The Structural Economic Roots of Liberal Democracy

Sam van Noort*

This version: December 7, 2019
(Latest version: **click here**)

Abstract

Recent studies on the modernization theory have found little evidence of a consistent effect of income on democracy. I argue that this is because only growth derived from manufacturing is conducive to democracy, while growth based on other economic activities is unlikely to affect democracy (e.g. farming) or may even harm it (e.g. mining). To test this theory I use the geological potential to have domestic access to coal as an instrument for industrialization. Until far into the 20th century manufacturing production strongly relied on domestic access to coal. Coal-induced differences in industrialization have since then largely persisted through time. Coal itself is meanwhile formed due to an exogenous process and is unlikely to affect democracy through factors besides industrialization. In line with my theory I find large and highly robust effects of industrialization on democracy. These results are confirmed in two-way fixed effects models using data from 148 countries over the 1845 to 2015 period.

Keywords: Industrialization, Democracy, Modernization Theory
JEL Codes: O14, P16, P51

*Ph.D. candidate at the University of Cambridge and Fox International Fellow at Yale University. E-mail address: sam.vannoort@yale.edu. Website: sites.google.com/view/samvannoort. I am grateful to Toke Aidt, Ben Ansell, Carles Boix, Dawn Brancati, Brian Burgoon, Ha-Joon Chang, Andrew Eggers, Douglas Gollin, Tanushree Goyal, Elizabeth Harper, Hanna Kleider, David Samuels, Milan Svolik, and seminar participants at the University of Oxford, Yale University, and University College London for helpful comments and suggestions. Financial support from the Vice-Chancellor's Award at the University of Cambridge and the Fox International Fellowship at Yale University is gratefully acknowledged.

1 Introduction

The relationship between economic development and political democracy has generated prolonged debate within political science and economics. The existing literature on this topic has focused on the statistical association between GDP per capita and regime type, and has found little evidence for the idea that a higher level of average per capita income tends to make the introduction and sustainment of democratic forms of government more likely. The seminal work of Przeworski and Limongi (1997) has found that GDP per capita is only correlated with democratic consolidations, not democratic transitions. Even more fundamentally the work of Acemoglu et al. (2008, 2009) suggests that the correlation between GDP per capita and democracy disappears all together ones time-invariant confounding factors are controlled for via country fixed effects. While other studies have challenged these findings a recent meta-analysis of all econometric studies conducted between 1999 to 2015 confirms that the existing literature on the whole indicates that GDP per capita has on average *no* effect on democracy (Broderstad, 2018).[1]

In this paper I step away from the (implicit) assumption that any increase in economic production (as captured by GDP per capita) is equally likely to induce transitions to and consolidations of democracy. Instead I argue that how increases in economic production affect democracy is conditional on the type of production activity involved, and thus that we should not expect to find an unconditionally positive effect of income on democracy. More specifically I argue that only increases in economic activity in the manufacturing sector, in the form of *industrialization*, are likely to trigger causal mechanisms that are conducive to democracy.[2] Increases in other types of production activities, in contrast, do contribute to GDP but are unlikely to affect democracy (e.g. tourism, family farming, retail services), or are in fact likely to negatively affect the long-run prospects of democracy (e.g. natural resource extraction, agricultural production by large landowners, tax havens).

I substantiate this hypothesis by systematically linking economic theory on the structural economic properties of different economic sectors with comparative politics literature on the proximate causes of democracy. In sum, I argue that the manufacturing sector has a number of distinctive structural economic properties (e.g. high-productivity, large capacity to absorb labor, economy-wide linkages, high-returns to education, strong reliance on labor cooperation) that makes industrialization likely to positively affect a wide-range of

[1] Examples of studies that do find a positive effect of GDP per capita on democracy are: Boix and Stokes (2003), Epstein et al. (2006), Boix (2011), Benhabib, Corvalan and Spiegel (2013), and Che et al. (2013).

[2] Throughout the paper I refer to "industrialization" as an increase, in both absolute and relative terms, of economic activity in the manufacturing sector. Hence I refer to the extractive/mining sector, which is sometimes grouped together with manufacturing in the category "industry", as a separate sector.

in the comparative politics literature well-established proximate causes of democracy (e.g. living standard working and middle classes, population-wide education levels, organizatory capacity civil society). Similarly, I derive directly from the economic properties of the agricultural, extractive, and (pre-industrial) service sectors that growth stemming from these other segments of the economy is unlikely to lead to much politically meaningful structural socio-economic change, or is in fact likely to lead to types of structural socio-economic change that negatively affect democracy. As such I open the blackbox of GDP and provide a detailed theoretical framework that shows how and why it matters for democracy, and political economic processes more generally, from which segments of the economy growth originates.

To test the effect of industrialization on democracy I combine standard data on regime type with newly collected data on the fraction of the population employed in manufacturing for 148 countries over the 170 years between 1845 and 2015. For econometric identification I introduce a novel instrumental variable for cross-country differences in industrialization: a country' geological potential to have domestic access to coal.[3] This identification strategy exploits that: (1) industrial production until far into the 1970s strongly relied on coal; (2) coal deposits are highly unequally distributed across countries; (3) coal was before the worldwide introduction of containerized shipping in the 1970s limitedly tradable due to its exceptionally heavy weight relative to its per kilogram value; (4) cross-country differences in industrialization have since the 1970s (and before) been subject to strong path dependency; and (5) the cross-country distribution of coal deposits is determined by a geological process that is exogenous to political development.

Identification in my instrumental variable approach relies on the assumption that geologically determined access to coal does not affect democracy through channels other than industrialization. As is true for any instrumental variable model this exclusion restriction cannot be proven empirically. I nonetheless argue that the exclusion restriction underlying my approach is plausible because the process of coal formation is unlikely to be correlated with other determinants of democracy. As I discuss in more detail later most coal was formed in countries that were subject to a tropical climate during the Carboniferous era (i.e. approximately 358.9 to 298.9 million years ago).[4] In countries that were subject to a tropical climate during the Carboniferous period climatical conditions were highly conducive to plant growth, which in combination with regularly changing sea levels, caused large amounts of plants to be rapidly buried underneath deltaic sands. Over time these buried tropical plants

[3] As I discuss in more detail later I instrument with the geological potential to have coal, rather than the stock of proven coal reserves, because: (1) it is almost certainly true that not all coal deposits of the world have yet been discovered; and (2) coal discovery may be endogenous because more industrialized and more democratic countries may have (had) a greater incentive and/or a greater capacity to discover coal.

[4] Carboniferous literally means "coal-bearing" in Latin.

from the Carboniferous era produced most of the coal we know today. Countries that were not located in the Carboniferous tropics, in contrast, did not experience this process and are therefore significantly less likely to have (had) coal desposits on their own territories. Importantly for my purpose climate in the Carboniferous period is uncorrelated with current latitude/climate because of tectonic dynamics over the past 300 million years.[5] Besides causing major cross-country difference in coal deposits, which in turn affected the historic opportunity cost of industrialization, tropical climate during the Carboniferous era does not appear to have affected any other factors relevant to democracy.

I complement my instrumental variable estimates with results from dynamic panel models with country fixed effects, time fixed effects, and lags of industrialization on the right-hand side.[6] Although correlational in nature such two-way fixed effects models do control by design (rather than by assuming the instrument to be valid) for: (i) reversed causality (i.e. democracy causing industrialization, rather than vice versa – North (1990)); (ii) all time-invariant factors (e.g. historical critical junctures – Acemoglu et al. (2008, 2009)); and (iii) all confounding factors that affect all countries at the same point in time (e.g. international shocks/spillovers – Huntington (1991)). As is true for any time-series fixed effects model identification relies on the assumption that there are no unobserved time-varying factors that affect both industrialization and democracy. To build confidence in this assumption I show that the results are robust to the inclusion of a wide-range of control variables (e.g. conflict, private property rights security, Cold War alignment, resource rents).

The results of both empirical strategies are very consistent and suggest that: (1) industrialization is strongly correlated with democracy in both the cross-sectional and temporal dimension; and (2) on the assumption that the geological process of coal formation is only related to democracy through industrialization this effect is causal. Running a large number of robustness checks I find that the effect is: (i) robust across measures of electoral (i.e. regular competitive elections) and liberal democracy (i.e. elections + rule of law); (ii) robust to dropping any particular world region from the sample (e.g. Western and East Asian countries); (iii) equally large on democratic transitions, as compared to democratic consolidations; (iv) is not contingent upon the nature of the international system at any point in time; and (v) is economically large in all three waves of democratization identified by Huntington (1991).

[5]Note, for example, that the United States and the United Kingdom, countries that today have a temperate climate, were located at the equator during the Carboniferous (hence explaining the abundance of coal deposits in these two countries) (Thomas, 2013).

[6]Given that the geological potential to have coal is determined long before my period of observation I am unable (at least without interacting with another meaningful, exogenous, and excludable variable) to exploit over time variation in my instrumental variable analysis.

The paper is related to the classic modernization theory of Lipset (1959). Lipset argued that economic development tends to lead to a wide range of socio-economic changes which in turn make the existence and stability of democracy more likely. Lipset did not, however, define clearly what he meant with "economic development". Instead he moved directly into operationalization by asserting that economic development can be best measured by a country' level of industrialization, urbanization, wealth, and education. He argued that: "[...] the factors of industrialization, urbanization, wealth, and education, are so closely interrelated as to form one common factor." (p. 80). While this empirical statement may have been true in 1959, it certainly is not true today. Today much of the developing world is rapidly urbanizing without industrializing (Gollin, Jedwab and Vollrath, 2016), today many countries are wealthy (at least on average) without being particularly industrialized, urbanized, or educated (e.g. Qatar, Macao, Brunei, the United Arab Emirates, and Kuwait, 5 of the 10 richest countries in the world, are not commonly thought of as particularly highly industrialized, urbanized, or educated), and some countries are relatively highly educated without being particularly wealthy, industrialized, or urbanized (e.g. average years of education is higher in Uzbekistan, Cuba, and Kazakhstan as compared to France, Belgium, and Austria).[7] Hence we have to think more closely which of Lipset's (1959) factors are important for democracy and how the factors themselves relate to each other (if at all). I contribute to this endeavor by arguing based on economic theory that it is likely to be industrialization that is the driving force behind the relationship between structural socio-economic change and democracy, and that some of the other factors that Lipset (1959) has highlighted are plausible causal mechanisms mediating the effect of industrialization on democracy (e.g. industrialization tends to induce "high-quality" urbanization and industrialization tends to increase a country' level of wealth while also making it more widespread, factors which in turn are likely to positively affect democracy).

The results are also related to two other bodies of literature. First, the literature on the "natural resource curse". Here the argument is that the production of oil, and possibly the extraction of natural resources more generally, negatively affects democracy because it allows states to receive income without relying on taxation (which may reduce demands for accountability), may provide autocratic governments with additional funds for repression, and may reduce the incentive to invest in education and occupational specialization (Ross, 2001, 2013). My argument is related to Ross' theory in the sense that we both agree that not all economic growth positively affects democracy, and that increases in economic production in the extractive sector are particularly unlikely to be conducive to democracy. My

[7]These comparisons are made based on GDP data from the 2018 World Bank Development Indicators and years of education data from the 2017 Human Development Index.

theory goes significantly further, however, by showing that also the agricultural and the (pre-industrial) service sectors, while significant contributors to GDP, are unlikely to trigger causal mechanisms that positively affect democracy.[8] Second, my argument is related to a smaller literature on the role of industrial workers in democratization movements. Here the main idea is that factory work lessens some of the collective action problems inherent in political mobilization, which may in turn make the introduction and sustainment of democracy more likely (e.g. Rueschemeyer, Stephens and Stephens (1992) and Collier (1999)). In line with this argument Dahlum, Knutsen and Wig (2019) find that protest movements dominated by industrial workers are significantly more likely to yield a democratic outcome in a sample of 147 countries over the 1900 to 2012 period. While my theory incorporates this particular causal mechanism, it is broader as the effect of industrialization on democracy is understood to be not only due to the organizational capacity of industrial workers, but also because of the effect industrialization has on a country' socio-economic structure as a whole (which affects everyone in a society, not only does directly engaged in industrial production).

The rest of the paper is structured as follows. In the next section I shortly discuss how the manufacturing sector differs from the agricultural, extractive, and (pre-industrial) service sectors, and how this explains why industrialization, but not economic growth in general, is likely to positively affect democracy. In the third section I describe the background and construction of the coal instrument. In the fourth section I describe the measurement of other key variables. I then report the instrumental variable and two-way fixed effects results in two separate sections. In the seventh section I examine the heterogeneity of the effect across time, development levels, and transitions to and consolidations of democracy. In the last section I conclude.

2 What is special about the manufacturing sector?

I argue that the manufacturing sector has a number of distinctive structural economic properties which makes industrialization likely to trigger a wide range of socio-economic changes which in turn tend to make the introduction and sustainment of democracy more likely.[9] I also argue that other economic sectors lack these type of structural economic properties, and are therefore unlikely to trigger political-economic processes that positively affect democ-

[8]This is in line with the existing literature on income and democracy which does not find that the effect of GDP per capita and democracy becomes robustly positive when controlling for oil rents (Broderstad, 2018).

[9]Naturally this is not to deny the existence of variation in structural economic properties within the manufacturing (and other) sector(s) (both across formal and informal manufacturing, and across different types of manufacturing). Nonetheless, this within-sector variation tends to be small relative to the between-sector variation (Kuznets, 1966; Rodrik, 2014).

racy. It are thus systematic differences in the nature of the production process that can explain why industrialization, but not economic growth in general, tends to be conducive to democracy. I present my argument by discussing the properties of each major economic sector in turn.

The *manufacturing sector* combines three clusters of economic properties which makes industrialization likely to trigger a wide-range of proximate causes of democracy. First, the manufacturing sector exhibits very high labor productivity, with a relatively large capacity to absorb labor, and strong linkages to the rest of the economy (Hirschman, 1959; Rodrik, 2012, 2014).[10] Growth in the manufacturing sector therefore tends to (directly or indirectly) lead to stark and sustained increases in population-wide living standards. Over time this creates the relatively well-off working and middle classes that have so often proven important for democracy (Ansell and Samuels, 2014; Collier, 1999; Rueschemeyer, Stephens and Stephens, 1992). By the same token it also reduces the stark economic inequalities between the (extreme) poor and rich, thereby structurally lowering the redistributional risk democracy poses for elites (Boix, 2003; Acemoglu and Robinson, 2006). Second, the returns to formal education are highest in manufacturing (Jones, 2001; Rodrik, 2014).[11] The returns to education and the relatively large capacity to absorb labor in manufacturing means that industrialization tends to strongly increase the incentive to invest in population-wide education, which in turn is an important determinant of democracy (Glaeser, Ponzetto and Shleifer, 2007; Evans and Rose, 2007; Dahlum, Wig and Knutsen, 2019; Zeira, 2019). Last, due to significant economies of scale, high transactions costs, and the possibility of a deep division of labor in manufacturing, industrialization typically necessitates the creation of large firms, or networks of firms, that rely on the cooperation of large groups of people concentrated in urban areas (Krugman, 1993; Marx, 1867). This in turn significantly lessens collective action problems and increases civil society' capacity to organize (e.g. labor unions), which in turn are important determinants of democracy (Dahlum, Knutsen and Wig, 2019; Collier, 1999; Bermeo and Nord, 2000).

The *agricultural sector*, in contrast, is an important contributor to GDP *but* is significantly less likely to induce the type of structural societal changes that positively democracy.[12]

[10]The labor absorbing capacity of manufacturing may be declining in the future due to automation and other labor-saving technologies (Rodrik, 2016). Although the jury is still out as to whether recent manufacturing technologies will indeed destroy more jobs than create (especially in developing countries) (Haraguchi, Cheng and Smeets, 2017), this trend is in any case too recent to substantively affect the argument being made here.

[11]This is because technology can be more widely applied in manufacturing (as compared to other sectors), which makes having knowledge and creating new knowledge more valuable in manufacturing.

[12]The fact that agriculture tends to have relatively low productivity (particularly in pre-industrial economies) should not be taken to imply that dynamics in the agricultural sector do not affect the level or growth of GDP per capita very much. Particularly in pre-industrial countries the agricultural sector

First, the agricultural sector has relatively few linkages to the rest of the economy so that economic growth stemming from increases in economic activity in the agricultural sector does not tend to increase living standards outside of those engaged in agriculture very much (Hirschman, 1959). Second, the returns to formal education are low in agriculture, so that agricultural growth does not strongly incentivize investments in education (Rodrik, 2014). Last, agricultural production, particularly in pre-industrial societies, tends to be organized in relatively small firms where only a small number of people (generally family members) work together. In addition, agricultural production is land-intensive and thus actively disincentives agglomeration/urbanization. As such the nature of work in agriculture is unlikely to increase civil society' capacity to organize/cooperate nearly as much as manufacturing does. Worse still at least some types of agriculture-induced growth is likely to *negatively* affect democracy. This is particularly the case when stark increases in GDP are due to the production of high-value agricultural products produced on farms owned by large landowners (e.g. Argentina in the beginning of the 20$^{\text{th}}$ century, plantation economies in 19$^{\text{th}}$ century North America). Such growth is likely to increase the political-economic power of large landowners, which have strong incentives to oppose democratization (Moore, 1966; Acemoglu and Robinson, 2006; Boix, 2003).[13]

As is widely appreciated in the political science literature the *extractive sector* is a very significant contributor to many country' GDP *but* such income derived from natural resource rents is in fact likely to *negatively* affect democracy (Tsui, 2010; Ross, 2013; Andersen and Ross, 2014; Ramsay, 2011; Ahmadov, 2014).[14][15] The extractive sector is an economic sector that combines very high-returns, with very little labor input, and very few linkages to the rest of the economy (Ross, 2001). These structural economic properties make that there is little incentive to invest in the organization and human capital of mine workers, and that the economic gains of natural resource extraction tend to predominantly accrue to the mine

tends to be a very large part of GDP so that even small increases/decreases in economic activity in the agricultural sector tend to strongly affect aggregate GDP (Gollin, 2010). Similarly, significant growth in the manufacturing sector often affects aggregate GDP in developing countries relatively little because the manufacturing sector typically represents only a small part of such economies (Rodrik, 2012).

[13]This is because the (immobile) assets of large landowners are easily taxed and expropriated under democracy, and because the income of large landowners has historically often relied on labor-repressive economic institutions that are difficult to sustain under democracy.

[14]According to data from the World Bank Development Indicators a staggering 51 countries received more than 10% of their GDP from natural resource rents in 2017. Another 22 countries received more than 20% of their GDP from natural resource rents in 2017.

[15]Nonetheless the "political resource curse" is certainly not uncontroversial. Haber and Menaldo (2011), Brooks and Kurtz (2016), and Liou and Musgrave (2014), among others, have found that natural resource extraction has no effect on democracy. Others find only a negative effect of natural resource extraction conditional on third factors (e.g. the quality of regulatory institutions). Importantly, however, few, if any, authors theoretically expect or empirically find an unambiguous positive effect of natural resource extraction on democracy, which is sufficient for the validity of the argument pursued here.

owners. As the owners of extraction firms tend to be the enfranchised elite, and mines represent highly immobile capital, expansion in the extractive sector typically significantly increases the redistributive risk democratization poses to autocratic elites (Boix, 2003). In addition, as mines in autocratic regimes are often state-owned extraction-induced growth typically significantly increase state revenues without relying on taxation, which in turn may lower popular demands for political representation and may well increase governments' funds for repression and buying-off the disenfranchised population (Ross, 2001; Sandbu, 2006).

The *service sector* is a highly heterogenous category, with significant variation across different types of services and across formal and informal service activities. Nonetheless the structural economic properties of the service sector in pre-industrial societies are quite consistent, and are unlikely to positively affect democracy. This is because the service sector in pre-industrial societies tends to be organized in small firms, producing simple services that require little education, cooperation, or technology to produce, and which have few input or output linkages to the rest of the economy (e.g. street sellers, domestic servants, cab drivers) (Rodrik, 2014). Such pre-industrial services are nonetheless highly consequential for the relationship between income and democracy because even in low-income countries services constitute more than 50% of GDP (Cali, Ellis and te Velde, 2008).

Clearly their do exist a relatively wide range of service occupations that do require significant investments in education, do tend to bring large groups of people together in modern, large, and typically urban-based firms, and do provide workers with a relatively high and secure standard of living. Importantly, however, such "modern service" activities only tend to arise *after* a country had already substantially industrialized (e.g. structural change towards high-skill/high-income service employment in OECD countries) (Herrendorf, Rogerson and Valentinyi, 2014). This is because many modern services are to a large extent inputs (e.g. banking, design, R&D, consultancy) or outputs (e.g. retail, marketing) of manufacturing, and thus only tend to arise *because* of widespread industrialization. In the absence of a large domestic manufacturing sector modern services tend to remain a small enclave of the economy enriching a small number of high-skilled service workers, leaving most of the rest of the economy untouched (e.g. the direct and indirect jobs generated by India's "high-tech" service sector employ at best 2% of the total Indian workforce) (Rodrik, 2014).

Taken together the above suggests that we should *not* expect that any increase in GDP per capita, regardless from which production process it stems, is equally likely to positively affect democracy. Instead it are only increases in economic activity in the manufacturing sector, in the form of industrialization, that are likely to trigger causal mechanisms that make democracy as a political equilibrium more likely. The rest of the paper aims to empir-

ically test the effect of industrialization on democracy. Given the already very considerable challenges that establishing this effect poses I leave it for future research to assess empirically the causal mechanisms through which this effect manifests itself. This implies that the validity of my empirical results do not dependent upon the three causal mechanisms described above being the only or most important factors linking industrialization with democracy. Similarly, the validity of my empirical results do not dependent upon non-manufacturing induced economic growth always having a negligible effect on democracy. There certainly may be exceptions to the general pattern I have sketched above (e.g. when revenues from commodity exports are effectively invested in public goods that increase population-wide living standards).[16]

3 Background Coal Instrument

Identifying the effect of industrialization on democracy is challenging as democracy may itself affect industrialization (i.e. reversed causality) and because other factors may simultaneously affect both democracy and industrialization (i.e. confounding). Clearly no researcher can hope to randomly assign industrialization in a randomized controlled trial. Identification therefore effectively relies on discovering an observable variable which has a quantitatively meaningful effect on industrialization but is not directly related to democracy (i.e. an instrumental variable).

In this paper I propose such an instrument in the form of a country' geological potential to have coal deposits on its own territory. The logic is here that a country' geological potential to have coal is likely to be strongly correlated with its' level of industrialization because cheap access to coal has historically been very important for industrialization, while coal itself is exogenously given by nature, and is unlikely to have affected democracy through channels other than industrialization. In the following two sections I describe the logic and construction of the instrument in more detail.

3.1 The importance, distribution, and tradability of coal

Access to coal was close to a necessary (but far from a sufficient) condition for the industrial revolutions of the 19$^{\text{th}}$ and early 20$^{\text{th}}$ centuries (Wrigley, 2010). Virtually all technologies of

[16]Exceptions to the general pattern can also exist in the other direction. This may, for example, be the case when state-planned industrialization in autocratic regimes creates monopsonistic industrial firms owned by local oligarchs connected to the political elite (e.g. Soviet industrialization – Lankina and Libman (2018)), or when other factors that negatively affect democracy swamp the effect of industrialization (e.g. pre-World War II Germany – Inglehart and Welzel (2009)).

the First Industrial Revolution (+/- 1760 to 1840) relied on coal as a key input in production (e.g. steam engine) and so did most of the technologies of the Second Industrial Revolution (+/- 1870 to 1913) (e.g. railways). Echoing this Vries (2001) argues that:

> "[...] the first industrial revolution, even in Britain, was not solely founded on coal. This applies even more to other countries like the United States and Japan and their industrialization. But still, the thesis can easily be defended that coal was necessary to prevent the growth from eventually petering out. Without it – and in that sense I think it indeed was absolutely fundamental and thereby crucial – energy bottlenecks undoubtedly would have arisen." (p. 425).[17]

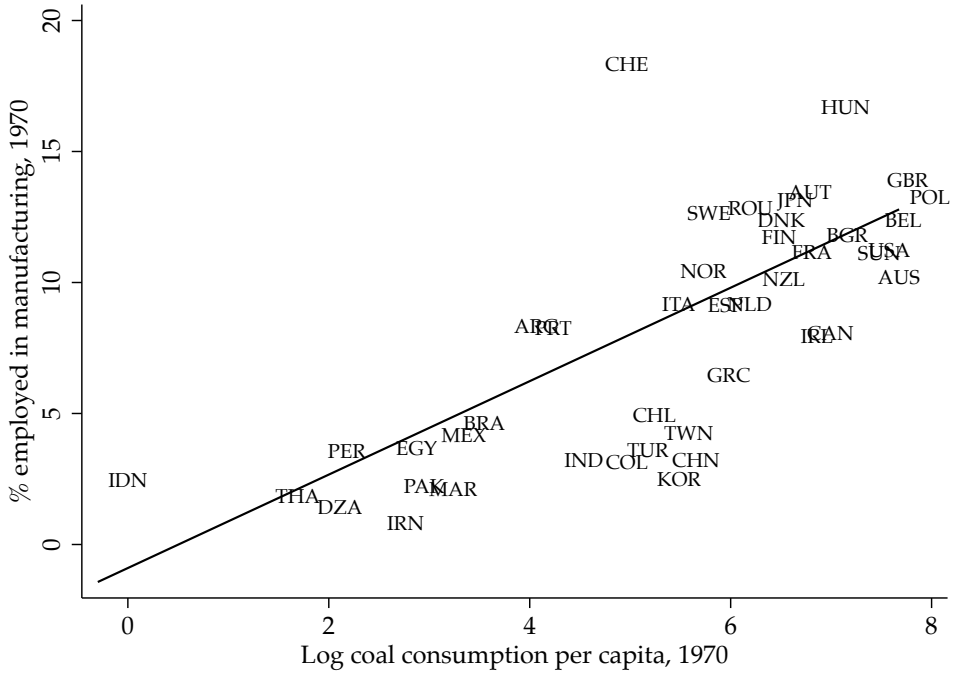
Although oil took over as the most important source of domestic and transport energy in the mid-1950s coal remained, at least until the 1970s, very important for industrialization. First, coal remained centrally important for the generation of electricity in the (extremely electricity-intensive) manufacturing sector, as it is cheaper to produce electricity from coal than from any other fossil fuel. Second, coal was before 1970, and still is, an important input in the production of steel (which in turn is necessary for many other industrial production processes). In 1970 no economically viable alternative for steel production existed, and even today more than 70% of global steel production is critically dependent on metallurgical/coking coal (Osborne, 2013). The importance of coal, even until as late as the 1970s, can be seen in figure 1 which plots coal consumption in 1970 (the first year for which systematic data is available) against the fraction of the total population employed in manufacturing in the same year. As can be seen from the scatter plot differences in per capita coal consumption explains approximately 54% of all cross-country differences in industrialization in 1970.

While coal was important for early industrialization *access* to coal was, at least before 1970, highly unequally distributed across countries. This is because: (1) coal deposits are very unequally distributed across countries. In fact, the country with most coal reserves (i.e. the United States) has more reserves than the bottom 180 countries *put together*, and 138 countries have no proven coal reserves at all (BP, 2017); and (2) coal was limitedly tradable before 1970 due to extremely high transportation costs. Coal is because of its extremely heavy weight/large volume relative to its per kg/m^3 monetary value one of the least tradable goods in the world (Bairoch, 1995; Hummels, 2007; Wårell, 2006). It is for this reason that coal was before the worldwide introduction of containerized shipping in the 1970s hardly internationally tradable (Hummels, 2007).[18] Coal became more internationally

[17] See for other work on the importance of coal for early industrialization: Wrigley (2010), Crafts and Mulatu (2006), and Fernihough and O'Rourke (2014).

[18] Wårell (2006) analyzes price differentials across national coal markets to assess the degree of world coal

Figure 1: Coal consumption and industrialization in 1970.



Note: Coal consumption per capita is the natural log of coal consumption (in kilograms) per person. Data comes from the BP (2017) Energy Consumption and Production Database. % in manufacturing is the percentage of the total population that is employed in the manufacturing sector. Manufacturing employment data comes from Mitchell (2013), the United Nations Industrial Development Organization, and the 10-Sector Database of the Groningen Growth and Development Centre. Population data comes from Bolt et al. (2018). The coefficient and R^2 of the linear regression line are 1.78 and 0.54, respectively.

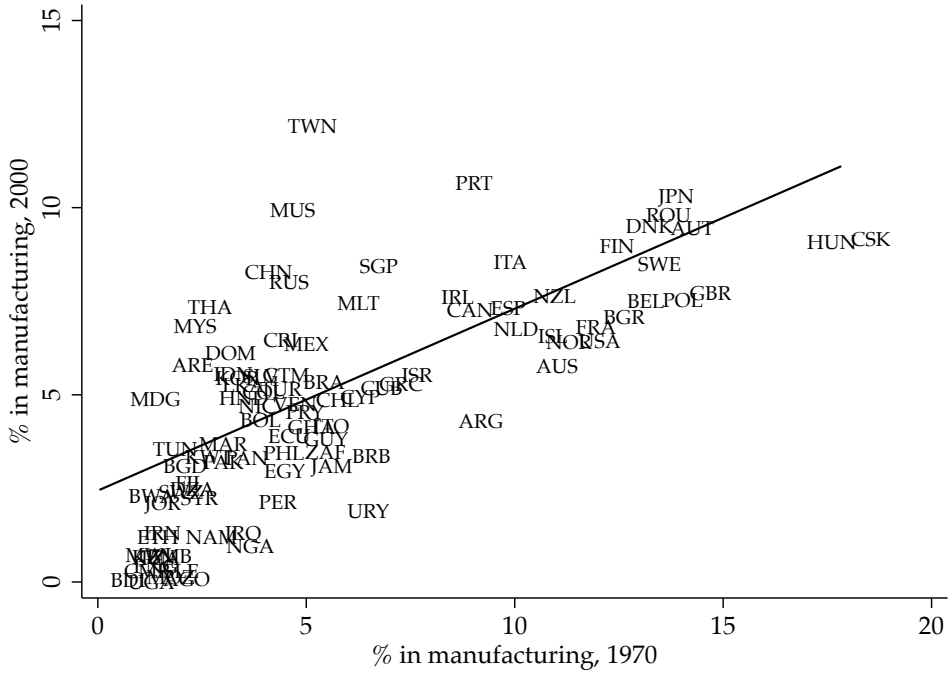
tradable after 1970, but even today less than 18% of all coal is traded internationally (WCA, 2019), and even today transportation costs *within* the United States are typically higher than the value of the transported coal itself (EIA, 2018).

Given that coal was very important for pre-1970 industrial production, and given that coal was limitedly tradable before the 1970s, differences in domestic access to coal are likely to have induced significant differences in pre-1970 industrialization levels. Since then these coal-induced differences in industrialization are likely to have largely persisted through time because of strong path dependency in cross-country industrialization patterns – i.e. countries that were highly (little) industrialized before 1970 have tended to remain highly (little) industrialized today (Acemoglu, 2009).[19] Figure 2 illustrates this well-known pattern of path

market integration and comes to the conclusion that: "In the 1960s the coal market could not be considered as international." (p. 99)

[19]Industrialization is likely to be subject to path dependency due to learning-by-doing and capital/skill "biased" technical change in manufacturing (Allen, 2011). My identification strategy is, however, not contingent upon the precise reason for path dependency in industrialization.

Figure 2: Industrialization in 1970 and 2000.



Note: % in manufacturing is the percentage of the total population that is employed in the manufacturing sector. Manufacturing employment data comes from Mitchell (2013), the United Nations Industrial Development Organization, and the 10-Sector Database of the Groningen Growth and Development Centre. Population data comes from Bolt et al. (2018). The coefficient and R^2 of the linear regression line are 0.49 and 0.53, respectively.

dependency in industrialization. As can be seen from figure 2 differences in industrialization in 1970 explain more than 50% of the differences in industrialization in 2000.

3.2 The geological potential to have coal as an instrument

Naturally it would be an overstatement to say that countries without domestic access to coal would before 1970 have been unable to industrialize at all (e.g. Switzerland became a highly industrialized society with very few proven coal deposits), or that no country managed to transform itself into a highly industrialized society after 1970 (e.g. South Korea's and Taiwan's industrialization mostly occurred after 1970). It is fair to say, however, that no domestic access to coal would, everything else equal, have substantially increased the opportunity cost of pre-1970 industrialization, and that at least some of these coal-induced differences in industrialization have persisted until today. It is this intensive margin that I exploit in my instrumental variable analysis.

To do this, best would be to construct a variable measuring the total amount of coal that

was ever in the ground of a country' territory. This, however, is not possible because: (1) even today it is extremely likely that some coal deposits are yet to be discovered (Thomas, 2013); and (2) there is no systematic data available on the amount of coal that countries have extracted/depleted in the past.

A second-best option could be to simply use a country' proven coal reserves today, or at some other point in time, as a rough proxy for the total amount of coal that was ever provided by nature. This, however, risks introducing endogeneity bias because industrialization and democratization may have affected the incentive and/or capacity to discover coal, and because past coal extraction/depletion is likely to be systematically correlated with (past) industrialization introducing systematic measurement error which biases both ordinary least squares (OLS) and two-stage least squares (2SLS) estimates (Wooldridge, 2013).

To avoid introducing these sources of bias I therefore instrument industrialization with the geological *potential* to have coal. Although a variable measuring a country' geological potential to have domestic coal deposits is likely to be less strongly correlated with industrialization (as compared to the hypothetical case of a variable which perfectly captures a country' total coal reserves ever provided by nature), such an instrument is clearly exogenous, because coal is formed as a result of a geological, not a man-made, process, and is plausibly excludable, because the geological process of coal formation is unlikely to have affected other determinants of democracy.

Finding a suitable instrument for industrialization thus comes down to creating one or several variables capturing well a country' geological potential to have coal. Luckily it is possible to construct such a variable because geologists have quite a good understanding of how coal is formed.[20] Fundamentally coal is formed when: (1) a large number of plants; (2) are quickly buried by sand, water, or any other relatively airtight material (to avoid dissolvement under the influence of bacteria); and (3) pressure and heat are applied over long periods of time (to dehydrate the coal, which increases its carbon concentration) (Thomas, 2013; Klein and Philpotts, 2013).

This three-step process has occurred successfully in several different places and times in earth' history. By far most of the earth' coal was, however, formed during one particular geological period: the upper/late Carboniferous era (approximately 323.2 to 298.8 million years ago).[21] Most coal stems from the Carboniferous period because of all geological periods in earth' history the conditions for coal formation were by far the most conducive during this particular time. First, temperatures were relatively high and the climate was relative

[20]I would here like to acknowledge the help of Dr. Elizabeth Harper (Department of Earth Sciences, Cambridge University).

[21]The upper Carboniferous period is known as the "Pennsylvanian period" in the United States.

wet. This meant that a tropical zone existed throughout this period; which is important for coal formation as significantly more plants grow in tropical environments. Second, the temperature varied significantly over the Carboniferous period and a large ice mass, called Gondwana, existed near the South pole. Due to increases (decreases) in temperate big parts of Gondwana melted (freezed) over relatively short periods of time. This caused major and frequent shifts in sea levels, which in turn caused the shoreline in the Carboniferous tropical zone to regularly transgress and regress. This constant flooding of tropical land enabled relatively large amounts of plants to get rapidly buried beneath deltaic sands. It are these area's in the Carboniferous tropics which over time produced most of the coal that we know today (Klein and Philpotts, 2013).

The importance of the Carboniferous tropical zone is well-illustrated by figure 3 which plots the location of proven coal deposits (green dots) against the outline of current countries (solid lines) and climatical zones (dotted lines) during the upper/late Carboniferous period. With the exception of several coal deposits in the cool temperature zones in the Northern- and Southern hemispheres virtually all coal during the Carboniferous was formed in the tropical zone.[22]

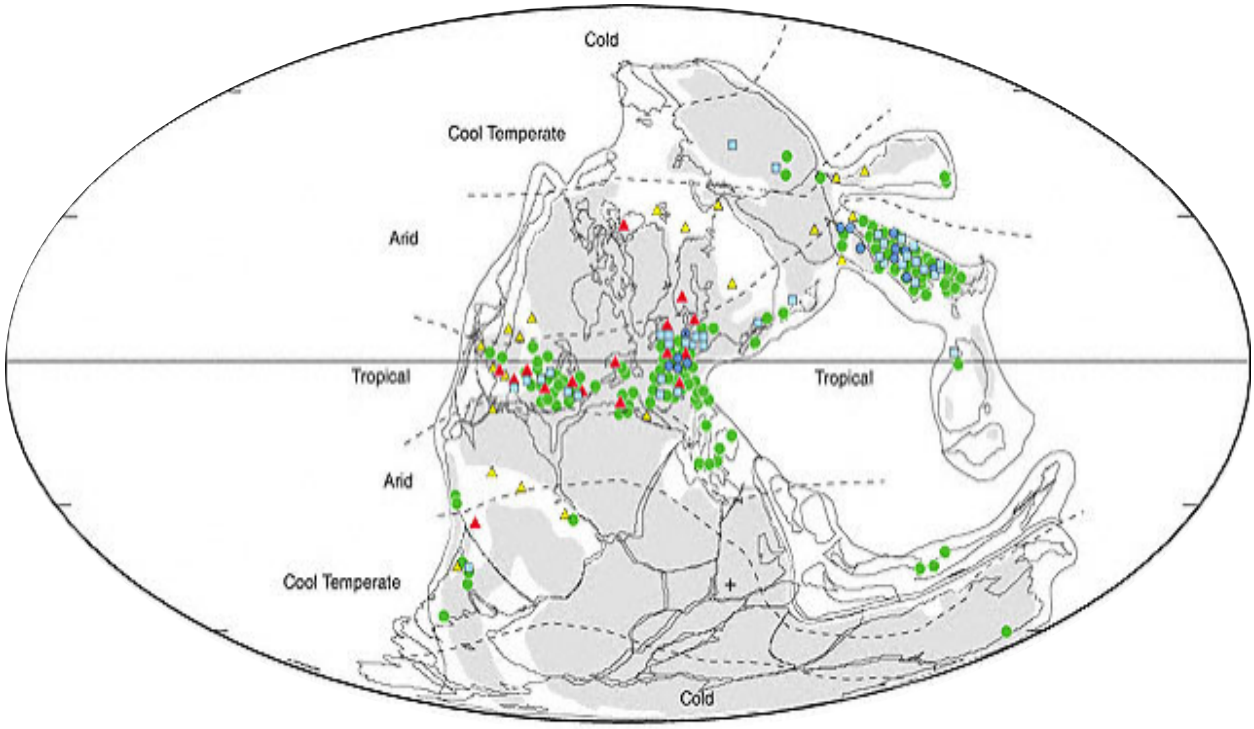
The above implies that I can construct a variable capturing a relatively large part of a country' geological potential to have coal by creating a dichotomous variable coding whether it was located in the tropical zone during the Carboniferous era, or not.[23] To do so I make use of data from Christopher Scotese (emeritus professor in Geology at Northwestern University) who provides geocoded data on climate zones and the location of current countries going back in geological time. All data is freely available at: www.globalgeology.com.

Clearly any geological model can only provide a reasonable approximation of the climatical conditions of current countries more than 300 million years ago. Importantly for my purpose, however, any measurement error in Scotese' so-called "continental drift" model is unlikely to be systematic with regard to industrialization and democracy. This is because Scotese derives the geo-location of climate zones and current countries during the Carbonif-

[22]The green dots in the cool temperature zone in the Northern hemisphere is the coal in Siberia. The green dots on the bottom left-side is the coal formation in South Africa. The green dots in the bottom right-side is the coal in Australia and India. I cannot predict the coal formation in these particular countries based on my simple instrument. This is because the coal formation in these countries took place under different circumstances (Thomas, 2013). As the first-stage relationship between the instrument and industrialization is strong regardless of this fact, and because the coal formation in these other countries can to the best of my knowledge not be captured by a simple set of additional instruments, I accept this as noise.

[23]Note that I only exploit the extensive margin of tropical climate during the Carboniferous because focusing on the percentage of a current country' territory located in the Carboniferous tropics would substantively misrepresent the cross-country distribution, which is almost entirely dichotomous in nature (see table 1). Focusing on the intensive margin would thus lead to a censored instrumental variable that for the far majority of countries would take the value 0 or 100.

Figure 3: Climate zones and coal formation during the late Carboniferous era.



Note: The figure shows the location of current countries (solid lines), the location of tropical zones (dashed lines), and the location of proven coal deposits (green dots) during the Upper Carboniferous. Data is for the Gzelian period, which is the latest subperiod within the Carboniferous. This geological period lasted from approximately 303.7 to approximately 298.9 million years ago. The blue dots, red dots, and yellow triangles are not relevant for the purpose at hand but reflect deposits of bauxite, calcrete, and evaporite, respectively. *Source*: `http://www.scotese.com/gzelclim.htm` (last assessed: March 20, 2019).

erous from global tectonic dynamics, fundamental geological principles, and some basic geological field research which has been done in virtually all world regions (Scotese, 2004). The estimates are therefore unlikely to be more or less precise for more industrialized and/or more democratic countries in which geologists may have done more field research. Even less likely is it that the measurement error systematically places more (less) industrialized countries inside (outside) the Carboniferous tropics, which is the only type of measurement error that would lead to bias, as opposed to attenuation, in my first-stage estimates (Wooldridge, 2013).

Given that the frontier of the tropical zone during the Carboniferous has shifted slightly over time as a result of temperature changes over millions of years I code countries as 1 if more than half of a current country' territory was according to Scotese' latest "snapshot" during the Carboniferous located inside the tropical zone, and 0 otherwise. In table 1 I list which countries were entirely, more than half, less than half, and not at all located in the

Carboniferous tropics at this particular point in time. All results hold and are substantively similar when including only countries which are entirely located in the Carboniferous tropics and when including all countries that are somewhat located in the Carboniferous tropics. The effect of industrialization on democracy is, if anything, slightly larger when using these alternative operationalizations.

As can be seen in table 1 the instrument is relatively strongly correlated within continents. This is because most of the current continents were already largely formed during the Carboniferous. To avoid picking up average differences in industrialization and democracy across continents I add in all my specifications fixed effects for America, Europe, Asia, and Africa. The results are, if anything, slightly larger when omitting these fixed effects.

Table 1: Degree to which modern day countries were located in the Carboniferous tropics.

Entirely	More than half	Less than half	Not at all	
Albania	China	Algeria	Angola	Mali
Austria	Norway	Canada	Argentina	Mozambique
Belgium	Sweden	Colombia	Australia	Namibia
Bulgaria	United States	Ecuador	Bangladesh	Nepal
Costa Rica		Egypt	Benin	New Zealand
Denmark		Finland	Bolivia	Niger
El Salvador		Iraq	Botswana	Nigeria
Estonia		Libya	Brazil	Pakistan
France		Mauritania	Burkina Faso	Paraguay
Germany		Russia	Burundi	Peru
Greece		Venezuela	Cameroon	Republic of the Congo
Guatemala			Central African Republic	Rwanda
Honduras			Chad	Senegal
Hungary			Chile	Sierra Leone
Ireland			Dem. Republic of Congo	Singapore
Israel			Djibouti	Somalia
Italy			Equatorial Guinea	South Africa
Jordan			Eritrea	South Sudan
Latvia			Ethiopia	Sri Lanka
Lebanon			Gabon	Sudan
Luxembourg			Ghana	Suriname
Mexico			Guinea	Swaziland
Morocco			Guinea-Bissau	Tanzania
Netherlands			Guyana	Thailand
Nicaragua			Hong Kong	The Gambia
Panama			India	Togo
Poland			Iran	Tunisia
Portugal			Ivory Coast	Uganda
Romania			Kenya	Uruguay
South Korea			Lesotho	Venezuela
Spain			Liberia	Zambia
Switzerland			Malawi	Zimbabwe
Syria			Malaysia	
Turkey				
United Kingdom				
West Germany				

Source: www.globalgeology.com (last accessed: April 24, 2019).

4 Measurement of democracy and industrialization

To measure a country' level of industrialization over time I collect annual time series data on the percentage of the population that is employed in the manufacturing sector from 1845 (i.e. the first year for which data is available) until 2015 (i.e. the last year for which data is available). I choose to focus on manufacturing employment, rather than manufacturing value added (another common measure of industrialization), for two reasons. First, and most importantly, the causal mechanisms described in the theory section are primarily about how many people in a society are affected by processes occurring in the manufacturing sector, rather than the productivity of a country' manufacturing sector as such (as would be better captured by manufacturing value added). Since manufacturing value added is strongly determined by labor-saving technologies increases in manufacturing value added tend to lead, holding everything else constant, to fewer people being engaged in the manufacturing sector (Rowthorn and Wells, 1987). Second, data on manufacturing value added is at best available since the year 1950, and for many countries only since 1970 or later. Measuring industrialization through manufacturing value added would thus severely limit the scope of the study, as well as the identifying variation available in the data to estimate the relationship between industrialization and democratization exploiting only the variation within countries over time (which naturally provides many advantages with regard to dealing with endogeneity).

To generate data on persons employed in the manufacturing sector I make use of three independent data sources: (1) Mitchell (2013), which provides census data for a large number of countries over the 1845 to 2008 period; (2) the 10-Sector Database of the Groningen Growth and Development Centre (GGDC), which provides data on persons employed in the manufacturing sector for 42 countries from 1950 to 2015. This data is constructed based on a country-to-country search for available national and international data sources[24]; and (3) the Industrial Statistics Database of the United Nations Industrial Development Organization (UNIDO), which provides data on employment in manufacturing for 170 countries from 1963 to 2015. Data for OECD countries is collected and provided by the OECD. Data for non-OECD countries is collected by surveying national statistical authorities using country questionnaires.

Estimates from all three sources are highly correlated as can be seen from the correlation matrix in table 2. Importantly, however, the UNIDO estimates for any given country-year tend to be slightly lower than the estimates from Mitchell (2013) and GGDC. This is because of the fact that Mitchell and GGCD classify workers as working in the manufacturing sector

[24] `https://www.rug.nl/ggdc/docs/10sector_sm_jan2015.pdf` describes the data sources and methods used by the GGDC to generate the data.

Table 2: Correlation between persons employed in manufacturing data from different sources.

	UNIDO	GGDC	Mitchell
UNIDO	1		
GGDC	0.942	1	
Mitchell	0.907	0.992	1

Note: Pearson r coefficients are reported. Data is included if variable is non-missing for a pair of sources in the same year.

on the basis of the nature of their job, while UNIDO classifies workers based on whether they work for a company that is predominantly engaged in manufacturing. As these differences are very small compared to the major cross-country differences in industrialization these slight differences in classification do not affect the cross-sectional instrumental variable estimates very much, and do not seem to warrant the exclusion of 48 cases that not including the UNIDO data would imply. As changes in industrialization over 5 to 10 year time periods are significantly smaller panel models with country fixed effects are more strongly affected by these classification differences. To avoid sudden jumps in the time-series due to changes in the data source, rather than real changes in a country' number of persons employed in manufacturing, I therefore use in all cases one manufacturing employment data source per country.[25] To choose which industrialization data to use for any individual country I simply select for each country the data source that maximizes the number of years for which data for that individual country is available. If necessary I linearly interpolate (but never extrapolate) the persons employed in manufacturing data.[26]

To account for differences in country size I calculate for each country-year the percentage of the total population that is employed in the manufacturing sector.[27] To do so I use data on population size from Bolt et al. (2018), Fink-Jensen (2015), and Haber and Menaldo (2011). Point estimates from these different data sources are virtually identical for most country-years and the data is always correlated with a Pearson r exceeding 0.99. Hence I combine the population data using simple imputation.

My baseline measure of democracy is the combined polity index of Polity IV. This variable is generated by subtracting for each country-year the institutionalized autocracy score

[25]Note that all results hold when creating the most extensive panel possible by imputing (in any order) the UNIDO, GGCD, and Mitchell (2013) data, and also when restricting the analysis to the use of only one data source for all countries.

[26]The UNIDO and GGCD data is in virtually all cases available on a yearly basis. The Mitchell (2013) data, being drawn from official censuses, is for most countries available on a 10 yearly basis (although the time intervals vary widely between countries and over time).

[27]Perhaps slightly more precise would have been to normalize with adult population. This data is, however, not systematically available for many countries before 1950.

from the institutionalized democracy score.[28] The index consists of five components: (1) the competitiveness of executive recruitment; (2) the openness of executive recruitment; (3) the degree of institutional constraints on the chief executive; (4) the regulation of political participation; and (5) the competitiveness of political participation.[29] To aid the interpretation of substantive effects I recode the polity index to range from 0 to 10. I use this measure as it is most widely used in long-run studies of democracy.

As a robustness check I replicate the results with the dichotomous (electoral) democracy measure of Boix, Miller and Rosato (2013). Countries are coded as 1 if they hold free and fair elections and have enfranchised the majority of the male population, and 0 otherwise.[30] This measure is chosen for three reasons. First, it is based on a relatively object criterion rather than on expert surveys which may be subject to cross-cultural bias (Silva and Littvay, 2019). Second, it does not assume that democracy is a concept that can be measured on the interval/ratio level (Munck and Verkuilen, 2002). Third, it allows to test whether the effect of industrialization differs across the electoral and liberal dimensions of the concept of democracy (Knutsen et al., 2019).

I include countries in my baseline sample if they (arguably) have the sovereignty to decide upon their own regime type and data is available. This first condition excludes all ex-colonies during their respective periods of colonialization, East-European countries during the Cold War, Hong Kong, and Macao. All results nonetheless hold when including these cases. Table 12 in the appendix lists the countries and years included in the analyses. Also in the appendix are the measurement and descriptive statistics of all (other) variables used in the analysis below.

5 Instrumental variable estimates

Before estimating the effect of industrialization on democracy using my instrumental variable approach I study the relationship between the tropical in the Carboniferous instrument and coal consumption (i.e. intention-to-treat), industrialization (i.e. first stage), and democracy (i.e. reduced form). The results are reported in table 3.

As can be seen in column (2) countries that were in the tropics during the late Carboniferous era, that is more than 300 million years ago, consumed 3.5 times (1.305*2.7) more coal per capita in 1970 (the first year for which systematic data is available), as compared to

[28] To facilitate time series analyses Polity recodes foreign "interruptions" as missing, cases of "interregnum" as a neutral polity score of 0, and prorates cases of "transition" across the span of the transition.

[29] See Marshall, Gurr and Jaggers (2016) for more information.

[30] This last condition, which is significantly less stringent than a threshold of full enfranchisement of both men and women, is chosen to capture the considerable variation in regime type before World War I.

Table 3: Intention-to-treat, first stage and reduced form.

	Log coal consumption per capita, 1970		% in manufacturing, 1990-2000		Democracy, 1990-2000	
	(1)	(2)	(3)	(4)	(5)	(6)
Tropical in Carboniferous	1.341**	1.305*	3.048***	2.049***	1.042*	-0.049
	(0.658)	(0.669)	(0.742)	(0.646)	(0.596)	(0.578)
Log GDP per capita		0.125		1.516***		
		(0.242)		(0.203)		
% in manufacturing						0.358***
						(0.097)
Continent fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.469	0.458	0.486	0.661	0.347	0.431
Observations	40	40	97	97	95	95

Notes: OLS regressions. All countries for which data is available enter the sample in columns (1) to (4). Countries which were independent throughout the 1990 to 2000 period and for which data is available enter the sample in columns (5) and (6). Dependent variable in column heading. Coal consumption per capita is the natural log of coal consumption (in kilograms) per person. Data comes from the BP (2017) Energy Consumption and Production Database. Democracy is measured by the combined polity index of Polity IV (rescaled to range from 0 to 10). % in manufacturing is the percentage of the total population that is employed in the manufacturing sector. Manufacturing employment data comes from Mitchell (2013), the United Nations Industrial Development Organization, and the 10-Sector Database of the Groningen Growth and Development Centre. Population data comes from Bolt et al. (2018). Tropical in Carboniferous is a dummy that takes the value 1 if a current country was located in the tropical zone during the late Carboniferous era (+/- 300 million years ago), and 0 otherwise. This variable is generated by the author using data from: `globalgeology.com`. All regressors are measured in the same period as the dependent variable. Dummies for American, European, Asian, and African countries are included in all regressions. Robust standard errors in parentheses. All regressions contain a constant which is not reported.

$***$ p<0.01, $**$ p<0.05, $*$ p<0.10.

countries that were not subject to a tropical climate during the Carboniferous, even when holding the level of GDP per capita constant.[31] This means that even when comparing countries with arguably equal demand for coal that those countries that were located inside (outside) the Carboniferous tropics consumed significantly more (less) coal per capita in 1970. The assumption underlying my instrumental variable approach is that this is because having been located in the Carboniferous tropics significantly increases the probability of the existence of coal deposits on a country' own territory.

Columns (3) and (4) show the important effect tropical climate during the Carboniferous has had on the average level of industrialization from 1990 to 2000.[32] Simply the difference

[31]Note the significance of this result as GDP per capita is likely to mediate some of the effect of tropical during Carboniferous on coal consumption in 1970 (i.e. because more industrialization made possible by better/cheaper access to coal is likely to have substantially increased GDP per capita before 1970, which in turn is likely to incentivize countries to consume more coal, even when holding the geological potential to have coal constant).

[32]In my main specifications I focus on the average level of industrialization from 1990 to 2000, as this

between being inside or outside the tropical zone more than 300 million years ago explains approximately 21.8% of the cross-country variation in industrialization levels in the 1990s (i.e. net of the continent fixed effecs). In terms of effect size the estimate in column (4) suggests that employment in the manufacturing sector in the 1990s was 2.1% higher (lower) in countries that were located inside (outside) the tropics during the Carboniferous, even when holding GDP per capita constant.[33] The assumption underlying my instrumental variable approach is that this is due to countries that were located inside (outside) the Carboniferous tropics having significantly greater (lesser) access to (cheap) coal, which has historically significantly lowered the opportunity cost of industrialization.

Column (5) of table 3 shows the reduced form results. Here I find that countries that were located inside (outside) the tropics during the Carboniferous score 1 entire scale point higher (lower) on the combined polity index (rescaled to range from 0 to 10). The exclusion restriction underlying my instrumental variable approach requires that this effect is entirely due to countries that were located inside (outside) the Carboniferous tropics being more (less) industrialized. Column (6) of table 4 provides suggestive evidence for this claim as the effect of tropical in the Carboniferous on democracy becomes negative, close to zero, and highly statistically insignificant after controlling for industrialization.

In table 4 I report OLS and 2SLS estimates of the effect of industrialization on democracy. The OLS coefficient in panel C column (1) indicates that a one percentage point increase in industrialization is estimated to lead to an average 0.355 scale points increase in democracy (on an 11-point scale). This is a substantial effect. It suggests, for example, that if a country would progress from a largely agricultural society of say less than 1% of the population employed in the manufacturing sector (e.g. Afghanistan, Angola, Eritrea, Kenya, Kyrgyzstan, Yemen, Zambia in 2010) to a highly industrialized society of approximately 11% employed in manufacturing (e.g. approximately the peak-level of industrialization in OECD countries) that its level of democracy would increase by slightly more than 32% of the full polity IV scale.

There are, however, two important reasons to not interpret OLS estimates of the relationship between industrialization and democracy as causal. First, more democratic countries may adopt better policies or provide a better institutional context for industrialization, leading to a reversed causality problem. Second, there may be other, perhaps currently unknown, determinants of democracy that may also cause countries to have different levels of industrialization, creating a confounding problem.

maximizes the sample size and instrument strength. As I show in table 5 the results are similar in earlier and later periods.

[33]Given that the average percentage of the population employed in manufacturing is only 4.7% in this sample this is an economically substantial effect.

Table 4: OLS and 2SLS estimates.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Democracy (second stage)						
% in manufacturing	0.342**	0.366**	0.328**	0.290*	0.347**	0.374**
	(0.164)	(0.185)	(0.146)	(0.173)	(0.158)	(0.149)
Tropical today		0.003				
		(0.007)				
Soil quality			-0.018			
			(0.017)			
Oil and gas rents (% of GDP)				-0.168***		
				(0.061)		
Democracy, 1800					-0.025	
					(0.160)	
% of years conflict, 1750-2000						-0.025*
						(0.014)
Panel B: % in manufacturing (first stage)						
Tropical in Carboniferous	3.045***	2.760***	3.429***	3.041***	3.031***	3.081***
	(0.725)	(0.688)	(0.685)	(0.806)	(0.768)	(0.730)
Montiel-Pflueger F statistic	17.6	16.1	25.1	14.2	15.6	17.8
Observations	95	95	94	91	95	95
Panel C: Democracy (OLS estimates of equivalent model)						
% in manufacturing	0.355***	0.367***	0.387***	0.321***	0.356***	0.352***
	(0.087)	(0.098)	(0.085)	(0.089)	(0.087)	(0.086)
Continent fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.438	0.433	0.455	0.421	0.431	0.456
Observations	95	95	94	91	95	95

Notes: OLS and 2SLS regressions. Countries enter the sample if they were independent throughout the 1990 to 2000 period and data is available. The dependent variable is the level of democracy in 2000. This variable is measured by the combined polity index of Polity IV (rescaled to range from 0 to 10). % in manufacturing is the percentage of the total population that is employed in the manufacturing sector. Manufacturing employment data comes from Mitchell (2013), the United Nations Industrial Development Organization, and the 10-Sector Database of the Groningen Growth and Development Centre. Population data comes from Bolt et al. (2018). Tropical in Carboniferous is a dummy that takes the value 1 if a current country was located in the tropical zone during the late Carboniferous era (+/- 300 million years ago), and 0 otherwise. This variable is generated by the author using data from: `globalgeology.com`. All regressors are averaged over the 1990 to 2000 period. See the appendix for the measurement of control variables. Dummies for American, European, Asian, and African countries are included in all regressions. Robust standard errors in parentheses. All regressions contain a constant which is not reported. Panel B and C always contain the same control variables as panel A, but these are not reported.

$***$ p<0.01, $**$ p<0.05, $*$ p<0.10.

To alleviate both concerns I instrument a country' average level of industrialization in the 1990s with whether it was subject to a tropical climate 300 million years ago during the Carboniferous era (which strongly captures a country' geological potential to have domestic access to coal). Panel A of table 4 reports the second stage estimates of interest and panel B displays the corresponding first stages.

The first-stage F statistics in panel B column (1) is 17.6, which reconfirms the strong relationship between tropical climate in the Carboniferous and industrialization (as previously shown in table 3). I regard this large F statistic as indicating that weak instrumentation bias is unlikely to be a concern in my analysis. The corresponding 2SLS estimate of the effect of industrialization on democracy is 0.342; which indeed is almost identical to the previously mentioned OLS estimate of 0.355. The 2SLS estimate thus suggests that simple OLS estimation *does not* overstate the effect of industrialization on democracy.

The validity of the instrumental variable estimates in table 4 depends on the assumption that tropical climate approximately 300 million years ago has no effect on democracy besides its effect through industrialization. Arguably the main concern with this assumption is that climate during the Carboniferous is somehow related to current climate and that the estimates are therefore unintentionally picking up the effect of contemporary tropical climate on industrialization and democracy. I regard this scenario as highly unlikely as the (latitude-longitude) location of countries today has very little, if anything, to do with the location of these same landmasses 300 million years ago during the Carboniferous era (because of continental drift). During the Carboniferous era, for example, the United States, the United Kingdom, and most of continental Europe were located at the equator (explaining the large amount of coal deposits in these regions), while most of Africa and South America was meanwhile located at the South Pole (Scotese, 2004). To nonetheless account for this possibility I control for the percentage of a country' territory that is currently subject to a tropical climate (data comes from Nunn and Puga (2012)). As can be seen in column (2) of table 4 this leaves the results unchanged.[34]

If not through climate today, tropical climate during the Carboniferous may have affected the soil quality/type of current countries which may have had an independent effect on industrialization and democracy at the end of the 20^{th} century. This scenario is also highly unlikely, however, as soil quality/type is determined by much more recent climate, as well as simply the surface material in place (which are both plausibly uncorrelated with climate during the Carboniferous). To nonetheless control for this possibility I use data from John Gallup on the percentage of a country' total land area that is very or moderately suitable

[34]The same applies when controlling for latitude (both linearly and quadratically).

to produce rainfed crops.[35] As shown in column (3) of table 4 including this variable also leaves the results unchanged.

A third possibility is that the geological determinants of coal are also correlated with other fossil fuels (i.e. oil and natural gas), which in turn may have an independent effect on democracy. Again, this scenario is highly unlikely as oil and gas are formed by dead sea animals (as opposed to dead plants) being rapidly buried by airtight material, and subsequently being subjected to heat and pressure over long periods of time. The process of oil and gas formation is thus fully independent of the process of coal formation, and has nothing to do with the tropical zone during the Carboniferous era (Klein and Philpotts, 2013). Not surprisingly therefore column (4) of table 4 shows that the results are unaffected by controlling for the % of GDP derived from oil and natural gas rents (data comes from Haber and Menaldo (2011)).

Another potential violation of the exclusion restriction comes from the fact that coal has been used as a source of fuel for hundreds of years prior to the Industrial Revolution. At least theoretically it could therefore be the case that better access to (cheap) coal has caused institutional differences before modern industrialization took place and that this is (partly) causing countries that were located inside (outside) the Carboniferous tropics to be more (less) democratic in the 1990s. I take this possibility into account by controlling for the combined polity index in the year 1800 (i.e. the first year for which data is available).[36] As can be seen in column (5) of table 4 this leaves the results unchanged.

A last concern is that of inter and/or intrastate conflict over coal resources. Although I am unaware of any qualitative-historical evidence that points in this direction it could theoretically be true that countries with better access to coal have due to this been subject to more conflict, which in turn may have had an independent (positive or negative) effect on democracy. To take this into account I control for the percentage of years from 1750 to 2000 in which a country was engaged in inter- or intrastate conflict, relative to the number of years a country was independent and data is available (data comes from Brecke (2001)). As can be seen from column (7) in table 4 controlling for this factor also leaves the results unchanged.

In table 5 I examine the robustness of these results. I here find that the results hold, and are quantitatively similar, when: (1) excluding all Western countries (incl. Greece); (2) excluding all East-Asian "miracle economies" (i.e. Japan, South Korea, Taiwan, and

[35]Data comes from: `https://www.pdx.edu/econ/jlgallup/country-geodata` (last assessed: March 9, 2019).

[36]As is standard in the literature I code all countries that were colonies in 1800 as 0. Without this intervention the 2SLS estimates are quantitatively similar (second-stage coefficient: 0.199) but the sample reduces to only 14 countries.

Singapore); (3) excluding all Middle Eastern countries (excl. Israel); (4) focusing on the years 1990 and 2010, rather than 2000; and (5) using the Boix, Miller and Rosato's (2013) dichotomous electoral democracy measure as an alternative operationalization of democracy.

Taken together the cross-sectional estimates show that: (1) the average level of industrialization over the 1990 to 2000 period is strongly and robustly correlated with the level of democracy in 2000; and (2) on the assumption that the geological process of coal formation is not related to democracy besides affecting industrialization this effect is causal.

Table 5: Robustness checks OLS and 2SLS estimates.

	Excluding West	Excluding East Asia	Excluding Middle East	Polity in 1990	Polity in 2010	Boix democracy
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Democracy (second stage)						
% in manufacturing	0.306*	0.285*	0.368**	0.312**	0.474*	0.092***
	(0.176)	(0.156)	(0.153)	(0.149)	(0.279)	(0.031)
Panel B: % in manufacturing (first stage)						
Tropical in Carboniferous	3.272***	3.228***	3.386***	3.588***	2.110***	2.877***
	(0.823)	(0.725)	(0.758)	(1.034)	(0.617)	(0.723)
Montiel-Pflueger F statistic	15.8	19.8	19.9	12.0	11.7	15.8
Observations	77	93	89	99	88	97
Panel C: Democracy (OLS estimates of equivalent model)						
% in manufacturing	0.319***	0.388***	0.333***	0.416***	0.297***	0.060***
	(0.096)	(0.086)	(0.083)	(0.070)	(0.106)	(0.016)
Continent fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.319	0.460	0.421	0.634	0.306	0.331
Observations	77	93	89	99	88	97

Notes: OLS and 2SLS regressions. Countries enter the sample if they were independent throughout the 1990 to 2000 period and data is available. The dependent variable is the level of democracy in 2000 (except for columns (4) and (5)). In columns (1) to (5) this variable is measured by the combined polity index of Polity IV (rescaled to range from 0 to 10). In column (6) this variable is measured by the dichotomous electoral democracy variable of Boix, Miller and Rosato (2013). % in manufacturing is the percentage of the total population that is employed in the manufacturing sector. This variable is averaged over the 1990 to 2000 period. Manufacturing employment data comes from Mitchell (2013), the United Nations Industrial Development Organization, and the 10-Sector Database of the Groningen Growth and Development Centre. Population data comes from Bolt et al. (2018). Tropical in Carboniferous is a dummy that takes the value 1 if a current country was located in the tropical zone during the late Carboniferous era (+/- 300 million years ago), and 0 otherwise. This variable is generated by the author using data from: `globalgeology.com`. Dummies for American, European, Asian, and African countries are included in all regressions. Robust standard errors in parentheses. All regressions contain a constant which is not reported.

$***$ p<0.01, $**$ p<0.05, $*$ p<0.10.

6 Panel fixed effects estimates

No instrument can be foolproof, as there is no way to empirically verify that the instrument (the geological potential to have coal) is only correlated with the outcome variable of interest (democracy) through the potentially endogenous independent variable (industrialization). Acknowledging this I complement my instrumental variable results with an entirely distinct empirical strategy: dynamic panel models with country fixed effects, time fixed effects, and lags of industrialization on the right-hand side. Although correlational in nature such two-way fixed effects models do control by design (rather than by assuming the instrument to be valid) for: (i) reversed causality; (ii) all time-invariant factors; and (iii) all confounding factors that affect all countries at the same point in time. Focusing on the over time variation in industrialization and democratization also contributes substantively as it allows me to assess whether the effect differs across time periods, and whether the effect differs across transitions to and consolidations of democracy.

As is standard in the literature I collapse the data into 5-year simple moving averages, lag the % employed in manufacturing variable with one time period, and include two lags of democracy on the right-hand side to account for serial autocorrelation and potential mean-reverting.[37] After including two lags of democracy I find no economically or statistically significant effect of further lags of democracy and no evidence of serial autocorrelation in a Wooldridge (2010) test.[38] Further in line with the existing literature I focus in my baseline models on the relationship between industrialization and democracy in levels, but find similarly large effects when differencing the data.[39][40]

In table 6 I report the results from the simple OLS "within" estimator. In column (1) I find an estimate of the long-run effect of industrialization on democracy of 0.268.[41] This effect is roughly similar, although slightly smaller, than the cross-sectional instrumental variable estimate of 0.342 in table 4.[42] This two-way fixed effects estimate is causal to the extent

[37]The results are substantially the same when using 10, 15, 20, or 25 year moving averages.

[38]Nonetheless all results hold with 5 lags of democracy on the right-hand side.

[39]Note that a fisher test comfortably rejects the null-hypothesis of non-stationarity in all panels (P-value: 0.000), and that the combined coefficient on the lagged democracy variables is always significantly smaller than 1, which is consistent with the dependent variable being stationary.

[40]All results also hold with linear and quadratic time trends on the right-hand side.

[41]The long-run effect in dynamic panel models is given by:

$$\beta_1 X_{it-1}/(1-(\beta_2 Y_{it-1}+\beta_3 Y_{it-2})) \quad (1)$$

where X is % in manufacturing, Y is democracy, and the β's are regression coefficients (Pickup, 2014).

[42]There is good reason to believe that OLS estimation in this context leads to a lower bound estimate of the effect of industrialization on democracy (holding endogeneity concerns constant). This is because the differencing out of the unit fixed effects induces by construction a correlation between the regressors and errors term. Nickell (1981) showed that this in turn causes the regressor of interest to be biased downwards

that there exist no time-varying factors that simultaneously determine both democracy and industrialization. Although I am unable to provide a conclusive design-based solution to this problem of time-varying confounders I attempt to alleviate it by adding a large number of time-varying control variables. I include: (1) whether a country was aligned to the West, the USSR, or unaligned during the Cold War; (2) inter and intra-state warfare; (3) openness to international trade; (4) oil rents as a fraction of GDP; (5) the degree of state ownership in the economy; (6) private property rights security; (7) whether a country was a colonial power during the period of observation; and (8) the number of years a current regime type is in place (see appendix A for the measurement of these variables).[43] As can be seen in columns (2) to (6) of table 6 the effect of industrialization on democracy remains large, and substantively roughly similar to the effect in column (1), after controlling for these factors (and country and time fixed effects).

In table 7 I implement the same robustness checks as in table 5 and find that the results hold and remain quantitatively very similar when applying different sample restrictions (e.g. excluding Western countries), when using different lag specifications (i.e. using 10 year periods), and when using the Boix, Miller and Rosato (2013) electoral democracy measure as the dependent variable. I also find similar, if not stronger, effects in a balanced sample over the 1960 to 2000 period (column (4) in table 5).

To summarize the time series results: (1) the effect of industrialization on democracy is robust to controlling for reversed causality, all time-invariant confounding factors, all confounding factors that affect all countries at the same time, and all time-varying variables controlled for in table 6; and (2) the OLS/2SLS estimates of the relationship between industrialization and democracy in the 1990s are in terms of magnitude roughly in line with the nature of that same relationship over the 170 years period from 1845 to 2015.

when the (combined) coefficient on the lagged dependent variable is less than 1 (as in my case). In line with this appendix table 13 finds slightly larger effects of industrialization when using the Anderson and Hsiao (1982) IV estimator, the Arellano and Bond (1991) difference-GMM estimator, the Blundell and Bond (1998) system-GMM estimator, or the bootstrap corrected estimator of Everaert and Pozzi (2007), which provide different econometric solutions to correct for the Nickell bias.

[43] I include these control variables at $t-2$ to avoid post-treatment bias. All results nonetheless hold when (also) including them at $t-1$.

Table 6: Two-way fixed effects estimates.

	(1)	(2)	(3)	(4)	(5)	(6)
% in manufacturing $_{t-1}$	0.102***	0.096***	0.112***	0.084**	0.100***	0.098***
	(0.029)	(0.027)	(0.029)	(0.033)	(0.031)	(0.034)
Democracy $_{t-1}$	0.816***	0.814***	0.814***	0.792***	0.815***	0.817***
	(0.051)	(0.051)	(0.050)	(0.057)	(0.051)	(0.051)
Democracy $_{t-2}$	-0.195***	-0.188***	-0.194***	-0.213***	-0.205***	-0.212***
	(0.034)	(0.035)	(0.034)	(0.039)	(0.035)	(0.038)
Aligned to West $_{t-2}$		0.116				
		(0.179)				
Aligned to USSR $_{t-2}$		0.618**				
		(0.310)				
Intra-state warfare $_{t-2}$			0.499***			
			(0.156)			
Inter-state warfare $_{t-2}$			-0.171			
			(0.231)			
Trade openness $_{t-2}$				-0.157***		
				(0.040)		
Oil rents (% of GDP) $_{t-2}$				-0.009		
				(0.009)		
State ownership $_{t-2}$					0.010	
					(0.072)	
Property rights security $_{t-2}$					0.055	
					(0.048)	
Colonial power $_{t-2}$						-0.022
						(0.313)
Regime duration $_{t-2}$						-0.003
						(0.002)
Long-run effect of	0.268***	0.256***	0.295***	0.200**	0.256***	0.250***
% in manufacturing	(0.074)	(0.072)	(0.074)	(0.079)	(0.077)	(0.087)
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1425	1425	1425	1175	1413	1415
Countries	148	148	148	128	148	146

Notes: OLS "within" regressions with time dummies. Panel is unbalanced and includes data from 1845 to 2015. Data is in 5 year simple moving averages. Countries enter the panel if they were independent throughout the previous 10 years and data is available. The dependent variable is democracy, measured by the combined polity index of Polity IV (rescaled to range from 0 to 10). % in manufacturing is the percentage of the total population that is employed in the manufacturing sector. Manufacturing employment data comes from Mitchell (2013), the United Nations Industrial Development Organization, and the 10-Sector Database of the Groningen Growth and Development Centre. Population data comes from Bolt et al. (2018). See the appendix for the measurement of control variables. Standard errors robust against heteroskedasticity and serial correlation at the country level are reported in parentheses. All regressions contain a constant which is not reported.

$***$ p<0.01, $**$ p<0.05, $*$ p<0.10.

Table 7: Robustness checks two-way fixed effects estimates.

	Excluding West (1)	Exluding East Asia (2)	Excluding Middle East (3)	Balanced 1960-2000 (4)	10 year averages (5)	Boix democracy (6)
% in manufacturing $_{t-1}$	0.093*	0.089***	0.106***	0.223***	0.130***	0.012***
	(0.048)	(0.032)	(0.030)	(0.069)	(0.047)	(0.004)
Democracy $_{t-1}$	0.741***	0.812***	0.827***	0.705***	0.537***	0.822***
	(0.051)	(0.052)	(0.053)	(0.070)	(0.060)	(0.039)
Democracy $_{t-2}$	-0.183***	-0.195***	-0.210***	-0.266***	-0.111**	-0.249***
	(0.036)	(0.034)	(0.035)	(0.070)	(0.049)	(0.033)
Long-run effect of	0.211*	0.231***	0.277***	0.398***	0.227***	0.028***
% in manufacturing	(0.109)	(0.082)	(0.075)	(0.109)	(0.081)	(0.010)
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1083	1367	1306	366	700	1465
Countries	129	144	135	54	147	152

Notes: OLS "within" regressions with time dummies. Panel is unbalanced and includes data from 1845 to 2015. Except for column (5) data is in 5 year simple moving averages. Countries enter the panel if they were independent throughout the previous 10 years and data is available. The dependent variable is the level of democracy. In columns (1) to (5) this variable is measured by the combined polity index of Polity IV (rescaled to range from 0 to 10). In column (6) this variable is measured by the dichotomous electoral democracy variable of Boix, Miller and Rosato (2013). % in manufacturing is the percentage of the total population that is employed in the manufacturing sector. Manufacturing employment data comes from Mitchell (2013), the United Nations Industrial Development Organization, and the 10-Sector Database of the Groningen Growth and Development Centre. Population data comes from Bolt et al. (2018). See the appendix for the measurement of control variables. Standard errors robust against heteroskedasticity and serial correlation at the country level are reported in parentheses. All regressions contain a constant which is not reported.

$***$ p<0.01, $**$ p<0.05, $*$ p<0.10.

7 Treatment effect heterogeneity

As this is an important theme in the literature on income and democracy I now shortly test whether the effect of industrialization on democracy varies across transitions to and consolidations of democracy (section 7.1), and whether the effect is conditional on different time periods and international orders (section 7.2). In addition I discuss how, if at all, my results are affected by the structural change towards modern service employment at very high levels of industrialization (section 7.3).

7.1 Transitions to and consolidations of democracy

In table 8 I see whether the effect of industrialization differs across democratic transitions and democratic consolidations (as suggested by Przeworski and Limongi (1997) and Przeworski et al. (2000) with regard to the relationship between GDP per capita and democracy). To do so I estimate two different models. In the first model I regress whether a country-year is democratic on lagged industrialization, while restricting the sample to countries that were autocratic (column (1)) or democratic (column (2)) in the previous year (all according to the dichotomous democracy measure of Boix, Miller and Rosato (2013)). For identification I instrument the lagged level of industrialization with whether a country was located in the tropical zone during the Carboniferous, or not. As can be seen in columns (1) and (2) of table 8 a one percentage point increase in industrialization is estimated to lead, on average, to a 18.4% increase in the probability that a previously autocratic country transitions to democracy in the following year, and a 19.3% increase in the probability that a previously democratic country remains a democracy in the following year.

To distinguish between transitions to and consolidations of democracy using the continuous Polity IV measure of democracy I estimate two-way fixed effects models whereby I interact the lagged level of industrialization with a dummy that takes that value 0 if a country had the lowest Polity IV score (i.e. score 0) in all of the previous 5 years, and 1 otherwise. This model thus effectively captures whether the effect of industrialization on democracy is conditional on being a "full" autocracy in the previous period. As can be seen in column (3) industrialization has also according to this specification a large effect on both transitions to and consolidations of democracy. The estimate suggest that a country that was fully autocratic in the previous 5 years would on average experience a long-run 0.453 scale point increase, on the 11-point combined polity index, as a result of a one percentage point increase in manufacturing employment. This effect on democratic transitions is, if anything, larger than the effect on democratic consolidations.

Table 8: Effect on transitions to and consolidations of democracy.

	Democratic transition (1)	Democratic consolidation (2)	Polity democracy (3)
% in manufacturing $_{t-1}$	0.184**	0.193***	0.259**
	(0.092)	(0.038)	(0.116)
% in manufacturing $_{t-1}$			-0.129
* Full autocracy $_{t-1}$			(0.099)
Full autocracy $_{t-1}$			0.046
[*Reference*: Polity $\neq 0$ $_{t-1}$]			(0.468)
Democracy $_{t-1}$			0.538***
			(0.060)
Democracy $_{t-2}$			-0.110**
			(0.049)
Long-run effect of	–	–	0.453**
% in manufacturing			(0.208)
Country fixed effects	No	No	Yes
Time fixed effects	Yes	Yes	Yes
Montiel-Pflueger F statistic	10.6	29.1	–
Estimation technique	IV probit	IV probit	OLS FE
Observations	2589	2964	700
Countries	87	76	147

Notes: Instrumental variable probit regressions with time dummies in columns (1) and (2). OLS "within" regressions with time dummies in column (3). Panel is unbalanced and includes data from 1845 to 2015. Data is in 5 year simple moving averages in column (3). Countries enter the panel if they were independent throughout the previous 5 years and data is available. The dependent variable in column (1) takes the value 1 if a country in that year became an electoral democracy according to Boix, Miller and Rosato (2013), and 0 otherwise. The dependent variable in column (2) takes the value 1 if a country in that year remained an electoral democracy according to Boix, Miller and Rosato (2013), and 0 otherwise. The sample in column (1) is restricted to autocratic country-years in $t-1$. The sample in column (2) is restricted to democratic country-years in $t-1$. The dependent variable in column (3) is the combined polity index of Polity IV (rescaled to range from 0 to 10). % in manufacturing is the percentage of the total population that is employed in the manufacturing sector. Manufacturing employment data comes from Mitchell (2013), the United Nations Industrial Development Organization, and the 10-Sector Database of the Groningen Growth and Development Centre. Population data comes from Bolt et al. (2018). Coefficients are average marginal effects. Standard errors robust against heteroskedasticity and serial correlation at the country level are reported in parentheses. All regressions contain a constant which is not reported.

$***$ p<0.01, $**$ p<0.05, $*$ p<0.10.

7.2 Time periods, world orders, and waves of democracy

Next I see whether the effect of industrialization on democracy varies across the pre- and post-1950 period and/or whether the effect is conditional on the nature of the international system at any point in time (as suggested by Boix and Stokes (2003), Huntington (1991), and Boix (2011) with regard to the relationship between GDP per capita and democracy).

As shown in table 9 I find that the effect of industrialization on democracy does not differ significantly across the pre- and post-1950 period (if anything the effect is stronger in the post-1950 period), and that the positive effect of industrialization on democracy is not conditional on the international regime classification of Boix (2011), or the three waves of democratization classification of Huntington (1991). More specifically the models estimate that a one percentage point increase in industrialization after 1950, during Boix's (2011) "polarized order", or Huntington's (1991) Third Wave (the least-likely cases for modernization to affect democracy, according to these authors) would still lead to a respective 0.323, 0.333, and 0.282 long-run increase in the 11-point combined polity index.

Table 9: Heterogeneity of effect across time.

	(1)	(2)	(3)
% in manufacturing $_{t-1}$	0.122***	0.126***	0.107**
	(0.031)	(0.030)	(0.041)
Democracy $_{t-1}$	0.814***	0.808***	0.812***
	(0.051)	(0.050)	(0.050)
Democracy $_{t-2}$	-0.193***	-0.187***	-0.191***
	(0.034)	(0.034)	(0.034)
Pre-1950	-1.607***		
[*Reference*: Post-1950]	(0.428)		
Pre-1950 * % in manufacturing $_{t-1}$	-0.056		
	(0.042)		
Prodemocratic order (1918–1932; 1991–2000)		-0.816	
[*Reference*: Polarized order (1933-1990)]		(0.969)	
Neutral order (1848–1917)		-8.091***	
		(2.205)	
Prodemocratic * % in manufacturing $_{t-1}$		-0.089***	
		(0.027)	
Neutral * % in manufacturing $_{t-1}$		-0.027	
		(0.036)	
First wave (1850–1920)			0.814
[*Reference*: Third wave (1975–)]			(0.678)
Second wave (1945–1975)			5.527***
			(1.321)
First wave * % in manufacturing $_{t-1}$			-0.071
			(0.061)
Second wave * % in manufacturing $_{t-1}$			0.044*
			(0.025)
Long-run effect of	0.323***	0.333***	0.282***
% in manufacturing	(0.082)	(0.078)	(0.108)
Country fixed effects	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes
Observations	1425	1425	1425
Countries	148	148	148

Notes: OLS "within" regressions with time dummies. Panel is unbalanced and includes data from 1845 to 2015. Data is in 5 year simple moving averages. Countries enter the panel if they were independent throughout the previous 10 years and data is available. The dependent variable is democracy, measured by the combined polity index of Polity IV (rescaled to range from 0 to 10). % in manufacturing is the percentage of the total population that is employed in the manufacturing sector. Manufacturing employment data comes from Mitchell (2013), the United Nations Industrial Development Organization, and the 10-Sector Database of the Groningen Growth and Development Centre. Population data comes from Bolt et al. (2018). The international order classification is based on Boix (2011). The democratization waves classification comes from Huntington (1991). Standard errors robust against heteroskedasticity and serial correlation at the country level are reported in parentheses. All regressions contain a constant which is not reported.

$***$ p<0.01, $**$ p<0.05, $*$ p<0.10.

7.3 Post-industrial employment at very high levels of development

A last issue is how the results are affected by the relatively recent trend of deindustrialization (in terms of employment) in many OECD countries.[44] The fact is that at least up until now no OECD country has seen a clear and sustained decline in its level of democracy in the past few decades (at least according to the data from Polity IV and Boix, Miller and Rosato (2013)).[45] This may suggest that employment deindustrialization seizes to have its negative effect on democracy if it is the result of a structural change towards modern service employment (a type of structural change which in itself is largely a result of successful industrialization – Rowthorn and Wells (1987)).[46] Alternatively, it may be the case that the modern service sector can in the long-run not fully sustain the same manufacturing-induced social structure that supports democracy (e.g. education, purchasing power labor- and middle classes, inequality, organizational capacity civil society), but that the negative effect of deindustrialization in highly developed countries still lies too far into the future to observe, or that the negative effect of deindustrialization in such cases only sets in when countries pass a particular threshold of employment deindustrialization.[47]

Most important for my particular purpose is, however, that remaining agnostic about this issue can only lead me to *underestimate* the true average treatment effect of industrialization on democracy. This is because if it is in fact the case that manufacturing-induced structural change towards modern services is able to sustain the same pro-democratic causal mechanisms as manufacturing does, then my econometric specification would "incorrectly" predict a lowering of democracy in the years that OECD countries' employment structure shifted from manufacturing to modern services.

[44]Note that OECD countries have (with the exception of the United Kingdom) only deindustrialized in terms of employment, not manufacturing value added (in constant prices) (Rodrik, 2016).

[45]The exceptions among the current OECD member states are Greece and Hungary after 2011, and Turkey after 2015. Given that these countries are certainly not among the OECD countries most severely affected by (employment) deindustrialization this democratic decline can, however, hardly be ascribed to deindustrialization.

[46]Rowthorn and Wells (1987) describes this process as: "[...] the long-term growth rate of output is normally about the same for industrial products as for services. The same is true for expenditure. Thus, in real terms, there is no structural shift in output or expenditure from industry to services. However, labour productivity rises consistently faster in industry than services. Thus, to keep output rising at the same rate in the two sectors requires a continuous shift in the pattern of employment; in relative terms, labour must be continuously transferred from the industrial sector into services." (p. 22)

[47]Several studies find, for example, that employment deindustrialization in the OECD has in many cases led to high unemployment rates and stagnant, or even declining, living standards for many working- and middle class workers previously employed in manufacturing (Kollmeyer and Pichler, 2013; Brady and Wallace, 2001). Such a dynamic may create problems for the sustainment of democracy in the long-run, particularly if such dynamics also negatively feed into other dimensions of the socio-economic structure that supports liberal democracy (e.g. inequality, educational attainment, civil society participation).

8 Conclusion

Understanding why some countries are well-functioning democracies while other countries are autocracies is a major concern in comparative politics, and is of paramount importance for understanding the root causes of the major differences in human welfare across countries and over time. This paper contributes to this endeavor by showing that industrialization is an important determinant of transitions to and consolidations of democracy.

The paper contributes to the long run academic debate on the two-way relationship between economic development and political institutions. The results challenge the (implicit) assumption in the existing modernization theory literature that all increases in economic production (as captured by GDP per capita) are equally important for democracy. Instead the results suggest that growth derived from different types of production activities affects democracy differently, and that taking this into account may explain why the existing literature finds little evidence for a consistent effect of GDP per capita on democracy.

The results are also related to the existing New Institutional Economics literature and the important issue of the political-economic future of China. Over the past 30 years the development economics literature has reached a widespread consensus suggesting that private property rights security is the most important cause of cross-country differences in economic development, and that private property rights security is in turn determined by the degree to which a country' political institutions credibly constrain and credible hold accountable state executive power (e.g. North (1990), Acemoglu, Johnson and Robinson (2001, 2005)). Seen from this perspective China, arguably the greatest development success in human history, is seen as a major outlier (Acemoglu and Robinson, 2012). The results from this paper show that, at least when one focuses on industrialization (rather than GDP per capita), China is no exception at all, and that China's political-economic development is at least up until now directly in line with that what we have observed over the past 170 years – industrialization first, democratization second.[48] This order of causality is also directly in line with the qualitative-historical evidence on the development experience of all countries that actually managed to transform themselves into highly industrialized societies. While Western early industrializers went through their respective industrial revolutions during the 19th century they generally only truly democratized, at least in terms of enfranchisement rates, in the first half of the 20th century, or even later. The story for the late-developers in

[48]Note that China still being a dictatorship at the moment is not vastly out of line, in terms of the level of industrial employment, with the historical record either. The best estimates indicate that today approximately 19.2% of the Chinese workforce works in the manufacturing sector. This is still significantly lower as compared to when currently highly-industrialized countries introduced universal adult suffrage and competitive multiparty elections: United States 23.4% in 1965, United Kingdom 25.8% in 1928, France 23.6% in 1946, South Korea 25.0% in 1988, and Taiwan 26.2% in 1996.

East-Asia is the same. Japan's industrialization started with the Meiji Restoration in 1868 while it democratized in 1952. South Korea started industrializing in 1963 and held its first democratic election in 1988. Taiwan started industrializing in the 1950s and democratized in 1996 (Chang, 2002). Although competitive multiparty elections and institutional checks and balances may be important for economic development after countries have already reached very high levels of industrialization, the introduction of these type of political institutions appears to be more an outcome, rather than a cause, of industrialization.

The results also have important policy implications. This is particularly so because many developing countries' governments have recently expressed the belief that industrialization is no longer an important goal in itself, and that poor countries may instead be able to reach Western standards of living by moving directly into a service economy (e.g. government of India). Besides the economic question whether it is in fact true that countries can reach high levels of development in the long-run without industrializing my results suggest that this "post-industrial" development strategy may have large unintended political consequences as the socio-economic structure that supports well-functioning democratic institutions largely comes about through the process of industrialization. It is useful to point out in this regard that my instrumental variable results do *not* imply that a country' regime type is predetermined by the degree to which it has access to coal. Industrialization is only one of the determinants of democracy, and access to coal is only one of the determinants of industrialization (the importance of which has decreased over time under the influence of cheaper shipping and alternative energy sources). Since a country' geological potential to have coal is plausibly exogenous, it is useful as an instrument to estimate the effect of industrialization on democracy. This, however, is not meant to say that countries are doomed to remain little industrialized and/or little democratic if they have little access to coal. In contrast, industrialization is an outcome that is largely determined endogenously through public policy and political economy factors which developing countries can attempt to systematically put in place (see, for example, the discussion on industrial policy in Lin and Chang (2009)). This endogeneity in the process of industrialization is exactly what necessitates the use of an instrument to identify the effect of industrialization on democracy in the first place.

References

Acemoglu, Daron. 2009. *Introduction to Modern Economic Growth*. Princeton University Press.

Acemoglu, Daron and James A. Robinson. 2006. *Economic Origins of Dictatorship and Democracy*. Cambridge University Press.

Acemoglu, Daron and James A Robinson. 2012. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. New York: Crown Business.

Acemoglu, Daron, Simon Johnson and James A. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91(5):1369–1401.

Acemoglu, Daron, Simon Johnson and James A. Robinson. 2005. Institutions as a Fundamental Cause of Long-Run Growth. In *Handbook of Economic Growth*, ed. Philippe Aghion and Stephen Durlauf. Amsterdam: Elsevier pp. 385–472.

Acemoglu, Daron, Simon Johnson, James A. Robinson and Pierre Yared. 2008. "Income and Democracy." *American Economic Review* 98(3):808–42.

Acemoglu, Daron, Simon Johnson, James A. Robinson and Pierre Yared. 2009. "Reevaluating the Modernization Hypothesis." *Journal of Monetary Economics* 56(8):1043–1058.

Ahmadov, Anar K. 2014. "Oil, Democracy, and Context: A Meta-Analysis." *Comparative Political Studies* 47(9):1238–1267.

Allen, Robert C. 2011. *Global Economic History: A Very Short Introduction*. Vol. 282 Oxford University Press.

Andersen, Jørgen J and Michael L Ross. 2014. "The Big Oil Change: A Closer Look at the Haber–Menaldo Analysis." *Comparative Political Studies* 47(7):993–1021.

Anderson, Theodore Wilbur and Cheng Hsiao. 1982. "Formulation and Estimation of Dynamic Models using Panel Data." *Journal of Econometrics* 18(1):47–82.

Ansell, Ben W. and David J. Samuels. 2014. *Inequality and Democratization: An Elite-Competition Approach*. Cambridge University Press.

Arellano, Manuel and Stephen Bond. 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Review of Economic Studies* 58(2):277–297.

Bairoch, Paul. 1995. *Economics and World History: Myths and Paradoxes*. Chicago: University of Chicago Press.

Barbieri, K. and O.M.G. Keshk. 2016. "Correlates of War Project Trade Data Set Codebook, Version 4.0." Available at: `http://correlatesofwar.org` (last accessed: April 24, 2019).

Benhabib, Jess, Alejandro Corvalan and Mark M. Spiegel. 2013. "Income and Democracy: Evidence from Nonlinear Estimations." *Economics Letters* 118(3):489–492.

Bermeo, Nancy and Philip Nord. 2000. *Civil Society Before Democracy: Lessons from Nineteenth-Century Europe*. Rowman & Littlefield Publishers.

Blundell, Richard and Stephen Bond. 1998. "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models." *Journal of Econometrics* 87(1):115–143.

Boix, Carles. 2003. *Democracy and Redistribution*. Cambridge University Press.

Boix, Carles. 2011. "Democracy, Development, and the International System." *American Political Science Review* 105(4):809–828.

Boix, Carles, Michael Miller and Sebastian Rosato. 2013. "A Complete Data Set of Political

Regimes, 1800–2007." *Comparative Political Studies* 46(12):1523–1554.
Boix, Carles and Susan C Stokes. 2003. "Endogenous Democratization." *World Politics* 55(4):517–549.
Bolt, Jutta, Robert Inklaar, Herman de Jong and Jan Luiten van Zanden. 2018. "Rebasing 'Maddison': New Income Comparisons and the Shape of Long-Run Economic Development." Maddison Project Working Paper 10.
BP, The British Petroleum Company. 2017. "Statistical Review of World Energy." Available at: `https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html` (last accessed: April 11, 2019).
Brady, David and Michael Wallace. 2001. "Deindustrialization and Poverty: Manufacturing Decline and AFDC Recipiency in Lake County, Indiana 1964–93." *Sociological Forum* 16(2):321–358.
Brecke, Peter. 2001. "The Long-Term Patterns of Violent Conflict in Different Regions of the World." Uppsala Conflict Data Conference.
Broderstad, Troy Saghaug. 2018. "A Meta-Analysis of Income and Democracy." *Democratization* 25(2):293–311.
Brooks, Sarah M and Marcus J Kurtz. 2016. "Oil and Democracy: Endogenous Natural Resources and the Political "Resource Curse"." *International Organization* 70(2):279–311.
Cali, Massimiliano, Karen Ellis and Dirk Willem te Velde. 2008. *The Contribution of Services to Development: The Role of Regulation and Trade Liberalisation*. London: Overseas Development Institute.
Chang, Ha-Joon. 2002. *Kicking Away the Ladder: Development Strategy in Historical Perspective*. London: Anthem Press.
Che, Yi, Yi Lu, Zhigang Tao and Peng Wang. 2013. "The Impact of Income on Democracy Revisited." *Journal of Comparative Economics* 41(1):159–169.
Collier, Ruth Berins. 1999. *Paths Toward Democracy: The Working Class and Elites in Western Europe and South America*. Cambridge University Press.
Crafts, Nicholas and Abay Mulatu. 2006. "How Did the Location of Industry Respond to Falling Transport Costs in Britain before World War I?" *Journal of Economic History* 66(3):575–607.
Dahlum, Sirianne, Carl Henrik Knutsen and Tore Wig. 2019. "Who Revolts? Empirically Revisiting the Social Origins of Democracy." *The Journal of Politics* 81(4):1494–1499.
Dahlum, Sirianne, Tore Wig and Carl Henrik Knutsen. 2019. "Educating Demonstrators: Education and Mass Protest in Africa." *Journal of Conflict Resolution* 63(1):3–30.
EIA, United States Energy Information Administration. 2018. "Coal Explained: Coal Prices and Outlook." Available at: `https://www.eia.gov/energyexplained/index.php?page=coal_prices` (last accessed: March 3, 2019).
Epstein, David L, Robert Bates, Jack Goldstone, Ida Kristensen and Sharyn O'Halloran. 2006. "Democratic Transitions." *American Journal of Political Science* 50(3):551–569.
Evans, Geoffrey and Pauline Rose. 2007. "Support for Democracy in Malawi: Does Schooling Matter?" *World Development* 35(5):904–919.
Everaert, Gerdie and Lorenzo Pozzi. 2007. "Bootstrap-Based Bias Correction for Dynamic Panels." *Journal of Economic Dynamics and Control* 31(4):1160–1184.
Fernihough, Alan and Kevin Hjortshøj O'Rourke. 2014. Coal and the European Industrial Revolution. Working Paper No. 19802 National Bureau of Economic Research.

Fink-Jensen, Jonathan. 2015. World Countries Total Population Size, 1500-2000. Working paper clio-infra.

Glaeser, Edward L., Giacomo A.M. Ponzetto and Andrei Shleifer. 2007. "Why does Democracy Need Education?" *Journal of Economic Growth* 12(2):77–99.

Gollin, Douglas. 2010. Agricultural Productivity and Economic Growth. In *Handbook of Agricultural Economics*, ed. Bruce L. Gardner and Gordon C. Rausser. Elsevier pp. 3825–3866.

Gollin, Douglas, Remi Jedwab and Dietrich Vollrath. 2016. "Urbanization With and Without Industrialization." *Journal of Economic Growth* 21(1):35–70.

Haber, Stephen and Victor Menaldo. 2011. "Do Natural Resources Fuel Authoritarianism? A Reappraisal of the Resource Curse." *American Political Science Review* 105(1):1–26.

Haraguchi, Nobuya, Charles Fang Chin Cheng and Eveline Smeets. 2017. "The Importance of Manufacturing in Economic Development: Has This Changed?" *World Development* 93:293–315.

Herrendorf, Berthold, Richard Rogerson and Ákos Valentinyi. 2014. Growth and Structural Transformation. In *Handbook of Economic Growth*. Vol. 2 pp. 855–941.

Hirschman, Albert O. 1959. *The Strategy of Economic Development*. New Haven: Yale University Press.

Hummels, David. 2007. "Transportation Costs and International Trade in the Second Era of Globalization." *Journal of Economic Perspectives* 21(3):131–154.

Huntington, Samuel P. 1991. *The Third Wave: Democratization in the Late Twentieth Century*. University of Oklahoma Press.

Inglehart, Ronald and Christian Welzel. 2009. "How Development Leads to Democracy: What we Know About Modernization." *Foreign Affairs* pp. 33–48.

Jones, Patricia. 2001. "Are Educated Workers Really More Productive?" *Journal of Development Economics* 64(1):57–79.

Klein, Cornelis and Anthony R. Philpotts. 2013. *Earth Materials: Introduction to Mineralogy and Petrology*. Cambridge University Press.

Knutsen, Carl Henrik, John Gerring, Svend-Erik Skaaning, Jan Teorell, Matthew Maguire, Michael Coppedge and Staffan I Lindberg. 2019. "Economic Development and Democracy: An Electoral Connection." *European Journal of Political Research* 58(1):292–314.

Kollmeyer, Christopher and Florian Pichler. 2013. "Is Deindustrialization Causing High Unemployment in Affluent Countries? Evidence from 16 OECD Countries, 1970-2003." *Social Forces* 91(3):785–812.

Krugman, Paul R. 1993. *Geography and Trade*. MIT press.

Kuznets, Simon. 1966. *Modern Economic growth: Rate, Structure, and Spread*. New Haven: Yale University Press.

Lankina, Tomila V and Alexander Libman. 2018. "Soviet Legacies of Economic Development, Oligarchic Rule and Electoral Quality in Eastern Europe's Partial Democracies: The Case of Ukraine." *Comparative Politics* 51(1):127–148.

Lin, Justin and Ha-Joon Chang. 2009. "Should Industrial Policy in Developing Countries Conform to Comparative Advantage or Defy It? A Debate between Justin Lin and Ha-Joon Chang." *Development Policy Review* 27(5):483–502.

Liou, Yu-Ming and Paul Musgrave. 2014. "Refining the Oil Curse: Country-Level Evidence from Exogenous Variations in Resource Income." *Comparative Political Studies*

47(11):1584–1610.

Lipset, Seymour Martin. 1959. "Some Social Requisites of Democracy: Economic Development and Political Legitimacy." *American Political Science Review* 53(1):69–105.

Marshall, Monty G., Ted R. Gurr and Keith Jaggers. 2016. *Dataset Users' Manual: Political Regime Characteristics and Transitions*. Centre for Systematic Peace.

Marx, Karl. 1867. *Capital: Critique of Political Economy (Vol. 1)*. London: Penguin Classics.

Mitchell, Brian. 2013. *International Historical Statistics*. Basingstoke: Palgrave Macmillan.

Moore, Barrington. 1966. *Social Origins of Democracy and Dictatorship*. Boston: Beacon.

Munck, Gerardo L and Jay Verkuilen. 2002. "Conceptualizing and Measuring Democracy: Evaluating Alternative Indices." *Comparative Political Studies* 35(1):5–34.

Nickell, Stephen. 1981. "Biases in Dynamic Models with Fixed Effects." *Econometrica* pp. 1417–1426.

North, Douglass C. 1990. *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press.

Nunn, Nathan and Diego Puga. 2012. "Ruggedness: The Blessing of Bad Geography in Africa." *Review of Economics and Statistics* 94(1):20–36.

Osborne, Dave. 2013. *The Coal Handbook: Towards Cleaner Production: Volume 1: Coal Production*. Elsevier.

Pickup, Mark. 2014. *Introduction to Time Series Analysis*. Sage Publications.

Przeworski, Adam and Fernando Limongi. 1997. "Modernization: Theories and Facts." *World Politics* 49(2):155–183.

Przeworski, Adam, Michael E. Alvarez, Jose Antonio Cheibub and Fernando Limongi. 2000. *Democracy and Development: Political Institutions and Well-being in the World, 1950-1990*. Cambridge University Press.

Ramsay, Kristopher W. 2011. "Revisiting the Resource Curse: Natural Disasters, the Price of Oil, and Democracy." *International Organization* 65(3):507–529.

Rodrik, Dani. 2012. "Unconditional Convergence in Manufacturing." *Quarterly Journal of Economics* 128(1):165–204.

Rodrik, Dani. 2014. The Past, Present, and Future of Economic Growth. In *Towards a Better Global Economy: Policy Implications for Citizens Worldwide in the 21th Century*, ed. Franklin Allen, Jere R. Behrman, Nancy Birdsall, Shahrokh Fardoust, Dani Rodrik, Andrew Steer and Arvind Subramanian. Oxford University Press.

Rodrik, Dani. 2016. "Premature Deindustrialization." *Journal of Economic Growth* 21(1):1–33.

Roodman, David. 2009. "How to Do Xtabond2: An Introduction to Difference and System GMM in Stata." *Stata Journal* 9(1):86–136.

Ross, Michael L. 2001. "Does Oil Hinder Democracy?" *World Politics* 53(3):325–361.

Ross, Michael L. 2013. *The Oil Curse: How Petroleum Wealth Shapes the Development of Nations*. Princeton University Press.

Rowthorn, Bob and John R Wells. 1987. *De-Industrialization Foreign Trade*. Cambridge University Press.

Rueschemeyer, Dietrich, Evelyne H. Stephens and John D. Stephens. 1992. *Capitalist Development and Democracy*. Cambridge University Press.

Sandbu, Martin E. 2006. "Natural Wealth Accounts: A Proposal for Alleviating the Natural Resource Curse." *World Development* 34(7):1153–1170.

Scotese, Christopher R. 2004. "A Continental Drift Flipbook." *Journal of Geology* 112(6):729–741.

Silva, Bruno Castanho and Levente Littvay. 2019. "Comparative Research is Harder Than We Thought: Regional Differences in Experts' Understanding of Electoral Integrity Questions." *Political Analysis* 27(4):599–604.

Thomas, Larry. 2013. *Coal Geology*. Oxford: John Wiley & Sons, Ltd.

Tsui, Kevin K. 2010. "More Oil, Less Democracy: Evidence from Worldwide Crude Oil Discoveries." *Economic Journal* 121(551):89–115.

Vries, Peer H. H. 2001. "Are Coal and Colonies Really Crucial? Kenneth Pomeranz and the Great Divergence." *Journal of World History* 12(2):407–446.

Wårell, Linda. 2006. "Market Integration in the International Coal Industry: A Cointegration Approach." *Energy Journal* pp. 99–118.

WCA, World Coal Association. 2019. "Coal Market and Pricing." Available at: `https://www.worldcoal.org/coal/coal-market-pricing` (last accessed: April 11, 2019).

Wooldridge, Jeffrey. 2013. *Introductory Econometrics: A Modern Approach (5th edition)*. Mason: South-Western, Cengage Learning.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. MIT press.

Wrigley, Edward A. 2010. *Energy and the English Industrial Revolution*. Cambridge University Press.

Zeira, Yael. 2019. "From the Schools to the Streets: Education and Anti-Regime Resistance in the West Bank." *Comparative Political Studies* 52(8):1131–1168.

— Online Appendix —

The Structural Economic Roots of Liberal Democracy

By: Sam van Noort

Appendix A: Measurement

Democracy: Expert coded measure of a country' level of democracy. Data comes from Polity IV, where the variable is originally called the "combined polity score". Democracy is conceived by Polity IV as consisting of three essential and interdependent elements. First, the presence of institutions and procedures through which citizens can express effective preferences about alternative policies and leaders. Second, the existence of institutionalized constraints on the exercise of power by the executive. Third, the guarantee of civil liberties to all citizens in their daily lives and in acts of political participation. The variable is computed by subtracting Polity IV's institutionalized autocracy score from the institutionalized democracy score. To facilitate time series analyses Polity IV recodes foreign "interruptions" as missing, cases of "interregnum" as a neutral polity score of 0, and prorates cases of "transition" across the span of the transition. The variable is rescaled to range from 0 to 10. See for the variable's full codebook: `http://www.systemicpeace.org/inscr/p4manualv2017.pdf` (pages 16-17) (last assessed: April 24, 2019).

Boix democracy: Dichotomous variable coding whether a country is an electoral democracy, or not. The variable takes the value 1 if a country holds free and fair elections and has enfranchised a majority of the male population, and 0 otherwise. Data comes from Boix, Miller and Rosato (2013).

% in manufacturing: The percentage of the total population employed in the manufacturing sector. Manufacturing employment data comes from Mitchell (2013) and the 10-Sector Database of the Groningen Growth and Development Centre. Population data comes from Bolt et al. (2018).

Tropical in Carboniferous: Dichotomous variable that takes the value 1 if a current country was located in the tropical zone during the late Carboniferous era (+/- 300 million years ago), and 0 otherwise. This variable is generated by the author using data from: `globalgeology.com` (last assessed: April 24, 2019). See appendix table 1 for the coding scheme.

Tropical today: The percentage of a country' total land area that is subject to any of the four Köppen-Geiger tropical climates. Data comes from Nunn and Puga (2012).

Soil quality: The percentage of a country' total land area that is very or moderately suitable to produce rainfed crops. Data comes from the website of John Luke Gallup: `https://www.pdx.edu/econ/jlgallup/country-geodata` (last assessed: March 9, 2019).

Oil and gas rents (% of GDP): The real value of a country' oil and gas production as a percentage of GDP. Data on oil and gas production comes from Haber and Menaldo (2011). The variable is normalized with GDP data from Bolt et al. (2018).

% in mining: The percentage of the total population employed in the mining/extractive sector. Mining employment data comes from Mitchell (2013) and the 10-Sector Database of the Groningen Growth and Development Centre. Population data comes from Bolt et al. (2018).

% of years conflict, 1750-2000: The percentage of years a country was engaged in an inter- or intra-state conflict over the 1750 to 2000 period, relative to all years that it was independent and data is available. Data comes from `clio-infra.eu` (last assessed: April 24, 2019).

Aligned to West: Dichotomous variable that take the value 1 if the country is Portu-

gal, Greece, Spain, Thailand, South Korea, South Africa, Turkey, United States, France, Australia, Italy, West Germany, Netherlands, Denmark, United Kingdom, New Zealand, Belgium, Canada, or Norway, and 0 otherwise.

Aligned to USSR: Dichotomous variable that take the value 1 if the country is Cuba, East-Germany, Poland, Czech Republic, Slovakia, Albania, Hungary, Romania, Bulgaria, Estonia, Latvia, Lithuania, Belarus, Ukraine, Russia, Georgia, Armenia, Azerbaijan, Turkmenistan, Uzbekistan, Kazakhstan, Tajikistan, Kyrgyzstan, China, Mongolia, North Korea, or Vietnam, and 0 otherwise.

Intra-state warfare: Dichotomous variable that takes the value 1 if the country is engaged in an intra-state conflict, and 0 otherwise. Data comes from `clio-infra.eu` (last assessed: April 24, 2019).

Inter-state warfare: Dichotomous variable that takes the value 1 if the country is engaged in an inter-state conflict, and 0 otherwise. Data comes from `clio-infra.eu` (last assessed: April 24, 2019).

Trade openness: The sum of imports and exports as a percentage of GDP. Import/export data comes from the Correlates of War dataset (see Barbieri and Keshk (2016)). GDP data comes from Bolt et al. (2018).

Oil rents (% of GDP): The real value of a country' oil production as a percentage of GDP. Data on oil production comes from Haber and Menaldo (2011). GDP data comes from Bolt et al. (2018).

State ownership: Expert coded variable of whether the state owns or directly controls virtually all, most, some, or very little of a country' valuable capital. Data comes from the Varieties of Democracy Project. See for the variable's full codebook: `https://www.v-dem.net/media/filer_public/e0/7f/e07f672b-b91e-4e98-b9a3-78f8cd4de696/v-dem_codebook_v8.pdf` (page 164) (last assessed: May 26, 2019).

Property rights security: Expert coded measure of the degree to which private property rights exist and are effectively enforced. Data comes from the Varieties of Democracy Project. See for the variable's full codebook: `https://www.v-dem.net/media/filer_public/e0/7f/e07f672b-b91e-4e98-b9a3-78f8cd4de696/v-dem_codebook_v8.pdf` (page 237) (last assessed: May 26, 2019).

Colonial power: Takes the value 1 before 1878 for Sweden, before 1953 for Denmark, before 1984 for the United Kingdom, before 1975 for the Netherlands, before 1918 for Germany, before 1962 for Belgium, before 1980 for France, before 1960 for Italy, and 0 otherwise.

Regime duration: The number of consecutive years the current regime type is in place. Data comes from Boix, Miller and Rosato (2013). Regime type refers to Boix, Miller and Rosato's (2013) coding of (electoral) democracy (see above).

Log GDP per capita: The natural log of GDP per capita at purchase power parity. Data comes from the World Bank Development Indicators. Data is in 2011 international dollars.

Appendix B: Descriptive statistics

Table 10: Summary statistics for cross-sectional sample.

Variable name	N	Mean	Std. Dev.	Min	Man
Democracy	95	7.242	2.915	0.500	10
Boix democracy	95	0.621	0.488	0	1
% in manufacturing	95	4.519	3.369	0.088	13.569
Tropical in Carboniferous	95	0.389	0.490	0	1
Tropical today	95	34.039	42.059	0	100
Soil quality	94	34.436	10.932	5.229	65.449
Oil and gas rents (% of GDP)	91	1.523	3.184	0	15.222
Democracy, 1800	95	0.247	1.044	0	7
% of years conflict, 1750-2000	95	17.646	19.051	0	96.364

Notes: See Appendix A for measurement of variables. All democracy variables are measured in the year 2000. All other variables are averaged over the 1990 to 2000 period.

Table 11: Summary statistics for time series sample.

Variable name	N	Mean	Std. Dev.	Min	Man
Democracy	1507	6.024	2.035	0	10
Boix democracy	1499	0.509	0.305	0	1
% in manufacturing	1507	5.486	1.524	0.023	27.386
Aligned to West	1507	0.150	0.260	0	1
Aligned to USSR	1507	0.014	0.069	0	1
Intra-state warfare	1507	0.093	0.193	0	1
Inter-state warfare	1507	0.091	0.188	0	1
Trade openness	1252	1.972	4.336	0.014	147.211
Oil rents (% of GDP)	1279	1.945	2.875	0	72.286
State ownership	1496	0.331	0.680	-4.151	2.651
Property rights security	1496	6.262	1.071	0.032	9.525
Colonial power	1507	0.080	0.136	0	1
Regime duration	1499	44.566	25.188	1	216

Notes: See Appendix A for measurement of variables. Data is in 5-year simple moving averages. Standard deviation reported is the "within" standard deviation.

Appendix C: Countries included in analysis

Table 12: Country-years included in sample.

Country	Time	Country	Time	Country	Time
Afghanistan	1973–2000	Greece	1951–2008	Pakistan	1972–2008
Albania	1988–2015	Guatemala	1950–2006	Panama	1940–2008
Algeria	1962–2004	Guyana	1966–2002	Papua New Guinea	1975–2001
Angola	1975–2014	Haiti	1950–1991	Paraguay	1950–2008
Argentina	1950–2011	Honduras	1950–2007	Peru	1940–2008
Armenia	1991–2014	Hungary	1869–2008	Philippines	1948–2008
Australia	1931–2008	India	1950–1991	Poland	1918–2008
Austria	1919–2008	Indonesia	1961–2012	Portugal	1940–2008
Azerbaijan	1991–2014	Iran	1963–2015	Qatar	1986–2014
Bahrain	1992–2005	Iraq	1957–2002	Romania	1930–2008
Bangladesh	1972–2003	Ireland	1936–2008	Russia	1897–2008
Belarus	2005–2015	Israel	1963–2015	Rwanda	1978–1989
Belgium	1846–2008	Italy	1881–2008	Saudi Arabia	1976–2014
Benin	1975–1981	Ivory Coast	1966–1997	Senegal	1974–2014
Bolivia	1950–2010	Jamaica	1959–2008	Sierra Leone	1963–2004
Botswana	1966–2010	Japan	1872–2008	Singapore	1965–2015
Brazil	1950–2011	Jordan	1963–2015	Slovakia	1993–2015
Bulgaria	1963–2015	Kazakhstan	1998–2014	Slovenia	1991–2014
Burkina Faso	1974–1998	Kenya	1963–2015	Somalia	1967–1986
Burundi	1969–2013	Kuwait	1967–2014	South Africa	1946–2008
Cambodia	1988–2000	Kyrgyzstan	1991–2014	South Korea	1963–2015
Cameroon	1970–2008	Laos	1975–1999	Spain	1940–2008
Canada	1891–2008	Latvia	1991–2015	Sri Lanka	1953–2008
Central African Republic	1973–1993	Lesotho	1982–2009	Sudan	1972–2001
Chile	1940–2008	Liberia	1962–1984	Suriname	1975–2004
China	1952–2011	Libya	1964–1996	Swaziland	1968–2011
Colombia	1938–2004	Lithuania	1992–2015	Sweden	1945–2008
Costa Rica	1950–2011	Luxembourg	1985–2014	Switzerland	1941–1980
Croatia	1991–2015	Macedonia	1991–2011	Syria	1963–2010
Cuba	1943–2008	Madagascar	1967–2006	Taiwan	1949–2000
Cyrus	1960–2008	Malawi	1964–2012	Tajikistan	1991–2013
Czech Republic	1987–2015	Malaysia	1968–2015	Tanzania	1961–2011
Democratic Republic of Congo	1968–1988	Mali	1976–2004	Thailand	1937–2008
Denmark	1945–2008	Mauritius	1968–2008	The Gambia	1975–2004
Dominican Republic	1950–2007	Mexico	1930–2008	Togo	1981
Ecuador	1950–2006	Moldova	1991–2014	Trinidad and Tobago	1962–2008
Egypt	1937–2008	Mongolia	1990–2014	Tunisia	1963–2014
El Salvador	1950–2007	Montenegro	2010–2014	Turkey	1935–2008
Eritrea	1993–2011	Morocco	1960–2012	Uganda	1963–2000
Estonia	1993–2015	Mozambique	1975–2000	Ukraine	1992–2015
Ethiopia	1961–1993	Myanmar	1983–1997	United Arab Emirates	1971–2010
Fiji	1970–2013	Namibia	1990–2000	United Kingdom	1841–2008
Finland	1917–2008	Nepal	1961–1999	United States of America	1870–2008
France	1856–2008	Netherlands	1849–2008	Uruguay	1968–2012
Gabon	1963–1995	New Zealand	1906–2008	Venezuela	1941–2008
Georgia	1998–2015	Nicaragua	1940–2006	Yemen	1988–2013
Germany (Post-1990)	1991–2015	Niger	1960–1977	Zambia	1965–2010
Germany (Pre-1945)	1882–1939	Nigeria	1960–2011	Zimbabwe	1970–1996
West Germany (1945–1990)	1950–1990	Norway	1930–2008		
Ghana	1960–2011	Oman	1993–2015		

Appendix D: Results when using dynamic panel estimators.

Table 13: Time series estimates using different dynamic panel estimators.

	Anderson-Hsiao IV (1)	Arellano-Bond GMM (2)	Blundell-Bond GMM (3)	Everaert-Pozzi BCFE (4)
% in manufacturing $_{t-1}$	0.133*	0.140	0.207***	0.078***
	(0.077)	(0.110)	(0.060)	(0.026)
Democracy $_{t-1}$	0.878***	0.651***	0.729***	1.035***
	(0.070)	(0.192)	(0.124)	(0.063)
Democracy $_{t-2}$	-0.218***	0.009	-0.050	-0.206***
	(0.044)	(0.169)	(0.125)	(0.042)
Long-run effect of	0.391*	0.411*	0.645***	0.453***
% in manufacturing	(0.232)	(0.247)	(0.215)	(0.155)
Country fixed effects	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes
Kleibergen-Paab F statistic	69.1	–	–	–
Hansen J-test, p-value	–	[0.839]	[0.996]	–
AR(2) test, p-value	–	[0.503]	[0.221]	–
No. of GMM instruments	–	74	111	–
Observations	1275	1275	1425	1382
Countries	146	146	148	146

Notes: Panel is unbalanced and includes data from 1845 to 2015. Data is in 5 year simple moving averages. Countries enter the panel if they were independent throughout the previous 10 years and data is available. The dependent variable is democracy, measured by the combined polity index of Polity IV (rescaled to range from 0 to 10). % in manufacturing is the percentage of the total population that is employed in the manufacturing sector. Manufacturing employment data comes from Mitchell (2013), the United Nations Industrial Development Organization, and the 10-Sector Database of the Groningen Growth and Development Centre. Population data comes from Bolt et al. (2018). Column (1) uses the instrumental variable estimator of Anderson and Hsiao (1982), column (2) uses the difference-GMM estimator of Arellano and Bond (1991), column (3) uses the system-GMM estimator of Blundell and Bond (1998), and column (4) uses the bootstrap-corrected estimator developed by Everaert and Pozzi (2007). The instrument matrix of both GMM estimators are collapsed according to Roodman (2009). In both cases the second lags onwards are used as instruments. Standard errors robust against heteroskedasticity and serial correlation at the country level are reported in parentheses. All regressions contain a constant which is not reported.

$***$ p<0.01, $**$ p<0.05, $*$ p<0.10.