DEPARTMENT OF ECONOMICS
UNIVERSITY OF COPENHAGEN

PhD thesis

Jeanet Sinding Bentzen

Why are Some Countries Richer than Others?

Cross-Country and Cross-Regional Empirical Studies

Academic advisors: Carl-Johan Dalgaard and Thomas Barnebeck Andersen

Submitted: October 11, 2011

Contents

Acknowledgements

I sincerely thank my supervisor, coauthor, and friend, Carl-Johan Dalgaard, for great support throughout my entire thesis and extremely enjoyable collaboration of which I have learnt so much. Carl-Johan has been excellent at taking my ideas and opinions seriously, guiding me on my way to becoming a critical researcher, and inspiring me tremendously within the field of Growth Economics. I also thank my second supervisor and coauthor, Thomas Barnebeck Andersen, for great support and collaboration and for always sharing his honest opinions about my ideas. Also a great thank to my remaining two coauthors, Paul Sharp and Pablo Selaya, for a truly enjoyable collaboration. Without all of my four coauthors, my PhD thesis would probably have been somewhat lonely and I would certainly not have obtained half the knowledge that I possess today about structuring my arguments, writing good papers, and doing convincing empirical work.

I have indeed benefited from the social and academic life here at the Department of Economics in Copenhagen. In particular, I thank all members of our Macro Reading Group for great discussions and fun afternoons and evenings at Krut's Karport. I also thank my office mate and friend, Asger, for pleasant days in the office and for hopefully productive discussions about ideas for future papers. I thank the rest of my colleagues at the Department of Economics for providing such a pleasant environment, including the always so cheerful and helpful administrative staff.

During the writings of my PhD thesis, I spent one semester at Universitat Pompeu Fabra in Barcelona, attending some truly great PhD courses. In particular, I thank my friends there for great discussions and hospitality, Joachim Voth for his excellent course "Research Seminar: Economic History", which was of great inspiration to me, and Antonio Ciccone for productive discussions.

I thank my husband, Christian, for invaluable loving support and for always pointing me back on track when I stumble. Last, I thank my little Oskar for making me laugh every day.

Jeanet Sinding Bentzen

Copenhagen, October 2011

Evaluation committee

Sascha O. Becker, The University of Warwick

David Dreyer, University of Copenhagen

Mark Gradstein, Ben-Gurion University of the Negev

Summary

Why are some countries richer than others? An immediate answer is that they invest more and innovate or adopt more new technologies. But why do some countries then invest more and experience more technological progress? The economic literature agrees on three deeper factors: Institutions, culture, and geography. The idea is that some countries have set up institutions that create better environments for investments and technological progress, or that some countries were simply enriched with a culture or a geography that created such a productive environment. With my PhD dissertation, I set out to explore subsets of these causes empirically. My main method is cross-sectional empirical analysis. The cross-sectional unit varies from countries to states or counties within one country. The latter has the advantage of being more likely to satisfy the "all-other-things-equal" assumption, but also the disadvantage of perhaps not being able to extend the conclusions to the World. My dissertation consists of four chapters, which are all independent journal articles, described briefly below.

Chapter 1 "Lightning, IT Diffusion and Economic Growth Across US States" (joint with Thomas Barnebeck Andersen, Carl-Johan Dalgaard, and Pablo Selaya) shows that since the 1990s, growth across US states has become more and more sensitive to a specific (perhaps at first glance, peculiar) climatic phenomenon; lightning. US states with more lightning experience lower growth rates of GSP per capita. Before the 1990s, there was no relation between the two. We show that this increased sensitivity towards lightning is not because lightning has increased over time, nor is it because lightning is correlated with another natural phenomenon exhibiting this pattern, and finally not because lightning is correlated with another important growth determinant that follows this pattern. Instead, we argue that it is because of the emergence of the Internet in 1991 and the following increased importance of digital technologies for the economy. Computer chips inherent in all IT capital are highly sensitive towards shocks to the electricity supply and more and more so as computers get smaller and smaller. Lightning activity can cause such shocks, a problem acknowledged by engineers, private sector firms, and the like. Consumers can buy surge protectors, which increases the user cost of IT capital. We show empirically that US states with more lightning undertake significantly fewer IT investments, which results in significantly lower growth of GSP per capita across US states. We conclude that this increasing macroeconomic sensitivity to lightning may be due to the increasing importance of digital technologies for the growth process.

Chapter 2 entitled "Religious Orders and Growth through Cultural Change in Pre-Industrial England" (joint with Thomas Barnebeck Andersen, Carl-Johan Dalgaard, and Paul Sharp) investigates another

determinant of long run productivity, namely culture. We go back to the roots of Protestantism, specifically to a group of Monks, who are said to be early proponents of Protestantism; the Cistercians. The Cistercians were known for their high work ethics and thrift, values usually associated with Protestantism. Indeed, we show that regions with more Cistercians are more likely to possess values of hard work and thrift today. We set out to test whether these values were beneficial for growth, as argued for by Max Weber. We show that English counties with more Cistercian Monasteries as a share of total Monasteries experienced higher population densities (a measure of prosperity in the Malthusian era) over the period 1377-1801. This finding is robust to accounting for various geographic features important for development, the remaining Monk Orders, regional effects, potential endogeneity of the location of Cistercian Monasteries and potential other explanations such as trade, technology adoption and human capital. We conclude that the Cistercian monks spread a culture to the surrounding society, which created growth advantages, even long after the dissolution of the monasteries.

Chapter 3, "How Bad is Corruption? Cross-Country Evidence of the Impact of Corruption on Economic Prosperity", investigates a third determinant of prosperity differences; corruption. I set out to identify the impact of corruption on GDP per capita across countries. The relation is potentially spurred by endogeneity; richer countries have more resources to combat corruption (reverse causality) and omitted factors are likely to influence both corruption and GDP simultaneously (omitted variables bias). I suggest to instrument corruption using specific cultural values as instruments. The idea is that cultures with more focus on the social group compared to the individual and cultures that emphasize more hierarchical power structures resulting in less questioning towards the people in power, will experience more corruption, since the corrupt rulers are faced with lower risks of getting caught and with higher benefits of being corrupt. I show that the cultural values Individualism and Power Distances (measured in the 1960s) are strong instruments for corruption in a regression on GDP per capita. The OID test also cannot reject that corruption is the only channel through which these particular culture dimensions influence GDP per capita. But the test is of low power and I attempt to fulfill the exclusion restrictions by tying my hands as much as possible when including the remaining determinants of GDP per capita; geography and the remaining dimensions of institutions and culture. I find that Individualism and Power Distances remain strong instruments for corruption and that corruption does indeed exert a negative significant influence on GDP per capita.

Chapter 4, entitled "Does the Internet Reduce Corruption? Evidence from U.S. States and across Countries" (joint with Thomas Barnebeck Andersen, Carl-Johan Dalgaard, and Pablo Selaya), explores corruption further by investigating one potential means through which to get rid of corruption; the Internet. The Internet increases the costs for corrupt rulers to engage in corruption as it makes it easier to spread information about corruption (e.g. journalists' blogs) and it enables e-governance which reduces the interaction between rulers and the populace. The aim of the paper is to identify the impact of the Internet on corruption. We do so empirically across countries and within one country; the US. The empirical challenge is that the relation is potentially biased by endogeneity; corrupt rulers might attempt to censor the Internet (like in China, for instance) and omitted factors might influence the Internet and corruption simultaneously. We use lightning as an instrument for the Internet, as argued for in Chapter 1. We find that lightning is a strong instrument for the Internet both across US states and across countries. Equipped with this strong instrument, we can identify the causal impact of the Internet on corruption. We find that the World Wide Web does indeed reduce the extent of corruption across countries and across US states.

Resumé (summary in Danish)

Hvorfor er nogle lande rigere end andre? Et umiddelbart svar er, at de investerer mere og opfinder eller anvender flere nye teknologier. Men hvorfor har nogle lande så investeret mere og oplevede mere teknologisk fremskridt? Økonomisk litteratur peger på tre dybereliggende faktorer: Institutioner, kultur og geografi. Ideen er, at nogle lande har oprettet institutioner, der skaber bedre grobund for investeringer og teknologiske fremskridt, eller simpelthen er blevet beriget med en kultur eller en geografi, der har skabt et sådant produktivt miljø. Med min ph.d. afhandling undersøger jeg empirisk nogle af disse årsager. Min primære metode er tværsnits empirisk analyse, hvor tværsnitsenheden er både lande, men også stater og amter indenfor et enkelt land. Sidstnævnte har den fordel, at opfylde "alt-andet-lige" antagelsen med større sandsynlighed, men også ulempen, at resultaterne måske ikke kan ekstrapoleres til at omfatte hele verden. Min afhandling består af fire kapitler, som alle er uafhængige tidsskriftsartikler, kort beskrevet nedenfor.

Kapitel 1 "Lightning, IT Diffusion and Economic Growth Across US States" (skrevet sammen med Thomas Barnebeck Andersen, Carl-Johan Dalgaard og Pablo Selaya) viser, at økonomisk vækst på tværs af amerikanske delstater siden 1990'erne er blevet mere og mere følsom overfor et bestemt (måske ved første øjekast ejendommeligt) naturfænomen; lynnedslag. Stater med mere lyn har oplevet lavere vækst. Før 1990'erne var der ingen sammenhæng mellem de to. Vi viser, at dette ikke skyldes en stigning i lyn intensiteten over tid, ej heller at lyn er korreleret med andre naturfænomener, der udviser dette mønster, og heller ikke kan forklares ved at lyn er korreleret med andre vigtige vækstdeterminanter, der følger et sådant mønster. I stedet viser vi, at det skyldes internettets frembrud i 1991 og den stigende betydning af digitale teknologier for økonomisk udvikling derefter. Computerchips, som er en del af al IT-kapital, er ekstremt følsomme overfor selv små stød til elforsyningen, og følsomheden bliver kun forstærket i takt med at computere bliver mindre og mindre. Lynnedslag kan forårsage sådanne stød; et problem, som anerkendes af ingeniører, private virksomheder og lignende. Man kan købe sig til en overspændingsbeskytter, hvilket øger omkostningerne. Vi viser empirisk, at amerikanske stater med højere lyn intensitet foretager markant færre IT-investeringer, hvilket resulterer i væsentligt lavere vækst i GSP per indbygger. Vi konkluderer, at denne stigende makroøkonomiske følsomhed overfor lyn meget vel kan skyldes den stigende betydning af digitale teknologier for vækst processen.

I kapitel 2 "Religious Orders and Growth through Cultural Change in Pre-Industrial England" (skrevet med Thomas Barnebeck Andersen, Carl-Johan Dalgaard og Paul Sharp) undersøger vi en anden afgørende faktor for langsigtsproduktivitet, nemlig kultur. Vi går tilbage til en af protestantismens rødder, mere specifikt en gruppe munke, der siges at være tidlige fortalere for protestantismen, Cistercienserne. Cistercienserne var kendt for deres høje arbejdsmoral og sparsommelighed, værdier, der som regel forbindes med protestantismen. Vi viser, at engelske regioner med flere cisterciensere er mere tilbøjelige til at have høj arbejdsmoral og sparsommelighed i dag. Vi tester Max Weber's hypotese om at netop disse værdier var gavnlige for kapitalismens frembrud. Vi viser, at engelske amter med flere cistercienser klostre som andel af det samlede antal klostre oplevede højere befolkningstæthed (et mål for velstand i Malthusiansk økonomi) i perioden 1377-1801. Resultatet er robust overfor at tage højde for forskellige geografiske forhold af betydning for velstand, forskellige andre munkeordener, regionale effekter, potentiel endogenitet i placeringen af cistercienser klostrene og andre potentielle forklaringer såsom handel, teknologi adoption og humankapital. Vi konkluderer, at cisterciensermunkene spredte en kultur til det omgivende samfund, der skabte vækstfordele, selv længe efter klostrenes opløsning.

Kapitel 3 "How Bad is Corruption? Cross-Country Evidence of the Impact of Corruption on Economic Prosperity" omhandler en tredje bestemmende faktor for velstandsforskelle på tværs af lande; korruption. Jeg ønsker at identificere effekten af korruption på BNP per indbygger. Forholdet er potentielt forvredet af endogenitet; rigere lande har flere ressourcer til at bekæmpe korruption (omvendt kausalitet) og udeladte faktorer påvirker muligvis både korruption og BNP samtidigt (udeladt variabel bias). Jeg foreslår at instrumentere korruption med specifikke kulturelle værdier. Tanken er, at kulturer, der har større fokus på individets sociale gruppe i forhold til det enkelte individ og kulturer med en hierarkisk magtstruktur, der ansporer til færre spørgsmålstegn overfor de mennesker ved magten, vil opleve mere korruption, da de korrupte ledere står over for lavere risiko for at blive fanget og højere afkast ved at være korrupt. Jeg viser, at de kulturelle værdier Individualism og Power Distance (målt i 1960'erne) er stærke instrumenter for korruption i en regression på BNP per indbygger. Et OID test kan ikke afvise, at korruption er den eneste kanal, hvorigennem disse specielle kultur dimensioner påvirker BNP. Men testet er af lav styrke, og jeg forsøger at opfylde udeladelsesrestriktionerne ved at binde mine hænder så meget som muligt, når jeg inkluderer de resterende determinanter af BNP per indbygger; geografi og de øvrige dimensioner af institutioner og kultur. Individualism og Power Distance forbliver stærke instrumenter for korruption, og jeg finder, at korruption har en signifikant og negativ indflydelse på BNP per indbygger.

Kapitel 4 "Does the Internet Reduce Corruption? Evidence from U.S. States and across Countries" (skrevet med Thomas Barnebeck Andersen, Carl-Johan Dalgaard og Pablo Selaya) udforsker korruption yderligere ved at undersøge en faktor, der muligvis kan reducere korruption, nemlig internettet. Internettet øger omkostningerne for korrupte ledere ved at engagere sig i korruption, da internettet gør det lettere at sprede information om korruption (f.eks. journalisters blogs), og det muliggør e-forvaltning, som reducerer samspillet mellem magthaverne og befolkningen. Formålet med kapitlet er at identificere effekten af internettet på korruption. Det gør vi empirisk på tværs af lande og inden for ét land, USA. Den empiriske udfordring er, at forholdet potentielt er påvirket af endogenitet; korrupte magthavere kan forsøge at censurere internettet (som i Kina, for eksempel), eller udeladte faktorer kan have indflydelse på internettet og korruption på samme tid. Vi bruger lyn som et instrument for internet i en regression på korruption, som argumenteret for i kapitel 1. Vi finder, at lyn er et stærkt instrument for internet både på tværs af amerikanske stater og på tværs af lande. Udstyret med dette stærke instrument kan vi identificere den kausale effekt af internettet på korruption. Vi finder, at internettet reducerer omfanget af korruption på tværs af lande og på tværs af amerikanske stater.

Chapter 1

Lightning, IT Diffusion and Economic Growth across US States

Thomas Barnebeck Andersen, Jeanet Sinding Bentzen, Carl-Johan Dalgaard, and Pablo Selaya

Lightning, IT Diffusion and Economic Growth across US States*

November 1, 2010

*Thomas Barnebeck Andersen, Jeanet Bentzen, Carl-Johan Dalgaard and Pablo Selaya***

Abstract: Empirically, a higher frequency of lightning strikes is associated with slower growth in labor productivity across the 48 contiguous US states after 1990; before 1990 there is no correlation between growth and lightning. Other climate variables (e.g., temperature, rainfall and tornadoes) do not conform to this pattern. A viable explanation is that lightning influences IT diffusion. By causing voltage spikes and dips, a higher frequency of ground strikes leads to damaged digital equipment and thus higher IT user costs. Accordingly, the flash density (strikes per square km per year) should adversely affect the speed of IT diffusion. We find that lightning indeed seems to have slowed IT diffusion, conditional on standard controls. Hence, an increasing macroeconomic sensitivity to lightning may be due to the increasing importance of digital technologies for the growth process.

Keywords: Climate; IT diffusion; economic growth
JEL Classification: O33, O51, Q54

---

* We thank Daron Acemoglu, Roland Benabou, Michael Burda, Raquel Fernandez, Oded Galor, Norman Loayza, Heino Bohn Nielsen, Ariel Reshef, Jon Temple, Ragnar Torvik, David Weil, Joseph Zeira and seminar participants at University of Birmingham, Brown University, the 2009 NBER summer institute, the 2009 Nordic Conference in Development Economics, the 3rd Nordic Summer Symposium in Macroeconomics, NTNU Trondheim, University of Southern Denmark, and the 10th World Congress of the Econometric Society for comments and suggestions.

** Contact information: Andersen: Department of Business and Economics, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark. Email: barnebeck@sam.sdu.dk. Bentzen, Dalgaard, and Selaya: Department of Economics, University of Copenhagen, Øster Farimagsgade 5, building 26, DK-1353 Copenhagen, Denmark. Email: jeanet.bentzen@econ.ku.dk, carl.johan.dalgaard@econ.ku.dk, and pablo.selaya@econ.ku.dk.

1. Introduction

We are by all accounts living in a time of global climate change. This is a good reason to explore the economic consequences of climate related characteristics. In particular, how does the climate influence the growth process?

There seems to be compelling evidence to suggest that climate and geography profoundly affected the historical growth record (Diamond, 1997; Olsson and Hibbs, 2005; Putterman, 2008; Asraf and Galor, 2008). Today, climate shocks, like temperature changes, still affect growth in poor countries (Dell et al., 2008). But are climate and geography also important in highly developed economies, where high-tech industry and services are dominant activities?

Some research suggests that geography is still a force to be reckoned with, even in rich places. Access to waterways, for instance, appears to matter (Rappaport and Sachs, 2003). However, a geographic characteristic that exhibits a *time-invariant* impact on prosperity is difficult to disentangle from other slow moving growth determinants that may have evolved under the influence of climate or geography. In particular, climate and geography probably influenced the evolution of economic and political institutions.[1]

The present paper documents that a particular climate related characteristic – lightning activity – exhibits a *time-varying* impact on growth in the world's leading economy. Studying the growth process across the 48 contiguous US states from 1977 to 2007, we find no impact from lightning on growth prior to about 1990. However, during the post 1990 period there is a strong negative association: states where lightning occurs at higher frequencies have grown relatively more slowly. What can account for an increasing macroeconomic sensitivity to lightning?

In addressing this question one may begin by noting that the 1990s was a period of comparatively rapid US growth; it is the period where the productivity slowdown appears to

[1] An apparent impact from diseases on comparative development may be convoluting the impact from early property rights institutions in former colonies (Acemoglu et al., 2001); the impact of access to waterways, as detected in cross-country data, may also be related to the formation of institutions (Acemoglu et al., 2005).

finally have come to an end. Furthermore, the 1990s is the period during which IT appears to have diffused throughout the US economy at a particularly rapid pace. In fact, IT investment is often seen as a key explanation for the US growth revival (e.g., Jorgenson, 2001). On a state-by-state basis, however, the process of IT diffusion (measured by per capita computers and Internet users as well as manufacturing firms' IT investments) did not proceed at a uniform speed.

An important factor that impinges on IT investment and diffusion is the quality of the power supply. That a high quality power supply is paramount for the digital economy is by now widely recognized. As observed in *The Economist*:[2]

For the average computer or network, the only thing worse than the electricity going out completely is power going out for a second. Every year, millions of dollars are lost to seemingly insignificant power faults that cause assembly lines to freeze, computers to crash and networks to collapse. […] *For more than a century, the reliability of the electricity grid has rested at 99.9%* […] *But microprocessor-based controls and computer networks demand at least 99.9999% reliability* […] *amounting to only seconds of allowable outages a year.*

Indeed, a sufficiently large power spike lasting only one millisecond is enough to damage solid state electronics such as microprocessors in computers. Therefore, as a simple matter of physics, an irregularly fluctuating power supply reduces the longevity of IT equipment, and thus increases the user cost of IT capital.

A natural phenomenon that causes irregular voltage fluctuations is lightning activity. Albeit the impulse is of short duration, its size is impressive. Even in the presence of lightning arresters on the power line, peak voltage emanating from a lightning strike can go as high as 5600 V, which far exceeds the threshold for power disruptions beyond which connected IT equipment starts being damaged (e.g., Emanuel and McNeil, 1997). Moreover, the influence from lightning is quantitatively important. To this day, lightning activity causes around one third of the total number of annual power disruptions in the US (Chisholm and Cummings,

[2]"The power industry's quest for the high nines", *The Economist*, March 22, 2001.

2006). Theoretically, it is therefore very plausible that lightning may importantly have increased IT user costs.[3] Consequently, in places with higher IT user cost one would expect a slower speed of IT diffusion; lightning prone regions may be facing a climate related obstacle to rapid IT diffusion. It is worth observing that the problems associated with lightning activity, in the context of IT equipment, has not gone unnoticed by the private sector. As *The Wall Street Journal* reports:[4]

Even if electricity lines are shielded, lightning can cause power surges through unprotected phone, cable and Internet lines - or even through a building's walls. Such surges often show up as glitches. "Little things start not working; we see a lot of that down here," says Andrew Cohen, president of Vertical IT Solutions, a Tampa information-technology consulting firm. During the summer, Vertical gets as many as 10 calls a week from clients with what look to Mr. Cohen like lightning-related problems. Computer memory cards get corrupted, servers shut down or firewalls cut out.

Even though a link between lightning and IT diffusion is plausible, it does not follow that the link is economically important in the aggregate. Nor is it obvious that IT can account for the lightning-growth correlation.

We therefore also study the empirical link between lightning and the spread of IT across the US. IT is measured from both the household side (Internet and computer use) and the firm side (manufacturing firms' IT investment rates). We find that the diffusion of IT has progressed at a considerably slower pace in areas characterized by a high frequency of lightning strikes. This link is robust to the inclusion of a large set of additional controls for computer diffusion. Moreover, lightning ceases to be correlated with growth post 1990, once controls for IT are introduced. While the lightning-IT-growth hypothesis thus seems well founded, other explanations cannot be ruled out *a priori*.

[3] Naturally, the "power problem" may be (partly) addressed, but only at a cost. The acquisition of surge protectors, battery back-up emergency power supply (so-called uninterruptable power supply, UIP) and the adoption of a wireless Internet connection will also increase IT user costs through the price of investment. Hence, whether the equipment is left unprotected or not, more lightning prone areas should face higher IT user cost.
[4] "There Go the Servers: Lightning's New Perils". *Wall Street Journal*, August 25, 2009.

An alternative explanation is that the correlation between growth and lightning picks up growth effects from global warming. If global warming has caused lightning to increase over time, and simultaneously worked to reduce productivity growth, this could account for the (reduced form) correlation between lightning and growth. We document that this is unlikely to be the explanation for two reasons. First, we show that from 1906 onwards US aggregate lightning is stationary; on a state-by-state basis, we find the same for all save two states. There is thus little evidence to suggest that lightning density is influenced by a global warming induced trend. Second, we attempt to deal with the potential omitted variables problem by controlling directly for climate shocks which also could be induced by climate change. We examine an extensive list of climate variables, including rainfall, temperature and frequency of tornadoes. None of these variables impacts on the correlation between lightning and state-level growth rates. Nor does any other climate variable exhibit the kind of time-varying impact on growth that we uncover for lightning.

Another potential explanation is that the lightning-growth correlation is picking up "deep determinants" of prosperity that exhibit systematic variation across climate zones, just as lightning does. For instance, settler mortality rates, the extent of slavery and so forth. However, the correlation between lightning and growth is left unaffected by their inclusion in the growth regression.

In sum, we believe the most likely explanation for the lightning-growth correlation is to be found in the diffusion mechanism. The analysis therefore provides an example of how technological change makes economies increasingly sensitive to certain climate related circumstances. This finding is consistent with the "temperate drift hypothesis" (Acemoglu et al., 2002), which holds that certain climate related variables may influence growth in some states of technology, and not (or in the opposite direction) in others.

The paper is related to the literature that studies technology diffusion; particularly diffusion of computers and the Internet (e.g., Caselli and Coleman, 2001; Beaudry et al., 2006; Chinn and Fairlie, 2007). In line with previous studies, we confirm the importance of human capital for the speed of IT diffusion. However, the key novel finding is that climate related circumstances matter as well: lightning influences IT diffusion. In this sense the paper complements the thesis of Diamond (1997), who argues for an impact of climate on technology diffusion. Yet, whereas Diamond argues that climate is important in the context of agricultural technologies, the present paper makes plausible that climate also matters to technology diffusion in high-tech societies.

The analysis proceeds as follows. In the next section we document the lightning-growth link. Then, in Section 3, we discuss likely explanations (IT diffusion, other forms of climatic influence, institutions and integration) for the fact that lightning correlates with growth from about 1990 onwards. Section 4 concludes.

2. Lightning and US growth 1977-2007

This section falls in two subsections. In Section 2.1 we present the data on lightning and discuss its time series properties. In particular, we demonstrate that lightning is stationary and that, for panel data purposes, it is best thought of as a state fixed effect. Next, in Section 2.2, we study the partial correlation between lightning and growth across the US states.

2.1 The Lightning Data

The measure of lightning activity that we employ is the *flash density*, which captures the number of ground flashes per square km per year. We have obtained information about the flash density from two sources. The first source of information is reports from weather stations around the US. From this source we have yearly observations covering the period 1906-1995 and 40 US states. From about 1950 onwards we have data for 42 states. The second source of information derives from ground censors around the US. This data is *a priori* much more reliable than the data from weather stations.[5] In addition, it is available for all 48 contiguous states, but it only comes as an average for the period 1996-2005.[6]

In order to understand the data better, we begin by studying its time series properties. Figure 1 shows the time path for *aggregate* US lightning over the period 1906-95.

[5] Lightning events recorded at weather stations are based on audibility of thunder (i.e., these are basically recordings of thunder days), whereas ground sensors measure the electromagnetic pulse that emanates from lightning strikes (i.e., these are recordings of actual ground strikes). In the context of IT diffusion it is ground strikes that matter, and not the type of lightning occurring between clouds, say.

[6] Further details are given in the Data Appendix.

>Figure 1 about here<

The aggregate flash density is calculated as the state size weighted average over the 40 states with data for this extended period. Visual inspection suggests that there is no time trend. To test whether lightning contains a stochastic trend, we use an augmented Dickey-Fuller (DF) test with no deterministic trend. Lag length is selected by minimizing the Schwarz information criterion with a maximum of five lags. For aggregate US lightning the optimal lag length is one and the DF statistic equals -4.516. Hence the presence of a unit root is resoundingly rejected.

At the state level the presence of a unit root is also rejected at the 5% level in 38 of the 40 states, cf. Table 1. In light of the fact that DF tests have low power to reject the null of a unit root (even more so when, as here, we do not include a deterministic trend), we are in all likelihood safe to conclude that state-level lightning is also stationary.

>Table 1 about here<

These findings are of some independent interest in that they suggest that global warming has not interfered with the evolution of lightning trajectories in the US in recent times. In other words, there is little basis for believing that the flash density has exhibited a trend during the last century.

In the analysis below we focus on the period from 1977 onwards, dictated by the availability of data on gross state product. Consequently, it is worth examining the time series properties of the lightning variable during these last few decades of the 20th century.

During this period the flash density is for all practical purposes a fixed effect. In the Appendix, Table A.1, we show state-by-state that the residuals obtained from regressing lightning on a constant are serially uncorrelated. That is, deviations of the flash density from time averages are, from a statistical perspective, white noise. To show this formally, we use the Breusch-Godfrey test and a Runs test for serial correlation. By the standards of the Breusch-Godfrey test, we cannot reject the null hypothesis of no serial correlation in 38 states out of 42 states;

using the Runs test, we fail to reject the null in 40 states. Importantly, no state obtains a p-value below 0.05 in both tests. This suggests that for the 1977-95 period lightning is best described as a state fixed effect.

As remarked above, we have an alternative source of data available to us, which contains information for the 1996-2005 period. How much of a concurrence is there between data for the 1977-95 period and the data covering the end of the 1990s and early years of the 21st century? Figure 2 provides an answer. Eyeballing the figure reveals that the two measures are very similar. In fact, we cannot reject the null that the slope of the line is equal to one. This further corroborates that lightning is a state fixed effect.

>Figure 2 about here<

These findings have induced us to rely on the data deriving from ground censors in the analysis below. As noted above, this latter lightning data is of a higher quality compared to the measure based on weather stations and it covers more US states. Moreover, since deviations from the average flash density are white noise, we lose no substantive information by resorting to a time invariant measure. Still, it should be stressed that using instead the historical lightning measure based on weather stations (or combining the data) produces the same (qualitative) results as those reported below. These results are available upon request.

The cross-state distribution of the 1996-2005 data is shown in Figure 3, whereas summary statistics for 1996-2005 are provided in Table 2.

>Figure 3 about here<

>Table 2 about here<

There is considerable variation in the flash density across states. At the lower end we find states like Washington, Oregon and California with less than one strike per square km per year. It is interesting to note that the two states which are world famous for IT, Washington and California, are among the least lightning prone. At the other end of the spectrum we find Florida, Louisiana and Mississippi with seven strikes or more. It is clear that lightning varies systematically across climate zones. Hence, it is important to check, as we do below, (i) that lightning's correlation with growth is not due to other climate variables like high winds, rainfall and so on; and (ii) that spatial clustering effects are not deflating standard errors.

2.2 The Emergence of a Lightning-Growth Nexus

Figures 4 and 5 show the partial correlation between growth in labor productivity and the flash density, controlling only for initial labor productivity.

>Figures 4 and 5 about here<

We have data on gross state product (GSP) per worker for the period 1977-2007.[7] Hence, for this first exercise we have simply partitioned the data into two equal sized 15 year epochs. As seen from the two figures, there is a marked difference in the partial correlation depending on which sub-period we consider. During the 1977-92 period there is no association between growth and lightning; the (OLS) point estimate is essentially nil. However, in the second sub-period the coefficient for lightning rises twenty fold (in absolute value) and turns statistically significant; places with higher flash density have tended to grow at a slower rate during the 1990s and the first decade of the 21st century.

While this exercise is revealing, there is no particular reason to believe that the lightning-growth correlation emerged precisely in 1992. Hence, to examine the issue in more detail, we study the same partial correlation by running "rolling" regressions over 10 year epochs, starting with 1977-87.[8] That is, letting G_{it} denote the percentage average annual (continuously compounded) growth rate of GSP per worker over the relevant 10 year epoch,[9] we estimate an equation of the following kind:

$$G_{it} = b_0 + b_1 \log(y_{it-10}) + b_2 \log(lightning_i) + \varepsilon_i,$$

[7] State level data on personal income is also available, and for a longer period. But personal income does not directly speak to productivity. By contrast, GSP per worker is a direct measure of state level labor productivity. Moreover, the GSP per worker series is available in constant chained dollar values, which is an important advantage in the context of dynamic analysis. See the Data Appendix for a description of the GSP per worker series.

[8] The exact choice of time horizon does not matter much; below we run regressions with 5, 10, and 15 year epochs that complement the present exercise.

[9] That is, $G_{it} = 100*(1/T)*\log(y_{it}/y_{it-T})$, where T = 10.

and examine the evolution of b_2 as t increases. Figure 6 shows the time path for b_2 as well as the associated 95% confidence interval.

>Figure 6 about here<

In the beginning of the period there is not much of a link between lightning and growth; if anything the partial correlation is positive. As one moves closer to the 1990s the partial correlation starts to turn negative and grows in size (in absolute value). By 1995 the lightning-growth correlation is statistically significant at the 5% level of confidence. As one moves forward in time the partial correlation remains stable and significant. Hence, this exercise points to the same conclusion as that suggested by Figures 4 and 5: the negative partial correlation between lightning and growth emerged in the 1990s.

Albeit illustrative, both exercises conducted so far are *ad hoc* in the sense that they do not allow for a formal test of whether the impact from lightning is rising over time. Hence, as a final check, we run panel regressions with period length of 5, 10, and 15 years. The results are reported in Table 3 below.

>Table 3 about here <

Since lightning, for all practical purposes, is a fixed effect (cf. Section 2.1), Table 3 reports the results from running pooled OLS regressions. Specifically, we estimate the following growth regression:

$$G_{it} = b_0 + b_1 \log(y_{it-T}) + b_{2t} \log(\text{lightning}_i) + \mu_t + \varepsilon_{it},$$

where T=5, 10, 15 and b_{2t} accordingly is allowed to vary from period-to-period by way of interaction with time dummies. This way we can track the statistical and economic significance of lightning over time. Note also that we include time dummies independently of lightning, so as to capture a possible secular trend in growth over the period in question.

Turning to the results we find that the impact of lightning increases over time, and turns statistically significant during the 1990s.[10] The significance of lightning is particularly noteworthy as it is obtained for the relatively homogenous sample of US states. As is well known, the growth process for this sample is usually fairly well described by the initial level of income alone, suggesting only modest variation in structural characteristics that impinge upon long-run labor productivity (e.g., Barro and Sala-i-Martin, 1992). As a result, the scope for omitted variable bias contaminating the OLS estimate for lightning is *a priori* much more limited than, say, in a cross-country setting.

Still, a potential concern is that the lightning-growth correlation could be due to the omission of human capital. As is well known the return on skills appears to have risen during the 1990s, which could suggest an increasing effect from education on growth. If, in addition, the level of education is negatively correlated with lightning intensity (and it is) the lightning-growth link might disappear once schooling is introduced.

In Table 4 we therefore add measures of human capital to the growth regression. In order to do so rigorously we add information on primary, secondary and tertiary education simultaneously. As the lightning correlation does not depend appreciably on whether we invoke 5, 10 or 15 year epoch length we have chosen to focus on 10 year epochs. Results for 5 and 15 year epochs are similar, and available upon request.

>Table 4 about here <

Columns 2-5 of the table reveal that the human capital measures have no bearing on the lightning-growth correlation relative to the baseline growth regression in column 1; lightning is always significant irrespective of whether the three human capital proxies are added one-by-one (cf. columns 2-4) or included jointly (cf. column 5).

Another concern relates to regional effects. As is visually clear from Figure 3, lightning density is characterized by a certain degree of geographical clustering. Such cluster effects may

[10] The general time dummies (not reported) corroborate the prior of a revitalization of productivity growth during the 1990s.

impinge on the analysis in several ways.[11] Most importantly, one may worry that the lightning-growth correlation simply reflects that the Southeast, a high lightning area, is growing more slowly for reasons unrelated to lightning during this period. This suggests that we should add regional fixed effects to the growth regression.

In this endeavor we rely on the economic areas classification used by the Bureau of Economic Analysis (BEA), which distinguishes between eight regions.[12] This classification is however very taxing for our results in the sense that regressing the eight BEA areas on (log) lightning explains 84% of the cross state lightning variation (cf. Appendix, Table A.2, column 4).

In columns 6 of Table 4 we add the eight regional fixed effects. The inclusion of the BEA regions does not impinge on the *size* of the partial correlation between lightning and growth, but it impacts on the *precision* of the OLS estimate in a major way, by doubling the standard error. This is no surprise in light of the strong degree of multicollinearity between the regional effects and lightning intensity. This interpretation is further supported by the fact that while neither lightning nor the set of fixed effects are significant separately, they are jointly significant. In order to examine whether regional effects are at the root of the lightning-growth correlation, we therefore also ran regressions where we add each of the regional fixed effects one-by-one to the specification in column 5 of Table 4. The results are found in the Appendix (Table A.3). The key result is that no single BEA region can render lightning imprecise enough to be rejected as statistically insignificant.

In sum, the time varying effect of lightning on growth is not produced by the growth performance of any particular region, is robust to the inclusion of human capital and time dummies. The specification in column 5 of Table 4 will serve as our baseline specification when we examine the robustness of the lightning-growth link in much greater detail.

Before addressing robustness in depth, however, it is worth commenting on the economic significance of lightning. Taken at face value, the point estimate for the 1990s imply that a one

---

[11] See Cameron and Trivedi (2005) or Angrist and Pischke (2009) for general discussions of clustering.

[12] The eight BEA regions are Far West, Great Lakes, Mideast, New England, Plains, Rocky Mountain, Southeast, and Southwest.

standard deviation increase in lightning intensity (about 2.4 flashes per year per sq km) induces a reduction in growth by about 0.2 percentage points ($\approx 0.2 \cdot \log(2.4)$), conditional on the level of initial labor productivity, human capital and the time effects. This is about 12.5 % of the gap between the 5th percentile and the 95th percentile in the distribution of GSP per worker growth rates for the period 1977-2007 (for the 48 states in our sample). By extension, variation in lightning by four standard deviations (roughly equivalent to moving from the 5th percentile to the 95th percentile in the lightning distribution across US states) can account for about 50% of the "95/5" growth gap.[13] Needless to say, this is a substantial effect.

3. Robustness of the Lightning-growth nexus

3. 1 Climate Shocks

At first glance, a reasonable objection to the lightning-growth correlation is that it is somehow spurious: perhaps other climate related variables exert an impact on growth and, at the same time, happen to be correlated with the flash density?

To be sure, lightning correlates with various kinds of weather phenomena that arise in the context of thunderstorms. Aside from lightning, thunderstorms produce four weather phenomena: tornadoes, high winds, heavy rainfall, and hailstorms. It seems plausible that these climate variables can induce changes in the growth rate in individual states in their own right. Each of them destroy property (physical capital), people (human capital), or both (Kunkel et al., 1999). By directly affecting the capital-labor ratio, the consequence of, say, a tornado could be changes in growth attributable to transitional dynamics. The nature of the transitional dynamics (i.e., whether growth rises or falls) is unclear as it may depend on whether the tornado destroys more physical or human capital (e.g., Barro and Sala-i-Martin, 1995, Ch.5).[14] Nevertheless, since the lightning-growth correlation pertains to a relatively short time span (so far), it is hard to rule out that the above reasoning could account for it.

[13] Log normality of lightning is not accurate; but on the other hand not terribly misleading either. It does exaggerate the actual variation in lightning slightly; the observed variation is about 7 flashes, compared to the "back-of-the-envelope" calculation implying roughly 9.

[14] In a US context one may suspect a relatively larger impact on physical capital compared to human capital; if so climate shocks would tend to instigate a growth acceleration in their aftermath, as a higher marginal product of capital induces firms to invest in physical capital.

In addition, lightning correlates with temperature: hotter environments usually feature a higher flash density. Temperature has been documented to correlate with economic activity within countries (e.g., Nordhaus, 2006; Dell et al., 2009); therefore, we cannot rule out *a priori* that the link between lightning and growth is attributable to the intervening influence of temperature.[15]

Hence, in an effort to examine whether climate shocks could account for the lightning-growth correlation, we gathered data on all of the above weather phenomena: temperature, precipitation, tornadoes, hail size and wind speed. In addition, we obtained data on topography (i.e., elevation) and latitude. The latter is a useful catch-all measure of climate. For good measure, we also obtained data on sunshine, humidity, and cloud cover (albeit it is not entirely clear why these weather phenomena should matter to growth). In total, we have data on ten alternative climate/geography variables; the details on the data are found in the Data Appendix.

With these data in hand, we ask two questions. First, ignoring lightning, do any of these weather phenomena exhibit a correlation with growth which is similar to that of lightning? That is, do any of them appear to become more strongly correlated with growth during the period 1977-2007? Second, taking lightning into account, do any of the above mentioned variables render lightning insignificant?

Tables 5 and 6 report the answers. Columns 2-11 of Table 5 examine the potentially time varying impact from each weather variable; column 1 reproduces the lightning regularity from Section 2.1. It is plain to see that none of the weather variables exhibit a similar growth correlation as that involving lightning. The only variable that influences growth in a statistically significant way in the final period is hail size; however, unlike lightning, hail size also had a statistically significant growth impact in the first period.

[15] Nordhaus (2006) and Dell et al. (2009) document a correlation between temperature and income *levels,* not growth*.* In fact, Dell et al. (2008) find that temperature is *not* correlated with growth in rich places, using cross-country data. Nevertheless, the link seems worth exploring.

>Tables 5 and 6 about here<

In Columns 2-11 of Table 6 we simultaneously include lightning and the various alternative climate/geography controls. In all cases, lightning remains significantly correlated with growth. In fact, when comparing the point estimate for lightning with or without (column 1) additional controls, it emerges that the point estimate is virtually unaffected.

In sum, these results suggest that the lightning-growth correlation is unlikely to be attributable to other weather phenomena.

3.2 Institutions and Integration

An extensive literature examines the impact from historical factors on long-run development. For instance, variation in colonial strategies seems to have an important impact on institutional developments around the world, thus affecting comparative economic development (e.g., Acemoglu et al., 2001). Similarly, initial relative factor endowments, determined in large part by climate and soil quality, may well have affected long-run development through inequality and human capital promoting institutions (Engerman and Sokoloff, 2002; Galor et al., 2008). Thus, in many instances the initial conditions that may have affected long-run developments are related to climate or geography. In the present context, therefore, it seems possible that the lightning-growth correlation may be picking up the influence from such long-run historical determinants of prosperity. Naturally, the conventional understanding would be that "deep determinants of productivity", e.g. determinants of political and economic institutions, should have a fairly time invariant impact on growth. As a result, it would not be surprising if such determinants do not exert a time varying impact on growth. But whether it is the case or not is obviously an empirical matter.

To examine whether the lightning-growth nexus is attributable to such effects, we obtained data on ten potential determinants of long-run performance for the US. The source of the data is Mitchener and McLean (2003), who examine the determinants of long-run productivity levels across US states. In addition, we collected state-level data on three dimensions of global integration, related to international movements of goods and capital. This leaves us with 13 different potential determinants of labor productivity growth, broadly capturing "institutions, geography and integration" (Rodrik et al., 2004).[16]

As in Section 3.1 we ask whether these determinants, individually, exhibit a time varying impact on growth, and whether their inclusion in the growth regression renders lightning insignificant.

>Table 7 and 8 about here<

In Table 7 we examine the impact from various historical determinants of productivity one-by-one. Of particular note is column 4, which involves the percentage of the population in slavery in 1860. This is the only variable which behaves much like lightning, with a partial correlation that seems stronger at the end of the 1977-2007 period as compared to the beginning of the period.

Table 8 includes both lightning and the individual controls. Since the population in slavery is the only variable we have found so far that exhibits a correlation with growth that is qualitatively similar to that of lightning, the results reported in column 4 is of central importance. When both variables enter the growth regression only lightning retains explanatory power. The point estimate for the last period is more or less unaffected, while the statistical significance of lightning is reduced a bit. But population in slavery does not statistically dominate lightning in the specification. More broadly, it is once again worth observing how stable the partial correlation between lightning and growth seems to be. Comparing the results reported in column 1 (no historical controls) for lightning to those reported in columns 2-11 it is clear that the coefficient for lightning is quite robust.

Finally, Table 9 examines the potential influence from integration. As seen by inspection of columns 4 and 5, integration proxies cannot account for the lightning-growth correlation either.

[16] See the Data Appendix for details.

>Table 9 about here<

The results of this and the previous subsection uniformly support the same qualitative conclusion: a macro economic sensitivity to lightning has emerged over time in the US. The question is why?

4. An explanation for the Lightning-Growth nexus: IT diffusion

We begin this section by examining the theoretical foundation behind the claim that lightning (or, more appropriately, the flash density) should have an impact on growth via IT diffusion. Subsequently we examine the hypothesis empirically.

4.1. Theory: why lightning matters to IT diffusion.

The simplest way to think about IT diffusion is via basic neoclassical investment theory. That is, IT diffusion occurs in the context of IT capital investments; higher investments are tantamount to faster IT diffusion.

According to neoclassical investment theory, the central determinant of the desired capital stock, and thus investments for the initial stock given, is the user cost of capital (Hall and Jorgenson, 1967). Two elements of (IT) user cost are plausibly influenced by lightning: the total price of IT investment goods and the physical rate of IT capital depreciation.

IT capital depreciation is influenced by lightning activity for the following physical reason. Solid-state electronics, such as computer chips, are constructed to deal with commercial power supply in the form of alternating current. The voltage of the current follows a sine wave with a specific frequency and amplitude. If the sine wave changes frequency or amplitude, this constitutes a power disruption. Digital devices convert alternating current to direct current with a much reduced voltage; digital processing of information basically works by having transistors turn this voltage on and off at several gigahertz (Kressel, 2007). If the power supply is disrupted, the conversion process may become corrupted, which in turn causes damage to the equipment, effectively reducing its longevity. It is important to appreciate that even extremely short lasting power disruptions are potentially problematic.

Voltage disturbances measuring less than one cycle (i.e., 1/60th of a second in the US case) are sufficient to crash and/or destroy servers, computers, and other microprocessor-based devices (Yeager and Stalhkopf, 2000; Electricity Power Research Institute, 2003). A natural phenomenon which damages digital equipment, by producing power disruptions, is lightning activity (e.g., Emanuel and McNeil, 1997; Shim et al., 2000, Ch. 2; Chisholm, 2000).[17] This avenue of influence is *a priori* highly plausible. In the US lightning produces a large fraction of the total number of power disruptions (Chisholm and Cummings, 2006); firms specializing in delivering power protection are another testimony to the same thing.

The latter point immediately raises the issue that firms can take pre-emptive actions so as to reduce the impact of lightning on the cost of capital. This can be done by investing in surge protectors, say. However, the crux of the matter is that this imposes an additional cost to be carried in the context of IT investments; it amounts to an increasing IT investment price. Hence, even if we take the likely pre-emptive measures into account, more lightning prone areas will face higher IT user costs.

In sum: in areas with a greater flash density, the speed of IT diffusion, as measured by IT capital accumulation, will proceed at a slower pace. The reason is that a higher lightning density increases the frequency of power disturbances, IT capital depreciation (or the price of IT investments), the user cost of IT capital, and thus lowers IT investments. Moreover, if output is increasing in the IT capital stock, growth in output will similarly tend to be slower in areas with greater lightning activity, conditional on the initial level of output.

[17] Note that lightning may enter a firm or household in four principal ways. First, lightning can strike the network of power, phone, and cable television wiring. This network, particularly when elevated, acts as an effective collector of lightning surges. The wiring conducts the surges directly into the residence, and then to the connected equipment. In fact, the initial lightning impulse is so strong that equipment connected to cables up to 2 km away from the site of the strike can be damaged (BSI, 2004). Technically speaking, this is the mechanism we are capturing in the simple model above. Second, when lightning strikes directly to or nearby air conditioners, satellite dishes, exterior lights, etc., the wiring of these devices can carry surges into the residence. Third, lightning may strike nearby objects such as trees, flagpoles, road signs, etc., which are not directly connected to the residence. When this happens, the lightning strike radiates a strong electromagnetic field, which can be picked up by the wiring in the building, producing large voltages that can damage equipment. Finally, lightning can strike directly into the structure of the building. This latter type of strike is extremely rare, even in areas with a high lightning density.

While the above theoretical considerations speak to a direct impact of lightning on IT investment, there could be an important complementary mechanism at work. The *choice of firm location* may depend on the quality of power supply, and thus lightning. Specifically, it may be the case that IT intensive firms choose to locate in areas where lightning intensity is modest, due to the resulting (slightly) higher power quality. Interestingly, the National Energy Technology Laboratory, operated by the US Department of Energy, reports that a recent firm level survey had 34% respondents saying that they would shift business operations out of their state if they experienced ten or more unanticipated power disturbances over a quarter of a year.[18] Hence, it seems plausible that this mechanism also could affect comparative IT penetration across US States.

To this one may add that in areas with frequent power disruptions and outages, the marginal *benefit* of owning a computer is probably lowered as well. Obviously, if consumers and firms face regular power outages it will be difficult to employ IT efficiently. But even if power disruptions are infrequent and of very short duration, power disruptions lead to glitches and downtime, which serves to lower the productivity of IT equipment. Hence, aside from increasing the marginal costs of IT capital, lightning may also work to lower IT productivity.

Schematically we may summarize the theoretical considerations above in the following way:

$$\text{Lightning density} \rightarrow \text{Power disturbances} \rightarrow \text{IT investments} \rightarrow \text{Growth},$$

where the second from last arrow subsumes the likely impact from (lightning induced) power disturbances on IT costs and benefits.

The mechanisms linking lightning to growth are likely to have become increasingly important over time for a number of reasons. First, IT capital investments accounted for a substantial part of output growth, starting in the 1990s (e.g., Jorgenson, 2001). Consequently, factors that impact on IT capital accumulation (e.g., the flash density) should also become more important to growth. Second, the 1990s was the era during which the Internet emerged (in the sense of

[18]The report is available at: http://www.netl.doe.gov/moderngrid/

the World Wide Web); a conceivable reason why firms chose to intensify IT investments during the same period.[19] From a physical perspective, however, the network connection is another way in which lightning strikes may reach the computer, in the absence of wireless networks (which have not been widespread until very recently). Third, the 1990s saw rapid increases in the computing power of IT equipment. In keeping with Moore's law, processing speed doubled roughly every other year. This is an important propagation mechanism of the lightning-IT investment link. The reason is that the sensitivity of computers to small power distortions *increases* with the miniaturization of transistors, which is the key to increasing speed in microprocessors (Kressel, 2007).[20] As a result, these factors would all contribute to increasing the importance of the flash density to IT investments, and thus to growth, during the 1990s. Whether this theory is relevant, however, is an empirical issue to which we now turn.

4.2. Empirical analysis: Lightning, IT diffusion and Economic Growth.

In order for the above theory to be able to account for the lightning-growth correlation, two things need be true. First, it must be the case that lightning is a strong predictor of IT across the US states. Second, there should be no explanatory power left in lightning vis-à-vis growth once we control for IT. We examine these two requirements in turn.

In measuring the diffusion of IT capital across the US we employ three different measures. Two measures derive from a supplement to the 2003 Current Population Survey, which contained questions about computer and Internet use; the third measure derives from the 2007 Economic Census (see Data Appendix for further detail). The first measure is percentage of households with access to the Internet; the second measure is percentage of households with a PC; and the third measure is manufacturing firms' capital expenditures on computers

[19] The WWW was launched in 1991 by CERN (the European Organisation for Nuclear Research). See Hobbes' Internet Timeline v8.2 http://www.zakon.org/robert/internet/timeline/ .

[20] This is well known in the business world: "*The spread of technology has spawned a need for lightning-security specialists." The computer chip, the smaller it's gotten, the more susceptible it is," says Mark Harger, owner of Harger Lightning and Grounding in Grayslake, Ill. "It's been a boon to our business". His company manufacturers systems that shield buildings from direct strikes and power surges from nearby lightning. With a steady stream of orders from financial and technology companies looking to protect their data centers, the company has gone from eight employees to 100 over the past 20 years*. ""There Go the Servers: Lightning's New Perils", *The Wall Street Journal*, August 25, 2009.

and related equipment as a percentage of total capital expenditures on machinery and equipment.[21] A few comments on the IT data are in order.

First, our IT measures allow us to explore IT penetration in the US economy from two different perspectives: the firm and the household side, respectively. Whereas the household data speaks exclusively to the level of IT investments, the firm data arguably speaks both to IT investments and location choice. In the end there are two reasons why the fraction of IT expenditures to total capital expenditure might be higher in some states compared to others. On the one hand there is the *investment effect*, which captures that structurally similar manufacturing firms have different levels of IT investments, depending on whether they locate in high versus low lightning density areas; this sort of information is also likely captured by our household data. However, on the other hand, there is a potential *composition effect*, which captures that areas with less lightning may attract more IT intensive firms, which drives up the IT expenditure/Total capital expenditure ratio. Both effects, which we admittedly cannot disentangle, would predict a negative relationship between lightning density and manufacturing IT investment intensity.

Second, one may worry about vintage capital effects. In a vintage growth setting a higher (lightning induced) rate of capital depreciation will in principle have two opposite effects on the IT capital stock. One the one hand, we expect lower overall investments. On the other hand, faster depreciation implies that more recent (more productive) vintages take up a larger share of the stock. As a result, one may worry about the net impact of lightning on IT capital and long-run productivity. Unfortunately we do not have access to information about IT quality, which would be ideal. Still, on *a priori* grounds, a higher rate of capital depreciation unambiguously lowers IT capital *intensity* in the standard neoclassical vintage growth model (Phelps, 1962). Hence, even allowing for vintage effects, higher depreciation should lower IT

---

[21] We did consider inferring IT capital intensity at the state level since the Bureau of Economic Analysis produces sector specific data on IT capital stocks. To exploit these data we would have to assume that the marginal product of IT capital is equalized *within* sectors, *across* states. Weighting the sector specific IT capital intensities by state specific sector composition would yield a guesstimate for state IT capital intensity. However, since (state specific) lightning affects the user cost of capital via the price of acquisition and/or the rate of capital depreciation, the assumption of within industry equalization of marginal products is implausible on *a priori* grounds. To put it differently, the main avenue through which lightning should affect IT capital intensity would be eliminated *by construction* had we used this procedure to generate state level IT capital. As a result, we have not pursued the matter further.

intensity and thereby long-run productivity. Moreover, if the IT variable is measured with gross error, it would tend to make it less likely that it appears as a significant growth determinant in the regressions to follow at the end of this section; i.e., it would make it less likely that IT (as measured here) can account for the lightning-growth correlation.[22]

Third, with only one observation for the IT variables, we have to settle for cross section regressions.

Finally, one may question whether there is value in using both household IT variables, since having access to a computer is a prerequisite for the use of the Internet. Yet, the emergence of the WWW is a much more recent technology than the PC, as the former derives from 1991. The personal computer started spreading earlier. Hence, the initial conditions that may matter to the speed of adoption are discernible by time. For instance, whereas educational attainment in the 1970s should influence the spread of the personal computer, the Internet is affected by education levels in the 1990s. Consequently, the two empirical models of IT diffusion will have to differ in terms of the dating of the right hand side IT diffusion determinants. As a result, we employ both.

A natural point of departure is the simple correlation between the flash density and the three IT measures for the 48 states in our sample. Figures 7 to 9 depict them.

>Figures 7 to 9 about here<

Visually, the strong negative correlations between the flash density and household and firm IT use, respectively, are unmistakable. By the middle of the first decade of the 21th century, states that experienced lightning strikes at a higher frequency also had relatively fewer users of computers and the Internet as well as lower IT investment intensity in manufacturing.

[22] If IT is poorly measured this would also make it less likely that we can establish a link between lightning and IT. Measurement error (in this case) is found in the dependent variable, for which reason it will (under standard assumptions) inflate the standard errors of the estimated parameters. It thus becomes less likely to observe a statistically significant correlation with lightning activity.

A more systematic approach involves more controls. Human capital is probably the first additional determinant of diffusion that comes to mind. The idea that a more educated labor force is able to adopt new technologies more rapidly is an old one, going back at least to the work of Nelson and Phelps (1966). Another natural control is the level of GSP per worker. Aside from being a catch-all control for factors that facilitate diffusion, it can also be motivated as a measure of the "distance to the frontier". The sign of the coefficient assigned to GSP per worker is therefore ambiguous. A positive sign is expected if initially richer areas are able to acquire IT equipment more readily. A negative sign could arise if richer areas, by closer proximity to the technology frontier, are less able to capitalize on "advantages of backwardness".

In addition to labor productivity and human capital, we chiefly follow Caselli and Coleman (2001) in choosing relevant additional determinants of IT diffusion (they also include human capital and income per capita). First, we use measures for the composition of production; it seems plausible that IT may spread more rapidly in areas featuring manufacturing rather than agriculture. Second, we employ proxies for global links, measured by international movements of goods and capital, and a measure of local market size: state population. Third, we employ various historical variables as controls. Caselli and Coleman, studying cross-country data, include a measure of economic institutions, which we are not able to do directly in our US sample. However, by including various plausible historical determinants of productivity (e.g., soldier mortality, the pervasiveness of slavery in the late 19th century, etc.) we hope to pick up much the same type of information. Of course, in US cross-state data one expects differences in institutional quality to be a great deal smaller than what is typically found in cross-country data. Finally, moving beyond the "Caselli-Coleman controls", we examine the impact from the age structure of the population, religiousness, ethnic composition and urbanization on IT diffusion. [23]

In Table 10 we report baseline results for all three IT measures. In columns 1, 5 and 9 of the table we examine the simple correlations between the flash density and computer use, Internet use and manufacturing firms' IT investments, respectively. The lightning variable is

[23] Details on all the data mentioned above are given in the Data Appendix.

always highly significant and it accounts for about 24% to 43% of the variation in the IT variables. In the remaining columns we add human capital controls and regional fixed effects progressively. Lightning is always highly significant, even with the inclusion of eight regional fixed effects. The only other variable that is consistently significant is the fraction of state population with a bachelor or higher; this is consistent with previous findings (e.g., Caselli and Coleman, 2001; Beaudry et al., 2006). It is also worth noting that we are able to span more than 80% of the variation in IT on the household side (columns 4 and 8) and more than 60% on the firm side (column 12).

Using the estimate from column 7 in Table 10 we find that a one standard deviation increase in lightning leads to a reduction in Internet users by about 1 percentage point.[24] In 2003 the states with the lowest Internet penetration (the 5th percentile) had about 44% of the population being able to access the Internet; at the other end of the spectrum (the 95th percentile) about 60% of the population was online. Hence the estimate for lightning implies that one standard deviation change in lightning can account for about 6% of the 95/5 gap; four standard deviations therefore motivates about 25% of the difference.

In an effort to check for robustness, Table 11 introduces additional controls to the long regressions in Table 10 (i.e., columns 4, 8 and 12), one by one. Nowhere is the influence from the flash density eliminated. Rather, the point estimate appears reasonably robust to the inclusion of alternative IT diffusion controls, economically as well as statistically.

>Table 11 about here<

The lightning-IT correlation can obviously not be ascribed to reverse causality. Moreover, since the remaining diffusion determinants are lagged, the risk that endogeneity of these variables is contaminating the OLS estimate for lightning is diminished. To be sure, it is impossible to completely rule out that the partial correlation between lightning and IT could be attributed to one or more omitted variables in the analysis above. Still, a causal interpretation is well founded on theoretical grounds: the empirical link between IT and

[24] Recall, the standard deviation of the flash density variable is 2.4 in our 48 state sample.

lightning is clearly robust to a reasonable set of alternative IT determinants, and it is robust to regional fixed effects. Moreover, the point estimate seems stable across specifications. These characteristics provide a sound basis for believing the estimates above can be taken to imply that lightning is causally impacting on the speed of IT diffusion. We can, however, push the matter further on two accounts. First, we can ask whether IT can account for the link between growth and lightning. This is basically an indirect check of the exclusion restrict in an IV setup, where lightning serves as instrument for IT. Second, we can simply perform such an IV exercise.

Table 12 shows the relevant regression results; in columns 1-13 we address the first issue, while in columns 14 and 15 we address the second issue. Our focus is specifically on the 1991-2007 period, as this is the period during which lightning is significantly correlated with growth.

>Table 12 about here<

In column 1 of Table 12 the lightning-growth correlation is reproduced. In the following three columns we add the IT measures. Individually, all three are significantly and positively correlated with growth, as expected. The interpretation of the household IT variables is slightly different though. As noted above, the Internet originated in 1991. As a result, the independent variable can be seen as a proxy for Internet investments over the period; in 1991 the number of Internet users inevitably was close to zero, so the 2003 value effectively captures *changes* in Internet users over the relevant period. Needless to say the same is not true for computers, which started diffusing far earlier. If the IT investment rate is the relevant control, the computer variable is therefore measured with error. This may account for the fact that the economic size of the impact of the Internet variable is larger than that of computers in Table 12.

A key result of the exercise is reported in columns 5-7. When the IT variables are added to the equation, the flash density looses significance. The loss of significance is mainly attributable to a much lower point estimate; in column 7 it is reduced by almost a factor 7. A reasonable interpretation is that lightning appears in the growth regression due to its impact on IT diffusion. In columns 8-10 we include all four variables at once; in column 9 all human capital controls are also included, whereas in column 9 regional fixed effects are added to the list. Despite the obvious multicollinearity in this experiment, manufacturing firms' IT investment share remains strongly significant. This means that this latter variable dominates household's computer and Internet use as a predictor of US cross state real GSP growth rates in the Internet era: 1991 onwards. This continues to be the case when we exclude lightning, as done in columns 11-13.

In columns 14 and 15, we turn to an IV exercise using lightning as an instrument for manufacturing firms' IT investment share. In light of column 7, we have good reason to be optimistic that lightning satisfies the required exclusion restriction. In the IV regressions, we always include the human capital controls; in column 14 we include in addition the eight regional fixed effects.

Turning to the results, we first note that lightning is significant in the first stage in both columns. Moreover, the 2SLS point estimate is very similar to what is found using OLS. As expected, lightning is only a moderately strong instrument when the eight regions are included, as in column 14. However, the weak-instrument robust Stock-Wright (2000) S statistic, which tests the null that the coefficient of the endogenous regressor in the second stage is equal to zero, deems the IT variable significant. Moreover, since the regional fixed effects jointly are not even marginally significant, they can safely be excluded, in which case the lightning instrument becomes strong (first stage F > 10). Figure 10 provides a visual representation of the IV results.

>Figure 10 about here<

What is the economic significance of IT diffusion on growth? One approach would consist of studying the impact effect on growth from an increase in the intensity of IT investments. If we use the estimate from Table 12 (column 15) we find that a one percentage point increase in IT investment intensity increases growth – on impact – by about 0.15 percentage points. Of course, the initial growth impact should then drop off as the economy converges towards steady state.

Another approach to is to study the impact from greater IT investment intensity on the long run level of GDP per worker, rather than IT investments impact on transitional growth. Taking the IV estimate at face value (again Table 12, column 15), we find that an increase in IT expenditures as a fraction of total expenditures by 1 percent increases long-run labor productivity by about 0.6 percent.[25] Hence, our estimates suggest that IT indeed exerted a positive influence on growth, consistent with previous micro (firm) level estimates (e.g., Brynjolfsson and Hitt, 2003).

Overall, we believe the above analysis builds a fairly strong case in favor of the IT diffusion hypothesis; i.e., the thesis that lightning appears as a growth determinant in the 1990s due to the growing influence of digital technologies on economic growth.

4. Concluding Remarks

In theory, lightning should impact on IT diffusion. Higher lightning intensity leads to more frequent power disruptions, which in turn reduces the longevity of IT equipment. As a result, by inducing higher IT user cost, a higher lightning frequency should hamper IT investments. By implications, high-tech societies may actually be quite vulnerable to climate shocks. Consistent with the temperate drift hypothesis, technological change may therefore render societies more sensitive to climate phenomena that previously were only of second order importance.

Empirically, we document that lightning activity is negatively correlated with measures of IT diffusion; computers and Internet hook-ups per household and IT investment rates by manufacturing firms. Conditional on standard controls, states with less lightning have adopted IT more rapidly than states where lightning activity is more intensive.

Consistent with a detrimental impact on IT diffusion, we find that states with more lightning have grown slower from about 1990 onwards. This pattern cannot be accounted for by other

[25] The implicit calculation proceeds as follows. Consider steady state where growth in GDP per worker is zero (or constant). Then we can work out the semi-elasticity of steady state GDP per worker with respect to IT investments as 0.15/1.25 = 0.12 (Table 12, column 15). Evaluated at the mean investment level (see Table 2), we then find the elasticity of roughly 0.6 (=0.12*5.4).

climate phenomena, nor can it be explained by a time varying influence from deep historical determinants of productivity.

REFERENCES

Acemoglu, D., S. Johnson, and J. Robinson, 2001. The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review*, 91, 1369-1401.

Acemoglu, D., S. Johnson, and J. Robinson, 2002. Reversal of Fortune: Geography And Institutions In The Making Of The Modern World Income Distribution. *Quarterly Journal of Economics*, 117, 1231-1294.

Acemoglu D., S. Johnson, and J. Robinson, 2005. The Rise of Europe: Atlantic Trade, Institutional Change, and Economic Growth, *American Economic Review*, 95, 546-579.

Angrist, J. and J. Pischke, 2009. Mostly Harmless Econometrics. *Princeton University Press*.

Asraf Q. and O. Galor, 2008. Dynamics and Stagnation in the Malthusian Epoch: Theory and Evidence". Working Paper (Brown University).

Barro R.J. and X. Sala-i-Martin, 1992. Convergence. *Journal of Political Economy*, 100, 223-51.

Barro R.J. and X. Sala-i-Martin, 1995. Economic Growth. McGraw Hill.

Beaudry, P., M. Doms, and E.G. Lewis, 2006. Endogenous Skill Bias in Technology Adoption: City-Level Evidence from the IT Revolution. FRB of San Francisco Working Paper No. 2006-24.

Brynjolfsson, E. and Lorin M. Hitt, 2003. Computing productivity: Firm level evidence. *Review of Economics and Statistics*, 85, 793-808.

BSI, 2004. Bundesamt für Sicherheit in der Informationstechnik, IT-Grundschutz Manual. Available online at http://www.bsi.de/english/gshb/index.htm.

Caselli, F. and W.J. Coleman II, 2001. Cross-Country Technology Diffusion: The Case of Computers. *American Economic Review*, 91, 328-35.

Cameron, A.C. and P.K Trivedi, 2005. Microeconometrics. *Cambridge University Press*.

Changnon, S.A., 2001. Development and analysis of data bases for assessing long-term fluctuations in thunderstorms in the United States, Final report, Climate Change Detection and Attribution Program, National Oceanic and Atmospheric Administration, US Department of Commerce.

Chinn, M.D. and R.W. Fairlie, 2007. The determinants of the global digital divide: a cross-country analysis of computer and Internet penetration. *Oxford Economic Papers*, 59, 16-44.

Chisholm, W., 2000. Lightning Protection. Chapter 4.10 in: *The Electrical Power Engineering Handbook*. Grigsby, L. (eds.), IEEE Press.

Chisholm, W., and K. Cummins, 2006. On the Use of LIS/OTD Flash Density in Electric Utility Reliability Analysis. *Proceedings of the LIS International Workshop, MSFC, Huntsville, AL (2006) Sept*.

Diamond, J., 1997. *Guns, Germs and Steel: The Fates of Human Societies*. New York: W. W. Norton.

Dell, M., B. Jones, and B. Olken, 2008. Climate Change and Economic Growth: Evidence from the Last Half Century. NBER Working Paper 14132.

Dell, M., B. Jones, and B. Olken, 2009. Temperature and Income: Reconciling New Cross-Sectional and Panel Estimates. American Economic Review Papers & Proceedings (forthcoming).

Emanuel, A.E. and J. A. McNeill, 1997. Electric Power Quality. *Annual Review of Energy and the Environment*, 22, 263–303

Engerman, S.L. and K.L. Sokoloff, 2002. Factor Endowment, Inequality, and Paths of Development among New World Economies. NBER Working paper 9259.

Galor, O., O. Moav and D. Vollrath, 2009. Inequality in Land Ownership, the Emergence of Human Capital Promoting Institutions, and the Great Divergence. *Review of Economic Studies*, 76, 143-179.

Hall, R.E. and D.W. Jorgenson, 1967. Tax Policy and Investment Behavior. *American Economic Review*, 57, 391-414.

Jorgenson, D., 2001. Information Technology and the US Economy. *American Economic Review*, 91, 1-32

Kunkel, K.E.K, R.A Pielke Jr. And S.A. Changnon, 1999. Temporal Fluctuations in Weather and Climate Extreemes That Causes Economic and Human Health Impacts: A Review. *Bulletin of the American Meteorological Society*, 80, 1077-98.

Kressel, 2007. *Competing for the Future: How Digital Innovations are Changing the World*. Cambridge University Press.

Mitchener, K.J and I.W. McLean, 2003. The Productivity of US States since 1880. *Journal of Economic Growth*, 8, 73-114.

National Energy Technology Laboratory, 2003. The Value of Electricity When It's Not Available. http://www.netl.doe.gov/moderngrid/docs/The_Value_of_Electricity_When_It's_Not_Available.pdf)

Nelson, R.R. and E. Phelps, 1966. Investment in Humans, Technological Diffusion, and Economic Growth. *American Economic Review*, 56, 69-75.

Nordhaus, W., 2006. Geography and Macroeconomics. *Proceedings of the National Academy of Sciences (US)*, 103, 3510-3517.

Olsson, O. and D. Hibbs, 2005. Biogeography and long-run economic development, *European Economic Review*, 49, 909-938.

Phelps, E., 1965. The new view of investment: a neoclassical analysis. *Quarterly Journal of Economics*, 76, p. 548-67

Putterman, L., 2008. Agriculture, Diffusion, and Development: Ripple Effects of the Neolithic Revolution. *Economica,* 75, 729-748.

Rappaport, J. and J. Sachs, 2003. The US as a Coastal Nation. *Journal of Economic Growth*, 8, 5-46.

Rodrik, D., A. Subramanian, and F. Trebbi, 2004. Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development. *Journal of Economic Growth*, 9, 131-65.

Shim, J.K., A. Qureshi, and J.G. Sieqel, 2000. *The International Handbook of Computer Security*. Glenlake Publishing Company Ltd.

Yeager, K. and K. Stahlkopf, 2000. Power for a Digital Society. Prepared for the Department of Energy CF-170/1-1-DOE . Available online at:
http://www.rand.org/scitech/stpi/Evision/Supplement/yeager.pdf

Data Appendix

Lightning. Our main measure of lightning density, originating from ground-based flash sensors, is from the US National Lightning Detection Network Database (NLDN). The NLDN consists of more than 100 remote, ground-based lightning sensors, which instantly detect the electromagnetic signals appearing when lightning strikes Earth's surface. The data is available as an average over the period 1996-2005 for the 48 contiguous US states from Vaisala's website: http://www.vaisala.com.

We find that lightning is not statistically different from a constant plus white noise (see main text for analysis). Therefore, we extend Vaisala's data to the period 1977-2007.

To investigate the time-series properties of lightning, we use data on the number of thunder days (TD) per year by state, available for the period 1901-1995. These data are collected as part of the Climate Change Detection and Attribution Program at the National Oceanic and Atmospheric Administration (NOAA). The raw data comes from 734 cooperative observer stations and 121 first order stations (see Changnon, 2001 for a detailed description). The data consists of monthly and yearly TD totals for 38 US states over the period 1901-1995, 40 states over the period 1906-1995 and 42 states over the period 1951-1995. It is available for purchase from the Midwestern Regional Climate Center:
http://mrcc.isws.illinois.edu/prod_serv/tstorm_cd/tstorm1.html.

From these data, we calculated the average yearly number of thunder days per state. Ultimately, we are interested in average *flash density* (FD) by state rather than thunder days per year. FDs are defined as the number of ground strikes per sq km per year. We converted yearly TDs into FDs using the following formula (Chisholm, 2000):

$$FD = 0.04 * TD^{1.25}$$

Temperature and Precipitation. Data from the United States Historical Climatology Network (USHCN) project, developed at NOAA's National Climatic Data Center (NCDC) to assist in the detection of regional climate change across the US. The USHCN project has produced a dataset of daily and monthly records of basic meteorological variables (maximum and minimum temperature, total precipitation, snowfall, and snow depth) from over 1000 stations across the 48 contiguous US states for the period 1900-2006.

The precipitation data we use is corrected by USHCN for the presence of outlier daily observations, time of data recording, and time series discontinuities due to random station moves and other station changes. The temperature data we use is additionally corrected for warming biases created by urbanization, and the replacement of liquid-in-glass thermometers by electronic temperature measurement devices during the mid 1980s.

We construct yearly average temperatures (expressed in degrees Celsius) and yearly average precipitation totals (expressed in cm per year) for each state, as simple averages of monthly data from 1221 stations across the country. The data is available at:
http://cdiac.ornl.gov/epubs/ndp/ushcn/newushcn.html.

Latitude. Latitude at the center of the state, calculated from geographic coordinates from the US Board on Geographic Names. The data is available at: http://geonames.usgs.gov/domestic/download_data.htm.

Altitude. Approximate mean elevation by state. Data source: US Geological Survey, Elevations and Distances in the United States, 1983. Available from the US Census Bureau at: http://www.census.gov/prod/2004pubs/04statab/geo.pdf.

Tornadoes, Wind, and Hail. The Storm Prediction Center of NOAA's National Weather Service Center provides data for tornadoes, wind, and hail for the period 1950-2007.

Data is available for the tornado occurrences and their damage categories in the Enhanced Fujita (EF) scale (assigning 6 levels from 0 to 5). We construct a measure of tornado intensity as the average damage category for all tornado occurrences during a year. For all the estimations, we rescale the EF categories from the original 0 to 5 scale to a 1 to 6 scale.

Wind is measured as the yearly average of wind speed, expressed in kilometers per hour.

Hail is measured as the average size of hail in centimeters.

The data is available at http://www.spc.noaa.gov/climo/historical.html.

Humidity, Sunshine and Cloudiness. Data from the "Comparative Climatic Data for the United States through 2007", published by NOAA.

(Relative) humidity is the average percentage amount of moisture in the air, compared to the maximum amount of moisture that the air can hold at the same temperature and pressure.

Cloudiness is measured as the average number of days per year with 8/10 to 10/10 average sky cover (or with 7/8 to 8/8 average sky cover since July 1996).

Sunshine is the total time that sunshine reaches the Earth's surface compared to the maximum amount of possible sunshine from sunrise to sunset with clear sky conditions.

The data is available at http://www1.ncdc.noaa.gov/pub/data/ccd-data/CCD-2007.pdf.

GSP per worker*.* Gross Domestic Product by state (GSP) per worker in chained 2000 US$.

US Bureau of Economic Analysis (BEA) offers two series of real GSP. The first is for the period 1977-1997, where industry classification is based on the Standard Industrial Classification (SIC) definitions. The second series covers the period 1997-2007 and relies on industrial classification based on the North American industrial Classification System (NAICS). Both GSP series are available at http://www.bea.gov/regional/gsp/.

We build a single measure of real GSP, extending levels of the series based on the SIC system with the yearly growth rates of the series based on the NAICS. This is equivalent to assuming

that from 1997 onwards, the growth rate of GSP per worker calculated with the SIC system equals the growth rate of real GSP calculated with the NAICS definitions.[26] Based on this estimate for real GSP, we construct a yearly series of real GSP per employed worker dividing real GSP by the number of employees per state. The growth rate is measured in percentages. State-by-state data for the number of employed workers is provided by the State Personal Income accounts at the US BEA (available at:
http://www.bea.gov/regional/spi).

Computers and Internet. Percentage of households with computer and percentage of households with Internet access at home in 2003. Data collected in a supplement to the October 2003 US Current Population Survey, available at:
http://www.census.gov/population/socdemo/computer/2003/tab01B.xls.

Manufacturing firms' IT investments. Capital expenditures on machinery and equipment for firms in the manufacturing sector are comprised by the following three categories: (1) Expenditures on automobiles, trucks, etc. for highway use. (2) Computers and peripheral data processing equipment. This item includes all purchases of computers and related equipment. (3) All other expenditures for machinery and equipment excluding automobiles and computer equipment. The variable we use is (2)/[(1)+(2)+(3)] ≡ Capital expenditures on computers and peripheral data processing equipment as a % of total capital expenditures on machinery and equipment of manufacturing firms. Data is from US Census Bureau, 2007 Economic Census. Detailed statistics for the manufacturing sector, by State, 2007 http://factfinder.census.gov/servlet/IBQTable?_bm=y&-geo_id=&-ds_name=EC0731A2&-_lang=en

Additional variables used in the paper

Variable	Definition and source
Human capital variables	This extended list of human capital variables is downloaded from www.allcountries.org.
Enrollment rate	Public elementary and secondary school enrollment as a percentage of persons 5-17 years old. From "Digest of Education Statistics", National Center of Education Statistics (NCES), Institute of Education Sciences, US Department of Education, http://nces.ed.gov/programs/digest/. Available at: http://www.allcountries.org/uscensus/266_public_elementary_and_secondary_school_enrollment.html.
High school degree or higher	Persons with a high school degree or higher as a percentage of persons 25 years and over. From "Digest of Education Statistics", National Center of Education Statistics (NCES), Institute of Education Sciences, US Department of Education, http://nces.ed.gov/programs/digest/d03/tables/dt011.asp.
Bachelor's degree or higher	Persons with a bachelor's degree or higher as a percentage of persons 25 years and over.

[26] BEA warns against merging the *level* of the two series of real GSP directly, since the discontinuity in the industrial classification system will obviously affect level and growth rate estimates. Our choice of merging the *growth rates* of the two series can be justified recalling both the SIC and the NAICS aim to classify production of all industries in each state, so that the growth rate of both GSP series in levels is comparable. As a check, we computed the correlation between the growth rate of aggregate US GDP and gross domestic income (GDI), since GDP corresponds to the NAICS-definition and GDI corresponds to the SIC-definition (BEA, http://www.bea.gov/regional/gsp/). The correlation is higher than 0.99 for different periods between 1929 and 2007.

	Same source as high school degree or higher.
College degree or higher	Persons with a college degree or higher as a percentage of persons 25 years and over. Same source as high school degree or higher and bachelor's degree or higher.
Graduate or professional degree	Persons with a graduate or professional degree as a percentage of persons 25 years and over. Same source as high school degree or higher, bachelor's degree or higher, and college degree or higher.
Additional determinants of IT diffusion	In addition to human capital, Caselli and Coleman (2001) suggest the following set of determinants of computer technology diffusion across countries: real income, GDP shares of different sectors, stock of human capital, amount of trade, and degree of integration to the world economy. We gathered similar data for US states, described below.
Shares of agriculture production, manufacturing production, and government spending in GSP	Agriculture, forestry, fishing, and hunting production as % of GSP; Manufacturing production as % of GSP, Total Government spending as % of GSP. The 3 variables constructed from US BEA's data of GSP by industry, in millions of current US$. Available at: http://www.bea.gov/regional/gsp/.
Agricultural exports per capita	Agricultural exports per capita (US$). Total value of Agricultural exports by state, from US Department of Agriculture, divided by population. Available at: http://www.ers.usda.gov/Data/StateExports/2006/SXHS.xls Population data from US Census Bureau.
FDI per capita	Gross value of Property, Plant, and Equipment (PPE) of Nonbank US Affiliates, per capita (US$). Data on PPE available from US BEA for the period 1999-2006 available at: http://bea.doc.gov/international/xls/all_gross_ppe.xls. For the year 1981 and the period 1990-1997 available at: http://allcountries.org/uscensus/1314_foreign_direct_investment_in_the_u.html. Population data from US Census Bureau.
Institutional and historical determinants of productivity	All variables are taken from Mitchener and McClean (2003).
% workforce in mining, 1880	Percentage of the workforce employed in mining in 1880.
Average no. cooling degree days	The average number of cooling degree days is computed as the number of days in which the average air temperature rose above 65 degrees Fahrenheit (18 degrees Celsius) times the number of degrees on those days which the average daily air temperature exceeded 65 over the year.
% of 1860 population in slavery	The total number of slaves as a percentage of the total population of each state in 1860.
% of 1860 population on large slave plantations	The number of slaves owned by slaveholders having more than 20 slaves as a percentage of the total population of each state in 1860.
Access to navigable water	An indicator variable that takes the value of one if a state borders the ocean/Great Lake /river, and zero otherwise.
Settler origin	A series of indicator variables which take on positive values if a state, prior to statehood, had ties with that colonial power.
Average annual soldier mortality in 1829-1838, 1839-1854, %	Soldier mortality rates at the state level are derived using US soldier mortality data for individual forts. Quarterly data were collected by the US Surgeon General and Adjutant General's Offices 1829-1838 and by the US Surgeon General's Office for 1839-1854. Mitchener and McClean obtained the yearly mortality rates by dividing the number of deaths each year by the average annual "mean strength" of soldiers.
Socio-demographic indicators	Data on religiousness, race and ethnicity, urbanization and age structure of the population; from various sources.
Church attendance, average 2004-2006	Data from a Gallup Poll analysis, conducted between January 2004 and March 2006, based on responses to the question, "How often do you attend church or synagogue – at least once a week, almost every week, about once a month, seldom, or never?" Available at: http://www.gallup.com/poll/22579/church-attendance-lowestnew-england-highest-south.aspx#2
% of white population, black population, and	Data for race and Hispanic origin for the US, regions, divisions, and states (100-Percent Data). Source: US Census Bureau.

population of Hispanic origin	Available at: http://www.census.gov/population/www/documentation/twps0056/tabA-03.xls (for 1980), and http://www.census.gov/population/www/documentation/twps0056/tabA-01.xls (for 1990).
% of urban population	Rural and Urban population 1900-1990 (released 1995). Source: US Census Bureau. Available at: http://www.census.gov/population/www/censusdata/files/urpop0090.txt
% of population 15 years or less, and % of population between 15-64 years	Population by broad age group. "Demographic Trends in the 20th Century", Table 7, parts D and E. Source: US Census Bureau. Available at: http://www.census.gov/prod/2002pubs/censr-4.pdf

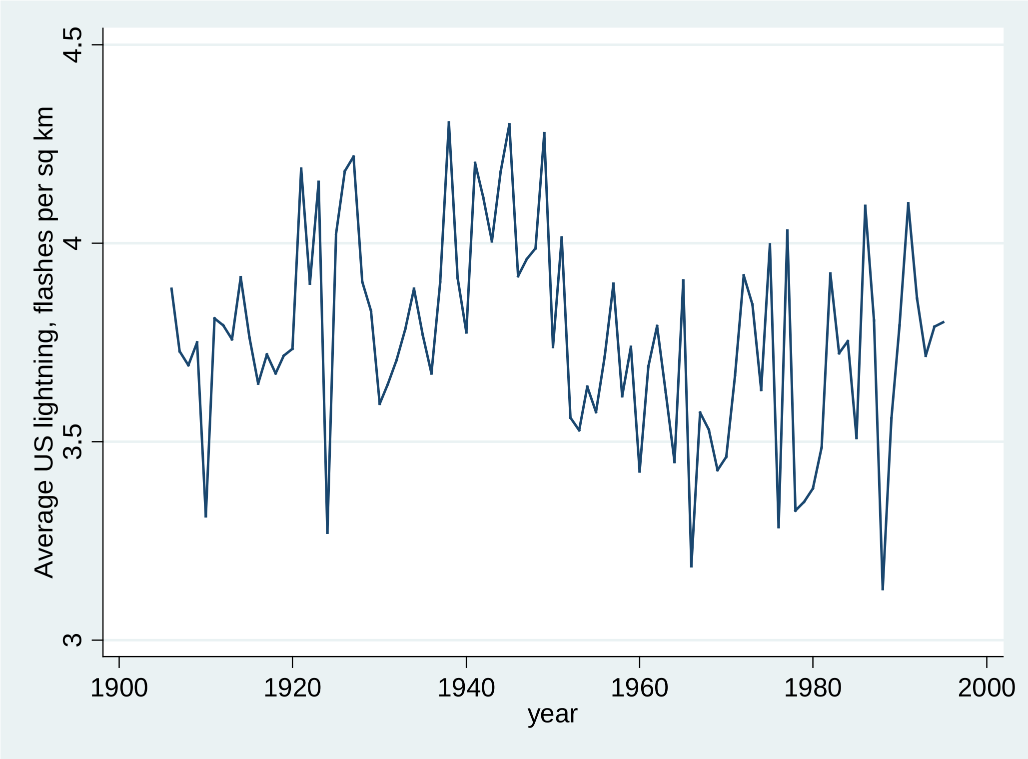


Figure 1. The average flash density in the US: 40 states

Source: Lightning observations from weather stations, transformed from thunder days (TD) into flash density (FD) using the formula FD = 0.04*TD1.25. See Data Appendix for details.

Notes: Only 40 states have complete information for the period 1906-1995. The "left-out" (contiguous) states are Connecticut, Delaware, New Hampshire, New Jersey, Rhode Island, Vermont, Mississippi, and West Virginia. The figure shows the weighted average, where the weight is determined by state size.

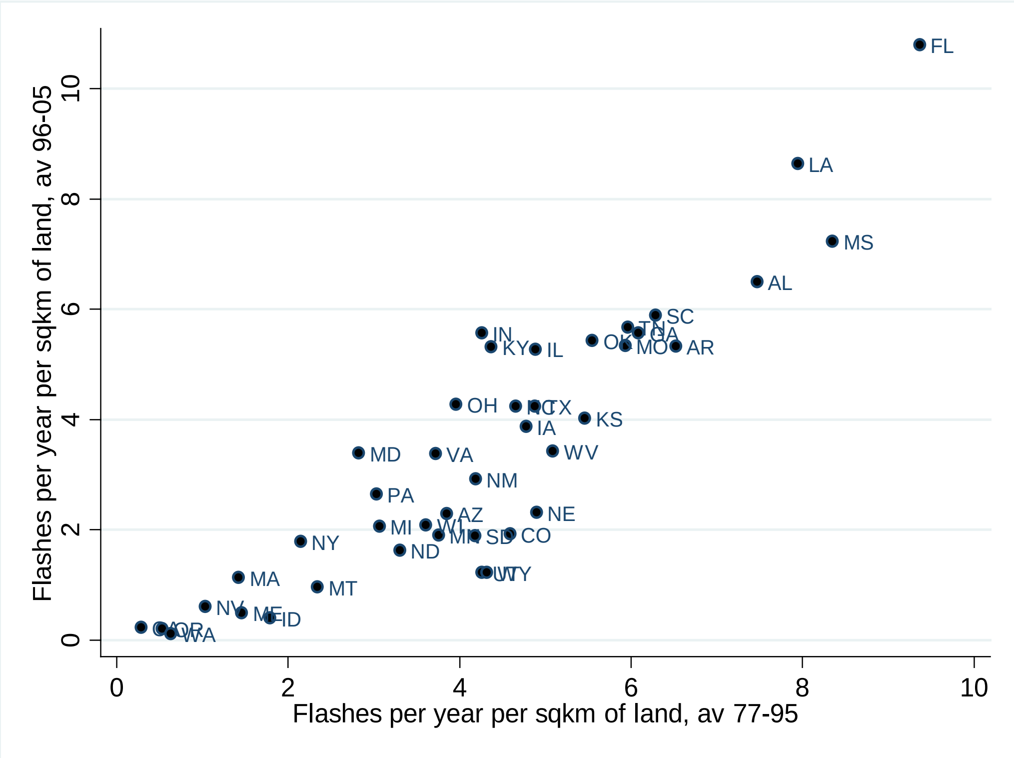


Figure 2. The average flash density 1977-95 versus 1996-2005: 42 states.

Sources: 1977-95 based on Thunder days (TD) from weather station observations, converted into flash density (FD) using the formula FD = 0.04*TD1.25. 1996-2005 data are based on ground detectors. See Appendix for further details.

Notes: The correlation is 0.90, and a regression, FL96-05 = a + bFL77-95 returns: a=-0.99, b=1.05, R2=0.81.

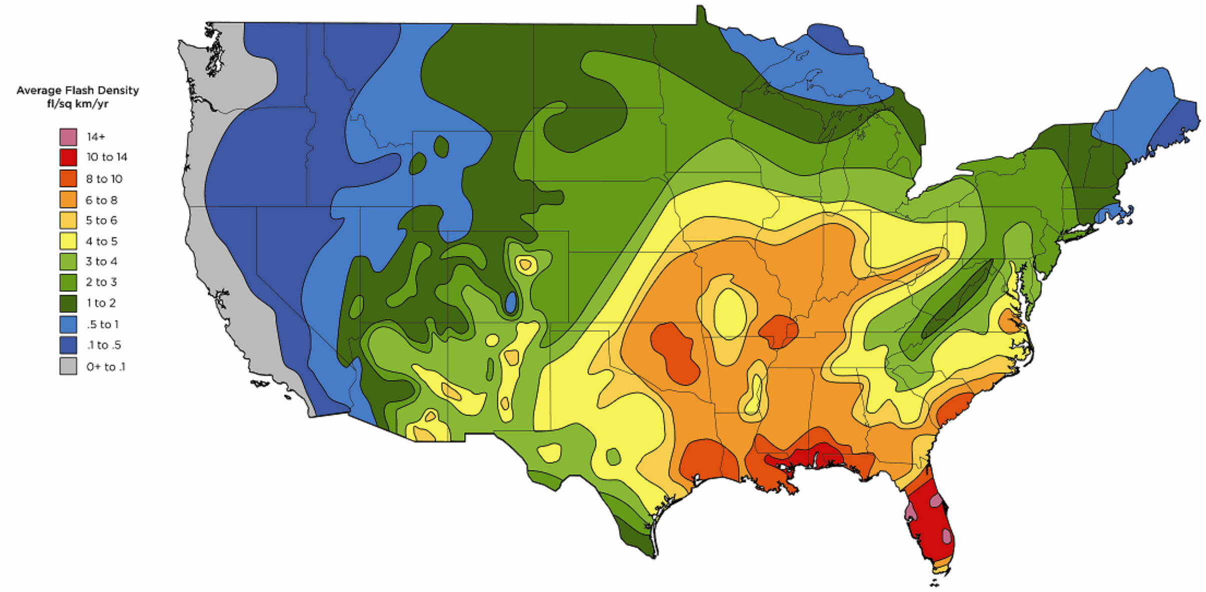


Figure 3. The distribution of flash densities across the US: 1996-2007.
Source: http://www.vaisala.com.

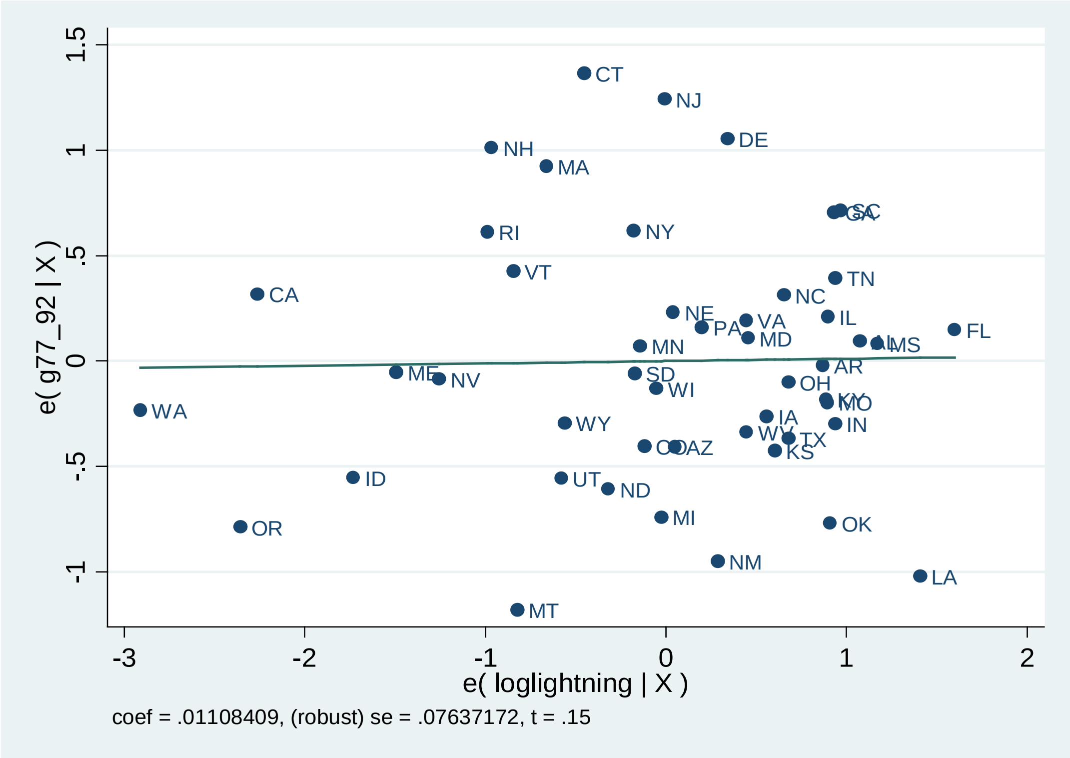


Figure 4. The correlation between state growth and (log) flash density, conditional on initial income per worker: 1977-1992.

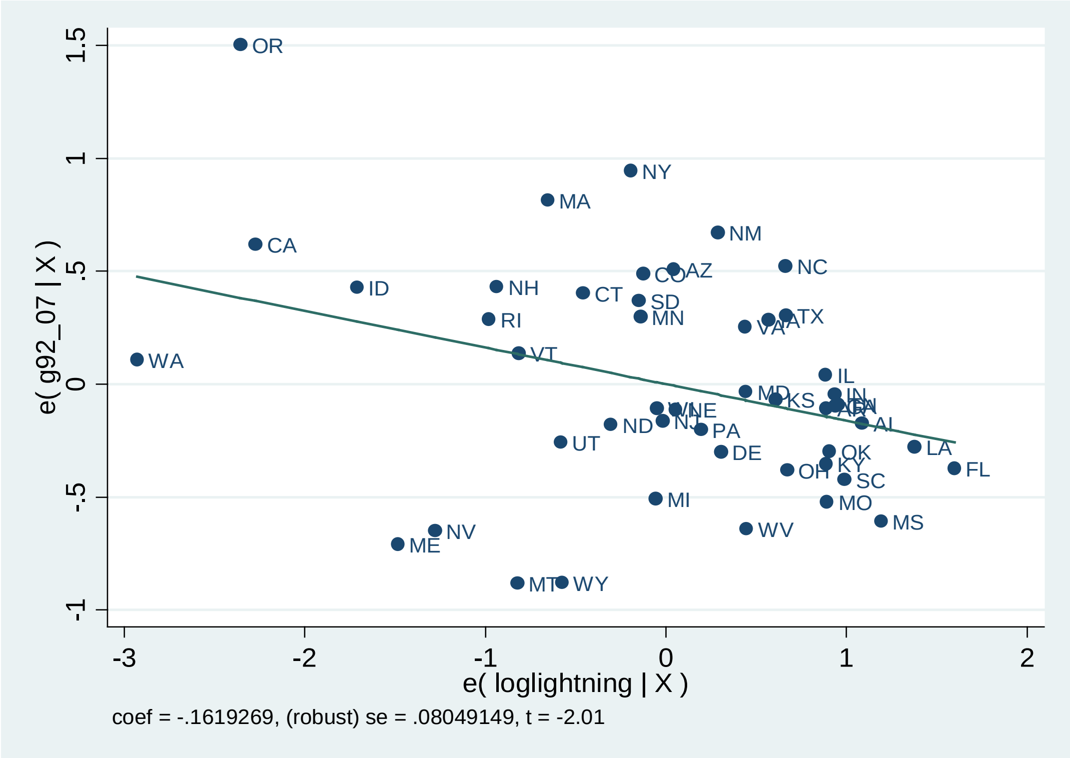


Figure 5. The correlation between state growth and (log) flash density, conditional on initial income per worker: 1992-2007.

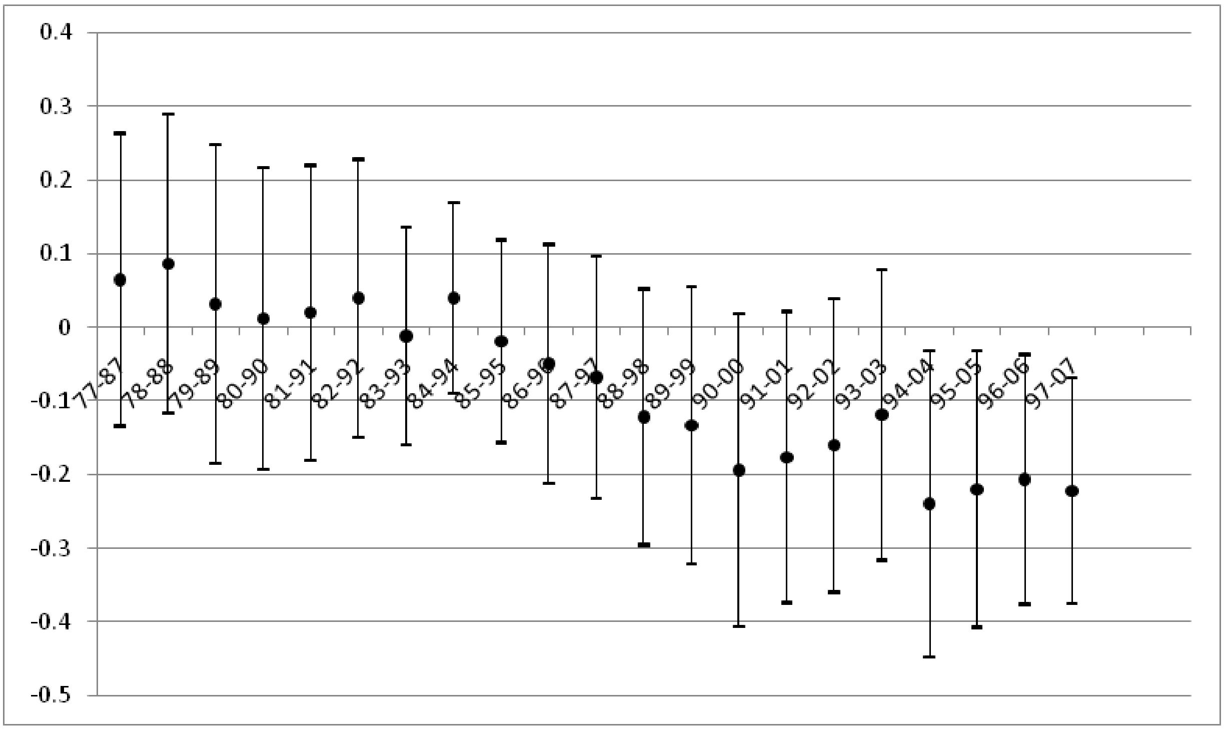


Figure 6. The lightning-growth nexus: 1977-2007.

Notes: The figure shows estimates for b2 (and the associated 95 percent confidence interval) from regressions of the form: G = b_0 + $b_1$log(y_{t-10})+ b2log(lightning)+e, where y is gross state product per worker and t=1987,…,2007. 48 states; estimated by OLS.

Figure 7. Lightning versus Internet users per 100 households in 2003.
Sources: See Data Appendix
Notes: The raw correlation between the two series is -0.62.

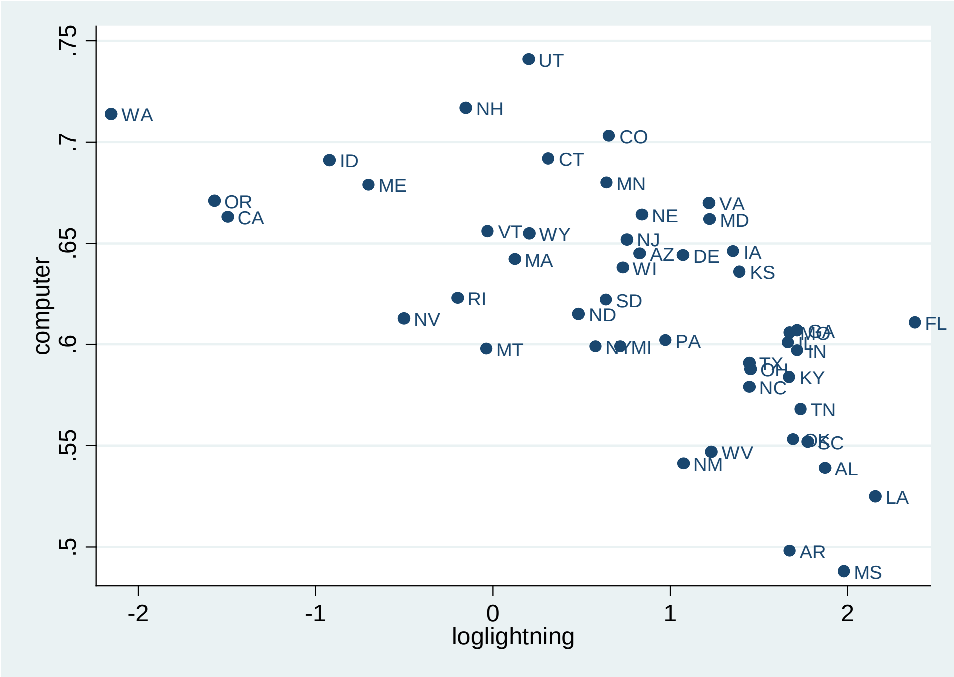


Figure 8. Lightning versus personal computers per 100 households in 2003.
Sources: See Data Appendix.

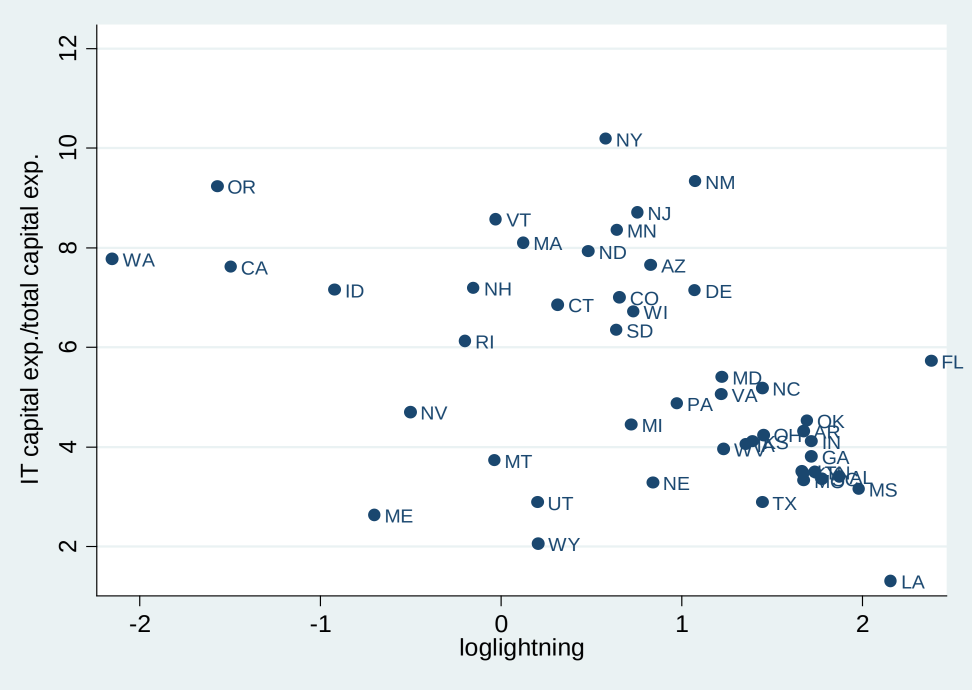


Figure 9. Lightning versus manufacturing firms' ICT capital expenditure to total capital expenditure.

Sources: See Data Appendix.

Notes: The raw correlation between the two series is -0.49.

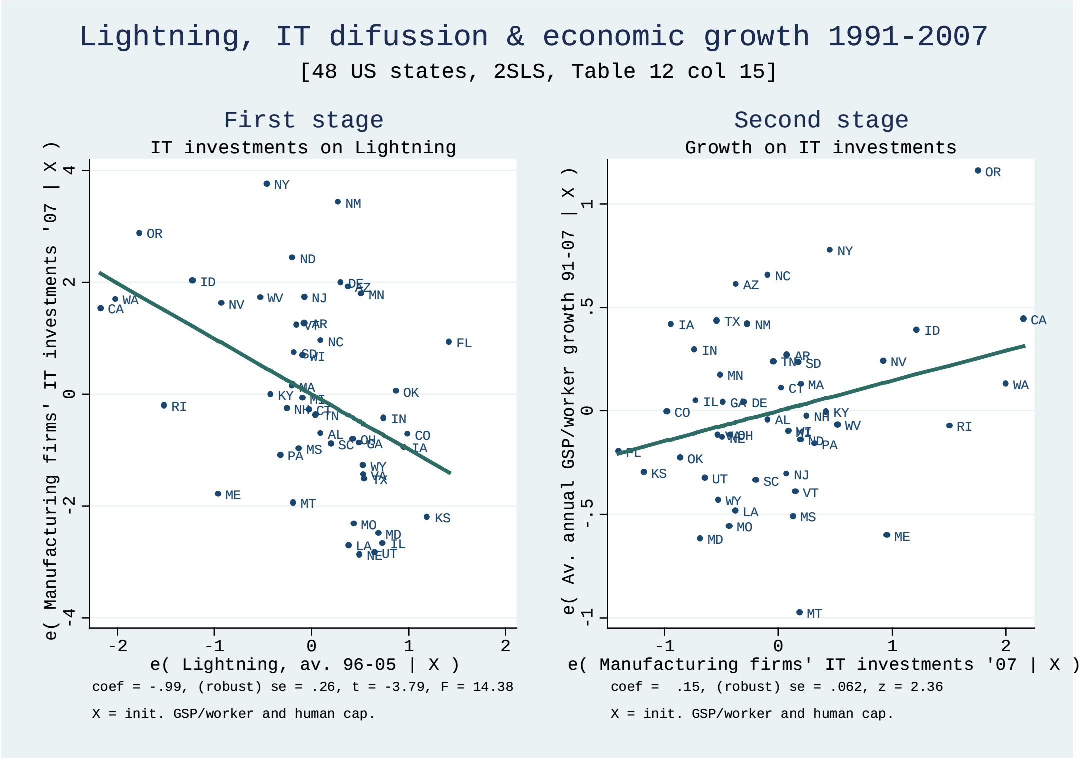


Figure 10. Exogenous component of manufacturing firms' ICT capital expenditure to total capital expenditure and economic growth, 1991-2007.

Sources: See Data Appendix.

Notes: Estimated by 2SLS.

Table 1. Dickey-Fuller tests for unit root in lightning

	test-statistic	p-value	No. obs.	No. lags
Aggregate US	-4.52	0.0000	88	1
Alabama	-5.31	0.0000	88	1
Arizona	-3.38	0.0118	87	2
Arkansas	-8.98	0.0000	89	0
California	-8.40	0.0000	89	0
Colorado	-8.69	0.0000	89	0
Florida	-8.19	0.0000	89	0
Georgia	-8.58	0.0000	89	0
Idaho	-3.48	0.0085	87	2
Illinois	-9.61	0.0000	89	0
Indiana	-8.24	0.0000	89	0
Iowa	-9.42	0.0000	89	0
Kansas	-4.46	0.0002	88	1
Kentucky	-2.94	0.0412	87	2
Louisiana	-4.62	0.0001	88	1
Maine	-2.75	0.0662	87	2
Maryland	-5.32	0.0000	88	1
Massachusetts	-9.25	0.0000	89	0
Michigan	-8.76	0.0000	89	0
Minnesota	-10.28	0.0000	89	0
Missouri	-9.92	0.0000	89	0
Montana	-9.01	0.0000	89	0
Nebraska	-3.64	0.0051	87	2
Nevada	-10.02	0.0000	89	0
New Mexico	-3.58	0.0062	87	2
New York	-4.01	0.0013	88	1
North Carolina	-5.40	0.0000	88	1
North Dakota	-7.84	0.0000	89	0
Ohio	-3.59	0.0059	87	2
Oklahoma	-11.61	0.0000	89	0
Oregon	-7.09	0.0000	89	0
Pennsylvania	-2.20	0.2045	86	3
South Carolina	-8.01	0.0000	89	0
South Dakota	-8.62	0.0000	89	0
Tennessee	-7.32	0.0000	89	0
Texas	-5.45	0.0000	88	1
Utah	-5.55	0.0000	88	1
Virginia	-7.41	0.0000	89	0
Washington	-8.75	0.0000	89	0
Wisconsin	-9.45	0.0000	89	0
Wyoming	-7.71	0.0000	89	0

Notes. The Augmented Dickey-Fuller test with no deterministic trend for each of the 40 states over the period 1906-1995. Lags selected by Schwarz's information criteria. Lightning is average number of flashes per year per square km, measured at weather stations.

Table 2. Summary statistics for the main variables

				Percentiles				
	Obs.	**Mean**	**Std. Dev.**	**99%**	**75%**	**50%**	**25%**	**1%**
Average annual growth rate of real GSP per worker (%):								
1977-1987	48	0.81	0.77	2.69	1.32	0.74	0.30	-0.76
1987-1997	48	1.21	0.58	2.67	1.50	1.22	0.82	-0.32
1997-2007	48	1.18	0.54	2.59	1.49	1.15	0.74	0.26
1977-2007	48	1.07	0.42	1.97	1.37	1.07	0.82	0.10
1991-2007	48	1.34	0.50	2.79	1.71	1.35	1.01	0.29
Lightning density, average 1996-2005 (flashes/year/sq km)	48	3.18	2.39	10.8	5.30	2.48	1.23	0.12
Manufacturing firms' IT investments, 2007								
(% of non-construction capital expenditures)	48	5.40	2.20	10.19	7.17	4.78	3.51	1.31
Access to Internet at home, 2003 (% of households)	48	54.39	5.88	65.50	58.10	55.00	51.20	39.50
Computer at home, 2003 (% of households)	48	62.10	5.71	74.10	66.25	61.85	58.95	48.80

Notes. Lightning defined as average number of flashes per year per square km over the period 1995-2006, measured by flash-detectors. IT capital expenditures defined as capital expenditures on computers and peripheral data processing equipment in all manufactuting firms in 2007, expressed as a percentage of all non-construction capital expenditures. Data sources and extended definitions are provided in the Data appendix.

Table 3. Growth and lightning

(1)	5-year periods	**1977-1982**	**1982-1987**	**1987-1992**	**1992-1997**	**1997-2002**	**2002-2007**	Observations	R-squared
		-0.04	0.17	-0.09	-0.04	-0.28**	-0.18*	288	0.20
		[0.10]	[0.16]	[0.09]	[0.12]	[0.11]	[0.09]		
(2)	10-year periods	**1977-1987**		**1987-1997**		**1997-2007**		Observations	R-squared
		0.07		-0.07		-0.22***		144	0.15
		[0.10]		[0.08]		[0.08]			
(3)	15-year periods	**1977-1992**			**1992-2007**			Observations	R-squared
		0.01			-0.16**			96	0.20
		[0.08]			[0.08]				

Notes. Pooled OLS estimates of the coefficient on lightning (b_{2t}). The dependent variable in regressions (1), (2) and (3) is the yearly average growth rate in GSP per worker over periods of 5, 10, and 15 years, respectively. All regressions include a constant, the initial level of (log) real GSP per worker and a full set of time-dummies. Lightning is the average number of flashes per year per square km, measured by flash-detectors. Robust standard errors in brackets, adjusted for clustering at state level. Asterisks ***, **, and * indicate significance at the 1, 5, and 10%, respectively.

Table 4. Growth and lightning - controlling for human capital and regional fixed effects

Dependent variable:	Average annual growth in GSP per worker over periods of 10 years (1977 - 1987, 1987 - 1997, 1997 - 2007)					
	(1)	(2)	(3)	(4)	(5)	(6)
(log, initial) Real GSP per worker	-0.72	-1.24***	-0.60	-1.25***	-1.80***	-1.97***
	[0.45]	[0.41]	[0.46]	[0.44]	[0.41]	[0.54]
(log) Lightning × $t_{77\text{-}87}$	0.07	-0.04	-0.14	0.13	-0.12	-0.04
	[0.10]	[0.11]	[0.12]	[0.11]	[0.11]	[0.15]
(log) Lightning × $t_{87\text{-}97}$	-0.07	-0.16**	-0.07	0.03	-0.12	-0.05
	[0.08]	[0.07]	[0.09]	[0.08]	[0.08]	[0.14]
(log) Lightning × $t_{97\text{-}07}$	-0.22***	-0.24***	-0.22**	-0.13*	-0.21**	-0.17
	[0.08]	[0.08]	[0.09]	[0.08]	[0.08]	[0.14]
(initial) Enrollment rate × $t_{77\text{-}87}$		-0.07***			-0.06***	-0.04*
		[0.02]			[0.02]	[0.02]
(initial) Enrollment rate × $t_{87\text{-}97}$		-0.07***			-0.07***	-0.05*
		[0.02]			[0.02]	[0.03]
(initial) Enrollment rate × $t_{97\text{-}07}$		-0.03			-0.01	0.01
		[0.02]			[0.02]	[0.02]
(initial) High school degree or higher × $t_{77\text{-}87}$			-0.04***		-0.06***	-0.05***
			[0.01]		[0.02]	[0.02]
(initial) High school degree or higher × $t_{87\text{-}97}$			-0.0016		-0.02	-0.01
			[0.015]		[0.02]	[0.02]
(initial) High school degree or higher × $t_{97\text{-}07}$			-0.00076		-0.05**	-0.03
			[0.019]		[0.02]	[0.03]
(initial) Bachelor's degree or higher × $t_{77\text{-}87}$				0.18	0.51***	0.50***
				[0.16]	[0.16]	[0.15]
(initial) Bachelor's degree or higher × $t_{87\text{-}97}$				0.06**	0.07**	0.06
				[0.02]	[0.03]	[0.04]
(initial) Bachelor's degree or higher × $t_{97\text{-}07}$				0.07***	0.10***	0.09***
				[0.01]	[0.02]	[0.02]
Observations	144	144	144	144	144	144
R-squared	0.15	0.28	0.20	0.24	0.44	0.47
Regional fixed effects (8 BEA economic areas)	No	No	No	No	No	Yes
Joint significance tests (p values):						
H_0: Regional FEs = 0	.	.	.	.	.	0.79
H_0: Regional FEs and lightning terms = 0	.	.	.	.	.	0.0065

Notes. Pooled OLS estimates. The dependent variable is the yearly growth rate of GSP per worker over the periods 1977-1987, 1987-1997, and 1997-2007. Lightning is the average number of flashes per year per square km, measured by flash-detectors. The different proxies for human capital are described in the appendix, and measured at the beginning of each 10-year period (1977, 1987 and 1997), except for enrollment rates (measured in 1980 instead of 1977 for the first period) and the % of population with a highschool degree or higher (measured in 1980, 1990 and 2000 instead of 1977, 1987 and 1997 for each respective period), due to data availability. The set of regional fixed effects in column (6) accounts for the 8 US Bureau of Economic Analysis' economic areas. All regressions include a constant and a full set of time-dummies. Robust standard errors in brackets, adjusted for clustering at the state level. Asterisks ***, **, and * indicate significance at the 1, 5, and 10%, respectively.

Table 5. Growth regressions with lightning and other geographical and climate variables

Dependent variable:	Average annual growth in GSP per worker over periods of 10 years (1977-1987, 1987-1997, 1997-2007)										
GEOGRAPHY:		Temperature (C degrees)	Precipitation (cm/year)	Tornado intensity (av EF-scale)	Hail size (cm)	Wind speed (km/h)	Humidity (% moisture in air)	Cloudiness (days/year)	Sunshine (days/year)	Elevation (meters above sea level)	Latitude (degrees)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(log, initial) Real GSP per worker	-1.80***	-1.71***	-1.82***	-1.83***	-2.01***	-1.83***	-1.76***	-1.73***	-1.93***	-1.81***	-1.72***
	[0.41]	[0.39]	[0.45]	[0.45]	[0.42]	[0.44]	[0.42]	[0.41]	[0.47]	[0.43]	[0.40]
(log) Lightning × t_{77-87}	-0.12										
	[0.11]										
(log) Lightning × t_{87-97}	-0.12										
	[0.08]										
(log) Lightning × t_{97-07}	-0.21**										
	[0.08]										
(log) GEOGRAPHY × t_{77-87}		-0.38	0.77*	1.11*	-1.36**	-0.41*	1.08	0.76	-0.91	-0.31**	1.07
		[0.26]	[0.41]	[0.60]	[0.66]	[0.20]	[1.05]	[0.50]	[0.67]	[0.13]	[0.93]
(log) GEOGRAPHY × t_{87-97}		0.31	0.14	0.082	-0.086	0.063	-1.06	-0.25	0.028	0.13	-0.34
		[0.29]	[0.39]	[0.48]	[0.71]	[0.11]	[0.88]	[0.42]	[0.50]	[0.093]	[0.99]
(log) GEOGRAPHY × t_{97-07}		-0.033	0.042	-0.25	-1.79*	0.32	-0.38	-0.11	-0.09	0.13	0.95
		[0.35]	[0.19]	[0.22]	[0.95]	[0.32]	[0.59]	[0.34]	[0.48]	[0.087]	[0.80]
Observations	144	144	144	144	144	144	144	144	141	144	144
R-squared	0.44	0.42	0.43	0.43	0.44	0.43	0.42	0.42	0.42	0.46	0.42
Human capital controls (enrollment, high school or higher, BA)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes. Pooled OLS estimates. The dependent variable is the annual growth rate in GSP per worker over the periods 1977-1987, 1987-1997 and 1997-2007. All regressions include a constant and a full set of time-dummies. Lightning is the average number of flashes per year per square km, measured by flash-detectors. The controls for human capital are the initial enrollment rate, percentage of population with a high school or higher degree, and percentage of population with a BA degree. All the human capital controls are measured at the beginning of each 10-year period (1977, 1987 and 1997), except for enrollment rates (measured in 1980 instead of 1977) and the % of population with a highschool degree or higher (measured in 1980, 1990 and 2000 instead of 1977, 1987 and 1997), due to data availability. All geographic/climate variables are averages taken over periods of 10 years. Robust standard errors in brackets, adjusted for clustering at state level. Asterisks ***, **, and * indicate significance at the 1, 5, and 10%, respectively.

Table 6. Growth regressions with lightning and geographical and climate controls

Dependent variable:	Average annual growth in GSP per worker over periods of 10 years (1977-1987, 1987-1997, 1997-2007)										
GEOGRAPHY:		Temperature (C degrees)	Precipitation (cm/year)	Tornado intensity (av EF-scale)	Hail size (cm)	Wind speed (km/h)	Humidity (% moisture in air)	Cloudiness (days/year)	Sunshine (days/year)	Elevation (meters above sea level)	Latitude (degrees)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(log, initial) Real GSP per worker	-1.80***	-1.80***	-1.85***	-1.85***	-1.96***	-1.84***	-1.78***	-1.77***	-1.98***	-1.86***	-1.81***
	[0.41]	[0.37]	[0.44]	[0.42]	[0.43]	[0.42]	[0.41]	[0.39]	[0.44]	[0.41]	[0.38]
(log) Lightning × $t_{77\text{-}87}$	-0.12	-0.085	-0.07	-0.13	-0.048	-0.19	-0.12	-0.059	-0.07	-0.16	-0.07
	[0.11]	[0.11]	[0.11]	[0.11]	[0.11]	[0.12]	[0.11]	[0.12]	[0.11]	[0.11]	[0.13]
(log) Lightning × $t_{87\text{-}97}$	-0.12	-0.16*	-0.11	-0.12	-0.13	-0.11	-0.11	-0.15	-0.14	-0.098	-0.20*
	[0.08]	[0.096]	[0.085]	[0.084]	[0.085]	[0.086]	[0.089]	[0.10]	[0.095]	[0.084]	[0.12]
(log) Lightning × $t_{97\text{-}07}$	-0.21**	-0.24***	-0.21**	-0.20**	-0.20**	-0.20**	-0.21**	-0.21***	-0.22***	-0.19**	-0.23**
	[0.08]	[0.078]	[0.079]	[0.079]	[0.086]	[0.083]	[0.084]	[0.077]	[0.077]	[0.081]	[0.095]
(log) GEOGRAPHY × $t_{77\text{-}87}$		-0.31	0.73*	1.14*	-1.24*	-0.49**	1.01	0.67	-0.76	-0.34***	0.76
		[0.28]	[0.39]	[0.61]	[0.69]	[0.20]	[1.06]	[0.51]	[0.66]	[0.11]	[1.05]
(log) GEOGRAPHY × $t_{87\text{-}97}$		0.44	0.10	0.17	0.23	0.055	-1.00	-0.42	0.28	0.11	-1.26
		[0.28]	[0.39]	[0.51]	[0.75]	[0.11]	[0.84]	[0.46]	[0.52]	[0.097]	[1.22]
(log) GEOGRAPHY × $t_{97\text{-}07}$		0.30	0.06	-0.13	-0.28	0.16	0.0031	-0.23	0.15	0.065	-0.43
		[0.39]	[0.19]	[0.21]	[0.81]	[0.29]	[0.73]	[0.28]	[0.36]	[0.082]	[0.95]
Observations	144	144	144	144	144	144	144	144	141	144	144
R-squared	0.44	0.46	0.47	0.47	0.46	0.47	0.45	0.46	0.46	0.49	0.45
Human capital controls (enrollment, high school or higher, BA)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes. Pooled OLS estimates. The dependent variable is the annual growth rate in GSP per worker over the periods 1977-1987, 1987-1997 and 1997-2007. All regressions include a constant and a full set of time-dummies. Lightning is the average number of flashes per year per square km, measured by flash-detectors. The controls for human capital are the initial enrollment rate, percentage of population with a high school or higher degree, and percentage of population with a BA degree. All the human capital controls are measured at the beginning of each 10-year period (1977, 1987 and 1997), except for enrollment rates (measured in 1980 instead of 1977) and the % of population with a highschool degree or higher (measured in 1980, 1990 and 2000 instead of 1977, 1987 and 1997), due to data availability. All geographic/climate variables are averages taken over periods of 10 years. Robust standard errors in brackets, adjusted for clustering at state level. Asterisks ***, **, and * indicate significance at the 1, 5, and 10%, respectively.

Table 7. Growth regressions with historical controls (geography and institutions)

Dependent variable:	Average annual growth in GSP per worker over periods of 10 years (1977-1987, 1987-1997, 1997-2007)										
HISTORY:		% of workforce in mining, 1880	Average no. of cooling degree days	% of 1860 population in slavery	Access to navigable water	% of 1860 population on large slave plantations	Settler origin: English	Settler origin: French	Settler origin: Spanish	Settler origin: Dutch	Average annual soldier mortality in 1829-1838, 1839-1854, %
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(log, initial) Real GSP per worker	-1.80***	-1.77***	-1.75***	-1.82***	-1.89***	-1.82***	-1.70***	-2.07***	-1.70***	-1.85***	-1.59***
	[0.41]	[0.43]	[0.41]	[0.42]	[0.46]	[0.42]	[0.40]	[0.34]	[0.40]	[0.44]	[0.42]
(log) Lightning × $t_{77\text{-}87}$	-0.12										
	[0.11]										
(log) Lightning × $t_{87\text{-}97}$	-0.12										
	[0.08]										
(log) Lightning × $t_{97\text{-}07}$	-0.21**										
	[0.08]										
HISTORY × $t_{77\text{-}87}$		-0.015*	-0.017*	0.0020	0.49	0.0041	0.49***	-0.33*	-0.45***	0.30	-0.20**
		[0.0082]	[0.0092]	[0.0065]	[0.31]	[0.010]	[0.17]	[0.16]	[0.17]	[0.22]	[0.097]
HISTORY × $t_{87\text{-}97}$		-0.0011	-0.0075	-0.0094	0.23	-0.017**	-0.022	-0.50***	-0.11	0.075	-0.017
		[0.010]	[0.0093]	[0.0058]	[0.31]	[0.0083]	[0.19]	[0.15]	[0.16]	[0.25]	[0.098]
HISTORY × $t_{97\text{-}07}$		0.0078	-0.00097	-0.0097*	0.15	-0.014*	-0.14	-0.18	0.092	0.22	-0.19*
		[0.0061]	[0.011]	[0.0049]	[0.17]	[0.0071]	[0.17]	[0.13]	[0.15]	[0.29]	[0.10]
Observations	144	144	144	144	144	144	144	144	144	144	144
R-squared	0.44	0.42	0.42	0.43	0.43	0.43	0.45	0.46	0.44	0.41	0.43
Human capital controls (enrollment, high school or higher, BA)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes. Pooled OLS estimates. The dependent variable is the annual growth rate of GSP per worker over the periods 1977-1987, 1987-1997 and 1997-2007. All regressions include a constant and a full set of time-dummies. Lightning is the average number of flashes per year per square km, measured by flash-detectors. The controls for human capital are the initial enrollment rate, percentage of population with a high school or higher degree, and percentage of population with a BA degree. All the human capital controls are measured at the beginning of each 10-year period (1977, 1987 and 1997), except for enrollment rates (measured in 1980 instead of 1977) and the % of population with a highschool degree or higher (measured in 1980, 1990 and 2000 instead of 1977, 1987 and 1997), due to data availability. HISTORY controls taken from Mitchener and McLean (2004). Robust standard errors in brackets, adjusted for clustering at state level. Asterisks ***, **, and * indicate significance at the 1, 5, and 10%, respectively.

Table 8. Growth regressions with lightning and historical controls (geography and institutions)

Dependent variable:	Average annual growth in GSP per worker over periods of 10 years (1977-1987, 1987-1997, 1997-2007)										
HISTORY:		% of workforce in mining, 1880	Average no. of cooling degree days	% of 1860 population in slavery	Access to navigable water	% of 1860 population on large slave plantations	Settler origin: English	Settler origin: French	Settler origin: Spanish	Settler origin: Dutch	Average annual soldier mortality in 1829-1838, 1839-1854, %
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(log, initial) Real GSP per worker	-1.80***	-1.80***	-1.76***	-1.80***	-1.92***	-1.79***	-1.73***	-2.03***	-1.73***	-1.86***	-1.65***
	[0.41]	[0.42]	[0.40]	[0.41]	[0.44]	[0.41]	[0.39]	[0.35]	[0.39]	[0.42]	[0.40]
(log) Lightning × t_{77-87}	-0.12	-0.14	-0.051	-0.14	-0.10	-0.14	-0.098	-0.072	-0.10	-0.11	-0.087
	[0.11]	[0.12]	[0.12]	[0.12]	[0.12]	[0.12]	[0.10]	[0.12]	[0.10]	[0.11]	[0.11]
(log) Lightning × t_{87-97}	-0.12	-0.12	-0.11	-0.072	-0.12	-0.067	-0.12	-0.047	-0.11	-0.12	-0.12
	[0.08]	[0.083]	[0.092]	[0.10]	[0.086]	[0.100]	[0.085]	[0.11]	[0.084]	[0.083]	[0.089]
(log) Lightning × t_{97-07}	-0.21**	-0.20**	-0.28***	-0.18*	-0.22***	-0.19**	-0.20***	-0.20**	-0.20**	-0.21**	-0.18**
	[0.08]	[0.089]	[0.079]	[0.092]	[0.079]	[0.089]	[0.076]	[0.092]	[0.077]	[0.080]	[0.088]
HISTORY × t_{77-87}		-0.016**	-0.015	0.0042	0.47	0.0068	0.48***	-0.29	-0.44**	0.28	-0.18*
		[0.0079]	[0.0098]	[0.0075]	[0.30]	[0.011]	[0.17]	[0.20]	[0.17]	[0.23]	[0.097]
HISTORY × t_{87-97}		-0.0035	-0.0017	-0.0076	0.23	-0.015	-0.022	-0.48***	-0.10	0.037	0.016
		[0.010]	[0.0092]	[0.0070]	[0.31]	[0.0096]	[0.19]	[0.18]	[0.16]	[0.25]	[0.10]
HISTORY × t_{97-07}		0.0015	0.017***	-0.0044	0.21	-0.0065	-0.10	-0.055	0.065	0.19	-0.12
		[0.0065]	[0.0058]	[0.0056]	[0.17]	[0.0079]	[0.14]	[0.15]	[0.13]	[0.27]	[0.12]
Observations	144	144	144	144	144	144	144	144	144	144	144
R-squared	0.44	0.46	0.46	0.45	0.47	0.46	0.48	0.48	0.48	0.45	0.46
Human capital controls (enrollment, high school or higher, BA)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes. Pooled OLS estimates. The dependent variable is the annual growth rate of GSP per worker over the periods 1977-1987, 1987-1997 and 1997-2007. All regressions include a constant and a full set of time-dummies. Lightning is the average number of flashes per year per square km, measured by flash-detectors. The controls for human capital are the initial enrollment rate, percentage of population with a high school or higher degree, and percentage of population with a BA degree. All the human capital controls are measured at the beginning of each 10-year period (1977, 1987 and 1997), except for enrollment rates (measured in 1980 instead of 1977) and the % of population with a highschool degree or higher (measured in 1980, 1990 and 2000 instead of 1977, 1987 and 1997), due to data availability. HISTORY controls taken from Mitchener and McLean (2004). Robust standard errors in brackets, adjusted for clustering at state level. Asterisks ***, **, and * indicate significance at the 1, 5, and 10%, respectively.

Table 9. Growth regressions with lightning and trade & integration controls

Dependent variable:	Average annual growth in GSP per worker over periods of 10 years (1977-1987, 1987-1997, 1997-2007)				
INTEGRATION:		Agricultural exports per capita	FDI per capita	Agricultural exports per capita	FDI per capita
	(1)	(2)	(3)	(4)	(5)
(log, initial) Real GSP per worker	-1.80***	-1.81***	-1.78***	-1.82***	-1.79***
	[0.41]	[0.38]	[0.51]	[0.38]	[0.50]
(log) Lightning × $t_{77\text{-}87}$	-0.12			-0.023	-0.12
	[0.11]			[0.11]	[0.11]
(log) Lightning × $t_{87\text{-}97}$	-0.12			-0.082	-0.12
	[0.08]			[0.088]	[0.085]
(log) Lightning × $t_{97\text{-}07}$	-0.21**			-0.24***	-0.21***
	[0.08]			[0.067]	[0.077]
(log) INTEGRATION × $t_{77\text{-}87}$		-0.13**	0.023	-0.13**	0.034
		[0.048]	[0.15]	[0.051]	[0.15]
(log) INTEGRATION × $t_{87\text{-}97}$		-0.094	0.11	-0.08	0.11
		[0.057]	[0.18]	[0.063]	[0.17]
(log) INTEGRATION × $t_{97\text{-}07}$		0.065	-0.17	0.10**	-0.19
		[0.040]	[0.13]	[0.039]	[0.13]
Observations	144	144	144	144	144
R-squared	0.44	0.46	0.41	0.49	0.45
Human capital controls (enrollment, high school or higher, BA)	Yes	Yes	Yes	Yes	Yes

Notes. Pooled OLS estimates. The dependent variable is the annual growth rate of GSP per worker over the periods 1977-1987, 1987-1997 and 1997-2007. All regressions include a constant and a full set of time-dummies. Lightning is the average number of flashes per year per square km, measured by flash-detectors. The controls for human capital are the initial enrollment rate, percentage of population with a high school or higher degree, and percentage of population with a BA degree. All the human capital controls are measured at the beginning of each 10-year period (1977, 1987 and 1997), except for enrollment rates (measured in 1980 instead of 1977) and the % of population with a highschool degree or higher (measured in 1980, 1990 and 2000 instead of 1977, 1987 and 1997), due to data availability. Robust standard errors in brackets, adjusted for clustering at state level. Asterisks ***, **, and * indicate significance at the 1, 5, and 10%, respectively.

Table 10. Lightning and IT diffusion

Dependent variable:	% of households with a personal computer at home, 2003				% of households with Internet access at home, 2003				Manufacturing firms' IT investments, 2007			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(log) Lightning	-3.68***	-3.68***	-1.15***	-2.66***	-3.57***	-3.57***	-1.14***	-1.69*	-1.06***	-1.06***	-0.99***	-1.69**
	[0.56]	[0.58]	[0.38]	[0.80]	[0.61]	[0.62]	[0.41]	[0.88]	[0.23]	[0.24]	[0.26]	[0.64]
(log) Real GSP per worker, 1991		5.57*	1.96	4.64		9.95***	4.83	5.32		1.39	-3.31	-4.05*
		[3.26]	[2.85]	[3.19]		[3.46]	[2.97]	[3.70]		[1.92]	[2.00]	[2.37]
Enrollment rate, 1991			0.066	0.037			-0.0054	0.05			-0.20***	-0.17
			[0.097]	[0.20]			[0.092]	[0.22]			[0.066]	[0.14]
High school degree or higher, 1990			0.60***	0.76***			0.53***	0.77***			-0.091	0.014
			[0.11]	[0.15]			[0.11]	[0.19]			[0.058]	[0.12]
Bachelor's degree or higher, 1991			0.35**	0.30**			0.45***	0.39**			0.32***	0.29**
			[0.15]	[0.13]			[0.15]	[0.17]			[0.073]	[0.14]
Observations	48	48	48	48	48	48	48	48	48	48	48	48
R-squared	0.43	0.45	0.79	0.84	0.38	0.45	0.77	0.82	0.24	0.25	0.50	0.63
Regional fixed effects												
(8 BEA economic areas)	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
H_0: Regional FEs = 0 (p value)	.	.	.	0.02	.	.	.	0.28	.	.	.	0.09

Notes. OLS estimates. The dependent variables are (a) the percentage of household with access to a personal computer at home in 2003, (b) the % of households with access to Internet at home, and (c) the level of IT investments in the manufacturing sector, measured as the amount capital expenditures on computers and peripheral data processing equipment, relative to all non-construction capital expenditures, respectively. Lightning is the average number of flashes per year per square km, measured by flash-detectors. The rest of the covariates are described in the Data Appendix. Robust standard errors in brackets. Asterisks ***, **, and * indicate significance at the 1, 5, and 10%, respectively.

Table 11. Lightning and IT diffusion - Additional controls

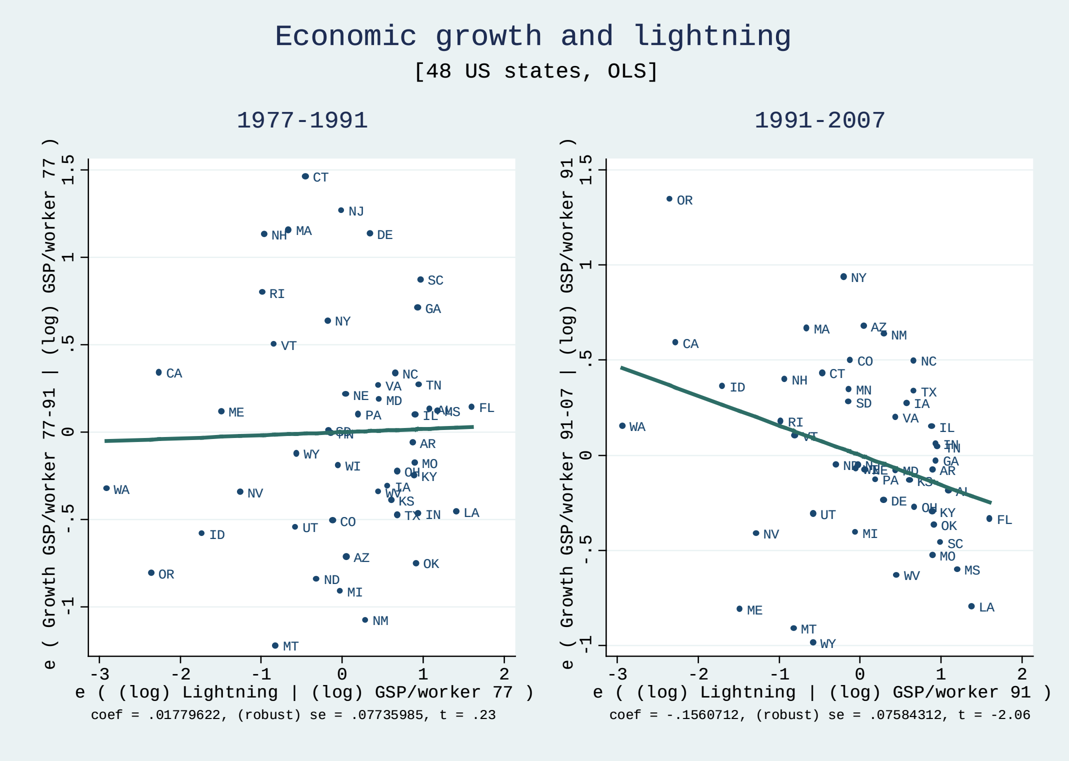
		Economy structure			Trade & Integration			Institutions			Religion	Race & ethnicity			Urbanization	Age structure	
ADDITIONAL CONTROL:		Share of agriculture in GSP, 1991	Share of government in GSP, 1991	Share of manufacturing in GSP, 1991	(log) FDI per capita, 1991	(log) Agricultural exports per capita, 1991	(log) Population, 1991	Soldier mortality, 1829-1854	% of workforce in mining, 1880	% of slavery, 1860	% population attending a church or a sinagogue almost every week, av. 2004-2006	% white population, 1990	% black population, 1990	% Hispanic origin population, 1990	% urban population, 1990	% population 15 years or less, 1990	% population 15-64 years, 1990
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
Dependent variable:	**Panel A: % of households with a personal computer at home, 2003**																
(log) Lightning	-2.66***	-3.16***	-2.71***	-2.58***	-2.58***	-2.68***	-2.75***	-3.78***	-2.63***	-2.57***	-2.63***	-2.20**	-2.12**	-2.66***	-3.01***	-2.62***	-2.69***
	[0.80]	[0.92]	[0.84]	[0.84]	[0.79]	[0.82]	[0.80]	[0.80]	[0.77]	[0.86]	[0.84]	[0.89]	[0.91]	[0.83]	[0.83]	[0.87]	[0.86]
ADDITIONAL CONTROL:		-0.35	-0.046	0.13*	-1.07	-0.22	0.83	1.76**	0.03	-0.015	0.073	0.18**	-0.13	-0.036	0.063	0.10	0.10
		[0.27]	[0.20]	[0.075]	[1.35]	[0.51]	[0.52]	[0.65]	[0.060]	[0.054]	[0.27]	[0.074]	[0.082]	[0.090]	[0.052]	[0.39]	[0.44]
R-squared	0.84	0.85	0.84	0.85	0.84	0.84	0.85	0.86	0.84	0.84	0.84	0.86	0.85	0.84	0.85	0.84	0.84
Regional fixed effects (8 BEA economic regions)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
H_0: Regional FEs = 0 (p value)	0.02	0.04	0.04	0.02	0.02	0.02	0.01	0.001	0.04	0.03	0.11	0.02	0.01	0.10	0.03	0.02	0.03
Dependent variable:	**Panel B: % of households with Internet access at home, 2003**																
(log) Lightning	-1.69*	-1.15***	-1.07**	-1.19**	-1.08**	-0.99**	-1.18***	-1.17***	-1.27***	-1.27***	-1.34**	-0.85*	-0.83*	-1.21**	-1.12**	-0.98**	-0.99**
	[0.88]	[0.41]	[0.40]	[0.46]	[0.41]	[0.44]	[0.41]	[0.42]	[0.45]	[0.44]	[0.63]	[0.46]	[0.48]	[0.49]	[0.43]	[0.37]	[0.41]
ADDITIONAL CONTROL:		-0.14	-0.18	0.12	-1.46	-0.36	0.40	0.16	-0.05	0.025	0.15	0.21***	-0.11	-0.14*	-0.031	-0.49	0.41
		[0.19]	[0.23]	[0.084]	[1.64]	[0.37]	[0.39]	[0.53]	[0.043]	[0.050]	[0.29]	[0.060]	[0.080]	[0.070]	[0.039]	[0.34]	[0.33]
R-squared	0.82	0.77	0.78	0.78	0.78	0.78	0.77	0.77	0.78	0.77	0.77	0.83	0.78	0.80	0.77	0.79	0.78
Regional fixed effects (8 BEA economic regions)	Yes	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
H_0: Regional FEs = 0 (p value)	0.28	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Dependent variable:	**Panel C: IT investments of manufacturing firms, 2007**																
(log) Lightning	-1.69**	-1.66**	-1.98***	-1.69**	-1.77**	-1.71**	-1.64**	-1.36*	-1.63**	-1.32*	-1.87***	-1.53**	-1.24*	-1.67***	-1.72**	-1.79***	-1.69**
	[0.64]	[0.74]	[0.55]	[0.64]	[0.65]	[0.63]	[0.68]	[0.71]	[0.60]	[0.66]	[0.57]	[0.66]	[0.64]	[0.56]	[0.64]	[0.59]	[0.66]
ADDITIONAL CONTROL:		0.021	-0.28**	-0.011	1.06	-0.21	-0.37	-0.52	0.048*	-0.059***	-0.43***	0.06	-0.10**	0.094**	0.006	-0.27**	0.028
		[0.20]	[0.13]	[0.055]	[0.79]	[0.31]	[0.43]	[0.42]	[0.025]	[0.020]	[0.14]	[0.037]	[0.042]	[0.042]	[0.032]	[0.11]	[0.21]
R-squared	0.63	0.63	0.68	0.63	0.65	0.63	0.64	0.64	0.64	0.67	0.70	0.64	0.68	0.67	0.63	0.66	0.63
Regional fixed effects (8 BEA economic regions)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
H_0: Regional FEs = 0 (p value)	0.09	0.10	0.00	0.09	0.05	0.09	0.04	0.06	0.04	0.01	0.01	0.05	0.03	0.09	0.10	0.10	0.10

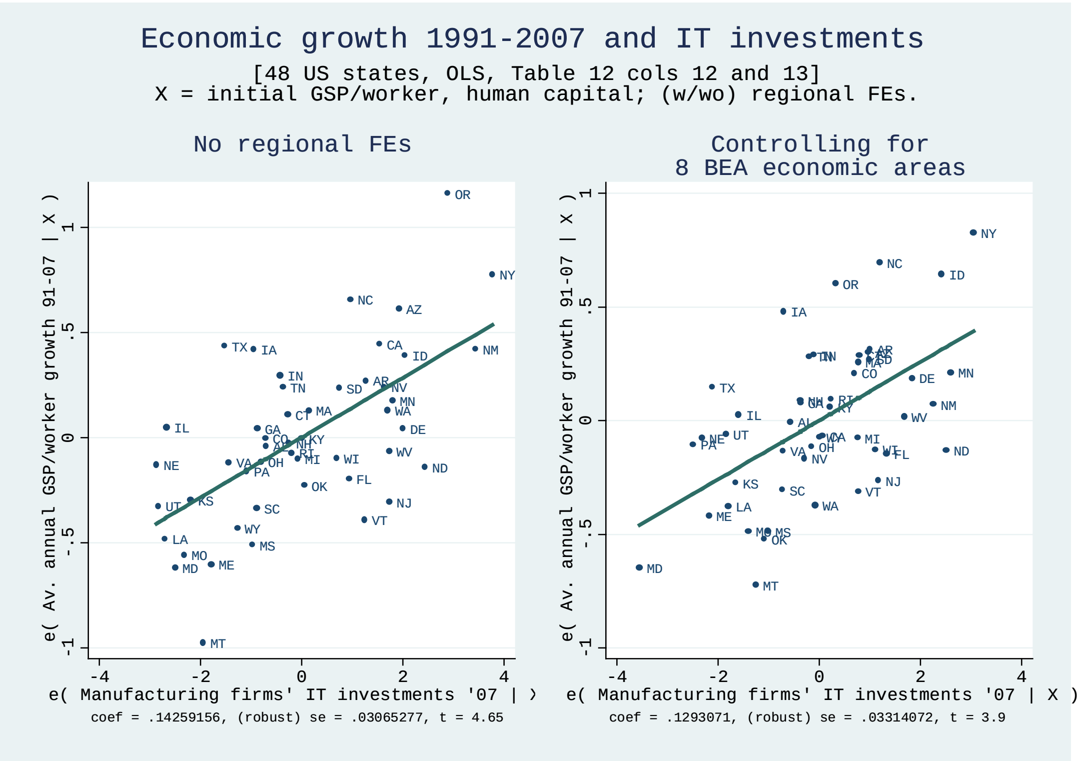
Notes. OLS estimates. All regressions have 48 observations, include a constant term, and control for the levels of (log) real Gross State Product per worker, enrollment rates, and % of population with a high school or higher degree in 1991, and the % of population with a BA degree in 1990. The dependent variables are (Panel A) the percentage of household with access to a personal computer at home in 2003, (Panel B) the % of households with access to Internet at home, and (Panel C) the level of IT investments in the manufacturing sector, measured as the amount capital expenditures on computers and peripheral data processing equipment, relative to all non-construction capital expenditures, respectively. Lightning is the average number of flashes per year per square km, measured by flash-detectors. Robust standard errors in brackets. Asterisks ***, **, and * indicate significance at the 1, 5, and 10%, respectively.

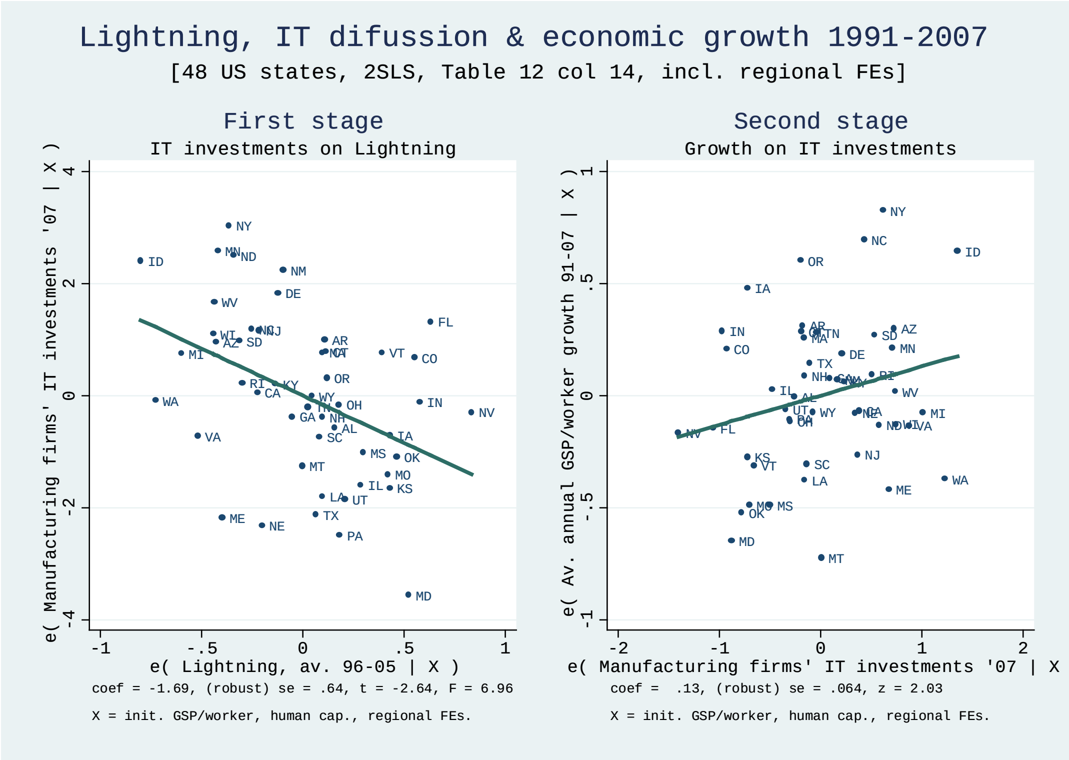
Table 12. Lightning, IT diffusion, and economic growth

Dependent variable:	Average annual growth in GSP per worker, 1991-2007														
	OLS													IV	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
(log) Real GSP per worker, 1991	-0.66	-0.82*	-0.99**	-0.89**	-0.76*	-0.92**	-0.90**	-0.99**	-1.29***	-1.35***	-1.30***	-1.23***	-1.26***	-1.23**	-1.25***
	[0.41]	[0.44]	[0.45]	[0.35]	[0.42]	[0.43]	[0.35]	[0.43]	[0.42]	[0.46]	[0.44]	[0.44]	[0.36]	[0.54]	[0.33]
(log) Lightning	-0.16**				-0.093	-0.064	0.024	0.066	0.029	0.15					
	[0.076]				[0.10]	[0.092]	[0.057]	[0.081]	[0.086]	[0.15]					
Computer at home, 2003		0.028**			0.017			0.0063	0.021	0.051	0.029				
		[0.012]			[0.016]			[0.036]	[0.037]	[0.039]	[0.036]				
Access to Internet at home, 2003			0.033***			0.026*		0.0064	0.0019	-0.0095	0.0051				
			[0.011]			[0.014]		[0.036]	[0.036]	[0.038]	[0.037]				
Manufacturing firms' IT investments, 2007				0.16***			0.17***	0.17***	0.15***	0.15***	0.13***	0.13***	0.14***	0.13**	0.15**
				[0.025]			[0.027]	[0.029]	[0.034]	[0.038]	[0.034]	[0.033]	[0.031]	[0.064]	[0.062]
Observations	48	48	48	48	48	48	48	48	48	48	48	48	48	48	48
R-squared	0.15	0.15	0.19	0.57	0.17	0.20	0.57	0.58	0.62	0.71	0.70	0.67	0.60	.	.
Instrumented variable	.	.	.	.	.	.	.	.	.	.	.	.	.	Manufacturing firms' IT investments, 2007	
Instrument	.	.	.	.	.	.	.	.	.	.	.	.	.	(log) Lightning	
1st stage F statistic (Kleibergen-Paap Wald F stat)	.	.	.	.	.	.	.	.	.	.	.	.	.	6.96	14.38
Weak-instrument robust inference: Stock-Wright LM S statistic (p val)	.	.	.	.	.	.	.	.	.	.	.	.	.	0.061	0.067
Human capital controls (enrollment, high school or higher, BA)	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional fixed effects (8 BEA economic regions)	No	No	No	No	No	No	No	No	No	Yes	Yes	Yes	No	Yes	No
H_0: Regional FEs = 0 (p value)	.	.	.	.	.	.	.	.	.	0.24	0.24	0.32	.	0.12	.

Notes. OLS and IV estimates. The dependent variable in all regressions is the average annual growth rate in GSP per worker over the period 1991-2007. Columns 14 and 15 report 2SLS regressions, where manufacturing firms' IT investments are instrumented by (log) lightning density. Access to a computer and Internet at home are measured as the % of households by state. IT investments at manufacturing firms are measured as the % of capital expenditures on computers and data processing equipment, relative to all non-construction capital expenditures for all firms in the manufacturing sector. Lightning is the average number of flashes per year per square km, measured by flash-detectors. All regressions include a constant. Robust standard errors in brackets. ***, **, and * indicate significance at the 1, 5, and 10%, respectively.


Economic growth and lightning
[48 US states, OLS]
1977-1991
1991-2007
e (Growth GSP/worker 77-91 | (log) GSP/worker 77)
e ((log) Lightning | (log) GSP/worker 77)
coef = .01779622, (robust) se = .07735985, t = .23
e (Growth GSP/worker 91-07 | (log) GSP/worker 91)
e ((log) Lightning | (log) GSP/worker 91)
coef = -.1560712, (robust) se = .07584312, t = -2.06


Economic growth 1991-2007 and IT investments
[48 US states, OLS, Table 12 cols 12 and 13]
X = initial GSP/worker, human capital; (w/wo) regional FEs.
No regional FEs
Controlling for
8 BEA economic areas
e(Av. annual GSP/worker growth 91-07 | X)
e(Manufacturing firms' IT investments '07 |)
coef = .14259156, (robust) se = .03065277, t = 4.65
e(Av. annual GSP/worker growth 91-07 | X)
e(Manufacturing firms' IT investments '07 | X)
coef = .1293071, (robust) se = .03314072, t = 3.9


Lightning, IT difussion & economic growth 1991-2007
[48 US states, 2SLS, Table 12 col 14, incl. regional FEs]
First stage
IT investments on Lightning
e(Manufacturing firms' IT investments '07 | X)
e(Lightning, av. 96-05 | X)
coef = -1.69, (robust) se = .64, t = -2.64, F = 6.96
X = init. GSP/worker, human cap., regional FEs.
Second stage
Growth on IT investments
e(Av. annual GSP/worker growth 91-07 | X)
e(Manufacturing firms' IT investments '07 | X)
coef = .13, (robust) se = .064, z = 2.03
X = init. GSP/worker, human cap., regional FEs.

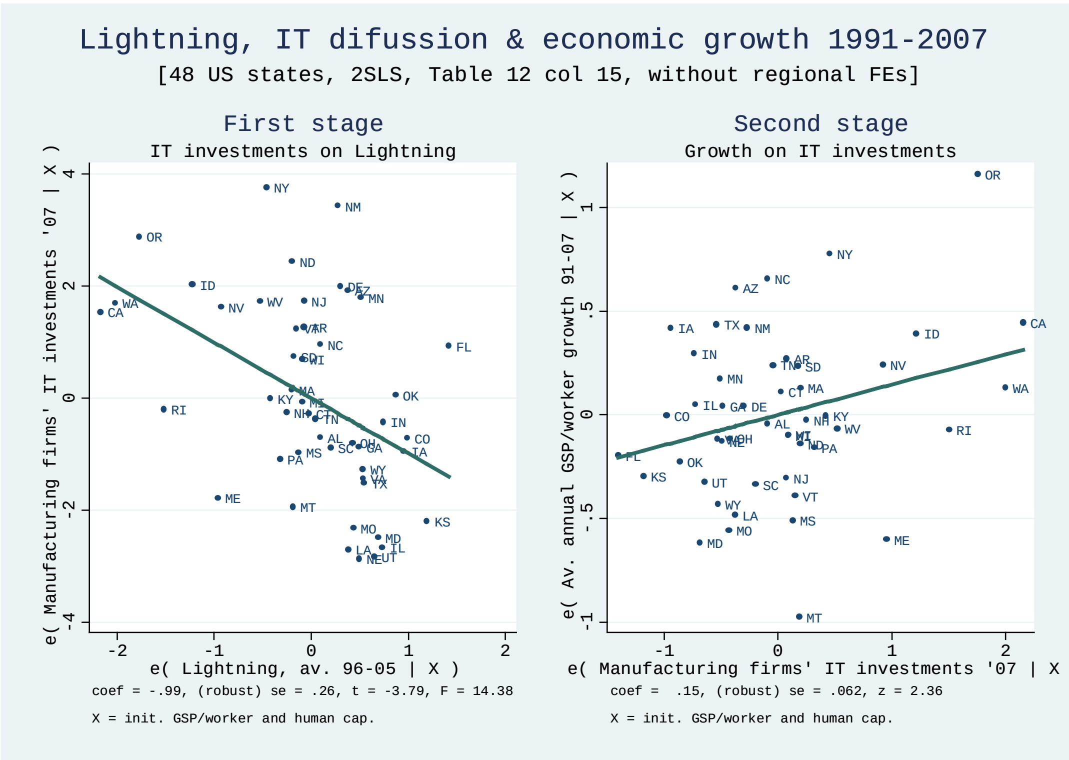

Lightning, IT difussion & economic growth 1991-2007
[48 US states, 2SLS, Table 12 col 15, without regional FEs]
First stage
IT investments on Lightning
e(Manufacturing firms' IT investments '07 | X)
e(Lightning, av. 96-05 | X)
coef = -.99, (robust) se = .26, t = -3.79, F = 14.38
X = init. GSP/worker and human cap.
Second stage
Growth on IT investments
e(Av. annual GSP/worker growth 91-07 | X)
e(Manufacturing firms' IT investments '07 | X)
coef = .15, (robust) se = .062, z = 2.36
X = init. GSP/worker and human cap.

Table A1. Tests for whether lightning is a constant plus white noise

	Breusch-Godfrey test			Runs test	
	test-statistic	p-value	N of lags*	test-statistic	p-value
Aggregate US	0.02	0.88	1	0.46	0.65
Alabama	0.61	0.43	1	-0.22	0.82
Arizona	0.16	0.69	1	-0.13	0.90
Arkansas	0.16	0.69	1	1.67	0.09
California	0.48	0.49	1	-0.12	0.91
Colorado	0.12	0.73	1	0.25	0.80
Florida	0.02	0.90	1	-0.70	0.49
Georgia	0.00	0.95	1	0.25	0.80
Idaho	0.02	0.90	1	0.72	0.47
Illinois	0.20	0.65	1	-1.64	0.10
Indiana	1.67	0.20	1	-0.22	0.82
Iowa	0.20	0.66	1	-0.22	0.82
Kansas	0.58	0.44	1	0.84	0.40
Kentucky	0.24	0.62	1	0.25	0.80
Louisiana	0.06	0.81	1	-0.70	0.49
Maine	1.05	0.31	1	0.25	0.80
Maryland	0.01	0.94	1	0.25	0.80
Massachusetts	1.29	0.26	1	0.72	0.47
Michigan	0.33	0.56	1	-0.70	0.49
Minnesota	0.00	0.98	1	-1.64	0.10
Mississippi	0.98	0.32	1	-2.12	0.03
Missouri	0.19	0.66	1	0.36	0.72
Montana	0.71	0.40	1	-2.12	0.03
Nebraska	0.22	0.64	1	-0.70	0.49
Nevada	0.02	0.88	1	0.72	0.47
New Mexico	1.25	0.26	1	-0.22	0.82
New York	7.52	0.02	2	0.36	0.72
North Carolina	0.74	0.39	1	-1.45	0.15
North Dakota	5.30	0.07	2	-0.22	0.82
Ohio	0.03	0.85	1	-0.70	0.49
Oklahoma	2.97	0.09	1	-1.64	0.10
Oregon	0.64	0.42	1	-1.45	0.15
Pennsylvania	5.25	0.07	2	0.72	0.47
South Carolina	0.23	0.63	1	-0.22	0.82
South Dakota	2.93	0.09	1	1.33	0.18
Tennessee	0.22	0.64	1	-0.22	0.82
Texas	3.79	0.05	1	-0.22	0.82
Utah	4.54	0.03	1	-0.70	0.49
Virginia	4.68	0.03	1	-0.22	0.82
Washington	0.48	0.49	1	-0.61	0.54
West Virginia	4.56	0.03	1	0.72	0.47
Wisconsin	0.57	0.45	1	-1.17	0.24
Wyoming	0.09	0.77	1	-0.22	0.82

Notes. The residuals are obtained from regressing lightning on a constant for each of the 42 states over the period 1977-1995. H_0: Residuals are not serially correlated. Lightning is average number of flashes per year per square km, measured at weather stations.
*: Number of lags selected by Schwarz's information criteria.

Table A2. Correlations: Lightning and US Bureau of Economic Analysis' 8 Economic Areas

Dependent variable:	Lightning		(log) Lightning	
	(1)	(2)	(3)	(4)
Far West	0.29***	(dropped)	-1.43***	(dropped)
	[0.10]	.	[0.32]	.
Great lakes	3.85***	3.56***	1.26***	2.69***
	[0.74]	[0.75]	[0.22]	[0.39]
Mid East	2.57***	2.29***	0.92***	2.35***
	[0.28]	[0.30]	[0.11]	[0.34]
New England	0.94***	0.65***	-0.11	1.32***
	[0.12]	[0.16]	[0.14]	[0.35]
Plains	2.99***	2.71***	1.00***	2.43***
	[0.54]	[0.55]	[0.18]	[0.37]
Rocky Mountain	1.15***	0.86***	0.021	1.45***
	[0.24]	[0.26]	[0.26]	[0.41]
Southeast	6.00***	5.71***	1.74***	3.17***
	[0.64]	[0.65]	[0.10]	[0.34]
Southwest	3.72***	3.43***	1.26***	2.69***
	[0.66]	[0.67]	[0.18]	[0.37]
Constant	.	0.29***	.	-1.43***
	.	[0.10]	.	[0.32]
Observations	48	48	48	48
R-squared	.	0.69	.	0.84

Notes: OLS regressions. Lightning is the average (1996-2005) number of flashes per year per square km, measured by flash-detectors. 8 economic areas defined by the US BEA. Robust standard errors in brackets. ***, **, and * indicate significance at the 1, 5, and 10%, respectively.

Table A3. Growth, lightning, and regional fixed effects (BEA economic areas) - controlling for human capital

Dependent variable:	Average annual growth in GSP per worker over periods of 10 years (1977 - 1987, 1987 - 1997, 1997 - 2007)							
BEA economic area:	Far West	Great Lakes	Mid East	New England	Plains	Rocky Mountain	Southeast	Southwest
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(log, initial) Real GSP per worker	-1.85***	-1.80***	-1.86***	-1.71***	-1.78***	-1.82***	-1.80***	-1.80***
	[0.41]	[0.43]	[0.44]	[0.42]	[0.43]	[0.41]	[0.41]	[0.41]
(log) Lightning × $t_{77\text{-}87}$	-0.098	-0.12	-0.12	-0.10	-0.13	-0.11	-0.12	-0.11
	[0.13]	[0.11]	[0.11]	[0.11]	[0.11]	[0.12]	[0.11]	[0.11]
(log) Lightning × $t_{87\text{-}97}$	-0.092	-0.12	-0.12	-0.098	-0.12	-0.11	-0.12	-0.11
	[0.099]	[0.083]	[0.083]	[0.084]	[0.084]	[0.085]	[0.085]	[0.085]
(log) Lightning × $t_{97\text{-}07}$	-0.18*	-0.21**	-0.21**	-0.19**	-0.21***	-0.21***	-0.22***	-0.21**
	[0.10]	[0.080]	[0.079]	[0.084]	[0.078]	[0.078]	[0.080]	[0.079]
BEA economic area	0.11	-0.0093	0.09	0.16	0.026	-0.33	0.054	-0.091
	[0.20]	[0.14]	[0.24]	[0.17]	[0.17]	[0.22]	[0.15]	[0.099]
Observations	144	144	144	144	144	144	144	144
R-squared	0.44	0.44	0.44	0.45	0.44	0.46	0.44	0.44
Human capital controls (enrollment, high school or higher, BA)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes. Pooled OLS estimates. The dependent variable is the yearly growth rate of GSP per worker over the periods 1977-1987, 1987-1997, and 1997-2007. Lightning is the average number of flashes per year per square km, measured by flash-detectors. The different proxies for human capital (enrollment rates, % of population with high school or higher, or BA degree) are described in the appendix, and measured at the beginning of each 10-year period (1977, 1987 and 1997), except for enrollment rates (measured in 1980 instead of 1977 for the first period) and the % of population with a highschool degree or higher (measured in 1980, 1990 and 2000 instead of 1977, 1987 and 1997 for each respective period), due to data availability. The set of region fixed effects accounts for the 8 US Bureau of Economic Analysis' economic areas. All regressions include a constant and a full set of time-dummies. Robust standard errors in brackets, adjusted for clustering at the state level. Asterisks ***, **, and * indicate significance at the 1, 5, and 10%, respectively.

Table A4. Correlations: Lightning and geographic/climate variables

Dependent variable:	(log) Lightning											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(log) Temperature (C degrees)	1.20***										0.091	0.056
	[0.17]										[0.36]	[0.39]
(log) Precipitation (cm/year)		0.084									-0.24	-0.22
		[0.49]									[0.30]	[0.27]
(log) Tornado intensity (av EF-scale)			0.088								0.78**	0.76**
			[0.33]								[0.32]	[0.28]
(log) Hail size (cm)				2.30*							1.68	1.54
				[1.24]							[1.16]	[0.99]
(log) Wind speed (km/h)					-0.077						0.046	0.036
					[0.31]						[0.20]	[0.17]
(log) Humidity (% moisture in air)						-0.32					-1.39	-1.37
						[1.25]					[1.49]	[1.42]
(log) Cloudiness (days/year)							-1.04**				0.68	0.88
							[0.47]				[1.13]	[0.73]
(log) Sunshine (days/year)								1.39***			-0.31	
								[0.49]			[0.98]	
(log) Elevation (m above sea level)									-0.20**		-0.097	-0.096
									[0.078]		[0.11]	[0.074]
(log) Latitude (degrees)										-4.30***	-5.40***	-5.47***
										[0.61]	[1.19]	[1.12]
Observations	48	48	48	48	48	48	48	47	48	48	47	48
R-squared	0.90	0.84	0.84	0.85	0.84	0.84	0.87	0.87	0.86	0.93	0.96	0.96
Regional fixed effects												
(8 BEA economic regions)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
H_0: Regional FEs = 0 (p value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Notes. OLS regressions. All regressions include a constant and control for the 8 BEA economic regions (Far West excluded). Robust standard errors in brackets. Lightning is the average number of flashes per year per square km, measured by flash-detectors. Temperature, precipitation, tornado intensity, hail size and wind speed are averages over the period 1997-2007. Humidity, cloudiness and sunshine are state averages through 2007, as reported by the US National Oceanic and Atmospheric Administration (NOAA). Data sources and definitions for all variables are provided in the Data appendix. Data for all variables are available for the 48 contiguous US states, except sunshine, which has missing data for Delaware. ***, **, and * indicate significance at the 1, 5, and 10%, respectively.

Table A5. Correlations: Lightning, historical and trade variables

Dependent variable:	(log) Lightning												
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Controls for history:													
% of workforce in mining, 1880	-0.56												-0.52
	[1.36]												[1.32]
Average no. of cooling degree days		0.030*											0.025
		[0.017]											[0.018]
% of 1860 population in slavery			1.40***										0.52
			[0.36]										[1.43]
Access to navigable water				0.038									-0.17
				[0.30]									[0.26]
% of 1860 population on large slave plantations					1.76***								-0.035
					[0.53]								[1.84]
Settler origin: English						-0.28							0.012
						[0.18]							[0.18]
Settler origin: French							0.17						0.26
							[0.16]						[0.17]
Settler origin: Spanish								0.24					-0.027
								[0.21]					[0.19]
Settler origin: Dutch									-0.14				-0.17
									[0.18]				[0.26]
Average annual soldier mortality in 1829-1838, 1839-1854, %										29.2***			20.3**
										[5.17]			[9.19]
Controls for trade:													
(log) Agricultural exports per capita											-0.11		-0.14
											[0.097]		[0.093]
(log) FDI per capita												0.27*	0.083
												[0.16]	[0.16]
Observations	48	48	48	48	48	48	48	48	48	48	48	48	48
R-squared	0.84	0.87	0.86	0.84	0.85	0.85	0.85	0.85	0.84	0.88	0.85	0.85	0.91
Regional fixed effects													
(8 BEA economic regions)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
H_0: Regional FEs = 0 (p value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Notes. OLS regressions. All regressions include a constant and control for the 8 BEA economic regions (Far West excluded). Robust standard errors in brackets. Lightning is the average number of flashes per year per square km, measured by flash-detectors. Historical variables taken from Mitchener and McLean (2004). Agricultural exports and FDI per capita are averages taken over the periods 1997-2007. Sources and definitions are provided in the Data appendix. ***, **, and * indicate significance at the 1, 5, and 10%, respectively.

Table A6. Correlations: Lightning, initial GSP per worker, and human capital

Dependent variable:	(log) Lightning				
	(1)	(2)	(3)	(4)	(5)
(log) Real GSP per worker, 1991	1.06**				1.11**
	[0.45]				[0.53]
Enrollment rate, 1991		-0.025			-0.019
		[0.030]			[0.028]
High school degree or higher, 1990			-0.0089		-0.014
			[0.026]		[0.030]
Bachelor's degree or higher, 1991				0.0096	-0.0032
				[0.030]	[0.036]
Observations	48	48	48	48	48
R-squared	0.86	0.85	0.84	0.84	0.86
Regional fixed effects					
(8 BEA economic regions)	Yes	Yes	Yes	Yes	Yes
H_0: Regional FEs = 0 (p value)	0.00	0.00	0.00	0.00	0.00

Notes. OLS regressions. All regressions include a constant and control for the 8 BEA economic regions (Far West excluded). Robust standard errors in brackets. Lightning is the average number of flashes per year per square km, measured by flash-detectors. Sources and definitions for the human capital variables are provided in the Data appendix. ***, **, and * indicate significance at the 1, 5, and 10%, respectively.

Table A7. Correlations: Lightning and Caselli and Coleman's (2001) additional detereminants of IT diffusion

Dependent variable:	(log) Lightning																
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
Share of agriculture in GSP, 1991	-0.11***																-0.17*
	[0.035]																[0.087]
Share of government in GSP, 1991		-0.058**															-0.039
		[0.026]															[0.034]
Share of manufacturing in GSP, 1991			-0.0058														-0.025
			[0.017]														[0.016]
(log) FDI per capita, 1991				0.20													-0.18
				[0.14]													[0.21]
(log) Agricultural exports per capita, 1991					-0.068												0.16
					[0.11]												[0.11]
(log) Population, 1991						0.078											-0.21
						[0.084]											[0.13]
Soldier mortality, 1829-1854							0.29***										0.19
							[0.052]										[0.12]
% of workforce in mining, 1880								-0.0056									-0.002
								[0.014]									[0.012]
% of slavery, 1860									0.014***								0.0026
									[0.0036]								[0.012]
% population attending a church or a sinagogue almost every week, av. 2004-2006										-2.21							1.01
										[4.49]							[4.22]
% white population, 1990											-1.79**						1.00
											[0.67]						[4.55]
% black population, 1990												2.76***					2.69
												[0.78]					[6.13]
% Hispanic origin population, 1990													0.66				0.50
													[1.28]				[2.33]
% urban population, 1990														1.27**			0.41
														[0.48]			[0.66]
% population 15 years or less, 1990															-4.03		-9.93*
															[4.73]		[5.50]
% population 15-64 years, 1990																4.71	-4.27
																[5.02]	[4.85]
Observations	48	48	48	48	48	48	48	48	48	48	48	48	48	48	48	48	48
R-squared	0.87	0.86	0.84	0.85	0.85	0.85	0.88	0.84	0.86	0.84	0.86	0.87	0.84	0.87	0.85	0.85	0.94
Regional fixed effects																	
(8 BEA economic regions)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
H_0: Regional FEs = 0 (p value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Notes. OLS regressions. All regressions include a constant and control for the 8 BEA economic regions (Far West excluded). Robust standard errors in brackets. Lightning is the average number of flashes per year per square km, measured by flash-detectors. The set of additional determinants of IT diffusion is chosen following the relevant group of variables in Caselli and Coleman (2001). Sources and definitions for all the data are provided in the Data appendix. ***, **, and * indicate significance at the 1, 5, and 10%, respectively.

Chapter 2

Religious Orders and Growth through Cultural Change in Pre-Industrial England

Thomas Barnebeck Andersen, Jeanet Sinding Bentzen, Carl-Johan Dalgaard, and Paul Sharp

Religious Orders and Growth through Cultural Change in Pre-Industrial England*

Thomas Barnebeck Andersen† Jeanet Bentzen Carl-Johan Dalgaard
Paul Sharp‡

February 16, 2011

Abstract

We advance the hypothesis that cultural values such as high work ethic and thrift, "the Protestant ethic" according to Max Weber, may have been diffused long before the Reformation, thereby importantly affecting the pre-industrial growth record. The source of pre-Reformation Protestant ethic, according to the proposed theory, was the Catholic Order of Cistercians. Using county-level data for England we find empirically that the frequency of Cistercian monasteries influenced county-level comparative development until 1801; that is, long after the Dissolution of the Monasteries. The pre-industrial development of England may thus have been propelled by a process of growth through cultural change.

Keywords: Protestant ethic, Malthusian population dynamics, economic development

JEL Classification codes: N13; O11; Z12

*We thank Daron Acemoglu, Philippe Aghion, Quamrul Ashraf, Roland Benabou, Raquel Fernández, Oded Galor, Thomas Markussen, Omer Moav, Stelios Michalopoulos, Ola Olsson, Enrico Perotti, Karl Gunnar Persson, Christoffer Sonne-Smith, Holger Strulik, Fabrizio Zilibotti as well as seminar participants at the 2011 AEA meeting in Denver, DEGIT XV in Frankfurt, the 2010 NBER Summer Institute, Gothenburg University, University of London Royal Holloway, University of Oslo and University of Zürich for comments and suggestions. Errors are ours.

†Andersen is affiliated with the Department of Business and Economics, University of Southern Denmark, Campusvej 55, DK-5230 Odense M. Contact: Thomas Barnebeck Andersen (barnebeck@sam.sdu.dk).

‡Bentzen, Dalgaard, and Sharp are affiliated with the Department of Economics, University of Copenhagen, Øster Farimagsgade 5, building 26, DK-1353 Copenhagen, Denmark. Contact: Jeanet Bentzen (jeanet.bentzen@econ.ku.dk), Carl-Johan Dalgaard (carl.johan.dalgaard@econ.ku.dk), and Paul Sharp (paul.sharp@econ.ku.dk).

1 Introduction

In what is surely one of the most famous works in all of social science, Max Weber (1905) argued that the Protestant Reformation was instrumental in facilitating the rise of capitalism in Western Europe. More specifically, Weber argued that Protestantism, in contrast to Catholicism, commends the virtues of hard work and thrift. These values, which Weber famously refers to as the "Protestant ethic", laid the foundation for the eventual rise of modern capitalism. A noteworthy study by Becker and Woessmann (2009) suggests that Weber was right, albeit for the wrong reasons: Protestants did not prosper as a result of their work ethic, say; rather, they prospered because instruction in reading the Bible generated the human capital so crucial to economic prosperity. Using data for Prussia, where some regions converted to Protestantism while others remained Catholic, Becker and Woessmann document the strength of the human capital channel. In fact, Becker and Woessmann argue that the human capital mechanism can account for most of the difference in comparative development between the Protestant and Catholic regions of Prussia.[1]

Nevertheless, one does not have to reject the original Weber thesis in order to support the human capital story. Landes (1999) is a case in point. While acknowledging the human capital mechanism, he maintains the importance of the Protestant ethic, both on empirical grounds (Protestant merchants and manufacturers played a leading role in banking, industry and trade) and on theoretical grounds (the Reformation created a new kind of man: rational, ordered, diligent and productive). More formally, using the World Values Survey, McCleary and Barro (2006) make probable that Weber was in fact right in emphasizing a link between religion on the one hand and work ethic on the other.[2] Nevertheless, it is clearly a difficult task to separately identify the importance of the human capital mechanism and the traditional "Weber mechanism", as both have arguably been at work simultaneously in the wake of the Reformation.

The present paper offers an attempt to separate the impact of the Weber mechanism and the human capital mechanism. We document below that the cultural virtues emphasized by Weber had a pre-Reformation origin in the religious Order of the Cistercians; a Catholic order which spread across England during the 12th century. We hypothesize that the Cistercians had a long term impact on development by encouraging a greater appreciation of hard work and wealth accumulation in local populations. That is, we argue the Cistercians encouraged growth by instigating the kind of cultural change that Weber attributed to Protestantism. Using cross-county data for England for the period 1377-1801, we find strong empirical support for a growth enhancing impact of Cistercian presence. Since the Cistercians did not encourage human capital accumulation, our findings suggests that the original Weber thesis, stressing the importance of cultural values like hard work and thrift, holds considerable explanatory power with respect to the pre-industrial growth

[1] Cantoni (2009), however, finds no effects of Protestantism on city growth across the German Lands of the Holy Roman Empire prior to industrialization.

[2] Thrift is not significant in their specification.

record of England.

The Cistercian order, a Benedictine offshoot, was established in France in 1098 as a reformist movement with the aim of returning to the literal observance of the "Rule of St. Benedict". They rejected the developments the Benedictines had undergone and tried to reproduce life exactly as it had been in St. Benedict's time; in fact, they often ventured beyond it in austerity. Put differently, the salient feature in the reform was a return to hard manual labor and the restrain from consumption (Kieser 1987). This meant that within the walls of the Cistercian monasteries one would find cultural values similar to those which, promulgated by the Protestant Reformation centuries later, is thought to have assisted the rise of capitalism outside the monastic walls. Several scholars have noted that the simplicity of the Order's lifestyle and their pursuit of wealth were in fact early manifestations of "the Protestant ethic" (e.g., Baumol 1990, p. 906; Collins 1986, p. 54; Kieser 1987, p. 116); Weber (1958, p. 118-19) himself singled out the Cistercians as encompassing values with a clear antecedent to the Protestant ethic.

We hypothesize that the cultural values embedded in the Cistercian order diffused to the local populations. Hence we argue that virtues associated with the Protestant movement started to spread in England long before Martin Luther posted his theses on the door of the All Saints' Church in Wittenberg. Of course, the cultural influence from the Cistercians was not immediate. Initially, the Cistercians may only have "convinced" a (potentially very) small group of people to "adopt" their attitudes towards hard work and thrift. But in a Malthusian setting work ethic and thrift translates into economic success, and ultimately into reproductive success. To the extent that cultural values carry over from parent to offspring, a cumulative process of growth through cultural change can be envisioned. If the pervasiveness of Protestant-type cultural values increases, this will stimulate work effort, investments and technological change; in turn, this works to encourage population growth and, as a consequence of selection, cultural change.

We construct a model that illustrates this cumulative process. To fix ideas, we focus on how Cistercians may have influenced the attitude towards hard work and thereby macroeconomic development. Using the model, we demonstrate that an initially small group of dynasties featuring a relatively strong preference for work effort could plausibly have come to dominate the population within the span of 500 years. Moreover, we show that small differences in the initial rate of "conversion" to a high work ethic could result in considerable cross-county variation in cultural values in the course of centuries. Finally, we derive an estimable equation from the model.

In order to proxy the initial cultural influence from the Cistercians on local populations, we employ information on the historic location of English Cistercian abbeys at the county level. With this data in hand, we proceed to document that the intensity of Cistercian presence left an important imprint on comparative development across English counties until 1801; that is, long after the Dissolution of the Monasteries, which took place between 1536 and 1540. As we focus on the pre-industrial period, and in keeping with our theoretical model, we rely on population density as our measure of productivity; in doing so we follow the lead of, among others, Ashraf and Galor (2009). Specifically, we show that, conditional on relevant exogenous controls, English counties with a higher share of Cistercian monasteries (as a fraction of all religious houses) experienced faster population growth during the period 1377-1801.

We believe the most plausible explanation for this finding is that the Cistercians influenced local cultural values, which subsequently took hold in the population. These new values in turn stimulated growth through attendant changes in work effort, investment behavior and technological progress. While we cannot document a link between Cistercian presence and *pre-industrial* cultural values across England, we are able to present evidence that the historic share of Cistercian abbeys is strongly correlated with *contemporary* work ethic and thrift at the regional level in England, as measured by the World Values Survey.

Naturally, a priori there could be other viable explanations for the observed link between intensity of Cistercian presence and population growth over the period in question. Perhaps this particular religious order simply managed to locate in areas with high growth potential; perhaps they influenced growth via international trade; or maybe the observed association is best motivated by technological change or human capital accumulation. We address these alternative narratives below. But we are led to the conclusion that they are unable to account for the observed relationship between the intensity of Cistercian presence and county-level population growth.

Our analysis contributes to several strands of literature. By demonstrating an impact from religious orders on economic development, we contribute to a literature which examines the religion-prosperity nexus (e.g., Landes 1999; Barro and McCleary 2003; McCleary and Barro 2006; Cavalcanti, Parente and Zhao 2007; Becker and Woessmann 2009; Cantoni 2009). In addition, by documenting a long lasting impact from Cistercian monasteries, our work contributes to a recent literature which suggests that past events (treatments) can permanently affect economic outcomes if they influence norms of behavior and/or culture (e.g., Guiso, Sapienza, and Zingales 2008; Nunn and Wantchekon 2009; Tabellini 2010).

The closest precursor to the argument developed below is the work of Clark (2007), which also takes as point of departure that cultural attributes, such as a high work ethic, breed economic success, and ultimately reproductive success in a Malthusian setting. Clark's theory is based on endogenous factors: The rich became rich because of certain favorable traits (cultural or perhaps even genetic); their children inherited these traits, and because the rich had such staggering reproductive success, their offspring were forced to move downward in the social hierarchy, implying that the "positive" traits eventually spread to the entire population. In a similar vein, Doepke and Zilibotti (2008) develop a theory of endogenous preference formation whereby cultural virtues conducive to growth flourish in parts of the population and facilitate a growth take-off. In contrast to Clark, however, Doepke and Zilibotti argue that the new cultural values emerged in the middle class, and not among the initially rich. We differ from both Clark and Doepke and Zilibotti in emphasizing a *shock* to cultural values: the settlement of the Cistercians. This allows us to test our argument statistically. In contrast to Doepke and Zilibotti, but similarly to Clark, we emphasize the reproductive advantage of high work ethic dynasties in explaining the diffusion of cultural values. In practise, deliberate investments and differential fertility probably both contributed to the diffusion of work ethic and thrift.

Finally, our analysis is related to evolutionary growth theory, as pioneered by Galor and Moav (2002). Galor and Moav demonstrate how dynasties with greater preference for child quality, relative to child quantity, hold a selective advantage in a Malthusian setting and come to dominate the population. Moreover, the theory predicts that the positive selection of quality oriented individuals stimulated long-run economic development. Similarly, the theory advanced below predicts that the epoch of Malthusian stagnation involved selection of individuals with a high work ethic, thus importantly influencing comparative development in England during the pre-industrial era.

The rest of the paper is organized as follows: Section 2 presents the theory, including the formal model, while Section 3 contains the empirical analysis. Section 4 provides a conclusion.

2 Theory

This section develops a theory of how Cistercians may have left a lasting imprint on comparative development in England. The following subsection provides details on Cistercian monks. We discuss their values, how these values manifested themselves in terms of work effort, capital accumulation and technological change, and how their values may have spread to the local population.

We argue that a major reason why the presence of Cistercian monasteries is apparent in comparative development long after the Order's disbandment is that they instigated a process of cultural change. To clarify how this process may have played out, we develop a model of growth through cultural change in Section 2.2. The model elucidates how an initially modest "cultural shock" to a population cumulates over time in a Malthusian setting, ultimately leaving a significant imprint on the growth record. The model also allows us to gauge the speed of the process of cultural change.

2.1 Cistercian Values

The Cistercian order was founded in 1098 in France; the first Cistercian monastery in England was founded in 1128 (Cooke 1893; Donkin 1963). During the 12th Century the Order spread rapidly across England, cf. Figure 1. By the end of the 14th century the expansion of the Order had essentially ceased. Hence, from the perspective of our regression analysis below, which involve the time period from 1377 onwards, we can treat Cistercian settlements as predetermined.

There is little doubt that the Cistercians held beliefs which were later to be associated with the Protestant

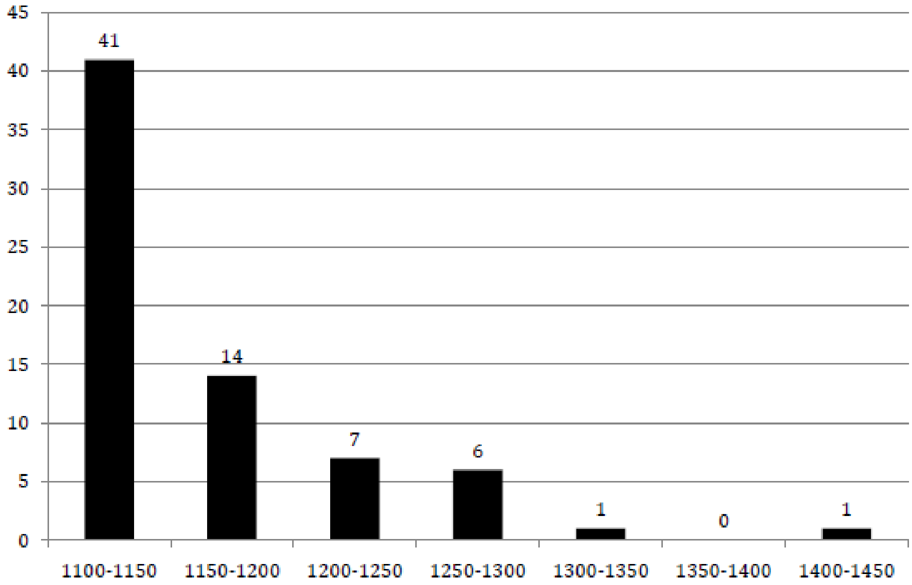


Figure 1: Frequency of founding years of Cistercian monasteries In England.

ethic. By seeking to return to a literal interpretation of the Rule of St. Benedict, the small book written in the sixth century by its namesake, they stressed the trinity of prayer, work and study, as well as the values of practicality, adaptability, simplicity and moderation (Hill 1968, p. 3). The Exordium Cistercii, written in the 1120s, and the statutes promulgated at the general chapter of 1134, stated that the monks were to work hard and live "*from the labour of their own hands, from cultivation and from their flocks*". They were also to live frugally, and were not permitted to have any possessions "*contrary to monastic purity*" such as parish churches, the tithes of other men's labour, dependent peasants, mills, ovens, or other income sources attached to the land. Hence, it is no surprise that Baumol (1990, p. 906) suggests that the monks of the Order of Cistercians may have embodied an earlier "Protestant ethic": "*Puritanical, at least in the earlier years, in their self-proclaimed adherence to simplicity in personal lifestyle while engaged in dedicated pursuit of wealth, they may perhaps represent an early manifestation of elements of 'the Protestant ethic'*". Collins (1986, p. 54) is slightly more direct when he notes that the Cistercians: "*had the Protestant ethic without Protestantism*".[3]

The simplicity of the Cistercians was thus only a liturgical simplicity, replacing long days of ritual with short prayers that could be said in pauses from labor (Bouchard 1991; Hill 1968). Moreover, "useless" labor, such as painting pictures, decorating books, breeding useless animals, etc. was banned (Kieser 1987). Some

[3] Kiefer (1987, p. 116) makes the same observation.

have suggested that they were attempting to reduce the need for manual labor in order to leave more time for prayer (Bloch 1935; Gimpel 1976; Ovitt 1986; Landes 1999). Whatever the case, from the very beginning the Cistercians were involved in the rapidly developing economic practices of the 12th century, and were in some cases initiators of these practices. Moreover, the monks' asceticism, by keeping down consumption, drove up levels of investment (Kiefer 1987; Baumol 1990).

Kaelber (1998) points out that Weber himself saw monastic asceticism as a clear precursor to ascetic Protestantism: the key driving force behind European capitalism according to Weber. More specifically, as argued by Weber (1958, p. 118-19): "*In the rules of St. Benedict, even more so in the case of the monks of Cluny and the Cistercians...[Christian asceticism] has become a systematically developed method of rational life conduct, with the goal to overcome the status naturae, to free man from the power of irrational impulses and his dependence on the world and on nature...It attempted to subject man under the supremacy of purposive will, to bring his action under constant self-control with a careful consideration of their ethical consequences. Thus it trained the monk, objectively, as a worker in the service of the Kingdom of God, and thereby further, subjectively, assured the salvation of his soul...[T]he end of this asceticism was to be able to lead an alert, intelligent life: the most urgent task the destruction of spontaneous, impulsive enjoyment, the most important means was to bring order into the conduct of its adherents. All these important points are emphasized in the rules of Catholic monasticism as strongly as in the principles of conduct of the Calvinists.*" Hence the idea that the Cistercians held values close to those promulgated by the Protestant Reformation has a long and distinguished tradition.[4]

The emphasis on hard work and thrift made the Cistercians entrepreneurial and ultimately very successful economically (Baumol 1990). They contributed much as agriculturists and as horse and cattle breeders. Their major contribution was the introduction of the grange system, whereby land was held in compact blocks, in contrast to the usual fragmented and unenclosed village holdings (Donkin 1963). Another contribution seems to have been advanced irrigation techniques, thus predating Rowland Vaughan's famous popularization of these methods by centuries.[5] Moreover, their high level of agricultural technology was matched by their industrial technology. Every monastery had a model factory, often as large as the church, with waterpower to drive the machinery (Gimpel 1976). This power was used for crushing wheat, sieving flour, fulling cloth and tanning (Baumol 1990). The Cistercians are known to have been skilled metallurgists (Gimpel 1976).

The Cistercian monastic system was one based on the principle of kinship, and thus Cistercian work practices and technology seem to have spread easily from house to house (Donkin 1978). These values

[4] As Weber points out, similar values were found among the Cluniacs. The impact of the Cluny order has received scant attention in the literature in comparison with the Cistercians. Yet, as we shall see, they too seem to have left a mark (albeit not as statistically robust as the Cistercians) on pre-industrial growth in England, conceivably for the same reasons the Cistercians influenced growth.

[5] Vaughan's Golden Valley was actually located in an area where the Cistercians had held extensive estates prior to the Dissolution (Cook, Stearne and Williamson 2003).

in turn spread into the local area partly due to the Cistercian practice of incorporating illiterate peasant lay brothers (known as conversi) for agricultural labor (Berman 2000). Lay brothers were bound by vows of chastity and obedience to their abbot, but were otherwise permitted to follow a less demanding form of Cistercian life. Work on Cistercian granges were also carried out by various classes of secular laborers. These included servi (servants), mercenarii (hired laborers), familiares (workers with intermediate status between hired workmen and lay brothers) and donate or oblate (pious laymen exchanging work for support). The exact fraction of lay brothers to these other types of labor is difficult to determine, but the latter seem to have become increasingly important at the turn of the 13th century (Noell 2006). Another important group of settlers in the abbeys were the corrodians, who spent their years of retirement there. Moreover, settled communities, including shopkeepers, formed outside the monasteries (Williams 1970). In this manner, the ways of the Cistercians spread beyond the Order itself.

An ideal check of this would be to study the correlation between Cistercian presence and preindustrial ethical values, like work ethic and thrift. Unfortunately, data constraints prevent us from carrying out such a check. What we can do instead is to study the relationship between the intensity of historical Cistercian presence and the pervasiveness of (proxies for) contemporary Protestant ethic across England.

In order to quantify differences in work ethic across countries, McCleary and Barro (2006) use the fraction of World Values Survey (WVS) respondents who indicated that they think that valuing "hard work" is an important trait for children to learn at home. To measure thrift they calculate the frequency of respondents indicating that "thrift, saving money and things" is an important trait for children to learn at home. These variables are also available for the United Kingdom in the WVS 2005. Unfortunately, it is only possible to disaggregate down to the regional level.[6] The "intensity" of Cistercian presence in a geographical area is proxied as the number of Cistercian monasteries relative to the total number of religious houses in the area.

Figures 2 and 3 depict the correlation between the intensity of Cistercian presence and work ethic and thrift, respectively. As expected, the three variables are positively correlated; Cistercian presence has a correlation of 0.62 with "work ethic" and 0.42 with "thrift". With only eight observations, an OLS regression returns a statistically significant (at the five percent level) correlation between work ethic and Cistercian presence. Statistical significance is not attained (at conventional levels) in the context of thrift; but the positive association between Cistercian presence and fraction of regional respondents emphasizing thrift is visually discernible and positive. Taken together, this exercise provides some support of our hypothesis that the Cistercian "treatment" influenced cultural values across England.

[6] The data only have regional identifyers. We have also contacted the British Values Survey, and the same holds for this survey.

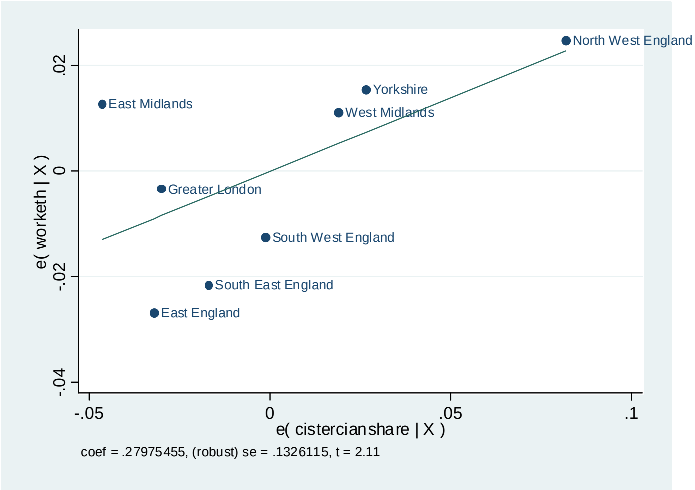


Figure 2: Cross regional correlation between Cistercian intensity and contemporary work ethic (source World Values Survey).

2.2 A Model of Growth through Cultural Change

In order to think more formally about how the ways of the Cistercians spread beyond the Order itself, and the ensuing macroeconomic impact, consider the following overlapping generations model for a closed economy in the process of development. Time is discrete and extends to infinity, $t = 0, 1, 2...\infty$.

People live for at most two periods: childhood and adulthood. During childhood individuals receive consumption from their unique parent; for simplicity of exposition, we assume that individuals only consume during period one.[7] If adulthood is reached, an individual decides on work effort and then reproduces. While each parent produces a fixed number of offspring, nutritional intake (deterministically) determines how many survive to adulthood. Hence, while fertility is exogenous, the number of surviving offspring (and thus population growth) is endogenous on account of the link between nutritional intake and the fraction of children that makes it into adulthood.

Cultural values are crudely represented by utility weights. Specifically, "high work ethic" dynasties are identified as dynasties that attach relatively low disutility to effort, and the impact from the Cistercians are conceptualized as a shock to the utility weights of a subset of the individuals. Moreover, we assume that in the absence of a cultural shock the offspring adopts the preferences of their unique parent. Needless to say these are strong assumptions. Yet the point of the model is not to assess the generality of the hypothesized

[7] This could be viewed as an implicit assumption that parental consumption only involves a fixed minimum consumption requirement, which we then for expositional simplicity normalize to zero. With a positive (exogenous) level of parental subsistence consumption we would have to take into account that there might be situations (i.e., parameter configurations and population levels) under which dynasties die out. We have no particular interest in studying this sort of scenario, which motivates the normalization.

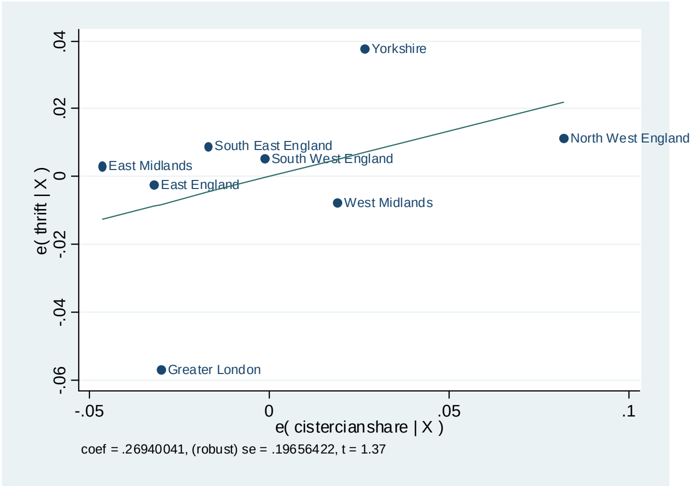


Figure 3: Cross regional correlation between Cistercian intensity and contemporary thrift (source World Values Survey).

trajectory, whereby a shock to the preferences of a small subset of individuals eventually causes a proliferation of these new preferences in the population at large. The point is rather to examine the implications of such a trajectory, which we a priori hypothesize is relevant for pre-industrial England, and then subsequently confront with data, and potentially reject.[8]

By focusing on the impact of changes in the attitude to hard work we suppress changes in cultural attitudes towards saving and investment; i.e., thrift. It is worth observing, however, that while the model focuses on the work ethic of individuals, similar results would likely arise if we instead examined thrift. As long as thrift implies a greater earnings potential, groups with high thriftiness will be selected in the Malthusian setting.[9]

A final observation worth making is that in the canonical Malthusian macro model fertility is endogenously determined, while child mortality is suppressed or implicit (see Ashraf and Galor 2009). Here it is the other way around. We follow this alternative route because it allows us to capture the taste for hard work in a

[8] In a more detailed model one would want to study the conditions under which preferences are in fact maintained over time within a dynasty, and distinguish between men and women in order to study matching. But it should be clear that the reduced form outcome studied below may well be viable in this richer environment if differences prevail across dynasties in perceived benefits of high work ethic values (for which reason such values are promoted by some dynasties but not by others; Doepke and Zillibotti 2008), and if there is sufficient assortative matching in the marriage market such that individuals with identical values choose to set up a family together (Becker, 1973). In such an environment one can think about the "Cistercian shock" as being represented by a change in perceived benefits to a high work ethic rather than by directly modifying preferences of some individuals. In order to maintain tractability, while studying the dynamic general equilibrium of the economy in the presence of cultural heterogeneity, we ignore this kind of micro-level behavior however interesting it may be in its own right.

[9] Becker (1980) explores a dynamic economy where agents differ in terms of the rate of time preference; i.e., in terms of "thrift". In the long run the most patient dynasty ends up "owning" the economy. Below we demonstrate a similar result in that dynasties with greater work ethics will end up dominating the population.

convenient way, while at the same time retaining the basic properties of the Malthusian framework as well as simplicity and tractability.

2.2.1 Individual's optimization problem

Individuals derive disutility from work effort and utility from the number of surviving offspring, with utility given as

$$u_t = \omega \log(1 - e_t) + \log(\sigma_t n). \tag{1}$$

In equation (1) $e_t \in (0, 1)$ is work effort, $n > 1$ is the (fixed) number of children coming into existence and σ_t is the fraction that survives to adulthood. The parameter $\omega > 0$ captures individuals' distaste for hard work. It can thus be seen as a simple manifestation of the work ethic of the individual, which we will think of as a cultural value that typically is fixed but may undergo occasional "mutation" in a population. In the absence of shocks to preferences, we assume children inherit the cultural value from their parent. For now we consider a population where everyone shares the same work ethic, ω. Finally, observe that since $\sigma_t n$ is ultimately linked to consumption the utility function is in a reduced form sense defined over consumption and work effort.

A higher level of effort allows for more income, which in turn facilitates a higher level of child consumption, c_t, according to

$$c_t = e_t y_t, \tag{2}$$

where y_t is potential per capita income. Finally, we assume that the survival rate to adulthood depends on child consumption in the following way

$$\sigma_t = \min\left\{\frac{c_t/n}{\eta}, 1\right\}, \tag{3}$$

where the parameter $\eta > 0$ is a reference consumption level.[10] As consumption per child falls below η, child mortality rises. Equation (3) is meant to capture that insufficient nutrition during early childhood weakens the offspring and thereby (deterministically) elevates the mortality rate. In this way we capture the notion of a Malthusian "positive check": In periods of plenty c_t/n will rise thus allowing for more surviving offspring; vice versa when income falls.[11]

---

[10] Assuming that maximum survival is 100% is a simplification. Nothing would change if we, at the costs of additional notation, were to assume a maximum survival rate of, say, $\bar{\sigma} < 1$ instead; that is, assume that $\sigma_t = \min\{[c_t/n]/\eta, \bar{\sigma}\}$.

[11] Kelly and O'Grada (2008) find that low real wages caused by bad harvests led to increased mortality in England during the 14th and 15th century. This link is however weakened from the 16th century onwards, quite possibly due to the introduction of the Poor Law.

Maximizing (1) subject to (2), (3) and $e_t \in (0,1)$, we find that

$$e_t^* = \begin{cases} \frac{\eta n}{y_t} & \text{if} \quad y_t \geq (1+\omega)\, n\eta \\ \frac{1}{1+\omega} & \text{if} \quad y_t < (1+\omega)\, n\eta \end{cases}, \quad c_t^* = \begin{cases} \eta n & \text{if} \quad y_t \geq (1+\omega)\, n\eta \\ \frac{1}{1+\omega} y_t & \text{if} \quad y_t < (1+\omega)\, n\eta \end{cases}. \tag{4}$$

The solution in (4) shows that for all t, $e_t^* \in \left(0, \frac{1}{1+\omega}\right]$. That is, it can never be optimal to supply an effort level larger than the upper bound $\frac{1}{1+\omega}$. This is so because with $y_t > (1+\omega)\, n\eta$ potential income is so high that providing an effort level at $\frac{1}{1+\omega}$ would make $\frac{c_t/n}{\eta} > 1$. This can never be optimal since individuals (by assumption) have no interest in consumption *per se*; consumption only matters insofar as it increases σ_t. Indeed, when $\frac{c_t/n}{\eta} > 1$ utility can be increased by lowering e_t (and thus c_t) without affecting σ_t. On the other hand, when $y_t < (1+\omega)\, n\eta$ the effort level is constant; i.e., $e_t^* = e^* = \frac{1}{1+\omega}$.

2.2.2 Production

There is a unique consumption good, Y_t, which is produced using labor input, L_t, as well as a fixed supply of land, X. Technology (or aggregate efficiency), A, is also parametrically fixed. Potential income is

$$Y_t = AL_t^{\alpha} X^{1-\alpha} \Rightarrow y_t = A\,(L_t/X)^{\alpha-1}, \tag{5}$$

where $y_t \equiv Y_t/L_t$.

Actual income of individuals depends on effort. Effort is thought to scale income up or down but is not subject to diminishing returns. This assumption ensures that dynasties who exert more effort will hold a permanent earnings advantage; if effort is subject to diminishing returns high-effort dynasties would not persistently be able to sustain larger family sizes than low-effort dynasties. The absence of diminishing returns to effort may be reasonable if "effort" is given a broad interpretation. That is, if effort is thought to capture the intensity at which individuals dedicate themselves mentally as well as physically to income enhancing activities, rather than being narrowly defined as the supply of working hours. Assuming labor absorbs all rents, the income of an individual therefore is

$$e_t y_t = e_t A\,(L_t/X)^{\alpha-1}. \tag{6}$$

2.2.3 The Evolution of the Economy

We now characterize the dynamic evolution of the economy. Initially we consider a setting with cultural homogeneity. Having characterized this we then introduce a cultural change, taking the form of a parametric change in preferences, and investigate how this cultural change plays out in the economy.

Cultural Homogeneity When all individuals share the same work ethic, ω, the size of the population evolves according to

$$L_{t+1} = \sigma_t n L_t, \ L_0 \text{ given.} \tag{7}$$

We therefore have the following lemma:

Lemma 1

(i) The time path for population is given by the law of motion

$$\begin{aligned} L_{t+1} &= \begin{cases} nL_t & \text{if} \quad L_t \leq \overline{L} \\ \frac{1}{1+\omega}\frac{1}{\eta}AL_t^{\alpha}X^{1-\alpha} & \text{if} \quad L_t > \overline{L} \end{cases}, \\ &\equiv G(L_t), \ L_0 \textit{ given,} \end{aligned}$$

and $\overline{L} \equiv \left(\frac{(1+\omega)n\eta}{A}\right)^{\frac{1}{\alpha-1}} X$.

(ii) For constant values of ω, A, η *and* X*, the model admits a unique steady state population size,* L^**, given by*

$$L^* = \left(\frac{A}{1+\omega}\frac{1}{\eta}\right)^{\frac{1}{1-\alpha}} X.$$

(iii) Steady state income per capita is

$$ey^* = \eta.$$

Proof. See Appendix.

Proposition 1

(i) The level of income per capita is independent of technology and preferences. (ii) A higher level of technological sophistication and greater preference for work effort increases long-run population size: $\partial L^*/\partial A > 0$, $\partial L^*/\partial \omega < 0$.

Proof. See Appendix.

The dynamic system works as follows. If initially the population is sufficiently small (i.e., if $L_t < \bar{L}$) the level of income per capita in the economy is high enough to ensure that all children survive.[12] Hence $\sigma = 1$, for which reason the population grows at the exogenous rate n. As $n > 1$, a steady state does not exist in

[12] Again, this follows since we have (to conserve on notation) defined the maximum survival to be 1.

the range $0 < L_t < \bar{L}$. Eventually, however, income drops below the threshold needed for an entire cohort to survive to adulthood, due to diminishing returns to labor input. Gradually, therefore, population growth grinds to a halt and the economy ends up in the steady state L^*. In the steady state the level of income per capita is η; *independent* of technology and preferences.

The intuition for the comparative statics is straightforward. Individuals with a greater work ethic will obtain a higher level of income, which in turn translates into more surviving offspring. Similarly, if A rises, income per capita increases, and this again leads to more surviving offspring. However, in both cases a larger population is associated with diminishing returns, which serves to equilibrate the system. In the long-run, therefore, greater work effort or technological advances will not elevate income per capita but are fully converted into a larger population. Aside from the impact of cultural values on long-run population density, these predictions coincide with those of the canonical Malthusian macro model (Ashraf and Galor 2009).

Cultural Heterogeneity Suppose now that a (small) subgroup of the population experience a parametric preference change; disutility from work effort declines. Subsequently there are two groups in society: group one and two, with $\omega_2 < \omega_1$. These new preferences are preserved within the dynasty, and thus transmitted from parent to offspring. By assuming a *unique* shock to preferences we are able to examine the pure selection channel by which the new preferences become more pervasive in the population. That is, by way of higher reproductive success, high effort dynasties grow as a share of the population. Naturally, a process of cultural change occurs more rapidly if low effort dynasties gradually choose to mimic the successful high effort dynasties. But in the present case we assume that there is a unique shock to preferences.

Hence, in the absence of further shocks we have the following lemma:

Lemma 2

(i) With cultural heterogeneity the law of motion for population size is given by

$$
\begin{aligned}
L_{t+1} &= \begin{cases} n\left(L_{1,t}+L_{2,t}\right) & \text{if} \quad L_t < \overline{L}_1 \\ nL_{2,t} + \frac{1}{1+\omega_1}\frac{1}{\eta}AL_{1,t}^{\alpha}X^{1-\alpha} & \text{if} \quad \overline{L}_1 \leq L_t < \overline{L}_2 \\ \left\{\frac{\omega_1-\omega_2}{(1+\omega_1)(1+\omega_2)}\left[1+\left(\frac{1+\omega_2}{1+\omega_1}\right)^t\frac{L_{1,0}}{L_{2,0}}\right]^{-1}+\frac{1}{1+\omega_1}\right\}\frac{1}{\eta}AL_t^{\alpha}X^{1-\alpha} & \text{if} \quad \overline{L}_2 \leq L_t \end{cases}, \\
&\equiv G\left(L_t, t\right), \; L_{1,0}, L_{2,0} \text{ given,}
\end{aligned}
$$

where total population $L_t = \sum_i L_{it}$, $\overline{L}_i \equiv \left(\frac{(1+\omega_i)n\eta}{A}\right)^{\frac{1}{\alpha-1}} X$, $i = 1, 2$.

(ii) For constant values of ω_i, A, η *and* X, *the model admits a unique steady state population size,* L^*,

given by

$$L^* = \left(\frac{A}{1+\omega_2}\frac{1}{\eta}\right)^{\frac{1}{1-\alpha}} X,$$

where $L^* > \overline{L}_2$.

Proof. See Appendix

Proposition 2

If initially the population holds a work ethic consistent with ω_1 *and then subsequently an arbitrarily small subgroup of the population changes cultural values to* $\omega_2 < \omega_1$ *then long-run aggregate population size rises.*

Proof. See Appendix

With two groups there are two thresholds to be distinguished, as ω is a co-determinant of the level of population at which $\sigma = 1$. As group two exerts more effort ($\omega_2 < \omega_1$) it is able to ensure survival of an entire birth cohort at a lower level of potential income, y, than group one. This creates the intermediate regime, $\bar{L}_1 < L_t < \bar{L}_2$, where all offspring of high work ethic parents survive until period two, whereas a fraction of the children of the low work ethic group dies during childhood. Eventually, however, income falls to a sufficiently low level to produce $\sigma < 1$ for both groups. The process ultimately stabilizes and allows for a unique steady state as described in Lemma 2.

The dynamic process works as follows. Initially a small group of citizens change cultural values. Since the new group works harder, their income is greater. This works to increase population density. However, if the high work ethic group initially is small then the immediate impact on aggregate population size will be miniscule. But since the high work ethic group can afford higher levels of child consumption, it holds a reproductive advantage in the Malthusian setting. This advantage implies that the group's population share gradually rises over time, thereby increasingly stimulating aggregate population size. Hence, after the initial shock to the economy, the growth process is driven by the changing composition of the population in terms of cultural values; a process of growth through cultural change is occurring.

Eventually the group with high work ethic will dominate the population, and the economy convergences to a steady state where population size reflects the preferences of the high work ethic group. It follows from Proposition 2 that the steady state level of population is higher in the new steady state compared to the original one, as $\omega_2 < \omega_1$.

The model thus shows how a change in a certain cultural attitude in a small subset of the population may rise in importance due to selective pressures and eventually influence the macroeconomy. The source of the change of preferences is left unexplained by the model. But it seems plausible that the Cistercians have influenced county populations in this manner, as argued in Section 2.1. Accordingly, our hypothesis is that Cistercians planted the seeds of change by affecting the cultural attitudes; or, more appropriately, the work ethic of a (in principle arbitrarily) small part of the county population. By so doing, they instigated a process of growth through cultural change.

2.2.4 Speed of Diffusion

A question of some relevance is how fast the cultural diffusion process played out if it only emanates from differential population growth rates across dynasties with different values. Naturally, the process would conceivably occur at a faster rate than what we find below if values gradually diffuse across dynasties as well, following the initial shock to a select group of dynasties. In practise one may well imagine that both mechanisms were at work.

In order to examine the speed of population-growth driven cultural change, note that the fraction of individuals with high work ethic at time t is

$$\frac{L_{2,t}}{L_t} \equiv \pi_t = \frac{1}{1 + \left(\frac{1+\omega_2}{1+\omega_1}\right)^t \frac{L_{1,0}}{L_{2,0}}}, \tag{8}$$

when $L_t > \bar{L}_2$. Hence we focus on the (more realistic) case where not all children survive to adulthood from either group. The speed at which ω_2 becomes dominant in society depends on how much more effort the high work ethic group exerts, $\frac{1+\omega_2}{1+\omega_1} = \frac{e_1}{e_2}$, as well as how many individuals were "persuaded" to change their values as of time $t = 0$. The ratio of ω's is hard to pin down in any precise manner. But suppose group two exerts 20% higher effort than group one.[13] In this case Figure 4 shows how the new cultural values grow in significance over time for different assumptions about the initial "infection rate"; that is, $\pi_0 = 0.1\%$, 1% and 10% of the population, respectively.

The spread of the new cultural values follows an S-shaped trajectory: the process is slow to begin with but accelerates over time and ultimately levels off. Consider the curve in the middle, associated with an initial "infection rate" of 1%. The first 10 generations only raise the fraction with strong work ethic modestly (to about 6%), the next 10 generations increase the share to 30% of the population, and another five to nearly 50%. If a generation is about 20 years, 25 generations (what it takes to go from 1% to 50%) is about 500 years. The point is that, within the window of observation available to us (about 500 years), it is possible for a small (initial) cultural shock from the Cistercians to accumulate into a major aggregate impact on the

[13] Clark and Van der Werf (1998) estimate that the number of days worked per year (standard deviation in parenthesis) rose in England from 266 (4.8) in 1560-99 to 280 (12.9) in 1771. Suppose this increase is attributable to the rise of the Protestant work ethic, resulting from the Cistercian presence and the Reformation. Then the estimated increase over time in work days provides a crude guesstimate for ω. Factoring in the statistical uncertainty we may note that working days in 1771 may have been between 5% lower and 23% higher than in 1560, with a mean around +10%. Hence, assuming a 20% higher work effort may not be outlandish; especially so since our notion of work effort is somewhat broader than the mere number of days worked.

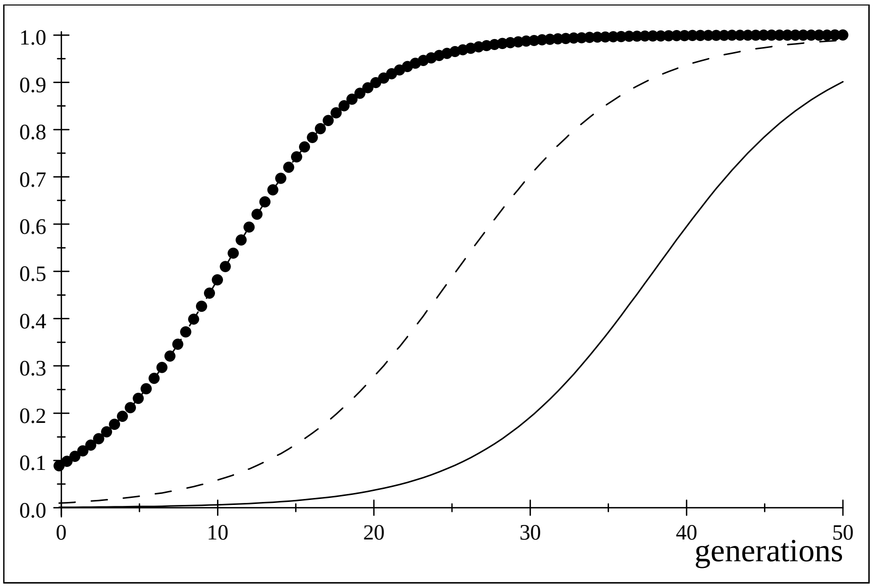


Figure 4: The rise of new cultural values in the population. *Assumptions*: (a) 20% higher work effort among individuals with "high" work effort". (b) Initial "infection" rate: 1/1000 (solid black), 1/100 (dashed), 1/10 (dotted).

composition of the population solely by way of selective pressure.

Another point worth emphasizing is the implied comparative differences in cultural beliefs that seemingly small initial differences translates into. With an initial infection rate, π_0, of one percent, 50 percent of the population holds a high work ethic after 25 generations; but only eight percent have high work ethic after 25 generations if the initial infection rate is 1/10th of a percentage point. This implies that variations in the intensity of Cistercian presence may have generated substantial comparative differences in cultural values across English counties over the period in question, by affecting π_0. It may therefore be possible to detect the legacy of the Cistercians on population dynamics over the period 1377-1801, which we examine below.

3 Empirical Analysis

This section proceeds in a series of steps. We begin by deriving an empirical model based on the theoretical model from the previous section. Subsequently, in Section 3.2, we discuss the relevance and interpretation of the resulting empirical model. Section 3.3 presents the data, while Section 3.4 contains our OLS regression results. Finally, Section 3.5 discusses potential pitfalls in ways of identification and reports our IV results.

3.1 From the Theoretical Model to an Empirical Model

The theoretical analysis established that, upon a small cultural shock to the population, areas with a larger fraction of their citizens having a higher work ethic should see faster population growth. To make this prediction amendable to empirical testing, we use Lemma 2 to write the law of motion for aggregate population as

$$\log(L_{t+1}) - \log(L_t) = -\log(\eta) - (1-\alpha)\log(L_t/X) + \log(\psi_t) + \log(A).$$

where

$$\psi_t = \left[\frac{\omega_1 - \omega_2}{(1+\omega_1)(1+\omega_2)}\pi_t + \frac{1}{1+\omega_1}\right],$$

while $\pi_t \equiv L_{2,t}/L_t$ captures the fraction of the population with high work ethic.[14] If we linearize $\log(\psi_t)$ around $\pi_t = 0$ we obtain $\log(\psi_t) \approx -\log(1+\omega_1) + \frac{\omega_1-\omega_2}{1+\omega_2}\pi_t$, which can be reinserted into the law of motion for population growth so as to obtain (approximately)

$$\log(L_{t+1}) - \log(L_t) = -\log(1+\omega_1) - \log(\eta) + \frac{\omega_1-\omega_2}{1+\omega_2}\pi_t - (1-\alpha)\log(L_t/X_i) + \log(A).$$

Finally, denoting a county by i and adding an error term, we may write the above as the empirical model:

$$\Delta\log(L_{it+1}) = a_0 + a_1\log(L_{it}/X_i) + a_2\pi_{it} + \mathbf{Z}_i'\mathbf{a} + \epsilon_i,$$

where $\mathbf{Z}_i$ contains time-invariant controls for productivity (A).

Naturally, we do not have data on π_{it}. But, according to the theory, we may proxy it using some measure of the *intensity* of Cistercian presence in the county, as it should influence π_{i0}, and thereby π_{it} (see equation (8)). We define this intensity as the *Cistercian presence relative to other moral influences.* Since the Church was the principal authority in matters of moral in medieval times, we construct π as the ratio of Cistercian monasteries to all religious houses; i.e. $\pi = C/R$. However, the counterfactual we are interested in is that of changing the *composition* of moral influences while at the same time holding constant its *level.* This dictates that we also control for the total number of religious houses, R, separately. Consequently, we take the following specification to the data:

$$\Delta\log(L_{it+1}) = a_0 + a_1\log(L_{it}/X_i) + a_2\frac{C_{it}}{R_{it}} + a_3R_{it} + \mathbf{Z}_i'\mathbf{a} + \varepsilon_i. \quad (9)$$

Ceteris paribus, areas with more Cistercians saw a larger fraction of the population initially being "persuaded" by the Cistercian work ethic. As seen from Figure 4 this should imply a higher π_{it} at any given

[14] More specifically, we use equation (10) in the proof of the lemma; see Appendix A.3.

point in time. As a result, we expect a_2 to come out with a positive sign. In addition, theory predicts that $a_1 < 0$, capturing convergence effects. The coefficient a_3 is a priori indeterminate.

3.2 Relevance and Interpretation of the Empirical Model

In the context of empirical relevance there are two issues worth considering. First, is the Malthusian perspective relevant for the period in question? Second, equation (9) is derived under a "closed economy" assumption. That is, the model does not allow for migration flows. Is this a reasonable approximation for the period in question?

Ultimately it is an empirical issue whether the Malthusian population theory has any bearing on developments in England from 1377-1801. However, a priori we believe a reasonable case can be made that Malthusian considerations were relevant. Clark (2007) builds a strong case that Malthusian dynamics were relevant until about 1800 in England. In a similar vein, Møller and Sharp (2008), employing time series data for England over the period 1560-1760, confirm the relevance of Malthusian population dynamics. Using cross-country data Ashraf and Galor (2009) also confirm the central predictions of the Malthusian model in pre-industrial times.

We also view the closed economy assumption as a reasonable (albeit crude) approximation to reality for the period in question. Although serfdom began to decline with the Black Death, and was practically obsolete in England by the sixteenth century, even as late as the early eighteenth century internal migration was characterized by limited geographical movement (Clark 1979). A contributing factor to the low degree of mobility was the Old Poor Law, which meant to supply relief to the temporarily unemployed, and which was administrated at the Parish level. In particular, the Settlement Act of 1662 imposed that only "established residents" of a Parish could receive relief. In practise this meant that only individuals who were able to prove an affiliation with the Parish through birth, marriage or apprenticeship were eligible for aid. Needless to say, this policy worked to lower mobility. The Poor Law Amendment Act overhauled the existing system. In particular, it established "Poor Law Unions" around groups of Parishes, which then administrated the poor relief. But this amendment did not take effect until 1834.

This is not to say that individuals did not move at all during the period in question. They did, and increasingly so over time. London, in particular, enjoyed a special status and always experienced substantial immigration. But it was the industrial revolution which saw the major break with the past and large scale migration in particular to the new industrial centres in the Northwest and the Midlands (Nicholas and Shergold 1987). As a result, suppressing internal migration seems like a reasonable approximation for most of the period in question.

Nevertheless, one may speculate what the implications would be if the assumption is not met. Suppose that counties characterized by individuals with high work ethic are (for this reason) more innovative. That is, suppose A is affected by culture (more on this below). If so, one would expect people to migrate to the high work ethic counties, which would stimulate growth in population density in areas where individuals have high work ethic. In terms of population dynamics this outcome is therefore observably equivalent to the no-migration scenario that we examined theoretically. The only way to distinguish the migration scenario from the no-migration scenario would be to study the impact of Cistercian presence on income per capita. With migration innovations would induce rising income per capita, wheras this is not the case when no migration is taking place (see Ashraf and Galor, 2009). Unfortunately, county level data on income per capita is not available for England during this period. Hence we are unable to distinguish whether the Cistercians, through instigation of cultural change, induced higher population growth either by increasing the number of surviving offspring at the county level (as suggested by our model), by stimulating inward migration, or by some combination of the two. However, regardless of the precise source of rising population density, a positive impact from Cistercians on population growth implies, in a Malthusian setting, a productivity enhancing effect from the Order. Hence, the empirical test of the impact from the Cistercians (using equation (9)) is meaningful whatever the "truth" may be about internal migration in England prior to 1801. But there are other reasons why the interpretation of a_2 might differ from what is implied by the theoretical model.

While equation (9) has a structural foundation in our model, we doubt very much that our estimations of a_2 will map into the structural parameters ω_1 and ω_2, for two reasons. First, the Cistercians almost certainly also influenced attitudes towards saving and investment; i.e., thrift. Hence, dynasties that were influenced by the Cistercian mindset also had an earnings advantage through this channel. A more fully articulated theoretical model (but also a much more complex one) would allow both cultural traits to emerge in the wake of Cistercian influence, and grow in pervasiveness over time. We conjecture that such a model would predict that the relationship between π_{it} and population growth reflects preferences for work effort as well as the willingness to postpone consumption. Second, it is probable that the changing cultural values influenced productivity, A. As explained in Section 2.1, the Cistercian order were at the forefront of technological change prior to the Reformation. Insofar as their cultural attitudes spread throughout local populations, it is conceivable that this also encouraged local technology adoption. If so, the level of A may have been influenced by the frequency of individuals valuing hard work and thrift.[15] While our regressions below do involve several controls for A, our estimates for a_2 may nevertheless also be capturing the indirect influence of cultural values on population growth through productivity.

In sum, one should be cognizant of the fact that the cultural values emphasized by the proposed theory represents "ultimate determinants" of prosperity, which served to stimulate key proximate sources of growth: labor effort, capital accumulation and technological change. Realistically, the estimate for a_2 is therefore

[15] Endogenous technological change is not inconsistent with Malthusian population dynamics; see Aiyar, Dalgaard and Moav (2008).

best viewed as the reduced form impact of cultural values on growth mediated by these individual channels.

3.3 Data

3.3.1 Cistercian presence (π_{it})

Researchers at the University College London (UCL) have constructed a database of 776 religious houses in England from the 10th to the 16th century. The database includes the name of the particular religious houses, the order of the monks, nuns etc., year of foundation and dissolution, and the county in which the monastery was located.[16] We gathered these data into one dataset, which we then used to calculate the number of religious houses in each county (*relhouses*) and the number of Cistercian monasteries as a share of total religious houses in each county (*cistercianshare*). In order to gauge robustness, we also construct the share of other major religious orders. We made one correction to the data with respect to the city of York, which was listed by UCL as a county. York was a walled city situated in North Yorkshire. To be able to match the data with the data on population density, we re-coded it as part of the county North Yorkshire. Table 1 lists the frequency distribution of the various religious houses in the UCL database, while Figure 5 maps the spatial distribution of the *cistercianshare*. In the analysis below we focus on the main religious orders: Benedictian monks, Augustinian canons, Cistercian monks, as well as the Premons and Cluniacs.

[Table 1 about here]

3.3.2 Population and population density (L_{it} and L_{it}/X_i)

We obtained data on population density for the year 1377 from Campbell (2008) (*popdens1377*). Campbell also provides the area of the counties; we transformed them from square miles into square kilometers. The distribution of the population in 1377 is based on 1.38 million adult males and females who contributed to the poll tax of 1377.[17] The level of the population is based on an estimate by Campbell (2000) of 4 million.[18] Campbell only reports population numbers for the aggregate of London and Middlesex, not for the two counties separately. In order to match the data, all data on all variables is aggregated in this way. Yet we end up excluding London and Middlesex in all regressions, since it is an outlier. We note for completeness, however, that including London and Middlesex makes no difference to our results.

Wrigley (2007) provides population estimates for 1761, 1771, 1781, 1791, and 1801. These are based on registered marriages, which were more completely recorded than baptisms and burials on which previ-

[16] The data are available online at: <http://www.ucl.ac.uk/history2/englishmonasticarchives/religioushouses/index.php>.
[17] These numbers are available in Dobson (1983).
[18] Campbell (2008) also reports population data for 1300. But since about 10% of Cistercian settlement occur around that year, or after, the risk of reverse causality tainting our estimations would be enhanced if we used 1300 as our initial year. As a result we stick with 1377 as the initial date. However, we will use the 1300 numbers in the context of our IV regressions below.

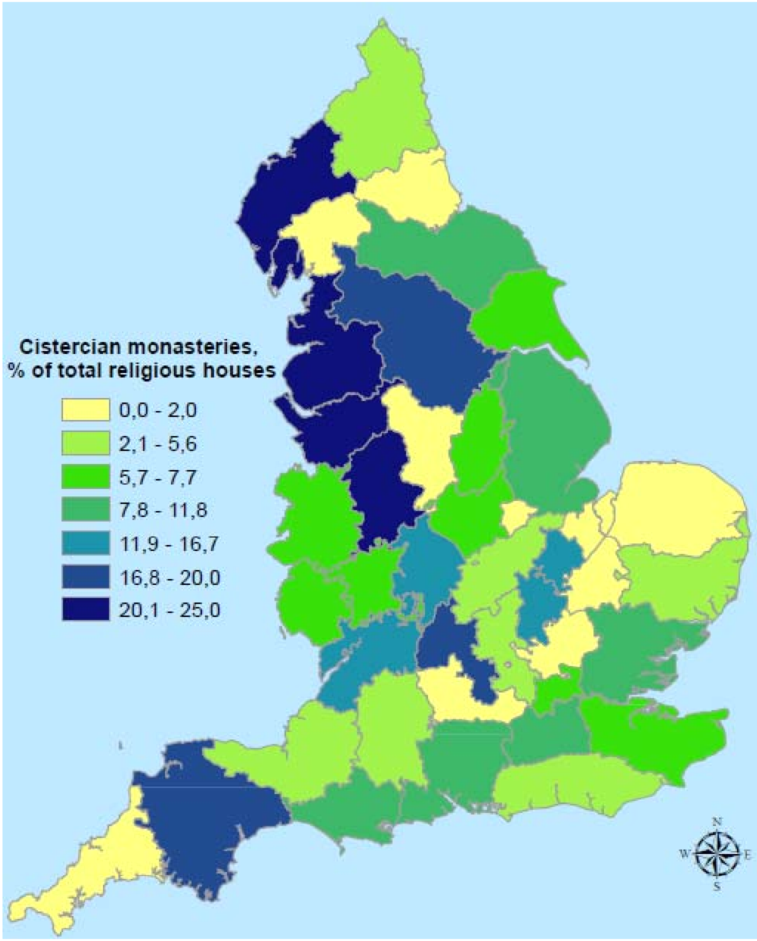


Figure 5: Cistercian monasteries as a share of all religious houses across England, 1098-1540

ous population estimates were based (Rickman 1802). While our preferred variable is population in 1801 (*popdens1801*), the choice of end-year is inconsequential to our results.

3.3.3 Time invariant productivity controls (Z_i)

Agricultural Land Classification Natural England provides a measure of agricultural land classified into five grades plus classifications for non-agricultural and urban land. Grade one is best quality and grade five is poorest quality, grade six is non-agricultural land and grade seven is urban. The measure is calculated by Natural England using information on climate (temperature, rainfall, aspect, exposure, frost risk), site (gradient, micro-relief, flood risk) and soil (depth, structure, texture, chemicals, stoniness). The source of the data is Raster Digital mapping with a scale of 1:250,000.[19] The data was gathered with coordinate precision of 1 meter. We used these data to create a measure of agricultural land quality within each county.

The earliest digital map of English counties is from 1851. These data were kindly provided to us by University of Portsmouth and the Great Britain Historical GIS Project. Combining the shapefile including the agricultural land quality and the shapefile including English county borders, we were able to create measures of the area in a county with agricultural land of quality level 1-5, each as a share of total county area.[20] Our preferred variable is the combination of qualities 1 and 2, which we shall denote *agrquality1_2*.[21]

Waterways As noted in Section 2.1, the Cistercian were strong exponents of water powered production and they employed advanced irrigation techniques, which could be responsible for their influence on English population growth. To control for this kind of influence from Cistercian presence we therefore add controls for waterways.

The German company Geofabrik freely provides shapefiles on various geographic features.[22] Of our interest is their data on waterways in Great Britain, where waterways are divided into canal, dock, drain, moat, river, and stream.[23] As with the data on agricultural land quality, we merge the shapefile describing waterways with the shapefile describing the county borders of England. The outcomes of interest from this procedure is the total length of rivers as a share of the total area in a county (*rivershare*) and the total length of streams as a share of county area (*streamshare*). The variable (*riverstreamshare*) measures the total length of rivers and streams as a share of the county area.

Regional fixed effects In an effort to control more rigorously for structural characteristics with bearing on population growth we add a full set of regional dummy variables. The regional classification is based on Government Office regions: East Midlands, East of England, London, North East, North West, South East, South West, West Midlands, and Yorkshire and the Humber. Observe that in a sample consisting of 40 counties it is a rather strong check of the relevance of Cistercian presence to allow for nine regional identifiers.

Table 2 provides summary statistics and a correlation matrix on the variables discussed above.

[19] Available online at: <http://www.gis.naturalengland.org.uk/pubs/gis/gis_register.asp>. Data description available online at: <http://www.magic.gov.uk/datadoc/metadata.asp?dataset=2&x=16&y=10> and <http://naturalengland.etraderstores.com/NaturalEnglandShop/product.aspx?ProductID=88ff926a-3177-4090-aecb-00e6c9030b29>.

[20] The total county area was here calculated by summing over the land quality variable, since this variable spans the entire area.

[21] None of the results change if we instead include agrquality1 and agrquality2 together or separately. If we include a variable measuring the aggregate agricultural quality over grades 1, 2, and 3, results are unchanged, except column 9 of Table 4 below, where the t-value on *cistercianshare* drops to 1.16.

[22] These shapefiles are based on maps created by the OpenStreetMap project using data from portable GPS devices, aerial photography, other free sources, or simply from local knowledge.

[23] Available online at: <http://download.geofabrik.de/osm/europe/great_britain/>

[Tables 2a and b about here]

3.4 Results

Table 3 reports our baseline results. In all columns of the table we control for initial population density, the total number of religious houses, the share of all religious houses which are Cistercian, and the productivity control agricultural land quality. The regression in column 1 shows that these variables collectively hold significant explanatory power with respect to population growth over the period 1377-1801; the regression explains roughly two thirds of the variation in the dependent variable.

The following eight columns add the additional controls discussed above, one by one. Finally, column 10 includes all the controls simultaneously. Several features of the results are noteworthy. First, the share of Cistercians is statistically significant in all columns. This means that the composition of religious houses matters, with more Cistercians being associated with higher population growth rates. In addition, Cistercian point estimates are fairly stable, always situated in the interval [1.67, 2.07].

Second, while initial population density displays the expected conditional convergence feature, it is surprising that agricultural land quality has a negative impact on population growth. However, the explanation for this is partly found in the high positive correlation between initial population density and agricultural land quality (corr. coef. = 0.43), cf. Table 2b. This means that initial population density picks up some of the effect of land quality on population growth.

Third, land area adds significant explanatory power. Yet the fact that the physical infrastructure of rivers and/or streams did not seem to matter for population growth suggests that neither irrigation nor aqua-based transportation were significant binding constraints to growth.

Fourth, in columns 6-9, where we add the share of the other dominant religious orders, only the Cluniac order adds significant explanatory power. This is in itself an interesting finding since the Cluniacs can be viewed as carriers of the same sort of cultural values embedded in the Order of the Cistercians. The Cluniacs were an earlier attempt (from the tenth century) to return to a more strict observance of the Rule of St. Benedict, although within the Benedictine order of monks (Southern 1970). The significance of the Cluniacs is all the more interesting in light of the fact that Weber highlighted this particular Order alongside the Cistercians as precursors of "Protestant values" (cf. Section 2.1). Accordingly, their influence on pre-industrial development in England represents further support for a pre-Reformation origin of the "Protestant ethic".

Finally, in column 10, where we include all control variables simultaneously, the Cistercian share remains significant and situated in the aforementioned interval. The association between Cistercian presence and population growth is therefore quite robust.

[Table 3 about here]

In Table 4 we add regional fixed effects to all columns of Table 3. To the extent that we have omitted certain time-invariant regional productivity factors, regional fixed effects will alleviate this problem provided the regional classification captures these omitted confounders. One case in point could be proximity to coal production, which Allen (2009) and Pomeranz (2000) argue was critical for British industrialization because it supplied an inexhaustible supply of cheap energy. Since the location of coal mines is a fixed effect, regional dummies will pick up this effect.

The first thing to notice in Table 4 is that the share of Cistercian houses remains significant in all columns save for columns 7 and 10. However, in both columns the regional dummies are jointly insignificant, for which reason these columns add nothing to columns 7 and 10 of Table 3. Consequently, disagreement between Tables 3 and 4 only arise in column 5 with respect to initial population density.

Another way to appreciate the findings reported in Table 4, as compared to those reported in Table 3, is by observing that the point estimate for Cistercians is virtually unaffected by adding regional fixed effects. The occasional change in statistical significance is thus solely due to a reduced precision in estimation, which may well be due to multicollinearity; multicollinearity almost inevitably becomes an issue when we introduce nine regional identifiers, on top of the other controls, in a $N = 40$ sample.

Overall, the results from Table 3 continue to hold up fairly well when regional dummies are added. In particular, only the Cistercians and the Cluniacs seem to exert a significant impact on population growth during the period; it is, however, only the Cistercian impact that survives inclusion of all controls. Finally, it appears that the simple baseline model associated with column 1 of Table 3 is sufficient for purposes of accounting for the association between Cistercian presence and comparative English population growth.

[Table 4 about here]

Figure 6 provides a visual depiction of the relationship between the share of Cistercians and the growth rate in population as estimated by column 1, Table 3. Inspection of the figure shows that Westmorland and Lancashire may exert some leverage on the estimated coefficient. Yet, neither of these two counties is driving the result: exclusion of either one changes nothing. In addition, the share of Cistercians stays significant (slope est. = 1.47 and std. err. = 0.76) when we perform robust regression analysis (more detailed results are available upon request).[24]

What is the economic effect of changing the composition of religious houses in the direction of one more Cistercian abbey, holding the total number of religious houses constant? To answer this question we differentiate (9) with respect to C to get a_2/R. Evaluated at the mean of R, we get that the said change will increase the proportional difference in population by approximately 0.1 log point. This change would have lifted the population of Cambridgeshire in 1801, the median county in terms of population in 1377, from an

[24] This is the **rreg** option in Stata 10.

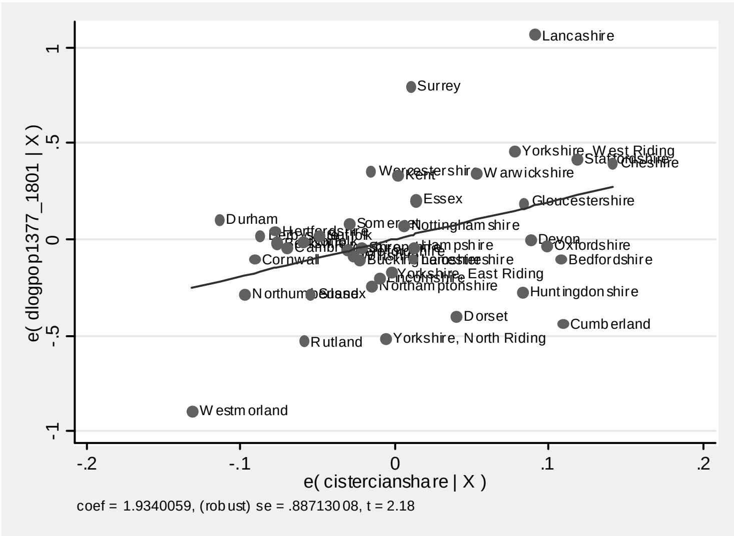


Figure 6: The impact of Cistercians presence on population growth.

actual size 93,440 to a counterfactual size of about 103,000.

3.5 Threats to Identification

3.5.1 Location of Cistercian Monasteries

An objection to the preceding results is that they could be spurious. That is, perhaps the Cistercians simply chose to locate in areas with a pronounced productive potential. Based on the historical evidence, however, this possibility does not seem likely. The Order had a stated preference for situating their monasteries in remote, even devastated, locations (Cooke 1893; Donkin 1963). Indeed, it has long been accepted by scholars that the Cistercians acted as transformers of wastelands into fertile farms, as mirrored in the poet Wordsworth's *Cistercian Monastery.*[25]

This conventional wisdom receives quantitative support in Table 2b, from which it is clear that Cistercian presence was lower in areas featuring high population density in 1377. Also supportive of the traditional view of the spatial distribution of monasteries is the negative correlation between the intensity of Cistercian presence and soil quality. Since we control for both initial population density and soil quality in our regressions, these regularities are unlikely to bias the parameter of interest. To this one may add that since the monasteries were largely in place from the beginning of our period of observation, as observed in Section 2.1, it is impossible that their location is endogenous to population growth during the ensuring centuries on

[25] "Where'er they rise, the sylvan waste retires, And aery harvests crown the fertile lea."

account of reverse causality.

Of course, one may worry that by chance Cistercian monasteries just happened to be located in areas that ultimately proved to be high growth regions. For instance, looking at Figure 5 it is clear that there is a cluster of Cistercian monasteries in the North West of England; areas that long after the Dissolution of Monasteries turned out to be rich on coal (Allen 2009). Yet coal played a relatively modest role *vis-à-vis* the growth process prior to industrialization, which is the time period we examine above. More generally, our regression analysis introduces nine regional dummies, which should ensure that omitted time-invariant confounders do not bias results. In spite of these consideration doubt may linger. As a consequence we provide a final check by invoking instrumental variables estimation.

The Cistercians had a particular preference for locating in secluded and sparsely populated areas, as explained above. At the time of arrival the most secluded areas may well have been the forests owned by the Crown. As Donkin (1963, p. 184) observes: "*..there is a really significant connection with the Royal Forests; one-third of all the English [Cistercian] houses lay at first within or very near their bounds [...]. In these areas there was a good deal of land of low value for endowments; nonroyal landowners were gravely hampered by the forest laws; and, as elsewhere, prospective founders undoubtedly responded to the willingness of the early generations of monks to exploit rough, undeveloped country.*" Thus, there may well have been a double coincidence of wants. Nonroyal landowners, wanting to save their souls from eternal damnation, had an interest in allowing Cistercians to settle at or near Royal Forests, which were of limited value beyond the occasional hunt with the Monarch. At the same time, this location satisfied the ascetic needs of the Cistercian settlers.[26] Hence the presence of a Royal Forest in a county could be a potentially viable instrument for Cistercian settlements. But is the presence of Royal Forest a plausibly excludable instrument for Cistercian presence?

The concept of Royal Forest was introduced in England by the Normans in the 11th century. They were protected areas of land (not necessarily woodland) where the king had privileged hunting rights under the "forest law", which offered strict penalties to anyone using these areas for hunting or farming. The system was at its height in the late 12th and early 13th centuries, but already in 1215 Magna Carta laid down limits to the power of the monarchy in the forests, and the "Great Perambulation" of 1300 vastly reduced the scale of the forests. Generally the system decayed after this time, and Henry VIII placed the forests under the Court of Augmentations in 1547, the body which was set up to administer the land and finances of the Roman Catholic Church after the Dissolution of the Monasteries. Although the designation "Royal Forest" still exists in contemporary England, Royal Forests were considered anachronistic after the Tudor period and the enforcement of any rights had completely died out by the mid-17th century (Grant 1991).

[26] Finally, the monarch may also have had an incentive to encourage the practise. Madden (1963) notes that the king likely granted rights of pasture over wide tracts of the royal lands and forests because the Cistercians were willing to pay for this service using revenue from sale of wool; wool which derived from sheep using the royal lands for grazing.

Towards the end Royal Forests were mainly associated with giving privileged access to timber, but even these rules were poorly enforced. Most of the protection for wooded areas within the Royal Forests was broken with the massive disafforestments of 1327 and the generally less effective means of enforcement (Young 1978, pp. 102-103). Accordingly, since the Royal Forest as an institution was of little importance after the 14th century, it seems plausible that while the settlement pattern of the Cistercians is partially explained by the location of Royal Forests, the location of Royal Forests as such cannot explain cross-county population growth into the nineteenth century.

We obtained data on the location of Royal Forests in the 13th Century from Bazeley (1921). Based on the maps constructed by Bazeley, we constructed a dummy variable, "Rforest", which is equal to one if a royal forest were to be found in the county in the 13th century.[27] Accordingly, we expect to find a positive partial effect of Royal Forest on the intensity of Cistercian settlements.

In order to better explain Cistercian settlements across areas where a Royal Forest was found in the 13th century, we also employ (log) population density in 1300; a point in time where not all monasteries had been founded (see Figure 1). The theory is that Cistercians would prefer to settle near Royal Forests, and even more so if the county in question was sparsely populated. As a result, we expect to find a negative effect from population density in 1300 in the first stage.

We also expect that population density in 1300 is excludable in our regressions. It should be of little relevance to growth in population density between 1377 and 1801 (i.e., above and beyond its influence via Cistercian settlements) since we control for initial population density in 1377 in the regressions below.

In sum, we believe the presence of Royal Forest and county level population density in 1300 both plausibly fulfill the exclusion restriction. Of course, with two instruments and one endogenous variable this prior can be subjected to formal tests, thus offering an opportunity to reject the identifying assumption.

Table 5 reports a summary of our results. In column 1 we estimate a stripped down model were the only independent variables are the intensity of Cistercian presence and initial population density in 1377. The two instruments have the expected sign in the first stage, and the Cistercian influence is estimated with high precision in the second stage. The obtained result is consistent with our OLS findings: more Cistercians seems to foster faster population growth from 1377 to 1801. Moreover, data does not allow us to reject the exclusion restriction.[28]

In the remaining columns we add more controls. In particular, in column 3 we add *all* the controls featured in Table 3 simultaneously; i.e., we estimate the equivalent of column 10 in Table 3. The impact of Cistercians remains significant and positive but the size of the point estimate shrinks. This is no surprise. As explained in Section 2, one may view Cistercian cultural values as an ultimate determinant of productivity,

[27] We also experimented with using forest area in the 13th century, but this instrument turned out to be weak.

[28] In column 1 the instruments are not statistically strong. But the Anderson-Rubin test (not shown), which is robust to weak identification, reveal that the Cistercian share is significant thus suggesting a causal impact.

which influences population growth through several more proximate pathways (e.g., technology, labor effort, and perhaps savings). The more controls we add, the more of these proximate determinants of population density are likely controlled for, which should cause the impact from Cistercian presence to decline. Of course, absent direct measurement of cultural values (or of all the proximate sources of growth), it should not be possible to eliminate the impact from the Order entirely, which is what we find. In columns 2 and 3 our instruments of choice are statisically strong and the exclusion restriction is not rejected by the data.

As a quick comparison between column 3 in Table 5 and column 10 in Table 3 makes clear the estimate rises somewhat when the intensity of Cistercian presence is instrumented. This is consistent with the underlying theory that Cistercian presence is capturing cultural values; if so, then Cistercian presence is a *proxy* variable for the fraction of the population carrying Protestant values, which therefore should be associated with an attenuation bias in the OLS setting.

[Table 5 about here]

Taken together, we believe that these results strengthen the case for a causal link between Cistercian presence and population growth 1377-1801. A greater Cistercian presence worked to promote population growth, consistent with the proposed theory of growth through cultural change. Still a causal impact could also arise for non-cultural reasons, as discussed next.

3.5.2 International Trade

There is an influential strand of literature which asserts that international interaction profoundly influences economic growth. For instance, the work of Frankel and Romer (1999), Alcalá and Ciccone (2004) and Andersen and Dalgaard (2011) suggests that geographic features that facilitate international international interaction (e.g., access to sea, small country size, etc.) hold a significant impact on actual international interaction (either via travel or trade) and ultimately on contemporary prosperity. But why might interaction, in ways of trade for instance, have stimulated growth in the very long run?

One possibility is that the intensity of international trade has influenced the type of policies and institutions that were implemented, and thereby long-run economic outcomes. More concretely, Acemoglu et al. (2005) argue that the Atlantic trade was a key driving force behind the rise of Western Europe after 1500. The argument is that the Atlantic trade enriched and strengthened commercial interests outside the royal circle in countries with non-absolutist initial institutions (such as England); this in turn shifted the balance of power away from the Crown, ultimately instigating significant pro-growth institutional reform.[29]

Naturally, this literature focuses on the development of nation states, whereas we are examining regional

[29] Another potential benefit from trade is that it enables knowledge spillovers between countries. We return to the issue of whether technology can account for our results in the next section.

development. Nevertheless, inspired by the theory of Acemoglu et al., a hypothesis that suggests itself is that counties particularly involved in international trade adopted *local* policies and *local* institutions more beneficial to growth than what was the case elsewhere.

Could such considerations impinge on the apparent link between Cistercians and population growth? It is certainly well known that the Cistercians were active in international trade; they were particularly involved in the trade of wool and cloth. Hence the Cistercians may have contributed to the establishment of *early* trade centers, which would then (partly) account for a persistent effect on *later* population growth. That is, long after the Cistercian presence ended, trade continued at high intensity in the places the Order supported early on. Observe that this theory would still support a causal impact from Cistercians on long-run growth, consistent with our 2SLS estimates above. However, according to this line of reasoning, cultural change is unimportant. In sum, the non-cultural trade story behind our findings would be

$$\text{Cistercians}_t \Rightarrow \text{Trade}_t \Rightarrow \text{Trade}_{t+T} \Rightarrow \text{Population density}_{t+T},$$

where the last arrow could be motivated via policy or institutional changes, following the logic of the Acemoglu et al. (2005) argument. We unfortunately lack data on early international trade flows at the county level. But what we can do instead is to examine whether the influence from Cistercians is reduced once we control for key geographical features that should support international trade; if actual historical trade, much like contemporary international trade, depend on geography then this is a viable proxy variable approach.

Specifically, we add a coastal dummy to our baseline regressions in Table 3. To be sure, if the Cistercians indeed instigated the rise of trade centers, one would expect this effect to be strong precisely along the coast, which is where goods arrive and are shipped off (recall that our regressions already control for inland waterways).

Table 6 presents the results from introducing the coastal dummy. Columns 1-9 reproduce Table 3 while adding on top the coastal dummy. In Column 10 we add all the controls simultaneously (11 in total). The basic message from the table is that the inclusion of the coastal dummy does not affect the partial correlation between the intensity of Cistercian presence and population growth. Hence, although there are some hints at a correlation between Cistercian presence and coastal areas (see Figure 5), and although their strong trading traditions may have left a lasting mark on economic growth, it does not seem that this can account for the reduced form link between the Order and long-run population growth.

[Table 6 about here]

3.5.3 Technology and Human Capital

As discussed in Section 2.1, the Cistercians were mediators of technological change, possibly driven by the desire to free up time for prayer. In addition, the principle that abbots from all Cistercian monasteries had to congregate annually at Cîteaux (the founding monastery) may have implied that new successful innovations (even those not originally due to the Cistercians themselves) were quickly communicated and diffused. While we view technological change as a plausible consequence of the values proliferated by the Cistercians, it cannot be ruled out *a priori* that innovations are in fact the full story. That is, perhaps the legacy left by the Cistercians was technological, and not cultural.

The question is therefore whether the following causal chain, ignoring any cultural effects, may in fact account for our previous findings:

$$\text{Cistercians}_t \Rightarrow \text{Technology}_t \Rightarrow \text{Technology}_{t+T} \Rightarrow \text{Population density}_{t+T}.$$

A theoretical difficulty with this account of our findings above is that the Cistercian impact exhibits such strong persistence. The Monasteries were all dissolved during the fourth decade of the 16th century; nevertheless, we find an effect on population growth reaching as far as the beginning of the 19th century. This requires strong persistence in technology across counties, as reflected in the postulated link between Technology_t and Technology_{t+T}. Recent empirical evidence does in fact suggest that 20th century cross-country technology diffusion may be slow when the intensive margin is considered.[30] However, by any stretch of imagination, it seems hard believe that pre-industrial technological diffusion *within* a country could proceed sufficiently slow so as to account for our results spanning several centuries after the Dissolution of the Monasteries, without some additional source of persistence. Nevertheless, it is worth examining the issue in some detail.

As a first exercise we attempt to control for a key technology that the Cistercians were renowned for diffusing and improving: watermills. If "technology" itself is the main "story", while the cultural channel which we have emphasized all along is largely irrelevant, one would expect the impact of the Cistercian presence to disappear (or at least diminish in a major way) once we control for the intensity of watermills at the county level.

We obtained data on the location of watermills during the 14th century from Campbell and Bartley (2007). Naturally, watermills are seldom preserved across the centuries. Hence, the data originates from records of so-called inquisitions post mortem (IPM) following the passing of lay tenants in chief of the Crown. Specifically, we assign intensity of watermill presence according to the fraction of the IPMs where watermills are mentioned. It should be understood that this is a very crude proxy for the presence of watermills

[30] That is, diffusion appears to be slow when one does not simply consider whether a particular technology is present or not (the extensive margin), but rather focus on the intensity of its use in an area. See Comin et al. (2008).

in local areas; in particular, it excludes watermills belonging to the Church. Nevertheless, if Cistercians were instrumental in increasing the intensity of this important technology prior to the Dissolution of the Monasteries, this variable may still serve as a useful proxy for this technology dimension.

Table 7 shows the effect of introducing watermills. Columns 1-9 reproduce Table 3 with watermill intensity as an added control, whereas column 10 shows the results when all controls, including watermills, are added simultaneously. The recurring theme of the table is that the impact from Cistercian presence is unaffected by the inclusion of the intensity of watermill presence. In fact, comparing the results of Table 7 to those of Table 3 makes clear that the OLS estimate from Cistercian presence is essentially unaltered by the inclusion of watermills. This is fully consistent with the cultural theory proposed above, while hard to reconcile with the notion that technological change in itself accounts for the relationship between Cistercian presence and growth.

[Table 7 about here]

An alternative interpretation of these results is that our measure of technology, watermills, is simply too crude and imprecisely recorded. Hence, in an effort to pursue the matter a bit further, we introduce human capital. The motivation is twofold. First, it seems possible that if the Cistercians managed to instigate technological change during their tenure then this may have increased the return to skill accumulation in the local area. By accounting for the end result of early technological change, i.e. human capital accumulation, we may implicitly be controlling for early technological change more fully than what is admitted by the watermill control. Second, if initial technological change stimulated early human capital accumulation then the latter could be the source of persistence. Perhaps what our results above are suggesting is the following causal chain:

$$\text{Cistercians}_t \Rightarrow \text{Technology}_t \Rightarrow \text{Skill accumulation}_{t+T} \Rightarrow \text{Population density}_{t+T}.$$

That is, perhaps Cistercian presence stimulated early technological change, which led to comparative differences in the speed of skill accumulation in the centuries following the Dissolution of the Monasteries. Insofar as early skill accumulation led to greater earnings it could have worked to elevate population density during the ensuing centuries, prior to the demographic transition. Still, human capital accumulation might also lead to *slower* population growth due to the presence of a quantity-quality trade-off (e.g., Galor and Weil, 2000). Whether the above chain motivates our findings is an empirical matter. Accordingly, by including measures of historical human capital accumulation alongside Cistercian presence we may try to gauge the viability of the cultural change hypothesis.

We obtained data on county-level literacy rates from Hechter (2001). The earliest year from which data on literacy rates is available is 1851. Although this is after the "closing" of our observation window on population growth we hope that 1851 is a sensible proxy for comparative differences across English countries circa 1800 as well. Notice that by introducing the literacy rate at the end of the period we are in effect proxying the rate of human capital accumulation over the preceding centuries, assuming literacy rates were close to zero in 1377. Naturally, one might expect most of this change to have occurred towards the end of the period in question.

Table 8 provides the results from adding the literacy rate. In column 1 we add literacy to our baseline specification which involves (aside from the Cistercian share) the total number of religious houses, initial population density, and land quality. As can readily be seen, the Cistercian share remains significant in spite of the human capital control.

[Table 8 about here]

Human capital enters with a negative sign, suggesting that areas where individuals invested relatively more in their children, at least in terms of basic reading and writing skills, were the areas with the slowest growth in population between 1377 and 1801. This is consistent with an operative quantity-quality trade-off. Naturally, one cannot rule out that the correlation is due to reverse causality. That is, areas with fast population growth may have been areas that later (in 1851) ended up with lower human capital levels. Whatever the right interpretation, the main point is that controlling for human capital does *not* eliminate the significance of the Cistercian share. The basic pattern from the first column is repeated when we add our additional controls one by one, and when we add all controls simultaneously. In every case the Cistercian share remains significant at conventional levels of significance, and the size of the point estimate appears stable. These results are consistent with the theory advanced above, which proposes that the Cistercian influence on Pre-industrial growth in England did not manifest itself through accelerated human capital accumulation but rather via cultural change.

Finally, observe that the results reported in Table 8 also suggest that the impact of Cistercian presence cannot be accounted for by the mechanism featured in Becker and Woessmann (2009): The lasting impact of the Cistercians does not seem to have involved *all* the same proximate sources of growth as Protestantism. Human capital accumulation is unlikely to be part of the story in the present case.

4 Concluding Remarks

The present paper documents that Cistercian monks left a persistent imprint on long-run comparative development across English counties during the pre-industrial era. In counties with greater Cistercian presence population growth was faster during the period 1377-1801, suggesting that the Cistercians stimulated local earnings. The remarkable aspect of this finding is that the Catholic monasteries were dissolved by 1540. Hence the influence from the Order was felt more than 300 years after they had disappeared from England.

These results are robust to a considerable number of controls for productivity, and our IV estimates suggest the correlation can be given a casual interpretation.

We have also offered a potential explanation for these facts, namely that the Cistercians ignited a process of growth through cultural change. That is, a gradual change in local populations in terms of taste for thrift and hard work; much like what Max Weber suggested was the end result of the Protestant Reformation. We believe this theoretical explanation is plausible for two reasons. First, a cultural concordance between the Cistercians and the Protestants, in the dimensions of work ethic and thrift, has already been observed by several scholars including Weber himself. Second, the cultural explanation has the virtue of being able to plausibly account for the long-term persistency of Cistercian influence on growth. Consistent with the cultural mechanism we find, using data from the World Value Survey, that regions in England which historically were influenced relatively more by the Cistercians tend to have populations with greater taste for hard work and thrift today.

Naturally, there are other potentially viable explanations beyond cultural change. For instance, we have examined whether the above facts alternatively can be accounted for by technology, human capital or international trade. While all three channels may be plausible alternatives to the cultural mechanism, we find that none of them are able to account for the Cistercians influence in the data. As a result, we are led to the conclusion that the long term Cistercian impact was most likely caused by a change in cultural values, which stimulated earnings at the local level. Hence, our research suggest that the original Weber thesis, stressing the importance of cultural values like hard work and thrift to economic growth, holds considerable explanatory power with respect to the pre-industrial growth record of England.

A Appendix

A.1 Proof of Lemma 1

(i): Note first that $L_t \lesseqgtr \overline{L}$ iff $y_t \gtreqless (1+\omega)\, n\eta$. Consequently, with $L_t \leq \overline{L}$ we have by (4) that $\sigma_t = 1$ in which case (7) gives that $L_{t+1} = nL_t$. With $L_t > \overline{L}$ we have that $\sigma_y = \frac{c_t/n}{\eta}$. Inserting this, the appropriate c^* from (4), and equation (5) into equation (7) then gives $L_{t+1} = \frac{1}{1+\omega}\frac{1}{\eta}AL_t^{\alpha}X^{1-\alpha}$.

(ii): When $L_t \leq \overline{L}$ we have that $L_{t+1} = nL_t$, which cannot cross the 45-degree line in (L_t, L_{t+1})-space since $n > 1$. However, when $L_t > \overline{L}$, $G'(L_t) > 0$, $G''(L) < 0$, and $\lim_{L\to\infty} G'(L) = 0$. This ensures a unique and globally stable steady state. The steady state population, L^*, is found by solving $L^* = \frac{1}{1+\omega}\frac{1}{\eta}A(L^*)^{\alpha}X^{1-\alpha}$ for L^*.

(iii): Steady state income per capita is obtained by inserting steady state population, L^*, and $e^* = \frac{1}{1+\omega}$ into (6). ■

A.2 Proof of Proposition 1

Follows from differentiating the relevant expression from Lemma 1. ■

A.3 Proof of Lemma 2

(i) The law of motion for the first two regimes follows from straightforward application of Lemma 1 to the two groups: ω_1 and ω_2. The law of motion for the final regime, $\bar{L}_2 < L_t$, is obtained as follows. The time path for group size is given by

$$L_{i,t+1} = \frac{1}{1+\omega_i}\frac{1}{\eta}y_t L_{i,t} \qquad i = 1, 2.$$

Using $L_t = L_{1,t} + L_{2,t}$ and the production function, we obtain after some rearrangements

$$L_{t+1} = \left[\frac{\omega_1 - \omega_2}{(1+\omega_1)(1+\omega_2)}\frac{L_{2,t}}{L_t} + \frac{1}{1+\omega_1}\right]\frac{1}{\eta}AL_t^{\alpha}X^{1-\alpha}. \tag{10}$$

Observe now that $L_{2,t}/L_t = (1 + L_{1,t}/L_{2,t})^{-1}$. Since the laws of motions for the individual groups are symmetrical, save for the value of ω, we have that $L_{1,t+1}/L_{2,t+1} = [(1+\omega_2)/(1+\omega_1)]\, L_{1,t}/L_{2,t}$. Solving this difference equation yields $L_{1,t}/L_{2,t} = \left(\frac{1+\omega_2}{1+\omega_1}\right)^t L_{1,0}/L_{2,0}$. Substituting this solution into the law of motion for L, yields the expression stated in the Lemma.

(ii) Let group i grows according to

$$L_{it+1} = G_i L_{it} \tag{11}$$

where G_i, $i = 1, 2$ equals n or $\frac{1}{1+\omega_i}\frac{1}{\eta}y_t$ depending on L_t.

The law of motion for total population is given as $L_{1t+1} + L_{2t+1} = G_1 L_{1t} + G_2 L_{2t}$, and so

$$\frac{L_{t+1}}{L_t} = G_1\left(1 - \frac{L_{2t}}{L_t}\right) + G_2\frac{L_{2t}}{L_t} \Leftrightarrow \frac{L_{t+1}}{L_t} = G_1 + (G_2 - G_1)\frac{L_{2t}}{L_t}. \tag{12}$$

Note also that $\frac{L_{2t}}{L_t} = \left(1 + \frac{L_{1t}}{L_{2t}}\right)^{-1} \equiv \frac{1}{1+z_t}$. We can therefore write (12) as

$$\frac{L_{t+1}}{L_t} = G_1 + \frac{G_2 - G_1}{1 + z_t}. \tag{13}$$

Finally, note that by (11) we have

$$z_{t+1} = \frac{G_1}{G_2} z_t, \tag{14}$$

where we recall that G_1 and G_2 are functions of L_t.

In this setting, a *steady state* is a pair (L^*, z^*) such that $0 < L^* = L_{t+1} = L_t$ and $z_{t+1} = z_t = z^*$, with $0 \le z^* \le 1$, and where (L^*, z^*) fulfills equations (13), (14).

We need to consider existence of a steady state in the three regimes stated in the lemma.

Case I: $L < \bar{L}_1$, where (it is recalled) $\bar{L}_1$ is defined as $y(L_1) = (1 + \omega_1) n\eta$. When $0 < L < \bar{L}_1$, $G_1 = n$ and $G_2 = n$. Hence, relative group size (and thus $\frac{L_{2t}}{L_t}$) is constant. But the aggregate population is rising, since (inserting into (13)) $\frac{L_{t+1}}{L_t} = n > 1$. Accordingly, for $L < \bar{L}_1$ there does not exist a steady state with $L_{t+1} = L_t = L^*$ and L^* fulfilling (13).

Case II. $0 < \bar{L}_1 \le L_t < \bar{L}_2$, with $\bar{L}_2$ is defined as $y\left(\bar{L}_2\right) = (1 + \omega_2) n\eta$. Note that $\bar{L}_2 > \bar{L}_1$ since $\omega_2 < \omega_1$. Relative group size in this interval.

$$z_{t+1} = \frac{G_1}{G_2} = \frac{\frac{1}{1+\omega_1}\frac{1}{\eta}y_t}{n} z_t \tag{15}$$

while aggregate growth is given by (13).

Working towards a contradiction, assume that a steady state exist. This requires (from (13)) that

$$1 = G_1(L^*) + (G_2(L^*) - G_1(L^*))\frac{1}{1 + z^*}.$$

or, since $\bar{L}_1 \le L_t < \bar{L}_2$,

$$\frac{1 - \frac{1}{1+\omega_1}\frac{1}{\eta}y(L^*)}{n - \frac{1}{1+\omega_1}\frac{1}{\eta}y(L^*)} = \frac{1}{1 + z^*}. \tag{16}$$

That is, existence requires constancy of relative group size z^*. But this demands, from (15), that

$$n = \frac{1}{1 + \omega_1}\frac{1}{\eta}y(L^*),$$

which contradicts $0 \leq z^* \leq 1$, as seen from (16). Hence, a steady state cannot exist in the interval $0 < \bar{L}_1 \leq L_t < \bar{L}_2$.

Case III. $\bar{L}_2 < L_t$. Observe that $G_t' > 0$ for $L_{1,0}$, $L_{2,0} > 0$. This follows from differentiation and noting that $\omega_2 < \omega_1$. Next, note that $\lim_{t\to\infty} \left(1 + \left(\frac{1+\omega_2}{1+\omega_1}\right)^t \frac{L_{1,0}}{L_{2,0}}\right)^{-1} = 1$, implying $\lim_{t\to\infty} G = \frac{1}{1+\omega_2}\frac{1}{\eta}AL_t^\alpha X^{1-\alpha}$. Given this result and continuity of $G(L_t, t\ )$, the same considerations as those laid out in the proof of Lemma 1 leads to existence and uniqueness of the steady state.■

A.4 Proof of Proposition 2

Before cultural mutation occurs the law of motion for population size is given by Lemma 1, with $\omega = \omega_1$. After the change, the law of motion is given in Lemma 2. Since $\omega_2 < \omega_1$ Proposition 2 follows from Proposition 1 and Lemma 2.■

References

[1] Aiyar, S., C.-J. Dalgaard and O. Moav, 2008. Technological Progress and Regress in Pre-Industrial Times. *Journal of Economic Growth*, 13, 125-44

[2] Alcalá, F. and A. Ciccone, 2004. Trade and Productivity. *Quarterly Journal of Economics,* 119, 613-646.

[3] Allen, R., 2009. *The British Industrial Revolution in Global Perspective.* Cambridge University Press

[4] Andersen, T.B. and C-J. Dalgaard, 2011. Flows of People, Flows of Ideas, and the Inequality of Nations. *Journal of Economic Growth*, forthcoming

[5] Angrist, J. and J.-S. Pischke, 2009. *Mostly Harmless Econometrics.* Princeton University Press

[6] Ashraf, Q. and O. Galor, 2009. Dynamics and Stagnation in the Malthusian Epoch: Theory and Evidence. Working Paper (Brown University). *American Economic Review* (forthcoming)

[7] Barro, R. and R. McCleary, 2003. Religion and Economic Growth across Countries. *American Sociological Review*, 68, 760-781

[8] Baumol, W., 1990. Entrepreneurship: Productive, Unproductive, and Destructive. *Journal of Political Economy*, 98, 893-921

[9] Bazeley, M.L., 1921. The Extent of the English Forest in the Thirteenth Century. Transactions of the Royal Historical Society, 4 , pp. 140-172.

[10] Becker, G.S., 1973. A Theory of Marriage: Part I. *Journal of Political Economy* 81, p.881-46.

[11] Becker, R.A., 1980. On the long-run steady state in a simple dynamic model of equilibrium with heterogeneous households. *Quarterly Journal of Economics*, 95, 375–82

[12] Becker, S. and L. Woessmann, 2009. Was Weber Wrong? A Human Capital Theory of Protestant Economic History. *Quarterly Journal of Economics*, 124, 531-596

[13] Berman, C.H., 2000, *The Cistercian Evolution.* Penn Press

[14] Bloch, M., 1935. *Avènement et conquêtes du moulin a eau. Annales d'Histoire Èconomique et Sociale,* 7, 538–563

[15] Bouchard, C., 1991. *Holy Entrepreneurs.* Cornell University Press

[16] Cameron, C. and P. Trivedi, 2005. *Microeconometrics.* Cambridge University Press

[17] Campbell, B., 2000. *English seigniorial agriculture, 1250-1450.* Cambridge University Press

[18] Campbell, B., 2008. Benchmarking medieval economic development: England, Wales, Scotland and Ireland, c. 1290. *Economic History Review*, 61, 896-945

[19] Campbell B. and K. Bartley, 2007. *England on the Eve of the Black Death: An Atlas of Lay Lordship, Land and Wealth, 1300-49*. Manchester University Press

[20] Cantoni, D., 2009. The Economic Effects of the Protestant Reformation: Testing the Weber Hypothesis in the German Lands. Manuscript (Harvard University)

[21] Cavalcanti, T.V., Parente, S.L., and R. Zhao, 2007. Religion in macroeconomics: a quantitative analysis of Weber's thesis. *Economic Theory*, 32, 105–123

[22] Clark, P., 1979. Migration in England during the Late Seventeenth and Early Eighteenth Centuries. *Past & Present*, 83, 57-90

[23] Clark, G., 2007. *A Farewell to Alms*. Princeton University Press

[24] Clark, G. and Y. Van Der Werf, 1998. Work in progress? the industrious revolution, *Journal of Economic History*, 58, 830–43

[25] Collins, R., 1986. *Weberian Sociological Theory*. Cambridge University Press

[26] Comin, D. and Hobijn, B. and Rovito, E., 2008. Technology usage lags. *Journal of Economic Growth*, 13, 237–256

[27] Cook, H., Stearne, K., and T. Williamson, 2003. The origins of water meadows in England. *Agricultural History Review*, 51, 155-162

[28] Cooke, A., 1893. The Settlement of the Cistercians in England. *English Historical Review*, 8, 625-676

[29] Dobson, R., 1983. *The Peasants' Revolt*. Palgrave Macmillan

[30] Doepke, M. and F. Zilibotti, 2008. Occupational choice and the spirit of capitalism. *Quarterly Journal of Economics*, 747-93

[31] Donkin, R., 1963. The Cistercian Order in Medieval England: Some Conclusions. *Transactions and Papers (Institute of British Geographers)*, 33, 181–198

[32] Donkin, R., 1978. *The Cistercians: Studies in the geography of medieval England and Wales*. Toronto: Pontifical Institute of Medieval Studies

[33] Galor, O. and D. Weil, 2000. Population, technology, and growth: From Malthusian stagnation to the demographic transition and beyond. *American Economic Review*, 90, 806–828.

[34] Galor, O. and O. Moav, 2002. Natural Selection and the Origin of Economic Growth. *Quarterly Journal of Economics*, 117, 1133–91

[35] Gimpel, J., 1976. *The medieval machine*. Penguin

[36] Grant, R. (1991). *The Royal Forests of England*. Stroud, UK: Alan Sutton Publishing

[37] Guiso, L., Sapienza, P., and L. Zingales, 2008. Long Term Persistence. EUI Working Paper 2008/30

[38] Hill, B.D., 1968. *English Cistercian monasteries and their patrons in the twelfth century*. University of Illinois Press

[39] Kaelber L., 1998. *Schools of Asceticism: Ideology and Organization in Medieval Religious Communities*. Penn State Press

[40] Kelly, M. and C. O'Grada, 2008. Living Standards and Mortality in England since the Middle Ages: The Poor Law versus the Positive Check. Manuscript (University College Dublin)

[41] Kieser, A., 1987. From Asceticism to Administration of Wealth. Medieval Monasteries and the Pitfalls of Rationalization. *Organization Studies*, 8, 103-123

[42] Landes, D., 1999. The Wealth and Poverty of Nations. Abacus

[43] Madden, J.E. (1963). Business Monks, Banker Monks, Bankrupt Monks: The English Cistercians in the Thirteenth Century. *The Catholic Historical Review*, Vol. 49, No. 3, pp. 341-364

[44] McCleary, R. and R. Barro, 2006. Religion and Economy. *Journal of Economic Perspectives*, 20, 49-72

[45] Møller, N. and P. Sharp, 2009. Malthus in Cointegration Space: A new look at living standards and population in pre-industrial England. Discussion Paper No. 08-16, Department of Economics, University of Copenhagen

[46] Nicholas, S. and P.R. Shergold, 1987, Intercounty labour Mobility during the Industrial Revolution: Evidence from Australian Transportation Records. *Oxford Economic Papers*, 39, 624-40

[47] Noell, B., 2006. Expectation and unrest among Cistercian lay brothers in the twelfth and thirteenth centuries. *Journal of Medieval History*, 32, 253-274

[48] Nunn, N. and L. Wantchekon, 2009. The Slave Trade and the Origins of Mistrust in Africa. Manuscript (Harvard University)

[49] Ovitt Jr, G., 1986. The Cultural Context of Western Technology: Early Christian Attitudes toward Manual Labor. *Technology and Culture*, 27, 477–500

[50] Ovitt Jr, G., 1987. *The Restoration of Perfection: Labor and Technology in Medieval Culture*. Rutgers University Press

[51] Pomeranz, K., 2000. *The Great Divergence: China, Europe, and the Making of the Modern World Economy*. Princeton University Press

[52] Southern, R.W., 1970. *Western Society and the Church in the Middle Ages*. Penguin

[53] Tabellini, G., 2010. Culture and institutions: economic development in the regions of Europe. *Journal of the European Economic Association,* 8, 677-716

[54] Young, C.R. (1978). 'Conservation Policies in the Royal Forests of Medieval England'. *Albion: A Quarterly Journal Concerned with British Studies*, 10, 95-103

[55] Weber, M., 1905. *The Protestant Ethic and the Spirit of Capitalism.* Allen & Unwin

[56] Weber, M., 1958. *The Protestant Ethic and the Spirit of Capitalism.* Charles Scribner's Sons

[57] Williams, D.H., 1970. *The Welsh Cistercians: Aspects of Their Economic History.* The Griffin Press

[58] Wrigley, E., 2007. English county populations in the later eighteenth century. *Economic History Review*, 60, 35-69

Table 1: Frequency distribution of monastic orders	Number	Share of total
Benedictine monks	239	30.8
Augustinian canons	208	26.8
Benedictine nuns	77	9.9
Cistercian monks	70	9.0
Premonstratensian canons	37	4.8
Cluniac monks	34	4.4
Cistercian nuns	28	3.6
Augustinian canonesses	24	3.1
Gilbertine canons	15	1.9
Trinitarian brothers	10	1.3
Gilbertine canons & nuns	9	1.2
Carthusian monks	8	1.0
Fontevraud nuns	3	0.4
Grandmontine monks	3	0.4
Premonstratensian canonesses	3	0.4
Bonhommes brothers	2	0.3
Cluniac nuns	2	0.3
Brigettine nuns & brothers	1	0.1
Fontevraud monks	1	0.1
Gilbertine nuns	1	0.1
unknown monks or brothers	1	0.1
Total	776	100.0

Table 2a: Summary statistics	Obs	Mean	Std.	Min	Max
cistercianshare	40	0.09	0.07	0.00	0.25
relhouses	40	19.03	12.93	2.00	73.00
popdens1377	40	31.55	11.83	8.98	52.98
popdens1801	40	60.45	24.82	20.92	143.77
agrquality1_2	40	0.18	0.16	0.01	0.73
riverstream	40	0.41	0.29	0.06	1.09
rivershare	40	0.12	0.03	0.03	0.18
streamshare	40	0.30	0.27	0.01	0.91
area	40	3256.94	1623.37	392.20	7423.75

Table 2b: Correlation matrix								
cistercianshare	1							
relhouses	0.0134	1						
popdens1377	-0.2791	0.3091	1					
popdens1801	0.3161	-0.1644	-0.0579	1				
agrquality1_2	-0.1234	0.2460	0.4287	-0.2149	1			
riverstream	0.0127	-0.1391	-0.6152	-0.1460	-0.2602	1		
rivershare	0.1589	-0.1271	-0.4816	-0.0063	-0.2092	0.6345	1	
streamshare	-0.0044	-0.1341	-0.6025	-0.1553	-0.2542	0.9962	0.5644	1
area	0.2625	0.6603	-0.2802	-0.0381	-0.0865	0.3191	0.2011	0.3181

Table 3: OLS estimation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					Dependent variable : dlogpop1377_1801					
logpopdens1377	-0.614***	-0.694***	-0.606***	-0.700***	-0.438**	-0.715***	-0.615***	-0.635***	-0.674***	-0.729***
	(0.171)	(0.207)	(0.186)	(0.206)	(0.173)	(0.172)	(0.169)	(0.155)	(0.160)	(0.249)
cistercianshare	1.934**	1.802**	1.931**	1.779**	1.751**	1.911**	1.982**	2.074**	1.666*	1.773*
	(0.887)	(0.825)	(0.916)	(0.827)	(0.798)	(0.847)	(0.927)	(0.876)	(0.852)	(0.968)
relhouses	-0.007*	-0.007+	-0.007*	-0.006+	-0.017***	-0.005	-0.007*	-0.008**	-0.006*	-0.014**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.005)
agrquality1_2	-0.634*	-0.633*	-0.634*	-0.632*	-0.536*	-0.543*	-0.631*	-0.488+	-0.675**	-0.279
	(0.313)	(0.321)	(0.315)	(0.320)	(0.297)	(0.288)	(0.326)	(0.301)	(0.326)	(0.324)
riverstream		-0.180								
		(0.233)								
rivershare			0.267							1.481
			(1.940)							(1.987)
streamshare				-0.208						-0.473
				(0.247)						(0.293)
logarea					0.303**					0.280
					(0.113)					(0.176)
augustinianshare						0.771				0.970
						(0.541)				(0.589)
benedictineshare							0.146			0.629
							(0.383)			(0.479)
cluniacshare								2.465**		1.848
								(1.159)		(1.486)
premonshare									-0.674	0.245
									(0.847)	(1.088)
Constant	2.799***	3.148***	2.742***	3.158***	-0.011	2.878***	2.754***	2.744***	3.060***	0.472
	(0.587)	(0.742)	(0.735)	(0.724)	(1.257)	(0.548)	(0.631)	(0.542)	(0.529)	(2.08)
Observations	40	40	40	40	40	40	40	40	40	40
R-squared	0.641	0.646	0.641	0.647	0.676	0.667	0.642	0.683	0.650	0.756

Notes. Robust standard errors in parentheses. Symbols ***, **, * indicate significance at the 1, 5, and 10%, respectively. In column 10, riverstream is omitted due to perfect multicollinearity.

Table 4: OLS estimation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					Dependent variable : dlogpop1377_1801					
logpopdens1377	-0.482**	-0.606**	-0.460*	-0.619**	-0.124	-0.521**	-0.469*	-0.287	-0.519*	-0.359
	(0.235)	(0.290)	(0.267)	(0.287)	(0.266)	(0.233)	(0.232)	(0.248)	(0.282)	(0.339)
cistercianshare	1.977*	1.849*	1.975*	1.823*	1.810**	1.662*	1.947+	2.214**	1.659*	1.741
	(1.078)	(1.039)	(1.112)	(1.044)	(0.825)	(0.859)	(1.176)	(0.958)	(0.876)	(1.303)
relhouses	-0.006*	-0.006+	-0.006+	-0.006+	-0.023***	-0.003	-0.006*	-0.007**	-0.005	-0.015*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.006)	(0.003)	(0.004)	(0.003)	(0.004)	(0.009)
agrquality1_2	-0.682*	-0.640*	-0.671*	-0.626+	-0.694**	-0.519	-0.665*	-0.471	-0.717*	-0.309
	(0.350)	(0.366)	(0.352)	(0.368)	(0.299)	(0.390)	(0.340)	(0.426)	(0.361)	(0.586)
riverstream		-0.158								
		(0.286)								
rivershare			0.362							1.294
			(2.128)							(2.204)
streamshare				-0.188						-0.329
				(0.309)						(0.278)
logarea					0.518***					0.328
					(0.182)					(0.237)
augustinianshare						1.137*				1.056
						(0.564)				(0.835)
benedictineshare							-0.154			0.475
							(0.487)			(0.649)
cluniacshare								3.238*		1.853
								(1.653)		(1.861)
premonshare									-0.524	0.342
									(1.374)	(1.665)
Constant	2.295**	2.787**	2.169*	2.815**	-2.649	1.986*	2.288**	1.372	2.462**	-1.274
	(0.971)	(1.153)	(1.200)	(1.126)	(2.151)	(0.977)	(0.980)	(1.072)	(1.152)	(2.709)
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F test, H0: Reg. dummies = 0 (p-val)	0.487	0.658	0.510	0.668	0.070	0.374	0.558	0.354	0.616	0.562
Observations	40	40	40	40	40	40	40	40	40	40
R-squared	0.678	0.681	0.679	0.682	0.750	0.723	0.680	0.737	0.682	0.799

Notes. Robust standard errors in parentheses. Symbols ***, **, *, + indicate significance at the 1, 5, 10, and 15%, respectively. In column 10, riverstream is omitted due to perfect multicollinearity.

Table 5: IV (Limited Information Maximum Likelihood) estimation

	IV		
	(1)	(2)	(3)
	Dependent variable : dlogpop1377_1801		
Second stage			
cistercianshare	4.77**	3.37**	2.57**
	-2.07	(1.54)	(1.19)
	Dependent variable: cistercianshare		
First stage			
Rforest	0.08**	0.10***	0.07**
	(0.04)	(0.02)	(0.03)
logpopdens1300	- 0.14**	- 0.17**	- 0.20***
	(0.07)	(0.06)	(0.07)
First stage F-static	4.81	15.68	7.49
Hansen J-stat (p-value)	0.55	0.46	0.75
Additional Controls	logpopdens1377	All Baseline	All controls
Observations	40	40	40

Notes. Robust standard errors in parentheses. Symbols ***, **, * indicate significance at the 1, 5, and 10%, respectively. All regressions include a constant. "Baseline controls" is controls from Table 3, Column 1: logpopdens1377, religious houses and agricultural land quality. "All controls" involve all the controls featured in Table 3: logpopdens1377, religious houses (total), agricultural land quality, rivershare, streamshare, logarea, augustinianshare, clunicshare, and premonshare. To avoid multicollinarity river and stream shares are controled for seperately, while the combined variable "riverstream" is ignored.

Table 6: OLS estimation

	Trade									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Dependent variable : dlogpop1377_1801									
cistercianshare	1.92**	1.77**	1.92**	1.74**	1.77**	1.90**	1.97**	2.08**	1.66*	2.02*
	(0.88)	(0.80)	(0.91)	(0.80)	(0.82)	(0.85)	(0.92)	(0.89)	(0.87)	(1.05)
coastal dummy	0.03	0.03	0.03	0.05	-0.11	0.01	0.01	-0.01	-0.01	-0.25*
	(0.13)	(0.14)	(0.13)	(0.14)	(0.13)	(0.11)	(0.12)	(0.13)	(0.13)	(0.14)
Controls	All Baseline	All Baseline + riverstream	All Baseline + rivershare	All Baseline + streamshare	All Baseline + logarea	All Baseline + augustinianshare	All Baseline + benedictineshare	All Baseline + cluniacshare	All Baseline + premonshare	All controls
Observations	40	40	40	40	40	40	40	40	40	40
R-squared	0.64	0.65	0.64	0.65	0.68	0.67	0.64	0.68	0.65	0.78

Notes. Robust standard errors in parentheses. Symbols ***, **, * indicate significance at the 1, 5, and 10%, respectively. All regressions include a constant. "Baseline controls" are the controls from Table 3, Column 1: logpopdens1377, religious houses, and agricultural land quality. "All controls" means all the controls featured in Table 3: logpopdens1377, religious houses, agricultural land quality, rivershare, streamshare, logarea, augstinianshare, benedictineshare, cluniacshare, and premonshare.

Table 7: OLS estimation

						Technology				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	[10]
	Dependent variable : dlogpop1377_1801									
cistercianshare	1.92**	1.78**	1.92**	1.76**	1.75**	1.90**	1.98**	2.07**	1.58*	1.79*
	(0.89)	(0.83)	(0.91)	(0.83)	(0.82)	(0.85)	(0.91)	(0.90)	(0.82)	(1.00)
watermills	0.22	0.24	0.22	0.22	0.26	0.14	0.30	0.05	0.39	-0.18
	(0.64)	(0.83)	(0.67)	(0.65)	(0.66)	(65)	(0.70)	(0.61)	(0.64)	(0.60)
Controls	All Baseline	All Baseline + riverstream	All Baseline + rivershare	All Baseline + streamshare	All Baseline + logarea	All Baseline + augustinianshare	All Baseline + Benedictineshare	All Baseline + cluniacshare	All Baseline + premonshare	All controls
Observations	40	40	40	40	40	40	40	40	40	40
R-squared	0.64	0.65	0.64	0.65	0.68	0.67	0.65	0.68	0.66	0.76

Notes. Robust standard errors in parentheses. Symbols ***, **, * indicate significance at the 1, 5, and 10%, respectively. All regressions include a constant. "Baseline controls" are the controls from Table 3, Column 1: logpopdens1377, religious houses and agricultural land quality. "All controls" means all the controls featured in Table 3: logpopdens1377, religious houses and agricultural land quality, rivershare, streamshare, logarea, augstinianshare, benedictineshare, clunyshare, and premonshare.

Table 8: OLS estimation — ***Human capital***

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Dependent variable : dlogpop1377_1801									
cistercianshare	1.64**	1.58**	1.64**	1.58*	1.53**	1.61**	1.64*	1.79**	1.57*	1.76**
	(0.79)	(0.79)	(0.79)	(0.79)	(0.70)	(0.76)	(0.84)	(0.83)	(0.79)	(0.82)
literacy1851	- 1.93**	- 1.86**	- 1.93**	- 1.86**	- 1.94**	- 1.96**	- 1.93**	- 1.64*	-1.84*	- 1.69*
	(0.91)	(0.88)	(0.94)	(0.88)	(0.80)	(0.93)	(0.94)	(0.83)	(1.07)	(0.97)
Controls	All Baseline	All Baseline + riverstream	All Baseline + rivershare	All Baseline + streamshare	All Baseline + logarea	All Baseline + augustinianshare	All Baseline + benedictineshare	All Baseline + cluniacshare	All Baseline + premonshare	All controls
Observations	37	37	37	37	37	37	37	37	37	37
R-squared	0.69	0.69	0.69	0.69	0.72	0.71	0.69	0.70	0.69	0.77

Notes. Robust standard errors in parentheses. Symbols ***, **, * indicate significance at the 1, 5, and 10%, respectively. All regressions include a constant. "Baseline controls" are the controls from Table 3, Column 1: logpopdens1377, religious houses, and agricultural land quality. "All controls" means all the controls featured in Table 3: logpopdens1377, religious houses, agricultural land quality, rivershare, streamshare, logarea, augstinianshare, benedictineshare, cluniacshare, and premonshare.

Chapter 3

How Bad is Corruption? Cross-Country Evidence of the Impact of Corruption on Economic Prosperity

Jeanet Sinding Bentzen

How Bad is Corruption? Cross-Country Evidence of the Impact of Corruption on Economic Prosperity

Jeanet Bentzen*

September 27, 2011

Abstract

Most people today would argue that corruption is bad for countries' economic development. Yet, we still lack a reliable empirical estimate of the effect. This study addresses the econometric shortcomings of the literature and provides an estimate of the causal impact of corruption on GDP per capita across countries. I use certain dimensions of a country's culture as instruments for corruption. These instruments stay strong when I control for the other deep determinants of economic development; geography and the remaining dimensions of institutions and culture. I tie my hands in the process of choosing controls by including the entire set of variables available in the online database QoG that includes all central variables from the literature on institutions and culture. I find that corruption *does* exert a significant and negative impact on countries' productivity levels.

1. Introduction

Most of us agree that corruption is a government failure that one would like to get rid of. (see Leff, 1964 for a counter argument). But how bad is it really? It is important to get a sense of the severity of the problem in order to be able to prioritize which government inefficiencies to fight first. Mauro (1995) provided an empirical estimate of the impact of corruption on investments and growth. Using the instrument variables approach, he aimed at solving the obvious endogeneity issues: Not only does economic performance influence corruption, but countries performing better might also

* Affiliated with the Department of Economics, University of Copenhagen, Øster Farimagsgade 5, building 26, 1353 Copenhagen, Denmark. Contact: Jeanet.bentzen@econ.ku.dk. I thank Thomas Barnebeck Andersen, Christian Sinding Bentzen, Carl-Johan Dalgaard, Joannes Jacobsen, Peter Norman Sørensen, an anononymous referee, and seminar participants at the PhD seminar at the Department of Economics, University of Copenhagen for useful comments and suggestions. Errors are my own.

have more resources to combat corruption. Also, other factors might influence both corruption and economic performance. Mauro found that corruption reduces investments and thereby economic growth. However, Mauro's study has two econometric shortcomings that the present paper tries to deal with.

Mauro's instrument for corruption was ethnolinguistic fractionalization (ELF). The idea is that bureaucrats may favor members of their own ethnolinguistic group and corruption is expected to be higher the higher the degree of ELF in a country. However, the assumption of exclusion restrictions does not hold: ELF might be correlated with other determinants of growth that are excluded from the analysis (see also Acemoglu et al., 2001). For instance, Easterly and Levine (1997) argue that ELF can influence performance by creating political instability. La Porta et al. (1999) show that countries closer to the Equator are more ethnolinguistically fractionalized than countries further away. If e.g. political instability and distance to the Equator influence growth, ELF is an invalid instrument. Indeed, it seems that these factors do influence growth (e.g. Gallup et al., 1999, Sachs, 2003, Easterly and Levine, 1997).

To alleviate this problem, one could include controls for geography and political instability in Mauro's regressions. However, then arises the second econometric problem: ELF becomes a weak instrument when including other determinants of economic outcomes in Mauro's regressions.

I attempt to address these issues by providing new instruments for corruption, namely cultural values of individualism and power distances. The idea is that corruption tends to be higher in cultures that encourage individuals to prioritize loyalty towards ones social group above the individual itself (e.g. Rose-Ackerman, 1999). I.e. in cultures with lower rates of Individualism. Another dimension of culture that influences the level of corruption is what Hofstede (2001) terms Power Distances (termed hierarchy by Schwartz, 1999, 2004). A worker in a culture with large power distances prefers large distances between the people in power and himself. These societies end up being more hierarchical and also tend to be more corrupt. I use these two dimensions of culture as instruments for corruption in cross-country regressions on GDP per capita.

However, these culture dimensions might be correlated with other factors that influence GDP per capita. This would render them invalid as instruments (the exclusion restrictions would not hold). Since I have two instruments, I can use the OID test to test this. In all regressions, this test accepts that corruption is the only channel through which these particular culture dimensions influence GDP per capita. While this is encouraging, it does not prove anything, since the OID test is a test of low power, meaning that it accepts "too many" wrong hypotheses (e.g. Wooldridge, 2002). Hence, whether the effect of corruption on economic prosperity can be estimated, hinges on whether it is possible to control for all relevant covariates. Note that it does not matter for the validity of the instruments that I have omitted factors influencing culture, as long as these factors do not influence GDP per capita. Hence, if I can control for everything with a bearing on GDP per capita, the exclusion restrictions hold. In order to choose these relevant controls, I rely on the framework of "deep determinants" of economic prosperity (e.g. Hall and Jones, 1999, Acemoglu et al., 2001, Rodrik, 2004, and Licht et al., 2007). The idea is that deep determinants of development such as institutions, geography, and culture affect proximate factors such as technological progress and investments in human and physical capital, which then determine the level of GDP per capita. Applied to the present paper; if the framework of deep determinants captures the entire variation in economic development, I can account for any link between culture and economic development, provided that I possess measures of all these deep determinants. In other words; if institutions, geography, and culture exhaust the list of possible deep determinants, what is left for me to do is to find good measures of these dimensions.

In order to reduce the degrees of freedom in choosing these control variables, I have included all measures of institutions and culture available in a large online dataset provided by The Quality of Government Institute, encompassing all central measures from the literature (Teorell et al., 2009).

Another concern, that one might have, is that the instruments could be influenced by the level of GDP per capita. However, this should not be a problem when using these particular culture measures, since they are from 1967 and 1973 (with an extension in 2001), whereas GDP and corruption are from 2006. Furthermore, Hofstede's culture measures are based on interviews of a group of people rather similar across countries, namely IBM workers. This makes the measures not so representative for a particular country's culture, but very suitable for value comparisons *across* countries, as they are rather independent of demographics, income etc. Nevertheless, I provide empirical evidence showing that culture is, indeed, exogenous to the level of economic development. Specifically, I use an instrument for Individualism provided by Licht et al. (2007); pronoun drop. The reasoning is that the language of less individualist cultures evolves to drop the pronoun (I, she, they etc.), since the individual is less important compared to more individualist countries.

My study differs from Mauro's in one last aspect. By linking corruption to *growth rates* of GDP per capita, Mauro explains the transitional dynamics towards steady state. Instead, Hall and Jones (1999) stress that the *level* of GDP per capita should be employed if the purpose is to explain the long run differences in economic performance (see also Acemoglu et al., 2001 and Rodrik et al., 2004). As this paper analyses the level of GDP per capita, the conclusions thus concern long run differences and not short run dynamics.

The paper is structured as follows. The following section outlines the framework of deep determinants of economic development in order to set up the regression equation. Section 3 introduces the data. The empirical analysis is performed in Sections 4 and 5; Section 4 providing the OLS analysis and Section 5 the IV analysis together with robustness checks. Section 6 concludes.

2. Framework: Deep Determinants of Economic Prosperity

I follow a broad literature seeking the fundamental causes of long run development (see Rodrik et al., 2004 for an overview). According to neoclassical growth theory, cross-country income differences arise from a combination of differences in the rates of technological progress and physical and human capital investment. More recently, authors have raised the deeper question: Why is it that some countries do not improve their technology, invest more in physical capital, and accumulate more human capital? (e.g. Hall and Jones, 1999) There must be some reasons underneath that prevent countries from improving the proximate causes of economic development. The literature has broadly agreed on three such reasons: Institutions (corruption, contracting institutions, and property rights institutions; Hall and Jones, 1999, Acemoglu et al., 2001), geography (Sachs, 2003), and culture (Tabellini, 2008).[1] This framework is consistent with a regression line of the form:[2]

$$GDPCAP_i = \beta_0 + \beta_1 CORRUPT_i + \beta_2 GEO_i + \beta_3 CULTURE_i + \beta_4 INST_i + \varepsilon_i, \qquad (1)$$

where β_1 is the parameter of interest in the present paper. I have split up institutions into corruption, CORRUPT, and remaining institutions, INST. $GDPCAP_i$ is real GDP per capita in country i, GEO is a vector of geographical factors, and CULTURE is a vector of culture dimensions.

If institutions, geography, and culture span the entire set of deep determinants of economic prosperity, regressions of type (1) will explain the entire variation in GDP per capita across countries. Hence, in the perfect world, where we have perfect data, we would expect an R^2 of 100%. This is obviously a rather naïve assumption, but in the empirical analysis I actually get an explanatory power of the model of around 80%. If the “deep determinants” framework is true, the deviation from 100% is simply due to measurement error.

3. Data

GDP per capita. GDP per capita is from Heston et al.’s (2009) Penn world Tables. I use real GDP per capita from PWT version 6.3 (the measure called rgdpl2) in 2006. 2006 is not the most recent year, but I expect that the most recent year might be subject to more measurement error.

Corruption. The measure of corruption used in the main regressions is control of corruption in 2006 (cci2006) by Kaufmann et al. (2009). The index runs from -2.5 to 2.5, where lower values indicate less control of corruption, meaning more corruption. This index includes various measures of corruption, which might reduce the scope for measurement error.

Remaining institutions. The Quality of Governance (QOG) dataset by Teorell et al. (2009) provides a long list of measures of institutions, gathered from some of the most central research.

Acemoglu and Johnson (2005) define institutions as encompassing property rights institutions and contracting institutions. Some authors also include rule of law (e.g. Licht et al., 2007). Therefore, I have included all institutional measures from QOG, except measures of labor market institutions, which gives me a total of 23 measures in addition to corruption (described in Web appendix A1, available on my webpage). As contracting institutions were not included in the QOG dataset, I added two measures of contracting institutions used by Acemoglu and Johnson (2005); "number of procedures" and "procedural complexity", also described in Web appendix A1.

Geography. Regional dummies capture a large share of the variation in geography. The ones used here are by World Development Indicators: Sub Saharan Africa (SSA), South Asia (SOA), East Asia & Pacific (EAP), Latin America & Caribbean (LAC), Middle East & North Africa (MENA), North America (NA), and Europe & Central Asia (ECA). The results are unchanged when using instead continent dummies. Other measures of geography: Absolute latitude, landarea in km^2, a dummy for whether the country is landlocked, the share of the total area that lies within the tropics, a dummy for whether some part of the country lies within the tropics, elevation (mean meters above sea level), mean distance to coast, mean distance to river, mean distance to coast or river.

Hofstede culture measures. The measures of culture that I use as instruments for corruption are from Hofstede (2001).[3] Hofstede performed value surveys of a group of people, similar across countries, namely IBM employees. This makes his survey suitable for value comparisons *across* countries, as opposed to studies like World Values Survey, which are more suitable for obtaining representative measures of culture *within* a country. From the value surveys, Hofstede developed a model that identifies five primary dimensions to differentiate cultures. Power Distances (PDI): the degree to which inequality in the distribution of power is accepted and expected in a society. Individualism (IDV): the extent to which a society reinforces individual or collective achievement. The three remaining dimensions are described in Web Appendix A1 on my webpage. The value surveys of the IBM workers were performed in 1967 and 1973 with an extension in 2001.

Remaining culture. In addition to the Hofstede measures, I gathered 26 measures of culture from the QOG dataset. I added all religious, language, and colonial origin variables, which have been argued to be proxies for culture. All described in Web Appendix A1 on my webpage.

4. OLS Estimation

Panel A of Table 1 provides OLS estimates of equation (1), regressing 2006 real GDP per capita on corruption and the various deep determinants.

In column (1), control of corruption is the only explanatory variable, exerting a significant and positive impact on GDP per capita, meaning that the impact of corruption is negative. Obviously, this simple correlation is biased. First of all; omitted factors potentially influence both corruption and GDP. A way to eliminate this particular bias is to use the deep determinants framework to include all factors that affect GDP per capita. Column (2) includes regional dummies, which should pick up a large share of the variation in GDP due to geographic differences. Corruption exerts a significant and negative impact on GDP per capita; taking the OLS estimate of column (2) at face value, a one unit increase in corruption (standard error 0.99) reduces real GDP per capita by 9,223 US $. The AV-plot corresponding to column (2) is shown in Figure 1. No single group of countries seems to be driving the result.[4]

Another way to account for omitted variables is TSLS estimation, which also eliminates bias due to reverse causality. TSLS is performed in panel B of Table 1, which I will comment on in section 5 below.

Table 1 about here

Figure 1 about here

Columns (3) through (6) of Panel A include the remainder of the three groups of deep determinants as control variables: Geography, institutions, and culture. There are numerous measures of these three determinants, and studies like the present are dependent on the particular list of control variables. For instance, it is always possible to find some measure of institutions that does not “kill” the corruption effect, if this is what the purpose is. To limit the degrees of freedom, I have taken all measures of institutions and culture from the large online dataset gathered by the Quality of

Governance Institute (QOG), included them one by one and chosen the measure yielding the highest explanatory power to the model.

Unfortunately, the QOG dataset does not include geographical measures. Most of the geographical variation is already accounted for by the regional dummies, but nevertheless I have included nine additional geographical covariates one by one in the column (2) regression.

Another choice increasing the degrees of freedom of the empirical researcher is the choice between taking logs or not. To reduce my degrees of freedom, I choose between taking logs or not by using the solution with the highest explanatory power.[5]

Geography. The top two geographic measures yielding highest explanatory power when included one by one in column (2) are also the only two significant measures; log of distance to coast or river (logdistcr) and log of elevation above sea level (logelev). These are included in column (3). The estimate of corruption remains significant at the 1% level, and the size of the estimate is statistically unchanged compared to column (2) (same result prevails when including them one by one).

Institutions. Similarly for the measures of institutions; I included the 25 institutional measures one by one in regression (2), taking logs whenever the explanatory power was higher.[6,7] Within each of the three categories of institutions (property rights institutions, rule of law, and contracting institutions), I pick the measure leaving the model with the highest explanatory power. Column (4) includes these three dimensions of institutions: The measure of property rights institutions providing highest R^2 was Henisz' (2000) political constraints measure, logh_polcon5, higher scores indicating more political constraint. The rule of law measure that gives highest explanatory power to the model was a measure by Freedom House, fh_rol, higher scores indicating more rule of law.[8] None of the six measures of contracting institutions turned out significant. Nonetheless, to complete the model in line with Acemoglu and Johnson (2005), I have included in column (4) the measure of contracting institutions yielding the highest explanatory power, procedural complexity, logproc_compl, higher scores meaning more procedural complexity.[9] Corruption is still significant at the 1% level, and the estimate increases above the column (2) estimate. The column (4) results might suffer from problems of multicollinearity; the rule of law measure, fh_rol, is highly correlated with both corruption (corr 0.77) and the property rights measure, logh_polcon5 (corr 0.73). Column (5) excludes the rule of law measure, including only property rights institutions and contracting institutions in line with the analysis by Acemoglu and Johnson (2005). The explanatory power of the model does not fall much. Excluding instead property rights institutions and keeping rule of law reduces the explanatory power even less (not shown), which indicates that maybe the rule of law dimension is a better fit than property rights institutions. In column (9), I deal with the multicollinearity problem using principal components analysis (described below).

Culture. The QOG dataset includes numerous measures of culture, mostly from the World Values Survey. I have not included all of them, but found 31 measures of various religious affiliations, ethnic fractionalization, and also some measures from World Values Surveys that measure similar culture dimensions as the Hofstede measures. Only one of all these measures had a significant impact on GDP per capita at the 10% level when included in the column (2) regression; fraction of Protestants in 1980, lp_protmg80.[10] Column (6) includes fraction of Protestants, leaving the estimate of corruption and the level of significance unchanged.

Columns (7) and (8) include all deep determinants, excluding rule of law in column (8) to reduce the potential multicollinearity issue. Corruption remains negative and significant at the 1% level.

The main purpose of this part of the analysis is to remove as much as possible of the variation in GDP per capita caused by the deep determinants. This will both reduce bias caused by omitted variables, but will also increase the validity of the instruments in the next part of the analysis. Therefore, I would like to get rid of the multicollinearity problem in another way than excluding variables. In Web Appendix A2 (on my webpage) I perform principal component analysis (pca) to replace the control variables by their principal components.[11] In short, I identify combinations of the control variables (so-called principal components) that explain as large a share of the variation in GDP per capita as possible and include them into the analysis in column (9) (corresponds to column (11) of Table A3 in Web Appendix A2). This method is useful, as I am not interested in the particular estimate of the control variables, but only in the estimate on corruption. Note that the estimate of corruption becomes more precise than in any other specification and is now higher than the column (2) estimate at the 10% level. The increased estimate is not due to the smaller sample; when reducing the sample to the 119 observations, the column (2) estimate *drops* to 8.703. Hence, it seems that the omitted factors in the column (2) regression are driving down the corruption estimate, making corruption seem less severe.

Note that the R-squared of column (7), (8), and (9) reaches as high as 81%, which is high for cross-country regressions.

5. Identification

Obviously, the estimates in Panel A of Table 1 may be biased by endogeneity. The level of GDP per capita might influence corruption; richer countries have more resources to combat corruption (e.g. Gundlach and Paldam, 2009). Or GDP could have a similar effect on any of the other dimensions of institutions, which will also bias the estimate of corruption.

I attempt to solve the endogeneity problem exploiting the idea that corruption is affected by deeply rooted cultural values (e.g. Rose-Ackerman, 1999, Treisman, 2000, and Licht et al., 2007). Two dimensions of culture in particular seem to be impacting the tendency to be corrupt; dimensions concerning individualism and hierarchy.

5.1. IV Strategy

Individualism distinguishes between collectivist societies on the one hand, where people attach great importance to their social networks, as opposed to individualist societies, where the *individuals'* rights and responsibilities are in focus (Hofstede, 2001, Schwartz, 1999). A collectivist orientated culture values tightly knit relations in which people expect their social group to look after them in exchange for unquestioning loyalty. When allocating resources, a public official favors his own social group in return for a share of the benefits. In relation to this, Rose-Ackerman (1999) observes that (p. 98) "*He [the public official] may do this not only because he cares for them [the social group], but also because they care about him and will be less likely than strangers to reveal* *the corrupt deal or renege on the agreement. The interdependency of utilities reduces the risks to both participants.*" Decisions made in more collectivist societies are, ceteris paribus, subject to less questioning from people within the same social network as the particular decision maker and the risk of getting caught is lower (Licht et al., 2007).

Several other authors have noted a relationship between Individualism and corruption (Husted, 1999 for an overview). Hooper (1995) links the tendency to favor one's ingroup to corruption in Spain. Banfield (1958) saw a connection between the "amoral familiarism" (favoritism for family members) of a small village in Italy and the tendency for its public office holders to accept bribes.

The other culture dimension, Hofstede's Power Distance measure, focuses on the degree to which inequality in the distribution of power is accepted and expected in a society. Hofstede notes that larger power distances in a society mean fewer checks and balances on the use of power, and thereby more corruption. He underlines this idea by a credited statement by Lord Acton, a 19th century British historian: "*Power tends to corrupt, and absolute power corrupts absolutely.*"[12] In societies, where power is expected and accepted to be distributed unequally, people tend to be less critical towards decisions made by authorities. This reduces transparency and means that authorities in high Power Distance societies are faced with lower costs of corruption - in terms of the risk of getting caught - compared to societies where equal distribution of power is expected and where people in power are kept accountable by a critical populace. Takyi-Asiedu (1993) has linked power distances to corruption in sub-Saharan Africa. He found that in high power distance countries, "*scandals involving people in authority are, almost always, covered up as long as they remain in power.*" Treisman (2000) argues that in places where more hierarchical religions prevail, critical positioning towards people in power is rarer.

Based on this reasoning, I use Hofstede's two culture measures Individualism (IDV) and Power Distances (PDI) as instruments for corruption in equation (1). Figure 2 provides the simple scatterplots of CORRUPT versus IDV and PDI, showing the expected relationships: Countries with more individualism or smaller power distances tend to be less corrupt.

Figure 2 about here

5.2. IV Estimation

Panel B of Table 1 shows the TSLS estimates of the same regressions as in panel A. The difference between the two panels is that corruption in Panel B is instrumented with the Hofstede culture dimensions Power Distance and Individualism.

As a test of whether the instruments are weak, I use the Kleibergen-Paap F statistic, which is the heteroskedasticity robust generalization of the Cragg-Donald F statistic (which is again equivalent to the first stage F statistic in the case of one endogenous regressor). There has still not been generated critical values for the Kleibergen-Paap F statistic, so it is custom to use the critical values for the Cragg-Donald F statistic, available from Stock and Yogo (2005). For comparison, I also provide the Cragg-Donald F statistic in the tables. I have indicated whether the weak instrument test statistic is above or below the critical values provided by Stock and Yogo (2005). Specifically, I use the TSLS maximal size distortion, which provides the most binding critical values, compared to TSLS *bias*. Stock and Yogo (2005) provide maximum size distortions of 10, 15, 20, and 25%.

In column (1), where corruption is the only explanatory variable, the Kleibergen-Paap F statistic equals 68, which is well above the highest critical value of 19.93. Hence, IDV and PDI are strong instruments for corruption.

Panel B of Table 1 also provides the p-value of the OID test, which tests whether corruption is the only channel through which the instruments influence GDP per capita. This is a test of low power and it should not be relied upon independently. Interestingly, in column (1) we can only accept at the 12% level that the two culture dimensions only influence GDP per capita through corruption. This is a very low level of acceptance and it supports the suspicion that there might exist other channels through which these dimensions of culture influence GDP per capita or that there might be some excluded factors influencing both culture and GDP. This points to the importance of including other deep determinants of economic productivity. Indeed, the OID p-value increases to 61.3% in column (2), where the regional dummies are included, indicating that the problem of exclusion restrictions might be solved by just including these dummies. But again, the test has low power, and I continue by operating within the framework of deep determinants.

Adding other geography variables in column (3) does not alter the size or level of significance of the corruption estimate compared to the column (2) estimate. Adding the remaining dimensions of institutions in column (4) makes the instruments weak, inflating the standard errors and potentially biasing the estimates. LIML estimation is more robust to weak instruments. When using LIML estimation instead (not shown), the instruments become fairly strong (critical value of 15% LIML size distortion is 5.33) and the corruption estimate is 10.45 (standard error 3.35), significant and unchanged compared to column (2).

But again, the column (4) estimate might be biased due to multicollinearity between rule of law and both corruption and property rights institutions. When excluding the rule of law measure in column (5), the culture instruments are again strong and the estimate of corruption is unchanged compared to column (2).

The instruments remain strong in column (6), where the culture measure, fraction of Protestants, is included and the estimate of corruption is unchanged compared to the column (2) regression.

Columns (7) and (8) include all deep determinants simultaneously. The instruments remain fairly strong in column (8), where rule of law is excluded. The corruption estimate remains significant at the 1% level and statistically not different from the estimate in column (2), albeit it is numerically somewhat higher.

Again, I use principal component analysis to produce principal components that span the entire set of control variables, solving the multicollinearity problem and increasing the precision of the corruption estimate. The analysis is performed in Web Appendix A2 (on my webpage) and the significant principal components are included in column (9) of Table 1. The instruments are fairly strong and the corruption estimate increases above the column (2) estimate. When running the column (2) regression on the column (9) sample, the corruption estimate is unchanged at 8.729. Hence the increase is not due to sample differences.

The heteroskedasticity robust Hausman test cannot reject that the estimate of corruption is the same as that produced by OLS. That is, there does not seem to be a problem of endogeneity.

To sum up, corruption exerts a statistically significant and negative impact on productivity levels across countries in the samples studied. The effect is economically significant. Taking the lowest estimate (column (3), Panel A, Table 1), a one unit reduction in the level of corruption increases the level of GDP per capita by 8,547 US$. Within this sample of 156 countries, the control of corruption index runs from -1.83 in Somalia to 2.56 in Finland. Hence, a one unit reduction of corruption amounts to 22.8% of the entire interval. On the same sample, GDP per capita runs from a low of 369 US$ in Liberia to a peak of 77,242 US$ in Qatar. Hence, an increase of 8,547 US$ amounts to 11.1% of the interval. This means that reducing corruption in a country with a fifth of what is possible can increase this country's GDP per capita by more than 10%. Taking instead the highest corruption estimate from column (9) of Table 1, panel B, a one unit reduction in the level of corruption increases the level of GDP per capita by 14,606 US$. Within this sample of 69 countries, the control of corruption index runs from -1.32 in Bangladesh to 2.56 in Finland. Hence, a one unit reduction of corruption amounts to 25.8% of the interval. On the same sample, GDP per capita runs from a low of 887 US$ in Tanzania to a peak of 49,391 US$ in United Arab Emirates. Hence, an increase of 14,606 US$ amounts to 30.1% of the interval. While being large, the effect is not implausibly large.

5.3. Robustness: Other endogenous regressors

The remaining dimensions of institutions could impose other problems of endogeneity, biasing the estimate of interest. In Table 2 therefore, I include instruments for the remaining dimensions of institutions; property rights institutions, rule of law, and contracting institutions. The instruments meant for corruption remain PDI and IDV throughout. Since the instruments for the remaining institutions are somewhat weaker, I use LIML estimation instead of TSLS.[13]

Table 2 about here

Property rights institutions. Column (1) of Table 2 includes property rights institutions and uses Albouy's (2008) corrected version of Acemoglu et al.'s (2001) instrument for property rights institutions, settler mortality, together with urbanization rates from 1500, also suggested by Acemoglu et al. (2001). The number of observations drops to 27 and the instruments become extremely weak, producing huge standard errors and making all the estimates unreliable. In column (2) I keep the same instruments, but use another measure of property rights institutions, namely the preferred measure used by Acemoglu and Johnson (2005); Constraint on Executive (p_xconst). This measure produced the second highest explanatory power of the model when I included the measures of property rights institutions one by one in the column (2) regressions of Table 1. The instruments become strong: The critical value for the weak instrument test provided by Stock and Yogo (2005) for 2 endogenous variables and 4 instruments based on LIML size distortions of a maximum of 10% is 4.7 < 9.6 (Kleibergen-Paap F statistic). The estimate of corruption remains highly significant and not numerically different from the OLS estimate in column (2) of Table 1. The estimate of property rights institutions is insignificant.

Rule of law. Column (3) includes the rule of law measure, fh_rol, and introduces as instrument another culture dimension identified by Hofstede; Uncertainty Avoidance (UAI). The Uncertainty Avoidance Index focuses on the level of tolerance for uncertainty and ambiguity within a society. It indicates to what extent people feel comfortable in unstructured situations. Unstructured situations are novel, unknown, surprising, different from usual. Uncertainty avoiding cultures try to minimize the possibility of such situations by strict laws and rules, safety, and security measures. Therefore, I expect countries with high UAI to have better rule of law institutions, which proves to be the case. The instruments are fairly strong (above the 15% size distortion critical value) and the estimate of corruption remains unchanged. The estimate of rule of law is insignificant.

Contracting institutions. Column (4) includes the last dimension of institutions, contracting institutions, instrumented with a dummy capturing French legal origin, suggested by Acemoglu and Johnson (2005). The instruments become weak and the estimates unreliable (the same is the case using English legal origin etc.). Therefore I instead include the instrument for contracting institutions, French legal origin, directly into the regressions in column (5). The idea is that French legal origin captures at least some of the variation in contracting institutions.[14] This makes the instruments for corruption strong and the estimate of corruption is unchanged compared to column (2) in panel B of Table 1.

Columns (6) and (7) include the property rights institutions and rule of law measures, respectively together with the French legal origin dummy, producing strong instruments (not so strong for rule of law) and unchanged estimate of corruption. Adding geography to the pool of controls produces weak instruments in the regression with property rights institutions (not shown), but strong instruments in the regression with rule of law (column 8) and statistically unchanged estimate of corruption (albeit it falls somewhat in absolute value). Column (9) adds the last deep determinant; culture, which produces fairly strong instruments. Column (10) includes the instruments for property rights institutions and rule of law directly into the regression, producing strong instruments and unchanged estimate of corruption.[15] Column (11) includes the three significant principal components produced from a principal components analysis of the control variables in column (10). The estimate of corruption becomes again more precise and is not different from the Table 1, column (2) estimate.

5.4. Robustness: Instrument Exogeneity

The Hofstede dimensions of culture that I use as instruments for corruption, IDV and PDI, are measured 30 years before corruption and GDP, and it is therefore not possible for either corruption or GDP to have an effect on the instruments. Further, since Hofstede constructed the measures from a survey of a very similar group of people, IBM workers, the measures should be independent of demographic differences etc. across countries. Nevertheless, one may still worry that some omitted factors influence GDP and culture. There are many factors influencing individuals' inherited cultural beliefs, but omitting these will only bias the results if these factors also influence GDP per capita. Using the deep determinants framework, I hoped to be able to include all factors with a bearing on GDP. If I succeeded, there is no reason to worry about instrument endogeneity. But if the deep determinants framework is wrong or if the measures are systematically biased, I could have omitted factors that influence both culture and GDP, invalidating the instruments.

I can test empirically whether IDV is exogenous in relation to GDP per capita, since I have an instrument for IDV. Licht et al. (2007) argue that less individualist cultures have a tendency to drop the pronoun (I, you, he, she etc.), since the individual is less important in these cultures and therefore it does not matter whether it was he, she or I who did this and that. They document that a dummy equal to one if the country has dropped the pronoun is a strong instrument for IDV. Table 3 shows the results from the TSLS estimation of IDV on GDP per capita, using pronoun drop as an instrument for IDV. No control variables are included in column (1), and pronoun drop is a strong instrument for IDV (F=62>16.38, the 10% critical value with one instrument). The Robust Hausman test cannot reject that IDV is exogenous in relation to GDP per capita.

Table 3 about here

It seems unlikely that pronoun drop influences GDP per capita through other channels than culture, but just to make sure, I include the exogenous deep determinants of GDP per capita in columns (3) and (4). IDV remains exogenous in relation to GDP per capita.

Unfortunately, I have no instrument for PDI. Instead, as a last check of instrument exogeneity, Table 4 shows that IDV and PDI explain a significant share of the variation in GDP per capita (col 1-3), but that the impact of IDV and PDI goes away once corruption is controlled for (col 4-6). I take this as indicating that the impact of culture works only through corruption. The same pattern emerges when including IDV and PDI separately (not shown).

Table 4 about here

5. Conclusion

This paper provides an estimate of the impact of corruption on economic productivity levels. The motivation for doing so is that the estimate from Mauro's (1995) study is still the estimate referred to in the literature, despite the wellknown econometric shortcomings of the analysis. My attempt at overcoming these shortcomings starts with providing new strong instruments for corruption. The hypothesis is that more hierarchical cultures that focus more on loyalty towards one's social group than the individual's own responsibility, are more prone to become corrupt compared to more individualist and egalitarian cultures. This is supported by the empirics. Further, these dimensions prove to be strong instruments for corruption in the regressions on GDP per capita.

However, one could imagine that these dimensions are correlated with other deep factors that influence productivity levels. Therefore, the reliability of the estimates hinges on inclusion of additional deep determinants of GDP per capita; geography and the remaining dimensions of both institutions and culture. To reduce the degrees of freedom arising when choosing between the numerous measures of these deep determinants, I have restricted myself to include all measures in an online dataset by Teorell et al. (2009) that attempt to gather the variables from the central research in the field. After including these factors, the instruments remain strong and the impact of corruption on GDP per capita remains statistically and economically significant and negative.

References

Acemoglu, D. and S. Johnson (2005) Unbundling Institutions, *Journal of Political Economy* 113(5), 949-995.

Acemoglu, D., S. Johnson, and J. Robinson (2001) The Colonial Origins of Comparative Development: An Empirical Investigation, *American Economic Review* 91(5), 1369-1401.

Albouy, D. Y. (2008) The Colonial Origins of Comparative Development: An Investigation of the Settler Mortality Data, *NBER Working Paper* 14130.

Banfield, E. C. (1958) The Moral Basis of a Backward Society, New York: *The Free Press*.

de Vaal, A. and W. Ebben (2011) Institutions and the Relation between Corruption and Economic Growth, *Review of Development Economics* 15(1), 108-123.

Easterly, W. and R. Levine (1997) Africa's Growth Tragedy: Policies and Ethnic Divisions, *Quarterly Journal of Economics*, 112(4), 1203-50.

Flores-Lagunes, A. (2007) Finite Sample Evidence of IV Estimators Under Weak Instruments, *Journal of Applied Econometrics* 22, 677-694.

Gallup, J. L., J. D. Sachs, and A. Mellinger (1999) Geography and Economic Development, *International Regional Science Review* 22, 179-232.

Guiso, L., P. Sapienza, and L. Zingales (2006) Does Culture Affect Economic Outcomes? *Journal of Economic Perspective* 20(2), 23-48.

Gundlach, E. and M. Paldam (2009) The Transition of Corruption: From Poverty to Honesty, *Economic Letters* 103(3), 146-148.

Hall, R. E. and C. I. Jones (1999) Why Do Some Countries Produce So Much More Output Per Worker Than Others? *Quarterly Journal of Economics* 114(1), 83-116.

Henisz, W. J (2000) The Institutional Environment for Economic Growth, *Economics and Politics* 12(1), 1-31.

Heston, A., R. Summers, and B. Aten (2009) Penn World Table Version 6.3, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, August 2009.

Hofstede, G. (2001) Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations, 2nd Ed. Beverly Hills CA: *Sage Publications*.

Hooper, J. (1995) The new Spaniards. London: *Penguin Books*.

Husted, B. W. (1999) Wealth, Culture and Corruption, *Journal of International Business Studies* 30(2), 339-359.

Jolliffe, I. T. (2002) Principal Component Analysis, Second Edition, *Springer Series in Statistics*.

Kaufmann, D., A. Kraay, and M. Mastruzzi (2009) Governance Matters VIII: Governance Indicators for 1996-2008, *World Bank Policy Research* June 2009.

La Porta, R., F. Lopez-De-Silanes, A. Shleifer, and R. W. Vishny (1999) The Quality of Government, *Journal of Law, Economics, and Organization* 15(1), 222-279.

Leff, N. (1964) Economic development through bureaucratic corruption, *American Behavioral Scientist* 8, 8-14.

Licht, A. N., C. Goldschmidt, and S. Schwartz (2007) Culture Rules: The Foundations of the Rule of Law and Other Norms of Governance, *Journal of Comparative Economics* 35, 659-688.

Mauro, P. (1995) Corruption and Growth, *Quarterly Journal of Economics* 110(3), 681-712.

Rodrik, D., A. Subramanian, and F. Trebbi (2004) Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development, *Journal of Economic Growth* 9, 131-165.

Rose-Ackerman, S. (1999) Corruption and Government: Causes, Consequences, and Reform, Cambridge, UK; New York: *Cambridge University Press*.

Sachs, J. (2003) Institutions Don't Rule: Direct Effects of Geography on Per Capita Income, *NBER Working Paper* 9490.

Schwartz, S. H. (1999) Cultural Value Differences: Some Implications for Work, *Applied Psychology: An International Review* 48, 23-47.

Schwartz, S. H. (2004) Mapping and interpreting cultural differences around the world. In: Vinken, H., J. Soeters, P. Ester (Eds.), Comparing Cultures. *Brill Academic Publishers*, Leiden, 43–73.

Staiger, D. and J. H. Stock (1997) Instrumental variables with Weak Instruments, *Econometrica* 65, 557-586.

Stock, J. H. and M. Yogo (2005) Testing for Weak Instruments in Linear IV Regression, In D. W. K. Andrews and J. H. Stock, eds. Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg. Cambridge: Cambridge U y Press, 2005. Working paper version: *NBER Technical Working Paper* 284.

Tabellini, G. (2008) Institutions and Culture, *Journal of the European Economic Association Papers and Proceedings* 6(2-3).

Takyi-Asiedu, S. (1993) Some socio-cultural factors retarding entrepreneurial activity in sub-Saharan Africa, *Journal of Business Venturing* 8, 91-98.

Teorell, J., N. Charron, M. Samanni, S. Holmberg, and B. Rothstein (2009) The Quality of Government Dataset, version 17June09. University of Gothenburg: The Quality of Government Institute, http://www.qog.pol.gu.se.

Treisman, D. (2000) The causes of corruption: A cross-national study, *Journal of Public Economics* 76, 399-457.

Wooldridge, J. M. (2002) Econometric Analysis of Cross Section and Panel Data, *The MIT Press*, Cambridge.

Notes

1 Some authors argue that trade should also belong with these deep determinants. However, trade can be viewed as belonging to the more proximate factors, being itself affected by both institutions, geography, and culture.
2 de Vaal and Ebben (2011) show in a theoretical model that the impact of corruption on growth depends on the institutional framework. This speaks for including an interaction term between corruption and institutions, which could form basis for future research.
3 See also his webpage: www.geert-hofstede.org.
4 Removing the two countries Qatar and Luxembourg reduces the corruption point estimate from 9.2 to 8.4, which is not a statistically significant difference.
5 Of course, I did not take logs where it does not make sense; e.g. indexes and dummies.
6 Seven of the 25 measures turned out to exert a significant impact on GDP per capita: va2006, logbti_prp, logbti_rol, fh_rol, logh_polcon3, logh_polcon5, and p_xconst (see Web Appendix A1 for description, available on my webpage). Including all measures gives an estimate of corruption of 9.7 with a t-statistic of 5. 6. That is, no change of the results.
7 One could use principal components analysis and insert the principal components into the regression instead of the variables one by one. However, the problem is that the variables do not have enough observations in common.
8 The estimate of rule of law is negative, but the raw correlation between GDP and rule of law is positive as one would expect. The negative estimate might reflect that countries that already have good institutions (in terms of low corruption, low political constraints, and low procedural complexities) might not benefit from more rule of law.
9 If I instead exclude the measure of contracting institutions from column (4), the estimate of corruption increases from 11.2 to 11.6.
10 Additional measures of culture turned out significant when excluding the regional dummies, though. This finding underlines the importance of regional dummies and indicates that these culture measures capture nothing more than regional differences.
11 Thank you to an anonymous referee for suggesting this.
12 Lord Acton made his statement in a letter to Bishop Mandell Creighton dated April 5, 1887.
13 A Monte Carlo study done by Flores-Lagunes (2007) suggests that LIML does at least as well as alternatives. LIML is less precise than TSLS, but also less biased.
14 When only regressing contracting institutions on GDP and instrumenting with French legal origin, the instrument is strong, suggesting that countries with French legal origin indeed have more procedural complexities.
15 Note that settler mortality is not included. Including this variable in column (10) reduces the number of observations to 26 and produces weak instruments.

Figures and Tables

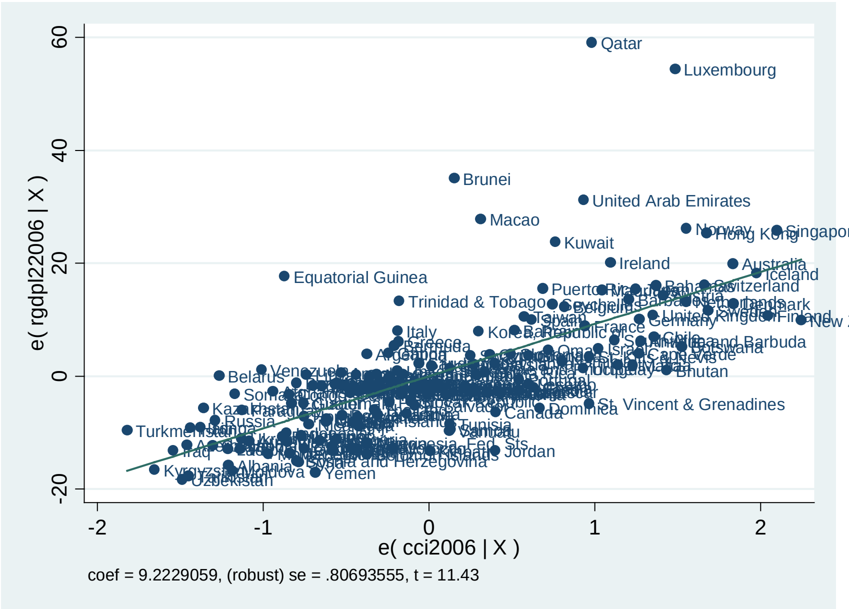


Figure 1. OLS estimation of column (2), panel A, Table 1.

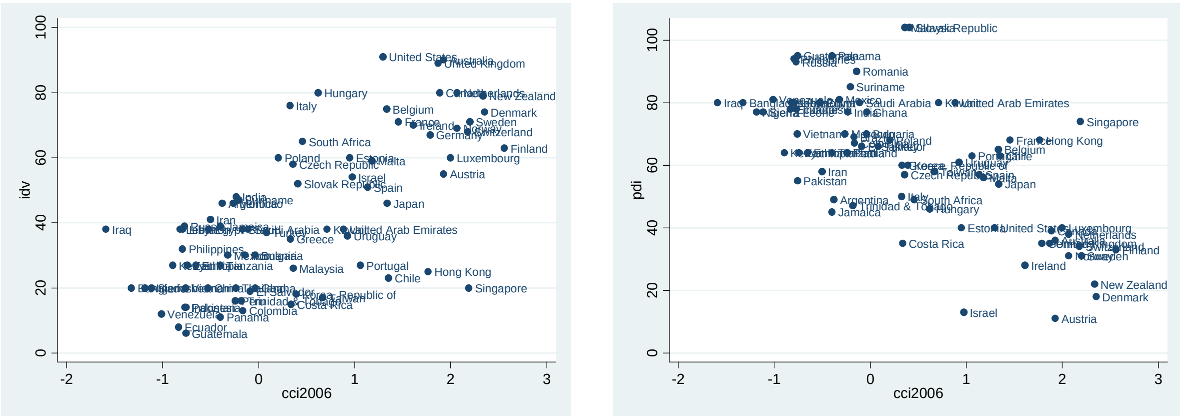


Figure 2. Simple plot of cci2006 vs IDV and cci2006 vs PDI (79 countries).

Table 1. OLS and IV (TSLS) estimates of the impact of corruption on GDP per capita

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)[a]
Dependent variable: rgdpl22006									
Regional dummies	no	yes	yes	yes	yes	yes	yes	yes	no
Panel A: OLS estimates									
cci2006	10.582***	9.223***	8.547***	11.182***	10.119***	9.637***	11.062***	9.644***	10.238***
	(0.766)	(0.807)	(0.901)	(0.982)	(0.857)	(0.993)	(1.147)	(0.950)	(0.567)
fh_rol				-0.435*			-0.523**		
				(0.233)			(0.253)		
logh_polcon5				-5.265	-8.811**		-4.856	-8.671**	
				(4.540)	(3.933)		(4.477)	(3.946)	
logproc_compl				2.295	2.092		2.741	2.493	
				(2.169)	(2.164)		(2.397)	(2.386)	
logdistcr			-1.148				-0.505	-0.309	
			(0.794)				(0.555)	(0.545)	
logelev			-1.288				-0.859	-0.901	
			(1.022)				(0.717)	(0.745)	
lp_protmg80						-0.070*	-0.009	-0.003	
						(0.036)	(0.048)	(0.046)	
Observations	186	186	156	122	122	178	119	119	119
R-squared	0.569	0.629	0.663	0.811	0.806	0.633	0.814	0.808	0.793
F-test cci2006 = 9.223, p	0.078	1.000	0.454	0.049	0.298	0.677	0.112	0.659	0.076
Panel B: TSLS estimates (corruption endogenous), second stage									
cci2006	10.760***	8.706***	8.613***	10.480***	10.059***	9.969***	15.062***	11.436***	14.606***
	(1.124)	(1.269)	(1.363)	(3.216)	(1.563)	(1.888)	(5.239)	(1.972)	(1.968)
fh_rol				-0.115			-1.061		
				(0.665)			(1.080)		
logh_polcon5				-9.545	-9.901		-8.617	-12.346**	
				(6.752)	(6.884)		(5.666)	(6.056)	
logproc_compl				4.072	3.835		8.627**	7.115**	
				(3.273)	(2.763)		(4.056)	(2.952)	
logdistcr			-1.476				-0.172	0.511	
			(1.267)				(1.098)	(0.757)	
logelev			-0.304				-1.980*	-1.902	
			(1.127)				(1.112)	(1.167)	
lp_protmg80						-0.067	-0.075	-0.050	
						(0.067)	(0.092)	(0.070)	
Observations	79	79	75	71	71	78	69	69	69
R-squared	0.669	0.714	0.726	0.811	0.810	0.724	0.816	0.819	0.800
Kleibergen-Paap F	68.11$^{\circ\circ\circ}$	48.12$^{\circ\circ\circ}$	48.37$^{\circ\circ\circ}$	7.595	23.63$^{\circ\circ\circ}$	30.99$^{\circ\circ\circ}$	3.314	16.50$^{\circ\circ}$	17.02$^{\circ\circ}$
Cragg-Donald F	51.16$^{\circ\circ\circ}$	30.77$^{\circ\circ\circ}$	36.80$^{\circ\circ\circ}$	5.330	15.69$^{\circ\circ}$	20.65$^{\circ\circ\circ}$	3.511	11.90$^{\circ\circ}$	13.82$^{\circ\circ}$
OID p-value	0.117	0.613	0.899	0.367	0.367	0.701	0.350	0.270	0.248
F-test cci2006 = 8.706, p	0.068	1.000	0.946	0.581	0.387	0.504	0.225	0.166	0.003
Robust Hausman test, p	0.725	0.348	0.561	0.710	0.584	0.723	0.786	0.938	0.551

Notes. OLS and second stage TSLS estimates. Dependent variable is real GDP per capita in 2006 from PWT. Robust standard errors in paranthesis. In all columns of Pabel B, corruption is instrumented with Hofstede s Power Distances (PDI) and Individualism (IDV). Constant included in all regressions. Asterisks *, **, and *** indicate significance at the 10, 5, and 1% level, respectively. Dots $^{\circ}$, $^{\circ\circ}$, and $^{\circ\circ\circ}$ indicate TSLS size distortions of a maximum of 20, 15, and 10%, respectively.

[a] Instead of the control variables included in column (7), the corresponding principal components are included. The principal components are based on the principal components analysis performed in Web Appendix A2, available on my website.

Table 2. Robustness: Second stage IV (LIML) estimates of the impact of corruption on GDP per capita

Dependent variable: rgdpl22006	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)[a]
cci2006	1.759	10.975***	10.286***	6.251**	8.409***	11.079***	11.237***	8.444**	12.805**	9.547***	9.496***
	(74.808)	(1.401)	(2.384)	(2.638)	(1.575)	(1.446)	(3.179)	(3.784)	(5.123)	(1.390)	(0.804)
logh_polcon5	93.574										
	(819.335)										
p_xconst		-2.049				-2.058					
		(1.360)				(1.268)					
fh_rol			-0.429				-0.702	0.011	-0.871		
			(0.584)				(0.667)	(0.899)	(1.107)		
logproc_compl				-12.229							
				(13.421)							
leg_french					-0.281	0.726	0.736	0.606	0.242	1.119	
					(3.418)	(3.095)	(3.499)	(3.248)	(2.792)	(1.809)	
logdistcr								-1.475	-1.558	0.574	
								(1.435)	(1.336)	(0.729)	
logelev								-0.358	-1.057	-0.030	
								(0.993)	(1.108)	(1.030)	
lp_protmg80									-0.096	-0.006	
									(0.085)	(0.072)	
uai										0.016	
										(0.044)	
urb1500										-0.011	
										(0.159)	
Regional dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	no
Instruments	pdi, idv, settler, urb1500	pdi, idv, settler, urb1500	pdi, idv, uai	pdi, idv, leg_french	pdi, idv	pdi, idv, settler, urb1500	pdi, idv, uai	pdi, idv, uai	pdi, idv, uai	pdi, idv	pdi, idv
Observations	27	27	78	69	76	27	75	73	73	58	58
R-squared	0.140	0.826	0.722	0.721	0.714	0.823	0.729	0.727	0.755	0.809	0.792
OID p-value	0.123	0.171	0.751	0.694	0.950	0.174	0.972	0.817	0.956	0.243	0.031
Kleibergen-Paap F	0.441	$9.569^{\circ\circ\circ}$	$4.403^{\circ\circ}$	2.513	$35.90^{\circ\circ\circ}$	$9.199^{\circ\circ\circ}$	3.509°	$5.576^{\circ\circ\circ}$	$4.414^{\circ\circ}$	$23.36^{\circ\circ\circ}$	$77.53^{\circ\circ\circ}$
Cragg-Donald F	0.832	$4.081^{\circ\circ}$	$6.400^{\circ\circ\circ}$	2.352	$26.03^{\circ\circ\circ}$	$4.715^{\circ\circ\circ}$	$4.955^{\circ\circ}$	$4.373^{\circ\circ}$	2.871	$25.20^{\circ\circ\circ}$	$54.98^{\circ\circ\circ}$
F-test cci2006 = 8.706, p	0.926	0.105	0.508	0.352	0.850	0.101	0.426	0.945	0.424	0.545	0.390

Notes. LIML estimates, second stage (endogenous variables: cci2006, logh_polcon5, p_xconst, and fh_rol). Dependent variable is real GDP per capita in 2006 from PWT. Robust standard errors in paranthesis. Constant included in all regressions. Asterisks *, **, and *** indicate significance at the 10, 5, and 1% level, respectively. Dots $^{\circ}$, $^{\circ\circ}$, and $^{\circ\circ\circ}$ indicate LIML size distortions of a maximum of 20, 15, and 10%, respectively.

[a] Instead of the control variables included in column (10), the corresponding principal components are included. The principal components analysis is similar to that performed in Appendix A2, but where the included variables are those of column (10) above.

Table 3. Robustness: Testing instrument exogeneity (IV)

	(1)	(2)	(3)	(4)
Second stage TSLS estimates, dependent variable: rgdpl22006				
idv	0.381***	0.371***	0.389***	0.330***
	(0.068)	(0.082)	(0.099)	(0.116)
logdistcr			-1.758*	-1.924**
			(0.980)	(0.952)
logelev			-1.216	-1.173
			(2.057)	(2.067)
lp_protmg80				0.059
				(0.067)
Regional dummies[a]	no	yes	yes	yes
First stage estimate of pronoun, dependent variable: idv				
pronoun	-35.913***	-21.666***	-18.942**	-20.502**
	(4.558)	(7.094)	(7.150)	(7.739)
OLS estimate of idv, corresponding to TSLS regressions				
idv	0.352***	0.333***	0.356***	0.332***
	(0.043)	(0.049)	(0.050)	(0.052)
Observations	69	69	66	66
R-squared	0.375	0.413	0.550	0.560
Kleibergen-Paap F	62.08$^{\circ\circ\circ}$	17.8$^{\circ\circ\circ}$	15.68$^{\circ\circ}$	9.858$^{\circ\circ}$
Cragg-Donald F	74.44$^{\circ\circ\circ}$	39.88$^{\circ\circ\circ}$	31.29$^{\circ\circ\circ}$	19.38$^{\circ\circ\circ}$
Robust Hausman test, p	0.537	0.562	0.667	0.981

Notes. TSLS estimates (endogenous variable: idv). Dependent variable is real GDP per capita in 2006 from PWT. Robust standard errors in paranthesis. Constant included in all regressions. Asterisks *, **, and *** indicate significance at the 10, 5, and 1% level, respectively. Dots $^{\circ}$, $^{\circ\circ}$, and $^{\circ\circ\circ}$ indicate TSLS size distortions of a maximum of 20, 15, and 10%, respectively.

[a] Due to singleton dummies, the regional dummies are gathered into larger groupings: SOA+EAP, SSA+MENA, NA+LAC, and ECA.

Table 4. Robustness: Testing instrument exogeneity (OLS)

Dependent variable: rgdpl22006	(1)	(2)	(3)	(4)	(5)	(6)
cci2006				11.327***	10.793***	10.693***
				(1.451)	(1.792)	(2.345)
idv	0.276***	0.112*	0.165***	0.046	-0.053	-0.017
	(0.053)	(0.056)	(0.059)	(0.051)	(0.061)	(0.070)
pdi	-0.169***	-0.197***	-0.165***	0.075	0.028	0.007
	(0.063)	(0.052)	(0.060)	(0.056)	(0.053)	(0.054)
logdistcr			-2.215			-0.796
			(1.341)			(1.096)
logelev			0.001			-1.236
			(1.632)			(1.375)
lp_protmg80			-0.010			-0.090
			(0.080)			(0.082)
Regional dummies	no	yes	yes	no	yes	yes
Observations	79	79	75	79	79	75
R-squared	0.374	0.518	0.591	0.676	0.723	0.739

Notes. OLS estimates. Dependent variable is real GDP per capita in 2006 from PWT. Robust standard errors in paranthesis. Constant included in all regressions. Asterisks *, **, and *** indicate significance at the 10, 5, and 1% level, respectively.

Chapter 4

Does the Internet Reduce Corruption? Evidence from U.S. States and across Countries

Thomas Barnebeck Andersen, Jeanet Sinding Bentzen, Carl-Johan Dalgaard, and Pablo Selaya

The World Bank Economic Review Advance Access published May 27, 2011

Does the Internet Reduce Corruption? Evidence from U.S. States and across Countries

*Thomas Barnebeck Andersen, Jeanet Bentzen, Carl-Johan Dalgaard, and Pablo Selaya**

We test the hypothesis that the Internet is a useful technology for controlling corruption. In order to do so, we develop a novel identification strategy for Internet diffusion. Power disruptions damage digital equipment, which increases the user cost of IT capital, and thus lowers the speed of Internet diffusion. A natural phenomenon causing power disruptions is lightning activity, which makes lightning a viable instrument for Internet diffusion. Using ground-based lightning detection censors as well as global satellite data, we construct lightning density data for the contiguous U.S. states and a large cross section of countries. Empirically, lightning density is a strong instrument for Internet diffusion and our IV estimates suggest that the emergence of the Internet has served to reduce the extent of corruption across U.S. states and across the world. JEL Classification codes: K4, O1, H0

Corruption is commonly perceived to be a major stumbling block on the road to prosperity. Aside from retarding growth (Mauro 1995), corruption entails fiscal leakage, which reduces the ability of poor countries to supply essential public services such as schooling and health care (Reinikka and Svensson 2004;

* Thomas Barnebeck Andersen (corresponding author) is professor at the Department of Business and Economics, University of Southern Denmark, Campusvej 55, DK-5230 Odense M. Jeanet Bentzen is Ph.D. student, Carl-Johan Dalgaard is professor, and Pablo Selaya is assistant research professor at the Department of Economics, University of Copenhagen, Studiestraede 6, DK-1455 Copenhagen K, Denmark. Contact: Thomas Barnebeck Andersen (barnebeck@sam.sdu.dk), Jeanet Bentzen (jeanet.bentzen@econ.ku.dk), Carl-Johan Dalgaard (carl.johan.dalgaard@econ.ku.dk), and Pablo Selaya (pablo.selaya@econ.ku.dk). We owe a special thanks to William A. Chisholm for generously sharing his expertise on the subject of lightning protection. We also thank three anonymous referees, Phil Abbott, Shekhar Aiyar, Mark Bills, Christian Bjørnskov, Matteo Cervellati, Areendam Chanda, Oded Galor, Chad Jones, Sam Jones, Phil Keefer, Pete Klenow, David Dreyer Lassen, Ross Levine, Norman Loayza, Ben Olken, Ola Olsson, Elena Paltseva, Nancy Qian, Martin Ravallion, Yona Rubinstein, Paolo Vanin, Dietz Vollrath, David Weil, Nadja Wirz, and seminar participants at the University of Bergen, Brown University, University of Copenhagen, the Jerusalem Summer School in Economic Growth at Hebrew University, ETH Zürich, the ZEW Conference on Economics of Information and Communication Technologies, the 2008 Nordic Conference in Development Economics, the Growth Workshop at the 2008 NBER Summer Institute, the Southern Economic Association Annual Meeting 2008, the Latin American Econometric Society Meeting 2009, and the World Bank. Errors are ours.

THE WORLD BANK ECONOMIC REVIEW, pp. 1–31 doi:10.1093/wber/lhr025

World Development Report 2004). Corruption is unquestionably a governance failure one would like to dispose of, yet combating it has not proven easy.

In the present paper we hypothesize that the Internet is a useful technology for combating corruption around the world. We test this hypothesis using data for the 48 contiguous U.S. states as well as for a cross section of countries. Our estimates support the proposition that the Internet has worked to reduce corruption since its inception.

There are several reasons why the Internet could serve as a corruption-reducing technology. First, the Internet is among a select group of innovations that are considered General Purpose Technologies (e.g., Jovanovic and Rousseau 2005). General Purpose Technologies are fundamental and pervasive innovations that (with time) hold a first order impact on economic growth. Insofar as economic growth works to lower corruption, the Internet is likely to act as a corruption suppressor due to its positive impact on growth.[1] Second, rapid technological change usually encourages investment in human capital, which may instigate lower levels of corruption.[2] Accordingly, human capital accumulation is another potential transmission channel linking Internet diffusion and reductions in corruption levels. Third, the World Wide Web is a major source of information. Spreading information about official wrong-doing inevitably increases the risk of detection for politicians and public servants, thus making corrupt behavior less attractive.[3] Fourth, the Internet is the chief vehicle for the provision of E-government worldwide (West 2005). By allowing citizens access to government services online, E-government obviates bureaucrats' role as intermediaries between the government and the public, thus limiting the interaction between potentially corrupt officials and the public. Moreover, online systems require standardized rules and procedures. This reduces bureaucratic discretion and increases transparency as compared to the

1. See Andersen *et al.* (2010) for evidence that the Internet has stimulated growth across U.S. states, and e.g. Gundlach and Paldam (2009) for evidence of a causal impact of income on corruption.

2. See Schultz (1975) and Foster and Rosensweig (1996) for evidence of a positive impact of technological change on the return to human capital investments, and Glaeser and Saks (2006) for evidence on the impact of human capital on corruption.

3. A nice illustration of this mechanism at work is found in a 2001 scandal from India, which nearly toppled the government. Reporters from the online news site <www.Tehelka.com> posed as arms dealers and documented negotiations with top politicians and bureaucrats over the size of required side payments to get a contract; in some instances the reporters even got the delivery of the bribe on camera. Consequently, numerous politicians and top officials had to resign, chief among them the defence minister George Fernandes. See The Sting That Has India Writhing by Celia W. Dugger, *The New York Times* (March 16, 2001).

arbitrariness available to civil servants when dealing with the public on a case-by-case basis.[4]

The present paper examines the *reduced form* effect of the Internet on corruption levels across the U.S. and across the World. Thus, our analysis addresses the key issue of whether the Internet has had a causal impact on corruption. At the same time, the analysis admittedly does not clarify the exact mechanism(s) through which the Internet affects corruption; it may be either one of the mechanisms mentioned above, or some combination of them.

The Internet is a new technology. Indeed, if we identify the Internet with the World Wide Web, it only emerged in 1991.[5] To examine the impact of the Internet we therefore estimate the impact of *changes* in Internet users on *changes* in corruption levels from the early 1990s to 2006. Using data for the 48 contiguous U.S. states as well as cross-country data we establish a strong partial correlation between the rate of changes in Internet users and the evolution of corruption, consistent with the proposed hypothesis. However, since the speed of Internet diffusion is likely endogenous, OLS estimates might be misleading.

In an effort to establish causality we develop an identification strategy designed to isolate exogenous variation in the speed of Internet diffusion. The theory underlying our instrument of choice is the following. Computer equipment is highly sensitive to power disruptions: power surges and sags lead to equipment failure and damage. Consequently, a higher frequency of power disruptions implies higher costs of IT equipment, either through elevated IT capital depreciation or due to incurrence of additional costs in order to protect equipment from power disruptions. Frequent power disturbances are also likely to reduce the productivity of IT capital (or its marginal benefit), as power disturbances produce downtime, generate data glitches, etc. A natural phenomenon which produces power disturbances is lightning activity. In fact, one third of all power disruptions in the U.S. are related to lightning activity. We therefore hypothesize that higher lightning intensity is a viable candidate instrument for Internet diffusion. Using state level measures of lightning density (ground strikes per square km per year) and global satellite data on lightning activity,

4. The celebrated Bhoomi program (located in the state of Karnataka in India) constitutes a good example of the effectiveness of E-government in limiting the interface between civil servants and the public. Starting in 1998 the program aimed to computerize land records, and at the time of writing more than 20 million landholdings belonging to the state's 6.7 million landowners have been registered. Before online registration was available citizens had to seek out village accountants to register, a process which involved considerable delays and the need for bribes to be paid; a typical bribe could range from Rs.100 to Rs. 2,000 (US $2 to $40) (Bhatnager 2003). With the online system there is no longer a need for the official middlemen, implying that the online system probably has saved locals hundreds of millions of Rs. in bribes. See also Andersen (2009).

5. TheWorld Wide Web was launched in 1991 by CERN (the European Organisation for Nuclear Research). See Hobbes' Internet Timeline v8.2 <http://www.zakon.org/robert/internet/timeline/>. In this paper, we define the Internet/WWW as the network of networks using the TCP/IP/HTTP protocols, which was spawned by the launch of WWW.

we establish its influence on Internet diffusion: areas (states and countries) with a higher flash density have experienced a slower speed of Internet adoption.

A concern of first-order importance is whether, conditional on Internet diffusion, lightning density acts as a stand-in for factors that are correlated with corruption. To fix ideas, suppose areas with high lightning density just happen to be characterized by an abundance of natural resources as well, which influences corruption in its own right (i.e., conditional on Internet penetration). If so, then lightning is not a valid instrument for the Internet.

Notice, however, that the proposed instrument is most likely to be (spuriously) correlated with factors that are exerting a time persistent effect on corruption. To stay with the example from before: lightning prone areas that were rich in natural resources in (say) 1970 most likely remain so, and if the presence of natural resources fueled corruption in the 1970s, it would probably also spur corruption at the beginning of the 21st century. Accordingly, if lightning density is picking up this sort of influence we would expect to see a fairly time invariant correlation between lightning and (changes in) corruption over the last three or four decades. However, if lightning density is picking up the influence from processes instigated by the emergence of the Internet (be that a growth surge, an acceleration in human capital accumulation, improved diffusion of information, or E-government), one would expect to see a time varying correlation between lightning and (changes in) corruption. In fact, instrument validity would require that there is *no* correlation between lightning density and changes in corruption before 1991 (the founding year of the World Wide Web). These considerations form the basis of a falsification test: If lightning is correlated with (changes in) corruption prior to 1991, it is unlikely to satisfy the exclusion restriction.

As documented below, lightning density exhibits a *time-varying correlation* with corruption; the reduced form relationship between lightning density and corruption does not exist prior to the inception of the World Wide Web. Using U.S. state-level data, as well as cross country data, we are able to establish that the lightning/corruption correlation only exists after 1991. This falsification test makes probable that the lightning instrument is capturing processes instigated by the emergence of the Internet, and it therefore supports the use of lightning as an instrument for Internet diffusion.

Against this background we employ lightning density as an instrument for Internet diffusion. Our 2SLS estimates corroborate the OLS results: Rising Internet use over the 1990s reduced corruption in the U.S. and across countries. Our results are admittedly stronger, statistically speaking, for the U.S. Still, our cross-section analysis does suggest that the U.S. state-level results generalize to an international setting.

The present research is related to the literature which studies the determinants of the level of corruption. Notable contributions include Ades and di Tella (1999), Treisman (2000), Brunetti and Weder (2003), Persson, Tabellini, and Trebbi (2003), Glaeser and Saks (2006), and Licht, Goldscmidt, and Schwartz (2007).

Since the Internet is a central source of information, the paper is also related to the political economy literature that studies the impact of information on governance more generally. This literature suggests that a better informed public serves to discipline the political establishment, thus affecting governance (e.g. Besley and Burgess 2002; Strömberg, 2004; Reinikka and Svensson 2004; Eisensee and Strömberg 2007; Ferraz and Finan 2008).

Finally, the paper is related to the literature which studies the determinants of the spread of the personal computer (e.g. Caselli and Coleman 2001) and the Internet (e.g. Chinn and Fairlie 2007) across countries. We add to this literature by documenting a link between lightning density and the speed of Internet diffusion.

The paper is structured as follows. In Section I we present our empirical specifications of choice. Section II outlines the identification strategy in detail; in particular, we explain how lightning activity impacts digital equipment. Section III provides an analysis of how the Internet has affected corruption across the U.S. states, whereas Section IV provides cross-country evidence. Section V concludes.

I. Specification

As argued above, we wish to understand the impact of the Internet (or World Wide Web) on the evolution of corruption, from the emergence of the former and onwards. Since the Internet is a General Purpose Technology it seems infeasible to try and isolate any particular mechanism linking the Internet to corruption (e.g., growth spurts, human capital, enhanced dissemination of information, etc.), for which reason we focus on the reduced form.

Accordingly, in the analysis to follow we mainly rely on the following parsimonious specification:

$$\Delta C_i = \alpha_0 + \alpha_1 \Delta INTERNET_i + \alpha_2 C_{initial,i} + \varepsilon_i, \qquad (1)$$

where ΔC_i is the change in corruption levels between an initial and a final year, $C_{final,i} - C_{initial,i}$, and $\Delta INTERNET_i$ is the change in Internet penetration, $INTERNET_{final,i} - INTERNET_{initial,i}$.[6] When $\Delta C_i > 0$ means *increasing* corruption, the key hypothesis under investigation is whether $\alpha_1 < 0$ or not.

There are several reasons why we adopt this lagged dependent variables specification as our preferred specification. First, corruption is a naturally bounded variable. In the U.S. setting we employ corruption convictions as our measure of C, which is bounded from below at zero. Similarly in the

6. In the cross-state sample we also condition on state population size. State population enters as a control because we use total corruption convictions as our measure of state corruption; by including state size we thereby ensure that all scale effects are pruned from the data in a simple way.

cross-country analysis we employ a corruption index (the *ICRG* index), which is confined to a particular interval (zero to six) by construction. Obviously, for all country observations near either of the two endpoints, and in states near the zero boundary, one should expect the evolution of corruption to be subject to *mean reversion*. Unless we thus control for $C_{initial,i}$, our empirical model is likely to be misspecified. To spell it out formally, assume that equation (1) is the true population model, and suppose that $Cov(\varepsilon, \Delta INTERNET) = 0$. If we mistakenly ignore the lagged dependent variable and instead focus on the first difference equation, in an attempt to kill off an imagined fixed effect, we would perform OLS on $\Delta C_i = \alpha_0 + \alpha_1^{FD} \Delta INTERNET_i + \upsilon_i$, where $\upsilon_i \equiv \alpha_2 C_{initial,i} + \varepsilon_i$. The probability limit of α_1 obtained via the first differenced estimator is

$$\text{plim } \alpha_1^{FD} = \alpha_1 + \alpha_2 \frac{Cov(C_{initial}, \Delta INTERNET)}{Var(\Delta INTERNET)}. \tag{2}$$

Unless $Cov(C_{initial}, \Delta INTERNET) = 0$ we have that $\alpha_1^{FD} \neq \alpha_1$. Indeed, *a priori* $\alpha_1^{FD} \approx 0$ would be a likely outcome. To see this, note that the mean reversion process implies that $\alpha_2 < 0$. Moreover, if places with high initial corruption levels tend to see a slower diffusion of the internet then $Cov(C_{initial}, \Delta INTERNET) < 0$. Taken together this implies that $\alpha_2 \cdot Cov(C_{initial}, \Delta INTERNET) > 0$, which implies that the expected *negative* point estimate for α_1 is biased towards zero (see equation (2)).

Second, another virtue of including the initial level of corruption is that it automatically controls for (a potentially large set of) variables which may influence the evolution of corruption. To see the latter point more clearly observe that equation (1) is equivalent to a levels regression with a lagged dependent variable:

$$C_{final,i} = \alpha_0 + \alpha_1 \Delta INTERNET_i + (\alpha_2 + 1)\, C_{initial,i} + \varepsilon_i. \tag{3}$$

Accordingly, all time invariant structural characteristics affecting the *level* of corruption will be picked up by $C_{initial,i}$. This reduces the scope for omitted variable bias in contaminating the estimate of α_1 (Wooldridge 2003, p. 300).[7] Third, the core part of our analysis involves 2SLS estimation, where lightning density is invoked as an instrument for $\Delta INTERNET_i$. Since lightning density and the *level* of corruption are highly correlated, one may harbor legitimate concerns about the exclusion restriction if the latter is omitted from the empirical model; if the evolution of corruption is subject to mean reversion (i.e., if $\alpha_2 < 0$), trouble arises (see again equation (2)).

Naturally, one might also worry that the exclusion restriction is jeopardized by the omission of fixed effects in equation (1). But in the analysis below we

7. See also Angrist and Pischke (2009, Ch. 5.3), for a discussion of the virtues of using the specification in equation (1) *vis-à-vis* the first difference model.

are able to gauge the likely severity of this problem by performing the falsification test mentioned above: If the true empirical model involves fixed effects (i.e., α_0 in equation (1) should be country specific), and if our instrument is correlated with these unobservables, we would expect to see that lightning density is correlated with ΔC_i before as well as after 1991 (the founding year of the World Wide Web). As shown below, however, we are unable to reject the absence of a correlation between our instrument and changes in corruption *prior* to 1991. This holds true across U.S. states as well as across countries. As a result, we view the lagged dependent variables specification as the relatively safe choice when trying to elicit information about the causal impact of the Internet on corruption. Nevertheless, we will in the interest of completeness report the results from the first difference specification as well; that is, we also report the results from assuming $\alpha_2 = 0$ in equation (1).

II. Identification Strategy

There is good reason to believe that the adoption of new technologies, such as the Internet, is endogenous to governance. New technologies may create political as well as economic losers, for which reason incumbent entrepreneurs and politicians may try to block adoption (Mokyr 1990; Parente and Prescott 1999; Acemoglu and Robinson 2001). It seems plausible that places with widespread corruption, for example, may have adopted the Internet later, due to the influence of politicians, civil servants, or both. This mechanism rationalizes a positive impact of governance on the number of Internet users. Consequently, OLS estimation is unlikely to identify the impact of the Internet on corruption. To address this concern we employ an IV approach, the logic of which we now describe.

Computers are highly sensitive to even ultra brief power disruptions. Such disruptions are likely to cause down-time, though sudden power surges may also damage the equipment and randomly destroy or alter data. As observed in *The Economist*:[8] "For the average computer or network, the only thing worse than the electricity going out completely is power going out for a second. Every year, millions of dollars are lost to seemingly insignificant power faults that cause assembly lines to freeze, computers to crash and networks to collapse."

The reason why IT equipment is so sensitive to power disturbances is that computers are constructed to work under a clean electrical current, featuring a particular frequency and amplitude of voltage. The alternating power emanating from the commercial power plant is converted into direct current, after which transistors turn this small voltage on and off at several gigahertz during digital processing (Kressel 2007). However, if the input, in the form of the alternating current, is disturbed or distorted the conversion process is

8. The power industry's quest for the high nines, *The Economist*, March 22, 2001.

corrupted, which may in turn result in equipment failure and damage. Indeed, voltage disturbances measuring less than one cycle are sufficient to crash servers, computers, and other microprocessor-based devices; that is, at a 60 Hz frequency (the standard in the U.S.) this means that a power disturbance of a duration less than 1/60 th of a second is enough to crash a computer (Yeager and Stalhkopf 2000; Electricity Power Research Institute 2003). Importantly, this issue is unlikely to diminish over time as the sensitivity to small power distortions increases with the miniaturization of transistors, which is the key to increasing speed in microprocessors (Kressel 2007).

Accordingly, in areas with more power disturbances, the user cost of IT capital will be higher due to a higher rate of IT capital depreciation (Hall and Jorgenson 1967). By implication, the desired IT capital stock will be lower, reducing IT investments and the speed of Internet diffusion. Of course, steps may be taken to protect the equipment from power disturbances. A high-quality surge protector provides protection against voltage spikes, for example. High-tech companies install generators to supplement their power needs, thereby insuring themselves against power failure. They also add uninterruptible power sources relying on batteries to power computers until generators kick in. However, these initiatives will in any case increase the costs of acquiring digital equipment, and thereby the user cost of IT capital. The crux of the matter is that if one lives in an environment with low power quality, this adds to the costs of a computer.[9]

To this one may add that in areas with frequent power disruptions and outages, the marginal benefit of owning a computer is probably lowered as well. Obviously, in countries where firms and consumers face regular power outages it will be difficult to employ IT efficiently. But even if power disruptions are infrequent and of very short duration, power disruptions lead to glitches and downtime which serves to lower the productivity of IT equipment. Both mechanisms, higher marginal costs and lower marginal return/benefit, imply that poorer power quality should lead to a slower speed of Internet diffusion.

Naturally, power quality is not exogenous; it may well be determined by governance.[10] As a result, we employ a variable which generates exogenous variation in power quality, and thus IT costs and benefits, as an instrument for Internet diffusion. A natural phenomenon which interferes with digital equipment, by producing power failures, is lightning activity (e.g., Shim *et al.* 2000, Ch. 2; Chisholm 2000). By all accounts, the influence of lightning on power quality is substantial. According to some estimates, lightning is the direct cause

9. Besides, the above mentioned protective devises are not necessarily enough to ensure against damage. According to the National Oceanic and Atmospheric Administration (NOAA), a typical surge protector will not protect equipment from a nearby lightning strike. Generators, in turn, do not react fast enough and can deliver dirty power; batteries are expensive to maintain and may also not react fast enough. See e.g., The power industry's quest for the high nines, *The Economist* (March 22, 2001).

10. See Fredriksson *et al.* (2004) for evidence that corruption affects energy supply.

of one third of all power quality disturbances in the United States (Chisholm and Cummins 2006). Moreover, the probability of lightning-caused power interruptions or equipment damage scales linearly with lightning density (Chisholm 2000; Chisholm and Cummins 2006).[11] As a result, in areas with greater lightning density (strikes per square km per year) the (expected) rate of IT capital depreciation will tend to be larger. This implies higher IT investment costs, and possibly lower IT productivity as well.

The problems associated with lightning activity in the context of IT equipment have not escaped the attention of the popular press. A recent article in *The Wall Street Journal* highlights the practical relevance of lightning activity, and stresses the difficulty in shielding IT equipment:[12] "Even if electricity lines are shielded, lightning can cause power surges through unprotected phone, cable and Internet lines - or even through a building's walls. Such surges often show up as glitches. "Little things start not working; we see a lot of that down here," says Andrew Cohen, president of Vertical IT Solutions, a Tampa information-technology consulting firm. During the summer, Vertical gets as many as 10 calls a week from clients with what look to Mr. Cohen like lightning-related problems. Computer memory cards get corrupted, servers shut down or firewalls cut out."

Against this background we propose lightning density as an instrument for the speed at which Internet use per capita changed over the period in question. Schematically, we can express the theory underlying our identification strategy in the following way:

$$\text{LIGHTNING DENSITY} \rightarrow \text{POWER DISTURBANCES} \rightarrow \text{INTERNET USE}, \quad (4)$$

where the second arrow implicitly subsumes the impact of power disturbances on the costs and benefits of IT capital.

Lightning is certainly external in the sense of Deaton (2010). But this, of course, does not imply that it fulfills the exclusion restriction required for instrument validity. In particular, it could conceivably correlate with fixed factors (natural resource endowment, say) which themselves exert a persistent effect on the evolution of corruption. In order to examine whether this is likely to be the case or not, we perform a falsification test below: Under the hypothesis that lightning influences the evolution of corruption, via the Internet, changes in corruption should only be correlated with lightning after the emergence of the World Wide Web. We can in fact reject that lightning and changes in corruption are correlated prior to the invention of the World Wide Web

11. This linear scaling can be expressed precisely. Let N_S denote the number of strikes to a conducter per 100*km* of power line length, h the average height (in meters) of the conducter above ground level, and GFD the ground flash density, then $N_S = 3.8 \cdot GFD \cdot h^{0.45}$ (see Chisholm 2000).

12. There Go the Servers: Lightning's New Perils. *The Wall Street Journal*, August 25, 2009.

(i.e., prior to 1991) using data for both the U.S. States and our world sample. This suggests quite strongly that lightning is not – spuriously – correlated with country specific factors that persistently affect the path of corruption, thus supporting the exclusion restriction.

III. Cross-State Evidence

The cross-state analysis proceeds as follows: We first provide details on the data used for the analysis. Next we provide evidence on the partial correlation between changes in Internet usage and changes in corruption levels. Finally, before proceeding to our 2SLS estimates, we provide an independent check of the validity of our identification strategy.

Data

In measuring corruption in the U.S. we follow Glaeser and Saks (2006) by employing corruption convictions. The data derives from the Justice Department's Report to Congress on the Activities and Operations of the Public Integrity Section, a publication which provides statistics on the nation-wide *federal* effort against public corruption, including the number of federal, state, and local public officials convicted of a corruption-related crime by state. As argued by Glaeser and Saks (2006), federal conviction levels capture the extent to which federal prosecutors have charged and convicted public officials for misconduct. There are potential problems with using conviction rates to measure corruption: in corrupt places, the judicial system is itself likely to be corrupt, meaning that fewer people will be charged with corrupt practices. This problem, however, is diminished when using federal convictions, the reason being that the federal judicial system is somewhat isolated from local corruption. Consequently, it should treat people similarly across states (Glaeser and Saks, 2006).

In concrete terms, we measure the *change* in corruption between 1991 and 2006 by the log difference in (one plus) the total number of corruption convictions in the two years; positive values in the rate of change reflect *increasing* corruption, and the choice of initial year follows from the fact that the Internet (in the sense of the World Wide Web) was introduced in 1990 (first Web page went online in 1991).[13]

The second key variable is Internet use, which we measure as the percentage of households with Internet access. It is based on data collected in a supplement to the October 2003 Current Population Survey (CPS), which includes

13. Note that an increased use of the Internet will both increase the *risk of detection* for a corrupt official (the detection technology is improved) as well as lower the *incentive* to commit corrupt acts. Hence, in theory, increased Internet use could increase the number of convictions if the former effect dominates. It might thus seem as if the Internet increases corruption. However, empirically the net effect is negative, implying that the incentive effect dominates, as documented below.

questions about computer and Internet use.[14] The CPS is a multi-stage probability sample with coverage in all states. The sample was selected from the 1990 Decennial Census files and is continually updated to account for new residential construction. To obtain the sample the United States is divided into 2,007 geographic areas, and about 60,000 households are eligible for interviews. Since U.S. corruption data goes back to 1991, the launch date of the WWW, we define the change in Internet use by state population as $\Delta INTERNET_i = INTERNET_{i,2003} - INTERNET_{i,1991} = INTERNET_{i,2003}$, since $INTERNET_{i,1991} = 0$ for all i (i.e., for all states).

In some specifications we also control for initial state population, or the growth of state population, which derives from the U.S. Bureau of Economic Analysis (BEA). The natural log of initial state population is denoted $\log(POP_{1991})$ and population growth is $\Delta \log(POP) = \log(POP_{2006} / POP_{1991})$. The purpose of including these controls is to make sure that all scale effects are pruned from the corruption data. Obviously, the *total* number of corruption convictions is likely greater in more populous states.

Finally, we measure lightning as the number of ground strikes per square kilometer per year. Accurate cloud-to-ground data exist for the 48 contiguous U.S. states, are measured by the U.S. National Lightning Detection Network (NLDN), and are provided by Vaisala for the period 1996-2005. NLDN consists of numerous remote, ground-based lightning sensors, which instantly detect (at a very high level of accuracy) the electromagnetic signals given off when lightning strikes Earth's surface.[15] Summary statistics are reported in Appendix Table A.1.

Partial Correlations

Table 1 documents the partial correlation between changes in Internet use and changes in corruption for the 48 contiguous U.S. states.

As a first step we report, in Columns 1 and 2, the results from the first difference specification where we thus omit initial corruption as a control; Column 1 is the basic specification, whereas population growth is included in Column 2 alongside Internet users. States that have seen larger increases in Internet use have also experienced larger reductions in corruption convictions, consistent with a corruption dampening effect of the Internet.

As explained in Section 2 we are concerned about the exclusion of the initial corruption level in the first difference specification, as changes in corruption convictions likely are subject to mean reversion. Specifically, in states with zero initial convictions one would almost inevitably expect an increase in convictions subsequently, regardless of Internet use.

In order to check whether this is an actual problem in the U.S. sample, we experiment, in Columns 3 and 4, with omitting states with zero corruption

14. <http://www.census.gov/population/socdemo/computer/2003/tab01B.xls>.

15. The data can be obtained, free of charge, here: <http://www.vaisala.com/thunderstorm>

TABLE 1. OLS Regressions, U.S. Sample

	Dependent variable: Change in corruption convictions 1991-2006					
	(1)	(2)	(3)	(4)	(5)	(6)
ΔINTERNET	−0.058**	−0.057**	−0.060**	−0.057*	−0.058**	−0.051***
	(0.027)	(0.028)	(0.029)	(0.030)	(0.024)	(0.019)
Δlog(POP)		−0.243		−0.914		
		(1.068)		(0.972)		
$\log(POP_{1991})$						0.849***
						(0.145)
$\log(1 + CC_{1991})$					−0.386***	−0.873***
					(0.096)	(0.135)
Constant	3.667**	3.655**	3.577**	3.543**	4.392***	−7.790***
	(1.448)	(1.469)	(1.584)	(1.578)	(1.312)	(2.201)
Observations	48	48	40	40	48	48
R-squared	0.09	0.09	0.11	0.13	0.29	0.55

Notes: Robust standard errors in paranthesis. Asterisks ***, ** and * signify, respectively, p-value < 0.01, p-value < 0.05 and p-value < 0.1. All variables are defined in the main text.

Source: Authors' analysis based on data described in the text.

convictions in 1991. This should in theory reduce the extent of mean reversion in the sample. Further, in Columns 5 and 6, we employ the lagged dependent variable specification on the full sample: our preferred specification.

As seen from Columns 5 and 6, initial corruption is a significant correlate with subsequent changes in corruption; it adds considerably to the explanatory power of the regression model, as evidenced by the increase in R^2 when comparing Columns 1 and 2 with Columns 5 and 6.[16] In Columns 3 and 4 we truncate the sample by excluding states with zero initial corruption. If one includes initial corruption in this setting, it is insignificant (not shown). This shows that the significance of initial corruption in Columns 5 and 6 is due to mean reversion in the U.S. sample. Nevertheless, despite the presence of mean reversion in corruption we obtain nearly identical results for $\Delta INTERNET$ in all columns. The reason is simply that, statistically speaking, $Cov(C_{initial}, \Delta INTERNET) = 0$ in the U.S. sample. As a result, by equation (2), the first difference estimator is consistent; i.e., plim $\alpha_1^{FD} = \alpha_1$.[17] In any event, the bottom line is that the partial correlation is robust to the exact choice of specification.

The partial correlation is also robust to a long list of additional controls beyond what is reported in Table 1. Maxwell and Winters (2005) propose four sets of fundamental traits of U.S. states, which should have predictable effects

16. We have also experimented with including the growth rate of population, rather than its level, in the lagged dependent variables specification (Column 5 and 6), but much like in Column 2 and 4 it comes out insignificant.

17. By including the lagged corruption level, we can, however, reduce the error variance.

on corruption: (i) number of corruptible government bodies, (ii) the size of the state, (iii) socio-ethnic homogeneity, and (iv) civic-minded and well-informed political cultures. The authors consider seven additional control variables including income growth, general tax revenue and campaign expenditure restrictions. We have experimented with the inclusion of all the proposed corruption determinants (alongside initial corruption), and we find that changes in Internet use (1991–2006) remains significant. The results are available upon request.

Despite these encouraging results concerns about causality may legitimately be raised. We therefore further scrutinize the Internet/corruption nexus by way of instrumental variables estimation.

Instrument Falsification Test

2SLS estimates are obviously no better than the invoked instrument. In the present context we propose the use of lightning density as an instrument for Internet use, based on the theoretical argument we presented in Section II.

According to the identification strategy lightning is only allowed to influence changes in corruption via Internet use. If, by contrast, it is correlated with unobserved determinants of the evolution of corruption, the exclusion restriction fails, and the 2SLS estimates are no better than the OLS estimates discussed above.

Under the null of instrument validity, however, lightning should exert a time varying impact on the path of corruption. The reason is simply that the Internet is a new technology which cannot possibly have affected corruption prior to its inception and widespread use. By extension, determinants of Internet use should not affect corruption prior to the emergence of the World Wide Web.

Hence, in an effort to try to falsify our instrument we run the following regression:

$$\Delta C_i = \alpha_0 + \alpha_1 \log(LIGHTNING_i) + \alpha_2 C_{initial,i} + \varepsilon_i, \qquad (5)$$

on two different time periods. First, we examine the link between lightning and changes in corruption during the Internet era: 1991 to 2006. In this period we expect a significantly positive point estimate for lightning, suggesting that high lightning states have experienced slower Internet diffusion and therefore smaller reductions in corruption levels than states with less lightning. Second, we examine the link between lightning and changes in corruption *prior* to the emergence of the World Wide Web. For instrument validity, the partial correlation between lightning and corruption should not be significantly different from zero. In the event of a significant correlation between lightning and changes in corruption we are forced to conclude that lightning is correlated with factors beyond the Internet (and the initial level of corruption), which influences changes in corruption. If so, the instrument is invalid.

The falsification test is potentially helpful in another respect: If changes in corruption are not correlated with lightning strikes in the pre-Internet age, the scope for lightning operating on corruption via other electronic technologies is limited. Indeed, it would appear unlikely that the impact date of such technologies coincides with the inception of the WWW. The introduction of the microprocessor, for instance, goes back to the early 1970s. By that time computers were widely used by the U.S. government. Therefore, if we think that increased and more efficient information storage capacity, say, reduces the scope for corruption, we should observe a correlation between lightning and corruption in the pre-Internet era. To the extent that this relationship is absent, the microprocessor *per se* is unlikely to influence the evolution of corruption.

Table 2 reports the results from this check. In Columns 1-4 we examine the pre-Internet periods 1976–1990, where we have missing observations for corruption in three states, and 1978–1990, where we have a full sample. The take-away is that lightning is not correlated with changes in corruption during these sub-periods, in keeping with the requirement that lightning is uncorrelated with other factors which exert a persistent impact on the evolution of corruption. Figure 1 provides a visual illustration of these findings.

In Columns 5 and 6 we shift attention to the period following the emergence of the World Wide Web. During this period lightning emerges as a strong correlate with changes in corruption levels: In states with more lightning we observe a slower rate of reduction in corruption compared to states with less lightning. This is consistent with the hypothesis that the Internet has worked to lower corruption, and that lightning has worked to slow down Internet diffusion. Figure 2 provides a visual illustration of the partial correlation between lightning and changes in corruption 1991–2006.

It is important to observe that the insignificance of lightning during the pre-Internet era is not simply a matter of imprecise estimates. As is evident from inspection of the results across columns, the size of the estimate itself rises many fold when moving from Columns 1–4 to Columns 5–6.

It may also be noted that the falsification test is consistent with lightning *not* operating through other electronic technologies, which would arguably have required that the correlation between lightning and corruption was in existence before 1991.

Overall, we find that these results constitute compelling evidence in favor of the exclusion restriction in the U.S. context. Hence, we now move to instrumental variables estimation where lightning is used as an instrument for Internet diffusion.

2SLS Estimates

Table 3 reports our 2SLS results; Panel A reports first stage results, whereas Panel B reports the second stage. In keeping with the approach taken in Table 1, we experiment with both the first difference specification and the lagged dependent variables specification. In the OLS setting (Table 1) we found

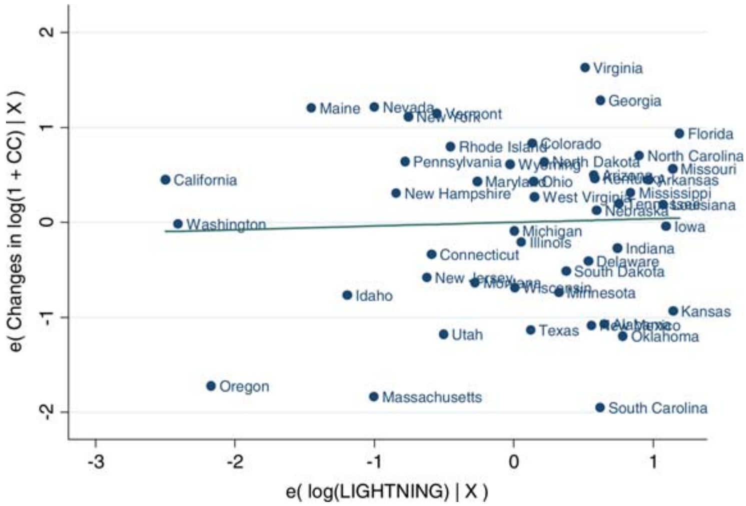
TABLE 2. Falsification test, U.S. Sample

	Change in corruption convictions 1976–1990		Change in corruption convictions 1978–1990		Change in corruption convictions 1991–2006	
Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)
log(LIGHTNING)	−0.068	−0.027	0.019	0.016	0.344**	0.278**
	(0.183)	(0.110)	(0.198)	(0.113)	(0.145)	(0.109)
$\log(1 + CC_{1976})$	−0.388***	−0.954***				
	(0.143)	(0.162)				
$\log(POP_{1976})$		0.908***				
		(0.185)				
$\log(1 + CC_{1978})$			−0.352**	−0.930***		
			(0.170)	(0.185)		
$\log(POP_{1978})$				0.908***		
				(0.153)		
$\log(1 + CC_{1991})$					−0.452***	−0.916***
					(0.109)	(0.142)
$\log(POP_{1991})$						0.833***
						(0.147)
Constant	1.654***	−11.220***	1.469***	−11.304***	0.787**	−10.723***
	(0.313)	(2.733)	(0.358)	(2.235)	(0.347)	(2.051)
Observations	45	45	48	48	48	48
R-squared	0.15	0.49	0.10	0.44	0.29	0.54

Notes: Robust standard errors in paranthesis. Asterisks ***, ** and * signify, respectively, p-value < 0.01, p-value < 0.05 and p-value < 0.1. All variables are defined in the main text.

Source: Authors' analysis based on data described in the text.

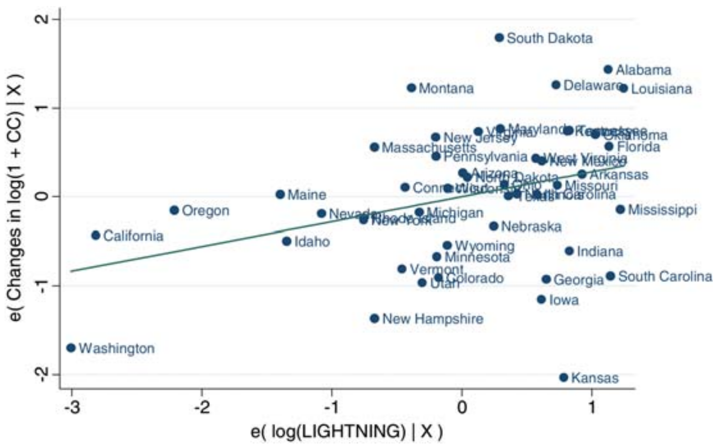
FIGURE 1. Reduced form in the pre-Internet era, U.S. sample



Note: The figure shows the association between lightning and changes in corruption convictions over the 1978–1990 period, with the influence of initial population partialled out.

Source: Figure is based on data described in the text.

FIGURE 2. Reduced form in the Internet era, U.S. sample



Note: The figure shows the association between lightning and changes in corruption convictions over the 1991–2006 period, with the influence of initial population partialled out.

Source: Figure is based on data described in the text.

virtually no difference in estimation results using either the first differenced or the lagged dependent variables specification. This is not true in the 2SLS setting.

In the first stage results are very similar across specifications. Lightning proves to be a strong instrument in explaining changes in Internet use; the first

TABLE 3. 2SLS regressions, U.S. sample

	Panel A: First stage					
	Dependent variable: Change in Internet users 1991–2006					
	(1)	(2)	(3)	(4)	(5)	(6)
log(LIGHTNING)	−3.583***	−3.505***	−3.821***	−3.764***	−3.801***	−3.805***
	(0.604)	(0.606)	(0.756)	(0.758)	(0.651)	(0.678)
Δlog(POP)		3.443		3.947		
		(5.099)		(5.935)		
log(POP_{1991})						0.050
						(0.794)
log(1 + CC_{1991})					0.721	0.693
					(0.526)	(0.703)
Constant	60.54***	59.83***	61.41***	60.70***	59.55***	58.86***
	(1.080)	(1.336)	(1.403)	(1.547)	(1.378)	(10.870)
F-test (H0: log(lightning) = 0)	35.23	33.49	25.54	24.66	34.11	31.48
	Panel B: Second stage					
	Dependent variable: Change in corruption convictions 1991–2006					
ΔINTERNET	−0.058	−0.056	−0.104***	−0.102***	−0.090**	−0.073**
	(0.043)	(0.046)	(0.035)	(0.037)	(0.036)	(0.030)
Δlog(POP)		−0.248		−0.606		
		(1.117)		(1.026)		
log(POP_{1991})						0.836***
						(0.146)
log(1 + CC_{1991})					−0.386***	−0.866***
					(0.097)	(0.137)
Constant	3.670	3.626	5.959***	5.928***	6.171***	−6.416**
	(2.343)	(2.413)	(1.924)	(1.924)	(1.949)	(2.708)
Observations	48	48	40	40	48	48
R-squared	0.09	0.09	0.05	0.07	0.26	0.54

Notes: Robust standard errors in paranthesis. Asterisks ***, ** and * signify, respectively, p-value < 0.01, p-value < 0.05 and p-value < 0.1. All variables are defined in the main text.

Source: Authors' analysis based on data described in the text.

stage F-statistic exceeds 10 in all columns, documenting that the 2SLS analysis in the U.S. sample is not plagued by weak identification (cf., Staiger and Stock 1997).

In the second stage results differ markedly across specifications (Panel B). In the first differenced specification Internet use is estimated imprecisely, with a point estimate virtually identical in size and sign to what we found in the OLS setting. In contrast, when we exclude either states with zero initial corruption (Columns 3 and 4), or when we control for initial corruption levels directly (Columns 5 and 6), changes in Internet use is estimated with high precision.

What should we make of the differences in size and significance of the Internet across specifications? The most straightforward explanation runs as follows: Recall from Section 1 that, under the maintained hypothesis that equation equation (1) is the true specification, the first difference equation is given by $\Delta C_i = \alpha_0 + \alpha_1^{FD} \Delta INTERNET_i + v_i$, where $v_i \equiv \alpha_2 C_{initial,i} + \varepsilon_i$. When we employ 2SLS with lightning as an instrument, the exclusion restriction is that $Cov(LIGHTNING, v) = 0$; since $Cov(LIGHTNING, \varepsilon) = 0$ by assumption, $Cov(LIGHTNING, v) = 0$ requires that $Cov(LIGHTNING, C_{initial}) = 0$. This latter condition is violated in the U.S. sample. The failure of the exclusion restriction implies that the 2SLS estimates are biased towards OLS, which explains the similarity of the results reported in Columns 1-2 in Tables 1 and 3.

In contrast, if we either eliminate the influence from initial corruption on changes in corruption by way of sample truncation (i.e., omission of states with zero initial corruption), or by including the initial level of corruption directly, the exclusion restriction is much more likely to be satisfied. The results reported in Columns 3-6 are therefore much more likely to identify the causal influence of Internet diffusion on changes in corruption. In all four columns where the exclusion restriction is plausible we find evidence that the Internet has worked to lower corruption in states where its diffusion was rapid, compared to states where Internet diffusion occurred at a slower rate.

What is the economic significance of the Internet? To answer this question we begin with the levels specification associated with equation (3):

$$\log(1 + C_{2006}) = \alpha_0 + \alpha_1 \Delta INTERNET + (\alpha_2 + 1)\log(1 + C_{1991}). \quad (6)$$

If we linearize, treating C_{1991} as a constant, the following simple approximation emerges:

$$\Delta C_{2006} \approx \alpha_1 (1 + C_{2006})\, INTERNET_{2003}, \quad (7)$$

where we have used that $\Delta INTERNET = INTERNET_{2003}$. Next, to gauge economic significance, consider moving from the median to the third quartile in the distribution of Internet users in 2003; this is equivalent to an increase of 3.1 Internet users per 100 people. Using the 2SLS results reported in Column 6 of Table 3 (the most conservative estimate) in equation (7) we find that $\Delta C_{2006} \approx (-0.073) \cdot (1 + 16) \cdot (3.1) = -3.85$ yearly convictions. This would correspond to moving from the median to the 33rd percentile in the U.S. state corruption convictions ranking in 2006. Accordingly, our results suggest that the introduction of the Internet has reduced U.S. corruption levels below what would otherwise have been observed absent this technology.

IV. Cross-Country Evidence

In this Section we examine whether the results obtained above generalize to a cross-country sample. After providing details on the data used in the cross-country setting, we discuss the partial correlation between changes in corruption and changes in Internet users, present an instrument falsification test, and discuss our 2SLS estimates.

Data

As our main measure of corruption we employ the well-known *ICRG* index. This index has the useful property that it is available from 1984, which enables us to perform the type of instrument falsification test employed in the context of the U.S. state-level analysis. The coverage for the pre-Internet era is shorter (six years, 1984-1990) than in the U.S. state-level analysis but hopefully long enough for us to clarify whether lightning is correlated with changes in corruption prior to the emergence of the World Wide Web. Changes in the *ICRG* index are calculated as the absolute change in the index between relevant years (1991-2006 and 1984-1990, respectively). The indicator is bounded between zero and six, with larger values of the index meaning *less* corruption.

As a matter of robustness we have also checked our main results (OLS and 2SLS) using the control of corruption index compiled by Kaufmann *et al.* (2007). This indicator is often used in cross-national studies (Treisman 2007). But it is unfortunately only available for the period 1996 onwards. Accordingly, the falsification test of our instrument cannot be performed with this indicator, for which reason we have chosen to stick with the *ICRG* index. But it should be noted that our OLS and 2SLS results for the Internet era are robust to the use of the control of corruption indicator; these results are available upon request.

Our key explanatory variable is the number of Internet users per 100 people. Increasingly, the number of Internet users is based on regular surveys. In situations where surveys are not available, an estimate can be derived based on the number of subscribers. Data is compiled by the International Telecommunication Union (ITU) and made available in the World Development Indicators (WDI) 2007. Changes in Internet users are calculated as in the context of our U.S. state-level analysis. When we examine the period 1991-2006, using the *ICRG* indicator, we use the approximation that $INTERNET_{i,1991} = 0$, as in the analysis of the U.S. states.

In the U.S. context data on lightning density derived from ground detectors. Naturally, such data is not available across the world. Instead we employ satellite data on lightning intensity. The raw data (strikes per km^2 per year) is provided by the National Aeronautics and Space Administration (NASA). Specifically, we rely on the data from the so-called Optical Transient Detector (OTD), a space based sensor launched on April 3, 1995. For a period of roughly 5 years the satellite orbited Earth once every 100 minutes at an altitude of 740 km. At any given instant it viewed a 1300 km × 1300 km region of Earth. Lightning is determined by comparing the luminance of adjoining frames of OTD optical data. When the difference was larger than a specified threshold value, an event was recorded.[18] These satellite-based data are archived and cataloged by the The Global Hydrology and Climate Center, where they are also made publicly available.[19] We apply the data from a high-resolution (0.5 degree latitude × 0.5 degree longitude) grid of total lightning bulk production across the planet, expressed as a flash density, from the completed 5 year OTD mission.[20] Figure 3 provides a world map of the average flash density over the 5 years period.

We construct average flash densities for each country by first mapping the corresponding geographic areas into the lightning data grid and then taking the average of flash densities within each of these areas. The coordinates describing the areas are taken from the GEOnet Names Server (GNS) at the U.S. National Geospatial-Intelligence Agency's (NGA),[21] and the U.S. Board on Geographic Names' (U.S. BGN) database of foreign geographic names and features.[22] We used the GNS database released on October 7, 2008.[23] The GNS data covers the entire planet with the exception of the U.S. and the Antarctica. The area for the U.S. was estimated using geographic features for the 48 contiguous U.S. states, contained in the database released on August 15, 2008 by the Geographic Names Information System at the U.S. BGN.[24]

A potential problem with the OTD data is that it only provides observations on *total* lightning events; i.e., intra-cloud, cloud-to-cloud, cloud-to-sky, and cloud-to-ground lightning. In other words, OTD data does not separate out *cloud-to-ground* lightning incidences. The pertinent characteristic of lightning in the evaluation of risk to electronic equipment and electric power systems is the cloud-to-ground flash density.

In an effort to examine whether the satellite data is likely to be a good proxy for ground strikes we compared our data on ground strikes for

18. Basically, these optical sensors use high-speed cameras designed to look for changes in the tops of clouds. By analyzing a narrow wavelength band (near-infrared region of the spectrum) they can spot brief lightning flashes even under daytime conditions.

19. <http://thunder.msfc.nasa.gov/data/#OTD_DATA>

20. <ftp://microwave.nsstc.nasa.gov/pub/data/lightning-satellite/lis-otd-climatology/HRFC/LISOTD_HRFC_V2.2.hdf>

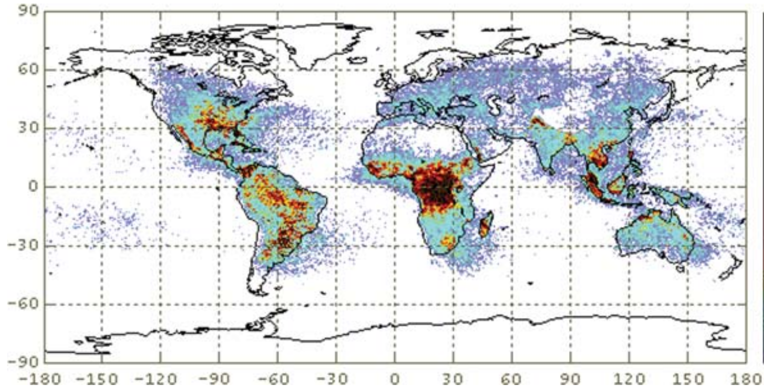
21. <http://earth-info.nga.mil/gns/html/namefiles.htm>

22. <http://geonames.usgs.gov/domestic/download_data.htm>

23. <ftp://ftp.nga.mil/pub2/gns_data/geonames_dd_dms_date_20081007.zip>

24. <http://geonames.usgs.gov/docs/stategaz/NationalFile_20080815.zip>

FIGURE 3. Global lightning map



Note: Average flash density (flashes per year per km^2). The figure is constructed using the OTD Global Lightning Distributions for the period April 12, 1995 to December 31, 1999.
Source: Figure is based on data described in the text.

individual U.S. states with corresponding satellite-based lightning data. Figure 4 provides a scatter plot of total lightning against cloud-to-ground lightning for the U.S. sample. Note that by definition total lightning should be at least as large as cloud-to-ground lightning, which is confirmed by the figure; all observations are below the 45 degree line.[25]

Figure 4 shows that there is agreement between the two measures of lightning. With a correlation above 0.95, total lightning appears to be a reasonable proxy for cloud-to-ground lightning in the contiguous U.S. states sample. This is also expected to be the case in the cross-country data (Chisholm and Cummins, 2006).

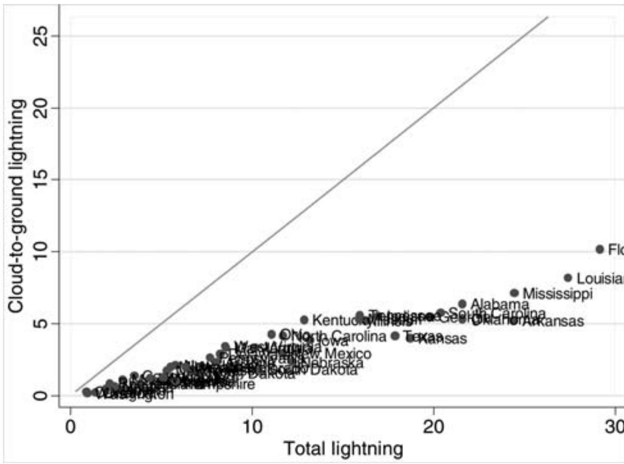
Partial Correlations

Table 4 reports the results from OLS estimation using cross-country data. As in Table 1 we experiment with various specifications: In Column 1 we use the first difference specification; in Column 2 we truncate the sample to limit the extent of mean reversion by excluding observations within one standard deviation from either endpoint of the 0-6 "*ICRG* interval"; in Column 3 we estimate the lagged dependent variables model; and finally in Column 4 we report the results from estimating our preferred model by way of an outlier robust estimator (Least Absolute Deviations, LAD).

Compared to the state-level analysis we find more variation in results when we move from the first difference specification to the lagged dependent variables specification. In Column 1 changes in Internet penetration is

25. NASA's flash densities of total lightning are calculated for the 1995-1999 period, while Vaisala's cloud-to-ground measures are calculated for the 1996-2005 period. In addition, Vaisala uses *mile*^{2} as the area unit; these where converted into *km*^{2} by dividing by the *mile*^{2} numbers by 1.609^2.

FIGURE 4. Total lightning versus cloud-to-ground lightning, U.S. sample



Note: The figure shows a scatter plot of total lightning flash rate density (flashes per km^2 per year) based on the Optical Transient Detector high resolution data (horizontal axis) versus cloud-to-ground flash rate density (flashes per km^2 per year) based on the U.S. National Lightning Detection Network (vertical axis). The full line is the 45-degree line. Number of observations is 48.

Source: Figure is based on data described in the text.

insignificantly correlated with changes in the *ICRG* index during the Internet era, whereas a positive correlation emerges when we employ the lagged dependent variables approach, thus suggesting a corruption suppressing influence of the Internet. It is worth observing that the initial *ICRG* level is highly significant and contributes considerably to the overall fit, as is evident from the increase in R^2 going from Column 1 to 3. The results from Column 2, where we truncate the sample, are somewhere in between. But even when we truncate the sample markedly to suppress mean reversion, the initial corruption level, *ICRG* in 1991, remains significant if included (not shown but results are available upon request); the partial correlation is weaker, as it should be in theory, but it is not rendered insignificant.

These results demonstrate that the evolution of corruption is subject to mean reversion in the world sample; perhaps even to a greater extent than in the U.S. sample, since even sample truncation is not sufficient to fully eliminate the correlation between initial corruption and subsequent changes in corruption. Now using both mean reversion ($\alpha_2 < 0$) and the fact that a high *ICRG* score in 1991 (low initial corruption) is associated with faster Internet diffusion (i.e., $Cov(C_{initial}, \Delta INTERNET) > 0$) in equation (2) immediately tells us that the first difference model is misspecified in the world sample; that is, since plim $\alpha_1^{FD} < \alpha_1$ the first difference model underestimates α_1. This is of course exactly what we observe in Table 4: the parameter estimate for $\Delta INTERNET$ rises monotonically when we go from Column 1 to 3.

TABLE 4. OLS and LAD regressions, world sample

	Change in ICRG 1991–2005			
Dependent variable:	(1)	(2)	(3)	(4)
ΔINTERNET	0.004	0.0138*	0.0361***	0.0352***
	(−0.003)	(−0.007)	(−0.006)	(−0.006)
$ICRG_{1991}$			−0.747***	−0.772***
			(−0.089)	(−0.097)
Constant	−0.923***	−1.044***	0.981***	1.072***
	(−0.140)	(−0.145)	(−0.235)	(−0.281)
Observations	102	67	102	102
R-squared	0.01	0.04	0.57	
Estimator	OLS	OLS	OLS	LAD

Notes: Robust standard errors in paranthesis. Asterisks ***, ** and * signify, respectively, p-value < 0.01, p-value < 0.05 and p-value < 0.1. In column 2 the sample is truncated; only countries with ICRG scores one standard deviation away from the ICRG index minimum (0) and maximum (6) value are included. Larger values for the ICRG index means *less* corruption. In column 4 the standard errors are bootstrapped with 1000 repetitions.

Source: Authors' analysis based on data described in the text.

There is a straightforward explanation of why the strong correlation between initial corruption and Internet diffusion, which we observe in the world sample, has no parallel in the U.S. sample. In the world sample, factors influencing the level of corruption, or the level of corruption *per se*, probably also influence the process of Internet diffusion; perhaps the over-all institutional quality of a nation influences not only corruption and its evolution, but also the speed of technology adoption. This seems plausible, and it can account for the positive correlation between the *ICRG* index and Internet diffusion in the world sample. It appears reasonable to expect that similar forces are *not* at play in the U.S. sample of relatively more homogenous states, for which reason a similar correlation between state-level corruption and state-level Internet use is absent. Therefore, in the world sample, but not in the U.S. sample, the end result of (erroneously) omitting the level of initial corruption is that the OLS estimate for Δ*INTERNET* is biased towards zero. The omitted variables bias is dampened when we truncate the sample (Column 2), as the correlation between the level of corruption and changes in corruption is reduced; the parameter estimate for Δ*INTERNET* therefore rises in absolute value. In sum, the lagged dependent variables specification should be considered the appropriate empirical model.

In an effort to check the robustness of the correlation we also report, in Table 4, the results from estimating our preferred model by way of an outlier robust estimator. As can be seen, comparing Column 3 and 4, the results are virtually the same whether we use OLS or LAD, implying that outliers are not carrying the correlation. This conclusion is further reinforced by invoking the Hadi (1992) outlier detection procedure; no observations are found to represent (multivariate) outliers.

As another check we have examined the robustness of the $\Delta INTERNET/\Delta ICRG$ correlation to the inclusion of the corruption determinants discussed in Treisman's (2007) survey. Treisman includes variables that constitute historical and controls; political controls; and finally a set of rents and competition controls. The correlation between $\Delta INTERNET$ and $\Delta ICRG$ remains when these variables are included in the lagged dependent variables model (one-by-one). Finally, the above results carry over if we employ the control of corruption indicator (Kaufmann *et al.*, 2007), instead of the *ICRG* index. These results are available upon request.

Instrument Falsification Test

Following the approach taken in the U.S. state-level analysis our 2SLS analysis on the World sample is preceded by an instrument falsification test.

Accordingly, Table 5 reports the results from estimating the reduced form on different time periods: pre-Internet (1984-1990) and post-Internet (1991-2005). As explained above, the pre-Internet period is limited to six years, due to lack of availability of the *ICRG* index beyond this. We report three sets of results on the two sub-periods: full sample OLS; OLS on a sample pruned for outliers detected by the Hadi (1992) method; full sample and outlier robust estimator (LAD).

Turning to the results, we find a statistically significant correlation between changes in *ICRG* and lightning during the Internet era (i.e., 1991 onwards). The partial correlation between lightning and changes in corruption reported in Column 1 can be inspected visually in Figure 5.

It is immediately obvious from the figure that Iceland (ISL) have considerable influence on the significance of lightning; i.e., Iceland may represent an outlier. Using the Hadi (1992) procedure, we confirm that Iceland is indeed an outlier; and the only one.

When we re-estimate the model, omitting Iceland (Column 2), we find essentially the same partial correlation as in the full sample. This remains true if we, rather than omitting Iceland from the sample, estimate the model by way of an outlier robust estimator (Column 3). The LAD point estimate is close to the OLS estimates. Hence, as in the U.S. sample, we find evidence that lightning prone regions have experienced lower reductions in corruption during the Internet era, consistent with the theory underlying the instrument.

In Columns 4-6 we turn to the pre-Internet era. At first sight the test seems to falsify the instrument as we detect a significant correlation between lightning and changes in *ICRG* during the pre-Internet era in Column 4. Taken at face value this means lightning is picking up forces that exert a time persistent influence on the evolution of corruption, which is inconsistent with the theory that lightning captures (solely) Internet diffusion. The OLS based partial correlation

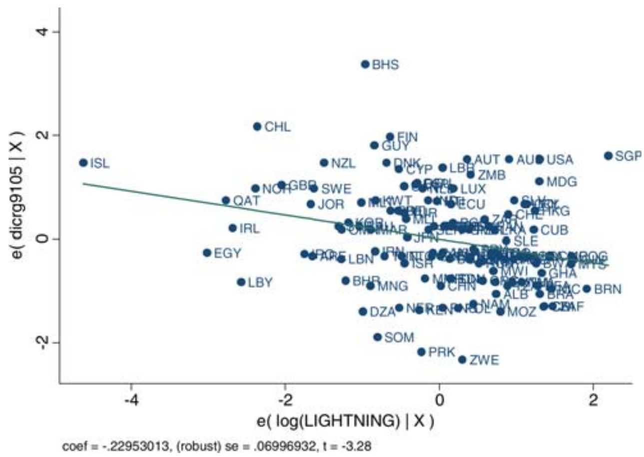
TABLE 5. Falsification test of instrument, world sample

	Change in ICRG 1991–2005			Change in ICRG 1984–90		
Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)
log(LIGHTNING)	−0.230***	−0.216***	−0.295***	−0.0931*	−0.0938	−0.0164
	(−0.070)	(−0.0785)	(−0.0777)	(−0.0527)	(−0.0619)	(−0.0588)
ICRG, 1991	−0.531***	−0.530***	−0.565***			
	(−0.0825)	(−0.0825)	(−0.0952)			
ICRG,1984				−0.232***	−0.232***	−0.0267
				(−0.0492)	(−0.0494)	(−0.101)
Constant	1.362***	1.331***	1.517***	1.040***	1.042***	0.132
	(−0.377)	(−0.385)	(−0.387)	(−0.244)	(−0.256)	(−0.469)
Observations	124	123	124	99	98	99
R-squared	0.36	0.36		0.16	0.16	
Estimator	OLS	OLS no outlier	LAD	OLS	OLS no outlier	LAD

Notes: Robust standard errors in paranthesis. Asterisks ***, ** and * signify, respectively, p-value < 0.01, p-value < 0.05 and p-value < 0.1. Column 2 and 4 omit an outlier (Iceland) detected using the Hadi (1992) procedure. Larger values for the ICRG index means *less* corruption. In columns 3 and 6 the standard errors are bootstrapped with 1000 repetitions.

Source: Authors' analysis based on data described in the text.

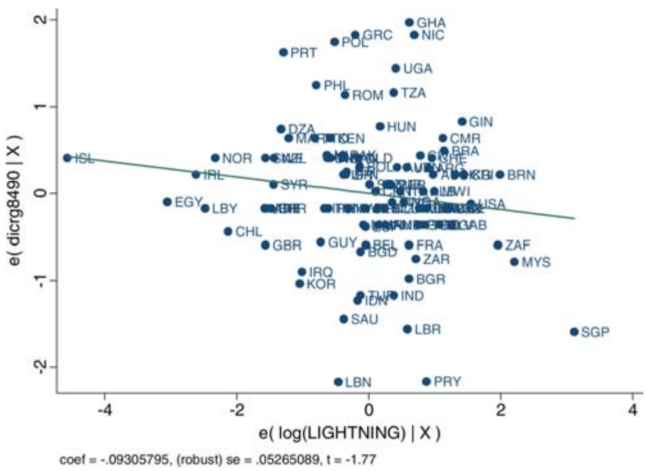
FIGURE 5. Reduced form in the Internet era, world sample



Note: The figure shows the association between lightning and changes in the ICRG index over the 1991–2005 period, with ICRG in 1991 partialled out.

Source: Figure is based on data described in the text.

FIGURE 6. Reduced form in the pre-Internet era, world sample



Note: The figure shows the association between lightning and changes in the ICRG index over the 1984–1990 period, with ICRG in 1984 partialled out.

Source: Figure is based on data described in the text.

between lightning and changes in the *ICRG* index from Column 4 is illustrated in Figure 6.

Visual inspection of the figure reveals that Iceland once again appears to represent an outlier, possibly even explaining the significance of lightning during the pre-Internet period. Employing the Hadi (1992) procedure one can confirm that Iceland indeed is a multivariate outlier; it is the sole observation (as in the post-Internet setting) that the Hadi (1992) test singles out.

When we subsequently re-estimate the model, omitting Iceland, the significance of lightning evaporates (Column 5). The fragility of the lightning /$\Delta ICRG$ correlation for the 1984-90 period is further illustrated by the LAD regression on the full sample (Column 6): In this setting, the size of the estimate is reduced considerably compared to the OLS result, and is far from being statistically significant at conventional levels.

Overall, these falsification tests are reassuring. When the outlier is omitted, or an estimation method that is robust to extreme observations is invoked, there is no statistically significant correlation between lightning and changes in corruption prior to the emergence of the World Wide Web. Moreover, moving beyond statistical significance, it is clear that lightning exhibits a much stronger correlation with changes in corruption during the post-Internet era, compared to the pre-Internet era. As can be seen from Table 5, the OLS estimate in Column 1 exceed that of Column 4 by no less than a factor of 2.5.

In sum, we find the evidence reported in Table 5 sufficiently supportive of the identification strategy (albeit perhaps less strong than in the U.S. case) so as to allow us to proceed to 2SLS estimation using lightning intensity as an instrument for Internet diffusion.

2SLS Estimates

Table 6 reports the 2SLS estimates of the impact from the Internet on the evolution of corruption around the World. Panel A reports the first stage, whereas Panel B displays second stage results. In keeping with the approach taken so far, we report results from three separate exercises: full sample, first difference specification (Column 1); truncated sample, first difference specification (Column 2); and our preferred lagged dependent variables specification on the full sample (Column 3).

Broadly speaking, the results mirror those attained using U.S. state-level data. First, the lightning instrument is statistically strong in the full sample of countries (Columns 1 and 3); the first stage features F-statistics well above ten. Second, the first difference specification produces insignificant estimates for $\Delta INTERNET$ in the second stage, whereas the lagged dependent variables specification leads to a statistically significant point estimate for $\Delta INTERNET$.

Our interpretation of this variation in results across specifications is the same as in the case of the U.S. state-level analysis. Changes in corruption display mean reversion, for which reason the initial level of corruption needs to be controlled for in the regression. When it is erroneously omitted, it jeopardizes identification since lightning and the level of corruption are correlated. A glance on the lightning map (Figure 3) above is enough to realize that high lightning intensity is more prevalent in the poorer regions of the World, near the equator, where countries also tend to be characterized by relatively poor

TABLE 6. 2SLS regressions, world sample

	Panel A: First stage		
	Dependent variable: ΔINTERNET		
	(1)	(2)	(3)
log(LIGHTNING)	−10.619***	−4.4499**	−6.9537***
	(1.2479)	(1.5087)	(1.2863)
$ICRG_{1991}$			6.5275***
			(1.1915)
Constant	38.985***	20.899***	10.252*
	(2.7145)	(3.4299)	(5.7635)
F-test (H0: log(LIGHTNING) = 0)	71.62	7.76	25.08
	Panel B: Second stage		
	Dependent variable: Change in ICRG 1991–2005		
ΔINTERNET	0.00011	0.0216	0.0435***
	(−0.006)	(−0.0285)	(−0.010)
$ICRG_{1991}$			−0.820***
			(−0.114)
Constant	−0.841***	−1.137***	1.079***
	(−0.176)	(−0.329)	(−0.244)
Observations	102	67	102

Notes: Robust standard errors in paranthesis. Asterisks ***, ** and * signify, respectively, p-value < 0.01, p-value < 0.05 and p-value < 0.1. In column 2 the sample is truncated; only countries with ICRG scores one standard deviation away from the ICRG index minimum (0) and maximum (6) value are included. Larger values for the ICRG index means *less* corruption.

Source: Authors' analysis based on data described in the text.

governance. Accordingly, identification is only plausible in the case of the lagged dependent variables specification.

The results from Column 3 of Table 6 corroborate our OLS findings (Table 4): From the early 1990s onward the Internet has worked to suppress corruption. Hence, it would appear that our main results, which pertain to the U.S., carry over to a cross-country setting. In closing, it should be noted that results are similar when we use the Kaufmann *et al.* (2007) corruption indicator instead of the *ICRG* data; these results are available upon request.

V. CONCLUDING REMARKS

In the present paper we have examined the hypothesis that technological change may cause improvements in governance. Specifically, we have studied the influence from the Internet on the evolution of corruption across U.S. states and across the World. The results are, statistically speaking, perhaps somewhat stronger in the U.S. setting but in both samples we find evidence that the Internet has worked to suppress corruption since its emergence.

Our results should be interpreted with some care: they are reduced form estimates of the total impact from the Internet on corruption. The analysis therefore does not speak to the mechanisms linking the Internet to corruption. As observed in the Introduction, the Internet is a General Purpose Technology, for which reason it probably influences the economy in a great many ways; growth, human capital accumulation, the dissemination of information, and E-government are the most obvious intervening factors that are both associated with the Internet and impinge on the level of corruption in a state or country. The task of disentangling these separate pathways of influence from the Internet technology to corruption is left open for future research.

The identification strategy developed in the present paper may be useful in other contexts. With a plausibly exogenous instrument for the Internet, researchers may potentially make new progress in the study of the impact from the Internet on other outcomes, such as the return to skills or productivity growth more broadly.

Appendix

Table A.1. Summary statistics

	Panel A: Cross-country sample					
	Obs.	Mean	Median	Std. Dev.	Min	Max
ICRG 2005	102	2.60	2.14	1.30	0	6
ICRG 1991	102	3.44	3	1.46	0	6
INTERNET 2005	113	22.56	15.24	21.66	0.21	86.94
Average lightning density	113	9.00	6.59	8.62	0.02	44.38
			Panel B: U.S. sample			
CC 2006	48	20.15	9.50	21.48	0	83
CC 1991	48	13.44	6.00	21.14	0	108
INTERNET 2003	48	54.39	55.00	5.88	39.50	65.50
POP 1991	48	52.22	36.46	56.16	4.59	305.00
Average lightning density	48	10.54	8.58	7.31	0.89	27.34

Source: Authors' calculations based on data described in the text.

References

Acemoglu, Daron, and James Robinson, 2001. Political Losers as a Barrier to Economic Development. *American Economic Review*, 90, 126–130.

Ades, Alberto, and Rafael di Tella, 1999. Rents, Competition and Corruption. *American Economic Review*, 89, 982–993.

Andersen, Thomas B., 2009. E-government as an anti-corruption strategy. *Information Economics and Policy*, 21, 201–210.

Andersen, Thomas B., Jeanet Bentzen, Carl-Johan Dalgaard, and Pablo Selaya, 2010. Lightning, IT Diffusion, and Economic Growth across U.S. States. *Review of Economics and Statistics*, forthcoming.

Angrist, Joshua, and Stefan Pischke, 2009. Mostly Harmless Econometrics, Princeton University Press.

Besley, Tim, and Robin Burgess, 2002. The Political Economy of Government Responsiveness: Theory and Evidence from India. *Quarterly Journal of Economics*, CXVII, 1415–1451.

Bhatnagar, Subhash, 2003. E-government and access to information. *Global Corruption Report 2003*, Transparency International.

BSI, 2004. Bundesamt für Sicherheit in der Informationstechnik, IT-Grundschutz Manual. Available online at <http://www.bsi.de/english/gshb/index.htm>.

Brunetti, Aymo, and Beatrice Weder, 2003. A free press is bad news for corruption. *Journal of Public Economics*, 87, 1801–1824.

Caselli, Francesco, and Wilbur J. Coleman III, 2001. Cross-Country Technology Diffusion: The Case of Computers. *American Economic Review*, 91, 328–335.

Chinn, Menzie D., and Robert W. Fairlie, 2007. The determinants of the global digital divide: a cross-country analysis of computer and Internet penetration. *Oxford Economic Papers*, 59, 16–44.

Chisholm, William A., 2000. Lightning Protection. Chapter 4.10 in: The Electrical Power Engineering Handbook. *Edited by* L Grigsby., IEEE Press.

Chisholm, William A., and Kenneth Cummins, 2006. On the Use of LIS/OTD Flash Density in Electric Utility Reliability Analysis. Paper presented at *The Lightning Imaging Sensor International Workshop*. Available online at <http://lis11.nsstc.uah.edu/Papers/Proceedings.html>.

Deaton, Angus, 2010. Instruments, Randomization, and Learning about Development, *Journal of Economic Literature*, 48, 424–455.

Eisensee, Thomas, and David Stömberg, 2007. New Droughts, New Floods and U.S. Disaster Relief. *Quarterly Journal of Economics*, 122, 693–728.

Electricity Power Research Institute, 2003. Electricity Technology Roadmap: 2003 Summary and Syntheses. Available online at <http://www.iea.org/Textbase/work/2004/distribution/presentations/Gellings_bkgd_%20paper.pdf>.

Ferraz, Claudio, and Frederico Finan, 2008. Exposing Corrupt Politicians: The Effect of Brazil's Publicly Released Audits on Electoral Outcomes. *Quarterly Journal of Economics*, 123, 703–745.

Glaeser, Edward L., and Raven E. Saks, 2006. Corruption in America. *Journal of Public Economics*, 90, 1053–1072.

Foster, Andrew D., and Mark R. Rosenzweig, 1996. Technical Change and Human-Capital Returns and Investments: Evidence from the Green Revolution. *American Economic Review*, 86, p. 931–53.

Fredriksson, Per G., Herman R. J. Vollebergh, and Elbert Dijkgraaf, 2004. Corruption and energy efficiency in OECD countries: theory and evidence. *Journal of Environmental Economics and Management*, 47, 207–231.

Gundlach, Erich, and Martin Paldam, 2009. The transition of corruption: From poverty to honesty. *Economics Letters*, 103, p.146–148.

Hadi, Ali S, 1992. Identifying multiple outliers in multivariate data. *Journal of the Royal Statistical Society* 54, 761–771.

Hall, Robert E., and Dale Jorgenson, 1967. Tax Policy and Investment Behavior. *American Economic Review*, 57, 391–414.

IEEE, 2005. *How to Protect Your House and Its Contents From Lightning*. IEEE Guide for Surge Protection of Equipment Connected to AC Power and Communication Circuits. Standards Information Network, IEEE Press.

Jovanovic, Boyan, and Peter L. Rousseau, 2005. General Purpose Technologies. In P. Aghion, and S. Durlauf (eds.) *Handbook of Economics Growth*, Elsevier, North-Holland.

Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi, 2007. Governance Matters IV: Governance Indicators for 1996–2006. Available online at <http://www.worldbank.org/wbi/governance/wp-governance.html>.

Kressel, Henry, 2007. *Competing for the Future: How Digital Innovations are Changing the World*, Cambridge University Press.

Licht, Amir N., Chanan Goldschmidt, and Shalom H. Schwartz, 2007. Culture Rules: The foundations of the rule of law and other norms of governance. *Journal of Comparative Economics*, 35, 659–88.

Mauro, Paolo. 1995. Corruption and Growth. *Quarterly Journal of Economics* 110, 681–712.

Maxwell, Amanda E., and Richard F. Winters, 2005. Political Corruption in America. Working Paper (Dartmouth College).

Mokyr, Joel, 1990. *The Lever of Riches: Technological Creativity and Economic Progress*, Oxford University Press.

NASA, 2007. Lightning Detection From Space: A Lightning Primer. Available online at <http://thunder.msfc.nasa.gov/primer>.

National Energy Technology Laboratory, 2003. The Value of Electricity When It's Not Available. Available online at <http://www.netl.doe.gov/moderngrid/docs/The_Value_of_Electricity_When_It's_Not_Available.pdf>.

Parente, Stephen L., and Edward C. Prescott, 1999. Monopoly Rights: A Barrier to Riches. *American Economic Review*, 1216–1233.

Persson, Torsten, Guido Tabellini, and Francesco Trebbi, 2003. Electoral rules and corruption. *Journal of the European Economic Association*, 1, 958–989.

Reinikka, Ritva, and Jakob Svensson, 2004. The power of information: Evidence from a newspaper campaign to reduce capture. Policy Research Working Paper No. 3239, World Bank.

Schultz, Theodore W., 1975. The Value of the Ability to Deal with Disequilibria. *Journal of Economic Literature*, 13, p. 827–46.

Shim, Jae K., Anique Qureshi, and Joel G. Siegel, 2000. *The International Handbook of Computer Security*. Glenlake Publishing Company Ltd.

Staiger, Douglas, and James H. Stock, 1997. Instrumental variables with Weak Instruments. *Econometrica*, 65, 557–586.

Strömberg, David, 2001. Mass Media and Public Policy, *European Economic Review*, 45, 652–663.

———, 2004. Radio's Impact on Public Spending, *Quarterly Journal of Economics*, 119, 189–221.

Svensson, Jakob., 2005. Eight Questions about Corruption. *Journal of Economic Perspectives*, 19, 19–42.

Treisman, Daniel, 2000. The causes of corruption: a cross-national study. *Journal of Public Economics*, 76, 399–457.

———, 2007. What Have We Learned About the Causes of Corruption From Ten Years of Cross-National Empirical Research? *Annual Review of Political Science*, 10, 211–244.

Yeager, Kurt, and Karl Stahlkopf, 2000. *Power for a Digital Society*. Prepared for the Department of Energy CF-170/1-1-DOE. Available online at <http://www.rand.org/scitech/stpi/Evision/Supplement/yeager.pdf>.

WDI, 2007. *World Development Indicators 2007*, World Bank.

West, Darrell M., 2005. *Digital Government*, Princeton University Press.

———, 2008. *Improving Technology Utilization in Electronic Government around the World*. Brookings Governance Studies.

Wooldridge, Jeffrey., 2003. *Introductory Econometrics: A Modern Approach*, 3E, South–Western.

World Development Report 2004. *Making Services Work for the Poor*, The World Bank.