

DISCUSSION PAPER SERIES

DP15417

PERSISTENCE – MYTH AND MYSTERY

Hans-Joachim Voth

ECONOMIC HISTORY



PERSISTENCE – MYTH AND MYSTERY

Hans-Joachim Voth

Discussion Paper DP15417
Published 30 October 2020
Submitted 29 October 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Economic History

Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Hans-Joachim Voth

PERSISTENCE – MYTH AND MYSTERY

Abstract

Is long-term economic persistence a myth? In this survey, I examine what has been learned from the wave of articles examining the persistence of economic phenomena. The article clarifies some of the concepts, distinguishes different strands of the literature, and questions the validity of recent critiques of persistence papers. While challenges in the form of misattribution and p-hacking abound, this survey concludes that persistence is real, and that future research should focus on mechanisms and the conditions under which persistence breaks down.

JEL Classification: N/A

Keywords: Long-run persistence, spatial correlation, Economic Geography, Culture

Hans-Joachim Voth - voth@econ.uzh.ch
UBS Center for Economics in Society and CEPR

Persistence – Myth and Mystery^{*}

Hans-Joachim Voth

First version: 2020-04-04

This version: 2020-09-24

How deep are the historical roots of the present? All around us, the world we live echoes with the footsteps of the past – from the houses we inhabit to the structure of the cities in which we dwell, the languages we speak and the religions many of us practice. Overwhelmingly, both the cultural universe and the built environment we inhabit was created before our time. The novelist William Faulkner famously observed: "The past is never dead. It's not even past." The notion that history can evolve in a particular direction, and not merely repeat itself, is relatively recent (Koselleck 2003). While cultural anthropologists and sociologists have long emphasized the sluggishness with which many important aspects of our lives change from generation to generation, economics has only recently become interested in the general stickiness of life – *homo oeconomicus*, as he emerges from most textbooks, has no past.

Once economists began to look, a growing number of rigorous studies uncovered a startlingly wide range of phenomena with deep roots. From gender attitudes, trust, investments in human capital, and within-country variation in riches to racial prejudice, the shadows of the past seem to loom large in every corner of intellectual inquiry pursued by economists (Dell 2010; Alesina, Giuliano, and Nunn 2013; Nunn and Wantchekon 2011; Guiso, Sapienza, and Zingales 2016; Valencia Caicedo 2019). At the same time, many historians (and economic historians) see persistence studies as "ahistorical". Persistence papers now form a distinct group of scholarly papers (Cioni, Federico, and Vasta 2020). Like many labels that gain wider currency for a research program, it actually refers to a wide range of different approaches and research ideas.

The *Oxford English Dictionary* defines persistence as the "continued or prolonged existence of something." Persistence papers typically measure some variable in the past, and then relate it to a later-day outcome, often in the present or recent past. The papers with the longest 'range' argue for historical dependence over millennia. Such long-term persistence seems surprising not least because of our everyday experience of massive and often rapid changes -- both in economic outcomes and cultural traits. Countries like South Korea or Italy were poor two generations ago, but rank amongst the richest today; attitudes towards

^{*} For detailed comments, I thank Sascha Becker, Joel Mokyr, and Felipe Valencia. Excellent research assistance by Marcel Caessman is gratefully acknowledged.

homosexuality, corporal punishment, smoking, women's work, and pre-marital sex have changed beyond all recognition in the last 40 years or so, a mere blink of an eye in the history of mankind.

How can we understand the juxtaposition of rapid change and remarkable persistence over the long run? If persistence is indeed a fact, and common across a wide range of outcomes, the implications for economic policy – and even for philosophical questions such as the concept of free will – can be profoundly disturbing. If, for example, a country's riches today are largely determined by its use of technology 1000 BC, its chances for development may be severely curtailed – at least as long as time travel remains a technological challenge.

This essay's purpose is to first clarify the meaning of persistence before providing an overview of what insights the buoyant literature has produced in the last decade. I will then discuss critically the empirical and conceptual challenges, and propose methods and standards for assessing the plausibility of persistence studies.

The range of historical variables used in persistence studies is wide -- from the self-governance of Italian city states to plough usage, medieval pogroms, and the adoption of Protestantism to gold discoveries, a very large number of different explanatory variables has been examined in persistence papers. The types of dependent variables used is also large, ranging from GDP per capita, city growth, human capital investment, and institutional quality to anti-Semitic violence, trust and gender attitudes. Since this chapter cannot survey every combination of variables that have been tried in the domain of persistence studies, it will aim to illustrate key themes and challenges.

Conceptual Challenges

Persistence studies arguably come in at least two and a half different “flavors”. Since the mechanisms underlying them are distinct, it would be misguided to conflate them by using the same rubric. The first “flavor” takes modern-day output or a related variable, and regresses it on a historical variable. Conceptual problems arise where the origin of the historical treatment is unclear – many shocks and outcomes in the past can create self-perpetuating equilibria. Numerous papers in this category can be classified as ‘apples-and-oranges’ papers – meaning that they regress outcomes (today) on historical variables that are quite different in nature. Many papers of this kind have economic outcomes on the left-hand side, and some non-economic variable on the right – not only arguing, in effect, that the past influences the economic presence, but that non-

economic variables are key drivers. While some of hypothesized linkages can seem ad-hoc, the more promising papers in this literature use theoretical frameworks from pre-existing studies to discipline the empirical analysis ('apples-and-oranges-with-theory'). This can help to alleviate concerns over extensive specification searches, with large numbers of combinations of variables being tried out until a significant result emerges.

A second flavor of persistence study focuses on cultural characteristics such as attitudes and beliefs, using non-economic factors as explanatory and dependent variable. I will refer to these as 'apples-on-apples' papers, using left-and-right-hand side variables that are conceptually close. This can seem more compelling. This second category of papers, while arguably easier to conceptualize, may suffers from possible confounders "along the way": many of the driving variables have direct economic effects; by influencing economic conditions in the past, their link with outcomes today may be intermediated by other, unrelated variables, creating a spaghetti-tangle of causal relationships.

The "History Curse" – An Illustration

The bible, in the book of Matthew (13:12), already had a view on the importance of persistence: "For whosoever hath, to him shall be given, and he shall have more abundance; but whosoever hath not, from him shall be taken away even that he hath." The biblical idea applies to individuals. It implies not just an enormous predictive power of starting positions, but even increasing polarization over time.

A first glance at long-run patterns in output per capita suggests that the Matthew principle broadly holds in the cross-section of countries. Figure 1 shows a scatterplot of per capita incomes in 1500, 1700, 1820, and 1998. While the correlation is not perfect, correlation coefficients are high. Past riches explain up to 64% of modern-day riches. While the effect of the past fades, it does so surprisingly slowly – per capita income in 1500 still explains 20% of the variation in incomes in 1998.

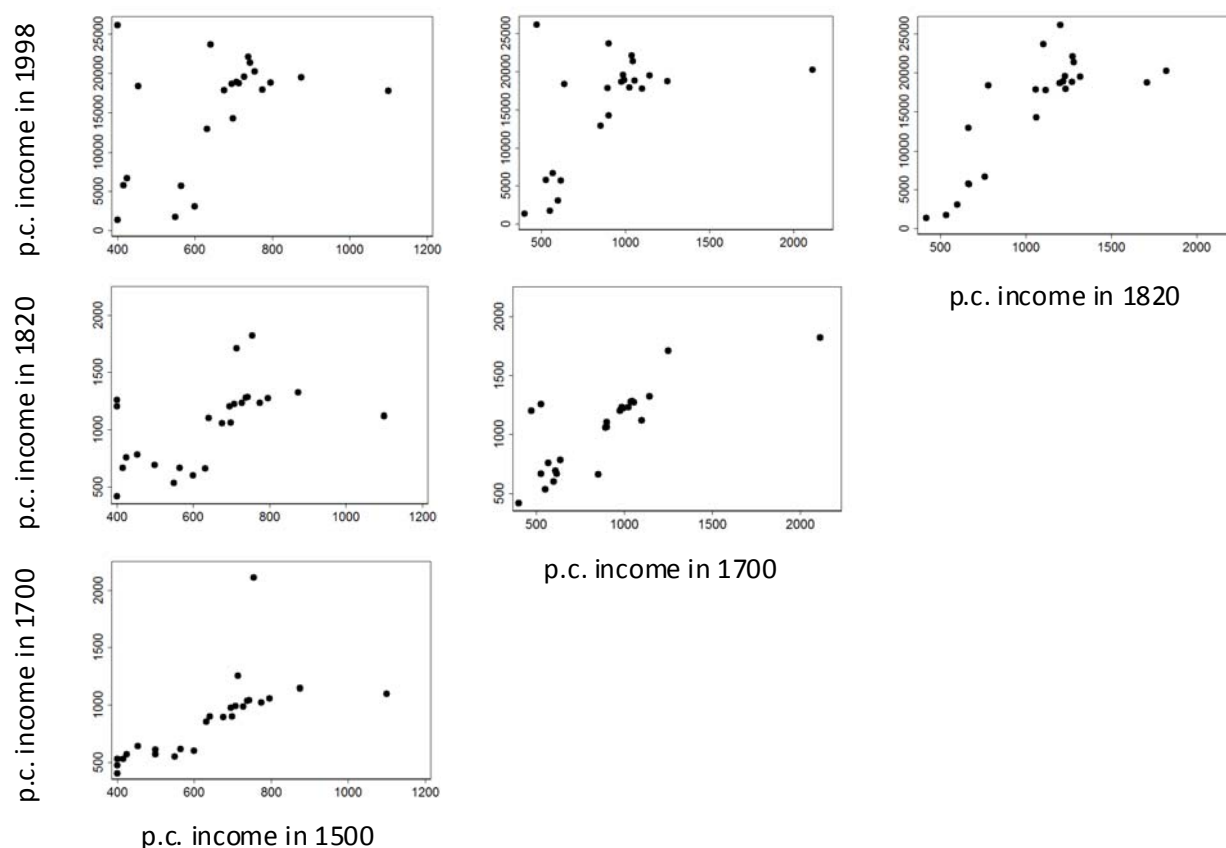


Figure 1: Scatterplot of Per Capita Incomes in 1500, 1700, 1820, and 1998
Source: Voigtländer and Voth (2013a)

The strong grip of the past is startling because it is hard to rationalize in workhorse models of economic growth. In the Solow-Swan model, for example, convergence between rich and poor countries should be relatively rapid. With low capital, the return to savings should be high. Consequently, in poor countries, high savings will drive a massive expansion of the capital stock, which in turn should create catch-up vis-a-vis leading countries in short order. Most of the convergence should occur within a few decades at most (Pritchett 1997).

First of all, the Matthew effect needs to be put into perspective – dramatic reversal can and have occurred over the course of history. China used to dominate most of the world in terms of the sophistication of its technology, the quality of its bureaucracy, and the effectiveness of its highly centralized state. Why it nonetheless failed to develop economically is the subject of the so-called “Needham Puzzle”, which is the subject of substantial recent research interest (De la Croix, Doepke, and Mokyr 2018; Voigtländer and Voth 2013b; Landes 2006; Lin 1995). Similarly, Islamic countries used to be ahead in terms of technology and science relative to European ones for centuries (Mokyr 1990), before falling behind. Sokoloff and Engerman (1997) showed how in the

Americas, areas that were ahead in the early days of colonization eventually fell behind. The sugar islands of the Caribbean are a case in point: once among the most prosperous places on earth, they have seen sharp declines in their relative economic performance (Sokoloff and Engerman 2000). Acemoglu, Johnson and Robinson (2002) argue that among the areas colonized by Europeans, those that were richest in 1500 are now particularly poor as the colonizers set up institutions designed to exploit natural resources.¹ Finally, a considerable number of once-poor East Asian countries have successfully escaped the “history curse” (Young 1995; Kim and Lau 1994; Hsieh 2002).

Despite these important exceptions, the Matthew effect holds on average, as Figure 1 shows. Generically, there can only be two broad classes of explanation for it—high incomes in the past may themselves facilitate growth, or the same factors that helped countries grow rich in the past are still important determinants of per capita income.

Persistence studies overwhelmingly argue for the latter channel, but not all long-range results are equally plausible. In a widely-cited paper, Comin, Easterly and Gong (2010) ask “Was the wealth of nations determined 1000 BC?”. The authors create indicators of technological development for different regions of the world, for 1000 BC, the year 0, and 1500, based on a detailed encyclopaedia of archaeology (for pre-historical times). These they use to explain income levels today. The authors find a high degree of predictability -- technologically advanced countries in 1000 BC, 0, or 1500 are all richer today. The finding is puzzling at several levels. First, because the predictive power of technology in 1000 BC for income and technology in 0 or 1500 is limited -- why would technology before the age of Homer predict incomes today if it is a poor predictor of incomes at the time of Christ or Luther? Second, it is challenging to think through underlying mechanisms because the authors use geographical units, not populations, as units of observation. Since not many descendents of the people living in Germany at the time of Tacitus are still living there today, it is difficult to see how technology adoption by the ancient tribe called the Cimbri, could influence technology use in modern-day Denmark, where they originated. The same is true of large parts of Greece, France, Eastern Europe, to name but a few areas.

Some of coding that drives results can be questioned. The authors, for example, award a point to countries building ships that cross the Atlantic, another point for crossing the Indian Ocean, another for the Pacific. This is, of course, not

¹ Relatedly, Maloney and Valencia (2012) show that at the sub-national level across Latin America, the persistence of fortunes is the rule – but that extractive institutions (and in particular, slavery) reduce local persistence.

"technology" in the strict sense of the word -- building large, ocean-going ships is a technological feat. To cross various oceans with them is an outcome, driven by a combination of political, economic and cultural factors. Relatedly, the 1500 coding awards a large number of points for firearms (small, large, mounted on large ships, etc.). These choices stack the odds in favor of finding technological advantages for European powers. The main puzzle pointed out by historians of technology over the last millennium is arguably not why China was consistently behind Europe in technology – but how it came to fall behind, having been ahead for so long (Needham 1976).

One of the earliest, path-breaking studies of history's long reach illustrates how challenging it is to untangle cause and effect. In their paper "Long-term persistence", Guiso, Sapienza and Zingales (2016) examine income differences across municipalities in modern-day Italy. They argue that the medieval experience of self-governance ("communes") is an important determinant of productivity today -- with cities that became independent creating a culture of civic responsibility and self-reliance that pays dividends to the present day. They document a persistent influence of commune status on an impressive range of outcomes -- from cheating on national math exams to financial intermediation, medieval self-governance is closely associated with desirable outcomes today.

There are indeed strong reasons to assume that self-governance would have affected efficiency and output in medieval times already -- as the literature on the importance of institutions would predict (Figure 2). Correlated outcomes creating a new, self-reinforcing equilibrium, however, will often create a good deal of stability in state variables over time.

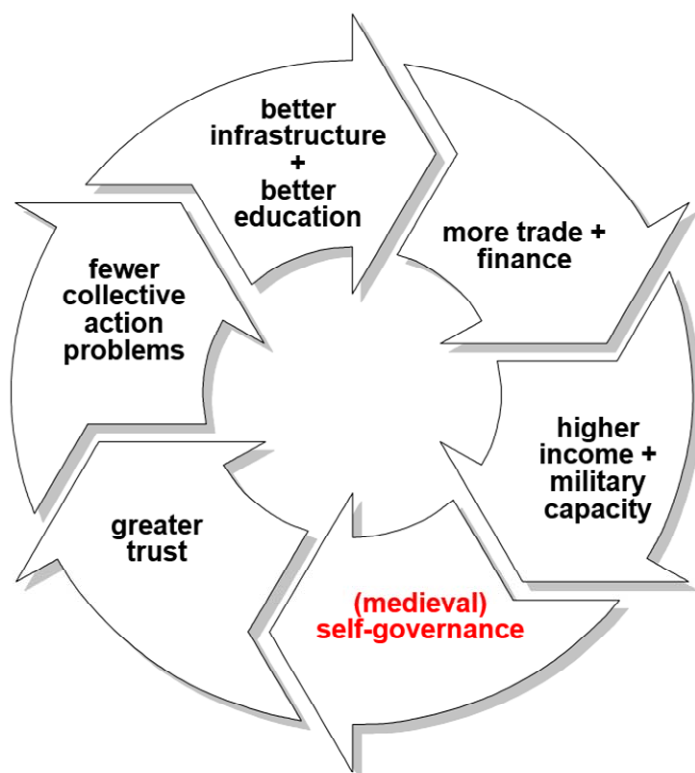


Figure 2: A medieval wheel of fortune – self-governance and economic success

Several studies argue in favor of seemingly deep roots of comparative development, but base their analysis on theoretical predictions from an earlier literature. One important example is Putterman and Weil (2010). The authors begin with the well-known connection between income per capita on the one hand, and state antiquity time as well as time since the agricultural revolution on the other -- where state-like structures have existed for centuries or millennia, incomes are higher. The same is true for the length of time since the invention of sedentary agriculture in a particular country. Putterman and Weil make the important point that state antiquity in a location is not necessarily relevant -- on the territory of the modern-day United States and Canada, there were no state-like structures until quite recently. However, the people living in the US today overwhelmingly come from countries where the state is quite old. Some 97 percent of the population of the New World is descended from people whose ancestors lived elsewhere. In other places, almost no adjustment is necessary -- the overwhelming majority (97.5%) of Europeans are descendents of people who already lived in Europe in 1500.

What happens if we adjust state antiquity and time since the start of agriculture by the effects of migration? Remarkably, their predictive power increases by leaps and bounds -- not state antiquity in a location matters, but the exposure to the existence of a state, transmitted from generation to generation. This points the finger at a plausible mechanism – the internalization of norms that

comes from abundant state power, leading to the growing “domestication” of humans along the lines of Pinker (2012) and Elias (1994).² Once “rendering unto Caesar” has become an ingrained habit, and the state can enforce the monopoly of violence, it is easier for growth to occur and riches to last. In this way, the results by Putterman and Weil do not only add credibility to an earlier persistence result, but provide additional insight into factors driving “stickiness”.³

This finding is in line with within-country evidence by Dell, Lane, and Querubin (2018). They show that in areas of Vietnam that were – for plausibly exogenous reasons – not brought under the control of the centralized Northern Dai Viet state, both civic society and local governance are better at creating public goods and redistribution. Just as in Africa (Gennaioli and Rainer 2007), pre-colonial centralization of state authority is systematically associated with better development outcomes, arguably because better governance crowds in pro-social attitudes and behaviors. The idea of „internalizing“ good social norms as a result of prolonged exposure to state power is examined more closely in a recent paper by Lowes et al. (2017). However, they find that among descendents of people who lived under the Kuba kingdom in the Central African Republic, honesty was no greater than amongst study participants from other areas. Whether this reflects the specificities of state-building under the Kuba kings, or the idiosyncracies of the setting (examining migrants to Kananga instead of locals under Kuba and alternative rule) is an open question.

A similarly spectacular illustration of long-term persistence is Ashraf and Galor’s (2013) argument that genetic diversity – largely determined at the time of mankind’s migration out of Africa – is a prime determinant of riches today. The authors first show that migration distances from Addis Abeba are a good predictor of genetic diversity today. Since there is ample archaeological evidence to suggest that humans first appeared in Africa, and then spread in small groups to the rest of the world, this is a plausible argument – reinforced by the fact distance to other locations has no predictive power. Genetic variation today is greatest today in Africa, and it is lowest amongst native tribes of the Amazon.

² The separate effect for agriculture is a little harder to explain. However, since the size of states historically increased rapidly once sedentary agriculture was invented, it may be hard to tell these two factors apart.

³ An even broader interpretation of Western European divergence and technological creativity argues that unusually high incomes led to greater individualism, pro-social attitudes, and higher rates of innovation (Baumard 2019).

Ashraf and Galor don't only show that genetic variation – and the part of the variation driven by the “out of Africa” component – predict income. They demonstrate that the relationship is non-linear, with intermediate levels of genetic diversity being best for GDP p.c. in 2000. Both too much and too little diversity can be bad for productivity. The argument is that ‘excessive’ ethnic diversity is bad because it undermines social cohesion and makes it more difficult for a capable state to emerge. Low levels of genetic diversity, on the other hand, are a barrier to creativity and technological progress. The ideal level of diversity follows a goldilocks logic – not too much and not too little.⁴

Two features distinguish the Ashraf and Galor study in methodological terms. First, they show that their mechanism is already at work in 1500. Using population density as a measure of production in the Malthusian epoch, they find the same effects of diversity and diversity squared. The fact that the same pattern is visible for incomes in 1500 and 2000 increases the likelihood that Ashraf and Galor have identified a deep parameter of the development process. Second, they show evidence consistent with these mechanisms: the number of scientific articles is higher in places with greater genetic diversity, but the level of interpersonal trust is lower. The value of the Ashraf and Galor study is in showing how one ‘deep’ parameter – genetic diversity – can lead to similar relative outcomes over the very long run, and in documenting plausible mechanisms. In other words, Ashraf and Galor's implied explanation of the Matthew paradox is that the key driving variable for incomes has remained the same since *homo sapiens* left Africa some 150,000 years ago.

In contrast to papers that present long-run persistence in economic outcomes as a black box, the Putterman-Weil and Ashraf-Galor papers have the benefit of documenting a detailed causal chain. Its plausibility derives from a well-documented literature demonstrating the importance of state capacity for economic development (Besley and Persson 2010; 2009; Diamond 1997; Spolaore and Wacziarg 2009) and the deleterious effects of ethnic heterogeneity on the size and strength of states (Alberto Alesina, Glaeser, and Sacerdote 2001).

Like Father, like Son

A second category of persistence studies is arguably more humble in ambition. Papers with this flavor (‘apples-on-apples’) employ variables on the left- and right-hand side that measure the same (or very similar) traits. There is little doubt that persistence in the strict sense of the word exists -- with the same attitude or belief being passed down the generations. When the variables in

⁴ The paper has been criticized by anthropologists, who contest the validity of the underlying data and assumptions about human behavior (Guedes et al. 2013), and declare themselves troubled by the potential social and political effects.

question are cultural features, such as attitudes and beliefs, the idea of persistence rests on relatively strong micro-foundations. This is what could be called "pure cultural persistence", especially in cases where there are no intervening or re-enforcing economic factors.

Important work by Dohmen et al. (2012) demonstrates that fundamental psychological characteristics show high correlations between parents and children. They use data from the German socio-economic panel (SOEP), and examine risk taking and trust of children as a function of parental responses. The authors find a highly significant relationship -- even in narrowly defined domains, such as risk taking when driving, as opposed to risk taking in financial matters. The average child in Dohmen et al.'s study is 25 years old, meaning that results are not driven by highly impressionable young children (who may change their views in later life). Both mothers' and fathers' attitudes matter for trust, while others are primarily transmitted only via one of the parents. Overall explanatory power is not overwhelmingly high, ranging from 7 to 11 % of overall variation -- but parents' attitudes on average matter a great deal. Both cultural and genetic factors could be driving such dependence. In either case, long-term transmission may be higher than implied by simple correlations across two generations -- for example, if latent genetic factors are activated by circumstances, or if cultural beliefs become relevant in particular contexts. A similar pattern is visible in intergenerational mobility of income and socio-economic status, where grandparents/grandchildren are more alike than is implied by parent-child transmission.

Similarly, Fernandez and Fogli (2009) pioneered the use of data on immigrants' behavior, showing striking correlations of migrants' fertility outcomes with fertility levels in the country of parental origin. Because the offspring of immigrants grew up in a very different environment than their parents, it is much more likely that we observe true cultural transmission and not a feedback effect from the environment to attitudes ("epidemiological approach"). How long attitudes can persist through vertical transmission is, however, not clear. Alesina and Fuchs-Schündeln (2007) in "Good-bye Lenin or Not?" argue that Germans who grew up under the Communist regime are much more likely to favor redistribution and social support (and to downplay individual responsibility) than their countrymen who grew up in the West. The more time Germans spent under Communist influence, the more pro-government people became. Because the environment changed radically after 1989, with the fall of the wall, the attitudes of young people growing up in the eastern part of Germany today are converging towards the West German norm -- but such convergence is relatively slow, and will take one to two generations to be complete.

The assumption that assignment to treatment in the past is often as “good as random” is unfortunately often wrong. Implicitly, the Alesina and Fuchs-Schündeln paper assumes that pre-1945, East and West Germany were largely identical in attitudes – and hence, any differences that appear today must be the result of the Communist “treatment”. This assumption is *prima facie* sensible, in the absence of evidence to the contrary – after all, they had formed part of the same country since 1871. However, there is strong evidence that the area of Germany that became the GDR was already strongly different in a number of important dimensions prior to 1945 – church attendance was already lower, female labor force participation was higher, and electoral support for the Communists was higher in the East in the interwar period (Sascha O. Becker, Mergele, and Woessmann 2020). In other words, the ‘treatment’ of Communist rule and Soviet occupation may only have preserved or amplified pre-existing differences. This does not imply that there was no persistence – it implies that it was so strong in this particular case that even the cultural and social effects of a brutal, long-lasting dictatorship did little more than to leave intact differences that existed long ago.

Both the Dohmen et al. and the Fernandez and Fogli studies provide conceptual microfoundations for *cultural* persistence – with attitudes and beliefs staying the same. Along the same lines, the *amplification* of existing prejudice through a synergy with policy can strengthen beliefs and biases, as Voigtländer and Voth (Voigtländer and Voth 2015) show in the case of exposure to Nazi schooling and anti-Semitism. Historical studies that have studied such persistence over the long run have often focused on human capital. Here, however, the possible ‘contamination’ with economic feedback effects is not trivial.

In one recent study, Kung et al. (2020) examines the long-run effects of keju, China's imperial exam system. They show that areas of China where more citizens passed the highest level exams (jinshi), educational attainment today is markedly higher. The imperial exam system itself first appeared in the Han Dynasty (206BC-220AD). Its core elements lasted for more than two millennia. The authors argue that social capital, educational infrastructure, and interactions with political elites led to continued investments in one trait – the value given to education. The fact that local continuity is visible in China despite the massive institutional, cultural and economic discontinuity of the Chinese civil war and the Communist takeover after 1948 is remarkable. A similar result emerges from recent work on earnings differences between descendents of elite and non-elite families, based on social stratification prior to the Communist take-over. Despite the best efforts of the Chinese Communist party to eliminate inter-generational transmission of status and wealth,

descendants of China's pre-revolutionary elite have higher education and higher earnings than the average Chinese (A. F. Alesina et al. 2020).⁵

Valencia (2019) examines educational outcomes in modern-day Paraguay, Brazil, and Argentina. In areas inhabited by Guaraní tribesmen, he finds that areas where the Jesuits proselytized, levels of education (and income) are markedly higher today. The Jesuits were more keenly interested in educating converts than other Catholic orders, building massive schools in the middle of the South-American rainforest. Similar effects do not exist for, say, areas of Guaraní settlement where the Franciscans, another order, built missions. Persistence of educational outcomes over 250 years is remarkable, not least because the Jesuits were expelled from South America, following an agreement between the Spanish and Portuguese crowns in 1759. In other words, just as in the Chinese case studied by Kung et al., Valencia finds that educational persistence is high over the very long run – despite the fact that in his case, the intervention leading to high investments in the past was itself relatively short-lived, compared with the prolonged Chinese exposure to a high-education environment in the past.

Both the Valencia and the Kung et al. studies have the advantage of studying the same variable's long-term evolution. At the same time, factors such as intellectual aptitude and the importance of education are typically partly passed down from parent to child within families, having both a genetic and cultural component – making for a plausible transmission mechanism. Along the same lines, Voigtländer and Voth (2012) analyze anti-Semitic attitudes and acts in Germany, comparing outcomes in the 1920s and 1930s with the geography of medieval pogroms. In these three cases, the concept being measured on the left-hand side of the regression is essentially the same as the one that is on the right-hand side, or the implicit model of causation is relatively straightforward. Sometimes, the concept itself is not straightforward to measure ("anti-Semitism"; "education"), requiring the use of multiple outcome indicators to confirm the similarity of basic patterns across time. Persistence papers in this vein are unlikely to suffer from the problem of multiple hypothesis testing and the attendant inflation of test statistics. They also - implicitly or explicitly - are based on a clear, micro-evidence based model of causation.

A second category of papers examining the roots of attitudes and beliefs relies on well-specified theoretical connections between historical and modern-day outcomes to guide and inspire statistical analysis. Here, the link is often

⁵ This finding echoes results by Clark on the long-term transmission of status in English society (Clark 2014).

indirect, but made plausible by a pre-existing narrative or guiding theory. One example of this is Alesina, Giuliano and Nunn (2013). They argue that plough adoption, with its need for greater upper-body strength, increased men's bargaining power within the household. This led to greater inequality between the sexes, with men being favored in terms of social status, access to food and education, and legal standing. The hypothesis goes back to Esther Boserup, whose book *Women's Role in Economic Development* (Boserup 2007) spelled out the idea that ancient agricultural practices influence gender roles today.

Alesina et al. then go on to show that plough usage in pre-modern agriculture correlates with lower labor force participation of women today - long after the role of agriculture as a source of employment has declined to low levels. To avoid issues of omitted variable bias, the authors use data on soil suitability for different crop types to show that the variation in plough usage driven by exogenous factors predicts gender roles today. Where productive agriculture required men to work long and hard in the fields, it is still the case that women mostly stay at home today. Attitudes in the past were probably changed by a new tool - the plough - and then perpetuated themselves over time. The mechanism's plausibility could have been enhanced further by showing that gender roles were more unequal in historical times. If, say, in 1900 countries with heavy plough usage did not show much evidence of traditional gender roles, and the correlation only emerged in the period after 1950, the plausibility of persistence as a key driver of the gender division of labor today would be called into question.

Similarly, Nunn and Wantchekon (2011) argue that slave catching in Africa from the 17th to the 19th century is a cause of lower trust today. Slave catching was pre-dominantly an activity that pitted Africans against Africans - few Europeans survived African disease environments before the 19th century. Groups that suffered from greater enslavement may have evolved traditions of distrusting outsiders, fellow tribesmen, and family, in order to avoid capture. Where slave catching was intense and long-lasting, this should have created a strong incentive for parents to pass on distrustful attitudes to children. The idea is not novel, but part of a long tradition of historical writings on slavery and its effects on African countries.

The Nunn and Wantchekon study has the advantage of a setting where the causal connection is highly plausible, coming from a rich historical literature. At the same time, because of the well-known feedback effect between income and trust may affect each other, it is harder to think of slavery's effect on trust as persistence in its simplest form. If slave-catching lowered trust historically, and thereby also incomes in the past, thinking about causal effects becomes more

challenging. Since low income environments typically breed distrust (arguably because people do not have the resources to invest in the social connections maintain exchange among equals), the continued co-existence of low incomes and low trust in areas with massive slave hauls in the past is harder to interpret. Instead of showing pure cultural persistence (perhaps as a result of parent to child or horizontal transmission), it could be part of an equilibrium outcome, with two variables reinforcing each other. Here, Nunn's and Wantchekon's use of the epidemiological approach is crucial. By showing that Africans who migrated to other areas still show lower trust today if they came from areas historically affected by slave-catching, they can rule out much of the effect of contemporaneous poverty on trust. What remains is the transmission of the effects of both poverty and trust "back home" on an individual's attitudes.⁶

The most questionable category of persistence papers regress a modern-day outcome on some variable from past, and then argues that a significant coefficient is a sign of persistence. This is all the more problematic if the explanatory variable is multi-faceted, touching on many beliefs and attitudes, like protestantism. Since culture travels in packages – as a set of interrelated characteristics – the actual insight gained by such regressions can be doubted. For example, a recent paper asked “Thank Martin Luther (Kenyon and Fatti 2020) that ciprofloxacin could cure your gonorrhoea?” The paper examines the consumption of antimicrobial medicine in 30 European countries and shows a strong ‘effect’ of Protestantism – which is then rationalized by uncertainty avoidance and a preference for cooperation over performance in Catholic countries are to blame (Kenyon and Fatti 2020). Apart from the doubts one might reasonably have about a cross-country study with 30 observations, it is clear that the association with (historical) Protestantism does not create particular insight into the origins of resistant “super-bugs”.

Mistakes and mechanisms

For a whole class of persistence papers, there exists a readily-available alternative interpretation that is hard to disprove -- economic geography. This is because cross-sectional differences are the main source of variation exploited in persistence studies. As long as it cannot be supplemented by data on movers as in the epidemiological approach, the underlying assumption has to be that geography only matters for creating variation in the cross-section but not in its own right. For example, if geography spelled riches for some Italian cities, but not others in the Middle Ages, then similar factors may still be at work today.

⁶ Note that the solution is necessarily imperfect -- because we know that people's attitudes are more malleable at a young age. If, say, the transmission from income was strong and took place entirely before age 10, looking at movers only would not attenuate the reflection problem.

While the exact set of geographical characteristics that lead to high incomes may have shifted over the centuries, basic features like coastal access, centrality in a network of cities, and a terrain suitable for housing an expanding population could all influence economic performance at the same moment in time. Given that both economic geography and cultural traits are often highly persistent, good practice arguably requires a concerted effort to rule out one in favor of the other.

Few papers explicitly try to explain (let alone model) the geographical distribution of their deep-rooted explanatory variable. In economic geography, the location of economic activity can be explained by two alternative approaches – either by increasing returns interacting with random shocks, as in the work by Krugman (1991) and Fujita (1999), or by locational fundamentals. Davis and Weinstein's (2002) study of Japanese population and production after 1945 shows that even a shock as large as the atomic bombing of Hiroshima and Nagasaki did not shift the distribution of economic activity significantly. This suggests that locational fundamentals are first-order. At the same time, numerous studies have found that temporary shocks can cause long-term shifts in economic activity. Bleakley and Lin's (2012) study of portage sites in the US, Redding and Sturm's work on Germany after 1945 – divided into two parts, with major consequences for economic activity – and Kline and Moretti's (2014) analysis of the Tennessee Valley Authority (TVA) all document that geographically well-defined local shocks can have repercussions over the long term.

Where such well-defined shocks are not a clear source for identification, the claim that 'persistence' is to blame can be hard to substantiate. For example, Fritsch et al. (2019) analyze patterns of self-employment in Kaliningrad -- the formerly German city of Königsberg, today an exclave of the Russian Federation. Comparing data from 1925 and 2010, they find that patterns of self-employment within each industry have remained broadly similar. This is striking because the area around Königsberg suffered heavily during the final phase of World War II. A high proportion of the housing stock, and of industrial installations, was completely destroyed. What remained of the original German population either fled the Soviets or was quickly expelled after 1945. The paper then argues that in areas where small German firms adopted advanced technology (in the form of electric motors), the historical success of small firms was self-perpetuating. As the authors put it:

"The most likely explanation for this persistence is the historical experience with entrepreneurship in a certain industry. In the case of the Kaliningrad area, this experience led to an awareness of the region-specific scope and viability of entrepreneurship among the new population. If this awareness was passed on across

generations, then it is likely to lead to a re-emergence of entrepreneurship after the dissolution of the Soviet Union, which implies high shares of small firms and self-employment rates."

The authors apparently rationalize the continued importance of self-employment as a form of updating in the minds of the new residents -- the Soviet citizens looking around the ruins of Königsberg in 1945 decided, in some places more than in others, that in this environment, self-employment is promising. This insight then lay dormant for half a century, before becoming relevant and useful once the Soviet Union fell. The authors conclude that cultural transmission from Germans to Soviets is an unlikely explanatory factor, as new Russian residents hardly overlapped with previous German inhabitants; but that the awareness of entrepreneurial activity was passed down the generations.

The effects of economic geography are a more likely interpretation. The authors argue that controlling for economic geography (which they proxy by coastal access) does not alter their results. A richer set of controls would likely capture more of the relevant variation. What is being produced where is often driven by regional specialization plus the stickiness of past infrastructure. While population changed entirely, existing factories and infrastructure did not. The Soviets repaired destroyed structures and production equipment on a considerable scale. The fact that industrial structures in 2010 are similar to those in 1925 is therefore not surprising -- not even the atomic blasts of Hiroshima and Nagasaki had major long-term effects on the distribution of economic activity across space in Japan (Davis and Weinstein 2002).⁷ As Fritsch et al.'s regressions show, adding a measure of industrial structure in 1925 already renders historical shares of self-employment insignificant predictors of self-employment in 2010.

Some useful insight into the role of economic geography comes from an innovative paper by Michaels and Rauch (2018). They examine changes in the urban structure of Britain and France after the end of the Roman Empire. While Britain saw a wholesale decline of her cities after the Roman legions withdrew, French cities never disappeared. The "reset" of the British urban system allowed for the creation of new centers in favorable locations during the middle ages. In contrast, French cities were "locked in"; the one-off cost of moving to more favorable locations was never paid. This was costly in the long run. As urban

⁷ In related research, Grosfeld and Zhuravskaya (2015) demonstrate that the quality of historical infrastructure in Poland in the 19th century still has predictive power for modern infrastructure and economic prosperity.

centers re-emerged in medieval Europe, British cities were much more likely to be in places that facilitated rapid growth, with good coastal or river access.

Michaels and Rauch's paper suggests that the shadows of history can be long indeed, and cause major inefficiency. However, if a major shock provides an opportunity for a "reset", locational fundamentals can reassert themselves, becoming key explanatory variables for the location of economic activity in space. The exact circumstances under which persistence is either strong, or fades away, have yet to be established.

Change and re-emergence

Persistence -- defined as the continued or prolonged existence of something -- also sits uncomfortably with the seeming re-emergence of historical traits and features. Individual human memory is not fundamentally different from that of other primates; but collective memory is a cultural construct: By definition, it is not based on the sensory experience of individuals, but an act of cultural production (Nora 1989). Many Israeli conscripts used to take their oath while serving in the IDF at Masada, where Sicarii rebels defied the Romans towards the end of the First Jewish-Roman war. Modern-day recruits could never have met anyone whose living memory included the siege, which concluded in 74 AD -- but it nonetheless defined their country's identity and outlook during the decades after 1948. Reinterpreting situations later in life through the lens of transmitted cultural focal points such as this one, there is 'persistence' -- but in the form of a latent trait or concept that can be activated when current circumstances are reminiscent of a situation in the past (imagined, collectively recalled, or personally experienced). In other words, Masada acquired relevance for Jews living in Palestine once more when faced with the overwhelmingly hostile environment of the Middle East after the Jewish state's founding.

One good example of the more complex interaction of political choices and a remembered past is Ochsner's and Roesel's (2017) paper on the Turkish sieges of Vienna. Turkish advances into central Europe in the 16th and 17th centuries were defining moments, not just for the affected areas, but for Europe as a whole. As was the case with all early modern armies, the Ottoman Empire's invasion of the Austrian lands was accompanied by arson and plunder, rape and pillage. Many cities and towns commemorate the hardships of Turkish occupation to the present day. Heraldic crests recall how cities were burned down; in other locations, street names remind citizens of the horrors of the Turkish invasion.

Interestingly, towns and cities affected by the Turkish advance on Vienna were not more right-leaning than others during the 20th century. This changed radically once the far-right freedom party of Austria began to campaign against a potential Turkish membership of the EU in 2005. While a remote political prospect, the perceived threat of such a move solidified support for the FPÖ -- and all the more so in locations that had suffered at the hands of the Turks. Hence, what appears like a clear case of persistence is actually the result of more complex interactions; instead of suffering at the hands of one group leading to enmity, a long-forgotten incident is revived in public consciousness only if and when doing so is propitious for one political player -- in this case "hate entrepreneurs" along the lines of Glaeser's classic (2005) paper. The standard persistence analysis, a simple regression of FPÖ voting on Turkish occupation in 1529 and 1683, would show a high correlation -- but there is no direct continuity. Instead, as the authors carefully document, it was collective recall, carefully engineered by far-right politicians in the FPÖ, that is created an effect as a result of being part of an "imagined community", in the words of Benedict Anderson (2006).

Some insightful studies look at change and persistence jointly. Bazzi et al. (2017), for example, test the "Turner hypothesis" about the American frontier experience. Turner (1893) had famously argued that living in the newly-settled areas of the American West was instrumental in creating a culture of rugged self-dependence, and that greater scepticism about the role of the government compared to other countries was a direct consequence of the frontier. Bazzi et al. now examine one indicator of individualism -- naming patterns. They show that in counties on the frontier, defined as those with population density below a low threshold of six people per square mile, rare first names back then and today are much more common. In other words, naming patterns reflected a willingness to stand out rather than blend in. Areas that experienced the frontier for a long time believe that the government should not redistribute income, prefer public spending cuts for the poor and welfare, oppose a ban on assault rifles and raising the minimum wage, and prefer to reduce public debt. Bazzi et al. then go on to show that in these places, there is more support for the Republican party, and greater distrust of government intervention, than in parts of the US without the frontier experience.

The frontier could have attracted people who were already more individualistic; and these could then have stayed, with their descendents still believing in the same values. This would be a case of direct cultural persistence. While transmission over a century alone would be interesting, it would not greatly shift priors relative to other papers. Importantly, Bazzi et al. also show that the frontier experience itself made people more individualistic -- families that

moved to the frontier increasingly gave highly unusual names to their children, and stopped using the names of fathers and mothers to name their children. Children whose parents moved to the frontier when they were young were much more likely to be highly individualistic in naming their children, too. Interestingly, they also find that individualists (based on their measure) did better economically in 19th century America – on the frontier. It is the combination of documenting cultural change, selection, and persistence that makes the Bazzi et al. paper a particularly compelling contribution, offering insight into mechanisms, economic feedback effects, and long-term cultural transmission.

Econometric challenges

Recently, both the size and significance of statistical persistence results have come to be questioned. We first discuss general challenges before turning to recent work on persistence and spatial autocorrelation, and the effect of long-term dependence on parameter estimates.

Potential econometric problems in persistence studies abound. Some of them are compounded by publication bias. Few papers estimate precisely estimated zero effects - in contexts where persistence might have been plausible -- or show that some seemingly immutable characteristics disappear. There are many historical events and features that could be investigated, and their effect on modern-day outcomes examined. Take, for example, the role of women in society today, and the geography of witch-hunting. Every argument made for why anti-Semitism might persist over long time periods could apply here; areas with more witch persecution will have had more negative views of women, on average, but will also have induced women to be less assertive, less insistent on equal rights, and less inclined to question traditional gender roles. Some of these mechanisms may have perpetuated themselves even without actual witch persecutions (the last of which still occurred in the 19th century). At the level of entire countries, the argument may have some empirical support -- the country with the most intense persecution of witches, Germany, has remarkably low rates of labor force participation, for example. In other words, there are strong reasons to argue that gender attitudes today should reflect witch-hunting -- but there are no papers making this point. Given the importance of the topic, this may well indicate that the link cannot be proven -- and yet, in surveys of persistence, non-results hardly play any role.

The very range of different variables and analytical categories employed creates the potential for false positives. If there are N different historical variables whose impact might be explored, and M modern-day outcomes that can be

analysed, researchers could theoretically run $N * M$ regressions). At standard levels of statistical significance, $0.1*(N*M)$ or $0.05*(N*M)$ regressions will find an effect that is different from zero. Of course, there are standard ways to adjust for multiple hypothesis testing, such as Bonferroni corrections. One problem is that for any one paper, it is hard to assess what the set of plausible variables is.. Researchers will typically iterate over different variables, looking for significant patterns. Once a pattern emerges, it becomes easier to think of a set of explanations why historical variable x is related to modern-day outcome y .

One simple way to address the multiple hypothesis testing problem would be to pre-register empirical analyses, as is common in development economics. Researchers would have to state *ex ante* which patterns in the data they expect, and why. They could even register the range of additional variables they expect to try, should the main hypothesized relationship not materialize. Researchers might start with the hypothesis that gold discoveries in an area lead to higher discount rates today, because gold mining rewards immediate effort and attracts the impatient. If this regression turned out to be insignificant, but the researchers had moved on to demonstrate that there were larger gay communities in areas of great mineral wealth of any kind (rationalized by the after-effects of a sizeable gender imbalance), it would become immediately apparent that authors were fishing for significance.⁸ Because the number of possible explanatory/dependent variables is specified *ex ante*, the appropriate adjustment could be applied, raising the bar for statistical significance.

The vast majority of persistence papers use geographical variation of a variable at some point in the distant past to explain a later – often modern-day – outcome. In a paper that has attracted substantial attention, Kelly (2019) argues that these papers severely understate the size of standard errors because of spatial correlation -- adjacent observations are not truly independent from each other, but display a great deal of dependence. Kelly examines the extent of spatial correlation analyzing the residuals of the key regressions in 23 papers, using Moran's I. He finds a significant degree of spatial dependence, which he takes to imply that significance levels have been severely overstated. To assess the extent of the problem, Kelly then uses spatial noise as both dependent and independent variable. Regressions using both produce levels of significance that in many cases exceed those for the main regression in the papers in question. While a few papers still show significance at standard levels, the vast majority do not.

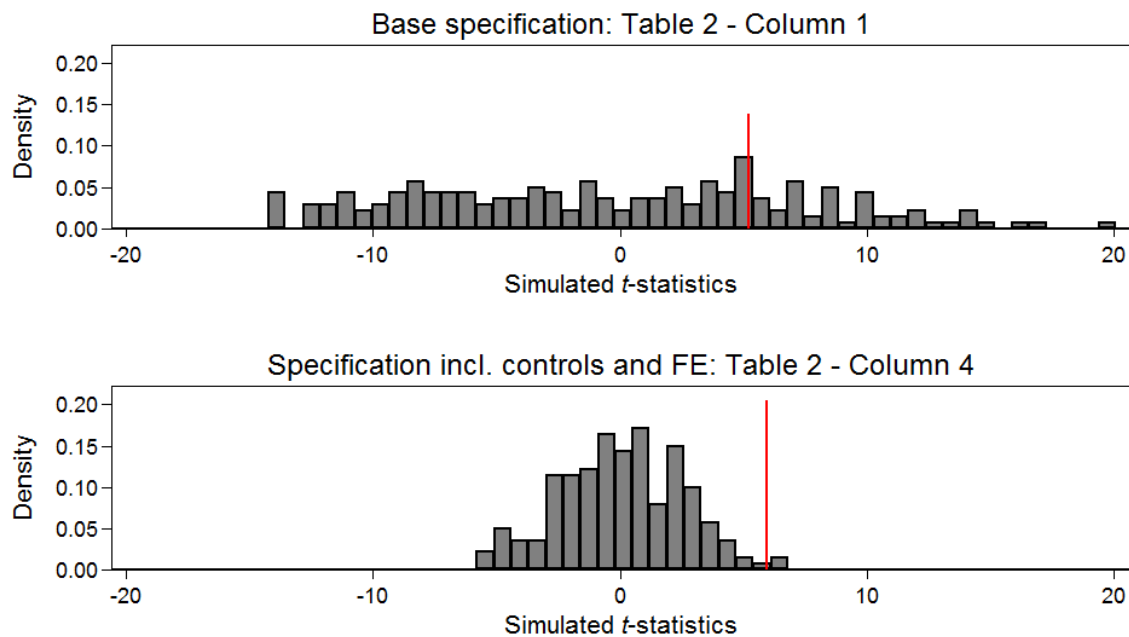
⁸ After this section was written, I was made aware of a recent paper by Brodeur and Haddad (2018). In no way do I wish to imply that the authors followed this path of discovery – this is an illustrative example only.

None of Kelly's analysis is specific to persistence studies as such - any paper using spatial variation will potentially suffer from the same problem.⁹ The potential problem of spatial correlation is also widely known, and subject of a large literature in econometrics that has developed a standard set of tools to deal with it. There appear to be no compelling reasons to cast aside the standard econometricians' toolbox for dealing with spatial error correlations. Equally importantly, there are strong reasons to reject the conclusions drawn by Kelly (2019). Several aspects of his implementation bias results in favor of finding overstated significance levels. First, where authors used a range of variables to capture outcomes, he invariably picks the one that makes results look more fragile. One example is the paper by Voigtländer and Voth (2012) about the link between medieval pogroms and anti-Jewish persecution in Germany. Kelly picks only one of the five different dependent variables used by the authors -- a regression with the percentage of Nazi votes in 1928 on the left-hand side. The paper uses another four additional outcomes that also measure interwar anti-Semitism -- anti-Jewish pogroms before 1933, the share of Jews deported, letters to the anti-Semitic newspaper *Der Stürmer*, and attacks during the Night of Broken Crystal in 1938. For each of these, results even in the baseline specification are markedly stronger when using the Kelly method. As Figure 5 (below) shows, for the example of synagogue attacks, using the Kelly method of simulated spatial noise yields only a 2% probability of noise outperforming.

Equally problematic is the fact that Kelly only picks the first specification presented in each paper. While this allows a certain degree of standardization, it creates a false impression of fragility -- the first regression in empirical papers typically shows only a simple bivariate relationship, demonstrating the strength of basic patterns. Papers that do not show such a basic relationship, but instead rely on saturated specifications to find significant coefficients, are often on shaky empirical grounds -- but according to Kelly's analysis, they will appear superior. All papers on Kelly's list as a matter of course conduct more sophisticated analysis using more extensive specifications. This matters because, if other explanatory variables capture a significant share of the spatial correlation, the error term itself may be largely free of spatially correlated noise.

⁹ Note that Kelly implicitly acknowledges as much, applying a related approach to non-historical papers (Kelly 2020).

Becker and Woessmann (2009) Spatial noise as explanatory variable



Notes: Vertical red line indicates t-statistic of observed explanatory variable.
Correlation range = 4

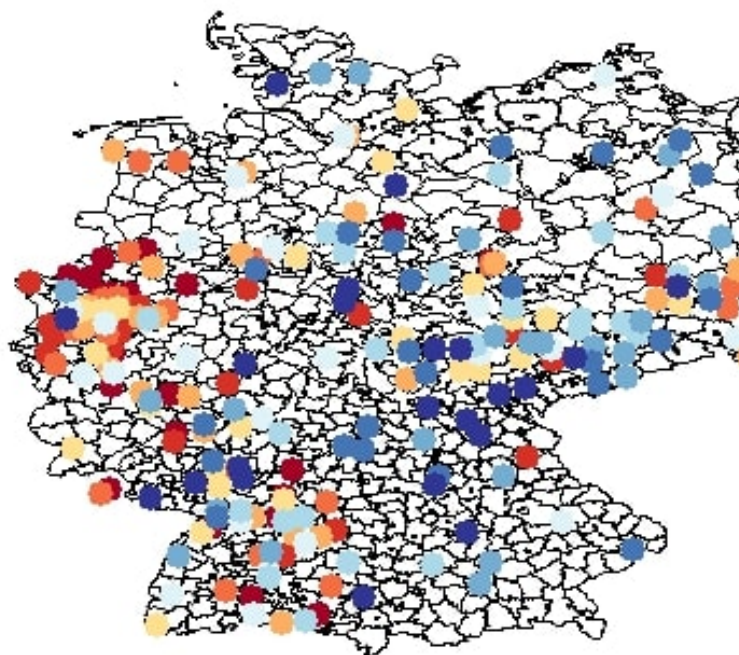
Figure 3: Simulated regression coefficients, education regressed on Protestantism, baseline regression and saturated specification, Kelly (2019 method)

Almost all of the papers in question also consider the possibility of spatially correlated variables directly. Standard remedies like regional fixed effects and clustered standard errors are common in many research papers on persistence. I will next illustrate the strength of the bias introduced by Kelly's focus on the first specification in each paper. Becker and Woessmann (2009) argue that early adoption of Lutheranism led to higher levels in education in 19th century Prussia. In other words, while Max Weber's famous hypothesis about the economic benefits of Protestantism looks right in the cross-section, the reasons for it may well be different -- since Protestants put particular emphasis on reading the bible, human capital was simply higher. In Kelly's exercise, spatial noise (used as an explanatory variable) outperforms the basic regression by Becker and Woessmann 56 percent of the time. This, taken at face value, would suggest that there is no explanatory power at all in their explanatory variable -- the share of Protestants in a Prussian county. Of course, Becker and Woessmann estimate full specifications including latitude and longitude, an interaction of latitude and longitude to control for non-linearities, plus distance to Berlin (Table 2, col 3), to which they add 35 district dummies (col 4). In the saturated specification, their coefficient on percentage Protestant improves

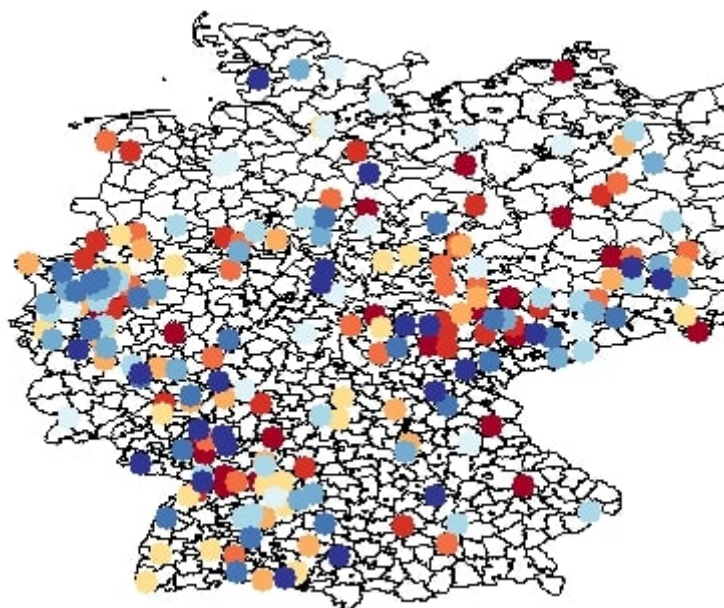
from 0.08 to 0.11, and the t-statistic from 5.15 to 10. Whatever overall similarity there may be in each district in terms of the share of Protestants is hence absorbed by the fixed effect. Put another way, when the authors effectively compare two counties in the same district with different shares of Protestants, holding constant latitude, longitude, latitude x longitude, and distance to Berlin, they find even larger effects. At the same time as their coefficient strengthens, the R^2 increases from 0.057 to 0.81, moving from the basic OLS to the saturated specification. In the spirit of Altonji et al./Oster, this implies that unobservables are unlikely to account for the main finding (Oster 2017; Altonji, Elder, and Taber 2005).

Once we re-estimate the Kelly approach with the full specification, including Becker and Woessmann's controls and fixed effects, noise only outperforms in one percent of simulations. The seeming problem of spatial correlation is well-addressed by what the authors of the original paper already did. Fig. 3 gives an overview of the significance of the share of simulations outperforming the paper's estimated coefficient. The distribution of simulated coefficients shrinks as controls and fixed effects are added. Since the coefficient estimated by Becker and Woessmann barely changes, its significance relative to the distribution of simulated coefficients increases sharply. This means that the overall conclusion in Becker and Woessmann is robust even when using Kelly's method to assess significance.

Figure 4 illustrates why including geographical fixed effects is a very effective tool to control for spatial correlation. It uses variation in Nazi voting in interwar Germany in 1928, and tries to explain it by the density of clubs and associations (Satyanath, Voigtländer, and Voth 2017). The upper panel plots the residuals in the naive regression, simply using a bivariate regression. There is a good deal of spatial dependence. Most of the Ruhr area shows positive deviations, meaning that there was more voting for the NSDAP than expected; the opposite is true in the Eastern part of Germany. The lower panel shows the effect of including province fixed effects. Now, both in the Ruhr and in Saxony positive and negative deviations at the local level sit side-by-side. As is readily apparent from eyeballing the residuals, spatial autocorrelation is well-controlled, with red and blue dots (indicating above and below average levels of Nazi support) alternating in close geographical proximity.



Panel a: Residuals, baseline OLS regression



Panel b: Residuals after using geographical fixed effects and controls

Figure 4: Spatial correlation in Nazi voting, residuals after baseline OLS specification (panel a), or with fixed effects (panel b). Data from Satyanath et al. (2017).

Figure 5 shows that most papers included in Kelly (2019) register a marked improvement once we use the more sensible, "full" specifications included in the original papers. In other words, the majority of researchers working on persistence were perfectly aware of the potential problems arising from spatial correlation, and already took effective steps to sidestep its impact. In doing so, they overcame most of the upward bias in significance; almost no paper suffers

from results that become insignificant in the key specification, once spatial dependence is properly controlled for. What remains are minor disagreements about the exact magnitude of statistical significance.¹⁰

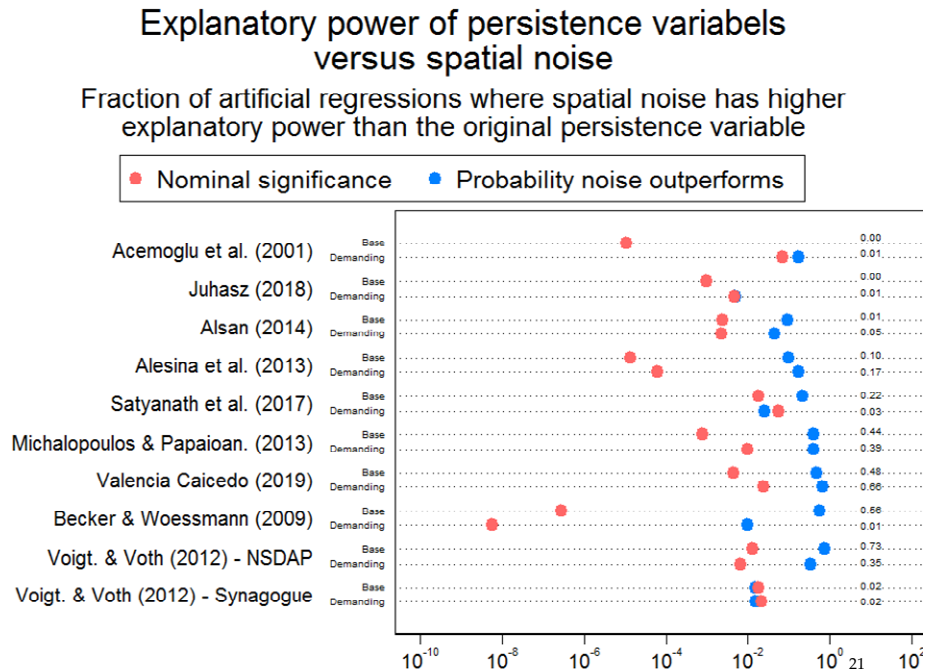


Figure 5: Significance of main parameter estimates, baseline vs full specification, OLS vs Kelly (2019) method

Note: The graph shows the statistical significance of the main variable of interest, for two estimation methods and two specifications (4 outcomes per paper) – simple OLS and the Kelly random noise method, using the baseline specification in the paper and the most demanding specification used. Example: In the Becker & Woessmann paper, using the baseline specification, spatial noise outperforms the estimated coefficient 56% of the time. In the most demanding specification, this is only the case 1% of the time.

There are also major doubts about the conceptual usefulness of Kelly's approach – the percentage of cases where noise outperforms the OLS estimation in Figure 5 is not an accurate guide to the validity of the estimation technique. Monte Carlo simulations by Colella et al. (2019) demonstrate that his horserace of spatial noise regressed on spatial noise is too restrictive, leading to a substantial exaggeration of any existing bias. The standard Conley correction of standard errors for spatial correlation is sufficient to assess the significance of findings. If randomization inference is to be used, it should only be used with spatial noise as the exogenous variable – a test which most persistence papers pass with flying colors. In other words, most of the problem identified by Kelly

¹⁰ The code and data for regressions and simulations in this essay are freely available at <http://www.jvoth.com/data/replication/handbook> and at the ICSPR data repository.

(2019) is a myth. His paper overestimates the scale of the problem, creates an artificially high bar for overcoming it; it is also not compelling in its insistence that persistence papers are particularly prone to suffer from spatial dependence to the point of insignificance.

A different potential objection concerns the size of parameter estimates. Casey and Klemp (2018) argue that the coefficients in persistence regressions suffer from systematic upward bias -- and the more so, the longer the period over a which a trait's persistence is being examined. Their argument is simple. If we are interested in the effect of A on B. A is changing as time goes by. If A is potentially endogenous, we need an instrument to estimate the causal effect of A on B. Now, A is often not observed in the past: instead we measure A', contemporaneous values of the variable. Crucially, historical values of A can influence modern-day values of A'. Casey and Klemp now argue that the OLS coefficient will over-estimate the effect of modern-day A if historical values of A' are not very similar to modern-day ones, i.e. if persistence is low. The intuition for this result is that the OLS coefficient indicates the long-run effect of an increase in historical A that leads to a unit rise in modern-day A. If persistence of A is low, then estimating with IV will systematically overestimate the size of effects.

Casey and Klemp make a point that is conceptually valid. At the same time, the corrections necessary appear relatively small. They apply their approach to deriving correct estimates to the Becker and Woessmann study, for example, and find a coefficient that is only 9% smaller than the one estimated by conventional IV. For the case of Acemoglu et al.'s (2001) famous study of settler mortality, institutions and income, the correction is somewhat larger -- reducing the main coefficient by 45%. However, conceptual challenges appear a potentially greater issue (Albouy 2012).

Conclusions

Econometric challenges in persistence studies undoubtedly have to be taken seriously. In 'apples-and-oranges' studies, there is a real risk of p-hacking because of the wide range of variables that can appear on the left-hand side of an econometric equation. This is especially true when modern-day surveys are used to generate dependent variables, since they typically contain dozens of questions. The potential effect of spatial correlation, on the other hand, has been exaggerated. The problem of autocorrelated spatial residuals biasing standard errors downwards is not specific to persistence studies, and is well-known. Recent work examining various persistence papers with this problem in mind (Kelly 2019) appears to overstate its severity by several orders of

magnitude, and offers no diagnostics that are conceptually or empirically superior to well-established routines (Colella et al. 2019).

Conclusion: Vanishing for Understanding

'My future will not copy fair my past'
- Elizabeth Barret-Browning

To escape one's origins, to leave behind the binding ties of family, place, class and gender, is a dream as old as mankind – as the Barret-Browning suggests. The rapidly growing literature in economics showing a surprising predictive power of the past touched a nerve within economics and beyond, because it seems to question this dream. Nothing can today be done to change a country's technology, or its government institutions, in 1000 BC; nor can African nations exposed to the slave-trade shed their past. One view is that, the tighter the handcuffs of the past shackle the present, the more fatalistic – and less hopeful – economics as a science will have to be. An alternative interpretation is that a deeper understanding of the determinants of riches and poverty can guide policy interventions towards more realistic 'levers of riches' (Mokyr 1990).¹¹

A significant share of papers claiming to find evidence of persistence regress some historical variable on a modern-day outcome. Many are almost entirely "black box", and do not go beyond finding a significant coefficient, or invoking vertical transmission from parent to child as a likely interpretation. This may make sense when the same trait, preference or attitude is measured historically and in modern times --- but this is arguably the case in only a minority of papers. As the conceptual distance between the historical and the modern-day measure increases, the passing on of culture from parent to child becomes less and less likely, and alternative mechanisms become more likely -- such as new equilibria driven by an interaction of economic outcomes with historical characteristics, or the simple stickiness of economic geography and man-made infrastructure. Many persistence studies invoke culture as an arguably plausible mechanism, but it is rare that its role is demonstrated in a rigorous fashion. Where complex interactions between historical variables combined to create differences in initial outcomes, the further scientific value-added of "reduced form" estimates, derived from regressing modern-day outcomes on past indicators, is going to be limited.

¹¹ A similar point is made by Nunn (2020).

As in other areas of empirical work in economics, more work on mechanisms will be helpful. To understand persistence of this (more narrow) kind, we arguably need to look at what it takes to make it go away. This will allow a better understanding of the mechanisms driving it in the first place. An ambitious paper that attempts to accomplish this is Giuliano and Nunn (2017). They argue that the stability of a group's environment encourages a high degree of vertical transmission. In particular, they examine whether areas that experienced considerable shifts in climatic conditions are more conservative in their attitudes, giving greater weight to the opinions of elders and the like. In each generation, people have a choice of either copying from previous generations, or of experimenting with new technologies, agricultural practices, or forms of interaction. The more stable the environment, the greater the similarity of conditions today relative to those of the previous generation. The extent of variability then affects the degree of stability of preferences and practices. To measure stability of the environment, Giuliano and Nunn (2017) look at climatic change over the period 500-1900. Across a range of empirical settings, they find support for their basic prediction: In the World Value Survey, people from countries with greater climatic variation in the past show less regard for tradition. When looking at immigrants to the US, there is also less outgroup marriage for those who come from countries where climate stayed large the same over the last 1,500 years. Similarly, immigrants from countries with stable climates were also less likely to speak English at home after migrating to the US. They then go on to examine stability of the same traits over time, and the extent to which this is attenuated by climatic variability. For three variables -- cousin marriage, polygamy, and female labor force participation -- they find a high degree of persistence, as well as an attenuating effect of climatic instability. Giuliano and Nunn conclude that persistence is an outcome, which can be rationalized based on the costs and benefits associated with it.

Along similar lines, Voigtländer and Voth (2012) document that the predictive power of medieval anti-Semitism in Germany in general was high for persecution in the interwar period. However, persistence is lower in towns and cities with a greater immigration in the 19th century. In other words, where modern-day inhabitants were largely unrelated to the people living in a location centuries ago, persistence – narrowly defined – vanishes. The same is true in old trading cities, where there was a particularly strong trade-off between prejudice on the one hand, and economic success on the other. Differential migration and its effect on the composition in general is arguably underappreciated as a driver of cultural change: Ichino and Maggi (2000) demonstrate that in Italy, the more reliable employees gravitated towards the North, where they would find co-workers with similar characteristics. Similarly,

East Germany's continued underperformance after reunification may echo the size of the brain drain prior to 1961 (Sascha O. Becker, Mergele, and Woessmann 2020).¹² Such findings reconnect with some of the theoretical work on the trade-offs involved in cultural transmission – in this case, the canonical Bisin-Verdier model of parental investment in childrens' attitudes (Bisin and Verdier 2000; 2001).¹³

The second meaning of “persistence” highlighted by the *Oxford English Dictionary* is the “fact of continuing in an opinion or course of action in spite of difficulty or opposition.” To make progress, persistence studies will have to grapple not just with scepticism by professional historians, but to find ways of dealing effectively with the challenges highlighted in this essay. In ‘apple-on-apple’ analyses, where the *same* or near-identical variables are on the right- and left-hand side, and where attitudes, beliefs, cultural preferences or social arrangements are often the focus, we can speak of persistence in its basic form. Going forward, a greater focus on persistence in this stricter sense of the word may be helpful.

References

- Acemoglu, Daron, Simon Johnson, and James A Robinson. 2001. ‘The Colonial Origins of Comparative Development: An Empirical Investigation’. *American Economic Review* 91 (5): 1369–1401.
- . 2002. ‘Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution’. *The Quarterly Journal of Economics* 117 (4): 1231–94.
- Albouy, David Y. 2012. ‘The Colonial Origins of Comparative Development: An Empirical Investigation: Comment’. *American Economic Review* 102 (6): 3059–76.
- Alesina, A, P Giuliano, and N Nunn. 2013. ‘On the Origins of Gender Roles: Women and the Plough’. *Quarterly Journal of Economics*, no. 2: 469–530.
- Alesina, Alberto F., Marlon Seror, David Y. Yang, Yang You, and Weihong Zeng. 2020. *Persistence through Revolutions*. NBER Working Paper Series, no. w27053. Cambridge, Mass: National Bureau of Economic Research.

¹² At the same time, if migration is slow enough, even substantial turnover may go hand-in-hand with stable cultural preferences, if horizontal transmission is sufficiently powerful (Sascha O. Becker et al. 2016)

¹³ For a related approach exploring interactions with economic incentives, cf. Doepke and Zilibotti (2008).

- Alesina, Alberto, Edward Glaeser, and Bruce Sacerdote. 2001. 'Why Doesn't the United States Have a European-Style Welfare State?' *Brookings Papers on Economic Activity*, 187–278.
- Altonji, Joseph G, Todd E Elder, and Christopher R Taber. 2005. 'Selection on Observed and Unobserved Variables'. *Journal of Political Economy* 113 (1): 151–84.
- Anderson, Benedict. 2006. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. Verso Books.
- Ashraf, Quamrul, and Oded Galor. 2013. 'The "Out of Africa" Hypothesis, Human Genetic Diversity, and Comparative Economic Development'. *American Economic Review* 103 (1): 1–46.
- Baumard, Nicolas. 2019. 'Psychological Origins of the Industrial Revolution'. *Behavioral and Brain Sciences* 42.
- Bazzi, Samuel, Martin Fiszbein, and Mesay Gebresilashe. 2017. 'Frontier Culture: The Roots and Persistence of "Rugged Individualism" in the United States'. Wp 0898–2937. National Bureau of Economic Research.
- Becker, Sascha O., Katrin Boeckh, Christa Hainz, and Ludger Woessmann. 2016. 'The Empire Is Dead, Long Live the Empire! Long - run Persistence of Trust and Corruption in the Bureaucracy'. *The Economic Journal* 126 (590): 40–74.
- Becker, Sascha O., Lukas Mergele, and Ludger Woessmann. 2020. 'The Separation and Reunification of Germany: Rethinking a Natural Experiment Interpretation of the Enduring Effects of Communism'. *Journal of Economic Perspectives* 34 (2): 143–71.
- Becker, S.O., and L. Woessmann. 2009. 'Was Weber Wrong? A Human Capital Theory of Protestant Economic History'. *Quarterly Journal of Economics* 124 (2): 531–96.
- Besley, T., and T. Persson. 2009. 'The Origins of State Capacity: Property Rights, Taxation, and Politics'. *American Economic Review* 99 (4): 1218–44.
- . 2010. 'State Capacity, Conflict, and Development'. *Econometrica* 78 (1): 1–34. <https://doi.org/10.3982/ECTA8073>.
- Bisin, Alberto, and Thierry Verdier. 2000. "Beyond The Melting Pot": Cultural Transmission, Marriage, And The Evolution Of Ethnic And Religious Traits'. *Quarterly Journal of Economics* 115 (3): 955–88.
- . 2001. 'The Economics of Cultural Transmission and the Dynamics of Preferences'. *Journal of Economic Theory* 97 (2): 298–319.
- Bleakley, Hoyt, and Jeffrey Lin. 2012. 'Portage and Path Dependence'. *Quarterly Journal of Economics* 127 (2): 587–644.
- Boserup, Ester. 2007. *Woman's Role in Economic Development*. Earthscan.
- Brodeur, Abel, and Joanne Haddad. 2018. 'Institutions, Attitudes and LGBT: Evidence from the Gold Rush'.

- Casey, Gregory, and Marc Klemp. 2018. 'Instrumental Variables in the Long Run'.
- Chen, Ting, James Kai-sing Kung, and Chicheng Ma. 2020. 'Long Live Keju! The Persistent Effects of China's Imperial Examination System'. *Economic Journal* forthcoming.
- Cioni, Martina, Giovanni Federico, and Michelangelo Vasta. 2020. 'The Long-Term Evolution of Economic History: Evidence from the Top Five Field Journals (1927–2017)'. *Cliometrica* 14 (1): 1–39. <https://doi.org/10.1007/s11698-019-00186-x>.
- Clark, Gregory. 2014. *The Son Also Rises: Surnames and the History of Social Mobility*. Princeton University Press.
- Colella, Fabrizio, Rafael Lalive, Seyhun Orcan Sakalli, and Mathias Thoenig. 2019. 'Inference with Arbitrary Clustering'. *IZA Discussion Paper*.
- Comin, Diego, William Easterly, and Erick Gong. 2010. 'Was the Wealth of Nations Determined in 1000 BC?' *American Economic Journal: Macroeconomics* 2 (3): 65–97.
- Davis, Donald R, and David E Weinstein. 2002. 'Bones, Bombs, and Break Points: The Geography of Economic Activity'. *American Economic Review* 92 (5): 1269–89.
- De la Croix, David, Matthias Doepke, and Joel Mokyr. 2018. 'Clans, Guilds, and Markets: Apprenticeship Institutions and Growth in the Preindustrial Economy'. *The Quarterly Journal of Economics* 133 (1): 1–70.
- Dell, Melissa. 2010. 'The Persistent Effects of Peru's Mining Mita'. *Econometrica* 78 (6): 1863–1903.
- Dell, Melissa, Nathan Lane, and Pablo Querubin. 2018. 'The Historical State, Local Collective Action, and Economic Development in Vietnam'. *Econometrica* 86 (6): 2083–2121.
- Diamond, Jared M. 1997. *Guns, Germs, and Steel: The Fates of Human Societies*. Norton.
- Doepke, Matthias, and Fabrizio Zilibotti. 2008. 'Occupational Choice and the Spirit of Capitalism'. *Quarterly Journal of Economics* 123 (2): 747–93.
- Dohmen, T., A. Falk, D. Huffman, and U. Sunde. 2012. 'The Intergenerational Transmission of Risk and Trust Attitudes'. *Review of Economic Studies* 79 (2): 645–77.
- Elias, Norbert, and Edmund Jephcott. 1994. *The Civilizing Process*. Vol. 2. Blackwell Oxford.
- Engerman, Stanley, and Kenneth Sokoloff. 1997. *Factor Endowments, Institutions, and Differential Growth Paths among New World Economies*. In *How Latin America Fell Behind: Essays on the Economic Histories of Brazil and Mexico* by Haber, S. Vol. 89. Stanford University Press.

- Fernandez, Raquel, and Alessandra Fogli. 2009. 'Culture: An Empirical Investigation of Beliefs, Work, and Fertility'. *American Economic Journal: Macroeconomics* 1 (1): 146–77.
- Fritsch, Michael, Alina Sorgner, Michael Wyrwich, and Evguenii Zazdravnykh. 2019. 'Historical Shocks and Persistence of Economic Activity: Evidence on Self-Employment from a Unique Natural Experiment'. *Regional Studies* 53 (6): 790–802.
- Fuchs-Schundeln, Nicola, and Alberto Alesina. 2007. 'Good-Bye Lenin (Or Not?)'. *American Economic Review* 97 (4): 1507–28.
- Fujita, Masahisa, Paul R Krugman, and Anthony Venables. 1999. *The Spatial Economy: Cities, Regions, and International Trade*. MIT press.
- Gennaioli, Nicola, and Ilia Rainer. 2007. 'The Modern Impact of Precolonial Centralization in Africa'. *Journal of Economic Growth* 12 (3): 185–234.
- Giuliano, Paola, and Nathan Nunn. 2017. 'Understanding Cultural Persistence and Change'. 0898–2937. National Bureau of Economic Research.
- Glaeser, E.L. 2005. 'The Political Economy of Hatred'. *Quarterly Journal of Economics* 120 (1): 45–86.
- Grosfeld, Irena, and Ekaterina Zhuravskaya. 2015. 'Cultural vs. Economic Legacies of Empires: Evidence from the Partition of Poland'. *Journal of Comparative Economics* 43 (1): 55–75.
- Guedes, Jade d'Alpoim, Theodore C Bestor, David Carrasco, Rowan Flad, Ethan Fosse, Michael Herzfeld, Carl C Lamberg-Karlovsky, Cecil M Lewis, Matthew Liebmann, and Richard Meadow. 2013. 'Is Poverty in Our Genes? A Critique of Ashraf and Galor, "The 'out of Africa' hypothesis, Human Genetic Diversity, and Comparative Economic Development," American Economic Review (Forthcoming)'. *Current Anthropology* 54 (1): 71–79.
- Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2016. 'Long Term Persistence'. *Journal of the European Economic Association* 14 (6): 1401–36.
- Hsieh, C. T. 2002. 'What Explains the Industrial Revolution in East Asia? Evidence from the Factor Markets'. *The American Economic Review* 92 (3): 502–26.
- Ichino, Andrea, and Giovanni Maggi. 2000. 'Work Environment and Individual Background: Explaining Regional Shirking Differentials in a Large Italian Firm'. *The Quarterly Journal of Economics* 115 (3): 1057–90.
- Kelly, Morgan. 2019. 'The Standard Error of Persistence'. *UCD Working Paper*.
- . 2020. 'Direct Standard Errors for Regressions with Spatially Autocorrelated Residuals'.
- Kenyon, Chris, and Geoffrey Fatti. 2020. 'Thank Martin Luther That Ciprofloxacin Could Cure Your Gonorrhoea? Ecological Association between Protestantism and Antimicrobial Consumption in 30 European Countries'. *Mimeo*.

- Kim, Jong-Il, and Lawrence J Lau. 1994. 'The Sources of Economic Growth of the East Asian Newly Industrialized Countries'. *Journal of the Japanese and International Economies* 8 (3): 235–71.
- Kline, Patrick, and Enrico Moretti. 2014. 'Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority'. *The Quarterly Journal of Economics* 129 (1): 275–331.
- Koselleck, Reinhard. 2003. 'Wiederholungsstrukturen in Der Geschichte'. *Divinatio*, no. 17: 17–32.
- Krugman, Paul. 1991. 'Increasing Returns and Economic Geography'. *Journal of Political Economy* 99 (3): 483–99.
- Landes, David S. 2006. 'Why Europe and the West? Why Not China?' *Journal of Economic Perspectives* 20 (2): 3–22.
- Lin, Justin Yifu. 1995. 'The Needham Puzzle: Why the Industrial Revolution Did Not Originate in China'. *Economic Development and Cultural Change* 43 (2): 269–92.
- Lowes, Sara, Nathan Nunn, James A Robinson, and Jonathan L Weigel. 2017. 'The Evolution of Culture and Institutions: Evidence from the Kuba Kingdom'. *Econometrica* 85 (4): 1065–91.
- Maloney, William F, and Felipe Valencia Caicedo. 2012. *The Persistence of (Subnational) Fortune: Geography, Agglomeration, and Institutions in the New World*. The World Bank.
- Michaels, Guy, and Ferdinand Rauch. 2018. 'Resetting the Urban Network: 117–2012'. *The Economic Journal* 128 (608): 378–412.
- Mokyr, Joel. 1990. *The Lever of Riches: Technological Creativity and Economic Progress*. New York: Oxford University Press.
- Needham, Joseph. 1976. *Science and Civilisation in China*. Vol. 5. Cambridge University Press.
- Nora, Pierre. 1989. 'Between Memory and History: Les Lieux de Mémoire'. *Representations*, 7–24.
- Nunn, Nathan. 2020. 'The Historical Roots of Economic Development'. *Science* 367 (6485).
- Nunn, Nathan, and Leonard Wantchekon. 2011. 'The Slave Trade and the Origins of Mistrust in Africa'. *American Economic Review* 101 (7): 3221–52.
- Ochsner, Christian, and Felix Roesel. 2017. 'Activated History-the Case of the Turkish Sieges of Vienna'.
- Oster, Emily. 2017. 'Unobservable Selection and Coefficient Stability: Theory and Evidence'. *Journal of Business & Economic Statistics*, 1–18.
<https://doi.org/10.1080/07350015.2016.1227711>.
- Pinker, Steven. 2012. *The Better Angels of Our Nature: Why Violence Has Declined*. Penguin Group USA.

- Pritchett, Lant. 1997. 'Divergence, Big Time'. *Journal of Economic Perspectives* 11 (3): 3–17.
- Putterman, L, and DN Weil. 2010. 'Post-1500 Population Flows and the Long Run Determinants of Economic Growth and Inequality'. *Quarterly Journal of Economics* 125 (4): 1627–82.
- Satyanath, Shanker, Nico Voigtländer, and Hans-Joachim Voth. 2017. 'Bowling for Fascism: Social Capital and the Rise of the Nazi Party'. *Journal of Political Economy* 125 (2): 478–526.
- Sokoloff, Kenneth L, and Stanley L Engerman. 2000. 'History Lessons: Institutions, Factors Endowments, and Paths of Development in the New World'. *The Journal of Economic Perspectives*, 217–232.
- Spolaore, Enrico, and Romain Wacziarg. 2009. 'The Diffusion of Development'. *The Quarterly Journal of Economics* 124 (2): 469–529.
- Turner, Frederick Jackson, and Allan G Bogue. 1893. *The Frontier in American History*. Courier Corporation.
- Valencia Caicedo, Felipe. 2019. 'The Mission: Human Capital Transmission, Economic Persistence, and Culture in South America'. *The Quarterly Journal of Economics* 134 (1): 507–56.
- Voigtländer, Nico, and Hans-Joachim Voth. 2012. 'Persecution Perpetuated: The Medieval Origins of Anti-Semitic Violence in Nazi Germany'. *Quarterly Journal of Economics* 127 (3): 1339–92.
- . 2013a. 'Gifts of Mars: Warfare and Europe's Early Rise to Riches'. *Journal of Economic Perspectives*.
- . 2013b. 'The Three Horsemen of Riches: Plague, War, and Urbanization in Early Modern Europe'. *The Review of Economic Studies*, no. Forthcoming..
- . 2015. 'Nazi Indoctrination and Anti-Semitic Beliefs in German'. *PNAS* 112 (26): 7931–36.
- Young, Alwyn. 1995. 'The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience'. *The Quarterly Journal of Economics* 110 (3): 641–80.