The Economic Impact of Universities: Evidence from Across the Globe

Anna Valero[a] and John Van Reenen[b]

April 2016

Abstract

We develop a new dataset using UNESCO source materials on the location of nearly 15,000 universities in about 1,500 regions across 78 countries, some dating back to the 11th Century. We estimate fixed effects models at the sub-national level between 1950 and 2010 and find that increases in the number of universities are positively associated with future growth of GDP per capita (and this relationship is robust to controlling for a host of observables, as well as unobserved regional trends). Our estimates imply that doubling the number of universities per capita is associated with over 4% higher future GDP per capita. Furthermore, there appear to be positive spillover effects from universities to geographically close neighboring regions. We explore some of the mechanisms behind the relationship between growth and universities. There is some role for the supply of educated individuals and spillovers from innovation, but no discernible role for demand effects. In addition we find that within countries, higher historical university presence is associated with stronger pro-democracy attitudes.

JEL No. I23, I25, J24, O10, O31

Keywords: Universities, growth, human capital, innovation

Acknowledgements:

We would like to thank Andrei Shleifer and Andy Feng and participants at seminars at the LSE and CEP for helpful comments. Nicola Gennaioli, Rafael La Porta and Florencio Lopez De Silanes were kind enough to make their regional data available to us. Financial support from the ESRC through the CEP is gratefully acknowledged.

[a] London School of Economics, Centre for Economic Performance

[b] London School of Economics, Centre for Economic Performance, NBER and CEPR

I. INTRODUCTION

A striking feature of the last century has been the large-scale expansion in university education worldwide. In 1900, only about one in a hundred young people in the world were enrolled at universities, but over the course of the Twentieth Century this rose to one in five (Schofer and Meyer, 2005). The term "university" was coined by the University of Bologna, founded in 1088, the first of the medieval universities. These were communities with administrative autonomy, courses of study, publicly recognised degrees and research objectives and were distinct from the religion-based institutions that came before (De Ridder-Symoens, 1992). Since then, universities spread worldwide in broadly the same form, and it has been argued that they were an important force in the Commercial Revolution via the development of legal institutions (Cantoni and Yuchtman, 2014) and the industrial revolution via their role in the building of knowledge and its dissemination (Mokyr, 2002).

While there is an extensive literature on human capital and growth, there is relatively little research on the economic impact of universities themselves. In this paper, we develop a new dataset on the location of universities in 1,500 regions across 78 countries, and explore how university formation has affected economic growth since 1950, when consistent sub-national economic data are first available. This period is of particular interest because in the years following World War II, university expansion accelerated in most countries; a trend partially driven by the new view that higher education is essential for economic and social progress. This was in contrast to the pre-War fears of "over-education" that were prevalent in many countries, should enrolments extend beyond limited sets of national elites (Schofer and Meyer, 2005; Goldin and Katz, 2008).

There are a number of channels through which universities may affect growth including skill supply, innovation, democracy and demand. Firstly, and most obviously, universities are producers of human capital; and skilled workers tend to be more productive than unskilled workers. Distance seems to matter for the probability of attending college and also the likelihood that students will remain as workers in the same area as the university itself (e.g. Card, 2001). The empirical macro literature has generally found that at the country level, human capital (typically measured by years of schooling) is important for development and growth. Growth accounting and development accounting relate educational attainment to economic performance and find some role.[1] Explicit econometric analysis usually, although not always, confirms this.[2] A problem with these analyses is that they are at the country level and subject to serious endogeneity concerns (e.g. Bils and Klenow, 2000; Hanusheck and Woessman, 2009). At the sub-national level, Gennaioli et al (2013) show that regional years of schooling is important for regional GDP per capita in the cross section and Gennaioli et al (2014) confirm this for growth. Furthermore, human capital appears to also have an indirect effect via externalities which are analysed *inter alia* by Gennaioli et al (2013) using international data (where regional human capital is found to raise firm level productivity, over and above firm level human capital) and Moretti (2004) in the US (where city level human capital is found to raise individual wages). In an historical setting, Squicciarini and Voigtländer (2014) show that "upper tail" knowledge was important in the industrial revolution, and they measure this type of knowledge using city level subscriptions to the Encyclopédie in mid-18th century France.

A second channel through which universities may affect growth is innovation. This effect may be direct as university researchers themselves produce innovations, or via universities' role as human capital producers, should graduates enter the workforce and innovate. A number of papers have found that universities increase local innovative capacity. [3] As useful as this literature is a drawback is it

[1] For example, Mankiw et al (1992), Hall and Jones (1994) and Caselli (2005).

[2] For example, Barro (1991) and more recently, de La Fuente and Domenech (2006) and Cohen and Soto (2007)

[3] Jaffe (1989) uses US state level data to provide evidence of commercial spillovers from university research in patenting (in specific technical sectors) and R&D spending by firms, and Jaffe et al (1993) provides more evidence for localization by comparing the distances between citations and cited patents. Belenzon and Schankerman (2013) look more closely at the distance between the location of university patents, and the firms that cite them in subsequent patents. They find that university knowledge spillovers are strongly localized. Hausman (2012) finds that universities stimulate nearby economic activity via the spread of innovation: long run employment and pay rises in firms closely tied with a local university's innovative strength, and this impact increases in proximity to university. Toivanen and Väänänen (2014) consider how universities affect innovation via their role as human capital producers: they use distance to a technical university as an instrument in estimating the effect of engineering education on patents in Finland (which they find to be positive and significant). They perform a counterfactual calculation which suggests that establishing three new technical universities resulted in a 20 per cent increase in USPTO patents in Finland. Kantor and Whalley (2014a) estimate broad based local agglomeration spillovers from research university activity in the US. They find that a 10% increase in higher education spending increases local non education sector income by a modest 0.8% (and that this effect is persistent), but that the effect is larger for firms that are "technologically closer" to universities. They suggest that the effects that universities have on their local economies may grow over time as the composition of local industries adjusts. While much of this effect is likely to be due to innovation spillovers, it may also capture other types of agglomeration externalities. Kantor and Whalley (2014b) explore agricultural productivity effects of proximity to research in US agricultural research stations, they find that effects peak between 20-30 years after the experiment stations opened, with some longer term persistence where stations focused on basic research and farmers were already at the technology frontier.

looks at an indirect cause of growth using proxies for innovation such as patents rather than at output directly. Moreover, the work is also focused on single countries, hence limiting its generalizability. Third, universities can contribute to the development of economic or democratic institutions which may matter for growth. Universities could promote strong institutions directly by providing a platform for democratic dialogue and sharing of ideas, through events, publications, or reports to policy makers. A more obvious channel would be that universities strengthen institutions via their role as human capital producers.

The relationship between human capital, institutions and growth are debated in the literature. Firstly, there is a debate around whether institutions matter at all for growth.[4] Some papers have argued that human capital is the basic source of growth, and the driver of democracy and improved institutions (e.g. Glaeser et al, 2004). But the relationship between education and democracy/institutions is contested by Acemoglu et al (2005b) who show that the effects found in the cross section of countries are not robust to including country fixed effects and exploiting within-country variation. Whether we believe that human capital or institutions are the basic driver of growth, it is clear that universities must have a role to play in this area as human capital producers and institutions in their own right.

Few papers have explicitly considered the direct link between university presence and economic performance. Cantoni and Yuchtman (2014) provide evidence that medieval universities in 14th century Germany played a causal role in the commercial revolution. In a contemporary setting, Aghion et al. (2009) consider the impact of research university activity on US state economies. They find that exogenous increases in investments in four year college education affect growth in all states, and exogenous shocks to research type education have positive growth effects only in states close to the technological frontier. This is explained by the potential beneficiaries of higher education migrating to the frontier states. They also show that innovation is likely to be an important channel for externalities from four year college and research education, since exogenous investments in both increase patenting activity. Feng and Valero (2016) show that firms that are closer to universities

[4] See, for example, Acemoglu et al. (2005a, 2014) who argue institutions matter a lot, and Gerring et al. (2005) for a summary of papers that conclude that they do not.

have better management practices, which they argue is due to a lower skill price faced by such firms, and complementarity between modern management "technologies" and skills.[5]

In this paper, we develop a new dataset using the World Higher Education Database (WHED) which is produced by the International Association of Universities in association with UNESCO. We map the location of over 15,000 universities into 1,500 regions across 78 countries to explore how changes in the number of universities within regions have affected subsequent growth.

Our first step is to estimate the total effect of university growth on regional GDP growth. We use a panel of annualized five yearly growth rates, and regress regional GDP growth on the lagged growth in universities. We include a host of controls (including country or region fixed effects to control for differential regional trends, and year dummies). We find that a doubling of universities in a region is associated with over 4% higher GDP per capita. This result is robust to a number of different specification checks. We show that reverse causality does not appear to be driving this, nor do mechanical demand (or multiplier) type effects. We also find that universities in neighboring regions or other regions in a country also affect a region's growth, and there appears to be a spatial element to this, with larger effects for regions that are close together. Our next step is to explore the mechanisms underlying the university effect on growth. We show that that both the supply of human capital and innovation channels are positive, but small in magnitude. In addition, we find that universities appear to affect views on democracy even when we control or human capital, consistent with a story that they may have some role in shaping institutions over longer time horizons.

This paper is organized as follows. Section II describes the data and some of its key features including interesting trends and correlations which give us a macro level understanding of the global rise in universities over time. Section III sets out our econometric strategy, and Section IV our results. Section V explores the mechanisms through which universities appear to affect regional growth and finally, Section VI provides some concluding comments.

II. DATA

[5] Management quality is measured in the World Management Survey (see Bloom et al., 2015).

Our regression analysis is based upon information on universities in some 1,500 regions in 78 countries. This represents the set of regions for which our university data can be mapped to a regional time series of key economic variables obtained from Gennaioli et al (2014), and covers over 90 per cent of global GDP. [6] We first describe the full World Higher Education Database (WHED) across all countries, with some key global trends and correlations. Then we focus on the 78 countries for which regional economic data are available, describing how we aggregate the WHED data into regions, and present some initial descriptive evidence.

II.1 World Higher Education Database

WHED is an online database published by the International Association of Universities in collaboration with UNESCO. [7] It contains information on higher education institutions that offer at least a three or four year professional diploma or a post-graduate degree. In 2010, there were 16,328 universities across 185 countries meeting this criterion. The database therefore excludes, for example, community colleges in the US and further education institutions in the UK and may be thought of as a sample of "higher quality" universities. The Data Appendix contains more discussion on countries and types of institution omitted. Key variables of interest include university location, founding date, subjects and qualifications offered and other institutional details such as how they are funded.

Our regional analysis is based on that sample of countries for which GDP and other data are available from 1955, which covers 78 countries, comprising 14,870 (or 91%) of the institutions from the full listing. Our baseline results simply use the year-specific count of universities by region as a measure of university presence, always controlling for region population. To calculate this, we first allocate each university to a region (for example, a US state), and then use the founding dates of universities in each region to determine the number of universities that were present at any particular date. [8] High rates of university exit would invalidate this type of approach, but we find that this does not appear

[6] Based on World Bank GDP in 2014 (US dollars, PPP).
[7] For more information, see http://www.whed.net/home.php.
[8] Of the full sample of 16,328 universities, we were unable to obtain founding date information for 670 institutions (4% of the total). 610 of these fall into our core analysis sample (in the 78 countries for which regional economic data are available). These institutions are therefore omitted from analysis.

to be an issue over the decades since the 1950s (see the Data Appendix for an extensive discussion). This is because there has been very little exit from the university sector over this period.

A disadvantage of the "university density" measure is that it does not correct for the size or quality of the university. Unfortunately, this type of data is not available on a consistent basis across all countries. So we present robustness results on sub-samples where we do have finer grained measure of university size and quality to make sure our baseline results are not misleading.

The Worldwide Diffusion of Universities

We begin by presenting some descriptive analysis of the university data at the macro level using the full university database. Figure 1 gives a snapshot of the geographical distribution of universities across countries in 2010. The US is the country with the largest share, accounting for 15% of the world's universities. The distribution is clearly skewed, with seven countries (US, Brazil, Philippines, Mexico, Japan, Russia and India, in descending order) accounting for over half of the universities in the world. Figure 2 shows how the total number of universities has evolved over time; marking the years that the number doubled. The world's first university opened in 1088 and growth took off in the 19th Century, growing most rapidly in the post-World War II period. Figure 3 shows the "extensive margin" – the cumulative number of countries that have any university, with the dates that some countries opened their first university marked on the line. Here we can see that the UK opened its first university in 1167 (University of Oxford) and the US in 1636 (Harvard University). The latest country to open a university in our data is Bhutan in 2003. By 2010, we can see that the vast majority of countries in the world have at least one university.

In Figure 4 we normalize the number of worldwide universities by the global population to show that university density also rose sharply in the 1800s. From the mid-20th Century, the rise in density slowed, as global population growth took off, and we see that it picked up again from the 1990s (the time when emerging markets like Brazil and India saw rapid expansions in universities).

Online Appendix 1 gives an historical overview of the diffusion of universities from the 1880s in four advanced economies: France, Germany, the UK and US, and two emerging economies: India and China. We compare the timing of historical university expansions to growth and industrialisation.

This analysis provides a visual "sense-check" for the thesis developed by Mokyr (2002) that the building and dissemination of knowledge played a major role in the Industrial Revolution.

Figure 5 gives the country level cross sectional correlations of universities with key economic variables. Unsurprisingly, we see that higher university density is associated with higher GDP per capita levels (Panel A). It is interesting that countries with more universities in 1960 generally had higher growth rates over 1960-2000 (Panel B). And there are strong correlations between universities and average years of schooling (Panel C), patent applications (Panel D) and democracy (Panel E).[9] While these are only associations, they provide a basis for us to explore further whether universities matter for GDP growth within countries, and to what extent any effect operates via human capital, innovation or institutions.

II.4 Regional Economic Data

We obtained regional economic data from Gennaioli et al (2014) who have collated a large dataset of key economic variables for growth regressions at the sub-national level. The outcome variable we focus on is regional GDP per capita. Since for many countries, regional GDP data and other variables such as population or years of education are not available annually we follow Barro (2012) and compute average annual growth rates in GDP per capita over five year periods.[10] We also gather patents data at the regional level as a measure of innovation. For the US only, we obtain USPTO data at the state level over the period 1965-1999 from the NBER (Hall et al, 2001). For 38 countries, we obtain region-level European Patent Office (EPO) patents from the OECD REGPAT database covering 1975 to 2005.

We examine the cross sectional correlations between universities and regional GDP per capita, based on the year 2000 – where data for 65 countries out of the full sample are available. Column (1) of Table 1 shows that there is a significant and positive correlation between GDP per capita and universities: controlling for population, a one per cent rise in the number of universities is associated with 0.7% higher GDP per capita. Column (2) includes country fixed effects which reduces the

[9] We use Polity scores as a measure of democracy, as is common in the literature. See for example Acemoglu et al (2014)
[10] We interpolate missing years, but do not extrapolate beyond the final year (or before the first year of data). Our results are robust to dropping interpolated data.

university coefficient from 0.681 to 0.214. We include a host of further geographic controls in column (3) - whether the region contains a capital city, latitude, inverse distance to ocean, malaria ecology and the log of cumulative oil and gas production[11]. This reduces the coefficient on universities still further to 0.160. In column (4) we add years of education. This reduces the coefficient on universities by around 65%, and the coefficient on years of education is highly significant and similar in magnitude to the cross section results in Gennaioli et al (2013). [12] In column (5) we restrict to the sample for which patents data are available, and add years of education in column (6). Again, this reduces the impact of universities- by 50% on this subsample. In column (7) we see that adding a measure of patent "stock" reduces our coefficient on universities by a further 28% to 0.0566, but it remains highly significant. This analysis suggests that, at least in the cross section, universities contribute to higher GDP per capita mainly as human capital producers, but also as producers of innovation. It also implies that even after controlling for human capital and innovation, universities still have an additional effect. This could be due to the fact that years of education or patents are imperfect proxies for human capital and innovation respectively, it could reflect other effects we are unable to capture, or just be the result of biases in the cross section.

There are a multitude of unobservables that can lead regions to have both higher GDP per capita, and more universities, which could result in a positive bias in our coefficient. Our focus therefore turns to growth rates for the bulk of our analysis which allows us to sweep out the time invariant factors. Figure 6 shows that the raw correlations between growth rates of universities and GDP per capita that we saw at the country level are also present within countries. Panel A simply plots the average annual growth in regional GDP per capita on the average annual growth in universities for each region, over the time period for which data are available (which differs by region). Average GDP per capita growth rates are plotted within 20 evenly sized bins of university growth, and country fixed effects are absorbed so that variation is within region. Panel B plots GDP per capita growth rates on lagged university growth for the 8,128 region-years (on which we conduct the core of our analysis that will follow). In both graphs it is clear that there is a positive relationship. In addition, these graphs show that there are observations in the top bin with very high university growth. We explore which region

[11] Specifically, we take the natural log of 1+ this value, so that we retain zeroes in our sample.
[12] In regressions of regional income per capita on years of education, controlling for geographic characteristics, Gennaioli et al (2013) estimate a coefficient of 0.2763, see Table IV column (2).

years are driving this in case they are outliers that should be excluded from our regression analysis, and find that in fact they are spread across 60 countries and different years. Around half of the observations involve regions opening their first universities, which clearly generates a large percentage change in universities per million people. Dropping the observations in the highest growth bin actually strengthens the correlation in this simple scatter plot, and later we show that are regression results are similarly robust to dropping these observations or winsorizing lagged university growth.

To round up this section, we describe the key variables used in our analysis. Table 2 shows that the average observation in our sample of 8,128 region-years has GDP per capita of just over $13,000, and the average growth rate is two per cent. It has nearly ten universities and this is driven by relatively few region-years with high numbers of universities (the median number of universities is only 2). In our robustness, we show that our results are not sensitive to dropping region-years with no universities. Similarly, the minimum and median growth rate in the number of universities is zero. In fact this is the case for 5,736 of the observations implying that the majority of region-years in our sample do not experience an increase in university numbers. As we set out in the next section, our core regressions will control for population levels and growth, and a number of geographic characteristics – including an indicator for whether a region contains a country's capital (this is the case for 5 per cent of the observations). Measures of regional human capital (college share and years of education) are available for sub-samples of region-years. In those samples, the average region-year has a college share of seven per cent and average years of education of just over seven.

III. EMPIRICAL FRAMEWORK

The underlying model we are interested in is the long run relationship between universities and economic performance:

$$(1) \qquad ln(Y/L)_{ic,t} = \alpha_1 \, ln(Uni_{ic,t-5}) + X'_{ic,t-5}\alpha_2 + v_{ic,t}$$

Where $(Y/L)_{ic,t}$ is the level of GDP per capita for region i, in country c, and year t and $Uni_{ic,t-5}$

is the 1 + the lagged number of universities in the region (so that we include observations where there are no universities in our analysis). We lag this to allow for the effect that the impact of universities is unlikely to be immediate and since we estimate in 5 year differences, using the fifth lag is natural (we also show longer distributed lags). In addition, using the lag means that we eliminate the effects of a contemporaneous demand shock that raises GDP per capita and also results in the opening of new universities. We also need to control for a number of observables X, that may be related to GDP per capita growth and also the growth in universities; in particular the population of the region, pop. $\upsilon_{ic,t}$ is the error term.

α_1 is interpreted as an elasticity: the percentage change in GDP per capita following a percentage change in the number of universities. This structure implies that the effects of adding one university to a region with relatively few universities (representing a large increase in percentage terms) will be larger in absolute terms than the effect of adding one university to a region with many universities: hence we are capturing diminishing marginal returns in university expansion.

The cross sectional relationship is likely to be confounded by unobservable region-specific effects. To tackle this we estimate the model in long (five-year) differences to sweep out the fixed effects. Our main estimating model is therefore:

$$(2) \qquad \Delta ln(Y/L)_{ic,t} = \alpha_1 \Delta\, lnUni_{ic,t-5} + X'_{ic,t-5}\alpha_2 + \alpha_3\, ln(Y/L)_{ic,t-5} + \alpha_4 \Delta\, ln\, pop_{ic,t-5} + \eta_c + \tau_t + \varepsilon_{ic,t}$$

We control for the lagged level of GDP per capita in the region; $ln(Y/L)_{ic,t-5}$ to allow for catch up (we expect α_3 to be negative). In the controls we include the lagged level of population, country level GDP per capita and the region specific time invariant controls. We control for the lagged growth in population because an increase in universities may simply reflect a greater demand for universities due to population growth. We do not initially include any other measure of human capital in these specifications, so that we can capture the total effect that universities have on growth. However, we explore the effect of adding human capital when we try to pin down the mechanism through which universities impact on growth. Finally, we include country fixed effects (η_c) to allow for country-specific time trends, time dummies (τ_t) and an error term $\varepsilon_{ic,t}$. Standard errors are clustered at the regional level. In robustness tests, we also estimate models where we include a full regional dummies and so allow for unobservable regional *trends*.

We also explore the extent to which GDP per capita growth in region *i* may be affected by growth of universities in *other* regions within the same country. We extend our estimating equation (2) to include the growth of universities in other regions, which may be the nearest region (*j*) or simply all other regions in the country (-*i*). Therefore, we include the growth in region *i*'s own universities ($\Delta lnUni_{ic,t-5}$) as well as a potential spillover effect from universities located in in neighboring regions ($\Delta lnUni_{jc,t-5}$), see equation (3). Lagged population and population growth in region *j* are in the controls, $X'_{jc,t-5}$.

(3) $$\Delta ln\left(Y/L_{ic,t}\right) = \theta_1 \Delta\, ln\left(Uni_{ic,t-5}\right) + \theta_2 \Delta ln\left(Uni_{jc,t-5}\right) + X'_{ic,t-5}\theta_3 + X'_{jc,t-5}\theta_4 + \eta_c + \tau_t + u_{ic,t}$$

We allow for spatial variation by interacting university growth with the distance between region *i* and its nearest region, and control for this distance separately, see equation (4). The variable $dist_j$ is the distance in kilometers between the centroids of regions *i* and *j*.

(4) $$\Delta ln\left(Y/L_{ic,t}\right) = \phi_1 \Delta\, lnUni_{ic,t-5} + \phi_2 \Delta lnUni_{jc,t-5} + \phi_3 dist_j * \Delta lnUni_{jc,t-5} + \phi_4 dist_j + X'_{ic,t-5}\phi_5 + X'_{jc,t-5}\phi_6 + \eta_c + \tau_t + u_{ic,t}$$

In the limit, if the nearest region were very close to region *i*, so that $dist_j$ = 0, we would expect that the effect of university growth in region *j* to be close to the effect of university growth in region *i,* so ϕ_2 should be close to ϕ_1. More generally, we would expect that ϕ_3 would be negative so that the effect of region *j* gets smaller as distance increases.

IV. RESULTS

IV.1 Basic Results

Table 3 presents our basic regressions. Column (1) is a simple correlation between the lagged growth rate of universities an regional GDP per capita, with no other controls. The estimated coefficient is about 0.05 and highly significant. To control for the fact that populous regions are more likely to require more universities, we add the lagged level of the population in column (2) which lowers the university coefficient. Adding country and year fixed effects raises the university coefficient column (3). In column (4) we add lagged regional GDP per capita, growth in population, and the regional covariates (latitude, inverse distance to the coast, malaria ecology, and the natural log of oil and gas production (1950-2010)) and a dummy for regions that contain a capital city. These combined raise the coefficient to 0.0458. In column (5) we control for country-level income with little effect. Columns (6) and (7) replicate (4) and (5) but include regional fixed effects, a very demanding specification which allows for regional trends. These do not much affect the university coefficient and in fact it is higher at 0.0468 in the most general specification. Overall, these results robustly suggest that on average, a doubling of universities in a region is associated with over 4% to 5% higher GDP per person.[13]

The other variables in the regressions take the expected signs. The coefficient on the regional convergence term is nearly 2% in columns (4) and (5), similar to Gennaoili et al (2014).[14] Country GDP per capita has a negative coefficient in these specifications. This becomes a positive relationship once regional fixed effects are included. Having a capital city in a region is associated with around one percentage point higher regional GDP per capita growth. The geographic controls generally have the expected signs (see Table A1 in the Appendix).

We explore different distributed lag structures, but find that a single five year lag is a reasonable summary of the data (Table A2)[15]. This is likely to be due to the fact that in longer time frames, there are more factors at play which are not captured in our estimation framework.

[13] Our analysis is carried out on a sample that drops 54 observations from China pre 1970, before and during the Cultural Revolution, when universities were shut down. Our effects survive if these observations are included, with the coefficient on university growth becoming 0.0308, still significant at the 1% level. We drop them because of the unique nature of this historical episode and the fact that this small number of observations (less than 1% of the full sample) seem to have a large effect on the coefficient.

[14] In the fixed effects specifications (7) and (8) this is larger, consistent with the bias towards faster convergence terms in short panels (see Barro (2012) for discussion of the Hurwiscz (1950) bias in growth regressions).

[15] There is some evidence that contemporaneous and 10 year lagged university growth has a positive significant effect in more advanced economies, but these results are not systematic. Interestingly, the contemporaneous (unlagged) effect of

IV.2 Robustness

Specification checks

We subject our core specification to a large number of robustness tests. Since our regional trends regression (e.g. column (7)) gives a very similar result to the covariates specification without trends we proceed with column (5) as our preferred and less-saturated model. In Table 4, we carry out a series of robustness tests on on this baseline specifiation, which is replicated in column (1). In column (2) we conservatively cluster standard errors at the country level, to account for correlation between the errors of regions within the same country. While the standard errors rise a little, the association between lagged university growth and GDP per capita growth is still significant at the 1% level[16]. In column (3) we weight the regression by the region's population as a share of total country population. Again, this weighting has little effect on the university coefficient. In column (4) we control for the current (as well as lagged) change in population to address the concern that the effect of the university is simply to pull in more people to the region, who spend or produce more and hence raise GDP per capita growth. Our university effect remains strong and therefore it does not appear to be driven by population growth. Column (5) drops regions which never have a university in the sample period and the coefficient remains unchanged. To make sure our results are not driven by extreme university growth observations we do two things. Column (6) drops region-year observations where a region opens its first university, and this actually strengthens the coefficient.[17] In addition, we winsorize the top and bottom five per cent of university growth which again strengthens the result, as shown in column (7). Column (8) uses similarly winsorized GDP per capita growth as the dependent variable, which dampens reduces our coefficient slightly but it still significant at the 1% level. In column (9) we include a dummy for regions where more than five percent of the universities have missing founding dates, reflecting the fact that university counts in those regions will be worse measured

---

university growth is zero or negative (though not particularly significant), suggesting that it takes some time for benefits to be felt, while presumably some costs are incurred at the regional level.

[16] In further analysis, not reported here, we use Driscoll-Kraay standard errors to allow for cross-sectional dependence in a panel (this is done in Gennaioli et al, 2014). This does not affect our results, and we report clustering at the country level which is a more conservative specification as it allows the standard errors in one region-year to be correlated with standard errors in any other region-year within the same country.

[17] In further robustness checks (available on request) we also explore if there are any heterogeneous effects for regions opening their first universities by including a dummy to indicate this, and interacting it with the university growth variable. The coefficients on both the dummy and interaction term are not significantly different from zero. Simply dropping all observations where the lagged university count is zero also does not affect our result.

compared to elsewhere, which has little effect. Finally, we explore whether the definition of university in WHED (i.e. only institutions that offer four year courses or postgraduate degrees) may be a problem, in the sense that there may be some countries that have a larger share of institutions outside this category which could be important for growth. For this purpose, we compare the most recent university numbers in our database to an external source, Webometrics. [18] Column (10) shows that our results are robust to dropping the 29 countries where there are more than double the institutions in Webometrics compared to the WHED listing.

Some further robustness tests are reported in the online Appendix. To investigate the potential concern that our results are driven by expectations of growth in the region we explore "Granger Causality" style tests. We use the growth in universities as the dependent variable and regress this on the lagged growth in regional GDP per capita, and the other controls (Table A3). We see that even as all controls are added, the lagged growth of regional GDP per capita has no relationship with current growth in unviersities and does not appear to "Granger cause" the opening up of univeristies.

The university effect is robust to different specification and sample selection. We subject the regression including regional trends (Table 3, column (7)) to the same robustness tests that we carried out for our baseline regressions, and find that the effect remains positive, significant and of similar order of magnitude (Table A4). We show that the university effect exists across continents (Table A5).[19] Finally, we find that lagged university levels have positive significant effects on growth in "Barro-style" regressions (see Tables A6 and A7, and discussion in the online Appendix).

Heterogeneity between universities

A concern with our econometric strategy is that our use of university numbers is not a perfect measure of university presence. Universities are not homogeneous, but vary in size and quality. Clearly, both

---

[18] http://www.webometrics.info/en/node/54

[19] We split the sample into Europe and North America, Latin America Asia (including Australia) and Africa. In further analysis, available on request, we examine more detailed splits of the data. Looking at North America alone, the coefficient seems low – we find that the effects are stronger for the US than Canada, and also dampened by low density regions. Weighting the regression by population, the coefficient for the US rises to 0.115, and becomes significant at the 5 per cent level. Further investigation of Europe reveals that the coefficient for the UK, France and West Germany is 0.0395 and signficant at the 5 per cent level, but that this is diluted by the rest of Europe (in particular Eastern Europe and Scandinavia) when all countries are aggregated. We also find that the significant effects observed in Latin America and Asia are not driven by any one particular country.

of these dimensions are likely to matter in terms of economic impact (although it is not obvious why this would generate any *upwards* bias in our estimates).

Using the crude count of universities may not be useful if the distribution of university size is uneven across regions. It is reasonable to expect that larger universities will have a larger economic impact; and thus size may be a better measure of university presence. For our university numbers to be an adequate measure of university presence, we would want them to be correlated with total enrolments at the region level, and the same with their growth rates. If, for example, some regions had a multitude of small universities and others had only a handful of large ones, we would not be comparing like for like. Ideally we would want to be able to run our regressions using changes in enrolments over time. Unfortunately, an international time series of regional enrolments is not available, but we focus on the United States where state level enrolments dating from 1970 are published by the National Center for Education Statistics (NCES)[20]. We find that university numbers and total enrolments are highly correlated (around 0.9 for any year over the period 1970-2010) that there is a strong positive relationship between the growth in universities and growth in students between 1970-2010 (Figure 7). This gives us some reassurance that the number of universities is a reasonable measure of university presence at the state level. We also include growth in enrolments in our growth regressions (results available on request), finding that the coefficient on enrolment growth is positive and of the same order of magnitude as our main university effect.[21] These correlations and regressions give us some comfort that the use of university numbers to reflect university presence is reasonable. Moreover, we consider that the number of universities is a better "supply side" measure of university presence in the sense that enrolments are more likely to reflect endogenous demand from students over time.

Another potential weakness in using the number of universities as our measure of university presence is that universities differ in their quality. Ideally to measure quality we would like to have global rankings for all our institutions, carried out annually throughout our sample period. However, university rankings tables only tend to cover the top few hundred institutions in the world, and tend

[20] See http://nces.ed.gov/fastfacts/display.asp?id=98.
[21] The effect of enrolments is positive but insignificant. This analysis is carried out on the 1970-2010 subsample of data, for which even our main university effect loses significance in the US.

to be available for recent years (for example the Shanghai Rankings have been compiled since 2003 and cover the world's top 500 universities).

Our data do contain some key attributes of universities which may be indicative of quality, specifically whether a university is a research institution which is more likely to have effects on innovation (we take whether or not a university is PhD granting as an indicator of research activity), whether it is public or private, whether it offers science, technology, engineering or mathematics (STEM) subjects, and whether it offers professional service related courses (business, economics, law, accounting, finance). We add these variables to the analysis by considering the effect of the growth in the share of each type of university over and above the growth in the number of all universities[22].

Table 5 examines whether these various proxies for university quality have differential effects on GDP growth. Here, we drop region-year observations with zero universities (where the share would be meaningless). Panel A shows the result for the full sample of countries. Each column includes one of these measures in turn. The effects are not significantly different from zero, suggesting that on the entire sample there seems to be a general university effect which does not vary much by type of university as defined here. We also perform this quality analysis on the more advanced economies of Western Europe (including Scandinavia) and the US. In Panel B we can see that now increases in the share of PhD granting institutions are significant though the other measures are not, suggesting that the research channel may be more important in countries at the technology "frontier" (Aghion et al., 2005). Growth in the shares of public, STEM subject and professional subject universities all have no impact on growth, though this may be due to measurement issues.[23] Equivalent analysis of the sample of all other countries (Panel C) shows no effect for any of the quality measures.

Summary on robustness

[22] We note that these characteristics apply to the universities' status in 2010. In the absence of a full time series of when universities begin to offer different courses or qualifications, we simply assume that these characteristics apply since the universities were founded.

[23] The way we ascertain subjects offered by each university is by extracting key relevant words from the information provided in WHED. For some universities the descriptions offered can be quite broad (e.g. it may specify "social sciences" instead of listing out individual subjects). We try to keep our STEM and professional course categories broad to account for this, but there are likely to be cases where we do not pick up the accurate subject mix at a university.

We have shown that our results are robust to different specification and, to the extent that the data allow, consideration of the size and quality dimensions. However, this framework does not allow us to address potential endogeneity due to time-varying unobservables. It could be that our results are driven by factors that vary at the region – year level (for example, it could be that some regions have good local governments in certain time periods who implement a number of policies that are growth enhancing, one of which happens to be opening up universities). There is no direct way to address this in our panel set-up since our variable of interest varies at the region-year level.

IV.3 Country-wide effects of universities

If the effects we are finding are real we would expect to see that universities do not just affect the region in which they are located, but also nearby regions, or even the whole country. Therefore, in addition to including own region university growth we include university growth in *other* regions to see whether this affect GDP per capita in our home region. Table 6 presents the results of this spillover analysis. Column (1) replicates our baseline result. Column (2) includes lagged university growth in the nearest region. This shows that universities in the nearest region have a positive but insignificant effect on home region growth. However, on closer inspection it appears that some "nearest regions" are actually very distant, and are dampening this result. In column (3) we drop observations where the nearest region is over 200km away (which constitute a fifth of all regions in the sample). In this column the nearest regions have an effect of nearly the same magnitude as the home region, though with larger standard errors. Therefore, using the full sample again in column (4), we control for the growth in universities in the nearest region interacted with the distance to that region (based on distance between centroids), and a linear term in distance. Consistent with column (3), the interaction is negative and significant and the linear "neighbouring universities" term is positive and significant. In column (5) we add the relevant controls for the neighbouring region – the lagged population and population growth (which should also control for a demand shock in the neighbouring region in the previous period). These have little effect on our coefficients or their significance. The magnitude of the effects is sensible: universities in a neighbour that has a distance of near zero has essentially the same effect on growth as a university in the region of interest (0.0430 vs. 0.0356).

Finally, we look at the effects of university growth in all other regions (including nearest region) in the country on our home region. Column (6) adds the lagged growth in universities in all regions of a country, excluding the home region. Column (7) also adds the relevant controls (lagged population and population growth for the other regions). These effects are now larger than our main effect and again highly significant[24]. The implication is that a doubling of universities in the rest of the country (which in most cases will represent a greater absolute increase than a doubling of home region universities) will on average increase home region's GDP per capita by nearly 6 per cent. Overall, this analysis suggests that universities not only affect the region in which they are built, but also their neighbours and that there does appear to be a spatial dimension to this, in the sense that the closer regions are geographically, the stronger the effects. This suggests that from a country perspective, universities generate local and macro growth. Therefore, the full effect of opening universities on a country must be greater than what we find in our core regressions.

Using the coefficients in Table 6 we can estimate a country-wide effect of a university expansion on the typical region in our dataset. The average region has nearly 10 universities (see Table 2), and the average country has 20 regions (and therefore 200 universities). Doubling the universities in one region (10 to 20) implies a 4% uplift to its GDP per capita according to our main result. For each other region, this represents a 5.3% increase in universities in the rest of the country (190 to 200), multiplied by 6% (the coefficient on other regions in column (7)), this implies an uplift to all other regions' GDP per capita of 0.3%. Assuming the regions in this hypothetical country are identical, the uplift to country-wide GDP per capita is simply the average of these effects: 0.5%.[25] While this seems like a significant amount, we also need to the costs of university expansion.[26]

Given that the costs of building and maintaining universities will vary widely by country, we choose to focus on a particular institutional setting for this calculation. In the UK in 2010, there were 173 universities across its 10 regions. As an experiment we add one university to each region, a total

---

[24] Standard errors in this analysis are clustered at the region level. Conservatively clustering at the country level does not affect the significance of the nearest region analysis at all. The coefficients on growth in all other regions (columns (6) and (7)) remain significant at the 10% level.

[25] As a sense check for this result, we collapse our regional dataset to the country level and run macro regressions of GDP per capita growth on lagged university growth. The coefficient on university growth is 0.03 (but insignificant). According to these results, an 100% increase in universities at the country level would be associated with a 3% increase in GDP per capita growth. Therefore a 6% increase in universities would imply a 0.2% uplift – this is smaller than the 0.5% we calculate using the results from our better identified regional analysis, but of the same order of magnitude.

[26] It is unlikely that these are controlled for in our regressions: university financing tends to be at the national level, and costs are incurred on an ongoing basis (e.g. property rental or amortisation and staff salaries are incurred every year) and so would not be fully captured by the inclusion of lagged country GDP per capita as a covariate.

increase of 10 universities (6%) at the country level. Using similar steps as in our hypothetical country above (but taking into account the actual numbers of universities in each UK region in 2010), we calculate that the overall increase to GDP per capita (or GDP, assuming population is held constant) is 0.7%.[27] Applied to UK GDP in 2010 (£1,614 billion according to the ONS[28]) this comes to an additional £11.3 billion per year. A crude approximation of the annual costs associated with a university can be made based on university finance data: in 2009-2010 the average expenditure per institution in the UK was around £160 million.[29] Multiplying this by the 10 universities in our experiment, the implied annual cost for the additional universities is £1.6bn, or 0.1% of GDP. While this calculation is highly simplified, it shows that there is a large margin between the potential benefits of university expansion implied by our regression results and likely costs. We note that the costs of setting up universities, and methods of university finance vary by country so we cannot generalize this result to other countries, nor make statements about the optimal number of universities in particular regions. Similar calculations can be made by delving into particular institutional settings.

V. Mechanisms

Having established a robust association of universities on GDP per capita we now turn to trying to understand the mechanisms through which universities may affect growth.

V.1 Human Capital

---

[27] For each of the ten regions in the UK in turn, we calculate the percentage increase implied by adding one university to that region's universities, and multiply this by the 4% (effect of a 100% increase). We then calculate the increase in the count of universities in all other regions (for that region), and raise their GDP per capita by that percentage increase multiplied by 6% (the effect of 100% growth of universities in "other regions" on home region GDP per capita from Table 6). We abstract from the 5 year lag in this calculation. We the add up the total GDP across regions, and divide by total population (assumed unchanged).

[28] Series ABMI, Gross Domestic Product: chained volume measures: Seasonally adjusted £m, Base period 2012

[29] Data on university finance, by institution, can be found at the UK Higher Education Statistical Authority (HESA) website (https://www.hesa.ac.uk/index.php?option=com_content&view=article&id=1900&Itemid=634). Total expenditure in the year 2009/10 was nearly £26 billion across 164 institutions listed in HESA, implying around £160 million per institution. University expenditure contains staff costs, other operating expenses, depreciation, interest and other finance costs. We checked if this figure has been relatively stable over time, finding that by 2013-14, average expenditure was £180 million. At this higher amount, the implied costs of our expansion rise to 0.11% of GDP. Note that the number of institutions present in 2010 was 173. The majority of institutions in WHED (156) correspond to those listed in HESA, but there are a small number of discrepancies due to differences in the classifications of some institutes or colleges between the two listings. This does not matter for our purposes, as are simply using the HESA data to calculate the average expenditure of a typical university.

We add measures of growth in human capital to our regressions to see how this influences the university coefficient. In Table 7 we consider the relationship between universities and college share (Table A8 shows a similar exercise for years of education). Column (1) replicates the core result, and column (2) the same specfication on the reduced sample where college share is non-missing, for which the university effect is a bit stronger at over 0.07. Column (3) adds the lagged growth in college share which in itself is highly significant, and reduces the university effect from 0.0711 to 0.0673, a reduction of five per cent. Column (4) uses contemporaneous growth in college share and column (5) adds in the lagged college share. In column (6) we include also the level with both lags, with little change in the university coefficient. In column (7) we look at the raw correlation between contemporaneous growth and the lagged growth in universities (with only country fixed effects as controls), and find it to be relatively small but highly significant. Adding all the other controls dampens this relationship further and this small effect of university growth on college share within a short time frame is what explains the fact that adding in growth in human capital makes only a small dent in the effect of universities. This analysis suggests that a 1% rise in universities gives rise to around a 0.4 percentage point rise in college share per annum in the next five year period, all else equal. It takes a while for a new university to affect the stocks of educated people in a region, since each year, the maximum direct effect it could have would be if all its graduates stayed and worked in the region. Of course there could be indirect effects too, if the university attracts skilled workers to the region to work in connected research institutes or firms.[30]

The effects of universities on human capital may appear small, but is unsurprising given the experiment we are using: it takes time for the stock of human capital in a region to build up. We construct a simple hypothetical experiment where we examine the effects of a new university assumed to have a full capacity 5,000 students, who all remain in the region to work post graduation (see the online Appendix). We explore the expected effects on both college share and years of education and show these are not so different to the regression results.

V.2 Innovation

[30] Table A8 in the Appendix uses another measure of human capital: years of education, which is available for a larger sample of countries and years. The messages are similar, including years of education does not dent the effect of universities on growth much, and this is largely due to the small correlation between the lagged growth in universities and growth in years of education over a five year period.

The best available measure of innovation output available consistently at the regional level over time is patents, though unfortunately patents with locational information are not available for our entire sample of countries or years. We consider the effects of adding growth in cumulative patent "stocks" [31] to our regressions, first for the US (where we use US PTO registered patents over the period 1965-1999), and then internationally (we use patents filed at the European Patent Office which are available for over 38 of our countries between 1975 and 2005). We note that these may not be a consistent measure of innovative activity across all countries in our sample, since for example, some inventors in the US may be less concerned about European markets.

Table 8, Panel A shows the results for the US. Column (1) runs the core regression for the US only, and over the time period that we have patents data.[32] Column (2) then includes the change in patents stock, which reduces the coefficient on university growth from 0.113 to 0.109 (a reduction of four per cent). Patents themselves have a positive impact on GDP per capita growth, but this is not significant. Column (3) considers the raw correlation between lagged university growth and current patent stock growth (including only year dummies), and shows it is positive and significant at the one per cent level. Column (4) then adds the standard controls, and this correlation becomes substantially smaller and insignificant. This analysis suggests that at least for the US, innovation is part of the story of why universities have an economic impact, though not the entire story. This may be because the effect of newer universities on patents takes a while to accumulate.

On a wider sample of countries, we also consider EPO patents (Panel B). Column (1) restricts the regression to the sample for which we have the EPO patents data giving us a university effect of 0.0220, signficant at the 10 per cent level. Adding in patents in column (2) reduces the effect of universities by around 13 per cent, and we see that growth in patents has a strong association with regional GDP growth in this international sample. A doubling of the patent stock is associated with 5% higher per capita GDP. Columns (4) and (5) shows the raw correlation between lagged university growth and patents growth, and shows that it is postitive, but insignificant.

[31] Patent stocks are simply the cumulative patents registered at the region level, and it is assumed that this is not depreciated. Results are not sensitive to alternative depreciation assumptions.
[32] This regression is weighted by population since we found that the effects of low density outliers were dampening the effects in the US.

V.3 Institutions and democracy

The use of country fixed effects throughout this analysis should rule out the possibilty that the effects of universities simply reflect different (time invariant) institutions, since these things tend to differ mainly at the country level. To the extent that institutions vary within countries, say at the US state level, our regional fixed effects analysis should address this.

However, institutions do vary over time, and it is possible that universities contribute to this – albeit over longer time horizons than those analysed in our core regressions. We saw in Figure 5 that there is a positive significant correlation between country level democratic institutions (as proxied by Polity scores) and universities. This correlation also exists when we consider the 1960-2000 change in universities and polity scores (see the online Appendix for more discussion). A time series of data on regional institutions to fit into our growth framework is not available, but we do explore the relationships between perceptions of democracy obtained from the World Values Survey and lagged university presence in the cross section.

Our chosen survey measure is a categorical variable which gives the approval of a democratic system for governing one's own country, as this is more widely available across survey waves compared with other questions on democracy. [33] We note however, that the experience in one's own country (for example, if corruption prevents democracy operating effectively) may affect this judgement. Therefore, in the robustness we test whether results hold for a another more general survey question[34] (available for fewer survey waves). World Values Survey data begins in the 1980s and we pool data into a cross section due to insufficient observations in some region–year cells to generate reliable variation over time.

[33] Specifically, the question asks respondents to say whether having a democratic political system is a (1) very good, (2) fairly good, (3) fairly bad, (4) bad way of governing their country. The scale is reversed for our estimation so that a higher score reflects higher approval.

[34] This question asks respondents if they (1) agree strongly, (2) agree, (3) disagree or (4) strongly disagree whit this statement "Democracy may have problems but is better than any other form of government". Again, the scale is reversed for our estimation so that a higher score reflects higher approval.

Table 9 shows the results of these regressions.[35] We start with a simple correlation between our measure of university density lagged by 15 years from the survey year, controlling for country and year fixed effects and clustering standard errors at the region level (column (1)).[36] This shows a highly signficant association between university presence in a region and approval of a democratic system. The relationship is robust to including a host of individual demographic characteristics such as gender, age, martial status, children, employment and relative incomes (column (2)). In column (3) we see that the result is also robust to including controls for the individual's own education: a dummy for whether or not they hold a university degree, and a dummy indicating whether or not they are a student. The result that one's own education is positively related to approval of democracy is consistent with Chong and Gradstein (2009) who use years of schooling. But the result that universities matter over and above their effect on an individual's education suggests that they may be a mechanism whereby democratic ideals spillover from those who have direct contact with universities, or that there is some form of direct diffusion from universities into their surrounding region. Further supporting this, we find that the result survives dropping students and graduates from the regression entirely (see Table A9 more robustness tests and further discussion in the online Appendix). Column (4) adds our standard geographic controls and shows that these actually strengthen our result.

This analysis shows that there is robust relationship between other lengths of lagged university presence in a region and approval of a democratic system, and that this operates over and above the human capital effect. While it is not possible to account for any potential impact of this type of mechanism on growth in our current framework, this analysis suggests that institutions could be part of the story, albeit on a longer term basis.

V.4 Mechanical demand effect

[35] This analysis is carried out on 58 of the 78 countries in our core sample, where World Values Survey data are available. World Values Survey data are available for Nigeria which is in our core sample, but it was not possible to map the regions to the regions used in WHED due to the fact that both sources used very aggregated but different regions.

[36] We explored different lag structures, and found that it takes time for universities to affect perceptions (see column (3) in Table D1 which shows a smaller positive, but insignificant effect of five year lagged university density on democratic approval). By contrast, on the full sample of countries there appear to be no effects for longer lags. When we consider the sub-sample of OECD countries where the results are stronger we see that the effects are similar in magnitude and significance for the 30 year lag.

By including changes in population in our regressions (lagged, and contemporaneous in the robustness tests), we believe that we have largely ruled out the possibility that universities simply contribute to growth through a mechanical demand channel; that the students and staff in a university, and the university itself simply consume more goods and services. Another angle to assess this possibility is to continue with the hypothetical experiment we used to simulate the expected effects of universities on human capital stocks. We use the same average region in the dataset in 1990, when average regional GDP per capita was $12,402. Again, we assume that our new university has a student capacity of 5,000. Now we account for staff – assuming staff-student ratio of 20 (so that there are 250 faculty and staff at any given time). We generously assume that students (and graduates when they finish their degree) and staff are all 50% richer than the average person in the region. Then we look at what happens over time as more and more students enter and leave the university (assuming they stay within the region), and the composition of GDP per capita in the region changes to reflect this. In fact, this experiment can explain only a small portion of the average annual growth in the following five year period (under one fifth of the effect predicted by our regressions). Given this, and the fact that our regressions are robust to including population growth, we believe it is unlikely that our results are simply driven by a mechanical demand effect.

V.5 Summary on mechanisms

In summary, it appears over the five year periods in our analysis, some of the effects of university growth on GDP growth work via human capital and innovation channels, though the effects of these are small in magnitude. In addition, universities may affect views on democracy but this appears to be on a longer term basis. We have shown convincingly that the university effect is not merely driven by demand from the increased population associated with the university.

VI. CONCLUSIONS

This paper presents a new dataset on universities in over 1,500 regions in 78 countries since 1950. We have found robust evidence that increases in university presence are positively associated with faster subsequent economic growth. Doubling the number of universities is associated with over 4% higher GDP per capita in a region. This is even after controlling for regional fixed effects, regional trends and a host of other confounding influences. The benefit of universities is not confined to the region where they are built but “spills over” to neighboring regions, having the strongest effects on those that are geographically closest. Using these results, we estimate that the economic benefits of university expansion are likely to exceed their costs.

Our estimates use within country time series variation and imply smaller effects of universities on GDP than would be suggested from cross sectional relationships (three quarters smaller). But we believe our effects are likely to be lower bounds as the long-run effect of universities through building the stock of human and intellectual capital may be hard to fully tease out using the panel data available to us. Nevertheless, the evidence seems compelling here that there is some effect of universities on growth.

Understanding the mechanisms through which the university effects works is an important area to investigate further. We find a role for innovation and human capital supply (but not demand), although these are small in magnitude. This might be due to econometric problems, but better data on the flow of business-university linkages, movements of personnel and other collaborations would help in unravelling the mechanisms. In addition, focusing on the relationships between universities and local economic performance in individual countries where better data on university size, subjects offered and quality over time are available may provide interesting results.

We provide suggestive evidence that universities play a role in promoting democracy, and that this operates over and above their effect as human capital producers. Understanding the extent to which this may account for part of the growth effect is another area for future research.

REFERENCES

Acemoglu, Daron, Simon Johnson and James A. Robinson (2005a) “Institutions as a Fundamental Cause of Development,” in *Handbook of Economic Growth*, edited by Philippe Aghion and Steven Durlauf, Amsterdam; North-Holland.

Acemoglu, Daron, Simon Johnson, James A. Robinson and Pierre Yared (2005b) “From education to democracy?,” *The American Economic Review*, 95, No.2, Papers and Proceedings of the One Hundred Seventeenth Annual Meeting of the American Economic Association, Philadelphia, PA, January 7-9, 2005 (May, 2005), pp.44-49.

Acemoglu, Daron, Suresh Naidu, Pascual Restrepo and James A. Robinson (2014) “Democracy does cause growth”, *NBER Working Paper No 20004*, March 2014

Adams, James D. (2002), “Comparative Localization of Academic and Industrial Spillover,” *Journal of Economic Geography* 2, 253-78.

Aghion, Philippe, Leah Boustan, Caroline Hoxby, and Jerome Vandenbussche (2009), ‘‘The Causal Impact of Education on Economic Growth: Evidence from U.S.,’’ *Harvard University working paper*

Aghion, Philippe, Meghir, Costas and Vandenbussche, Jérôme (2005), Growth, Distance to Frontier and Composition of Human Capital, No 4860, *CEPR Discussion Papers*, C.E.P.R. Discussion Papers.

Bairoch, Paul (1982), "International industrialisation levels from 1750 to 1980", *The Journal of European History*, 2, pp. 269-333.

Baldwin, Richard (2012), Global Supply Chains: Why they emerged, why they matter, and where are they going? *CEPR Discussion Paper No.9103*, August 2012

Barro, Robert J. (1991) “Economic Growth in a Cross Section of Countries,” *The Quarterly Journal of Economics*, Vol. 106, No. 2, pp. 407-443.

Barro, Robert J. (1999). “Determinants of democracy”, *Journal of Political Economy*, 107, 158–183.

Barro, Robert J. (2012) “Convergence and Modernization Revisited,” *NBER Working Paper No. 18295*.

Belenzon, Sharon and Schankerman, Mark (2013) “Spreading the word: geography, policy and knowledge spillovers”. *Review of Economics and Statistics*, 95 (3). pp. 884-903. ISSN 0034-6535

Besley, Tim and Van Reenen, John (2013), “Investing in Prosperity, A Manifesto for Growth”, London School of Economics

Bils, Mark and Klenow, Peter (2000), "Does Schooling Cause Growth?", *American Economic Review*, 90, 1160-1183

Bloom, Nick, Sadun, Raffaella and Van Reenen, John (2015), "Management as a Technology", London School of Economics, Mimeo

Cantoni, Davide and Yuchtman, Noam (2014) "Medieval universities, legal institutions, and the Commercial Revolution", *The Quarterly Journal of Economics*, 823–887.

Card, David (2001) "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems", *Econometrica*, Econometric Society, vol. 69(5), pages 1127-60, September.

Caselli, Francesco (2005) "Accounting for Cross-Country Income Differences" in Philippe Aghion and Steven Durlauf (ed.) *Handbook of Economic Growth I*, 1(9) pages 679-741 Amsterdam: Elsevier.

Chong, Alberto, and Mark Gradstein (2009) "Education and Democratic Preferences", Inter-American Development Bank, Research Department Working Papers; 684, June 2009

Cohen, Daniel, and Marcelo Soto (2007), ''Growth and Human Capital: Good Data, Good Results,'' *Journal of Economics Growth*, 12, no. 1 51–76.

de La Fuente, Angel, and Rafael Domenech (2006), ''Human Capital in Growth Regressions: How Much Difference Does Data Quality Make?'' *Journal of the European Economics Association*, 4, no. 1, 1–36.

De Ridder-Symoens, Hilde, (1992) "A History of the University in Europe", Cambridge University Press, 16 Oct 2003

Feng, Andy, and Valero, Anna (2016), "Skill biased management: Evidence from Manufacturing firms", Mimeo

Gennaioli, Nicola, La Porta, Rafael, Lopez de Silanes, Florencio, Shleifer, Andrei (2014) "Growth in Regions", *Journal of Economic Growth,* 19:259-309

Gennaioli, Nicola, La Porta, Rafael, Lopez de Silanes, Florencio, Shleifer, Andrei (2013) "Human Capital and Regional Development", *Quarterly Journal of Economics*, 128(1) 105-164

Gerring, John, Philip Bond, William Barndt, and Carola Moreno (2005) "Democracy and Growth: A Historical Perspective," *World Politics*, 57 (3), 323-64

Glaeser, Edward L., LaPorta, Rafael, Lopez-de-Silanes, Florencio, and Shleifer, Andrei (2004) "Do institutions cause growth?", *Journal of Economic Growth*, 9, 271–303.

Glaeser Edward L., Ponzetto Giacomo A.M., Shleifer, Andrei (2007), "Why does democracy need education?", *Journal of Economic Growth*.

Goldin, Claudia and Lawrence Katz (2008), *The Race between Education and Technology,* Harvard: Harvard University Press

Hall, Robert E. and Charles I. Jones (1999) “Why Do Some Countries Produce so Much More Output per Worker than Others?,” *Quarterly Journal of Economics*, CXIV, 83-116.

Hall, Bronwyn H., Jaffe Adam B. and Trajtenberg, Manuel (2001), The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools, *NBER Working Paper 8498*, National Bureau of Economic Research, Cambridge, MA.

Hanushek, Eric A. and Woessmann, Ludger (2009), “Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation”, *NBER Working Paper No. 14633*, National Bureau of Economic Research, Cambridge, MA.

Hausman, Naomi, (2012), “University Innovation, Local Economic Growth, and Entrepreneurship”, US Census Bureau Center for Economic Studies Paper No. CES-WP- 12-10. Available at SSRN: http://ssrn.com/abstract=2097842 or http://dx.doi.org/10.2139/ssrn.2097842

Hurwicz, Leonid. (1950), “Least-Squares Bias in Time Series,” in Tjalling C. Koopmans, ed., *Statistical Inference in Dynamic Economic Models*, Wiley, New York.

Jaffe, Adam B, (1989) "Real Effects of Academic Research," *American Economic Review*, American Economic Association, vol. 79(5), pages 957-70, December.

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson (1993), “Geographic Localization of Knowledge Spillovers As Evidenced By Patent Citations,” *Quarterly Journal of Economics* 108, 577-98.

Jäeger, Simon (2013): “Labour market returns to higher education, Evidence from University Openings in Germany”

Kantor, Shawn, Whalley, Alexander (2014a), “Knowledge spillovers from research universities: evidence from endowment value shocks”, *Review of Economics and Statistics*, 96 (1) (2014), pp. 171 – 188.

Kantor, Shawn, Whalley, Alexander (2014b), “Research Proximity and Productivity: Long-Term Evidence from Agriculture”, Mimeo.

Mankiw, N.G., Romer, D. and Weil, D.N. (1992) “A Contribution to the Empirics of Economic Growth”, *The Quarterly Journal of Economics*, 107 (2) 407-437.

Mokyr, Joel, (2002), “The Gifts of Athena: Historical Origins of the Knowledge Economy”, Princeton University Press

Moretti, Enrico, (2004), “Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data”, *Journal of Econometrics*, Elsevier, vol. 121(1-2), pages 175-212

Mankiw, N. Gregory, David Romer, and David Weil (1992), “A Contribution to the Empirics of Economic Growth,” *Quarterly Journal of Economics* 107 (2): 407-438.

Nickell, Stephen (1981), “Biases in Dynamic Models with Fixed Effects,” *Econometrica,* 49 (5): 1417-1426.

Papaioannou, Elias and Siourounis, Gregorios (2008), “Economic and social factors driving the third wave of democratization”, *Journal of Comparative Economics*, Volume 36, Issue 3, Pages 365 – 387

Robbins, Lionel (1963). “Higher Education”, Report of the Committee appointed by the Prime Minister under the Chairmanship of Lord Robbins, London: Her Majesty's Stationery Office 1963

Schofer, Evan and Meyer, John W. (2005), “The Worldwide Expansion of Higher Education in the Twentieth Century,” *American Sociological Review*, Vol. 70. No. 6 (Dec., 2005), pp.898-920

Squicciarini, Mara and Voigtländer, Nico (2014), “Human Capital and Industrialization: Evidence from the Age of Enlightenment”, *NBER Working Paper No. 20219*, National Bureau of Economic Research, Cambridge, MA.

Toivanen, Otto and Väänänen, Lotta (2015). “Education and Invention”, *The Review of Economics and Statistics*, forthcoming

Table 1: CROSS-SECTIONAL REGRESSIONS

Dependent variable: Regional GDP per capita	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(1 + # universities)	0.681***	0.214***	0.160***	0.0555***	0.159***	0.0789***	0.0566***
	(0.124)	(0.0422)	(0.0388)	(0.0206)	(0.0444)	(0.0255)	(0.0202)
ln(population)	-0.469***	-0.105**	-0.112***	-0.0692***	-0.0947**	-0.0582**	-0.0949***
	(0.0997)	(0.0412)	(0.0325)	(0.0233)	(0.0350)	(0.0225)	(0.0277)
Years of Education				0.292***		0.287***	0.262***
				(0.0280)		(0.0345)	(0.0394)
ln(1 + EPO Patent "stock")							0.0524***
							(0.0124)
Observations	1213	1213	1182	1182	658	658	658
Adjusted R-squared	0.247	0.890	0.904	0.932	0.916	0.947	0.951
# clusters	65	65	62	62	34	34	34
country dummies	no	yes	yes	yes	yes	yes	yes
region controls	no	no	yes	yes	yes	yes	yes

Notes: OLS estimates based on data in 2000. Standard errors clustered at the country level. Column (1) shows the relationship between universities and GDP per capita, controlling for population. Column (2) includes country dummies. Column (3) includes regional controls (a dummy indicating whether the region contains a capital city, together with latitude, inverse distance to ocean, malaria ecology, log(oil and gas production) 1950-2010, these are not reported here). Column (4) includes years of education. Column (5) is identical to column (4) but restricts the sample to the regions for which OECD REGPAT patents are available. Column (6) includes years of education, and column (7) includes the natural log of the regional patent "stock".

Table 2: DESCRIPTIVE STATISTICS, GROWTH ANALYSIS

	Mean	S.D	Min	p50	Max	Obs
Regional GDP per capita	13,055.75	11,958.30	262.15	8,463.02	105,648.25	8,128
Growth in regional GDP per capita	0.02	0.03	-0.20	0.02	0.30	8,128
Country GDP per capita	14,094.16	11,525.30	690.66	9,157.66	64,198.29	8,128
# universities	9.61	23.71	0	2.00	461.00	8,128
Growth in # universities	0.02	0.03	0	0	0.28	8,128
Population (millions)	2.78	7.97	0.01	1.01	196.00	8,128
Growth in population	0.01	0.02	-0.14	0.01	0.25	8,128
Latitude	27.74	25.65	-54.33	37.75	69.95	8,128
Inverse distance to ocean	0.03	0.07	0	0.01	1.89	8,128
Malaria index	0.89	2.31	0	0.01	25.51	8,128
log(oil and gas production) 1950-2010	1.72	2.86	0	0	12.05	8,128
Dummy for capital in region	0.05	0.22	0	0	1.00	8,128
College share	0.07	0.07	0	0.04	0.45	5,744
Years of education	7.37	3.08	0.39	7.42	13.76	6,640

Notes: Each observation is region-year. *Source: WHED and Gennaioli et al (2014) for regional economic data.*

Table 3: GROWTH ON LAGGED GROWTH REGRESSIONS

Dependent variable: Regional Growth of GDP per capita	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Lagged growth in #universities	0.0470***	0.0363***	0.0401***	0.0458***	0.0445***	0.0448***	0.0468***
	(0.0104)	(0.0105)	(0.0112)	(0.0106)	(0.0105)	(0.0118)	(0.0118)
Lagged level of regional GDP per capita				-0.0153***	-0.0127***	-0.0581***	-0.0776***
				(0.00125)	(0.00131)	(0.00359)	(0.00529)
Lagged level of country GDP per capita					-0.0213***		0.0378***
					(0.00422)		(0.00677)
Lagged level of population /100		0.178***	-0.0301	-0.0765*	-0.0855**	-1.095***	-0.850**
		(0.0324)	(0.0351)	(0.0395)	(0.0387)	(0.368)	(0.390)
Lagged growth in population				-0.0987**	-0.113***	-0.209***	-0.183***
				(0.0383)	(0.0385)	(0.0494)	(0.0500)
Dummy for capital in region				0.0125***	0.0110***		
				(0.00170)	(0.00168)		
Observations	8128	8128	8128	8128	8128	8128	8128
Adjusted R-squared	0.003	0.008	0.215	0.253	0.260	0.331	0.347
# clusters	1498	1498	1498	1498	1498	1498	1498
year dummies	no	no	yes	yes	yes	yes	yes
country dummies	no	no	yes	yes	yes	no	no
region controls	no	no	no	yes	yes	no	no
region dummies	no	no	no	no	no	yes	yes

Notes: OLS estimates, 78 countries. Column (1) is a simple correlation between regional GDP per capita growth and the lagged growth in university numbers. Column (2) controls for the lagged log of population. Column (3) includes country and year dummies. Column (4) controls for lagged regional GDP per capita, the lagged growth in population, the lagged log population density level, and lagged growth in average years of education, a dummy for whether the region contains a capital city, together with latitude, inverse distance to ocean, malaria ecology, log(oil and gas production) 1950-2010 (not reported here). Column (5) adds lagged country GDP per capita. Column (6) includes regional fixed effects, and the time varying controls of column (4). Column (7) adds lagged country GDP per capita. Standard errors are clustered at the regional level. Levels of GDP per capita and population are in natural logs.

Table 4: ROBUSTNESS

Dependent variable: Regional Growth of GDP pc	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Lagged growth in #unis	0.0445***	0.0445***	0.0411***	0.0441***	0.0434***	0.0510***		0.0343***	0.0422***	0.0530***
	(0.0105)	(0.0138)	(0.0126)	(0.0104)	(0.0106)	(0.0133)		(0.00775)	(0.0106)	(0.0130)
Lagged growth in #unis, winsorized							0.0615***			
							(0.0147)			
Lagged level regional GDP pc	-0.0127***	-0.0127***	-0.0134***	-0.0124***	-0.0116***	-0.0124***	-0.0127***	-0.00980***	-0.0128***	-0.0125***
	(0.00131)	(0.00218)	(0.00144)	(0.00134)	(0.00135)	(0.00133)	(0.00132)	(0.000929)	(0.00131)	(0.00159)
Lagged level country GDP pc	-0.0213***	-0.0213	-0.0241***	-0.0216***	-0.0223***	-0.0209***	-0.0213***	-0.0222***	-0.0213***	-0.0457***
	(0.00422)	(0.0180)	(0.00458)	(0.00424)	(0.00470)	(0.00430)	(0.00422)	(0.00324)	(0.00422)	(0.00374)
Lagged level of population/100	-0.0855**	-0.0855	0.00264	-0.0933**	-0.107**	-0.0802**	-0.0927**	-0.0779***	-0.105***	-0.127**
	(0.0387)	(0.0548)	(0.0481)	(0.0387)	(0.0435)	(0.0395)	(0.0388)	(0.0298)	(0.0404)	(0.0529)
Lagged growth in population	-0.113***	-0.113**	-0.105**		-0.167***	-0.118***	-0.114***	-0.129***	-0.114***	-0.104***
	(0.0385)	(0.0503)	(0.0427)		(0.0300)	(0.0392)	(0.0385)	(0.0234)	(0.0386)	(0.0365)
Current growth in population				-0.146***						
				(0.0511)						
Dummy for capital in region	0.0110***	0.0110***	0.0102***	0.0110***	0.0102***	0.0106***	0.0110***	0.00923***	0.0110***	0.0129***
	(0.00168)	(0.00210)	(0.00130)	(0.00169)	(0.00160)	(0.00169)	(0.00169)	(0.00143)	(0.00168)	(0.00242)
Observations	8128	8128	8128	8128	6642	7897	8128	8128	8128	5357
Adjusted R-squared	0.260	0.260	0.299	0.261	0.259	0.267	0.260	0.313	0.260	0.259
# clusters	1498	78	1498	1498	1110	1497	1498	1498	1498	939
weights	no	no	yes	no	no	no	no	no	no	no
standard errors clustered at	region	country	region	region	region	region	region	region	region	region

Notes: Column (1) replicates column (5) from Table 3. Column (2) is identical to column (1), but standard errors are clustered at the country level. Column (3) is a weighted regression, using the region's population as a share of total country population. Column (4) uses the current, rather than lagged, growth in population. Column (5) drops regions that have no universities. Column (6) drops regions that got their first university in the last period. Column (7) winsorizes the top 5% and bottom 5% the of observations of regional university growth. Column (8) winsorizes the top 5% and bottom 5% the of observations of regional GDP per capita growth. Column (9) adds a dummy for regions where more than 5 per cent of the universities have missing locational information and are not regionalized. Column (10) omits countries where the total number of universities in WHED in 2010 is less than half of the number according to a recent external source (Webometrics in 2015).

Table 5: EXPLORING UNIVERSITY QUALITY MEASURES

Dependent variable: Regional Growth of GDP per capita	(1)	(2)	(3)	(4)
Panel A: Full sample				
Lagged growth in #universities	0.0513***	0.0520***	0.0498***	0.0488***
	(0.0136)	(0.0135)	(0.0141)	(0.0136)
Lagged growth in PhD share	-0.00129			
	(0.00479)			
Lagged growth in public share		0.00241		
		(0.00444)		
Lagged growth in STEM share			0.00374	
			(0.00339)	
Lagged growth in professional share				0.00489
				(0.00574)
Observations	5612	5612	5612	5612
Panel B: US and "Western Europe"				
Lagged growth in #universities	0.0302**	0.0322**	0.0387**	0.0348**
	(0.0140)	(0.0141)	(0.0150)	(0.0151)
Lagged growth in PhD share	0.0108**			
	(0.00498)			
Lagged growth in public share		0.00472		
		(0.00583)		
Lagged growth in STEM share			-0.00292	
			(0.00433)	
Lagged growth in professional share				-0.000206
				(0.00451)
Observations	1547	1547	1547	1547
Panel C: All other countries				
Lagged growth in #universities	0.0557***	0.0572***	0.0547***	0.0535***
	(0.0159)	(0.0159)	(0.0163)	(0.0158)
Lagged growth in PhD share	-0.00344			
	(0.00622)			
Lagged growth in public share		0.00263		
		(0.00552)		
Lagged growth in STEM share			0.00645	
			(0.00447)	
Lagged growth in professional share				0.00661
				(0.00746)
Observations	4065	4065	4065	4065

Notes: Panel A includes the full sample of countries, and Panel B restricts to the US and "Western Europe" (defined as Austria, Belgium, Switzerland, West Germany, Denmark, Finland, France, UK, Ireland, Netherlands, Norway and Sweden). Within each panel, Column (1) replicates our core regression (Column (5) from Table 3), but drops regions with zero universities. Columns (2) to (5) add in the lagged growth of the shares of universities of different types as labelled.

Table 6: UNIVERSITY SPILLOVERS FROM OTHER REGIONS

Dependent variable: **Regional Growth of GDP per capita**	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Lagged growth in #universities	0.0445***	0.0427***	0.0405***	0.0432***	0.0430***	0.0354***	0.0380***
	(0.0105)	(0.0104)	(0.0116)	(0.0105)	(0.0105)	(0.0104)	(0.0105)
Lagged growth in #universities, nearest region		0.0120	0.0270**	0.0355**	0.0356**		
		(0.0104)	(0.0111)	(0.0139)	(0.0140)		
Lagged growth in #universities X distance to nearest region				-0.177**	-0.180**		
				(0.0760)	(0.0756)		
Distance to nearest region				0.00407**	0.00415**		
				(0.00205)	(0.00205)		
Lagged growth in #universities in other regions within country						0.0556***	0.0570***
						(0.0132)	(0.0133)
Observations	8128	8128	6544	8128	8128	8128	8128
Adjusted R-squared	0.260	0.260	0.288	0.261	0.261	0.262	0.266
# clusters	1498	1498	1257	1498	1498	1498	1498
Nearest / other region controls	no	no	no	no	yes	no	yes

Notes: Column (1) replicates our core regression (column (5) from Table 3). Column (2) adds in the lagged growth in universities in the nearest region. Column (3) replicates column (2) but conditions the sample to regions whose nearest region is less than 200km away. Column (4) returns to the full sample, but adds an interaction term of universities with distance to that region (in km), and distance to that region as a separate variable. Column (5) adds controls from the nearby region: namely the lagged population and population growth (not reported here). There were a small number of observations where the population in the nearest region was missing, relating to early years in the sample period. In this case, population was extrapolated back in time, using a log-linear trend, and a dummy variable included to indicate this. Column (6) includes the lagged growth in universities in all other regions of the country, and column (7) also adds the relevant controls from all other regions in the country: namely the lagged population and population growth (again with a dummy to indicate where the population in the rest of the country has been calculated with missing values for any regions that year).

Table 7: UNIVERSITIES AND COLLEGE SHARE

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	Δ GDPpc	Δ GDPpc	Δ GDPpc	Δ GDPpc	Δ GDPpc	Δ GDPpc	Δ % college	Δ % college
Lagged growth in #universities	0.0445***	0.0711***	0.0673***	0.0679***	0.0654***	0.0656***	0.00546***	0.00373***
	(0.0105)	(0.0137)	(0.0136)	(0.0136)	(0.0136)	(0.0136)	(0.00118)	(0.000984)
Lagged growth in college share			1.963***		1.698***	1.631***		
			(0.323)		(0.304)	(0.327)		
Current growth in college share				0.870***	0.651***	0.663***		
				(0.146)	(0.131)	(0.129)		
Lagged level of college share						0.0102		
						(0.0181)		
Lagged level of regional GDP per capita	-0.0127***	-0.00993***	-0.0112***	-0.0105***	-0.0114***	-0.0116***		0.000689***
	(0.00131)	(0.00165)	(0.00165)	(0.00164)	(0.00165)	(0.00171)		(0.0000899)
Lagged level of country GDP per capita	-0.0213***	-0.0452***	-0.0431***	-0.0436***	-0.0421***	-0.0424***		-0.00190***
	(0.00422)	(0.00439)	(0.00434)	(0.00437)	(0.00433)	(0.00444)		(0.000363)
Lagged level of population/100	-0.0855**	-0.169***	-0.191***	-0.178***	-0.195***	-0.198***		0.0110***
	(0.0387)	(0.0466)	(0.0471)	(0.0463)	(0.0471)	(0.0480)		(0.00354)
Lagged growth in population	-0.113***	-0.0226	-0.0242	-0.0238	-0.0249	-0.0259		0.00137
	(0.0385)	(0.0491)	(0.0495)	(0.0493)	(0.0496)	(0.0497)		(0.00233)
Dummy for capital in region	0.0110***	0.00952***	0.00790***	0.00864***	0.00746***	0.00716***		0.00102***
	(0.00168)	(0.00189)	(0.00194)	(0.00192)	(0.00197)	(0.00196)		(0.000204)
Observations	8128	5118	5118	5118	5118	5118	5118	5118
Adjusted R-squared	0.260	0.286	0.292	0.290	0.294	0.294	0.278	0.349
# clusters	1498	1089	1089	1089	1089	1089	1089	1089

Notes: Growth in college share is simply the percentage point difference: (college share (t) – college share (t-5))/5. Column (1) replicates Column (5) from Table 3. Column (2) restricts to the sample for which the change in college share is available. Column (3) drops the lagged growth in college share. Column (4) adds the contemporaneous change in college share. Column (5) includes both lagged and contemporaneous changes. Column (6) further adds the lagged level of college share (unlogged). Column (7) regresses the change in college share on the lagged growth in universities, with country dummies, but no other controls. Column (8) adds all the other controls.

Table 8: UNIVERSITIES AND PATENTS

	(1)	(2)	(3)	(4)
Dependent Variable	Δ GDPpc	Δ GDPpc	Δ patents	Δ patents
Panel A: USPTO				
Lagged growth in #universities	0.113*	0.109*	2.679***	0.0804
	(0.0616)	(0.0637)	(0.657)	(0.0654)
Growth in USPTO patent "stock"		0.0429		
		(0.0407)		
Lagged level of regional GDP per capita	-0.0303***	-0.0293***		
	(0.00773)	(0.00816)		
Lagged level of population/100	0.156*	0.165*		-0.431***
	(0.0883)	(0.0862)		(0.0864)
Lagged growth in population	-0.0888	-0.115*		0.510***
	(0.0568)	(0.0624)		(0.0755)
Dummy for capital in region	-0.0142*	-0.0133		-0.0113
	(0.00783)	(0.00811)		(0.0115)
Observations	306	306	306	306
Adjusted R-squared	0.350	0.349	0.049	0.994
# clusters	51	51	51	51
Panel B: EPO				
Lagged growth in #universities	0.0220*	0.0192	0.0531	0.0392
	(0.0122)	(0.0117)	(0.0457)	(0.0457)
Growth in EPO patent "stock"		0.0514***		
		(0.00528)		
Lagged level of regional GDP per capita	-0.0127***	-0.0146***		
	(0.00234)	(0.00241)		
Lagged level of country GDP per capita	-0.000438	-0.00245		
	(0.00653)	(0.00609)		
Lagged level of population/100	-0.00831	-0.0930*		1.661***
	(0.0541)	(0.0547)		(0.227)
Lagged growth in population	-0.137***	-0.111***		-0.451***
	(0.0445)	(0.0426)		(0.145)
Dummy for capital in region	0.0104***	0.0104***		0.0176*
	(0.00222)	(0.00230)		(0.00930)
Observations	3747	3747	3747	3747
Adjusted R-squared	0.320	0.342	0.306	0.343
# clusters	802	802	802	802

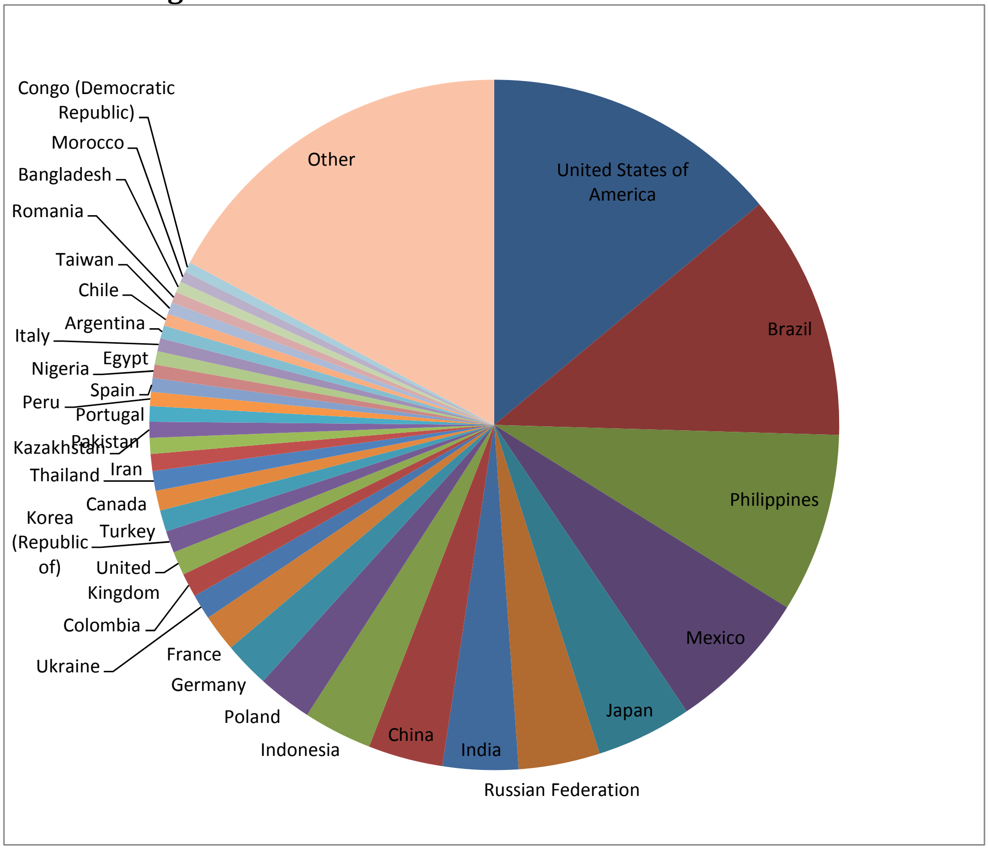
Notes: Panel A relates to the US and the time period for which USPTO patents have been assigned to state (region) in the NBER dataset (1964-1999). Panel B is a larger sample: the countries for which regionalized EPO patents are available in OECD REGPAT (1975-2005). Column (1) replicates our core regression (column (5) from Table 3), but restricts to the relevant sample for patents data. Column (2) adds in the contemporaneous growth in cumulative patent "stock" to the regression. Column (3) regresses the growth in patent stock on the growth in universities as a raw correlation, with no other controls. Column (4) then adds the standard time varying controls (reported) and geographic controls (not reported).

Table 9: UNIVERSITIES AND APPROVAL OF DEMOCRACY

Dependent variable: Approval of Democracy	(1)	(2)	(3)	(4)
15 year lagged ln(1+#universities per capita)	0.0293***	0.0257***	0.0230**	0.0229**
	(0.00934)	(0.00960)	(0.00946)	(0.00997)
Dummy for Male		0.0376***	0.0338***	0.0337***
		(0.00477)	(0.00468)	(0.00468)
Age (years)		0.00177***	0.00222***	0.00221***
		(0.000300)	(0.000302)	(0.000302)
Dummy for married		-0.00119	-0.00420***	-0.00407***
		(0.00127)	(0.00126)	(0.00123)
Children		-0.00696***	-0.00342*	-0.00348*
		(0.00200)	(0.00196)	(0.00194)
Employed (full, part time, self-employed)		0.0173***	0.0239***	0.0239***
		(0.00551)	(0.00616)	(0.00613)
Income scale		0.0116***	0.00540	0.00562*
		(0.00339)	(0.00329)	(0.00323)
Dummy for holds university degree			0.135***	0.135***
			(0.00775)	(0.00768)
Dummy for student			0.0854***	0.0856***
			(0.0116)	(0.0116)
Observations	138511	138511	138511	138511
Adjusted R-squared	0.075	0.079	0.085	0.085
# clusters	693	693	693	693
Country and year dummies	yes	yes	yes	yes
Geographic controls	no	no	yes	yes

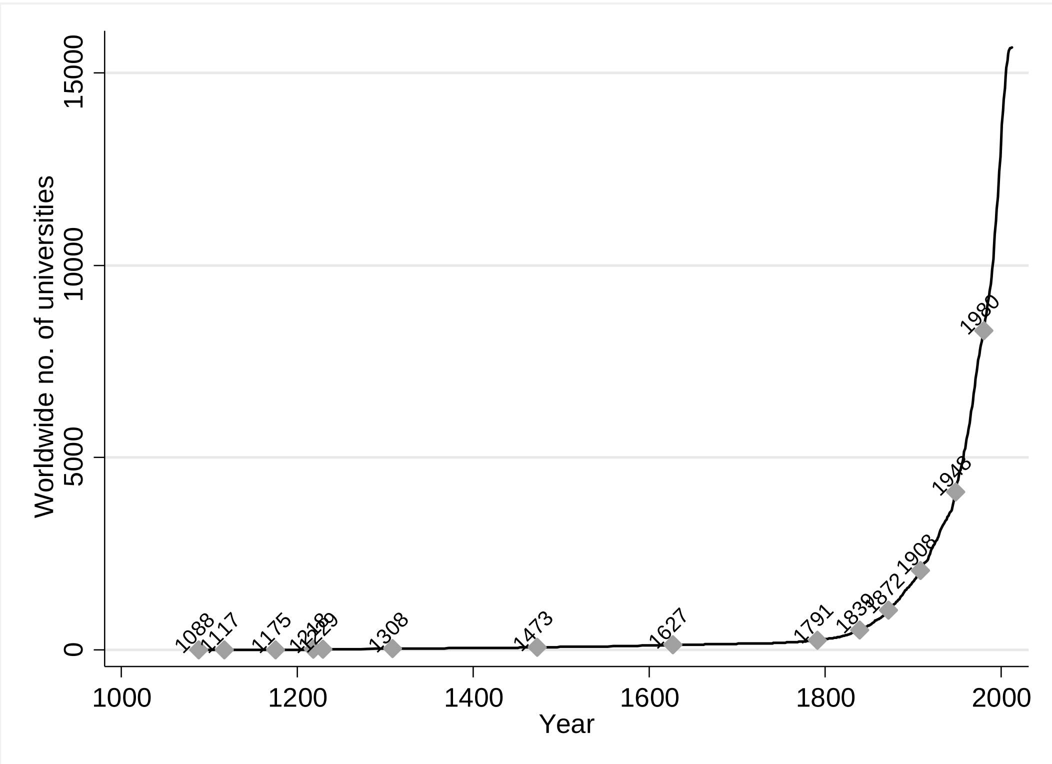
Notes: OLS estimates, 54 countries. Standard errors are clustered at the regional level. **Region controls** include latitude, inverse distance to ocean, malaria ecology, log(oil and gas production) 1950-2010 and a dummy for if a region contains the country's capital city.

Figure 1: LOCATION OF UNIVERSITIES IN 2010



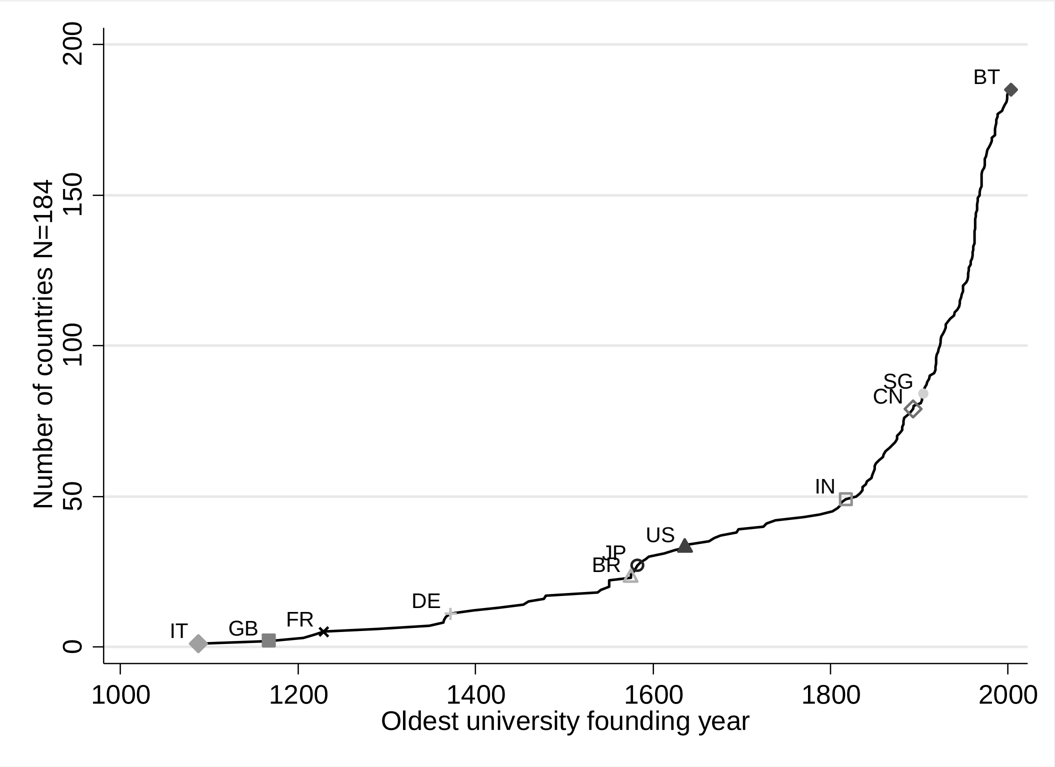
Notes: Pie chart shows the share of worldwide universities in each country, as at 2010. *Source: WHED.*

Figure 2: WORLDWIDE UNIVERSITIES OVER TIME



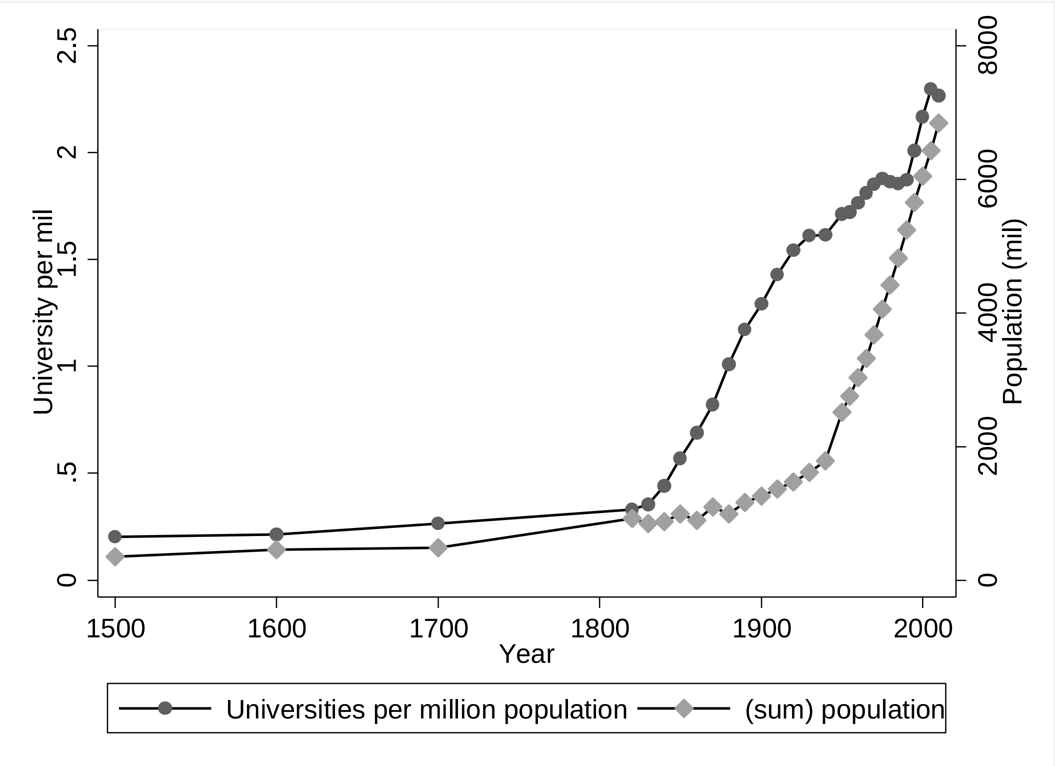
Notes: The evolution of global universities over time; years where the total number doubled are marked. *Source: WHED.*

Figure 3: DIFFUSION OF UNIVERSITIES ACROSS COUNTRIES



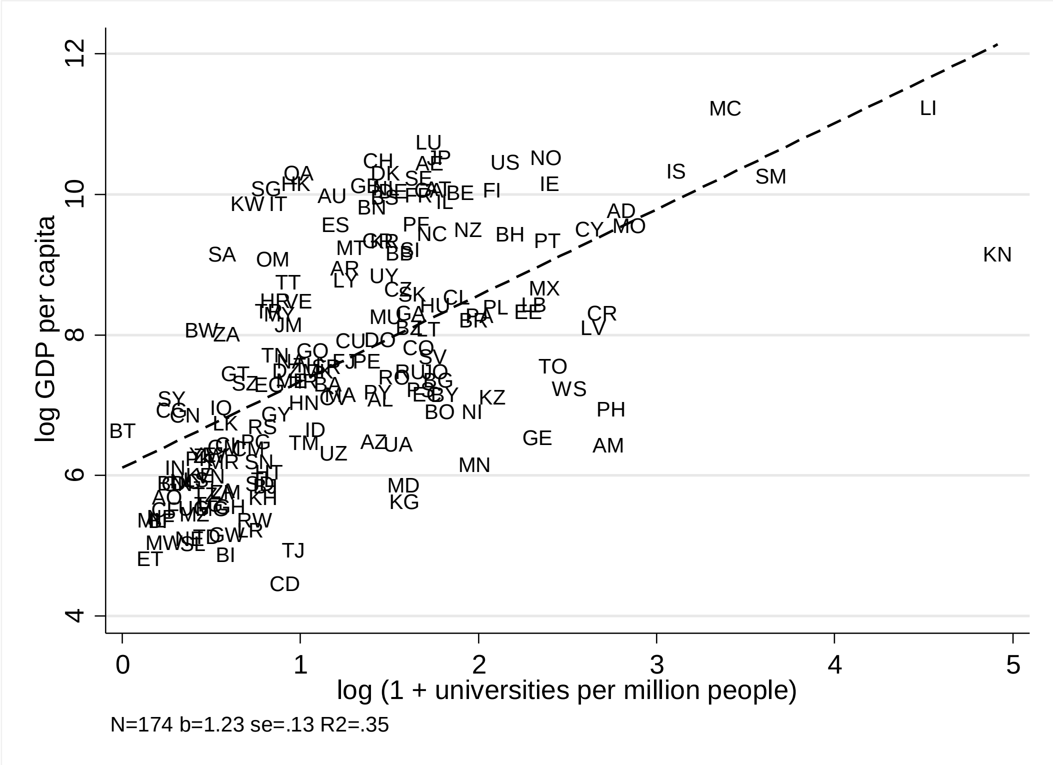
Notes: This chart shows the total number of countries that have universities over time, with some key countries marked in the year they opened their first universities marked. *Source: WHED.*

Figure 4: UNIVERSITY DENSITY AND POPULATION



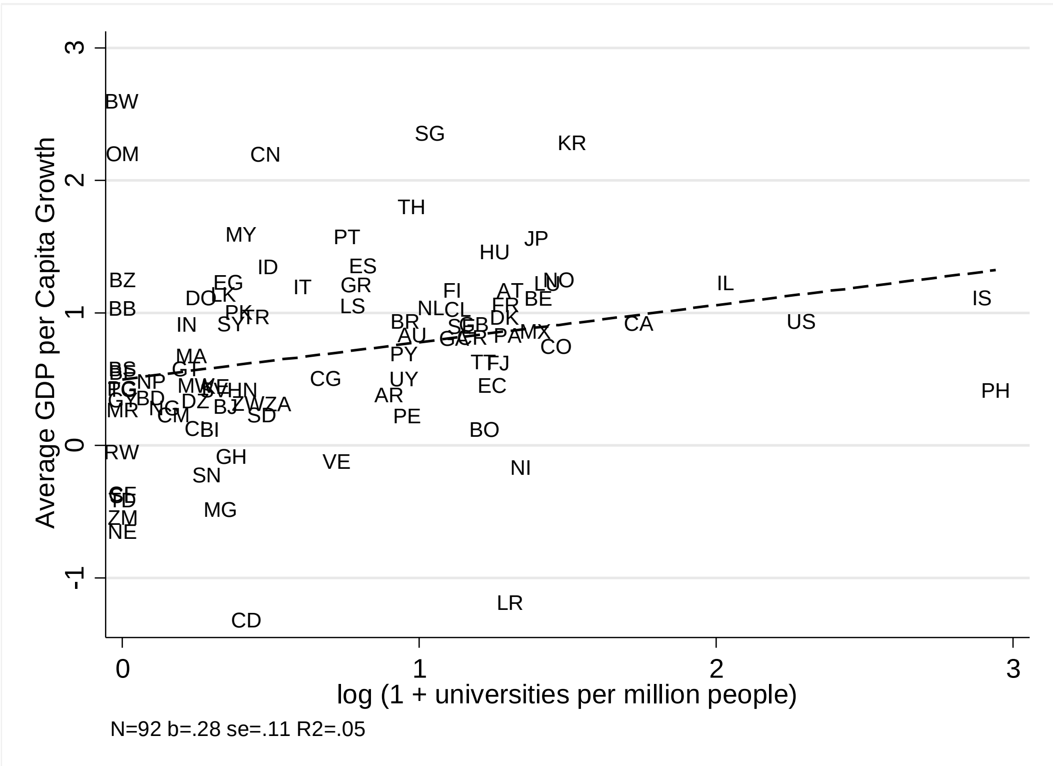
Notes: This chart shows the evolution of global university density (universities per million people) and population over time. *Source: WHED and Maddison population data.*

Figure 5: SCATTER PLOTS AT COUNTRY LEVEL, CROSS SECTION IN 2000
Panel A: Universities and income in 2000



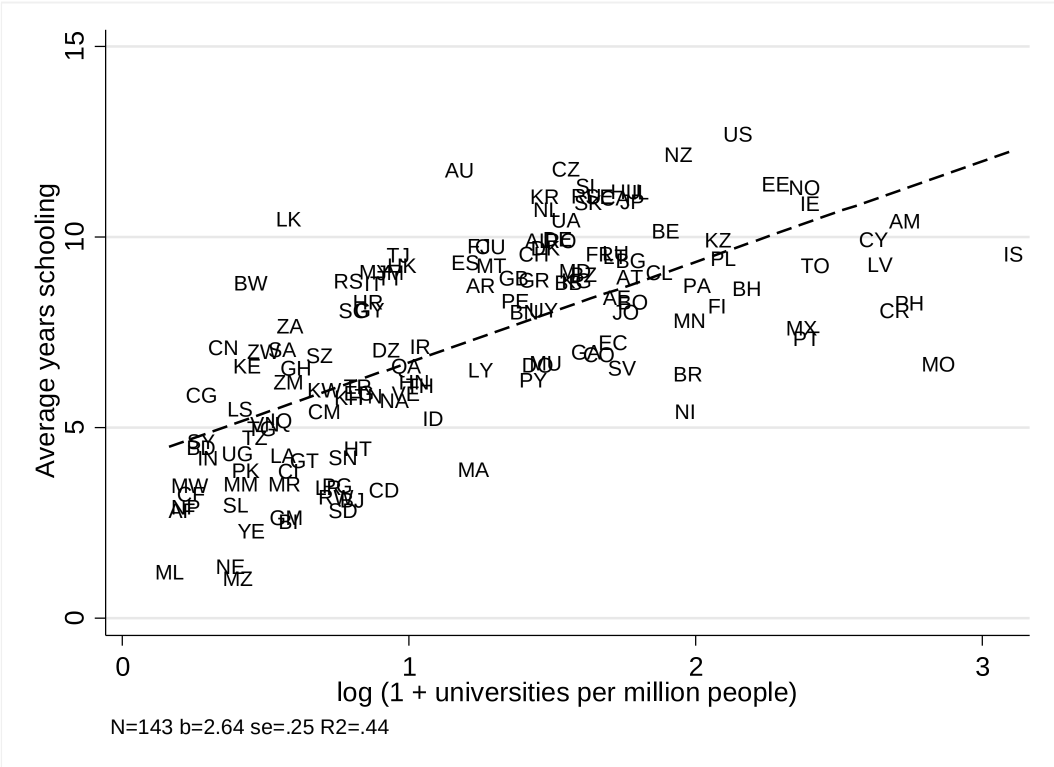
Notes: Each observation is a country in 2000. *Source: WHED and World Bank GDP per capita*

Panel B: Universities in 1960 and GDP per capita growth (1960-2000)



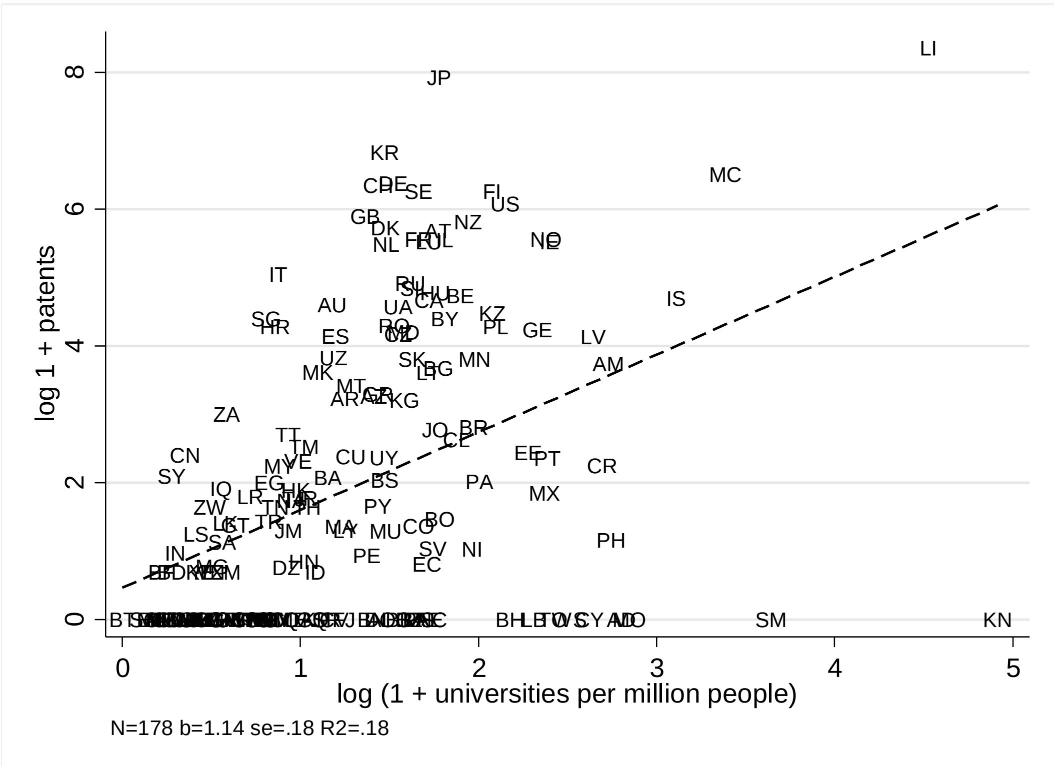
Notes: Each observation is a country. Average annual growth rates over the period 1960-2000 on the y axis. *Source: WHED and World Bank GDP per capita*

Panel C: Universities and average years of schooling in 2000



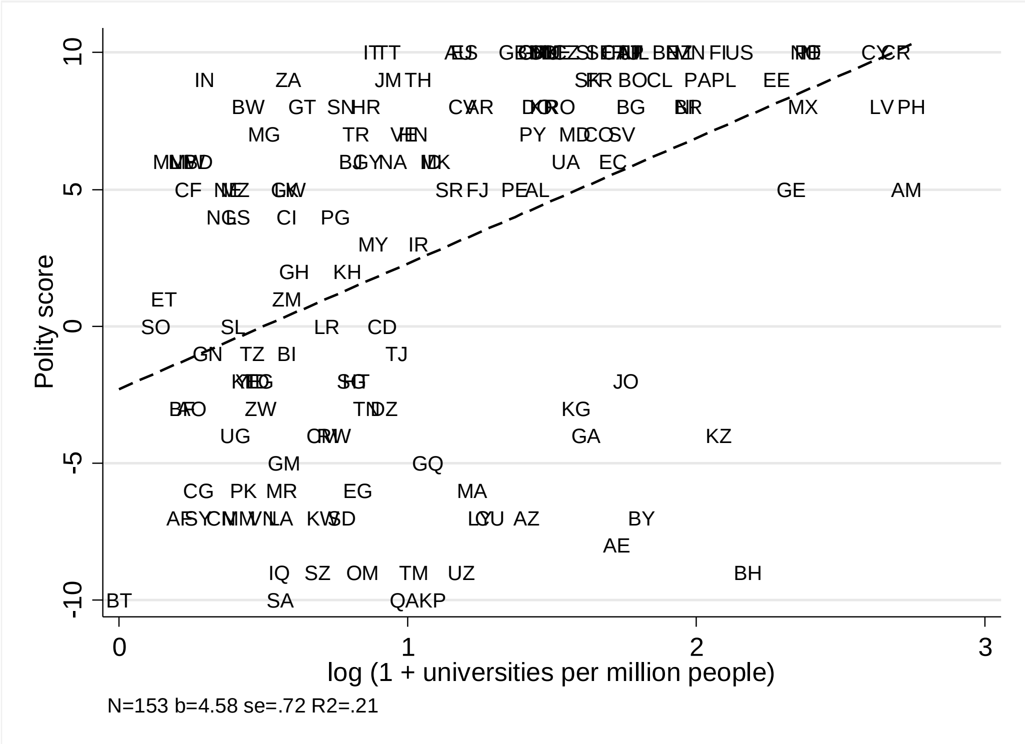
Notes: Each observation is a country in 2000. *Source: WHED and years of schooling obtained from Barro-Lee dataset*

Panel D: Universities and patents in 2000



Notes: Each observation is a country in 2000. *Source: WHED and patents from WIPO*

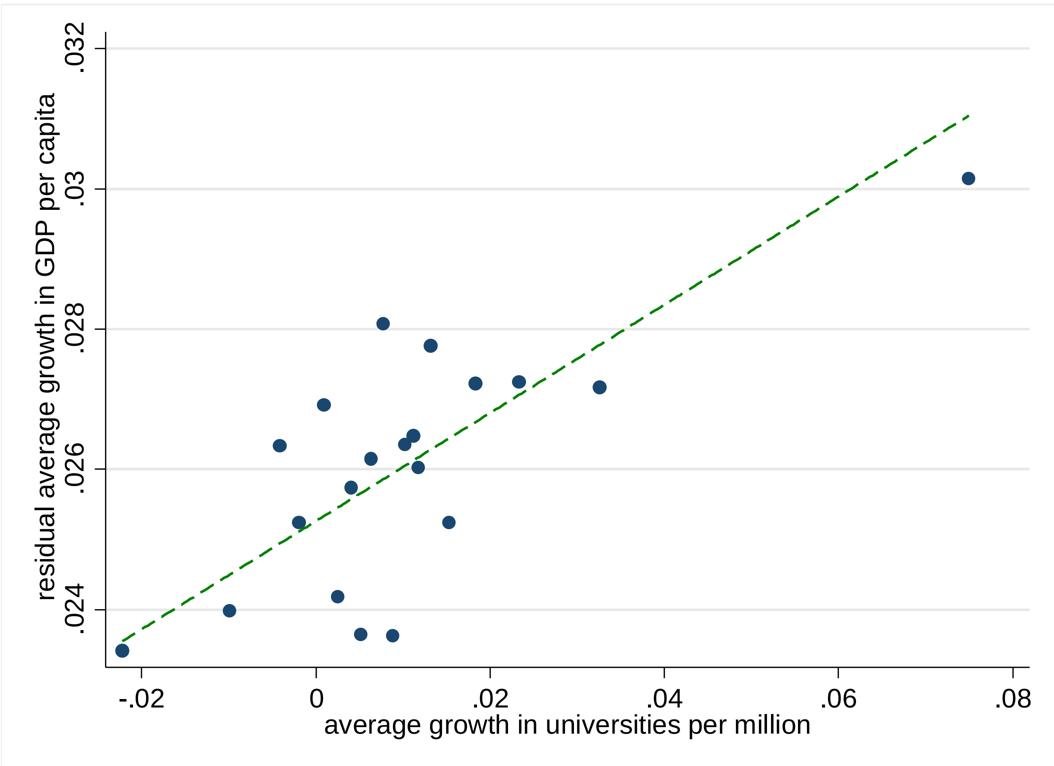
Panel E: Universities and democracy in 2000



Notes: Each observation is a country in 2000. *Source: WHED and Polity2 scores from Polity IV project*

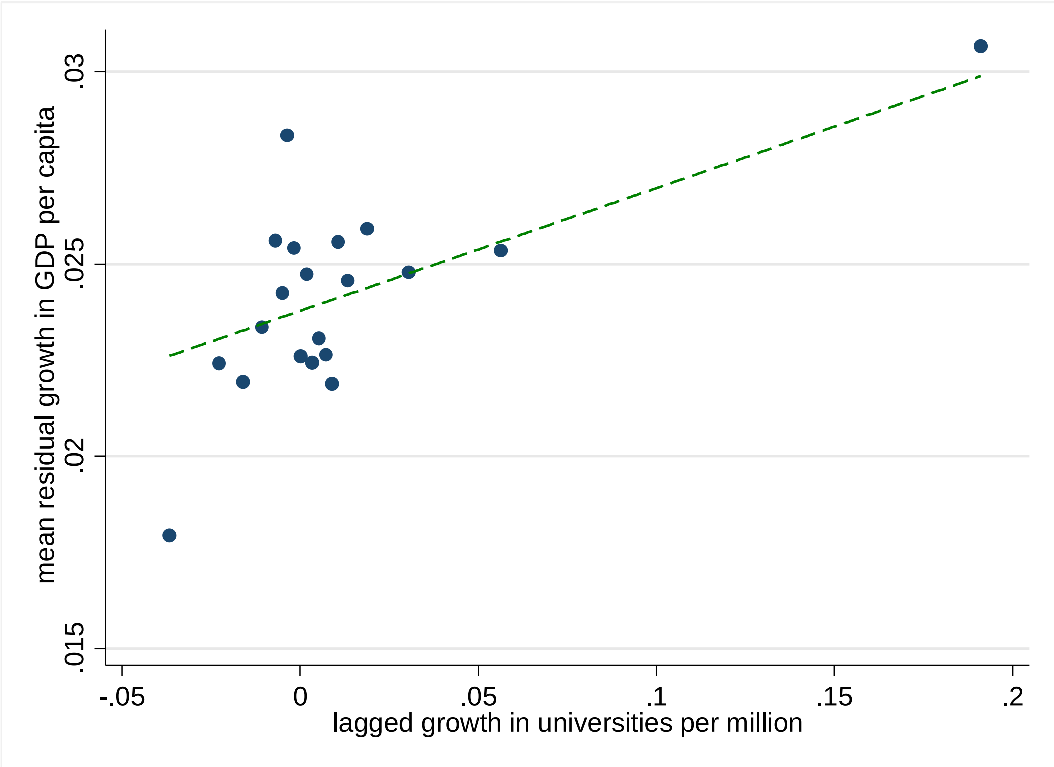
Figure 6: REGIONAL GDP PER CAPITA GROWTH AND UNIVERSITY GROWTH

Panel A: Average growth rates, one observation per region



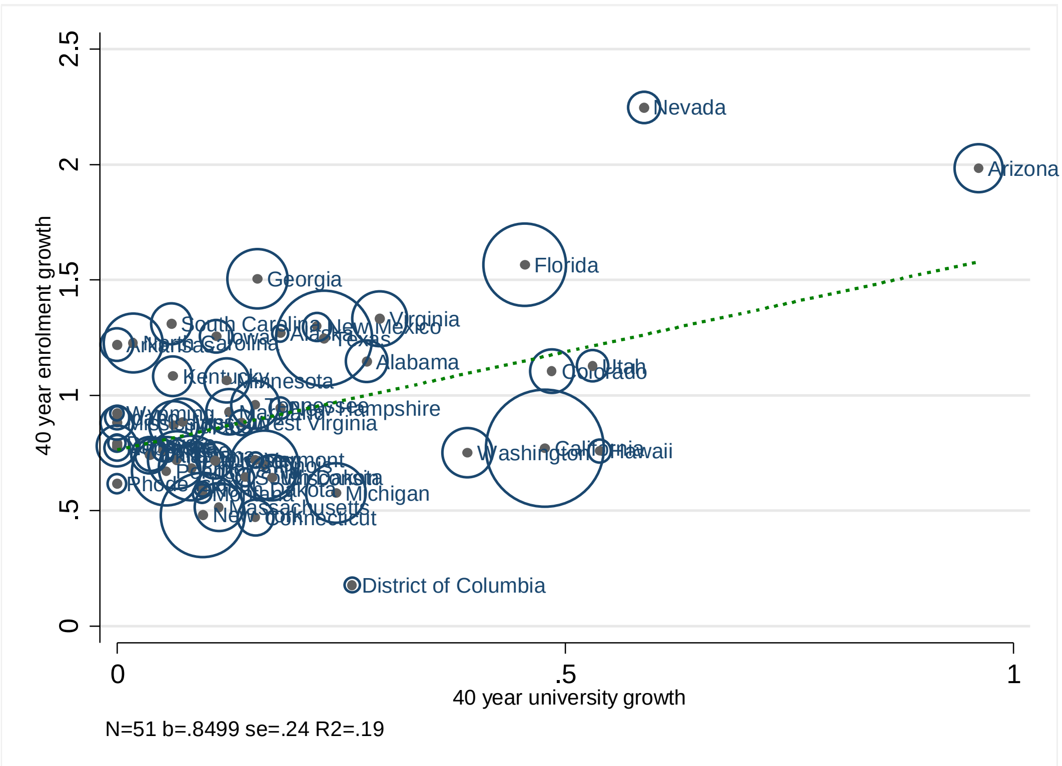
Notes: 1,498 region observations are grouped equally into 20 bins, variation is within country. *Source:* WHED and Gennaioli et al (2014) for regional GDP per capita and population

Panel B: Average growth rates, region-year observations



Notes: 8,128 region-year observations are grouped equally into 20 bins, variation is within country. *Source:* WHED and Gennaioli et al (2014) for regional GDP per capita and population

Figure 7: GROWTH IN US ENROLMENTS VS GROWTH IN UNIVERSITIES



Notes: Each observation is a region (US state), weighted by the region's share in total US population in 2010. 40 year growth relates to the period 1970-2010. Dropping Arizona, b=0.62 and se=0.28. *Source: WHED and NCES.*

APPENDICES: INTENDED ONLY FOR ONLINE PUBLICATION

APPENDIX A1: DATA APPENDIX

A.1.1 WHED COVERAGE

WHED contains data on 185 countries (which includes 176 countries plus 9 administrative regions/dependencies: Hong Kong, Macao, Curaçao, French Guyana, French Polynesia, Guadeloupe, Martinique, New Caledonia and Reunion). We cross-check the 176 countries to a full list of independent states (from the US State Department, http://www.state.gov/s/inr/rls/4250.htm) and find that there are only 16 more independent states not included in the database. These are Antigua and Barbuda, Comoros, Djibouti, Dominica, Eritrea, Grenada, Kosovo, Micronesia, Nauru, Palau, Saint Lucia, Saint Vincent and the Grenadines, Sao Tome and Principe, Seychelles, Solomon Islands and South Sudan.

WHED contains information on higher education institutions that offer at least a post-graduate degree or a four year professional diploma. It therefore excludes, for example, further education institutions in the UK or community colleges in the US and may be thought of as a sample of "higher quality" universities.

We compare the country totals in WHED as at 2010 to data from "Webometrics" (http://www.webometrics.info/en/node/54), a source where higher education institutions (including ones that would not qualify for inclusion in WHED) are ranked by their "web presence". This source puts the total number of universities worldwide at 23,887 in 2015 (part of this difference will be due to growth over the 2010-2015 period). In the results section, we discuss a robustness check where we drop countries from our regressions with a very large divergence between the two sources.

A1.2 VALIDATING OUR APPROACH

Our approach for calculating university presence by region uses the founding dates of universities to determine the number of universities that were present at any particular date. We consider that a "university" is founded on this initial founding date, even if it was a smaller higher education institute or college at that date. This is often the case, but our approach is reasonable since only the better quality institutions are likely to subsequently become universities. Furthermore, there are many cases where a number of universities or higher education institutes were merged together into what is today recorded as one university in WHED. Our approach avoids the apparent reduction that would occur in such cases if we were merely counted the number of institutions present at any given date.

One key concern with this strategy is that it would not be suitable in a world where university exits are commonplace. Say a number of universities were present in the past and closed down before WHED 2010. A region could have actually seen a fall in universities, but our method would not capture this since it includes only surviving universities. Anecdotally we know that the period since the 1960s has been one of university growth across the globe, but we investigate this issue further in order to gain more comfort on the validity of our approach. We do this by obtaining historical records of the universities and higher education institutions present in the 1960s, and assess whether significant numbers of these are missing from WHED 2010.

The appropriate sources are the predecessors to WHED: "The International Handbook of Universities" (1959, published by the IAU annually); "American Universities and Colleges" (1960, published by The American Council on Education and "The Yearbook of Universities of the Commonwealth" (1959), published by the Association of Universities of the British Commonwealth). As the name suggests, the "American Universities and Colleges" yearbook contains fully fledged universities, but also smaller colleges (including religious institutions), many of which would not be included in WHED today. The international handbook lists universities and other institutions not considered of "full university rank" separately. We include all of these institutions because the distinction is not consistent between countries – for example in France, these latter institutions contain all the grandes écoles which are considered to be of very high quality but are outside the framework of the French university system; and in China only one institution is listed as a full university while other institutions include a number of institutions with the name "university". The Commonwealth yearbook contains only fully fledged universities.

The main exercise we carry out is to name match between 1960 yearbooks and WHED 2010. There are 2,694 institutions listed across 110 countries in the three yearbooks in 1960, this compares with 5,372 institutions (in 132 countries) which according to WHED 2010 were founded pre 1960 – this is higher because WHED counts universities from the date they are founded, even if they are not founded as a fully-fledged university (as discussed above). The country level correlation of the number of universities present in 1960 in the two sources is 0.95. The matching process involves a number of iterations: exact matching, "fuzzy" matching, and manual matching. The process is complex because name changes and mergers are commonplace, therefore internet searches on Wikipedia or university websites were necessary. Where an institution was found to have been merged into a university that is present in WHED 2010 we considered it a match. The results of this process are summarized in Table A10. We find that university closure is extremely rare, and we only find evidence of this in the US, where 33 small (mostly religious) colleges are present in the 1960 yearbook and were found to have closed down, mainly due to bankruptcy. 155 institutions worldwide were found to still be in existence but not be listed in WHED. This was usually because they do not meet the WHED listing criteria (a university that offers at least a four year degree or postgraduate courses). Indeed, of the 155 institutions in this category, 115 were not considered fully fledged universities in 1960, and 33 of the remaining 40 were US colleges (mostly religious).

Based on these facts, we believe that it is reasonable to use the WHED founding dates as an (albeit imperfect) basis for a time series of university presence by region.

A1.3 DESCRIBING UNIVERSITY GROWTH IN SELECTED COUNTRIES

This section gives a historical overview of the diffusion of universities from the 1880s in four advanced economies: France, Germany, the UK and US, and two emerging economies: India and China. We compare the timing of historical university expansions to growth and industrialisation (see Figure A1 for a measure of industrialisation over time in the UK, US, France and Germany sourced from Bairoch (1982)). This analysis provides a visual "sense-check" for the thesis developed by Mokyr (2002) that the building and dissemination of knowledge played a major role in the Industrial Revolution.

In the United Kingdom, universities have been established in waves: the "Ancient universities" starting with Oxford in 1100s were the first seven universities which were founded before 1800. Then a number of universities were chartered in the 19th Century, followed by the "Red Brick" universities before World War I. A large expansion occurred after World War II, following the influential Robbins Report into Higher Education (1963). Former polytechnics were converted to universities in 1992, but in our data these higher education institutions are counted from when they opened in their original form. These waves can be seen in the university density line as shown in Figure A2, Panel A, which also plots national GDP per capita data (from Maddison), suggesting that the first expansions coincided with industrialisation in the 1800s (Figure A1 shows that industrialisation picked up from the 1830s in the UK). The raw university count trend is shown in Panel B.

In the US, the first university was Harvard, founded in 1636. By the American Revolution there were nine colleges modelled on the Oxford and Cambridge in England. However these were very small, exclusive and focused on religion and liberal arts. At that time, there were no law or medical schools, so one had to study these subjects in London. It was Thomas Jefferson who had a vision for state education, separate from religion, but this only took hold after the Civil War with the land grant colleges. This sharp rise in university density can be seen in Figure A3. Industrialization in the US began to pick up in the 1860s (see Figure 1). University density reached much higher levels than in Britain: at 13 universities per million people in 1900 versus just over 2 in the UK. The difference is that in the US, density came down again as population growth outpaced the opening of new universities which continued to grow as shown in Panel B; though the downward trend did slow during the post war period (we can see the slight kick in university numbers from the 1950s in Panel B). However, the fall in university density must be considered in the context that over the same period, university size has also been increasing in the US and (this can be seen in Figure A4 and in our analysis in Section IV on enrolments). Furthermore, there has been a sharp rise in "Community Colleges" in the US, which provide college access qualifications, and are not counted in our dataset.

In France, Figure A5 shows that university density really started picking up in the 1800s with the opening of the "Grande Écoles" which were established to support industry, commerce and science and technology in the late 19th Century. Indeed industrialization in France was more gradual, and started picking up in the late 1880s, early 1900s. The next dramatic increase in universities numbers and density occurred in the 1960s during de Gaulle's reforms of the French economy.

Cantoni and Yuchtman (2014) discuss the opening of the first universities in Germany following the Papal schism in the late 14th Century. However, during the 1800s, Figure A6 shows that university density actually fell, as population growth outpaced the gradual increases in university numbers which can be seen in Panel B. Historically, Germany had a low share of college graduates as higher shares of the population were educated via the apprenticeship system. A deliberate push to expand university education began in the 1960s, with new public universities founded across the country (Jäger, 2013). This was motivated by economic reasons; in particular the need to compete in technology and science against the backdrop of the Cold War; but also social reasons, namely the notion that education is a civil right to be extended beyond the elites, and is crucial for democracy.

China and India saw much later expansions as shown in Figure A7 and Figure A9. China started opening up to Western advances in science in the 1800s, and followed Soviet influence in the 1950s with centrally planned education. We can see a sharp rise in university density from the 1900s to 1960. The spike in the 1960s is due to the Cultural Revolution, when higher education institutions were shut down for 6 years, and all research terminated. When the universities were reopened, they taught in line with Maoist thought. It was from the 1980s that institutions began to gain more autonomy and when China began its rapid growth trajectory, though so far growth in universities has not outpaced population growth. In India, expansion occurred after independence in 1947. During the colonial era, the upper classes would be sent to England for education. The British Raj oversaw the opening of universities and colleges from the late 1800s, but university density only started rising more rapidly after 1947 and recently has picked up pace again. We note that the in both countries, there are around 0.4 universities per million people, which is still a lot lower than in the UK or US.

Finally, we note that in general, expansions in university numbers have been accompanied by increases in university size. As we saw in Figure A4 (using UNESCO data that are only available from 1970), university students normalized by population have been growing overall in the US and the UK since the 1970s (with a dip in the late 1990s in the US) and more recently in China and India.

APPENDIX A2: FURTHER RESULTS

A2.1 GROWTH ON LAGGED LEVELS REGRESSIONS

In Table A6 we replicate as closely as possible[37] the results in Gennaioli et al (2014) but add in the lagged university level into our analysis. Column (1) follows their Table (5), column (8), omitting years of education, and column (2) includes years of education. The coefficients are very similar: the convergence term is between 1.4% and 1.8%, and the coefficient on years of education is nearly identical at around 0.004. Adding in universities and population does not affect the other coefficients much. Column (3) suggests that doubling the level of universities leads to a 0.24% rise in the GDP per capita growth rate. Universities have a positive and significant effect over and above years of education. As we would expect if some of the effect of universities is via their production of human capital, the effect of universities is higher when years of education are omitted (column (3)). Table A7 presents a similar analysis, but in long difference format; so for each region there is one observation with the average annual growth rate over the 50 years to 2010, 40 years and 30 years respectively; regressed on starting period universities and other controls. Overall, this shows that even in this simplified specification on the reduced sample where the relevant data are available for the time periods, there is a positive significant relationship between initial period univerisities and subsequent growth once country fixed effects are included; and the magnitude is comparable with the conventional Barro-style results.

A2.2 SIMULATION OF THE EFFECTS OF HUMAN CAPITAL

We take the average region in the dataset in the year 1990: one with a population of 3 million people, a college share of 6%, average years of education of 6.7, and 9 universities. We want to assess how the opening of a new university will affect the growth in college share and years of education. For comparability to the regressions, we assume that there are 10 universities in the region, so that adding one university in a five year period represents a 10% increase over that period, or an average of 2% per year. We assume that the university enrolls 5,000 students (in cohorts of 1,250 which join each year), and we assume that after 4 years, each cohort

[37] Our sample is larger because for the purposes of our analysis we interpolate GDP per capita, and not just years of education and population as in their paper.

graduates with a bachelors degree and stay in the region; adding to its human capital stock. We assume a staff-student ratio of 20, based on enrollments in a given year. We do not account for post-graduates, or additional graduates as other skilled people move to the region following the opening of the university. From this simple example, we generate impacts on college share and years of education growth in the next five year period and compare these to our coefficients in column (8) of Tables 6 and A9.

Our calculation gives results of similar order of magnitude to those implied by the regressions: an average annual rise in college share of 0.0003 resulting from a two per cent rise in universities, which is actually smaller than the effect implied by the regression result of 0.007 (0.00373 x 2). The average annual rise in years of education of 0.06% which slightly larger than the 0.04% implied by the regression (0.0196% x 2) (Table A8).

Overall, we can see that in the main specification that gives us a relationship between lagged growth of universities and regional GDP per capita growth, human capital is unlikely to be the main mechanism. This is due to the time required for a change in the stock of universites to change the stock of human capital. We saw that in the cross sectional analysis in Table 3, and in the Barro-style regressions in Table A6 that adding in human capital measures made a larger "dent" in our university effect due to the fact both measures are accumulated "stocks".

A2.3 UNIVERSITIES AND DEMOCRATIC APPROVAL

Figure A10 shows that there is a positive and signifcant correlation between the change in university density and change in polity scores over 1960-2000.

Table A9 reports a number of robustness tests around the regressions of approval of democracy on lagged university presence reported in Table 9. Column (1) repeats Table 9 column (4). Column (2) shows that this effect appears to be driven by OECD countries, as an interaction term between an OECD dummy and the lagged university presence is positive and significant. Column (3) shows that the main result is much smaller in magnitude and insignificant for 5 year lagged university presence, and actually negative for a 30 year lag. We note however, that the results are robust across lags on the OECD subsample (available on request). Column (5) shows that our main result can be closely replicated using a different survey measure for approval of democracy, "democracy is best" which asks respondents whether they agree with the statement that democracy is better than any other form of government. Column (6) does not include country fixed effects. This shows that the positive relationship we find between universities and approval of democracy is valid within countries. Across countries, factors not controlled for in these regressions (such as levels of corruption) appear to influence the result. We investigated which countries appear to be driving this negative relationship and found, for example, that the Philippines (a country with high levels of corruption) has high university density but low approval of democracy. Column (7) clusters at the country level and significance holds. Column (8) weights by population, to account for the fact that some regions with low population may have less representative responses. Column (9) drops students and graduates and the main result gets stronger. Finally, column (10) shows that the results are robust to estimation using an ordered-probit model.

APPENDIX TABLES

Table A1: FULL GROWTH ON LAGGED GROWTH REGRESSION

Dependent variable: Regional Growth of GDP per capita	
Lagged growth in #universities	0.0445***
	(0.0105)
Lagged level of regional GDP per capita	-0.0127***
	(0.00131)
Lagged level of country GDP per capita	-0.0213***
	(0.00422)
Lagged level of population	-0.0855**
	(0.0387)
Lagged growth in population/100	-0.113***
	(0.0385)
Dummy for capital in region	0.0110***
	(0.00168)
Latitude	-0.000318***
	(0.0000875)
Inverse distance to ocean	0.00456
	(0.00373)
Malaria ecology	0.000736**
	(0.000292)
log(oil and gas production) 1950-2010	0.000293**
	(0.000142)
Observations	8128
Adjusted R-squared	0.260
# clusters	1498

Notes: This table replicates Table 3, column (5) to show the geographic controls. Note, log(oil and gas production) 1950-2010 is not normalized by population.

Table A2: DISTRIBUTED LAG SPECIFICATIONS

Dependent variable: Regional Growth of GDP per capita	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Current growth in #universities		-0.000701	-0.00195	-0.0179	-0.0212*	-0.0140	-0.00499
		(0.0108)	(0.0109)	(0.0121)	(0.0128)	(0.0144)	(0.0164)
5 year lagged growth in #universities	0.0445***		0.0446***	0.0507***	0.0458***	0.0486***	0.0675***
	(0.0105)		(0.0105)	(0.0117)	(0.0120)	(0.0123)	(0.0144)
10 year lagged growth in #universities				0.00916	0.00646	-0.00160	0.00442
				(0.0106)	(0.0110)	(0.0118)	(0.0130)
15 year lagged growth in #universities					-0.00663	-0.0210*	-0.0213
					(0.0109)	(0.0110)	(0.0131)
20 year lagged growth in #universities						-0.0138	-0.0189
						(0.0145)	(0.0158)
25 year lagged growth in #universities							0.00509
							(0.0130)
Observations	8128	9246	8128	6863	5635	4604	3638
Adjusted R-squared	0.260	0.253	0.261	0.252	0.245	0.237	0.283
# clusters	1498	1527	1498	1391	1104	1013	906

Notes: Column (1) is replicates column (5) from Table 3. The subsequent columns add contemporaneous and further lagged growth in universities, and corresponding population growth. The level of population at the furthest lag is also controlled for.

Table A3: UNIVERSITY GROWTH AS DEPENDENT VARIABLE

Dependent variable: Regional Growth in Number of Universities	(1)	(2)	(3)	(4)	(5)	(6)
Lagged growth in regional GDP per capita	-0.00523	0.00159	0.00439	0.00625	0.00995	-0.00293
	(0.0125)	(0.0147)	(0.0147)	(0.0147)	(0.0145)	(0.0148)
Lagged growth in country GDP per capita		-0.0224	-0.0230	-0.0251	-0.0383*	-0.0271
		(0.0224)	(0.0225)	(0.0225)	(0.0222)	(0.0222)
Lagged growth in population			0.0775*	0.0809**	0.0808*	0.0394
			(0.0407)	(0.0411)	(0.0418)	(0.0412)
Lagged level of population/100				0.262***	0.772***	0.813***
				(0.0419)	(0.0801)	(0.0893)
Lagged #universities					-0.00802***	-0.00983***
					(0.000980)	(0.00128)
Lagged level of regional GDP per capita						0.00528***
						(0.00113)
Lagged level of country GDP per capita						-0.00589*
						(0.00312)
Dummy for capital in region						0.00714**
						(0.00284)
Observations	7746	7746	7746	7746	7746	7746
Adjusted R-squared	0.133	0.133	0.134	0.138	0.154	0.159
# clusters	1489	1489	1489	1489	1489	1489

Notes: All columns include country and year fixed effects, and standard errors clustered by region. Column (1) is simple correlation between regional growth in universities and the lagged growth in regional GDP per capita. Columns (2) to (5) include the variables shown. In addition, column (6) includes geographic controls which are not reported: latitude, inverse distance to ocean, malaria ecology, log(oil and gas production) 1950-2010. Levels of GDP per capita and population are in natural logs.

Table A4: ROBUSTNESS ON FIXED EFFECTS SPECIFICATIONS

Dependent variable: Regional Growth GDP pc	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Lagged growth in #unis	0.0468***	0.0468**	0.0422***	0.0452***	0.0474***	0.0625***		0.0356***	0.0468***	0.0579***
	(0.0118)	(0.0179)	(0.0151)	(0.0118)	(0.0117)	(0.0157)		(0.00880)	(0.0118)	(0.0144)
Lagged growth in #unis, winsorized							0.0645***			
							(0.0164)			
Lagged level regional GDP pc	-0.0776***	-0.0776***	-0.0690***	-0.0767***	-0.0793***	-0.0764***	-0.0776***	-0.0543***	-0.0776***	-0.0829***
	(0.00529)	(0.00762)	(0.00608)	(0.00531)	(0.00619)	(0.00529)	(0.00529)	(0.00375)	(0.00529)	(0.00614)
Lagged level country GDP pc	0.0378***	0.0378*	0.0307***	0.0365***	0.0404***	0.0372***	0.0378***	0.0185***	0.0378***	0.0140*
	(0.00677)	(0.0199)	(0.00676)	(0.00684)	(0.00781)	(0.00681)	(0.00677)	(0.00483)	(0.00677)	(0.00744)
Lagged level of population/100	-0.850**	-0.850	-0.537	-1.173***	-0.798*	-0.763*	-0.862**	-0.544*	-0.850**	-2.047***
	(0.390)	(0.798)	(0.507)	(0.423)	(0.408)	(0.394)	(0.390)	(0.288)	(0.390)	(0.514)
Lagged growth in population	-0.183***	-0.183**	-0.120**		-0.205***	-0.175***	-0.184***	-0.185***	-0.183***	-0.177***
	(0.0500)	(0.0748)	(0.0591)		(0.0513)	(0.0499)	(0.0498)	(0.0377)	(0.0500)	(0.0643)
Current growth in population				-0.255***						
				(0.0740)						
Observations	8128	8128	8128	8128	6642	7897	8128	8128	8128	5357
Adjusted R-squared	0.347	0.347	0.355	0.349	0.339	0.355	0.347	0.371	0.347	0.371
# clusters	1498	78	1498	1498	1110	1497	1498	1498	1498	939
weights	no	no	yes	no	no	no	no	no	no	no
standard errors clustered at	region	country	region	region	region	region	region	region	region	region

Notes: Column (1) is replicates column (7) from Table 3. Column (2) is identical to column (1), but standard errors are clustered at the country level. Column (3) is a weighted regression, using the region's population as a share of total country population. Column (4) uses the current, rather than lagged, growth in population. Column (5) drops regions that have no universities. Column (6) drops regions that opened their first university in the last period. Column (7) winsorizes the top 5% and bottom 5% the of observations of regional university growth. Column (8) winsorizes the top 5% and bottom 5% the of observations of regional GDP per capita growth. Column (9) adds a dummy for regions where more than 5 per cent of the universities have missing locational information and are not regionalized. Column (10) omits countries where the total number of universities in WHED in 2010 is less than half of the number according to a recent external source (Webometrics in 2015).

Table A5: GROWTH REGRESSIONS BY CONTINENT

Dependent variable: Regional growth in GDP per capita	(1)	(2)	(3)	(4)	(5)
Sample	All	Europe & North America	Latin America	Asia	Africa
Lagged growth in #universities	0.0445***	0.0173*	0.0722***	0.0574**	0.0271
	(0.0105)	(0.00938)	(0.0208)	(0.0267)	(0.0642)
Lagged level of regional GDP per capita	-0.0127***	-0.0162***	-0.0137***	-0.0105***	-0.0325***
	(0.00131)	(0.00259)	(0.00247)	(0.00242)	(0.00507)
Lagged level of country GDP per capita	-0.0213***	-0.0338***	-0.0247***	-0.0149*	-0.157***
	(0.00422)	(0.00534)	(0.00688)	(0.00819)	(0.0505)
Lagged level of population	-0.0855**	0.0197	-0.335***	-0.0199	-0.486
	(0.0387)	(0.0450)	(0.0936)	(0.0793)	(0.326)
Lagged growth in population	-0.113***	-0.174***	-0.109**	-0.0879	0.0334
	(0.0385)	(0.0491)	(0.0524)	(0.100)	(0.0556)
Dummy for capital in region	0.0110***	0.0116***	0.0117***	0.0168***	0.0186***
	(0.00168)	(0.00217)	(0.00343)	(0.00370)	(0.00648)
Observations	8128	3815	1821	2249	243
Adjusted R-squared	0.260	0.445	0.206	0.211	0.357
# clusters	1498	674	295	462	67

Notes: Column (1) replicates column (5) from Table 3. The other columns carry out an identical regression, but restricting the sample to the sample continent as labelled. Levels of GDP per capita and population are in natural logs. Latin America contains Mexico, Central America and South America.

Table A6: BARRO-STYLE REGRESSIONS WITH UNIVERSITIES

Dependent variable: Regional Growth of GDP per capita	(1)	(2)	(3)	(4)
Lagged #universities			0.00243***	0.00189**
			(0.000759)	(0.000756)
Lagged level of population/100			-0.234***	-0.200***
			(0.0679)	(0.0673)
Lagged level of regional GDP per capita	-0.0141***	-0.0178***	-0.0150***	-0.0184***
	(0.00142)	(0.00165)	(0.00148)	(0.00169)
Lagged level of country GDP per capita	-0.0361***	-0.0321***	-0.0356***	-0.0318***
	(0.00353)	(0.00373)	(0.00353)	(0.00372)
Lagged level of population density	-0.000562	-0.00109**	-0.000221	-0.000731
	(0.000418)	(0.000424)	(0.000482)	(0.000488)
Lagged years of education		0.00399***		0.00386***
		(0.000569)		(0.000570)
Observations	8010	8010	8010	8010
Adjusted R-squared	0.273	0.279	0.274	0.280
# clusters	1504	1504	1504	1504

Notes: Column (1) replicates Gennaioli et al (2014) Table (5), column (8), with geographic controls, year and country fixed effects, but omits years of education. There are more observations because we have interpolated GDP per capita in the sample (Gennaioli et al only interpolate years of education and population). Column (2) adds years of education. Column (3) replicates column (1),but adds the five year lagged level of universities in a region, and lagged population. Column (4) then adds years of education to the specification in column (3). Standard errors are clustered at the regional level. Levels of GDP per capita, population and population density are in natural logs. Years of schooling are not logged.

Table A7: LONG DIFFERENCE GROWTH ON LEVELS REGRESSIONS

Dependent Variable:	50 year differences		40 year differences		30 year differences	
Average annual GDP per capita growth	(1)	(2)	(4)	(5)	(6)	(7)
Lagged #universities	-0.000295	0.00346**	-0.000763**	0.00198	0.000266	0.00238**
	(0.000396)	(0.00151)	(0.000319)	(0.00128)	(0.000603)	(0.00105)
Lagged level of regional GDP per capita		-0.0125***		-0.00650***		-0.0127***
		(0.00239)		(0.00169)		(0.00164)
Lagged level of population		-0.00258*		-0.00208*		-0.000817
		(0.00143)		(0.00116)		(0.000876)
Change in population		0.0842		0.0227		-0.0624
		(0.0644)		(0.0565)		(0.0715)
Observations	188	188	250	250	464	464
Adjusted R-squared	-0.003	0.423	0.014	0.289	-0.002	0.748
Country fixed effects	no	yes	no	yes	no	yes

Notes: This table shows long differences to 2000: so columns (1) to (2) show regressions for the sample where data are available for the period 1960-2000; and columns (3)-(4) show regressions for the period 1970-2000. Column (1) is a simple correlation of the average annual growth in regional GDP per capita over 1960-2000 on the natural log of 1+ the number of universities in 1960. Column (2) adds country fixed effects, the 1960 natural log of the level of regional GDP per capita, the 1960 natural log of the level of population, the 1960-2000 change in population and country fixed effects. Columns (3) and (4) do the same for the 30 year difference to 2000, for which more data are available. Robust standard errors are shown in parentheses.

Table A8: UNIVERSITIES AND YEARS OF EDUCATION

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	Δ GDPpc	Δ GDPpc	Δ GDPpc	Δ GDPpc	Δ GDPpc	Δ GDPpc	Δ Years Educ.	Δ Years Educ.
Lagged growth in #universities	0.0445***	0.0532***	0.0518***	0.0518***	0.0511***	0.0517***	0.0425***	0.0196**
	(0.0105)	(0.0122)	(0.0121)	(0.0122)	(0.0121)	(0.0121)	(0.0118)	(0.00774)
Lagged growth in years of education			0.0606**		0.0489**	0.0538**		
			(0.0235)		(0.0237)	(0.0242)		
Current growth in years of education				0.0678	0.0502	0.138***		
				(0.0415)	(0.0417)	(0.0447)		
Lagged years of education						0.00489***		
						(0.000711)		
Lagged level of regional GDP pc	-0.0127***	-0.0112***	-0.0110***	-0.0110***	-0.0109***	-0.0154***		-0.00335***
	(0.00131)	(0.00154)	(0.00155)	(0.00156)	(0.00156)	(0.00181)		(0.000518)
Lagged level of country GDP pc	-0.0213***	-0.0447***	-0.0461***	-0.0453***	-0.0463***	-0.0422***		0.00870***
	(0.00422)	(0.00431)	(0.00433)	(0.00438)	(0.00436)	(0.00457)		(0.00212)
Lagged level of population/100	-0.0855**	-0.155***	-0.153***	-0.149***	-0.149***	-0.184***		-0.0866***
	(0.0387)	(0.0459)	(0.0458)	(0.0458)	(0.0458)	(0.0467)		(0.0250)
Lagged growth in population	-0.113***	-0.0221	-0.0243	-0.0226	-0.0243	-0.0284		0.00710
	(0.0385)	(0.0460)	(0.0460)	(0.0461)	(0.0460)	(0.0456)		(0.0141)
Dummy for capital in region	0.0110***	0.0106***	0.0108***	0.0107***	0.0109***	0.00714***		-0.00215***
	(0.00168)	(0.00189)	(0.00189)	(0.00189)	(0.00190)	(0.00181)		(0.000723)
Observations	8128	6117	6117	6117	6117	6117	6117	6117
Adjusted R-squared	0.260	0.288	0.289	0.289	0.290	0.299	0.203	0.627
# clusters	1498	1343	1343	1343	1343	1343	1343	1343

Notes: Growth in years of education is the log difference. Column (1) replicates column (5) from Table 3. Column (2) restricts to the sample for which the change in years of education is available. Column (3) drops the lagged growth in years of education. Column (4) adds the contemporaneous change in years of education. Column (5) includes both lagged and contemporaneous changes. Column (6) further adds the lagged level of years of education (unlogged). Column (7) regresses the change in years of education on the lagged growth in universities, with country dummies, but no other controls. Column (8) adds all the other controls.

Table A9: ROBUSTNESS ON WORLD VALUES SURVEY ANALYSIS

Dependent Variable: View of democracy	(1) Approval	(2) Approval	(3) Approval	(4) Approval	(5) Best	(6) Approval	(7) Approval	(8) Approval	(9) Approval	(10) Approval
15 year lagged ln(1+#universities per capita)	0.0229**	0.0137			0.0332**	-0.0525***	0.0229*	0.0547**	0.0290***	0.0397**
	(0.00997)	(0.0120)			(0.0163)	(0.0164)	(0.0120)	(0.0218)	(0.0110)	(0.0162)
OECD dummy X 15 year lagged ln(1+#universities per capita)		0.0246								
		(0.0177)								
5 year lagged ln(1+#universities per capita)			0.0126							
			(0.00822)							
30 year lagged ln(1+#universities per capita)				-0.00217						
				(0.00913)						
Observations	138511	138511	138511	138511	48181	138511	138511	138511	100782	138511
Adjusted R-squared	0.085	0.085	0.085	0.085	0.099	0.018	0.085	0.071	0.083	
# clusters	693	693	693	693	335	693	58	693	691	693
Estimator	OLS	OLS	OLS	OLS	OLS	OLS	OLS	weighted OLS	OLS	Oprobit
Standard errors clustered at	region	region	region	region	region	region	country	region	region	region
Country and year dummies	yes	yes	yes	yes	yes	no	yes	yes	yes	yes
Sample	all	all	all	all	all	all	all	all	drop students, graduates	all

Notes: Column (1) replicates column (4) from Table 9. Column (2) includes an OECD dummy (not reported) and interaction between this and lagged university density. Column (3) is identical to column (1), but uses the five year lagged university density. Column (4) uses the thirty year lagged university density. Column (5) has a different dependent variable: the view that democracy is “best”. Column (6) omits country and year dummies. Column (7) clusters standard errors at the country level. Column (8) uses weighted OLS, weighting each region by its population as a share of the country’s total population. Column (9) drops graduates and students from the sample. Column (10) is estimated using an Ordered Probit model.

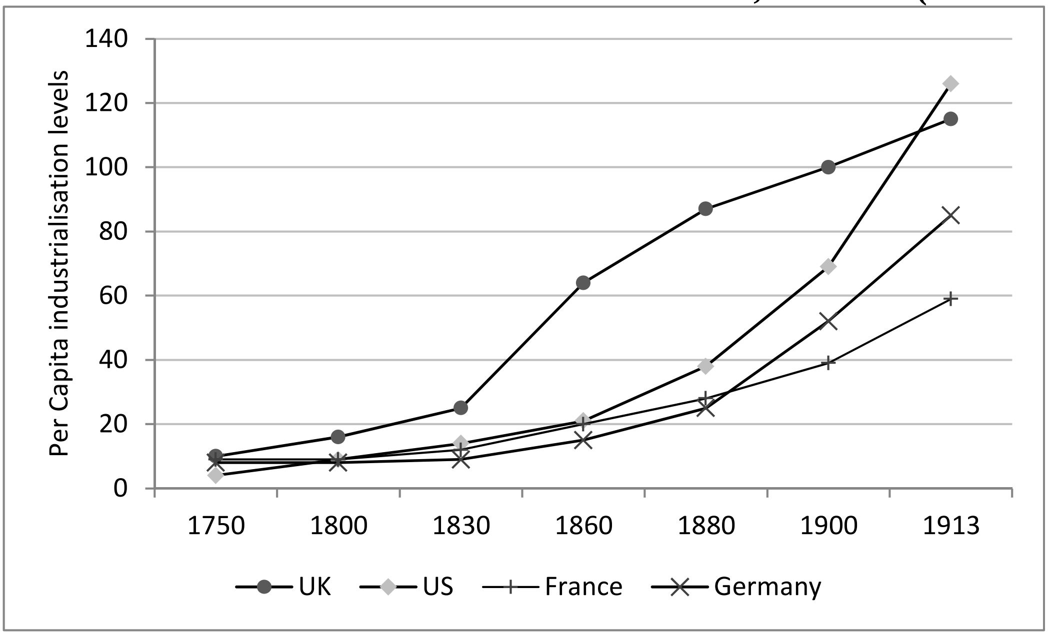
Table A10: MATCHING WHED AND 1960 YEARBOOKS

Outcome	"University"	Other	Total
Match - exact	570	65	635
Match - fuzzy	653	138	791
Match - manual	384	696	1080
Not in WHED 2010	40	115	155
Death	33	0	33
Total	1680	1014	2694

Notes: This table reports the outcome of the matching process between WHED and historical yearbooks, by universities and other types of institution.

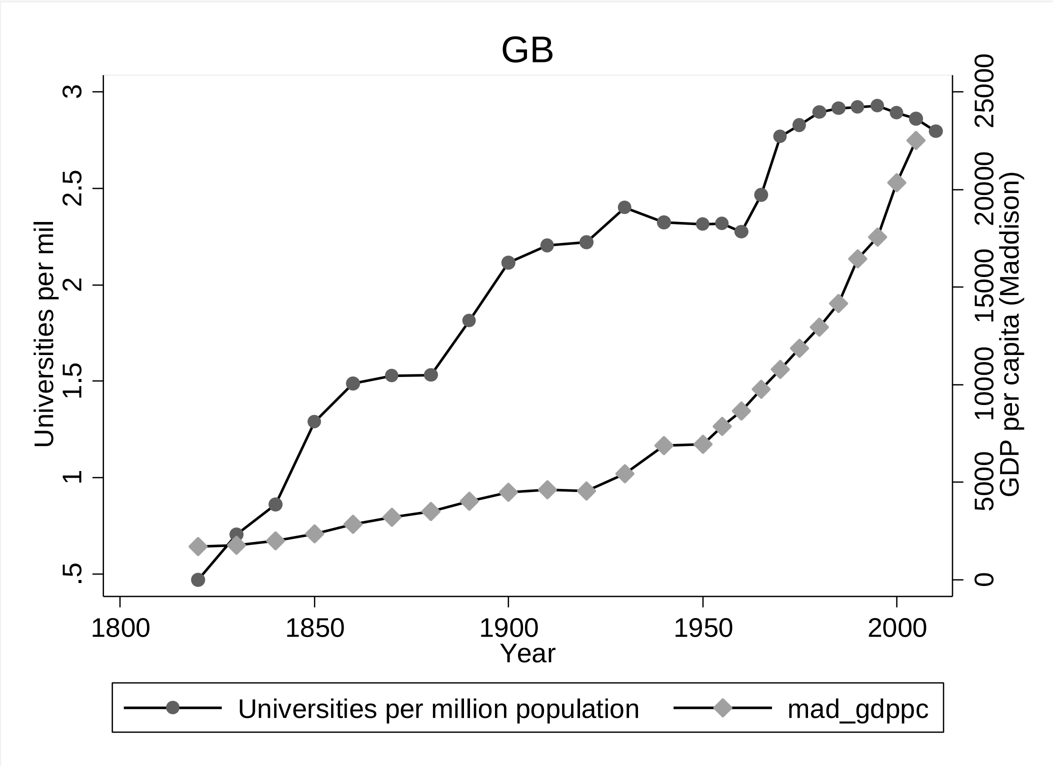
APPENDIX FIGURES

Figure A1: PER CAPITA INDUSTRIALISATION LEVELS, 1959-1913 (UK 1900=100)



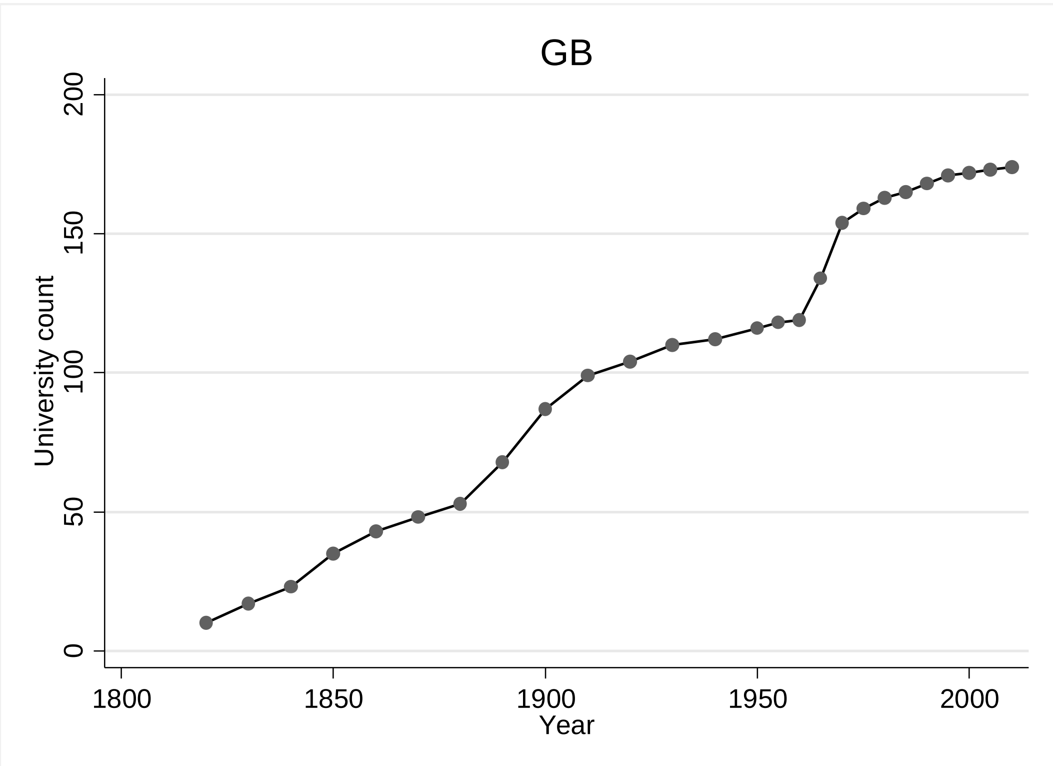
Notes: Graph based on Table 9, Bairoch (1982); taken from Baldwin (2012)

Figure A2: UNIVERSITIES AND INDUSTRIALISATION IN THE UK
Panel A: University density and GDP per capita trends



Notes: This chart shows the evolution of university density (universities per million people) and GDP per capita over time. *Source: WHED and Maddison GDP per capita data.*

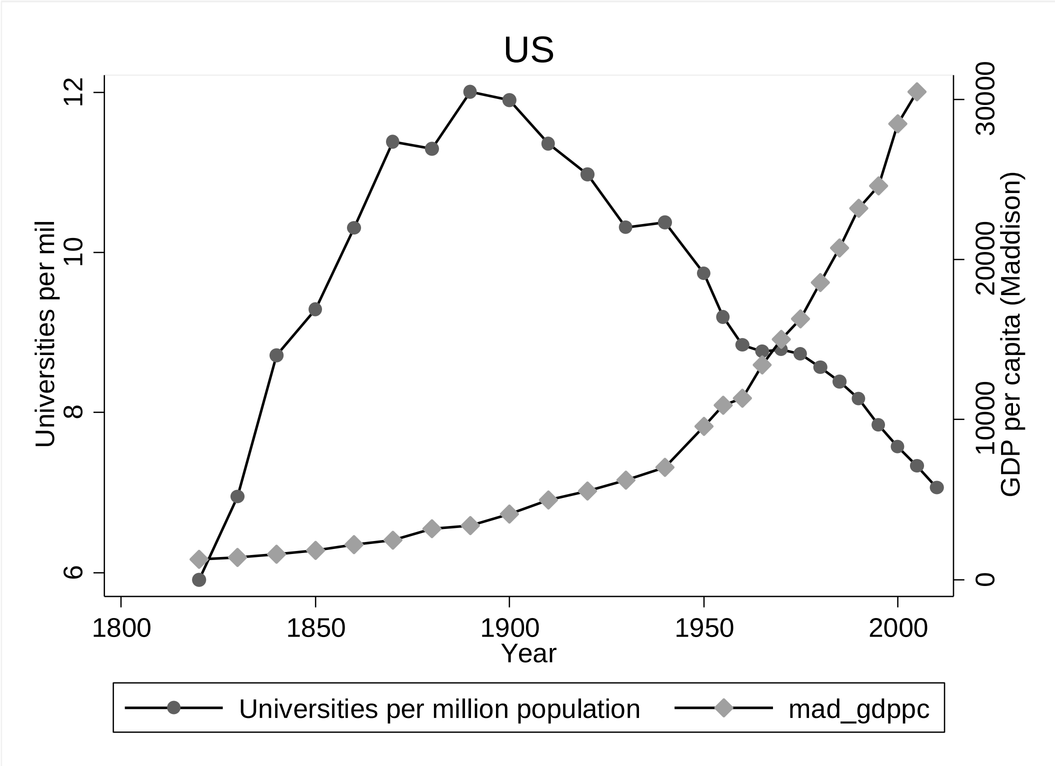
Panel B: University count trend



Notes: This chart shows the evolution of university count over time. *Source: WHED*

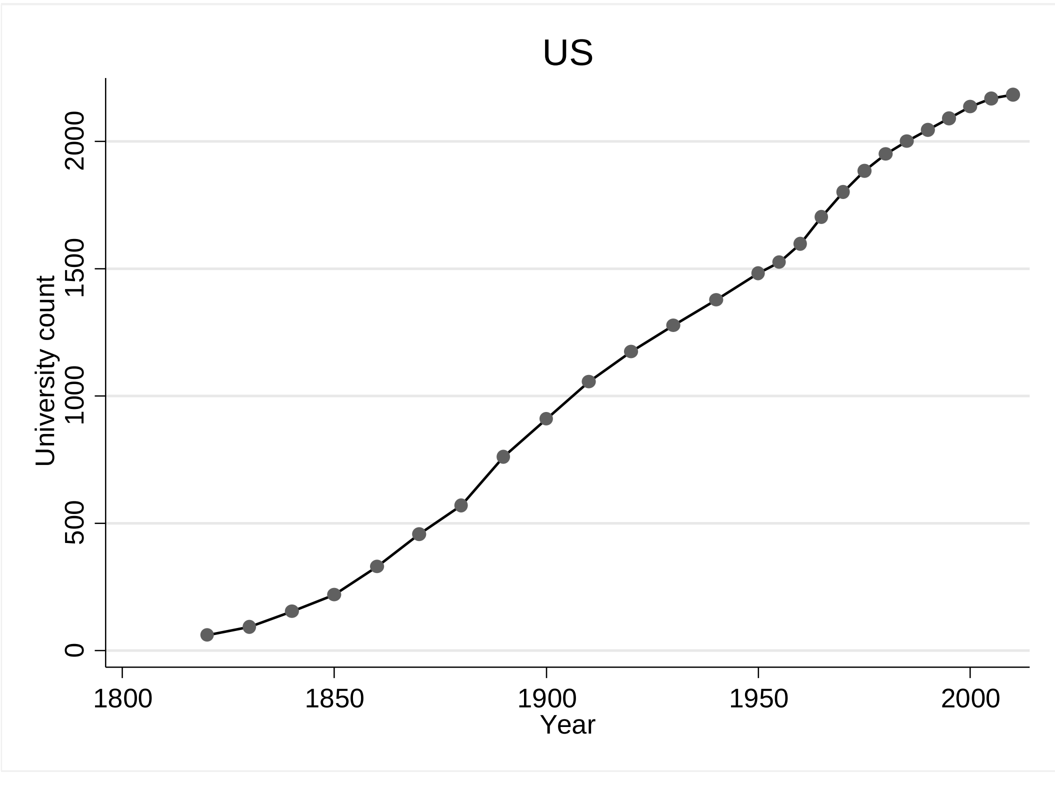
Figure A3: UNIVERSITIES AND INDUSTRIALISATION IN THE US

Panel A: University density and GDP per capita trends



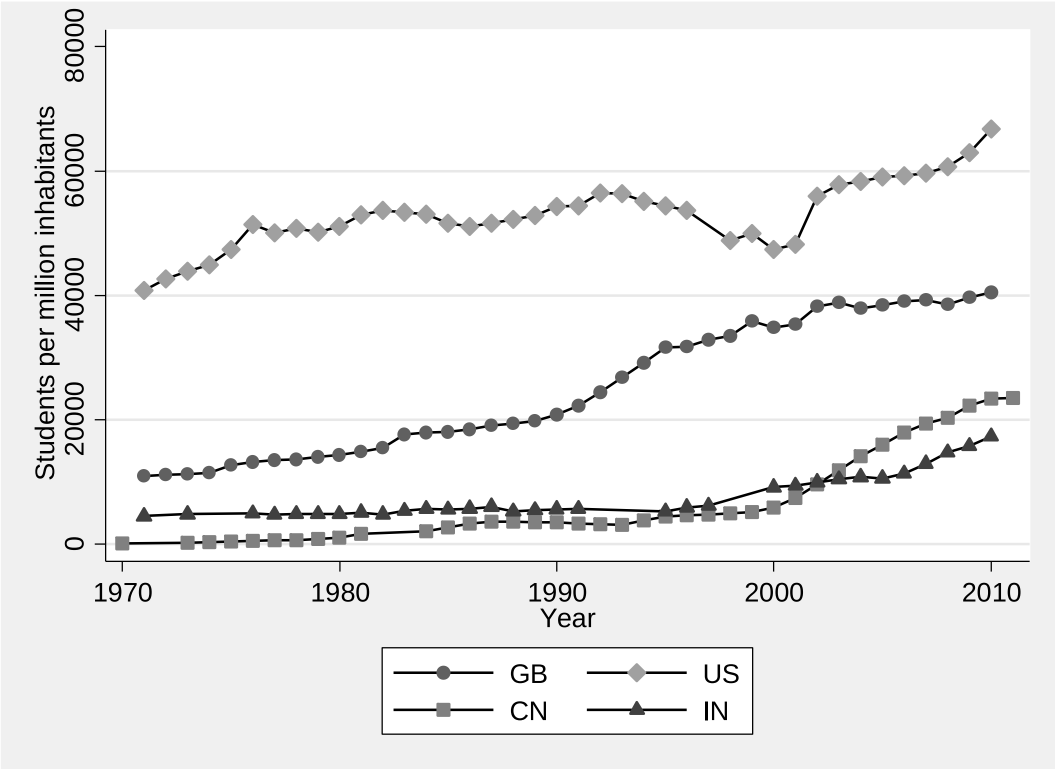
Notes: This chart shows the evolution of university density (universities per million people) and GDP per capita over time. *Source: WHED and Maddison GDP per capita data.*

Panel B: University count trend



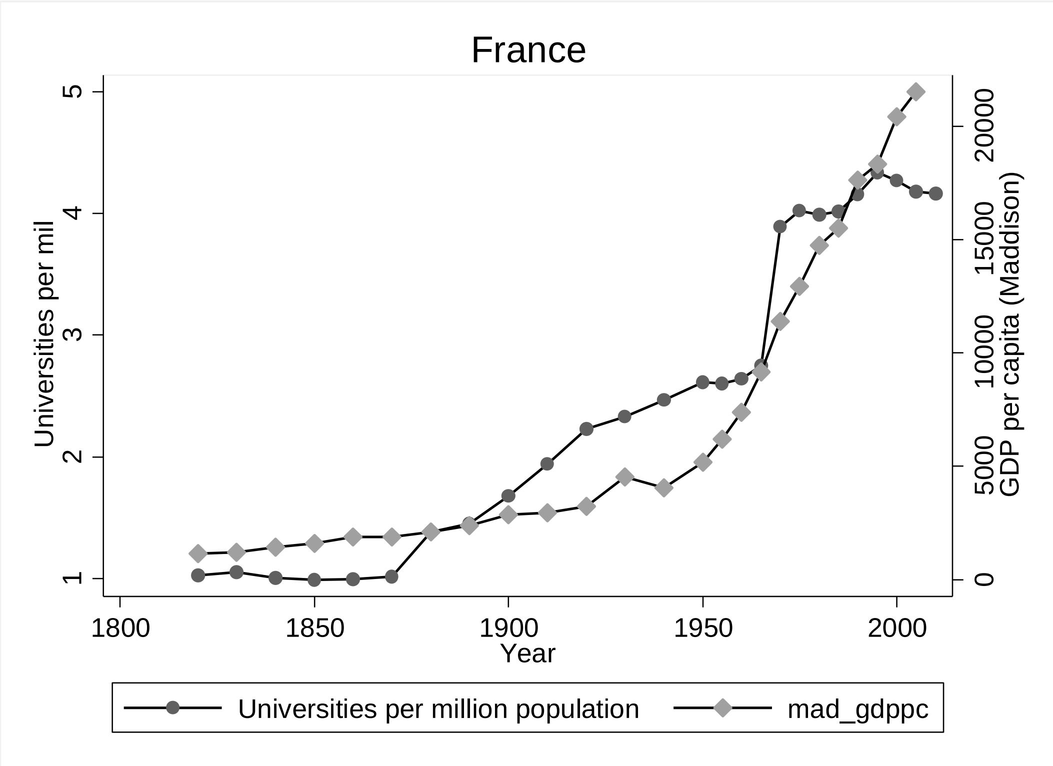
Notes: This chart shows the evolution of university count over time. *Source: WHED*

Figure A4: TRENDS IN STUDENT NUMBERS NORMALISED BY POPULATION



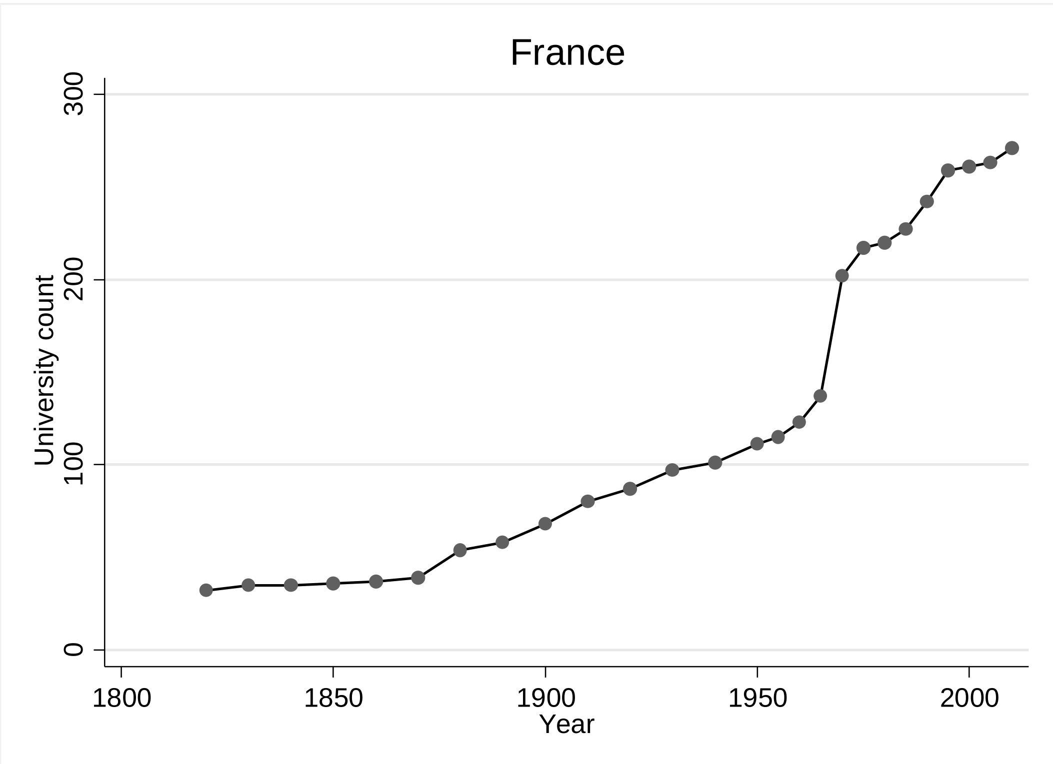
Notes: Number of students in tertiary education per million inhabitants. *Source: UNESCO*

Figure A5: UNIVERSITIES AND INDUSTRIALISATION IN FRANCE
Panel A: University density and GDP per capita trends



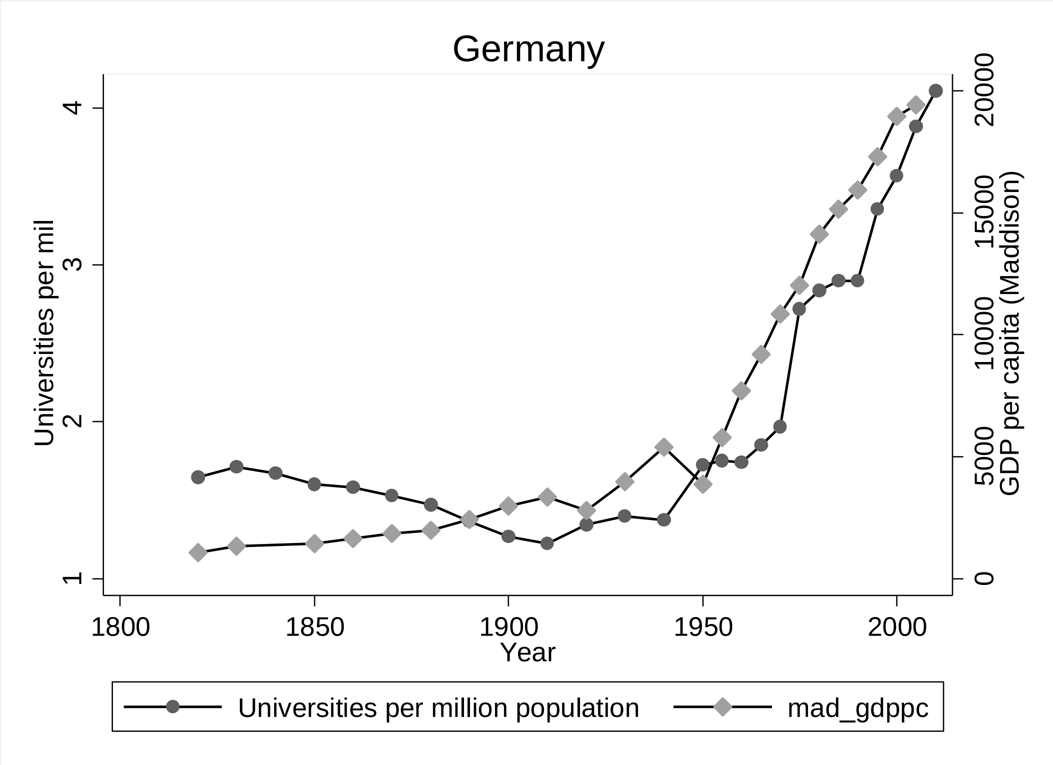
Notes: This chart shows the evolution of university density (universities per million people) and GDP per capita over time. *Source: WHED and Maddison GDP per capita data.*

Panel B: University count trend



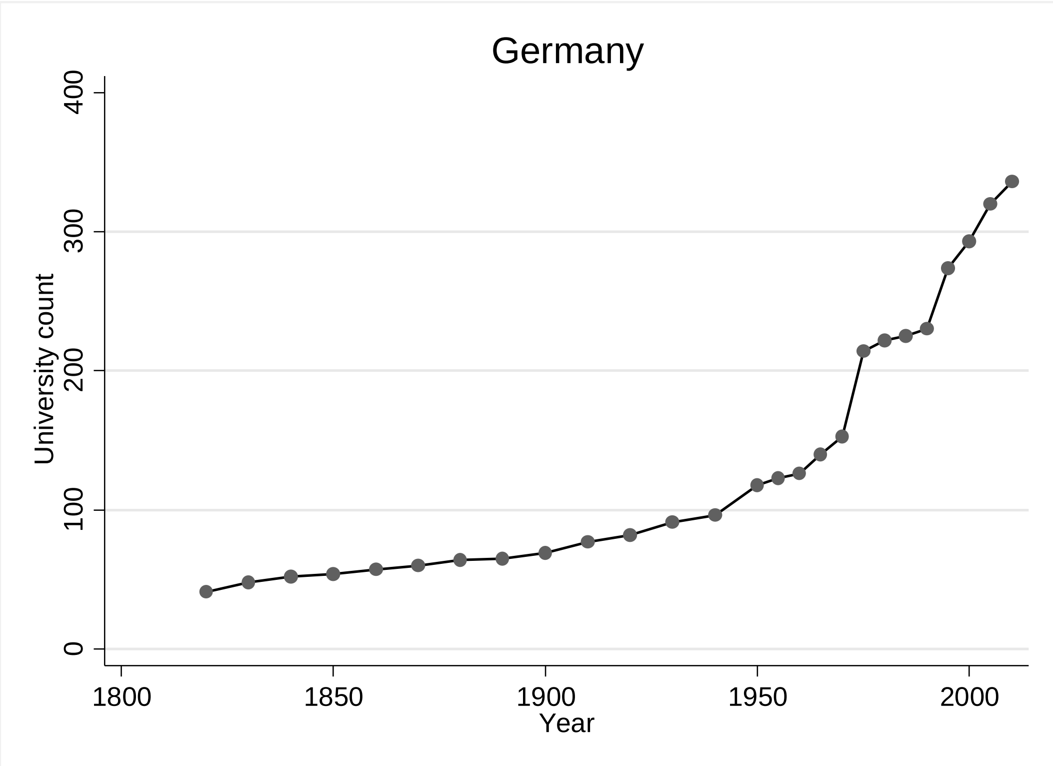
Notes: This chart shows the evolution of university count over time. *Source: WHED*

Figure A6: UNIVERSITIES AND INDUSTRIALISATION IN GERMANY
Panel A: University density and GDP per capita trends



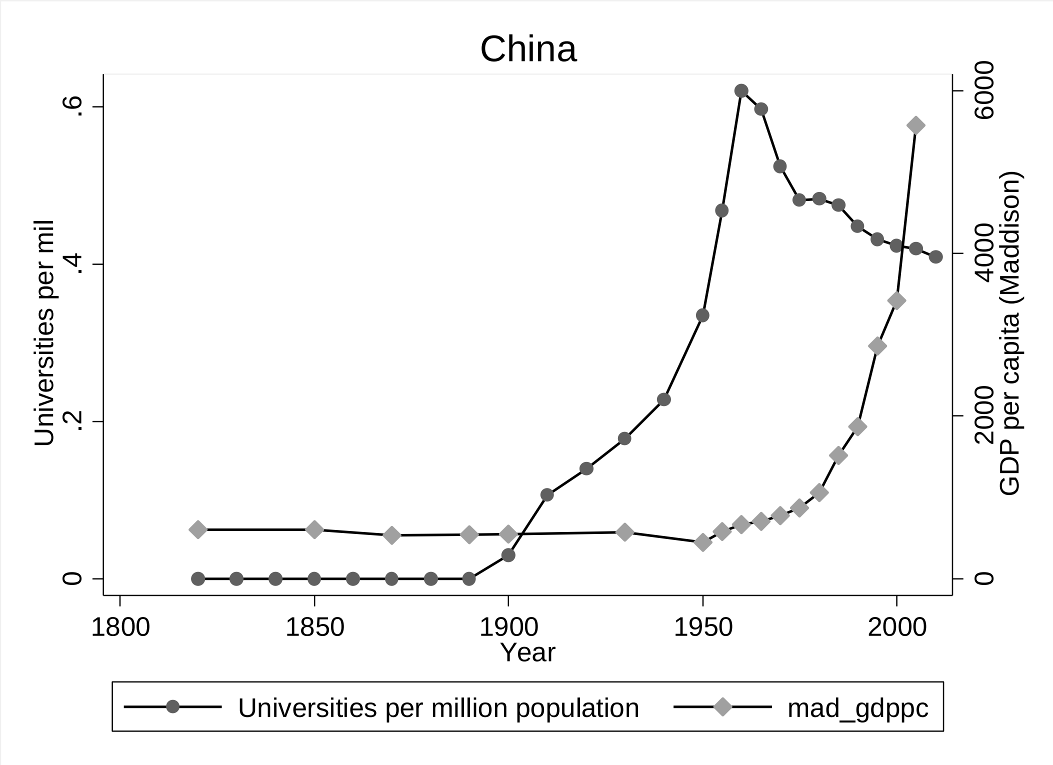
Notes: This chart shows the evolution of university density (universities per million people) and GDP per capita over time. *Source: WHED and Maddison GDP per capita data.*

Panel B: University count trend



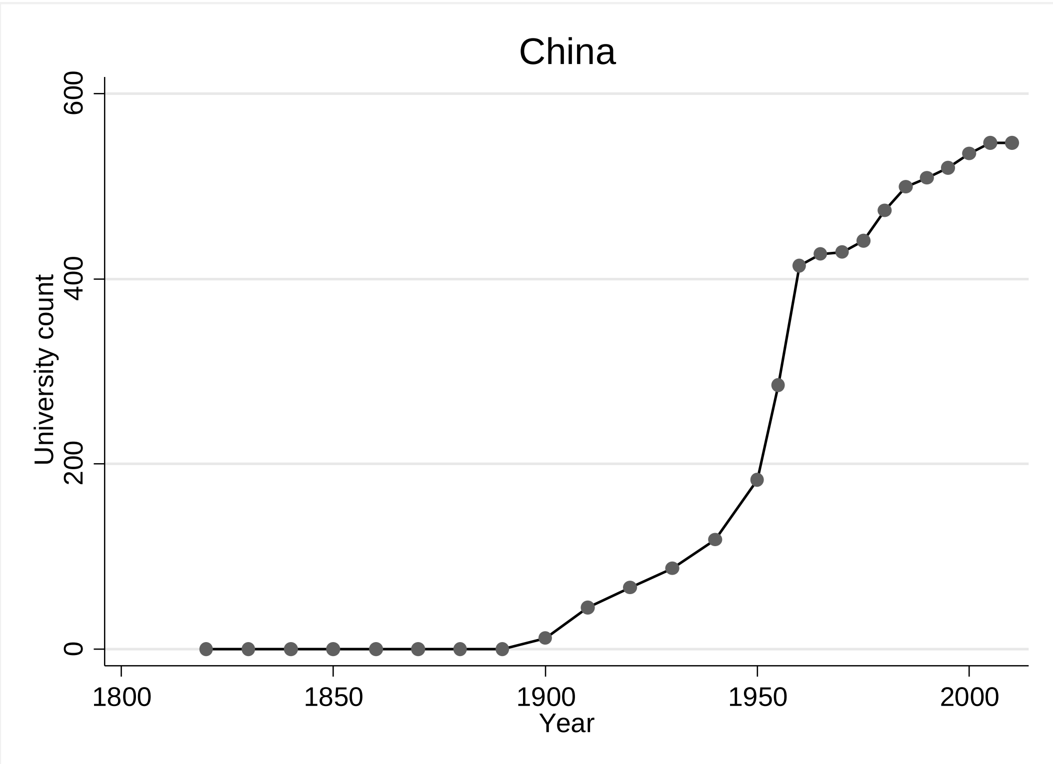
Notes: This chart shows the evolution of university count over time. *Source: WHED*

Figure A7: UNIVERSITIES AND INDUSTRIALISATION IN CHINA
Panel A: University density and GDP per capita trends



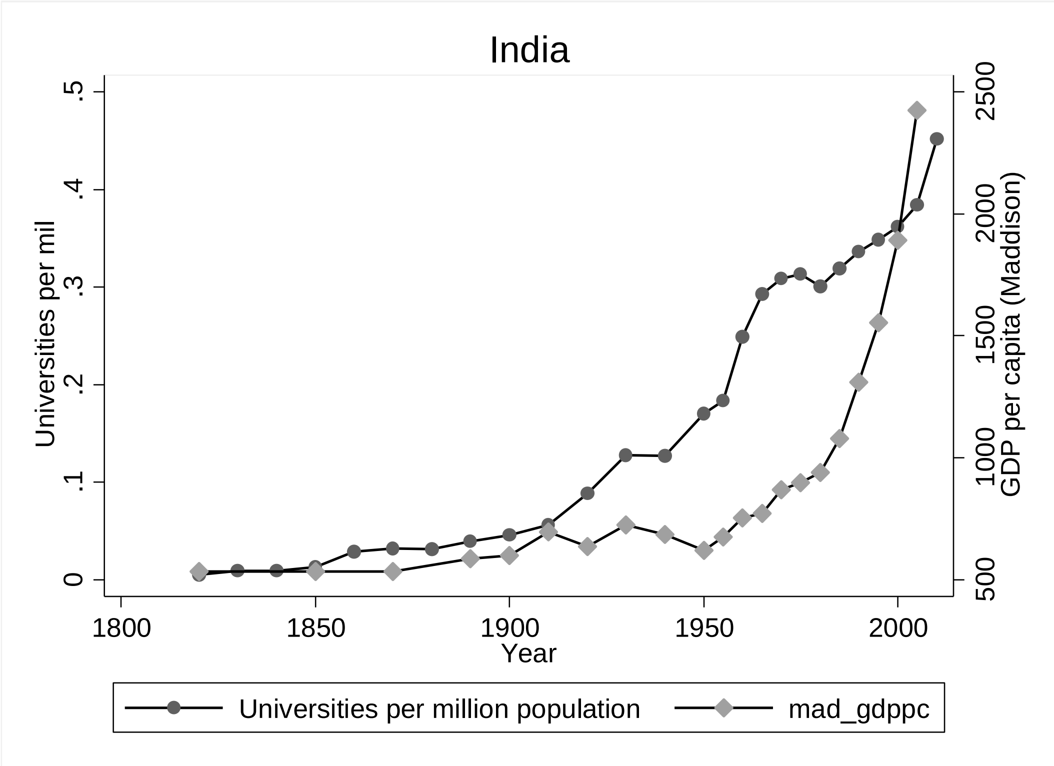
Notes: This chart shows the evolution of university density (universities per million people) and GDP per capita over time. *Source: WHED and Maddison GDP per capita data.*

Panel B: University count trend



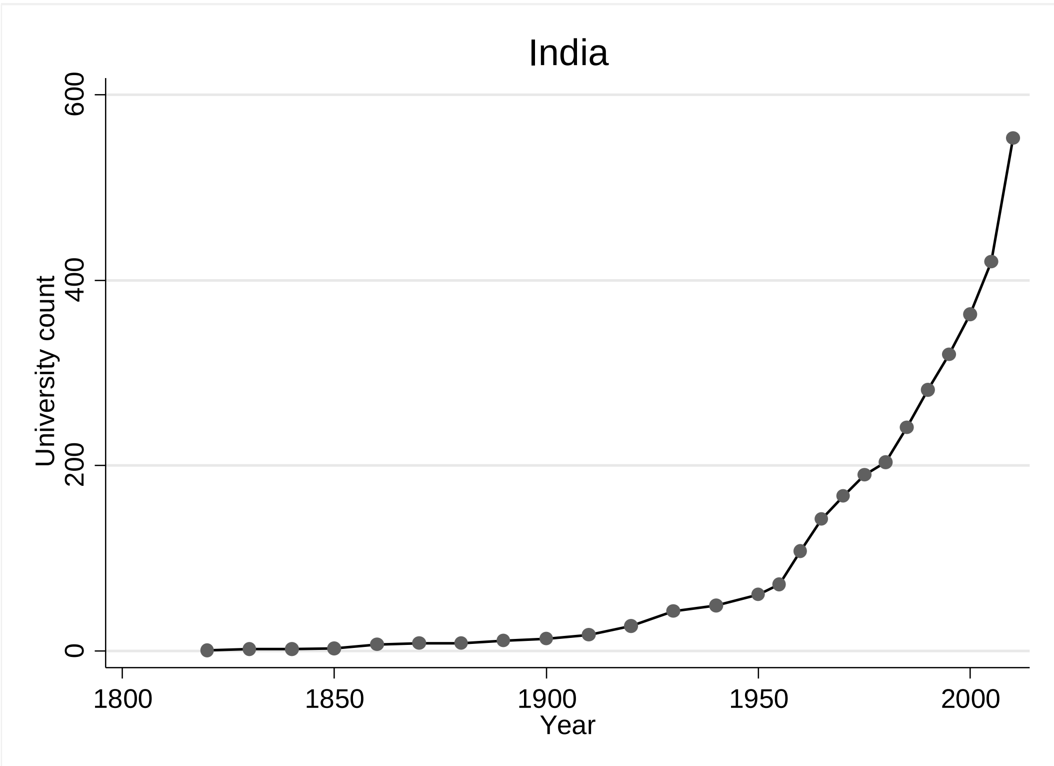
Notes: This chart shows the evolution of university count over time. *Source: WHED*

Figure A9: UNIVERSITIES AND INDUSTRIALISATION IN INDIA
Panel A: University density and GDP per capita trends



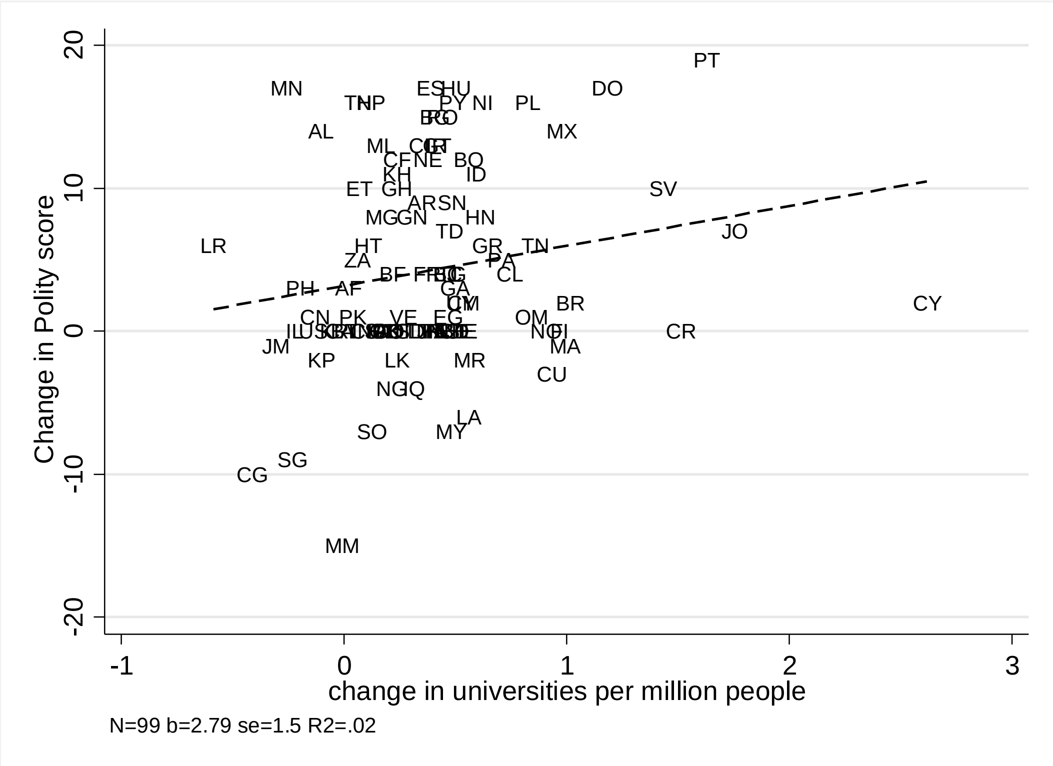
Notes: This chart shows the evolution of university density (universities per million people) and GDP per capita over time. *Source: WHED and Maddison GDP per capita data.*

Panel B: University count trend



Notes: This chart shows the evolution of university count over time. *Source: WHED*

Figure A10: CHANGE IN UNIVERSITIES AND CHANGE IN DEMOCRACY



Notes: Each observation is a country in 2000. *Source: WHED and Polity 2 variable from Polity IV*