescaled diversification measure allows us to examine the extent to which countries have changed their rank on the development ladder across the two globalizations.[13]

The maps (Figure 1a and 1b) show countries' diversification ranks during our two periods (1897-1906 and 1998-2007). Blank areas are missing from our data: these are mainly landlocked countries. The maps suggest that East and South Asia mostly ranked in the upper third across both periods with China and Thailand as examples of catching up. Central Asia, the Middle East and Africa have largely fallen behind starting mostly from medium ranks, while Latin American

[11] This treatment also applies to all control variables.
[12] See appendix for a formal derivation of the diversification measurement.
[13] See appendix for a formal definition of normalized diversification.

countries have become more polarized. Europe and the United States have sustained top diversification ranks across both periods.

High levels of diversification for Europe and North America contrast with the conventional wisdom that the late 19th century was a period of Great Specialization with the European countries and settler colonies specializing in manufacturing and the rest of the world specializing in agricultural goods and raw materials (Findlay & O'Rourke, 2007, p. 411). Our data does confirm the previous observation that colonies and independent countries were mostly limited to exporting such low value-added goods. The manufacturing share in exports was very low for most countries (mean 0.09) and reached a maximum of 0.71 per cent for Great Britain (Table 2). However, instead of specializing, exporters of manufacturing goods were diversified and competitively exported also primary products. Some prominent examples in this regard are Great Britain, Germany, Belgium and France. The fact that the more diversified a country the higher its manufacturing share is shown by a strong correlation (0.74) between the two measures for the 1897-1906 period which is considerably higher than the 0.44 we find for the contemporary period (Table 3).

We calculate the Gini coefficient over countries' export lines for the peak of the last and present eras of globalization (Figure 2). The two Ginis are surprisingly close with 0.40 for 1897-1906 and 0.36 for 1998-2007. The relative inequality in diversification in the present era of globalization is of a similar magnitude as that in the period that has previously been considered one of Great Specialization. Lorenz curves plotted for both periods in Figure 2 confirm this. While the Gini has fallen slightly, the lack of Lorenz dominance (the two curves intersect) leaves the evaluation of the change in relative inequality across the distribution ambiguous. To the extent that catching up has taken place it is skewed to the upper half of the distribution. The rest of the paper explores the persistence of countries' relative productive capabilities. For this, it is relevant that the overall distribution of our basic measure, diversification, is roughly stable across the two periods.

Beyond the simple count of distinct commodities competitively exported by a country that characterizes the diversification measure, the quality of exports can be assessed to approximate the underlying productive capacities employing complexity measures. To go beyond diversification, we can consider the ubiquity of export products. Hausman and Hidalgo (2011) define ubiquity as the number of countries competitively exporting a certain product (RCA$\geq$0.5 at SITC-3), where the larger this number of countries, the more ubiquitous a good.[14] The idea is that more ubiquitous commodities are those that require less capabilities and can therefore be competitively produced by a large number of countries. In order to approximate the underlying productive capabilities of a country, we therefore want to take both diversity and ubiquity into account. After all, diversity could be a misleading measure if one country exports many ubiquitous goods that require little productive capabilities where another country exports a small number of capability intensive non-ubiquitous goods.

Hausmann and Hidalgo (2011) find a robust relationship between ubiquity and diversification: less diversified countries tend to export more ubiquitous goods, whereas the most diversified

[14] See appendix for a formal definition.

countries tend to export both ubiquitous and non-ubiquitous goods that are out of reach for the countries further down in the diversification hierarchy. This relationship appears as a triangular adjacency matrix between countries and products for recent years (Hidalgo & Hausmann, 2009). This same triangular structure also appears in our data for the previous globalization (Figure 3a) and is confirmed for the present period (Figure 3b). Then and now less diversified countries are constrained to a small range of goods exported by many competitors.

Economic complexity is a measure that combines ubiquity and diversity in an iterative process weighing countries' diversity with the ubiquity of its products and the ubiquity of products in turn by the diversity of the exporting countries (see appendix for a formal definition using matrix notation) (Hidalgo & Hausmann, 2009). This measure gives a scalar index of how similar countries' export baskets are to each other (Mealy, Farmer, & Teytelboym, 2019). The economic complexity ranking is an approximation of countries' relative levels of productive capabilities at any point in time that considers not only the number of goods exported but also what and how many other countries can export these goods. As such complexity contains more information than diversity. Figures 4a-b map countries economic complexity rankings by decile for 1897-1906 and 1998-2007. The overall picture is similar to that for the maps of the diversification ranking. Europe and the United States are on top of the complexity ranking in the historical period and continue to lead in 1998-2007 with the United States, however, moving down to the second highest decile. Africa, the Middle East and Eastern Europe (including Russia) have fallen behind also in terms of economic complexity, while Latin America has become more polarized staying overall in medium ranks. Asia has caught up more in terms of economic complexity than diversification, reaching medium to high ranks.

For a set of exceptional cases, diversity performs better as an indicator for productive capabilities. This is the case for countries with low levels of diversification that export some non-ubiquitous goods unique to a few countries due to a scarce natural endowment. Consider for example Guyana, which was reasonably undiversified, but exported diamonds, a specialization shared by only two other countries in the historical period. Combining ubiquity and diversification in a measure of economic complexity, ranks the productive capabilities of such a country more highly than when considering only diversification. This is a similar "statistical illusion" as that found for more recent data in relation to developing countries (S. Lall, 2000; Lectard & Rougier, 2018).[15] A higher complexity rank in cases like this may be just a reflection of the uniqueness of a country's soil, climate or raw materials and not of its skill level or degree of industrialization—though it may be a good omen for the terms of trade, since a country that monopolizes the entire supply of a particular good will have market power. For this reason, we use both diversification and complexity to approximate productive capabilities.

An alternative proxy of productive capabilities is export sophistication (Lectard & Rougier, 2018). Not only countries but also goods can be ranked according to their complexity using an analogous method to the one just described. Sophistication ranks countries by weighing the product complexity of a country's export products by their share in the export basket.[16] In

[15] There the issue is that when countries with low levels of industrialization contribute labor-intensive steps in the global value chain for highly complex products the export basket can make it appear as if the industrial structure was more advanced than it actually is.
[16] See appendix for a formal definition of Sophistication.

addition to using an RCA threshold, this measure is sensitive to the relative importance of goods in a country's exports in ways that economic complexity is not. This can make a marked difference. For example, both China and the United States reach the top decile according to sophistication. But the two countries rank in the second highest decile for 1998-2007 in terms of economic complexity. For robustness we therefore consider both economic complexity and sophistication in our analysis.

All three proxies for productive capabilities, diversity, economic complexity and sophistication are comparable across periods thanks to our translation of the historical data into the SITC system. Table 2 reports summary statistics for both periods. In Table 3 we analyze correlation patterns across our key variables and find that they are all positively and highly significantly correlated. This indicates that all three measures proxy a common underlying feature, productive capabilities.

III. MAIN RESULTS

In this section we present our empirical strategy and main results. We demonstrate that the values of all our proxies of productive capabilities for the previous era of globalization are highly statistically significant and quantitatively important in relation to their respective contemporary values. Our measures of productive capabilities in 1896-1907 are also a highly significant predictor of real GDP per capita in the current period. Unlike most persistence studies (Bisin & Moro, 2021), we can measure the same variables for both the historical reference period (1897-1906) and the contemporary one (1998-2007). This is possible since all proxies for productive capabilities (diversification, economic complexity and export sophistication) rely solely on export data which we have available for both periods thanks to the construction of our new database (see section II). We can therefore use standard regression techniques to test the persistence in productive capabilities over time. We verify our results by varying the SITC digit levels, subperiods (Tables 5 & 6) and RCA cutoff points of 0.1 and 1 .[17] We also employ principal component analysis after adding the manufacturing share as an additional measure. We rule out re-emergence for both the persistence in productive capabilities as well as the relation between historical productive capabilities and contemporary income levels by running our regression for intermediate periods.

A Productive Capabilities across a Century

Prior to running regressions, Figures 5a-c illustrate our main finding: all three measures of productive capabilities point to strong persistence across the two eras of globalization. Each figure plots one of the three key variables (normalized diversification, economic complexity, and sophistication) for the previous globalization (1897-1906) on the horizontal and the present globalization (1998-2007) on the vertical axis. For each variable we observe a strong and positive correlation: the higher a country was up on the ladder of productive capabilities a century ago, the higher up it tends to be in the present globalization. The strong positive

[17] Varying SITC digit levels and RCA is standard practice in the diversification and economic complexity literature. Results available from the authors upon request.

correlations are robust to varying levels of minimum RCAs and for different subperiods in the contemporary data.[18]

The correlation between past and present diversification levels is strongest out of the three measures (0.73). Figure 5a also suggests more than third of the countries in our data, mostly former European colonies in Africa, are stuck across the two periods with less than 25% of the maximum diversification. Some countries have nevertheless managed to catch up and the mean of the normalized diversification has increased across the two periods from 0.24 to 0.42 (Table 2). However, as the Lorenz curve in Figure 2 illustrates, most of this catching up in diversification has taken place near the top of the hierarchy, a phenomenon described for GDP by Pritchett (1997). In 1897-1906 the UK was far ahead of all other countries. Its closest neighbors, France and Germany only reached 70 and 72 percent of the British diversification level. In 1998-2007, eight countries – all European or North American – scored above 80 per cent of the maximum diversification level.

For economic complexity (Figure 5b), the best linear fit has a slope of lower than 1 (0.70, Table 5, Column 3) which suggests that despite the general persistence some degree of catching up has occurred in terms of economic complexity. Notably, Japan and South-Korea have overtaken the European countries and the United States that continue to dominate in terms of diversification, while China and Mexico, for example, have caught up to the ranks of countries that dominated the top a century ago. For economic complexity, too, most African countries have either been stuck at the bottom of the ranking or have fallen even further behind compared to the historical period (1897-1907). For Latin America the picture is mixed with Chile as the country that has fallen furthest behind compared to where it started at the turn of the 20th century and Brazil as a notable exponent of catching up to upper-middle ranks.

Figure 5c illustrates the correlation (0.61) between countries' sophistication in 1897-1906 and 1998-2007, and points to a slightly greater potential for catch up compared to the very high persistence in economic complexity and diversification. This shows that to the extent that countries have managed to climb up the productive capabilities ladder, they have done so predominantly by increasing their exports in more sophisticated (or less ubiquitous) goods rather than relying primarily on exporting more different goods. Korea, the Philippines and China stand out as countries that have moved from upper middle to top ranks.

In section II.C we show that the persistence in our three measures of productive capabilities is statistically significant.

B Empirical Strategy

We broadly follow the approach in the pure persistence study by Voigtländer and Voth (2012). To estimate the persistence in productive capabilities, we use the following econometric model:

$$\boldsymbol{Y}_i^{1998-2007} = \alpha + \beta \boldsymbol{Y}_i^{1897-1906} + \gamma \boldsymbol{X}_i^{1998-2007} + \varepsilon_i \quad (1)$$

[18] We find strong correlations for minimum RCAs of 0.1 and 1.0 using the same periods and when combining 1897-1906 with the five year periods 1992-1996, 1997-2001, 2002-2007 and 2008-2012 as well as 1962-1971 and 1979-1989 applying a minimum RCA of 0.5.

where $\boldsymbol{Y}_i^{1998-2007}$ denotes one of the three proxies for productive capabilities in the current globalization at the country level and $\boldsymbol{Y}_i^{1897-1906}$ its counterpart for the previous globalization. $\boldsymbol{X}_i^{1998-2007}$ is a vector of control variables. These have been chosen from standard controls in the export diversification literature to test whether our historical measure is significant when accounting for the variables commonly considered as drivers of export patterns.[19] To account for the degree of economic liberalization we include trade openness as the trade/GDP ratio and FDI, measured as the ratio to GDP, for financial openness.[20] In light of the great importance attributed to human capital for growth (Barro, 2001), we control for the mean years of schooling. Following the large literature on the resource curse which finds that raw material abundance undermines the prospects for economic diversification (Frankel, 2010), we control for resource rents measured as the share of total natural resource rents in GDP. We control for the quality of institutions as possible alternative explanation of productive capabilities using POLITY V data. See Table 4 for an overview of all data sources.

C Persistence in Productive Capabilities

Table 5a reports the ordinary least square (OLS) regression based on the specification in III.B. Columns (1), (3), and (5) report the bivariate regressions of our main variables diversification, economic complexity and sophistication in 1998-2007 on the respective variables in 1897-1906. The association between the historical variables and their contemporary counterparts are all positive and significant at the 1 per cent significance level. This significance carries over after adding standard control variables in the trade diversification and growth literature in columns (2), (4), and (6). Resource rent and population are also highly significant in relation to all three key variables; human capital is only significant in relation to diversification and complexity not sophistication while the reverse is the case for trade openness. FDI is significant at 5 per cent in relation to diversification, and polity is significant at 10 per cent, in relation to diversification and sophistication.

Based on all three proxies of productive capabilities, persistence across the two waves of globalization is quantitatively vastly more important than the whole range of standard variables combined. The historical values of the three key variables can account by themselves for a major share of the variance in their contemporary counterparts: 48 per cent for diversification, 49 per cent for complexity and 39 per cent for sophistication. All other controls add 27, 23, and 34 percentage points respectively to the R-squared statistic.

The positive, quantitatively important and highly significant association between historical and contemporary diversification, complexity and sophistication subject to the control variables in Table 5, is robust when using SITC-2 and SITC-4 instead of SITC-3; when applying a low RCA

[19] For studies on trade diversification that each employs a subset of these controls see for example: Agosin, 2007; Agosin, Alvarez, & Bravo-Ortega, 2012; Cadot, Carrère, & Strauss-Kahn, 2013; Imbs & Wacziarg, 2003; Lectard & Rougier, 2018; Parteka & Tamberi, 2013. We have only added variables that measure levels rather than a rate of change since we are not interested in the changes within our contemporary reference period but the persistence across the two periods.

[20] The question whether economic liberalization aids or hinders diversification is a contested policy issue and studies reach inconsistent conclusions (Carrere, Cadot, & Strauss-Kahn, 2011; Chang & Andreoni, 2020; Dennis & Shepherd, 2011; Parteka & Tamberi, 2013; Rodrik, 2016).

of 0.1 and a high RCA of 1.0; as well as when substituting the contemporary reference period (1998-2007) with alternative subperiods (1962-1971, 1979-1988, see Table 5b).[21]

D Principal Component Analysis and Re-Emergence

To ensure the validity of our main regression results, we verify that our measures are approximating underlying productive capabilities and are not driven by the structure of the Standard International Trade Classification system. To this end we have varied the SITC digit levels as described. In addition, we add the manufacturing share in countries' export baskets as a more traditional measure of industrialization levels to our regression and perform a principal component analysis (PCA). Compared with the other three measures the manufacturing share is less sensitive to the structure of our data.

The contemporary manufacturing share is also positively and significantly correlated with its historical counterpart when controlling for covariates (Table 5a, column 8). 8 per cent of the variance are explained by the historical value of the manufacturing share. The small R-squared compared to our three main variables reflects that these other measures produce a full ranking of all countries while 30 countries did not export any manufacturing goods competitively (RCA≥ 0.5) in the historical period, ranking them all equally low. It is still remarkable that after decades of deindustrialization in Europe and North America and despite the globalization of industrialization and production chains in the current period, the manufacturing share a century earlier has a significant bearing on current outcomes.

We calculate the first principal component of the normalized scores of all four measures. We use the first principal component for 1998-2007 as dependent and the one for 1897-1906 as independent variable in columns (9) and (10) of Table 5a. The PCA confirms our previous result: productive capabilities are persistent and past ranks significantly impact those we observe today. In the bivariate regression, the historical principal component explains 49 per cent of the variation in the contemporary counterpart. Adding the control variables increases the R-squared by 27 percentage points. The first principal component is significant at the 1 per cent level (see Table 5a, columns (9) and (10).

The current era of globalization constitutes a break from the more protectionist and state-led policy regimes of the post-war period. We therefore also test whether the persistence in productive capabilities is not in fact a re-emergence (Voth, 2021). To this end, in Tables 5b and 7, we present the same regression as for 1998-2007 using our proxies for productive capabilities for 1962-1971 and 1979-1988 as dependent variables instead. We find that the significance across all five measures (diversification, complexity, sophistication, manufacturing share and first principal component) is consistently positive and highly significant at 1 per cent except for manufacturing in 1979-1988. The variance explained by our historical measures is even higher for 1962-1971 and 1979-1988 compared to 1998-2007. Across all five measures the variance accounted for by the historical correlates decreases slightly across the three periods. The coefficients for all our main explanatory variables are slightly lower for the 1998-2007 period

[21] Robustness results available from the authors on request.

compared with the previous periods. This confirms that productive capabilities are persistent rather than re-emergent.

Our results so far show that the higher a country was up on the productive capabilities ladder at the turn of the 20th century, the higher its position in today's era of globalization as well as in the intermediate post-war periods. Yet, this does not imply that there is no scope for path defiance. After all, about half of the variance remains unexplained by the historical position. This raises the question of the mechanism of persistence and the potential for defiance. In section IV. we examine the possibility of increasing returns to capabilities as possible explanation for persistence, explore possible omitted variables that could be an alternative reason for persistence (e.g. geography) and consider factors that enable path defiance.

E Persistent Productive Capabilities and the Wealth of Nations

An old insight and robust stylized fact in development economics is that the ways in which countries are integrated into global trade matters for their prosperity: higher levels of export diversification and complexity, are closely linked to higher levels of growth and GDP per capita.[22] Our new dataset allows us to test this relation for the first time for the long run of a century. Figure 6 plots our first principal component of the measures of productive capabilities for 1897-1906 on the horizontal axis and GDP per capita measured in purchasing power parity (PPP). We find a strong Pearson correlation coefficient of 0.69 suggesting that countries' position in the global division of labor a century ago is a strong predictor of today's per capita income. Some salient outliers that performed worse than their productive capabilities under the previous globalization indicated are Congo-Kinshasa, Ethiopia and Mozambique. They are part of a group of sub-Saharan African countries that were at the bottom of the rankings in terms of the productive capability measures under colonial rule in 1897-1906 and have fallen behind even further in terms of their income today. The most salient upward outliers are the Bahamas, Bermuda and the Seychelles – all known to be tax havens and Grenada, a tax friendly jurisdiction (see e.g. Alstadsæter, Johannesen, & Zucman, 2018). India and China have performed worse in per capita terms than their export structures in 1897-1906 would suggest.

In Table 8, we test whether this relation between productive capabilities then and per capita GDP now is statistically significant. Average GDP per capita in PPP for the years 1998-2007 is our dependent variable. The three proxies of productive capabilities (diversification, complexity and sophistication), the manufacturing share and the first principal component all measured for 1897-1906 are our explanatory variables. We use the same controls as before (trade openness, resource rent, FDI, population log, human capital and polity). We find that all our measures of productive capabilities in the previous era of globalization are highly significant at the 1 per cent significance level in relation to per capita GDP in the present period. This finding is robust to varying SITC digit levels and RCA cutoffs as well as to measuring GDP at market exchange rates instead of PPP. This suggests that the more diversified, complex and sophisticated a country's exports at the turn of the 20th century, the higher the wealth of the nation today.

[22] See for example Hausmann et al., 2007; Hausmann & Hidalgo, 2011; Hesse, 2009; Hirschman, 1958; Hidalgo, Winger, Barabási, & Hausmann, 2007; Lederman & Maloney, 2003; Prebisch, 1950; Rosenstein-Rodan, 1943; Zhu & Li, 2017.

The productive capabilities in the previous era of globalization are not only highly significant in relation to today's per capita income but are quantitatively very large. 56 per cent of the variance in today's per capita GDP can be explained by countries' economic complexity or the first principal component in 1897-1906. This is followed by 50 per cent for sophistication, 31 per cent for diversification and 27 per cent for the manufacturing share. After including the historical proxies for productive capabilities in our regression out of all control variables only human capital retains significance in relation to GDP per capita. Institutional quality and the economic openness measures – two variables of central importance in the recent growth literature and development policy making (Baldwin, 2004) – are consistently insignificant in relation to both GDP measures in our regressions.

We test whether the relation between historical productive capabilities and contemporary GDP per capita is a case of persistence or re-emergence by using average GDP per capita for the periods 1962-1971 and 1979-1988 as dependent variables in Table 6b. We find that our historical proxies of productive capabilities are also highly significant and quantitatively important for the intermediate periods. The only exception is sophistication for the 1962-1971 period, which could be due to the uneven availability of GDP data at the time. The coefficients are of comparable magnitudes across our three periods. Overall, our results point to the relationship between productive capabilities and income being persistent rather than re-emergent.

IV. PERSISTENCE MECHANISMS

Our main results show a high degree of persistence in productive capabilities across the past century. In this section, we present a simple framework to analyze the persistence mechanism. We test whether persistence is a result of geography and institutions and show that the persistence in productive capabilities is largely driven by complementarities between past and future capabilities. The only factor that significantly impacts productive capabilities in both periods independent from persistence itself is a history of colonization. We explore shocks that undermine persistence and confirm the resource curse hypothesis.

A A Framework for the Evolution of Productive Capabilities

With our main results, we have shown that productive capabilities are broadly persistent across the past century. We use the model for technology adoption dynamics in Comin, Easterly, and Gong (2010) to help us think through the mechanism underpinning persistence in productive capabilities and apply it in our empirical analysis. The idea is that the mechanisms for persistence in technology also apply to the full range of productive capabilities beyond production techniques more narrowly.

In their review of the economic history literature, Comin, Easterly, and Gong (2010, pp. 67–68) identify the following key drivers for persistence in technology: complementarities between existing technologies; recombination of old technologies into new ones; feedback and spillover of technology to science and vice versa as well as across different sectors; economies of scale and scope; and learning by doing. They measure technology directly as the adoption of certain production techniques. Our measures of productive capabilities instead measure outcomes in terms of export competitiveness that reflect underlying capacities. The basic intuition is that the

drivers that perpetuate technology advantages also apply to productive capabilities more broadly. As a general tendency, the more know-how, production experience, management skills, cutting-edge infrastructure, science, world-leading firms etc. a country already has, the more it can acquire. In other words, there are positive returns to productive capabilities.[23]

We can analyze the persistence mechanism in terms of this simple model:

$$\boldsymbol{Y}_{it} - \boldsymbol{Y}_{it-1} = \boldsymbol{Y}_{it-1}^{\gamma} e^{\theta it} + \varepsilon_{it} \quad (2)$$

Where generations are indexed by t and countries by i. $\boldsymbol{Y}_{it}$ denotes the level of productive capabilities up to time t. The term $\boldsymbol{Y}_{it-1}^{\gamma}$ reflects the complementarity between past and future productive capabilities. The strength of this complementarity is approximated by the parameter γ. θ_{it} is a parameter that measures factors other than this complementarity that also affect the enhancement of productive capabilities. This could for example include the institutional environment or natural endowments. ε_{it} is the error term. See Comin, Easterly, and Gong (2010) for a more detailed discussion of the properties of this model.

Depending on the value of γ this model indicates convergence, divergence or persistence. Assuming θ_{it} to be constant across time and countries for now, a γ bigger than one implies divergence. Countries with more productive capabilities would pull ahead of countries with less productive capabilities at an accelerating pace. This would quickly lead to an explosive form of bursting apart. If γ is equal to 1, productive capabilities are growing at a stationary rate and countries relative positions in the global economy are persistent. Finally, if γ is smaller than 1, there will be convergence in the long run. The productive capabilities of more advanced countries are growing slower than those of less advanced countries. But this process of convergence may be very slow. So, even after long periods initial levels of productive capabilities may still matter. Independent of the value of γ, factors other than initial productive capabilities are of course relevant: Empirically, we do not expect θ_{it} to be invariable.

B Persistence in Productive Capabilities versus Persistence in Correlates

The model presented in (2) is consistent with two alternative persistence mechanisms: First, complementarity between past and future productive capabilities could drive persistence – the complementarity view (Comin, Easterly, and Gong, 2010, 90). Second, some factor other than the productive capabilities themselves could be persistent and drive the relative stability in the country rankings across long periods of time – the correlates view. For the correlates view, γ would be close to zero and the persistent difference in productive capabilities would instead be solely explained by a persistent variance across countries in θ_{it}. Note that the complementarity view is consistent with persistence in different values for θ_{it}. In other words, the complementarity view does not suggest a mono-causal story. The key for the complementarity view is that γ is not equal to zero. To rule this out, we first identify the correlates of our historical proxies of productive capabilities and then test whether the significance of the

[23] This interpretation of productive capabilities is broadly in line with Weitzman (1998) and other studies using combinatorial arguments related to innovation and growth.

historical levels of productive capabilities in relation to their contemporary counterparts still holds when controlling for these correlates.

In light of the debate over whether institutions or geography drive growth in the long run (Rodrik et al., 2004; Sachs, 2003), variables that capture either of the two dimensions are important candidates to substantiate the first correlates view. The idea that climate, geography and ecology are key to explaining differences in economic performance has been recurring e.g. in the writings of Montesquieu (1899) or Marshall (1890). Diamond (1999) and Sachs (Sachs, 2001) have revived the hypothesis that geography drives growth and a number of recent persistence studies also highlight the importance of geographic characteristics (Spolaore & Wacziarg, 2013). Institutions have been an equally prominent explanation of differences in long run economic performance at least since the seminal contribution of (Acemoglu et al., 2021; Acemoglu, Johnson, & Robinson, 2001). The origins of today's institutions have been traced to colonial times in this literature and institutional quality has been found to be persistent. It might therefore be the case that the persistence in productive capabilities we observe is only an expression in the persistence of institutions.

A large and growing empirical literature confirms long-term effects of colonization (e.g. Dell, 2010; Nunn & Wantchekon, 2011) on growth performances. Reflecting that European overseas colonialism was important in shaping the global division of labor in the previous era of globalization, we use Hadenius and Berg-Schlosser's (2007) colonial heritage dummy that records if a country was *ever* a European overseas colony and combine this with a dummy for European colonial powers. Missing observations have been added using data from the ICOW Colonial History Data Set (Hensel, 2018). For robustness, we also construct an additional dummy of colonial status which measures whether a country was a colony or a colonizer in the year 1900 from the ICOW as well.

We have tested the significance of a range of geographic and institutional variables in relation to our productive capability proxies for 1897-1906 using OLS regressions. We have also included variables that measure other fundamentals such as the population size and used data from Pascali (2017) that measures shipping times in our historical period after the invention of the steamship to account for ease of market access. Population data was taken from the History Database of the Global Environment (HYDE) 3.2 dataset (Klein Goldewijk, Beusen, Doelman, & Stehfest, 2017). The shipping time variable gives the hypothetical time taken by a steamship to travel between two ports (using, if necessary, the Suez Canal). Country-level data was constructed by averaging the shipping times for each potential destination port by their share in global exports, taken from our dataset. The resulting variable can be interpreted as a measure of a given country's physical centrality in the network of global commodity trade. We also construct two variables to proxy for land and labor productivity in subsistence agriculture. We combine cell-level FAO-GAEZ agricultural suitability indices for sixteen major food crops with HYDE 3.2's estimates of cropland and rural population for all countries in our dataset. The labor productivity variable measures the maximum caloric output of all cropland in a given country in 1900 divided by the rural population; the land productivity variable uses the total extent of cropland as the denominator instead.

All variables that we have tested but found insignificant are reported in the notes to Table 7. Notably, the polity index for the year 1900 – a measure of institutional quality – turned out to not significantly affect the diversification, complexity or sophistication of countries' exports. Polity measured in 1998-2007 is also largely insignificant in our main regressions (Table 5 a-b). This suggests that persistence in productive capabilities is not simply an expression of persistence in institutions. Ease of market access measured in terms of shipping times, the most important means of transportation in that period, is also not significant in relation to any of our proxies. All variables that we have found to be significant in relation to our measures are reported in Table 7. See Table 4 for the data sources of all variables included in the regression. To prevent collinearity issues, we have grouped geographic variables and variables that capture dimensions of countries' economic structures and institutions separately.

Soil suitability, ruggedness, latitude, temperature, humidity and coal are the geographic variables that are significant in relation to at least some of our historical proxies. Population, land and labor productivity as well as countries' colonial status proved significant in the second group of variables. Being a colony was significantly negatively correlated with our measures of productive capabilities for both our measures of colonial status, whereas we find a significant positive relation for colonial powers. In sum, our findings suggest that productive capabilities in the previous globalization depended on a combination of geographic factors, population and labor productivity as well as a country's position in the global political system.

In Table 8 we test whether the correlates of our proxies of productive capabilities in the previous era of globalization also explain today's productive capabilities and erase the effect of the historical proxies ($\gamma \approx 0$). We find that the only historical correlate of productive capabilities that is consistently highly significant across all proxies for productive capabilities is the dummy that measures whether a country was an overseas colony. European overseas colonies suffer from lower productive capabilities in 1998-2007 compared to all other countries even when accounting for their low ranks in 1897-1906. This difference is significant at the 1 or 5 per cent level and robust when using the alternative dummy that measures whether a country was a colony in the year 1900.

All other correlates are only significant in relation to some but not other proxies. For diversification, labor productivity in the historical period still have some bearing today but this is not robust when using the alternative colonial status measure. A negative relation between temperature and economic complexity is significant only when we do not control for colonial status.

The coefficients of our historical proxies of productive capabilities estimate γ. They range between 0.39 and 0.75 and are 0.71 and 0.58 for the first principal component (Table 8, columns 9 and 10). The fact that they are not equal to 1 suggests that there is some degree of convergence. But our main result, that after a century, historical productive capabilities still matter, suggests that this convergence must be very slow. Taken together, this supports the complementarity view. Productive capabilities are persistent not because of some third factor but because more productive capabilities in the past help countries to gain more productive capabilities in the future. This is, however, subject to an important qualification: countries that were European

overseas colonies at some point in history have fallen further behind compared to their already low levels of productive capabilities in the previous globalization.

C Path-defiance: Which Countries Catch Up and Which Ones Fall Behind?

To further explore the mechanism that drives persistence of productive capabilities in the long run, let's consider cases of path-defiance. With this approach we follow Voigtländer and Voth (2012) who have pioneered the study of failed transmission. When have past levels of productive capabilities not determined countries' future performance? One possible answer to this question is that there have been shocks to factors that affect productive capabilities which changed countries' trajectories. That is θ_{it} might have been time variant for some countries but not others. As shown in the previous section, we find that most geographic variables as well as institutional quality beyond a legacy of colonization do not explain the persistence in productive capabilities. This does not rule out, however, that changes to institutions or the natural environment could prevent persistence.

We analyze two types of shocks in relation to path-defiance: democratization and the discovery of oil. Democratization has long been considered key in countries' development process. Based on recent contributions that report a positive relation between growth and democratization (Acemoglu, Naidu, Restrepo, & Robinson, 2019; Papaioannou & Siourounis, 2007), we expect that countries that democratized improved their performance more than other countries. An upgrade in institutional quality thanks to democratization should bring about an enhancement of productive capabilities. Samuel Huntington (1991) has famously distinguished three waves of democratization. His focus is on the most recent wave from 1974-1990 which has about doubled the number of democracies in the world and culminated with the fall of Communism. The first wave started with the 19th century expansion of suffrage and coincides roughly with the previous era of globalization. The second wave was unleashed by the victory of the allied forces in World War II and reversed in 1962.

In Table 9, we use the second and third waves of democratization to test whether an improvement in institutional quality has changed countries' productive capabilities relative to where they were in the previous globalization. Our dummy for the third wave of democratization is based on data of Papaioannou and Siourounis (Papaioannou & Siourounis, 2008).[24] For the second wave, we have constructed a democratization dummy using the Polity V index. We define countries that increased their polity score above the threshold for democracy starting from a persistently lower level as democratized and have cross-checked our classification with Huntington's own classification.

We find that second wave countries significantly improved their position in the global productive capabilities ranking in the current globalization compared to where they started at the turn of the 20th century (Table 9b). The only exception is the diversification measure. The interaction terms with our proxies for the previous era of globalization are not significant. Hence, this does take the form of a shock that undermined persistence: the countries of the second wave were lifted up in relation to all other countries but the ranking amongst them did not change significantly.

[24] See Papaioannou and Siourounis (2007) for a detailed data documentation.

However, the same result is less significant when we use 1962-1971 as reference decade which suggests that the full effect of democratization might unfold with a time lag. This is also consistent with our findings for the third wave. Countries that democratized in the most recent wave moved up in terms of economic complexity. But their sophistication and diversification ranks were not significantly altered (Table 9c). For diversification, the interaction term is significant and positive which suggests that the importance of the historical initial conditions is amplified in this dimension for the third wave countries.

A large literature suggests that countries that have an abundance in terms of natural resources are cursed when it comes to industrial development and economic diversification (Auty, 1998; Frankel, 2010; Humphreys, Sachs, & Stiglitz, 2007). The finding in Table 5a-c, that resource rents significantly lower countries' productive capabilities confirms this resource curse hypothesis. Discovery or new demand for some natural resources due to technological change are therefore possible factors that could change θ_{it} for some countries. In particular oil is a prominent candidate in the 20th century. In Table 9a, we test whether membership in the Organization of the Petroleum Exporting Countries (OPEC) affects the persistence in productive capabilities. We find that based on three proxies with the exception of diversification, OPEC members have significantly moved down in the productive capabilities ranking compared to where they started in the previous globalization. In contrast, the relative positions to one another have been largely unaffected as indicated by the insignificant interaction term. This confirms the resource curse hypothesis from a long run perspective.

The falling behind in terms of productive capabilities of the OPEC countries is also illustrated by the scatter plots in figures 5a-c: oil exporters are among the lowest ranking countries in terms of diversification, economic complexity and export sophistication and have moved down the ladder several steps compared to where they started in 1897-1906.

Another possible channel that could alternate persistence is development policy. After all, industrialization and export diversification have been long-standing goals in economic development. As a preliminary step, we have tested whether (1) land reforms, (2) membership in the World Trade Organization, or (3) adjustment programs by the World Bank or International Monetary Fund have affected the persistence in productive capabilities. The rational for this choice of variables is that (1) is a proxy for big push development programs as common in the post-war period, whereas (2) and (3) capture whether countries have undergone economic liberalization. None of these three variables have been significant in the same exercise as the once reported in Table 9. This does not suggest that development policy is futile in general, but rather that more research is required to examine how development programs have succeeded in upgrading productive capabilities. It might also turn out that there is no one big solution that fits all: countries' success in escaping the overall tendency of persistence might depend on strategies tailored to specific local conditions.

V. CONCLUSION

We document the persistence of productive capabilities across two periods of globalization at the turn of the 20th and 21st century. Our proxies of capabilities (diversification, complexity, and sophistication) are derived from export data with the same method across both periods. This is

possible thanks to the new global commodity export dataset we have constructed. We show that countries' position on the development ladder a century ago significantly determines their present position. This effect is quantitatively important. To illustrate, 49 per cent of the variance in countries' economic complexity today can be explained by their economic complexity in the previous globalization. Furthermore, historical productive capabilities have a statistically significant and large bearing on today's wealth of nations. 56 per cent in the variance in GDP per capita in the period 1998-2007 can be explained by economic complexity in 1897-1906.

We analyze the mechanism driving the persistence in productive capabilities and find that the most important factor is the complementarity between past and future productive capabilities. Countries that are already competitive in exporting a wide variety of goods, enjoy a complex economic structure and largely export sophisticated items, tend to perpetuate their productive capabilities. Conversely, countries starting from poor initial conditions, tend to be stuck with relatively low productive capabilities. This interpretation is consistent with predictions from the recombinant growth literature. Our results suggest that this is not simply driven by persistent geography or institutions. With one important exception: former European overseas colonies are performing significantly worse in terms of all proxies for productive capabilities. This confirms the great importance persistence studies have attributed to colonial origins of today's economic performance (e.g. Dell, 2010; Nunn & Wantchekon, 2011).

We explore two cases of defying persistence. First, we confirm the resource curse hypothesis (Frankel, 2010) for the long run: OPEC members have significantly lost out in terms of productive capabilities compared to where they started a century ago. Second, we build on the literature on democratization and growth (Acemoglu et al., 2019) and find that democratization has to some extent helped improve countries' productive capabilities. Weaker results for the quantitatively more important third wave compared to the second wave suggest that there might be a long time lag until the full effect has unfolded.

The persistence in productive capabilities documented in this paper is a possible candidate for the mechanism that underpins the non-convergence in countries' income levels (Johnson & Papageorgiou, 2020). What you exported matters and history has a long shadow on development. However, this should not lead to the conclusion that there is no scope for change. While still strong after a century, the effect of our historical proxies for productive capabilities does decrease over time even if the pace is slow. Yet, standard policy variables such as trade openness, FDI etc. have limited explanatory power once we account for the persistence in productive capabilities. Further research is needed that explores long-term trajectories of path-defying countries' to identify policy strategies that succeeded in undermining persistence and enabled catching up.

References

Acemoglu, D., Egorov, G., & Sonin, K. (2021). Institutional Change and Institutional Persistence. In *Handbook of Historical Economics*. Elsevier.

Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review*, *91*(5), 1369–1401. https://doi.org/10.1257/aer.91.5.1369

Acemoglu, D., Naidu, S., Restrepo, P., & Robinson, J. A. (2019). Democracy Does Cause Growth Suresh Naidu Pascual Restrepo. *Journal of Political Economy*, *127*(1), 47–100.

Agosin, M. R. (2007). Export Diversification and Growth in Emerging Economies. *Economics Department, University of Chile, Working Paper Series*, *233*.

Agosin, M. R., Alvarez, R., & Bravo-Ortega, C. (2012). Determinants of Export Diversification Around the World: 1962-2000. *World Economy*, *35*(3), 295–315. https://doi.org/10.1111/j.1467-9701.2011.01395.x

Alstadsæter, A., Johannesen, N., & Zucman, G. (2018). Who owns the wealth in tax havens? Macro evidence and implications for global inequality. *Journal of Public Economics*, *162*(September 2017), 89–100. https://doi.org/10.1016/j.jpubeco.2018.01.008

Auty, R. M. (1998). *Resource Abundance and Economic Development: Improving the Performance of Resource-Rich Countries*. Helsinki.

Bairoch, P. (1982). International industrialization levels from 1750 to 1980. *Journal of European Economic History*, *11*(2), 269–333.

Bairoch, P., & Burke, S. (1989). European trade policy, 1815–1914. In P. Mathias & S. Pollard (Eds.), *The Cambridge Economic History of Europe from the Decline of the Roman Empire: Volume 8: The Industrial Economies: The Development of Economic and Social Policies* (pp. 1–160). https://doi.org/DOI: 10.1017/CHOL9780521225045.002

Bairoch, P., & Kozul-Wright, R. (1996). Globalization Myths: Some Historical Reflections on Integration Industrialization and Growth in The World Economy. *UNCTAD Discussion Paper*, (113).

Balassa, B. (1965). Trade Liberalisation and "Revealed" Comparative Advantage. *The Manchester School*, *33*(2), 99–123.

Baldwin, R. E. (2004). Openness and Growth: What's the Empirical Relationship? In *Challenges to Globalization: Analyzing the Economics* (pp. 499–526). Retrieved from http://www.nber.org/chapters/c9548

Barro, R. J. (2001). Human Capital and Growth. *American Economic Review*, *91*(2), 12–17. https://doi.org/10.1257/aer.91.2.12

Bisin, A., & Moro, A. (2021). LATE for History. In *Handbook of Historical Economics*.

Bolt, J., & van Zanden, J. L. (2020). Maddison style estimates of the evolution of the world economy. A new 2020 update. *Maddison-Project WP*, *15*.

Cadot, O., Carrère, C., & Strauss-Kahn, V. (2013). Trade diversification, income, and growth: What do we know? *Journal of Economic Surveys*, *27*(4), 790–812. https://doi.org/10.1111/j.1467-6419.2011.00719.x

Carrere, C., Cadot, O., & Strauss-Kahn, V. (2011). Trade Diversification: Drivers and Impacts. In M. Jansen, R. Peters, & J.-M. Salazar-Xirinachs (Eds.), *Trade and Employment: from Myths to Facts* (pp. 253–307). ILO-EC International Labour Office - European Commission.

Chang, H. J., & Andreoni, A. (2020). Industrial Policy in the 21st Century. *Development and Change*, *51*(2), 324–351. https://doi.org/10.1111/dech.12570

Comin, D., Easterly, W., & Gong, E. (2010). Was the Wealth of Nations Determined in 1000 BC? *American Economic Journal: Macroeconomics*, *2*(3), 65–97. https://doi.org/10.1257/mac.2.3.65

Dahlman, C. J., Ross-Larson, B., & Westphal, L. E. (1987). Managing technological development: Lessons from the newly industrializing countries. *World Development*, *15*(6), 759–775. https://doi.org/https://doi.org/10.1016/0305-750X(87)90058-1

Dell, M. (2010). The Persistent Effects of Peru's Mining Mita. *Econometrica*, *78*(6), 1863–1903. https://doi.org/https://doi.org/10.3982/ECTA8121
Diamond, J. (1999). *Guns, Germs, and Steel: The Fate of Human Society*. New York ; London: W. W. Norton & Company.
Estevadeordal, A., Frantz, B., & Taylor, A. M. (2003). The Rise and Fall of World Trade, 1870-1939. *The Quarterly Journal of Economics*, *CXVIII*(2), 359–407. https://doi.org/10.1093/qje/qjs020.Advance
Fagerberg, J., & Srholec, M. (2008). National innovation systems, capabilities and economic development. *Research Policy*, *37*(9), 1417–1435. https://doi.org/https://doi.org/10.1016/j.respol.2008.06.003
Fajgelbaum, P. D., Goldberg, P. K., Kennedy, P. J., & Khandelwal, A. K. (2020). The Return to Protectionism. *The Quarterly Journal of Economics*, *135*(1), 1–55. https://doi.org/10.1093/qje/qjz036
Federico, G., & Persson, K. G. (2007). Market Integration and Convergence in the World Wheat Market, 1800-2000. In T. Hatton, K. H. O'Rourke, & A. M. Taylor (Eds.), *The New Comparative Economic History: Essays in Honor of Jeffrey G. Williamson* (pp. 87–114). Cambridge, MA: MIT Press.
Federico, G., & Tena-Junguito, A. (2019). World trade, 1800-1938: A new synthesis. *Revista de Historia Economica - Journal of Iberian and Latin American Economic History*, *37*(1), 9–41. https://doi.org/10.1017/S0212610918000216
Felipe, J., Kumar, U., Abdon, A., & Bacate, M. (2012). Product complexity and economic development. *Structural Change and Economic Dynamics*, *23*(1), 36–68. https://doi.org/10.1016/j.strueco.2011.08.003
Ferguson, N. (2002). *The Rise and Demise of the British World Order and the Lessons for Global Power*. London: Allen Lane.
Findlay, R., & O'Rourke, K. H. (2003a). Commodity Market Integration, 1500–2000. In M. D. Bordo, A. M. Taylor, & J. G. Williamson (Eds.), *Globalization in Historical Perspective* (pp. 13–65). Chicago: University of Chicago Press.
Findlay, R., & O'Rourke, K. H. (2003b). Commodity Market Integration, 1500 – 2000. In M. D. Bordo, A. M. Taylor, & J. G. Williamson (Eds.), *Globalization in Historical Perspective* (pp. 13–65). Chicago: University of Chicago Press.
Findlay, R., & O'Rourke, K. H. (2007). *Power and Plenty: Trade, War and the World Economy in the Second Millennium*. Princeton University Press.
Frankel, J. A. (2010). The Natural Resource Curse: A Survey. *NBER Working Papers*, *15836*.
Frankel, J. A., & Romer, D. (1999). Does Trade Cause Growth ? *American Economic Review*, *89*(3).
Frankema, E., & Jerven, M. (2014). Writing history backwards or sideways: towards a consensus on African population, 1850–2010. *The Economic History Review*, *67*(4), 907–931. Retrieved from http://www.jstor.org/stable/24027257
Frankema, E., Williamson, J., & Woltjer, P. (2018). An economic rationale for the west African scramble? The commercial transition and the commodity price boom of 1835-1885. *Journal of Economic History*, *78*(1), 231–267. https://doi.org/10.1017/S0022050718000128
Gallup, J. L., Sachs, J. D., & Mellinger, A. D. (1998). Geography and Economic Development. *NBER Working Papers*, *6849*.
Galor, O. (2005, January 1). Chapter 4 From Stagnation to Growth: Unified Growth Theory. *Handbook of Economic Growth*, Vol. 1, pp. 171–293. https://doi.org/10.1016/S1574-

0684(05)01004-X
Galor, O. (2011). *Unified Growth Theory*. https://doi.org/10.2307/j.ctvcm4h7m
Hadenius, A., & Berg-Schlosser, D. (2007). Success and Failuers of the New Democracies. In D. Berg-Schlosser (Ed.), *Democratization: The Sate of the Art* (2nd ed., pp. 97–110). Opladen and Farmington Hills: Barbara Budrich Publishers.
Hausmann, R., & Hidalgo, C. A. (2011). The network structure of economic output. *Journal of Economic Growth*, *16*(4), 309–342. https://doi.org/10.1007/s10887-011-9071-4
Hausmann, R., Hidalgo, C. A., Bustos, S., Coscia, M., Simoes, A., & Yildirim, M. A. (2014). *The Atlas of Economic Complexity: Mapping Paths to Prosperity*. Cambridge, MA: MIT Press.
Hausmann, R., Hwang, J., & Rodrik, D. (2007). What you export matters. *Journal of Economic Growth*, *12*(1), 1–25. https://doi.org/10.1007/s10887-006-9009-4
Headrick, D. R. (1981). *Tools of Empire: Technology and European Imperialism in the Nineteenth Century*. Oxford University Press.
Hensel, P. R. (2018). ICOW Colonial History Data Set, version 1.1. Retrieved April 10, 2020, from http://www.paulhensel.org/icowcol.html
Hidalgo, C. A., & Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(26), 10570–10575. https://doi.org/10.1073/pnas.0900943106
Hirschman, A. O. (1958). *The Strategy of Economic Development*. New Haven: Yale University Press.
Hobsbawm, E. (1989). *The Age of Empire 1875-1914*. New York: Random House.
Horlings, E. (2002). The International Trade of a Small and Open Economy. Revised Estimates of the Imports and Exports of Belgium, 1835-1990. *NEHA-Jaarboek*, *65*, 110.
Humphreys, M., Sachs, J. D., & Stiglitz, J. E. (Eds.). (2007). *Escaping the Resource Curse*. New York: Columbia University Press.
Huntington, S. P. (1991). *The Third Wave: Democratization in the late Twentieth Century*. Norman: University of Oklahoma Press.
Imbs, J., & Wacziarg, R. (2003). Stages of diversification. *American Economic Review*, *93*(1), 63–86. https://doi.org/10.1257/000282803321455160
Jacks, D. S., Meissner, C. M., & Novy, D. (2009). Trade Booms, Trade Busts, and Trade Costs. *NBER Working Papers*, *15267*.
Jerven, M. (2013). *Poor Numbers: How We Are Misled by African Development Statistics and What To Do About It*. Cornell University Press.
Johnson, P., & Papageorgiou, C. (2020). What Remains of Cross-Country Convergence? *Journal of Economic Literature*, *58*(1), 129–175. https://doi.org/10.1257/jel.20181207
Jones, M. T., & Obstfeld, M. (1997). Saving, Investment, and Gold: A Reassessment of Historical Current Account Data. *NBER Working Papers*, *6103*(July).
Klein Goldewijk, K., Beusen, A., Doelman, J., & Stehfest, E. (2017). Anthropogenic land use estimates for the Holocene--HYDE 3.2. *Earth System Science Data*, *9*(2), 927–953.
Lall, S. (2000). The technological structure and performance of developing country manufactured exports, 1985-98. *Oxford Development Studies*, *28*(3), 337–369. https://doi.org/10.1080/713688318
Lall, Sanjaya. (1992). Technological capabilities and industrialization. *World Development*, *20*(2), 165–186. https://doi.org/https://doi.org/10.1016/0305-750X(92)90097-F
Lectard, P., & Rougier, E. (2018). Can Developing Countries Gain from Defying Comparative

Advantage? Distance to Comparative Advantage, Export Diversification and Sophistication, and the Dynamics of Specialization. *World Development*, *102*, 90–110. https://doi.org/10.1016/j.worlddev.2017.09.012
Lederman, D., & Maloney, W. (2012). *Does What You Export Matter?: In Search of Empirical Guidance for Industrial Policies*. https://doi.org/10.1596/978-0-8213-8491-6
Lewis, A. (1969). *Aspects of tropical trade 1883-1965*. Stockholm: Almqvist and Wiksell.
Maddison, A. (1989). The World Economy in the Twentieth Century. *OECD*.
Marshall, A. (1890). *Principles of Economics*. London: Macmillan.
Mealy, P., Farmer, J. D., & Teytelboym, A. (2019). Interpreting economic complexity. *International Journal of Agricultural and Statistical Sciences*, *14*(2), 1–9. https://doi.org/10.1126/sciadv.aau1705
Meissner, C. M., & Tang, J. P. (2018). Upstart industrialization and exports: Evidence from Japan, 1880-1910. *Journal of Economic History*, *78*(4), 1068–1102. https://doi.org/10.1017/S0022050718000517
Michaely, M. (1984). *Trade, income levels, and dependence*. Amsterdam: North-Holland.
Mitchener, K. J., & Weidenmier, M. (2008). Trade and empire. *Economic Journal*, *118*(533), 1805–1834. https://doi.org/10.1111/j.1468-0297.2008.02192.x
Mitchie, J., & Kitson, M. (1995). Trade and Growth: A Historical Perspective. *Managing the Global Economy*.
Montesquieu, C. de S. (1899). *The Spirit of Laws*. New York: The Colonial Press.
Nayyar, D. (2006). Globalization, History and Development: A Tale of Two Centuries. *Cambridge Journal of Economics*, 137–159.
Nayyar, D. (2013a). *Catch Up: Developing Countries in the World Economy*. https://doi.org/10.1093/acprof:oso/9780199652983.003.0006
Nayyar, D. (2013b). *Catch Up: Developing Countries in the World Economy*. Oxford University Press.
Nunn, N. (2020). The historical roots of economic development. *Science*, *367*(6485). https://doi.org/10.1126/science.aaz9986
Nunn, N., & Wantchekon, L. (2011). The Slave Trade and the Origins of Mistrust in Africa. *American Economic Review*, *101*(7), 3221–3252. https://doi.org/10.1257/aer.101.7.3221
O'Rourke, K. H. (2018). Two Great Trade Collapses: The Interwar Period and Great Recession Compared. *IMF Economic Review*, *66*(3), 418–439. https://doi.org/10.1057/s41308-017-0046-0
O'Rourke, K. H., & Williamson, J. G. (2002). When did globalisation begin? *European Review of Economic History*, *6*(1), 23–50. https://doi.org/10.1017/s1361491602000023
O'Rourke, Kevin H., & Williamson, J. G. (1999). *Globalization and history : the evolution of a nineteenth-century Atlantic economy*. MIT Press.
Obstfeld, M., & Taylor, A. M. (2004). *Global Capital Markets: Integration, Crisis, and Growth*. Cambridge: Cambridge University Press.
Papaioannou, E., & Siourounis, G. (2007). Initial factors behind the third wave of democratization. *Retrieved on September*. Retrieved from http://isites.harvard.edu/fs/docs/icb.topic173303.files/data_democratization_oct07.pdf
Papaioannou, E., & Siourounis, G. (2008). Democratisation and growth. *Economic Journal*, *118*(532), 1520–1551. https://doi.org/10.1111/j.1468-0297.2008.02189.x
Parteka, A., & Tamberi, M. (2013). What determines export diversification in the development process? Empirical assessment. *World Economy*, *36*(6), 807–826.

https://doi.org/10.1111/twec.12064
Pascali, L. (2017). The wind of change: Maritime technology, trade, and economic development. *American Economic Review*, *107*(9), 2821–2854. https://doi.org/10.1257/aer.20140832
Piketty, T., & Saez, E. (2003). Income Inequality in the United States, 1913–1998. *The Quarterly Journal of Economics*, *118*(1), 1–41. https://doi.org/10.1162/00335530360535135
Pomeranz, K. (2000). *The Great Divergence*. https://doi.org/10.2307/j.ctt7sv80
Prebisch, R. (1950). The economic development of Latin America and its principal problems. Paper for the Economic Commission for Latin America. *Santiago de Chile*. Retrieved from https://repositorio.cepal.org/bitstream/handle/11362/10079/S6200129.pdf?sequence=1
Pritchett, L. (1997). Divergence: big time. *Journal of Economic Perspectives*. Retrieved from https://www.econ.nyu.edu/user/debraj/Courses/Readings/Pritchett.pdf?seq=14
Rodrik, D. (2016). Premature deindustrialization. *Journal of Economic Growth*, *21*(1), 1–33. https://doi.org/10.1007/s10887-015-9122-3
Rodrik, D., Subramanian, A., & Trebbi, F. (2004). Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development. *Journal of Economic Growth*, *9*(2), 131–165. https://doi.org/10.1023/B:JOEG.0000031425.72248.85
Rosenstein-Rodan, P. N. (1943). Problems of Industrialisation of Eastern and South-Eastern Europe. In *Source: The Economic Journal* (Vol. 53). Retrieved from https://about.jstor.org/terms
Sachs, J. D. (2001). Tropical Underdevelopment. *NBER Working Papers*, *8119*.
Sachs, J. D. (2003). Institutions Don't Rule: Direct Effects of Geography on Per Capita Income. *NBER Working Papers*, *9490*.
Schularick, M. (2006). A tale of two "globalizations": Capital flows from rich to poor in two eras of global finance. *International Journal of Finance and Economics*, *11*(4), 339–354. https://doi.org/10.1002/ijfe.302
Spolaore, E., & Wacziarg, R. (2013). How deep are the roots of economic development? *Journal of Economic Literature*, *51*(2), 325–369. https://doi.org/10.1257/jel.51.2.325
Sutton, J., & Trefler, D. (2016). Capabilities, Wealth, and Trade. *Journal of Political Economy*, *124*(3), 826–878. https://doi.org/10.1086/686034
Taylor, A. M., Elliott, G., Hatton, T., Jones, M., Lothian, J., & Mokyr, J. (2002). A century of current account dynamics. *Journal of International Money and Finance*, *21*, 725–745.
Touchelay, B. (2019). British and French Colonial Statistics: Development by Hybridization from the Nineteenth to the Mid-Twentieth Centuries. In *British and French Colonialism in Africa, Asia and the Middle East* (pp. 249–274). Springer.
United Nations Conference on Trade and Development (UNCTAD). (2020). UNCTAD Stat. Retrieved September 15, 2020, from Goods and Services (BPM6): Trade Openness Indicators website: https://unctadstat.unctad.org/wds/TableViewer/tableView.aspx?ReportId=90759
van Dam, A., & Frenken, K. (2020). Variety, complexity and economic development. *Research Policy*. https://doi.org/10.1016/j.respol.2020.103949
van Zanden, J. L., & Lindblad, J. T. (1989). De buitenlandse handel van Nederland, 1872-1913. *Economisch- En Sociaal-Historisch Jaarboek*, *52*, 231–269.
Voigtländer, N., & Voth, H.-J. (2012). Persecution Perpetuated: The Medieval Origins of Anti-Semitic Violence in Nazi Germany. *Quarterly Journal of Economics*, 1339–1392.
Voth, H. (2021). Persistence - Myth and Mystery. In *Handbook of Historical Economics*.
Wacziarg, R., & Welch, K. H. (2003). Trade Liberalization and Growth: New Evidence. *NBER*

Working Papers, *10152*.
Weitzman, M. L. (1998). Recombinant Growth. *The Quarterly Journal of Economics*, *113*(2), 331–360. https://doi.org/10.1162/003355398555595
Williamson, J. G. (2002). Land, labor, and globalization in the Third World, 1870-1940. *Journal of Economic History*, *62*(1), 55–85.
Williamson, J. G. (2006). *Globalization and the Poor Periphery Before 1950*. Cambridge, MA: MIT Press.
Williamson, J. G. (2011). Trade and Poverty: When the Third World Fell Behind. In *The New Palgrave Dictionary of Economics*. https://doi.org/10.1057/9780230226203.1724
Yates, L. (1959). *Forty Years of Foreign Trade: A Statistical Handbook*. London: Allen and Unwin.
Zhu, S., & Fu, X. (2013). Drivers of Export Upgrading. *World Development*, *51*, 221–233. https://doi.org/10.1016/j.worlddev.2013.05.017
Zhu, S., & Li, R. (2017). Economic complexity, human capital and economic growth: empirical research based on cross-country panel data. *Applied Economics*, *49*(38), 3815–3828. https://doi.org/10.1080/00036846.2016.1270413

Appendix: Measurement of Diversification, Ubiquity, Complexity and Sophistication

The measurement of export diversification in this paper is based on Hausmann et al. (2014, 24-5). Here we present a matrix version for easier replication. Let there be m varieties of commodities at a certain digit level (here SITC-3), and n countries in a certain period. Let E be the $m \times n$ matrix in which each column records a country's export values by commodity. Based on E, We calculate the revealed comparative advantage (RCA) of a country's exports by commodity, and then set to zero those elements of E that fall below a given threshold (here 0.5). Let M be the accompany matrix of E (after applying the RCA cutoff), the elements of which take on the value 1 or 0 depending on whether the corresponding element of E is positive or zero. Let e_c be an $n \times 1$ unit vector, and e_p an $1 \times m$ unit vector, then diversity is a $1 \times n$ vector $k_{c,0} = e_p M$, and ubiquity is a $m \times 1$ vector $k_{p,0} = M e_c$. Our diversification measurement that measures the variety of goods that country i exports, normalized by the maximum variety is calculated as:

$$diversification^i \equiv k_{c,0}^i / \mathbf{max}\,(k_{c,0})$$

where $\mathbf{max}\,(k_{c,0})$ takes the maximum of the elements of the vector $k_{c,0}$.

Let $\boldsymbol{diag}(v)$ be the transformation of the vector $v \equiv (v^1, v^2, \cdots, v^z)$ into a diagonal matrix, i.e.,

$$\boldsymbol{diag}(v) \equiv \begin{pmatrix} v^1 & 0 & \cdots & 0 \\ 0 & v^2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & v^z \end{pmatrix};$$

and use superscript T to denote transposition; then the iterative reweighting algorithms are given as: $k_{c,N} = k_{p,N-1}^T\, M\, \boldsymbol{diag}^{-1}(k_{c,0})$, and $k_{p,N} = \boldsymbol{diag}^{-1}(k_{p,0}) M k_{c,N-1}^T$. By iteration, we have $k_{c,N} = k_{p,N-1}^T M \boldsymbol{diag}^{-1}(k_{c,0}) = k_{c,N-2}\, M^T \boldsymbol{diag}^{-1}(k_{p,0})\; M\, \boldsymbol{diag}^{-1}(k_{c,0})$, and $k_{p,N} = \boldsymbol{diag}^{-1}(k_{p,0}) M k_{c,N-1}^T = \boldsymbol{diag}^{-1}(k_{p,0}) M\, \boldsymbol{diag}^{-1}(k_{c,0}) M^T k_{p,N-2}$. At convergence, $k_c \equiv \lim_{N\to\infty} k_{c,N}$ is the eigenvector of the matrix $\widetilde{M_{cc}} \equiv M^T \boldsymbol{diag}^{-1}(k_{p,0})\; M\, \boldsymbol{diag}^{-1}(k_{c,0})$, and $k_p \equiv \lim_{N\to\infty} k_{p,N}$ is the eigenvector of the matrix $\widetilde{M_{pp}} \equiv \boldsymbol{diag}^{-1}(k_{p,0}) M\, \boldsymbol{diag}^{-1}(k_{c,0}) M^T$. Following

Hausmann et al. (2014, 24), we use the eigenvector that is associated with the second largest eigenvalue for k_c and k_p. The economic complexity index of country i is therefore:

$$complexity^i \equiv \frac{k_c^i - \boldsymbol{mean}(k_c)}{\boldsymbol{sd}(k_c)},$$

where $\boldsymbol{mean}(\cdot)$ and $\boldsymbol{sd}(\cdot)$ take the average and standard deviation of all the elements in a vector, respectively.

The sophistication measure of country i we use as robustness check is:

$$sophistication^i \equiv \Sigma_j (E^{ji}/\Sigma_j E^{ji}) \frac{k_p^j - \boldsymbol{mean}(k_p)}{\boldsymbol{sd}(k_p)},$$

in which $\frac{k_p^j - \boldsymbol{mean}(k_p)}{\boldsymbol{sd}(k_p)}$ is known as the product complexity index (PCI) of product j. It measures the weighted average PCI of a country's export basket, using the product export shares of the basket as weights.

Tables and Figures

Table 1: Data Coverage of Export Value, Countries and Share in World Trade, 1897–1906

Year	Total covered trade (current billion USD)	Number of reporting countries	Percentage of world exports (compared to Federico-Tena dataset)
1897	5.706	53	68.7
1898	6.867	53	81.4
1899	7.467	55	80.7
1900	7.991	63	81.2
1901	8.553	94	87.9
1902	8.759	98	86.5
1903	9.329	99	87.3
1904	9.913	100	89.0
1905	10.593	99	87.5
1906	11.485	73	86.2

1. Species and bullions are not included.

Table 2: Summary Statistics for Proxies of Productive Capabilities, 1897–1906 and 1998–2007

Variable	Unit	Max	Mean	Min	Sd
Diversification (1897:1906)	Ratio	1.00	0.24	0.06	0.18
Diversification (1998:2007)	Ratio	1.00	0.42	0.02	0.26
Complexity (1897:1906)	Index	2.23	0.07	−1.94	1.02
Complexity (1998:2007)	Index	2.37	0.06	−2.49	1.02
Sophistication (1897:1906)	Index	0.68	−0.91	−2.12	0.77
Sophistication (1998:2007)	Index	1.07	−0.72	−2.06	0.86
Manufacturing (1897:1906)	Ratio	0.71	0.09	0.00	0.18
Manufacturing (1998:2007)	Ratio	0.99	0.46	0.00	0.27
First principal component (1897:1906)	Index	2.89	0.14	−2.16	1.27
First principal component (1998:2007)	Index	2.69	−0.24	−3.10	1.33
Trade openness	Ratio	2.05	0.71	0.01	0.36
Resource rent	Ratio	0.61	0.08	0.00	0.11
FDI	Ratio	0.24	0.04	0.00	0.04
Population	Million	1282.67	61.98	0.27	182.66
Human capital	Index	3.60	2.37	1.14	0.67
Polity	Index	10.00	4.89	−9.00	5.63

1. Observations = 86. Measured for 1998:2007 where not indicated.
2. First principal component calculated based on normalized scores of diversification, complexity, sophistication, and manufacturing share.

Table 3: Correlations Across Proxies of Productive Capabilities, 1897:1906 and 1998:2007

Panel (a)	1897:1906			
	Diversification	Complexity	Sophistication	Manufacturing
Diversification	1			
Complexity	0.69***	1		
Sophistication	0.72***	0.93***	1	
Manufacturing	0.74***	0.63***	0.64***	1

Panel (b)	1998:2017			
	Diversification	Complexity	Sophistication	Manufacturing
Diversification	1			
Complexity	0.73***	1		
Sophistication	0.80***	0.84***	1	
Manufacturing	0.44***	0.72***	0.58***	1

1. Presented are Pearson's product moment correlation coefficients.
2. *p<0.1; **p<0.05; ***p<0.01.

Table 4: Variables and Data Sources

Variable	Source	Measure	Code
FDI	World Bank[1]	FDI (net inflows)/GDP	BX.KLT.DINV.WD.GD.ZS
Trade openness	World Bank[2]	Trade/GDP	NE.TRD.GNFS.ZS
Resource rent	World Bank	Total natural resources rents/ GDP	NY.GDP.TOTL.RT.ZS
Population	World Bank	Population, total	SP.POP.TOTL
Polity	POLITY V data set[3]	Democracy score - autocracy score	polity2
Human capital	Penn World Table 9.1[4]	See "Human capital in PWT 9.0"	hc
Soil suitability	Center for Int'l Development	Percent of very suitable land	soilsui1
Ruggedness	Nunn and Puga (2012)	See source for details	
Latitude	Acemoglu et al. (2002, Table 3)	See source for details	lat_abst
Temperature	Acemoglu et al. (2002, Table 3)	See source for details	temp4
Humidity	Acemoglu et al. (2002, Table 3)	See source for details	humid4
Coal	Acemoglu et al. (2002, Table 3)	See source for details	coal
Population (historical)	HYDE 3.2 (Goldewijk et al., 2017)	See source for details	
Labor productivity Land productivity	FAO-GAEZ, HYDE 3.2	See text for details	
European overseas colony	Bernhard et al. (2004, Figure 1)	See source for details	
European colonial power	Bernhard et al. (2004, Figure 1)	See source for details	
OPEC	www.opec.org	Member or not	
Democratization	Constructed from polity	Score reaches 6 and above first time	

1. For Cuba, FDI (1993-2009) are from NACLA (https://www.nordeatrade.com/en/explore-new-market/cuba/investment), and GDP data are from World Bank (NY.GDP.MKTP.CD).
2. For Trinidad and Tobago, trade openess = merchandise trade/GDP + trade in services/GDP (World Bank, TG.VAL.TOTL.GD.ZS, BG.GSR.NFSV.GD.ZS).
3. Polity index for Bahamas, Iceland, and Belize are coded by the authors using the polity index score of similar countries according to the Freedom House Index (https://freedomhouse.org/).
4. Mean years of schooling data from human development data (the United Nations) are used to compute the human capital index according to "Human capital in PWT 9.0" for the following countries: Cuba, Georgia, Guinea, Guyana, Oman, Papua, New Guinea, Bahamas, Grenada, Somalia, Seychelles, Tonga, Samoa.

Table 5: Main Results—Persistence in Productive Capabilities 1897:1906 vs. 1998:2007

	Dependent variable (1998:2007)									
Panel (a)	N.m	N.m	ECI	ECI	SOP	SOP	MAN	MAN	PC1	PC1
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable (1897:1906)	1.01***	0.45***	0.70***	0.28***	0.71***	0.32***	0.43***	0.27***	0.73***	0.31***
	(0.14)	(0.15)	(0.07)	(0.10)	(0.09)	(0.11)	(0.10)	(0.10)	(0.07)	(0.10)
Trade openness		0.03		0.23		0.61***		0.36***		0.61*
		(0.05)		(0.30)		(0.18)		(0.08)		(0.33)
Resource rent		−0.81***		−3.24***		−3.79***		−1.82***		−5.15***
		(0.20)		(1.00)		(0.75)		(0.27)		(1.20)
FDI		0.90**		−0.05		0.12		−1.48***		−0.05
		(0.41)		(1.52)		(1.29)		(0.49)		(1.87)
Population (log)		0.05***		0.19***		0.22***		0.05***		0.30***
		(0.01)		(0.05)		(0.04)		(0.01)		(0.06)
Human capital		0.07**		0.52***		0.14		−0.02		0.47**
		(0.03)		(0.18)		(0.14)		(0.05)		(0.21)
Polity		0.01*		0.01		0.02*		−0.004		0.02
		(0.004)		(0.01)		(0.01)		(0.005)		(0.02)
Constant	0.18***	−0.03	0.01	−1.69***	−0.07	−1.61***	0.42***	0.28*	−0.34***	−2.34***
	(0.04)	(0.07)	(0.08)	(0.49)	(0.13)	(0.42)	(0.03)	(0.15)	(0.11)	(0.56)
Observations	86	86	86	86	86	86	86	86	86	86
R^2	0.48	0.75	0.49	0.72	0.39	0.73	0.08	0.55	0.49	0.76

	Dependent variable (1979:1988)					Dependent variable (1962:1971)				
Panel (b)	N.m	ECI	SOP	MAN	PC1	N.m	ECI	SOP	MAN	PC1
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable (1897:1906)	0.47***	0.43***	0.43***	0.21	0.45***	0.66***	0.60***	0.39***	0.36***	0.52***
	(0.12)	(0.09)	(0.10)	(0.14)	(0.08)	(0.15)	(0.14)	(0.08)	(0.10)	(0.11)
Trade openness	0.20**	0.57*	0.71***	0.13	0.90**	0.30***	0.88*	0.92***	0.31**	1.29***
	(0.08)	(0.34)	(0.23)	(0.20)	(0.36)	(0.11)	(0.46)	(0.23)	(0.16)	(0.49)
Resource rent	−1.33***	−4.14***	−5.28***	−1.90**	−6.79***	−1.75***	−4.92***	−5.54***	−2.48***	−7.58***
	(0.23)	(1.09)	(0.89)	(0.75)	(1.21)	(0.39)	(1.61)	(1.08)	(0.58)	(1.73)
FDI	0.86	5.08	−3.94	−3.38	1.46	−0.58	−5.04	−1.54	−1.18	−5.23
	(2.30)	(7.10)	(5.26)	(4.03)	(7.58)	(1.23)	(5.98)	(3.94)	(1.85)	(6.06)
Population (log)	0.06***	0.18***	0.17***	0.03	0.25***	0.04**	0.10	0.15***	0.05**	0.18**
	(0.01)	(0.05)	(0.04)	(0.03)	(0.06)	(0.02)	(0.08)	(0.04)	(0.02)	(0.09)
Human capital	0.10***	0.22	0.31**	−0.07	0.34*	0.12***	0.20	0.39***	−0.11	0.42
	(0.03)	(0.18)	(0.14)	(0.06)	(0.19)	(0.04)	(0.29)	(0.13)	(0.08)	(0.29)
Polity	0.003	0.02*	0.01	0.01	0.02	−0.003	0.01	−0.002	0.01	0.01
	(0.002)	(0.01)	(0.01)	(0.01)	(0.01)	(0.003)	(0.02)	(0.01)	(0.01)	(0.02)
Constant	−0.16**	−1.02**	−1.70***	0.53***	−1.59***	−0.20*	−0.69	−1.78***	0.36**	−1.41**
	(0.08)	(0.48)	(0.40)	(0.20)	(0.52)	(0.11)	(0.62)	(0.33)	(0.18)	(0.62)
Observations	69	69	69	69	69	62	62	62	62	62
R^2	0.85	0.78	0.83	0.34	0.85	0.82	0.76	0.82	0.38	0.80

1. N.m = diversification, ECI = complexity, SOP = sophistication, MAN = manufacturing, PC1 = first principal component.
2. Robust standard errors (HC3) are used. See Long and Ervin (2000).
3. *p<0.1; **p<0.05; ***p<0.01.

Table 6: Productive Capabilities 1897:1906 vs. GDP per capita 1998:2007

Panel (a)	Dependent variable: Real GDP (1998:2007)									
	N.m		ECI		SOP		MAN		PC1	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable (1897:1906)	3.71***	1.92**	0.88***	0.38***	1.10***	0.36**	3.50***	1.28***	0.70***	0.29***
	(0.62)	(0.82)	(0.07)	(0.13)	(0.11)	(0.14)	(0.44)	(0.43)	(0.05)	(0.10)
Trade openness		0.12		0.04		0.02		0.06		0.03
		(0.30)		(0.37)		(0.31)		(0.28)		(0.35)
Resource rent		0.36		0.80		0.64		0.21		0.78
		(0.97)		(0.91)		(0.96)		(0.95)		(0.93)
FDI		−2.32		−0.48		−0.05		−0.56		−0.43
		(2.41)		(2.58)		(2.56)		(2.39)		(2.61)
Population (log)		−0.13*		−0.06		−0.04		−0.06		−0.06
		(0.07)		(0.06)		(0.06)		(0.06)		(0.06)
Human capital		1.21***		1.01***		1.14***		1.31***		1.03***
		(0.17)		(0.23)		(0.20)		(0.14)		(0.22)
Polity		0.01		0.02		0.02		0.01		0.02
		(0.02)		(0.02)		(0.02)		(0.02)		(0.02)
Constant	7.98***	5.83***	8.81***	6.47***	9.88***	6.48***	8.56***	5.76***	8.78***	6.41***
	(0.20)	(0.35)	(0.09)	(0.51)	(0.12)	(0.54)	(0.13)	(0.33)	(0.09)	(0.49)
Observations	83	83	83	83	83	83	83	83	83	83
R^2	0.31	0.75	0.56	0.76	0.50	0.74	0.27	0.74	0.56	0.75

Panel (b)	Dependent variable: Real GDP (1979:1988)					Dependent variable: Real GDP (1962:1971)				
	N.m	ECI	SOP	MAN	PC1	N.m	ECI	SOP	MAN	PC1
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable (1897:1906)	1.52**	0.33***	0.28*	0.78**	0.24***	1.48**	0.28***	0.17	0.83**	0.19**
	(0.70)	(0.12)	(0.15)	(0.33)	(0.09)	(0.60)	(0.11)	(0.14)	(0.34)	(0.08)
Trade openness	−0.05	0.12	0.14	0.10	0.10	−0.61*	−0.41	−0.28	−0.45	−0.39
	(0.35)	(0.33)	(0.35)	(0.33)	(0.33)	(0.35)	(0.36)	(0.37)	(0.38)	(0.36)
Resource rent	1.31	1.37	1.16	1.08	1.30	1.40	1.06	1.02	1.54	1.01
	(1.48)	(1.50)	(1.64)	(1.59)	(1.55)	(1.46)	(1.70)	(1.91)	(1.67)	(1.78)
FDI	3.22	4.05	3.96	6.66	4.25	−3.26	1.67	−1.33	−2.75	0.83
	(8.97)	(9.10)	(8.44)	(9.00)	(8.76)	(5.44)	(8.43)	(8.51)	(6.29)	(8.92)
Population (log)	−0.11	−0.06	−0.03	−0.04	−0.05	−0.20**	−0.12**	−0.09	−0.13*	−0.11*
	(0.09)	(0.07)	(0.08)	(0.08)	(0.08)	(0.08)	(0.06)	(0.06)	(0.07)	(0.06)
Human capital	0.84***	0.63***	0.77***	0.95***	0.67***	0.98***	0.79***	1.05***	1.15***	0.87***
	(0.13)	(0.18)	(0.23)	(0.14)	(0.19)	(0.17)	(0.24)	(0.24)	(0.18)	(0.24)
Polity	0.03**	0.04***	0.04**	0.03**	0.04***	0.02	0.02	0.01	0.01	0.02
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Constant	6.73***	7.25***	7.16***	6.53***	7.17***	6.98***	7.34***	6.98***	6.75***	7.18***
	(0.37)	(0.48)	(0.67)	(0.36)	(0.50)	(0.37)	(0.48)	(0.55)	(0.40)	(0.47)
Observations	69	69	69	69	69	62	62	62	62	62
R^2	0.76	0.77	0.75	0.74	0.76	0.77	0.76	0.74	0.75	0.75

1. N.m = diversification, ECI = complexity, SOP = sophistication, MAN = manufacturing, PC1 = first principal component.
2. Real GDP is log-transformed.
3. Robust standard errors (HC3) are used. See Long and Ervin (2000).
4. *p<0.1; **p<0.05; ***p<0.01.

Table 7: Correlates of Productive Capabilities, 1897:1906

	Dependent variable (1897:1906)									
	N.m		ECI		SOP		MAN		PC1	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Soil suitability	4.28*		16.40***		9.03*		0.67		18.88**	
	(2.38)		(6.18)		(5.17)		(2.22)		(7.73)	
Ruggedness	−0.04***		−0.003		0.01		−0.02		−0.002	
	(0.01)		(0.05)		(0.03)		(0.02)		(0.06)	
Latitude	0.62***		5.19***		3.99***		0.51***		6.59***	
	(0.19)		(0.55)		(0.44)		(0.18)		(0.66)	
Temperature	0.001		0.01**		0.01**		−0.0005		0.02**	
	(0.002)		(0.01)		(0.01)		(0.002)		(0.01)	
Humidity	0.001		−0.01**		−0.01*		0.002*		−0.01**	
	(0.001)		(0.004)		(0.004)		(0.001)		(0.01)	
Coal	0.08***		0.31***		0.18*		−0.004		0.36***	
	(0.03)		(0.11)		(0.10)		(0.03)		(0.14)	
Population (log)		0.04***		0.12***		0.07*		0.02**		0.14***
		(0.01)		(0.04)		(0.04)		(0.01)		(0.06)
Labor productivity		3.05**		25.32***		17.47**		0.27		30.79***
		(1.20)		(6.24)		(7.26)		(1.02)		(8.99)
Land productivity		−8.14*		−57.25**		−45.88*		−7.49		−74.14**
		(4.53)		(27.17)		(23.78)		(5.51)		(34.86)
European overseas colony (dummy)		−0.05*		−0.88***		−0.59***		−0.06*		−1.06***
		(0.03)		(0.16)		(0.15)		(0.04)		(0.21)
European colonial power (dummy)		0.27***		0.76***		0.61***		0.27***		1.02***
		(0.07)		(0.22)		(0.18)		(0.09)		(0.28)
Constant	−0.001	0.25***	−0.93***	0.68***	−1.67***	−0.45*	−0.17**	0.15**	−1.15**	0.90**
	(0.09)	(0.04)	(0.35)	(0.26)	(0.36)	(0.25)	(0.08)	(0.07)	(0.45)	(0.35)
Observations	82	82	82	82	82	82	82	82	82	82
R^2	0.50	0.71	0.79	0.62	0.75	0.52	0.40	0.52	0.80	0.61

1. N.m = diversification, ECI = complexity, SOP = sophistication, MAN = manufacturing, PC1 = first principal component.
2. Robust standard errors (HC3) are used. See Long and Ervin (2000).
3. *p<0.1; **p<0.05; ***p<0.01.
4. On top of the control variables presented here, we have also tried including area, distance to coast and river, percentage of land in the tropical area, land-lock or not, island or not, other measures of temperature and humidity, other measures of land suitability, constraint on executive power, European descendants, reserves of gold, iron, silver, zinc and oil, European migrants, urbanization rate, population density, real wage rate, and steamship travel time, which are all insignificant at the 10% level across models.

Table 8: Correlates in 1897:1906 vs. Productive Capabilities in 1998:2007

	Dependent variable (1998:2007)									
	N.m		ECI		SOP		MAN		PC1	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable (1897:1906)	0.75***	0.39	0.61***	0.55***	0.61***	0.54***	0.57***	0.40**	0.71***	0.58***
	(0.19)	(0.26)	(0.16)	(0.17)	(0.20)	(0.15)	(0.14)	(0.20)	(0.17)	(0.16)
Soil suitability	3.00		5.97		8.67		0.82		8.63	
	(2.80)		(7.16)		(8.54)		(3.51)		(10.08)	
Ruggedness	−0.01		−0.03		0.02		0.08***		−0.01	
	(0.03)		(0.15)		(0.10)		(0.03)		(0.17)	
Latitude	0.23		−0.86		−0.30		−0.33		−1.51	
	(0.25)		(1.01)		(1.08)		(0.35)		(1.46)	
Temperature	−0.0005		−0.02**		−0.01		−0.003		−0.02	
	(0.003)		(0.01)		(0.01)		(0.005)		(0.01)	
Humidity	0.001		0.01		0.01		0.0003		0.01	
	(0.003)		(0.01)		(0.01)		(0.003)		(0.01)	
Coal	0.06		0.28		0.31*		0.02		0.38	
	(0.05)		(0.22)		(0.19)		(0.06)		(0.28)	
Population (log)		0.03		0.05		0.07		0.01		0.08
		(0.02)		(0.07)		(0.05)		(0.02)		(0.08)
Labor productivity		4.85***		−6.29		−6.43		−5.42*		−10.07
		(1.86)		(8.17)		(7.70)		(3.04)		(9.46)
Land productivity		7.41		51.10		85.86***		35.87***		100.97**
		(8.47)		(34.70)		(28.18)		(10.77)		(43.41)
European overseas colony (dummy)		−0.15***		−0.54**		−0.42**		−0.15**		−0.67**
		(0.06)		(0.25)		(0.21)		(0.06)		(0.32)
European colonial power (dummy)		0.14		−0.06		0.14		−0.04		0.02
		(0.12)		(0.21)		(0.26)		(0.11)		(0.31)
Constant	0.07	0.26**	−0.42	−0.13	−0.86	−0.78**	0.34*	0.22**	−1.01	−0.85**
	(0.20)	(0.10)	(0.80)	(0.30)	(0.74)	(0.32)	(0.20)	(0.11)	(1.02)	(0.41)
Observations	82	82	82	82	82	82	82	82	82	82
R^2	0.53	0.61	0.52	0.55	0.42	0.54	0.21	0.26	0.51	0.57

1. N.m = diversification, ECI = complexity, SOP = sophistication, MAN = manufacturing, PC1 = first principal component.
2. Robust standard errors (HC3) are used. See Long and Ervin (2000).
3. *p<0.1; **p<0.05; ***p<0.01.

Table 9: Path-defiance—Resource Curse and Democratization

	Dependent variable (1998:2007)				
Panel (a)	N.m	ECI	SOP	MAN	PC1
	(1)	(2)	(3)	(4)	(5)
Dependent variable (1897:1906)	1.07***	0.67***	0.61***	0.37***	0.68***
	(0.15)	(0.06)	(0.10)	(0.10)	(0.07)
OPEC	−0.11	−0.79*	−1.70***	−0.36***	−1.39***
	(0.12)	(0.46)	(0.37)	(0.08)	(0.48)
Interaction	−0.63	−0.13	−0.57***	−0.65	−0.36
	(0.69)	(0.45)	(0.20)	(0.64)	(0.38)
Constant	0.18***	0.07	−0.03	0.46***	−0.20**
	(0.03)	(0.07)	(0.12)	(0.03)	(0.09)
Observations	99	99	99	99	99
R^2	0.61	0.54	0.44	0.24	0.55

	Dependent variable (1998:2007)				
Panel (b)	N.m	ECI	SOP	MAN	PC1
	(1)	(2)	(3)	(4)	(5)
Dependent variable (1897:1906)	1.02***	0.69***	0.65***	0.39***	0.70***
	(0.15)	(0.06)	(0.10)	(0.10)	(0.07)
Democratization (second wave)	0.06	0.84***	0.74**	0.16	0.91***
	(0.13)	(0.16)	(0.33)	(0.14)	(0.21)
Interaction	−0.02	0.05	0.34	0.07	0.16
	(0.33)	(0.35)	(0.25)	(0.40)	(0.32)
Constant	0.16***	−0.07	−0.14	0.40***	−0.41***
	(0.03)	(0.08)	(0.13)	(0.03)	(0.10)
Observations	96	96	96	96	96
R^2	0.51	0.55	0.41	0.11	0.53

	Dependent variable (1998:2007)				
Panel (c)	N.m	ECI	SOP	MAN	PC1
	(1)	(2)	(3)	(4)	(5)
Dependent variable (1897:1906)	1.05***	0.73***	0.75***	0.45***	0.77***
	(0.14)	(0.07)	(0.10)	(0.10)	(0.07)
Democratization (third wave)	0.06	0.38**	0.06	0.06	0.53**
	(0.08)	(0.17)	(0.31)	(0.07)	(0.24)
Interaction	0.59*	−0.13	−0.26	0.31	−0.19
	(0.34)	(0.17)	(0.26)	(1.10)	(0.20)
Constant	0.12***	−0.09	−0.08	0.39***	−0.45***
	(0.03)	(0.08)	(0.14)	(0.04)	(0.10)
Observations	96	96	96	96	96
R^2	0.61	0.52	0.41	0.09	0.52

1. N.m = diversification, ECI = complexity, SOP = sophistication, MAN = manufacturing, PC1 = first principal component.
2. Robust standard errors (HC3) are used. See Long and Ervin (2000).
3. *p<0.1; **p<0.05; ***p<0.01.

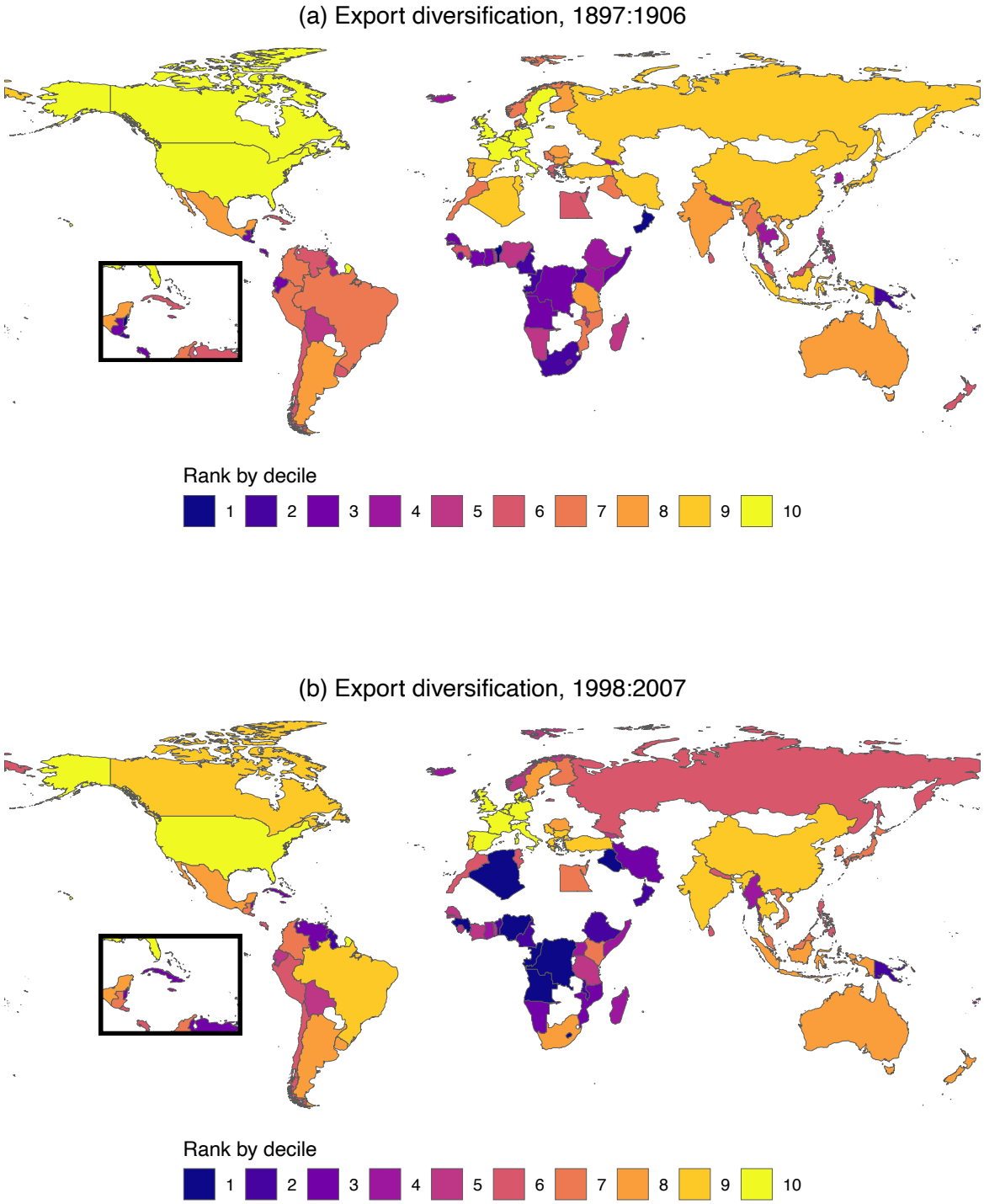

Figure 1: World Map of Export Diversification in 1897:1906 and 1998:2007

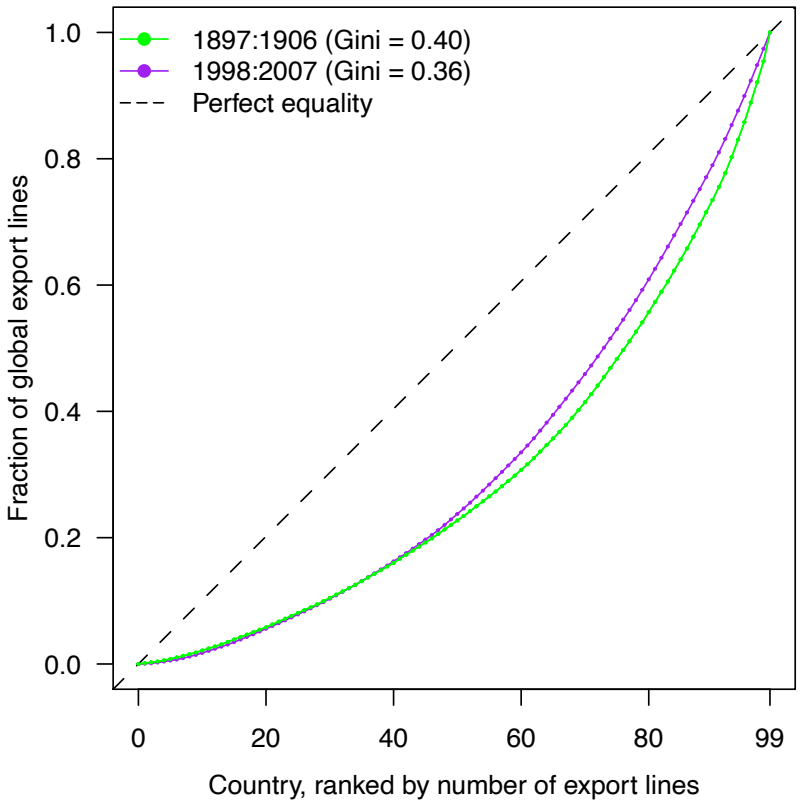

Figure 2: Lorenz Curve of Inequality in Export Diversification 1897–1906 and 1998–2007

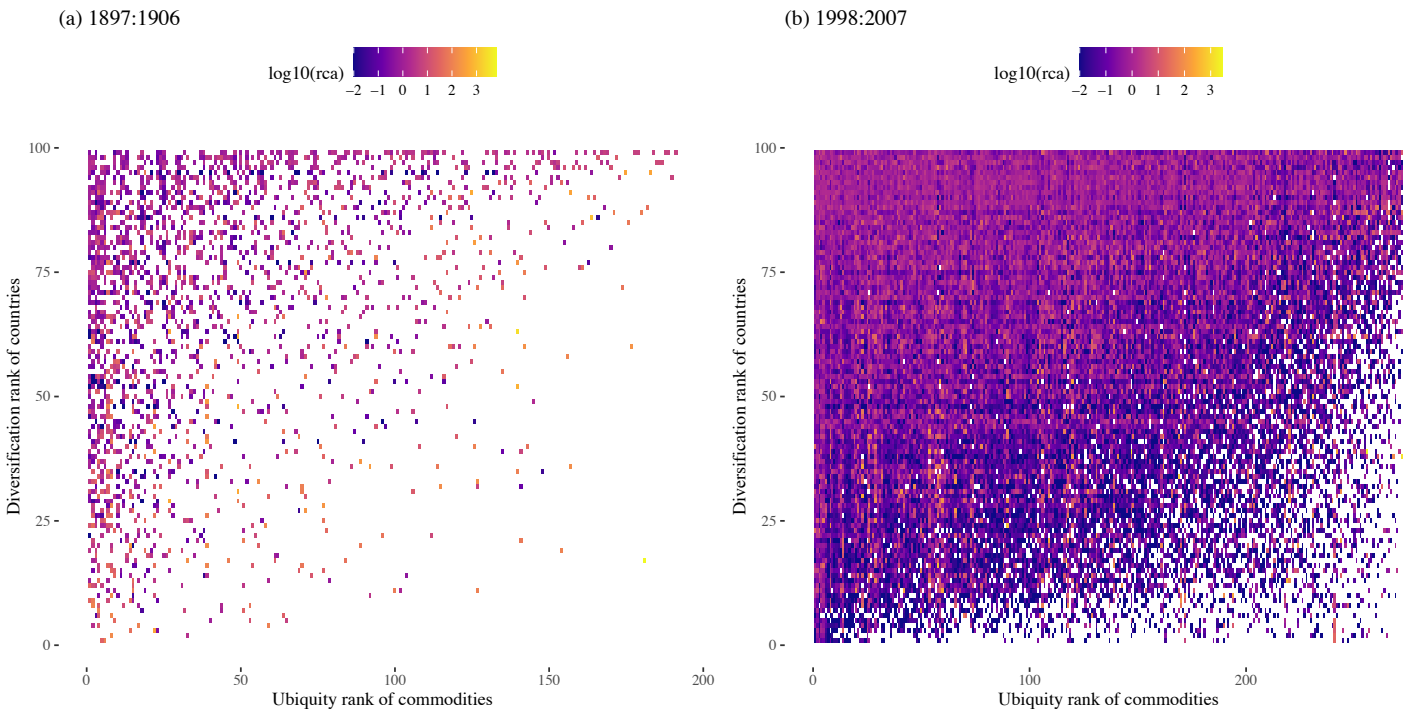

Figure 3: Matrices of Ubiquity and Diversification in 1897:1906 and 1998:2007

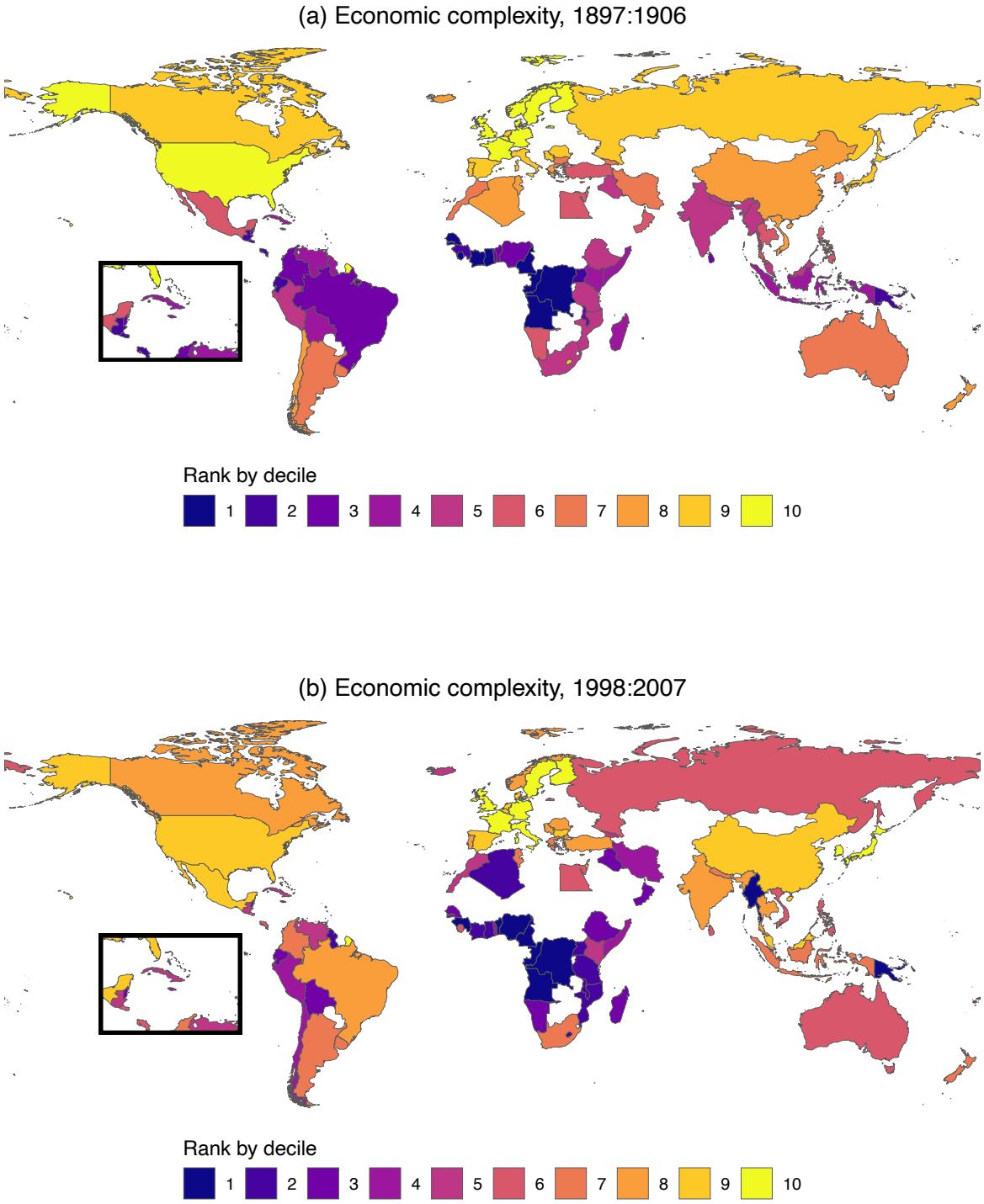

Figure 4: World Map of Economic Complexity in 1897:1906 and 1998:2007

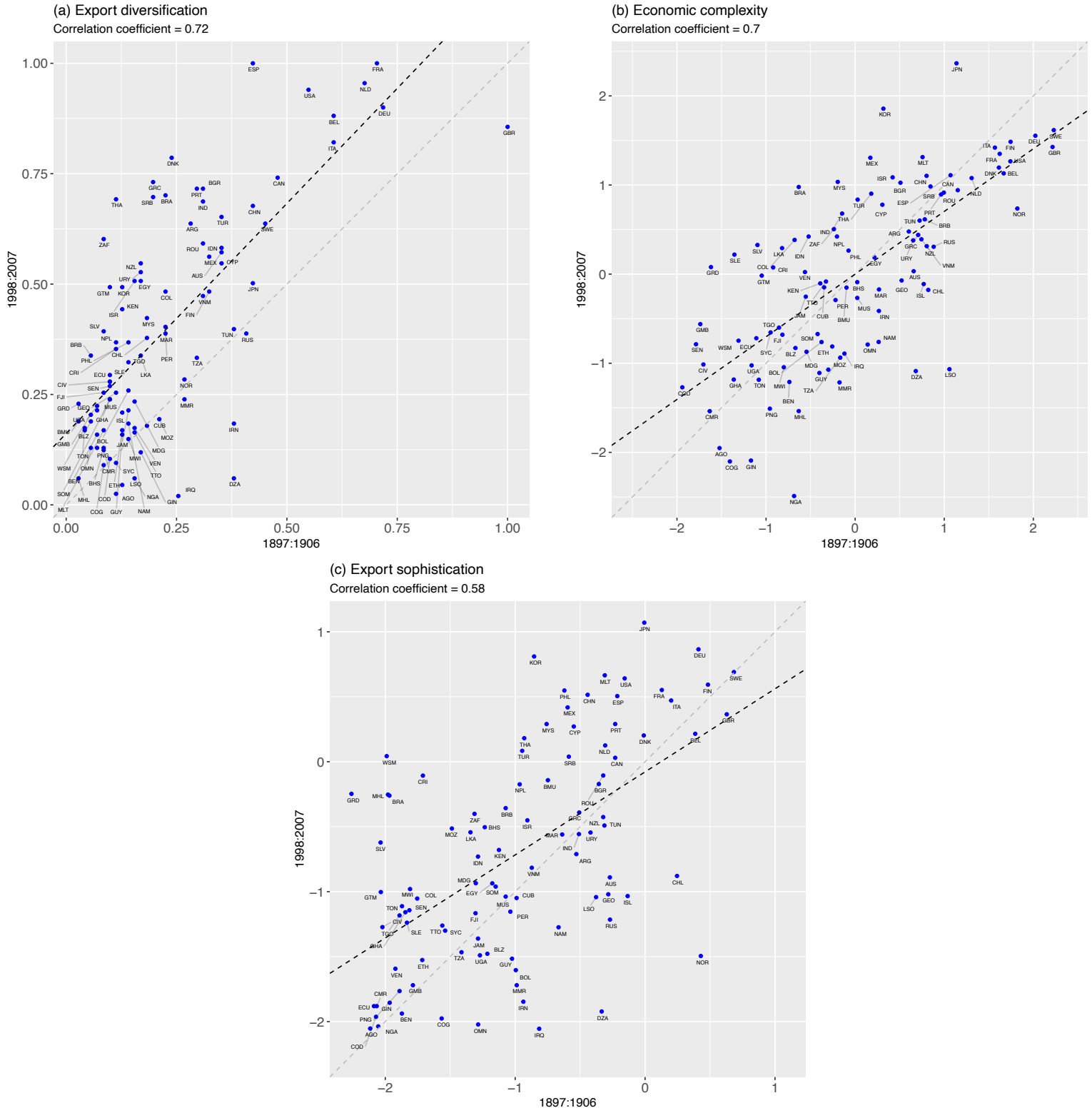

Figure 5: Export Diversification, economic complexity, and export sophistication, 1897-1906 vs. 1998-2007

1. The gray dashed line is the 45 degree line, and the black dashed line is the OLS fitted line.
2. Correlation coefficient is Pearson's product moment correlation coefficient.

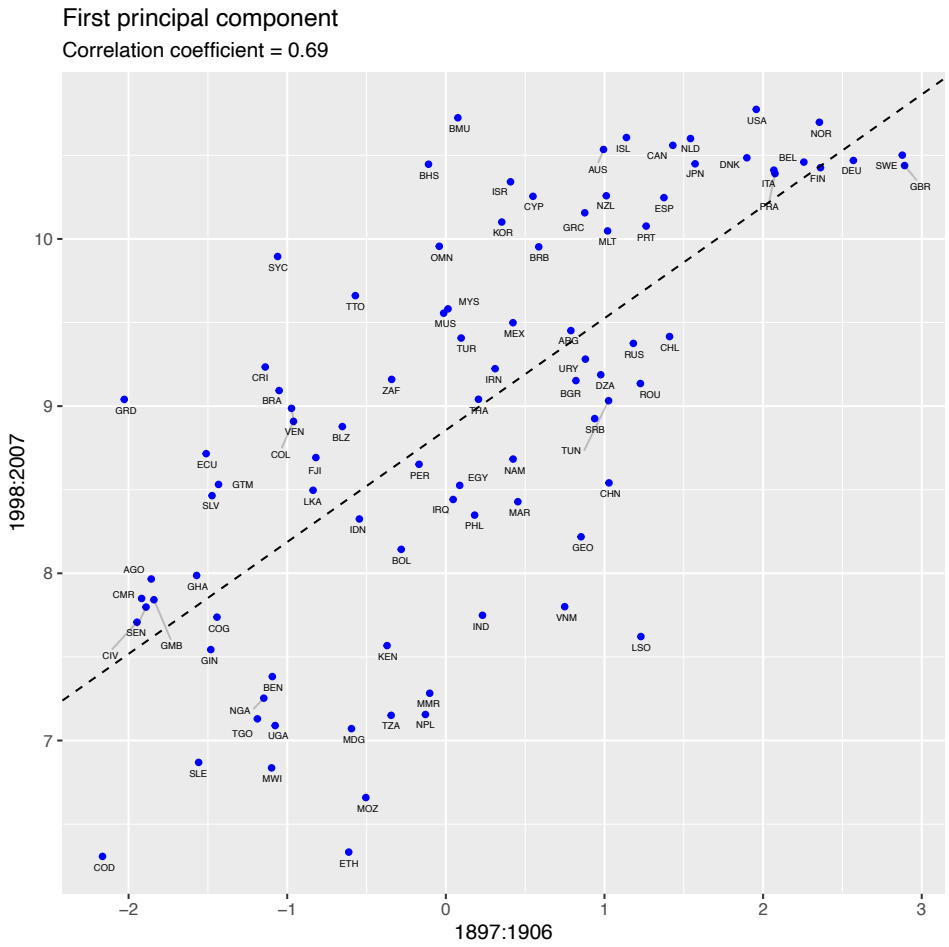

Figure 6: First principal Component 1897:1906 vs. GDP per capita 1998:2007

1. The gray dashed line is the 45 degree line, and the black dashed line is the OLS fitted line.
2. Correlation coefficient is Pearson's product moment correlation coefficient.