NBER WORKING PAPER SERIES

LATE FOR HISTORY

Alberto Bisin
Andrea Moro

Working Paper 28113
http://www.nber.org/papers/w28113

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2020, Revised February 2021

Prepared for the Handbook of Historical Economics, by Alberto Bisin and Giovanni Federico, eds, Elsevier North-Holland, forthcoming 2021. We thank Thierry Verdier for countless exchanges on these topics over the years. We also thank Jess Benhabib, Giovanni Federico, Atsushi Inoue, Franco Peracchi, and Pedro Sant'Anna for helpful comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2020 by Alberto Bisin and Andrea Moro. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

LATE for History
Alberto Bisin and Andrea Moro
NBER Working Paper No. 28113
November 2020, Revised February 2021
JEL No. C36,C51,N01

ABSTRACT

In Historical Economics, Persistence studies document the persistence of some historical phenomenon or leverage this persistence to identify causal relationships of interest in the present. In this chapter, we analyze the implications of allowing for heterogeneous treatment effects in these studies. We delineate their common empirical structure, argue that heterogeneous treatment effects are likely in their context, and propose minimal abstract models that help interpret results and guide the development of empirical strategies to uncover the mechanisms generating the effects.

Alberto Bisin
Department of Economics
New York University
19 West 4th Street, 5th Floor
New York, NY 10012
and NBER
alberto.bisin@nyu.edu

Andrea Moro
Department of Economics
Vanderbilt University
VU Station B #351819
Nashville, TN 37235
andrea.moro@vanderbilt.edu

LATE for History[‡]

Alberto Bisin[§] Andrea Moro[¶]

February 16, 2021

Abstract

In Historical Economics, Persistence studies document the persistence of some historical phenomenon or leverage this persistence to identify causal relationships of interest in the present. In this chapter, we analyze the implications of allowing for *heterogeneous treatment effects* in these studies. We delineate their common empirical structure, argue that heterogeneous treatment effects are likely in their context, and propose minimal abstract models that help interpret results and guide the development of empirical strategies to uncover the mechanisms generating the effects.

Keywords: Local average treatment effects Persistence studies Instrumental variables Identification Institutions Culture

1 Introduction

Many studies in Historical Economics document the persistence of some historical phenomenon, while others leverage this persistence to identify causal relationships of interest in the present. These are generally referred to as *Persistence studies* (Voth, 2021, in this book). In this chapter, we analyze the implications of allowing for *heterogeneous treatment effects* in these studies, we delineate their common empirical structure, argue that heterogeneous treatment effects are likely in their context, and

[‡]Prepared for the *Handbook of Historical Economics*, by Alberto Bisin and Giovanni Federico, eds, Elsevier North-Holland, forthcoming 2021. We thank Thierry Verdier for countless exchanges on these topics over the years. We also thank Jess Benhabib, Giovanni Federico, Atsushi Inoue, Franco Peracchi, and Pedro Sant'Anna for helpful comments

[§]Department of Economics, New York University, and NBER, alberto.bisin@nyu.edu

[¶]Department of Economics, Vanderbilt University, andrea@andreamoro.net

propose simple abstract models that help interpret results and guide the development of empirical strategies to uncover the mechanisms generating the effects.

Persistence studies focus empirically on the effects of a treatment variable in the present, assuming its persistence from the historical past. Consider as an illustration the effects of cultural norms or institutions, e.g., on economic development. The adoption of cultural norms or the process of institutional change can be viewed, using the language of the causal inference literature, as taking-up treatment in the past. High-quality norms or institutions may then persist over time and thereby realize their effects on economic development in the present. An exogenous historical factor may be available that directly affects the treatment variable and can be exploited as an instrumental variable to identify a causal effect of norms or institutions on development.

The adoption of norms or institutions may be correlated with the returns in terms of economic development. If these returns are heterogeneous across countries - that is, if treatment effects are heterogeneous - countries with higher returns may be more likely to have adopted higher quality institutions or norms.[1] Heterogeneous treatments may then change the interpretation of the identified relationship between norms or institutions and economic development. For example, if higher values of the instrument induce countries with relatively higher returns to adopt higher quality norms or institutions, then the instrumental variable procedure identifies the effects produced only in high-return countries, possibly overstating the average returns of all countries. Similarly, the instrument could activate institutional and cultural changes over time, interacting with successive independent historical phenomena, inducing a special selection of heterogeneous treatment effects. Finally, even if the impact of institutions and norms in the present is homogeneous, the persistence of institutional and cultural changes in the historical past activated by treatment can be heterogeneous, affecting the interpretation of the nature of persistence.

When treatment effects are heterogeneous, therefore, the causal arguments remain generally unaffected; but the interpretation of estimated coefficients may pose new empirical questions to analyze further the mechanisms underlying the identified causal relationships. More specifically, when treatment effects are heterogeneous, the research design identifies a *Local Average Treatment Effect* (LATE), rather than the Average (subject-level) Treatment Effect (ATE); see Imbens and Angrist (1994). In the investigation of the effects of cultural norms or institutions, for example, the *local*

[1]This is often labeled a "selection on returns" or "selection on gains," see Heckman et al. (2010).

effect identified by the instrumental variable procedure is the effect in countries that take-up treatment (higher quality institutions or norms) only when instrumented.[2]

Heterogeneous effects are bound to be important in Persistence studies. In fact, the persistent variable in Persistence studies is typically a cultural (or genetic) trait or an institutional feature or norm.[3] Because these variables are often defined by such sweeping concepts such as institution quality, cultural, or civic norms, they are likely affected by several underlying heterogeneous mechanisms. Furthermore, because treatment is generally taken-up in the historical past, various complex and heterogeneous dynamical processes may affect it, intervening in the determination of the objective of the analysis in the present.

Since heterogeneous treatment effects do not generally affect the causal identification argument, in this chapter we shy away from discussing the validity of the data and econometric procedures adopted in these studies.[4] We focus instead on the interpretation of the estimated coefficients under heterogeneous treatment effects. While the identification of causal effects in Persistence studies has produced significant first-order results, it has also highlighted how little we know about the mechanisms driving these effects. The paucity of data in historical contexts makes mechanisms hard to identify both with a quasi-experimental design and with a structural econometrics approach. For the same reason, it is hard to identify the distribution of treatment heterogeneity.

To guide our understanding about what moments of this treatment parameters

[2]In environments where this *essential heterogeneity* is present Heckman and Vytlacil (2005) show that the IV procedure identifies a weighted average of the returns and characterize the weights in terms of the relationship between the IV and the returns. In the Appendix we introduce a simple, informal primer on the distinction between LATE and ATE in the context of labor economics, to level the field between historians, economists, and other social scientists who might not be uniformly attuned to these concepts.

[3]Recent examples have exploited variation in a wide range of historical variables: from colonial settlers' mortality to the self-governance of Italian cities, the adoption of the plow in agriculture, medieval pogroms, the size of the slave trade from Africa. Some of these studies have identified causal relationships between phenomena of interest in the present, for example, the effect of institutional quality or civic capital on economic performance. Others have documented the persistence of various dimensions of institutional and cultural characteristics, including trust, gender attitudes, and anti-semitism. Persistence studies are a sizeable component of Historical Economics, about 10%, according to Cioni et al. (2021, in this book)'s classification; see their Table 2 and 4.

[4]Persistence studies have been the subject of great scrutiny. The credibility of the causal parameter estimates stands on the reliability of historical data, assumptions about the econometric model's error structures, and, in most cases, on the adoption of a valid instrumental variable (Casey and Klemp (2021)). Because most of these studies exploit variation across geographic dimensions, spatial correlations of residuals can be a severe econometric issue; see Kelly (2019) and Voth (2021, in this book).

distribution the estimates identify, it is therefore important to study the relationship between the instrument and treatment and how the mechanism responsible for the persistence of treatment over time correlates with the values of treatment effects. Explicit models of these relationships and mechanisms, in the context of the specific empirical analysis, help to clarify the interpretation of the identified causal effects and help the formulation of interesting new sets of questions, which can be addressed empirically, possibly with new data. To this end, we develop simple abstract models linking treatment take-up and treatment's persistence over history with treatment effects. These models delineate the underlying causal relationships and the role of the research design in the identification problem. Explicit models of these relationships and mechanisms represent the outcome of political equilibrium processes or the aggregation of individuals' relevant behavioral choices. This is typically the case, in particular, when treatment involves institutional change or change in cultural attitudes and traits as just described. To illustrate the role of these models in the abstract, without imposing a fully developed structure (which would be necessarily context dependent), we will construct minimal reduced-form models of treatment take-upwithout being explicit about their behavioral and equilibrium micro-foundations. These models provide us with interpretations of the estimated parameters, opening new empirical questions, possibly with new data.

We proceed, in the next section, by formalizing Persistence studies' common structure. In Section 3 we apply this framework to persistence studies whose main goal is to identify the causal relationship between variables in the present. These studies exploit the persistence of a variable in the historical past to provide an instrument. In Section 4, we focus instead on a set of persistence studies that are directly interested in investigating the persistence of historical variables.

2 Persistence studies

In this section, we introduce the issue of identification of causal relationships in Persistence studies. We delineate their common empirical structure by constructing a simple formal framework that encompasses most papers in the literature.[5]

[5] For a book-length treatment of causal analysis in econometrics, see Angrist and Pischke (2008, 2014). For a more abstract approach to causality, see Pearl (2009); Pearl et al. (2016).

2.1 Empirical model

Let any variables measured in present time be indexed by t and any historical variables by $t-h$, where h is the historical lag considered. Let $i = 1, 2, \ldots, N$ index the cross-section in the data, typically defined along political or geographic dimensions: countries, cities, ethnic groups, etc... Consider the following cross-sectional relationship between variables at present time t:

$$y_t = \alpha + \beta x_t + u_t, \tag{1}$$

where y_t, x_t, u_t are $N-$dimensional vectors indexed by location $i = 1, 2, \ldots, N$; the parameter β is also an $N-$dimensional vector: distinct β_i, across locations $i = 1, 2, \ldots, N$, represent the heterogeneous effects which are the focus of our analysis.[6] The explanatory variable x_t is generally endogenous, e.g., because of a common factor affecting both x_t and y_t or because of two-way causation between y_t and x_t.

History enters the empirical model through the underlying (unobservable) historical dynamics of the explanatory variable x_t, governed by a stochastic process, $\{x_\tau\}_{\tau \in T}$, where T denotes the historical sequence of time until the present $\{t - h, \ldots, \tau, \ldots t\}$. We model the persistence of the process $\{x_\tau\}$ assuming

$$cov(x_{t-h}, x_t) = \rho; \tag{2}$$

where ρ is also an $N-$dimensional vector, allowing for heterogeneity of persistence across locations $i = 1, 2, \ldots, N$ (for simplicity we assume that ρ does not depend on h or t).

The econometrician observes an instrument for x_t that we will assume to be valid throughout the chapter, in the sense defined below (Assumption 1). The characterizing feature of Persistence studies is that the instrument, an $N-$dimensional vector z_{t-h}, is realized in the historical past, at $t-h$. The empirical structure underlying the IV strategy depends crucially on the historical persistence of the instrumented variable, x_t. The instrument z_{t-h} is assumed to affect causally and directly x_{t-h} and hence indirectly x_t through the persistence of the process $\{x_\tau\}_{\tau \in T}$ in history. This structure does not require the econometrician to observe the realizations of the stochastic process at any time other than t; in particular, x_{t-h} is generally not

[6] We abuse notation by not distinguishing the random variables in the population from their sample realizations. Also, in our notation, products between vectors are to be intended as Hadamard products; so that, in Equation 1, $\beta x_t = [\beta_i x_{i,t}]_{i=1}^N$.

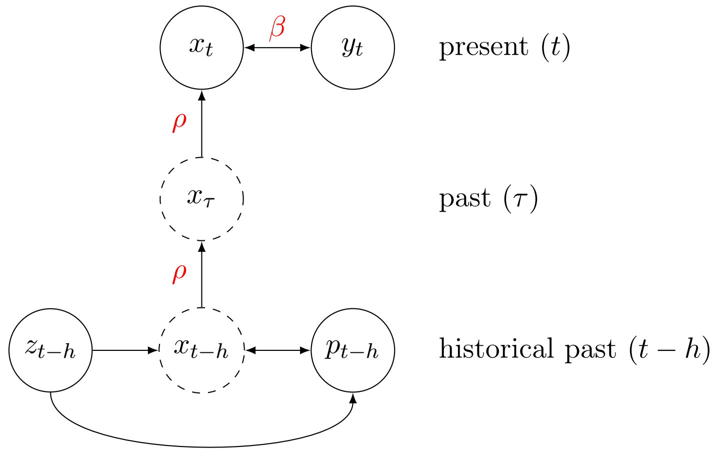


Figure 1: Persistence studies: general case.

Note: Circles indicate variables observed by the investigator. Dashed circles indicate unobserved variables. Solid arrows indicate directions of causality. Double arrows indicate endogeneity or any other factor preventing the identification of a causal effect (omitted variables, selection bias, etc...).

observable to the econometrician. However, if a proxy p_{t-h} for x_{t-h} is observable, z_{t-h} can be an instrument for p_{t-h}, even though p_{t-h} might be endogenous with respect to x_{t-h}. Fig. 1 illustrates the relationship between variables in this empirical model.[7]

The examples from the literature that we study in the rest of the paper focus on two separate objectives of the empirical analysis. One is to estimate some moment of the cross-section of treatment effects, β_i, or another relevant feature of their distribution. When data on a proxy p_{t-h} for x_{t-h} are available, another objective of the analysis is to estimate some relevant feature of the distribution of the cross-section of persistence effects, ρ_i.

2.2 Heterogeneity and LATE effects

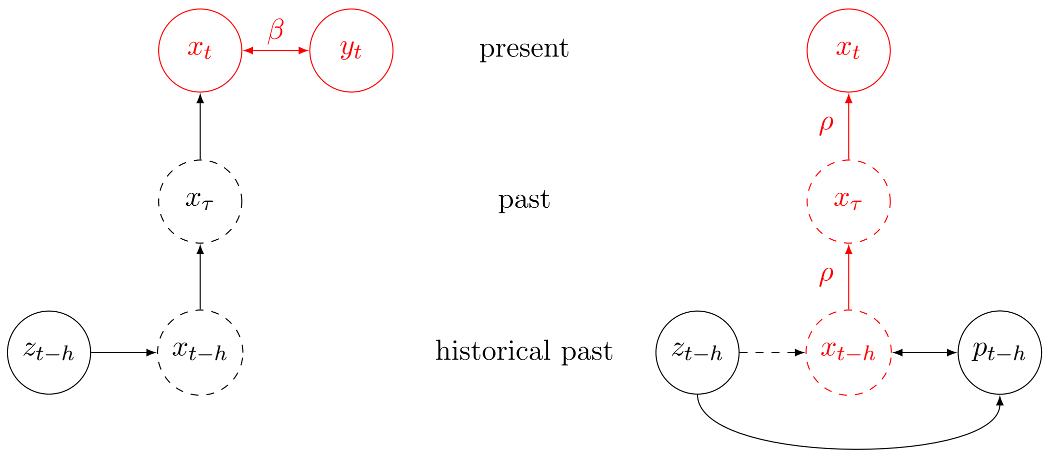
Allowing for heterogeneous treatment effects might have important implications with respect to how these effects, as they are identified in Persistence studies, are inter-

[7]It is difficult to represent models with endogenous, equilibrium relationships, or to represent the distinction between LATE and ATE, using Directed Acyclical Graphs (see Imbens (2019)). This chart and the next ones are meant to illustrate the relationships between variables in the various empirical models, not as a tool to discuss the causal identification designs, which we study analytically in the text.

preted. Consider the empirical model delineated in Section 2.1 and represented in Fig. 1. Assume that either the relationship between x_t and y_t and/or the persistence effects from x_{t-h} to x_t are heterogeneous; that is, β, ρ are heterogeneous across locations $i = 1, 2, \ldots, N$. As long as z_{t-h} is a valid instrument for x_t or for p_{t-h}, the IV procedure identifies a causal effect; that is, the causal arguments which are the objective of Persistence studies remain unaffected. But consider the case in which the mechanism inducing location i to take-up treatment after being instrumented, or to maintain the treatment over time, depends on the values of β and ρ assumed by that location. In this case, heterogeneity matters for the interpretation of the identified causal effect. For instance, if locations taking-up treatment have on average high β, e.g., because a high β lowers the cost of treatment, then the IV procedure will tend to identify a high β; that is, the *local* effect of treatment for those locations in fact induced by the instrument to adopting treatment (the *complier* locations in the jargon of LATE studies; see later).

To push the empirical analysis of Persistence studies forward it is therefore important to study the relationship between the instrument z_{t-h} and treatment x_t and how this relationship can be filtered through the values of β and/or ρ. Similarly, it is important to study how the mechanism responsible for the persistence of treatment over time, from x_{t-h} to x_t, correlates with the values of β and/or ρ. Explicit models of these relationships and mechanisms, in the context of the specific empirical analysis, help clarifying the interpretation of the identified causal effects and help the formulation of interesting new sets of questions, which can be addressed empirically, possibly with new data. To this end, we develop below minimal abstract models linking treatment take-up in a location, and the treatment's persistence over history with the location's value of β and/or ρ.

Importantly, the mechanisms leading to treatment take-up and its persistence over history cannot generally be represented by behavioral choice problems because the units of observation in most of these studies are countries, cities, ethnic groups. Consequently, whether treatment is adopted or not, and whether it persists over history, rests on the outcome of political equilibrium processes or on the aggregation of the relevants behavioral choices of individuals. This is typically the case when treatment involves institutional change or change in characteristic cultural attitudes and traits. For the sake of simplicity and abstraction, our analysis will be limited to reduced-form models of treatment take-up, without being explicit about their equilibrium micro-foundations. Furthermore, since the context determines the ap-



(a) Current variables (b) Pure persistence

Figure 2: Persistence studies: empirical models.

Note: Circles indicate variables observed by the investigator. Dashed circles indicate unobserved variables. Solid arrows indicate directions of causality. Double arrows indicate endogeneity or any other factor preventing the identification of a causal effect (omitted variables, selection bias, etc...). Dashed arrows indicate potential causal links. Highlighted in red (gray in print version) are the relationships of interest.

propriate modeling assumptions, we will illustrate our approach in the context of several classic Persistence studies.

In the next two sections, we introduce two empirical models, special cases of the model introduced above, to highlight different strands of the literature.[8] In Section 3 we consider first the special case of the empirical model in Section 2.1 in which the econometrician aims at uncovering a causal relationship between current variables. The instrument in the historical past and the persistence of the independent variable are, in a sense, means to this end. The relationships between variables in this class of models are represented in Fig. 2(a). As an illustration of this empirical model it will be useful to consider two classic papers among the most-well known in the field: Acemoglu et al. (2001) on the colonial origins of economic development, and Ashraf and Galor (2013) on the effect of human genetic diversity on development. These papers provide distinct causal explanations of comparative economic development (affected by different dependent variables). Acemoglu et al. (2001) focus on the effect of the quality of institutions; specifically, the level of protection of property rights. Ashraf

[8]This categorization is related to the distinction between *apples-on-apples* and *apples-on-oranges* models by Voth (2021, in this book).

and Galor (2013) study how human genetic diversity affects economic development through the contrasting effect of social conflict and innovation or creativity.

Next, in Section 4 we consider the special case of the empirical model in Section 2.1 in which the econometrician aims at uncovering the relationship between the same variable at two distant times in history. The instrument in the historical past identifies the persistence of a relevant variable, which is, in a sense, the objective of the analysis; see Fig. 2(b). As an illustration of this type of models we will consider four other classic well-known papers providing evidence for the persistence of some form of either institutions or cultural traits over the long-run historical past: Voigtländer and Voth (2012), Nunn and Wantchekon (2011), Guiso et al. (2016), and Alesina et al. (2013), studying, respectively, the persistence of anti-semitism, social trust, civic capital, and gender attitudes.

3 Persistence studies that analyze relationships between current variables

Consider the special case of the empirical model of Section 2.1 in which the econometrician does not observe any proxy p_{t-h} for x_{t-h} and z_{t-h} is an instrument for x_t, through the persistence of the process $\{x_\tau\}_{\tau \in T}$. We assume that z_{t-h} is a valid instrument to concentrate our analysis on the interpretation of the estimates of β or ρ when they are allowed to be heterogeneous across locations $i = 1, 2, \ldots, N$. Besides requiring that z_{t-h} be correlated with x_t, validity also requires z_{t-h} to be as good as randomly assigned and to satisfy the exclusion restriction, that is, z_{t-h} affects y_t only through x_t.[9]

Assumption 1. *z_{t-h} is a valid instrument for x_t; in particular,*

$$cov\,(z_{t-h}, x_t) \neq 0$$
$$cov\,(z_{t-h}, y_t \mid x_t) = 0.$$

In Acemoglu et al. (2001), for instance, the process $\{x_\tau\}_{\tau \in T}$ represents the quality of institutions, from colonial times $t-h$ to the present t. The instrument z_{t-h} is settler's mortality: the authors argue that low settlers' mortality facilitated the set-up of inclusive institutions by colonial powers, whereas high settlers' mortality

[9]See Angrist and Pischke (2008), ch. 4.4.1 for formal assumptions.

Table 1: Variables in selected current relationship studies.

Article	y_t	x_{t-h}	z_{t-h}
AJR (2001)	GDP per capita, $t = 1995$	Property rights	Settlers' mortality
AG (2013)	Population density, $t = 1500$	Genetic diversity	Distance from Africa

Note: See Fig. 2(a) for relationships between variables. Abbreviations: AJR: Acemoglu et al. (2001), AG: Ashraf and Galor (2013).

caused extractive institutions. The parameter β represents the returns to institutional quality in terms of economic development, measured by current per capita GDP. In Ashraf and Galor (2013) the process $\{x_\tau\}_{\tau \in T}$ is human genetic diversity, from the first migration of *Homo sapiens* "Out of Africa" until the present. The instrument is distance of the current geographical location from East Africa, which is causal to human genetic diversity as a consequence of the *serial-founder effect*, whereby genetic diversity is reduced at any successive migration event. The parameter β represents the returns to human genetic diversity in terms of economic development, measured by population density in 1500. The underlying hypothesis, confirmed by their empirical analysis, is that diversity has a hump-shaped effect on development, because it has both beneficial and detrimental effects by increasing social conflict, but fostering innovation or creativity. Table 1 presents concisely the variables adopted in the main specifications of these studies.[10]

To simplify the intuition and to present results with minimal algebra (the argument generalizes without these assumptions), consider an environment where both the instrument z_{t-h}, and the treatment x_τ, for any $\tau \in T$, are binary. Under these assumptions, the instrumental variable estimator coincides with the Wald estimator,

[10]These examples illustrate well how heterogeneous treatment effects may play a role in the analysis. Acemoglu et al. (2001), e.g., interpret their measure of institutional quality a "cluster of institutions, including constraints on government expropriation, independent judiciary, property rights enforcement, and institutions providing equal access to education and ensuring civil liberties, that are important to encourage investment and growth. Expropriation risk is related to all these institutional features." This interpretation allows for different mechanisms connecting expropriation risk to economic performance, leading naturally to the possibility of heterogeneous effects, which depend on which mechanism is activated. This is the case in Ashraf and Galor (2013) as well, where genetic diversity, or lack thereof, affects development by a combination of creativity, increased cooperation, trust, socioeconomic order, adaptability, specialization, and so on.

whose population analog is:

$$\frac{E(y_t|z_{t-h}=1)-E(y_t|z_{t-h}=0)}{E(x_t|z_{t-h}=1)-E(x_t|z_{t-h}=0)}=\beta \tag{3}$$

This representation of the IV estimator has an intuitive interpretation: assuming the instrument is random with respect to the treatment, the effect of the instrument on the dependent variable (the *Reduced form*) divided by the effect of the instrument on the treatment (the *First stage*) uncovers the causal effect of a unit change of the independent variable on the treatment.

Allowing for β to be heterogeneous across locations $i = 1, 2, \dots, N$ opens the door for the possibility that the relationship between the instrument z_{t-h} and the independent variable, the treatment x_t, is filtered through the effect of β. Instrumented locations, with $z_{i,t-h} = 1$, might take treatment or not, $x_{i,t} = 1$ or $= 0$, depending on β_i. This is the fundamental factor inducing relevant LATE effects that are different from ATE effects. Importantly, note that this does not affect the validity of the instrument.[11] To formalize the result we want to focus on, assume β_i affects differently the relationship between $z_{i,t-h}$ and the treatment $x_{i,t}$ across locations $i = 1, 2 \dots, N$. Then we can categorize four conceptually different types of locations, which we label, following the causal identification literature, *Always takers, Never takers, Compliers, and Defiers*:

Always takers: all i such that $x_{i,t} = 1$ for all values of $z_{i,t-h}$;

Never takers: all i such that $x_{i,t} = 0$ for all values of $z_{i,t-h}$;

Compliers: all i such that $x_{i,t} = 1$ when $z_{i,t-h} = 1$ and $x_{i,t} = 0$ when $z_{i,t-h} = 0$;

Defiers: all i such that $x_{i,t} = 1$ when $z_{i,t-h} = 0$ and $x_{i,t} = 0$ when $z_{i,t-h} = 1$.

The instrumental variables estimator in Eq. (3) identifies the LATE effects of the instrument, that is, the effect of the instrument for the Compliers:

Theorem 1. *(Imbens and Angrist, 1994) Suppose Assumption 1 holds and that the*

[11] Validity is guaranteed if $x_t \mid_{z_{t-h}=1} \neq x_t \mid_{z_{t-h}=0}$ and $y_t \mid_{x_t, z_{t-h}=0} = y_t \mid_{x_t, z_{t-h}=1}$, where $x_t \mid_{z_{t-h}}$ and $y_t \mid_{x_t, z_{t-h}}$ indicate, respectively, the random variable x_t conditional to the realization of z_{t-h} and y_t conditional to the realization of x_t, z_{t-h}. Weaker conditions in terms of conditional means are sufficient.

treatment induces no Defiers, then

$$\begin{aligned} & \frac{E(y_t|z_{t-h}=1)-E(y_t|z_{t-h}=0)}{E(x_t|z_{t-h}=1)-E(x_t|z_{t-h}=0)} \\ = & E\left(y_{i,t}\mid_{x_{i,t}=1}-y_{i,t}\mid_{x_{i,t}=0} \mid i \in \textit{ compliers}\right) \\ = & E(\beta_i \mid i \in \textit{ compliers}) \end{aligned}$$

This theorem helps qualifying the interpretation of the estimate of β in the empirical implementation of Persistence studies when reasonable structures can be hypothesized which separate Compliers as a subset of all the locations and on whether in these structures Compliers can be characterized by a distinct distribution of β with respect to the other types.

3.1 Abstract models of treatment take-up and persistence

To identify Complier locations and their characteristics it is generally useful to model how the mechanisms which induces treatment take-up and persistence, so that $x_t = 1$, can be filtered through β. We illustrate this generally, by developing four minimal abstract models linking z_{t-h} to x_t via β. These models represent reduced-form behavioral-equilibrium relationships, as we already noticed, without explicit micro-foundations. It is our aim to show how this analysis is especially relevant in an historical context and how it may guide the development of empirical strategies to uncover the mechanisms underlying the identified causal effect. We associate these models narratively to the papers we have selected as an example in this section.

3.1.1 Treatment take-up at $t - h$

Let b_i denote the benefits of treatment for location i in historical times, $t - h$; that is, the benefits of $x_{i,t-h} = 1$ relative to $x_{i,t-h} = 0$. Assume b_i represents a relevant parameter which characterizes location i, but is not observable to the econometrician. Let $c(z_{t-h})$ denote the cost of treatment at $t-h$; and let the effect of the instrument be a reduction in the cost of treatment,

$$c(z_{i,t-h} = 1) < c(z_{i,t-h} = 0).$$

In Acemoglu et al. (2001) lower settler mortality, which we denote with $z_{t-h} = 1$, is assumed to lower the cost of creating and maintaining inclusive institutions that protect property rights in the colony. In Ashraf and Galor (2013) the distance from East Africa affects treatment because a shorter distance facilitates higher genetic diversity being associated with fewer migration events after the original one "Out of Africa."

Consider a behavioral-equilibrium relationship postulating treatment at $t-h$ to be determined by a simple cutoff condition: each location $i = 1, 2, \ldots, N$ takes-up treatment if its benefit is greater than the cost:

$$x_{i,t-h} = \begin{cases} 1 & \text{if } b_i \geq c(z_{i,t-h}) \\ 0 & \text{otherwise} \end{cases} .$$

To focus on the relationship between z_{t-h} and x_t, assume that the persistence of the explanatory variable is homogeneous (and perfect); that is, $\rho_i = 1$, for all $i = 1, 2 \ldots, N$, and

$$x_{i,\tau} = x_{i,\tau-1}, \quad t-h \leq \tau \leq t.$$

Treatment can only occur at $t-h$ and is never undermined in the course of history. Consequently, locations can be categorized depending on their characteristic b_i. *Always takers* are all those locations i such that $b_i \geq c(z_{i,t-h} = 0) > c(z_{i,t-h} = 1)$; and *Never takers* are all i such that $c(z_{i,t-h} = 0) > c(z_{i,t-h} = 1) > b_i$. Most importantly, the *Compliers* whose effect is identified by the IV are those locations whose benefits of treatment b_i are higher than the cut-off when treated by the instrument, but lower when not-treated:

$$c(z_{i,t-h} = 1) \leq b_i < c(z_{i,t-h} = 0).$$

Finally, there are no *defiers* (as required by Theorem 1).

According to this abstract behavioral-equilibrium model of treatment, for LATE effects to be distinct from ATE it is sufficient to hypothesize that β_i and b_i are correlated, i.e. that the benefits of treatment b_i are at least in part obtained through the effects of treatment on the relevant dependent variable $y_{i,t}$, β_i. In the context of the effects of institutional quality on economic development in Acemoglu et al. (2001), it seems natural to think of the returns of institutional quality in terms of economic development for a country i, β_i, as a measure of the benefits of treatment b_i for this country. In this context, this simple behavioral-equilibrium model of institutional formation can be used to explicit conditions, that may in principle be validated empirically, and refine the interpretation of the IV estimate of the returns of institutional quality by identifying the mechanisms leading different locations to take up treatment or not.

Consider the following somewhat extreme case as further illustration.[12] Assume for simplicity that there are two types of countries, l, h with $\beta_h > \beta_l$, and that their proportions are $\pi_h, \pi_l = 1 - \pi_h$ respectively. Further, assume that $c(z_{i,t-h})$ satisfies $c(z_{i,t-h} = 1) < \beta_l < c(z_{i,t-h} = 0) < \beta_h$. High settlers' mortality, $z_{t-h} = 0$, increases the cost for the colonial power to implement the inclusive institutions that reduce future expropriation risk. Under this assumption, this cost is generally lower than the benefit, with the exception of l countries when settlers' mortality is high, which represent then the *compliers*.[13] In this example, therefore, the Wald estimator identifies

$$\frac{E(y_t|z_{t-h} = 1) - E(y_t|z_{t-h} = 0)}{E(x_t|z_{t-h} = 1) - E(x_t|z_{t-h} = 0)} = \frac{\beta_h\pi_h + \beta_l\pi_l - \beta_h\pi_h}{1 - \pi_h} = \beta_l \tag{4}$$

i.e., the gains for countries with low values of β, the average value of β among the *compliers*:

$$E\left(\beta_i \mid i : c(z_{i,t-h} = 0) > b_i \geq c(z_{i,t-h} = 1)\right) = \beta_l.$$

Note however that, with a different cost structure, for example in the opposite extreme case where high-quality institutions are adopted only by h countries and only when their settlers' mortality is low, the instrumental variable analysis would on the contrary identify β_h. We conclude that this analysis of a simple behavioral-equilibrium model of treatment take-up, while extremely stylized, suggests the importance of evaluating various specific dimensions institutional change. It suggests notably the importance of proposing an empirical strategy that tries to test empirically the underlying assumptions regarding the benefits from institutional change to better understand what effects are identified. It also provides some nuance about the interpretations of the contemporaneous effects of institutional quality, because different "instruments" favoring improvements in institutional quality may operate on a different set of compliers.

Related arguments can be used to illustrate the possible role of heterogeneous

[12] We follow Rosenzweig and Wolpin (2000) who use a similar approach to interpret IV estimates in several labor economics studies; see the Appendix for an illustrative example

[13] All that is needed in general is for settlers' mortality to affect disproportionately one group of countries. Note that settlers' mortality only affects economic performance by changing institutional quality directly and therefore is by assumption a valid instrument.

treatment effect in the context of Ashraf and Galor (2013). The interesting difference, in this case, consists in the fact that the treatment (genetic diversity) is not directly the result of a political economy equilibrium choice, but is rather the (perhaps unintended) consequence of the migration process via the *serial founder effect*. This mechanism cannot be directly conceived as the outcome of an equilibrium model. On the other hand, genetic diversity can also be thought of as the consequence of evolutionary processes, for example driven by mating patterns in society. Assortative mating along any phenotypical trait, notably e.g., along ethnic dimensions, may reduce genetic diversity by producing ethnic cleavages which might turn into distinct homogeneous populations.

Consider then a behavioral-equilibrium model of mating patterns along migration events after the original one "Out of Africa" as an illustration. Let instrumented locations i, such that $z_{i,t-h} = 1$, be those which are at the economic development-maximizing level distance from East Africa.[14] These locations, by construction, are treated by the *serial founder effect*: their genetic diversity is $x_{i,t} = 1$; that is, genetic diversity is such that the combination of innovation and creativity, on the one side, and conflict and trust, on the other, induces the highest economic development. Suppose there are two types of populations defined by their ethnicities, culture, etc... Type h has a hump-shaped benefit from diversity, type l also has hump-shaped benefits, but with smaller effects. Let the benefits at the peak be denoted b_h and b_l, respectively. Suppose that mating strategies conducive to genetic diversity $x_t = 1$ are costly - say they have a cost c - and assume that $b_l < c < b_h$. Therefore,

$$x_{i,t-h} = \begin{cases} 1 & \text{if } z_{i,t-h} = 1; \text{ or if } z_{i,t-h} = 0 \text{ and } b_i \geq c \\ 0 & \text{otherwise} \end{cases}.$$

Consequently, the locations with benefits b_h are *Always takers*: they achieve "optimal diversity" regardless of location. Locations with returns b_l are instead *Compliers*; that is, they obtain the optimal diversity $x_{i,t} = 1$ only if moving "Out of Africa." As in Acemoglu et al. (2001), it seems natural to hypothesize that β_i and b_i are correlated, that the benefits of treatment b_i are at least in part obtained through the effects of treatment on the relevant dependent variable $y_{i,t}$, β_i. In this case, the IV strategy identifies the LATE effect, that is, the average β_i of the locations that

[14] This is just an innocuous change-of-variable re-normalization of the IV procedure.

have a different adoption rule when z_{t-h} changes value[15]:

$$E\left(\beta_i \mid i : b_i = b_l\right) = \beta_l.$$

Indeed, in the context Ashraf and Galor (2013), as in Acemoglu et al. (2001), conditions can be assumed (in this case on the costs and benefits of evolutionary selection through mating patterns) which give rise to distinct implications with respect to LATE effects. If for instance evolutionary selection towards the optimal diversity does not operate unless migration occurs, and if even with migration it occurs only when there are high incentives to do so, then it is the combined effect of migration and high incentives that generates development. In this case, the IV would identify β_h, even if diversity may have low or zero effects on development for some populations. We conclude that this simple behavioral-equilibrium model in the context of Ashraf and Galor (2013) suggests the importance of empirically evaluating different mating patterns, and how the structure of benefits (e.g., from isolating ethnically) depends on migration, to shed light on the mechanisms linking genetic diversity to economic development. While the stylized model we used for illustration abstracts

[15]The Wald arithmetic is identical to (4). In fact, the running hypothesis in Ashraf and Galor (2013), documented in the data, is more nuanced. Genetic diversity has a hump-shaped effect on economic development, reflecting a combination of beneficial and detrimental factors: diversity is associated with a higher likelihood of innovation and creativity, but also increases conflict and decreases trust. The simple model we have delineated can be extended to an environment in which the instrument and the treatment take three values for x_{t-h} and z_{t-h}, say $0, 1, 2$, representing respectively ordered distances from East Africa and measures of genetic diversity. Under the assumptions that locations with benefits b_h have the evolutionary incentive to converge to the genetic diversity which induces the highest economic development, independently of their distance from East Africa, and that b_i and β_i are correlated, the IV strategy identifies β_l. In this case, the benefits are identified by calculating the Wald estimator, comparing pairwise locations. First consider locations i with $z_{i,t-h} = 0, 1$:

$$\frac{E(y_t|z_{t-h}=1) - E(y_t|z_{t-h}=0)}{E(x_t|z_{t-h}=1) - E(x_t|z_{t-h}=0)} = \frac{(\pi_l\beta_l + \pi_h\beta_h) - (\pi_l 0 + \pi_h\beta_h)}{1-\pi_h} = \beta_l$$

Similarly, calculate the benefit comparing locations i with $z_{i,t-h} = 1, 2$:

$$\frac{E(y_t|z_{t-h}=2) - E(y_t|z_{t-h}=1)}{E(x_t|z_{t-h}=2) - E(x_t|z_{t-h}=1)} = \frac{(\pi_h\beta_h + \pi_l 0) - (\pi_l\beta_l + \pi_h\beta_h)}{\pi_h + 2\pi_l - 1} = -\beta_l$$

Suggesting that under the assumptions made on the cost structure, the strategy identifies the least pronounced hump-shaped relationship between diversity and development.

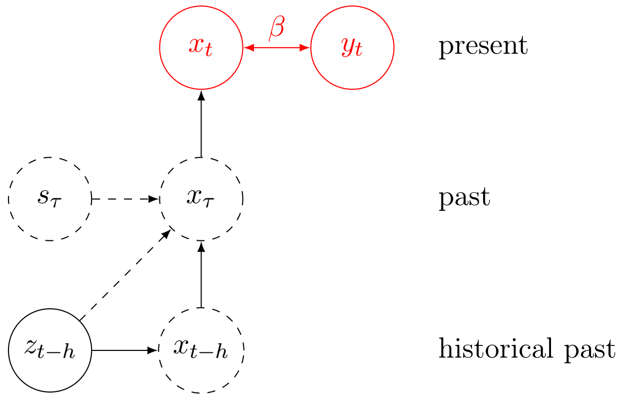


Figure 3: Current variables empirical model: interacting historical process.

Note: Circles indicate variables observed by the investigator. Dashed circles indicate unobserved variables. Solid arrows indicate directions of causality. Double arrows indicate endogeneity or any other factor preventing the identification of a causal effect (omitted variables, selection bias, etc...). Dashed arrows indicate potential causal links. Highlighted in red (gray in print version) is the relationships of interest.

from fertility, fertility patters would also relate to marriage and migration patters in fundamental ways affecting the interpretation of the identified causal effects of genetic diversity.

3.1.2 Treatment take-up at τ

In this section we leverage even more than in the previous one the particular role of history in the understanding of LATE effects. Consider the abstract behavioral-equilibrium models delineated in the previous section but assume now that locations can enter treatment at any time τ, with $t - h \leq \tau \leq t$, that is, after the realization of the instrument. One way to introduce the role of history in the determination of treatment take-up is to let the cost of treatment depend on the instrument z_{t-h} as well as on the realization of a stochastic process $\{s_\tau\}_{\tau \in T}$: $c_\tau = c(z_{t-h}, s_\tau)$. The empirical model, with the addition of the effects of the stochastic process $\{s_\tau\}_{\tau \in T}$ is represented in Fig. 3.

In the context of the effects of institutional quality on economic development, for instance, it is natural to assume the process $\{s_\tau\}_{\tau \in T}$ capturing the dynamics of relevant cultural variables interacting with the dynamics of institutions.[16] Importantly,

[16] Bisin and Verdier (2017) model these interactions in related contexts.

the validity of the instrument z_{t-h}, via the exclusion restrictions, is not hindered by the correlation of z_{t-h} with the cultural process $\{s_\tau\}_{t-h\leq\tau\leq t}$, as long as culture does not have a direct effect on economic development, that is, as long as s_τ affects y_t only through $\{x_\tau\}_{t-h\leq\tau\leq t}$. Let again the effect of the instrument be a reduction in the cost of treatment,

$$c(z_{i,t-h} = 1, s_{i,\tau}) < c(z_{i,t-h} = 0, s_{i,\tau}).$$

Maintain the assumption that the benefits of treatment across locations i, b_i, are correlated with β_i; i.e., that the returns of institutional quality in terms of economic development are a measure of the benefits of treatment. Suppose for simplicity a form of perfect persistence, where treatment is never reversed in the course of history,

$$x_{i,\tau} = 1,\ \tau \in T,\ \text{implies } x_{i,\tau'} = 1,\ \tau < \tau' \leq t.$$

Treatment is then still determined by a simple cutoff condition, for each location $i = 1, 2, \ldots, N$:

$$x_{i,\tau} = \begin{cases} 1 & b_i \geq c(z_{i,t-h}, s_{i,\tau}) \text{ or if } x_{i,\tau-1} = 1 \\ 0 & \text{otherwise} \end{cases}.$$

Treatment could occur at any $t-h < \tau \leq t$. The Compliers whose effect is identified by the IV are those locations whose benefits of treatment b_i are higher than the cut-off when treated by the instrument, but lower when not-treated; but whether this is the case might depend in general on the dynamics of s_τ. As in the previous case, since β_i and b_i are correlated, the IV strategy identifies the LATE effect, that is, the average β_i of the compliers. In this case, however, the instrument z_{t-h} interacts with the process $\{s_\tau\}_{\tau\in T}$. In general, b_i and the process $\{s_{i,\tau}\}_{\tau\in T}$ will be correlated. Interpreting β as a measure of the returns of institutional quality effectively disregards the effect of the dynamics of s_τ, which could represent, as we have already noted, cultural traits or social capital in the historical process of institutional change.[17] The IV strategy identifies in this case the returns to the institutional quality of the compliers that are activated by the historical processes. Different, counterfactual

[17]In fact, s_τ could also be thought of as a contributing causal factor, e.g., if a different realization of $s_{i,\tau}$ would have induced an instrumented location i not to set-up high quality institutions $x_{i,\tau} = 1$. In this case we could say that x_τ, s_τ are jointly causal to y_t.

histories may have activated processes with different returns.

To illustrate the role of the interaction between different processes in the course of history, consider the analysis of the economic development of the sample of countries colonized by European powers after 1500 in Acemoglu et al. (2002). This paper documents how i) colonial powers developed high-quality institutions disproportionally in initially poorer countries, (a "Reversal of Fortune"); ii) the inclusive institutions developed by colonial powers manifested their effects on economic development only after the Industrial Revolution in 1800-1900, and not before.

Consider the two following possible interpretations of these results.[18] One, assuming wealth in 1500 as exogenous with respect to economic development, is that historical poverty causes growth. Another interpretation, considering poverty in 1500 as an instrument for beneficial institutional change, is that inclusive institutions established by colonies cause economic growth. Both interpretations are valid in principle as long as they are qualified in terms of their effect being *local*. Poverty in 1500 acted *locally*, through institutional change in colonial times. But institutional change in colonial time acted *locally* on economic development through the Industrial Revolution.[19] Both of these *local* effects in principle have selected a subset of Compliers whose effect the empirical analysis identifies and which depend on both institutional quality and the Industrial Revolution. More specifically, the heterogeneity of the effects could be intrinsic to the quality of institutions as we argued in Section 3.1.1. But could also be due to the different nature of industrialization in time or place.[20] Even if the mechanism generating development from good institutions had homogeneous effect, heterogeneity could arise from different returns to industrialization.[21]

We conclude that this analysis suggests that the interpretation of the causal effects in Persistence studies depends in a fundamental manner not only from the historical process of the treatment variable but also from any other intervening his-

[18]See Cantoni and Yuchtman (2021, in this book) for a related argument.

[19]Following the logic exposed in Footnote 17, we could reasonably consider institutional quality and the Industrial Revolution as jointly causal.

[20]In fact, in Acemoglu et al. (2002), the identifying variation in the regression between post-industrialization production per capita and the interaction of institutions quality and opportunities to industrialize is the variation of U.K. industrial output over time, between 1750 and 1980.

[21]A similar discussion could pertain, for instance, to the negative relationship between past slave exports and economic performance within Africa, uncovered by Nunn (2008), if this effect appears when the dependent variable, economic performance, is measured in 1980, but not when measured in 1960, as suggested by Bottero and Wallace (2013). This would indicate the interaction of the effects of slave trade with a more recent phenomenon, like e.g., de-colonization.

torical process correlated with treatment. While generally difficult, this calls for the importance of historical narratives to identify possibly important intervening processes. This is exactly what Acemoglu et al. (2002) do in their study of the effects of colonization, isolating the Industrial Revolution as the main intervening factor in the process of development. The introduction of possible heterogeneous treatment effects adds a layer of complexity and interest, suggesting the importance to better identify the relationship between the returns to institutional change and industrialization and their interaction in the development process.

3.1.3 Treatment take-up and reversals

In this section we suggest the existence of circumstances in Persistence studies when one of the assumptions of Theorem 1 may not hold, namely the absence of Defiers, and study how this changes the interpretation of the estimates. In the previous sections we have developed minimal abstract behavioral-equilibrium models of Treatment take-up that consist of simple cutoff conditions: each location $i = 1, 2, \dots, N$ is treated if the benefits b_i are greater than the cost of adoption, which is reduced by the treatment and possibly an interacting process, as in the previous section. In these examples, as we observed, there are no Defiers: a Defier-location i would have to be characterized by benefits b_i such that $c(z_{i,t-h} = 0) \leq b_i < c(z_{i,t-h} = 1)$, which contradicts $c(z_{i,t-h} = 1) < c(z_{i,t-h} = 0)$.[22]

Consider as in the previous subsection the case where treatment can be adopted at $t-h \leq \tau < t$ and the cutoff cost are also affected by an interacting process $\{s_\tau\}_{\tau \in T}$. Suppose the instrument z_{t-h} - and possibly the treatment x_{t-h} - affect $\{s_{i,\tau}\}_{\tau \in T}$. In this case, interestingly, non-monotonic effects of z_{t-h} are possible, without impinging on the validity of z_{t-h} as an instrument for x_t. These relationships in the empirical model are represented in Fig. 4.

For instance, it is possible for the process $\{s_{i,\tau}\}_{\tau \in T}$ to act selectively on locations with $z_{i,t-h} = 0$, fostering high-quality institutions in a subset of these, against historical odds. E.g., in the context of Acemoglu et al. (2001), extractive institutions at $t - h$ could foster, by a mechanism of substitutability between culture and institutions, the development of cultural traits which over time lead to a reduction

[22]The existence of *Defiers* is ruled out by assumption in Theorem 1, a special case of the *Monotonicity* assumption in Imbens and Angrist (1994). To be precise, what is required is that Defiers be simultaneously present with Compliers: Defiers would be Compliers after reversing the definition of treatment.

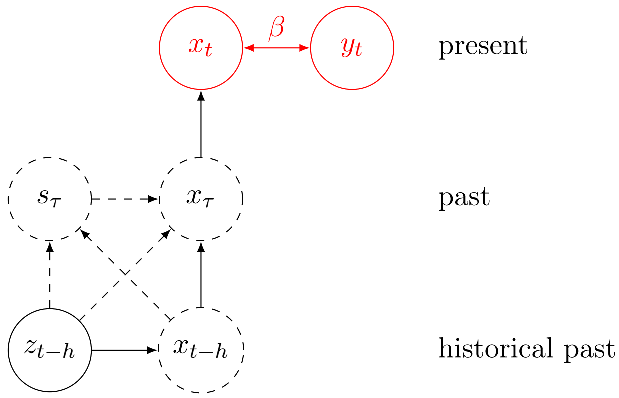


Figure 4: Current variables empirical model: non-linear interacting historical process.

Note: Circles indicate variables observed by the investigator. Dashed circles indicate unobserved variables. Solid arrows indicate directions of causality. Double arrows indicate endogeneity or any other factor preventing the identification of a causal effect (omitted variables, selection bias, etc...). Dashed arrows indicate potential causal links. Highlighted in red (gray in print version) are the relationships of interest.

of the cost of high-quality institutions and hence to treatment even location with relatively low benefits b_i.[23] In this case, $b_i < c(z_{i,\tau} = 1, s_{i,\tau})$ for any $\tau \in T$ but $b_i > c(z_{i,\tau} = 0, s_{i,\tau})$ for some $\tau \in T$. Keeping our assumption of correlation between β_i and b_i would induce the estimate $\hat{\beta}$ to weigh both Compliers and Defiers (but not Always-takers and Never-takers). Let π_c and π_d denote the proportion of Compliers and Defiers in the population. Then, the estimator identifies:

$$\frac{E(y_t|z_{t-h}=1) - E(y_t|z_{t-h}=0)}{E(x_t|z_{t-h}=1) - E(x_t|z_{t-h}=0)}$$

$$= E(\beta_i \mid i \in \text{ Compliers})\frac{\pi_c}{\pi_c - \pi_d} + E(\beta_i \mid i \in \text{ Defiers})\frac{-\pi_d}{\pi_c - \pi_d}. \tag{5}$$

Expression (5) is a weighted average of the treatment effects, but note that one of the two weights must be negative, therefore the procedure estimates not a return, but a net effect which is difficult to interpret.[24]

[23]See Bisin and Verdier (2017) for a formal discussion of substitutability between culture and institutions.

[24]See Heckman et al. (2006) for a discussion of the implications of violations to the monotonicity assumption. Heckman and Vytlacil (2005) show in a general framework that the IV identifies a weighted average of the treatment effects. In the absence of Monotonicity the weights may be negative. Eq. (5) characterizes this average in the context of our simple model.

Examples of the possible role of Defiers in historical contexts typically include cases of historical reversals in development. Consider Acemoglu et al. (2002) once again, for instance. Relative wealth in 1500 might have fostered future growth, in and of itself, but relative poverty as well, through institutional change in colonial times and the Industrial revolution. In this interpretation, countries which have maintained growth from 1500 would be Compliers, while the protagonists of the *reversal of fortunes* would be Defiers. Another interesting case of reversal is documented in Ashraf et al. (2010), when it is shown that more isolated locations in the Paleolithic are more developed today. It is argued that isolation was more conducive to innovation, through a mechanism combining less free-riding on inventions elsewhere and lower predisposition to being invaded. More generally, we conclude that the interpretation of causal effects in Persistence studies requires a careful analysis of the distribution of treatment take-up over historical time, once again suggesting an important role for historical narratives to guide the formal econometric IV strategy.

3.1.4 Persistence of treatment

In this subsection we entertain the analysis of the role of heterogeneous treatment effects operating through the persistence mechanisms itself rather than directly through treatment take-up as in the previous sections. Consider a behavioral-equilibrium model in which treatment at $t-h$ is perfectly determined by the instrument, $x_{t-h} = z_{t-h}$, but persistence is heterogeneous and depends on the value of x_τ: for instance treatment is more persistent than lack of treatment. Continuing our parallel with colonial origins in Acemoglu et al. (2001), this could occur for example because inclusive institutions, for example, are less costly to maintain once established.[25] In the context of (our interpretation of) Ashraf and Galor (2013), where the persistence of human genetic diversity is the result of the composition of mating patterns and the serial founder effect, heterogeneity of persistence could be the consequence of mating patterns depending on the number of migration events, e.g., because of the resulting distribution of phenotypical diversity, e.g., ethnic fractionalization.

A simple formalization of heterogeneous treatment effects operating through the persistence mechanisms is obtained if the transition matrix $\Pr(x_\tau \mid x_{\tau-1})$ depends on β. This could induce different compliance across values of β_i. To illustrate,

[25] See Przeworski (2004) for evidence regarding the heterogeneity of institutional persistence over historical times.

Table 2: Heterogeneous persistence with binary x: $\Pr(x_t|x_{t-h})$.

	x_t	β_1: 0	β_1: 1	β_2: 0	β_2: 1
x_{t-h}	0	p_1	$1-p_1$	p_2	$1-p_2$
	1	$1-q_1$	q_1	$1-q_2$	q_2

assume β_i only takes two values, $\beta_1 < \beta_2$, respectively in fractions $\pi_1, \pi_2 = 1-\pi_1$ of locations, and the correlation between x_t, x_{t-h} with $\Pr(x_t|x_{t-h})$ is as in Table 2.

The population analog of the Wald estimator is:

$$\frac{E(y_t|x_{t-h}=1)-E(y_t|x_{t-h}=0)}{E(x_t|x_{t-h}=1)-E(x_t|x_{t-h}=0)} =$$

$$\frac{(\beta_1\pi_1 q_1+\beta_2\pi_2 q_2)-(\beta_1\pi_1(1-p_1)+\beta_2\pi_2(1-p_2))}{\pi_1 q_1+\pi_2 q_2-(\pi_1(1-p_1)+\pi_2(1-p_2))} =$$

$$\frac{\sum_{i=1,2}\pi_i\beta_i(p_i+q_i-1)}{\sum_{i=1,2}\pi_i(p_i+q_i-1)}.$$

It follows that, as long as $p_1+p_2 \neq q_1+q_2$, the LATE effect is different from the ATE. In particular, consider the (natural) case in which if the persistence of the treatment is correlated with the returns of the treatment β, that is, $\Pr(x_\tau = 1 \mid x_{\tau-1} = 1)$ increases with β. In the example of the table this would be the case if, e.g., $q_2 > q_1$ and $p_2 = p_1$. In this case, the Wald estimator would identify a LATE effect β greater than the ATE effect, $\sum_{i=1,2}\pi_i\beta_i$.[26] We conclude that a formal behavioral-equilibrium model of persistence suggests dimensions and directions of inquiry useful to better qualify the mechanisms driving the causal relationships identified in Persistence studies that rely on the persistence of the treatment effect in history.

4 Pure persistence studies

In this section we consider the special case of the empirical model in Section 2.1 in which the objective of the analysis is to prove the persistence over time of a variable x_τ. In these studies, the econometrician observes a proxy p_{t-h} for x_{t-h}, and, possibly, z_{t-h}, an instrument for p_{t-h}. This empirical model is illustrated in Fig. 2(b) which is reproduced here for convenience, as Fig. 5. We assume that z_{t-h} is a valid

[26] If $p_1 = p_2 = 1/2$ the LATE effect would be β_2.

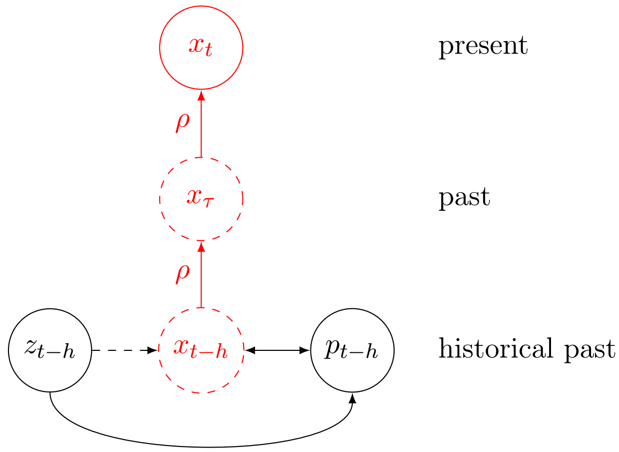


Figure 5: Pure persistence empirical model.

Note: Circles indicate variables observed by the investigator. Dashed circles indicate unobserved variables. Solid arrows indicate directions of causality. Double arrows indicate endogeneity or any other factor preventing the identification of a causal effect (omitted variables, selection bias, etc...). Dashed arrows indicate potential causal links. Highlighted in red (gray in print version) is the relationships of interest.

instrument for p_{t-h} to concentrate our analysis on the interpretation of the estimates of ρ when they are allowed to be heterogeneous across locations $i = 1, 2, \ldots, N$. Besides requiring that z_{t-h} be correlated with p_{t-h}, validity also requires z_{t-h} to be as good as randomly assigned and to satisfy the exclusion restriction, that is, z_{t-h} affects x_t only through p_{t-h}.

Assumption 2. *z_{t-h} is a valid instrument for p_{t-h}; in particular,*

$$cov\left(z_{t-h}, p_{t-h}\right) \neq 0$$
$$cov\left(z_{t-h}, x_t \mid p_{t-h}\right) = 0.$$

The examples of empirical studies that represent this model and we use as an illustration are listed in Table 3. All of these papers are interested in demonstrating the persistence in history of important factors related to current socio-economic performance: anti-semitism (Voigtländer and Voth (2012)), the role of women in society Alesina et al. (2013), civic trust (Nunn and Wantchekon (2011)), and social capital Guiso et al. (2016).

While it is possible to find or collect detailed measures of these variables in

Table 3: Variables in selected Pure persistence studies.

Article	x_t	p_{t-h}	z_{t-h}
VV (2012)	Anti-semitism	Pogroms in 1349	-
AGN (2013)	Gender attitudes	Plow use	Plow suitability
NW (2011)	Trust	Slave trade	Distance from sea
GSZ (2016)	Civic capital	City-state	Bishop city

Note: See Fig. 2(b) for relationships between variables. Abbreviations: VV: Voigtländer and Voth (2012), AGN: Alesina et al. (2013), NW: Nunn and Wantchekon (2011), GSZ: Guiso et al. (2016).

current times, the values of these variables in the past can only be measured with proxies. In addition, in some cases, the proxies are endogenous to the process and an historical instrument is introduced to disentangle the causal effect.

Consider for simplicity (but the arguments extend) an environment where both z_{t-h}, and p_{t-h}, are binary and assume that z_{t-h} is a valid instrument for p_{t-h}.[27] The Instrumental Variable estimator coincides with a Wald estimator:

$$\frac{E(x_t|z_{t-h}=1)-E(x_t|z_{t-h}=0)}{E(p_{t-h}|z_{t-h}=1)-E(p_{t-h}|z_{t-h}=0)}=\rho. \tag{6}$$

Analogously to the cases in Section 3, we distinguish:

Always takers: All i such that $p_{i,t-h}=1$ for all values of $z_{i,t-h}$;

Never takers: All i such that $p_{i,t-h}=0$ for all values of $z_{i,t-h}$;

Compliers: All i such that $p_{i,t-h}=1$ when $z_{i,t-h}=1$ and $p_{i,t-h}=0$ when $z_{i,t-h}=0$;

Defiers: All i such that $p_{i,t-h}=1$ when $z_{i,t-h}=0$ and $p_{i,t-h}=0$ when $z_{i,t-h}=1$.

A version of Theorem 1 can be stated in this case.

[27] The validity of the instrument assumption is formally guaranteed by the following assumption:

$$E\left(p_{t-h} \mid z_{t-h}=1\right) \neq E\left(p_{t-h} \mid z_{t-h}=0\right) \qquad \text{First stage}$$
$$x_t \mid_{p_t, z_{t-h}=0} = x_t \mid_{p_t, z_{t-h}=1} \qquad \text{Exclusion restrictions.}$$

Theorem 2. *(Imbens and Angrist, 1994) Suppose that Assumption 2 holds and that the treatment induces no Defiers, then*

$$\begin{aligned} &\frac{E(x_t|z_{t-h}=1)-E(x_t|z_{t-h}=0)}{E(p_{t-h}|z_{t-h}=1)-E(p_{t-h}|z_{t-h}=0)} \\ &= E\left(x_{i,t}\mid_{p_{i,t-h}=1} - x_{i,t}\mid_{p_{i,t-h}=0} \mid i \in \ \textit{compliers}\right) \\ &= E(\rho_i \mid i \in \ \textit{compliers}) \end{aligned}$$

The estimated parameter is generally different from $E(\rho_i)$ when persistence ρ_i is heterogeneous across locations. We analyze two special cases.

4.1 Abstract models of persistence of treatment

In this section we follow the analysis of Section 3.1.4, studying in some detail the role of heterogeneous treatment effects operating through persistence mechanism. We develop behavioral-equilibrium models in which the persistence of the treatment at $t-h$ is heterogeneous. We distinguish environments with and without a valid observed instrument z_{t-h}.

4.1.1 No instrument

Consider first the simplest case in which the econometrician does not observe any instrument in historical time z_{t-h}. Being interested in documenting persistence ρ in the process $\{x_\tau\}_{\tau\in T}$, the econometrician only observes x_t and a proxy p_{t-h} for x_{t-h}. Suppose for simplicity the underlying process for x_t takes the form

$$x_t = \rho x_{t-h} + \epsilon,$$

where ϵ is a random shock.[28] Voigtländer and Voth (2012)'s investigation of the historical persistence of anti-semitism in German cities serves well as an illustration of this type of Persistence studies. Lacking detailed evidence of anti-semitism in history, this article uses pogroms in 1348-50CE in a cross-section of cities as an imperfect proxy for it. The presumption is that the Black Death of 1348-50CE lowered the threshold for violence which resulted in Jews being blamed and in some cases being mass-executed in pogroms as a result. The study finds that pogroms in

[28] Recall that $\rho x_{t-h} = [\rho_i x_{i,t-h}]_{i=1}^N$.

1348-50CE are indeed positively correlated with various detailed measures of anti-semitism in the 20th century in the cross-section of cities (vote shares for the Nazi Party, number of deportees from each city, anti-semitic letters to newspapers, etc. . .). In this context, the shock ϵ captures all determinants of anti-semitism not operating through its cultural persistence.

Consider a behavioral-equilibrium cut-off model of the relationship between proxy p_{t-h} and treatment x_{t-h}:

$$p_{t-h} = \begin{cases} 1 & \text{if } x_{t-h} > \widetilde{x} \\ 0 & \text{if } x_{t-h} \leq \widetilde{x} \end{cases}; \tag{7}$$

that is, activation of the proxy is associated to treatment being sufficiently high, even though the relationship is not necessarily causal. In the context of Voigtländer and Voth (2012), indeed following their logic, pogroms $p_{t-h} = 1$ occur in cities with high anti-semitism x_{t-h}. But it is not excluded that pogroms might have induced a reinforcement of anti-semitic attitudes. Lacking data on x_{t-h} it is only possible to estimate

$$x_t = \alpha^R + \rho^R p_{t-h}, \tag{8}$$

which identifies

$$\begin{aligned} \rho^R &= E(x_t|p_{t-h} = 1) - E(x_t|p_{t-h} = 0) \\ &= \rho \cdot (E(x_{t-h}|x_{t-h} > \widetilde{x}) - E(x_{t-h}|x_{t-h} \leq \widetilde{x}))\,; \end{aligned} \tag{9}$$

that is, the true persistence parameter ρ times the difference in treatment between locations with $p_{t-h} = 1$ and $p_{t-h} = 0$. This kind of empirical analysis therefore reaches its objective of identifying the presence of long-run persistence: a positive estimate of ρ^R is obtained only if $\rho > 0$. As long as pogroms are positively correlated with anti-semitism x_{t-h}, using the pogroms variable produces an estimate of ρ^R that converges, in limit probability, to a positive number. The estimated ρ^R can be interpreted as the difference between treated and non-treated locations on the average level of x_t which is due to the persistence effect of x_{t-h}.

This interpretation of the estimate of ρ^R can be refined if we allow persistence to be heterogeneous; that is, if we allow ρ_i to differ across locations $i = 1, 2, \ldots, N$. As in the examples from the previous section, heterogeneity may occur because persistence is somehow correlated with the occurrence of pogroms, or because some

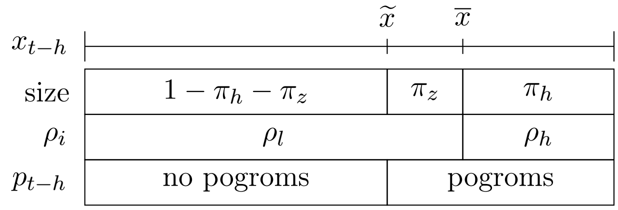


Figure 6: A model of Voigtländer and Voth (2012) with heterogeneous persistence.

other variable s_τ affects persistence itself over time. Allowing for heterogeneity in ρ_i across locations, estimating (8) identifies

$$\rho^R = E(\rho_i x_{t-h} | x_{t-h} > \widetilde{x}) - E(\rho_i x_{t-h} | x_{t-h} \leq \widetilde{x}). \tag{10}$$

An abstract behavioral-equilibrium model of the inter-generational transmission of cultural traits is helpful in interpreting (10) to illustrate the implications of heterogeneous persistence. Suppose cultural traits survive when they are sufficiently strongly held in the population; more specifically, e.g., a fraction of locations i, π_h with $x_{t-h} > \overline{x}$ have a high $\rho_i = \rho_h$ so that the trait persists more easily over time than in the remaining $(1-\pi_h)$ cities, with persistence $\rho_l < \rho_h$ (see Fig. 6).[29] In this case (10) reduces to

$$\rho^R = \rho_l E(\widetilde{x} < x_{t-h} \leq \overline{x}) + \rho_h E(x_{t-h} > \overline{x}) - \rho_l E(x_{t-h} < \widetilde{x}) \tag{11}$$

It follows then that the identified effect of persistence ρ^R can still be interpreted as a difference between treated and non-treated cities of the average level of x_t due to the persistence of x_{t-h}. On the other hand, the size of the effect depends on the distribution of x_{t-h} and on the relationship between x_{t-h} and persistence ρ across locations. For instance, assuming instead $\widetilde{x} > \overline{x}$, we obtain

$$\rho^R = \rho_h E(x_{t-h} > \widetilde{x}) - (\rho_l E(x_{t-h} < \overline{x}) + \rho_h E(\overline{x} < x_{t-h} \leq \widetilde{x}))$$

Under different hypotheses, which, e.g., in Voigtländer and Voth (2012)'s context can ultimately be reduced to how high anti-semitism in the XIVth century needed to be to trigger a pogrom, the size of the identified effect changes. Pogroms identify

[29]See Bisin and Verdier (2001, 2011) for models of inter-generational cultural transmission with implications along these lines.

the persistence of anti-semitism in cities where they took place, not the persistence in cities where they did not occur, because they did not occur in a random sample of cities.[30] We conclude that while it might be difficult to distinguish a priori between the different process we have hypothesized drive persistence of anti-semitism in relations to pogroms, it would be interesting in principle and with additional historical evidence we may be able to support one case over the other to improve our understanding of the nature of persistence. More generally, the simple behavioral-equilibrium model indicates novel and interesting empirical directions to pursue the analysis of the persistence of cultural traits and values.

4.1.2 Instrumenting for persistence

Consider now the case in which the econometrician does observe an instrument in historical time z_{t-h} for p_{t-h}. That is, consider the case where z_{t-h} affects causally x_{t-h} and p_{t-h}, while x_{t-h} and p_{t-h} are in principle linked by two-way causation, as in the empirical model in Fig. 2(b). The econometrician, interested in measuring the process' persistence ρ, observes x_t and a proxy p_{t-h} for x_{t-h} (but not x_{t-h} directly).

Three examples from the literature clarify this empirical model. Nunn and Wantchekon (2011) study the persistence of social trust, an important determinant of current development, using data from African regions. Modern surveys, such as the Afrobarometer, provide measures of trust in the present. The historical proxy is the size of slave trade, which varies by location and ethnicity, and may have affected local trust in the past by generating "an environment of ubiquitous insecurity caused individuals to turn on others". Reverse causality may also hold, however, because communities with lower trust may experience a lower cost to "kidnap, trick or sell each other to slave"). To address this endogeneity the geographical distance from the coast is used as an instrument for the size of trade. Distance from the coast correlates with the size of slave trade and, arguably, does not affect directly current levels of trust. The study finds that the level of social trust can be traced back to the slave trade.

[30]This model's assumptions can be mapped into the assumptions we used to build the model in Subsection 3.1.4. In that model, the persistent variable is binary, but the correlation between variables over time is more general than the correlations that are possible imposing the structure of Fig. 6. For example, in the model in this subsection, assume there is an underlying continuous latent variable χ_{t-h} that determines both p_{t-h} and $x_{t-h} \in \{0, 1\}$ according to different thresholds. Then, knowing the distribution of χ_{t-h} one can derive the conditional probabilities that are analog to $\{p_i, q_i\}, i = 1, 2$ in Table 2 and compute what magnitudes the Wald estimator identifies.

In Guiso et al. (2016) the process $\{x_\tau\}_{\tau \in T}$ represents civic capital in Italian cities, measured, in modern times, with indicators such as the proportion of student cheating in mathematics tests, the prevalence of non-profit associations, or measures of blood donations. The proxy for this variable in the past is whether the city experienced self-government in the middle ages, which arguably generated a culture of cooperation and trust that persists until current times. Indeed, the evidence shows that cities that experienced self-government have better measures of current social capital. Because this correlation may be due to omitted endogenous factors, the paper supports these results by showing that the correlation remains when instrumenting self-government with the presence of a bishop seat in the city before 1400 CE. The instrument's validity is motivated by the argument (advanced by several historians) that bishops facilitated the adoption of self-government by morally sanctioning the citizens' agreement. The presence of bishops lowers the coordination cost required to achieve independence without affecting modern social capital directly. The study finds that indeed, self-government in the late middle ages determines higher civic capital today.

In our last example of papers that conform to this framework, Alesina et al. (2013), trace the origin of current cross-cultural differences regarding the role of women in society (measured by labor-force participation, political representation, and women entrepreneurship, x_t) to agricultural practices in pre-industrial periods, proxied by plow cultivation (p_{t-h}). The process $\{x_\tau\}_{\tau \in T}$ represents gender attitudes and the proxy p_{t-h} is the adoption of plow cultivation in pre-industrial times. The argument is that practices requiring more physical strength encouraged specialization of production by gender, affecting the perception of gender roles, which persisted in current times even if modern market production does not require a physical strength advantage. The adoption of plow agriculture is therefore a proxy for the perception of gender roles, but could be affected by it, because societies that believe women should be confined to home production are more prone to adopt agricultural practices that comparatively advantage men. The study addresses this endogeneity using the suitability of locations for plow cultivation as an instrument. Plow suitability disproportionally induces the adoption of the plow, arguably without affecting the perception of gender roles.

In all these examples, because a valid instrument is observed for the adopted proxy, the empirical model leads to the identification of

$$\frac{E(x_t|z_{t-h}=1)-E(x_t|z_{t-h}=0)}{E(p_{t-h}|z_{t-h}=1)-E(p_{t-h}|z_{t-h}=0)}=\rho. \tag{12}$$

Following the logic of our discussion of Voigtländer and Voth (2012), an instrument is not necessary to identify a positive ρ. As long as $p_{i,t-h}=1$ and $x_{i,t-h}$ are correlated, the estimate ρ will have the correct sign. The endogeneity of the adoption of the plow in Alesina et al. (2013), for instance, does not invalidate the fact that locations which adopted the plow in $t-h$ are characterized by more conservative gender attitudes at t: locations i with $p_{i,t-h}=1$ have higher $x_{i,t}$. These attitudes in the present are the consequence of the persistence of cultural traits and attitudes. The same is true for the identification of the persistence of trust in Nunn and Wantchekon (2011) and of civic capital in Guiso et al. (2016), that we argued have the same structural relationship between variables.

However, the IV estimate of ρ can be interpreted as a LATE estimate of the persistence of the process $\{x_\tau\}_{\tau\in T}$; that is, a *local* effect through the specific instrument adopted. For instance, the IV in Nunn and Wantchekon (2011), identifies the persistence of community trust. Assume the benefits of trust b_i are heterogeneous across ethnic groups, for simplicity assume two levels $b_h > b_l$. Trust is then the consequence of a behavioral-equilibrium outcome: adopted when its benefits are greater than the benefits from slave trade. Let the benefit of slave trade be denoted by $d(z_{t-h})$ and denote with $z_{t-h}=1$ high distance from the coast, with $d(1)<d(0)$. Assume $b_l<d(1)<b_h<d(0)$ so that the ethnic groups with low value of trust $b_i=b_l$ always experience slave trade, whereas groups with $b_i=b_h$ experience trade only if they are close to the coast. In this case distance from coast induces a change in treatment only for ethnic groups with the highest returns to trust, those where, presumably, incentives for trust to persist are higher.[31] This would imply that the IV parameter identifies the persistence of trust in locations where trust is more beneficial and consequently where its persistence is higher.

Similarly, consider Guiso et al. (2016), and assume cities have different returns to civic capital, and persistence is higher when returns are higher. Further, assume that cities with high returns adopt self-government in the middle ages regardless of

[31] For example assume that the cost of acquiring trust is random, but on average higher than the average cost of maintaining it once acquired. Then ethnic groups with higher benefits will display higher persistence of trust.

the triggering effect of a coordinating factor (the presence of a bishop, z_{t-h}), whereas cities with $b_i = b_l$ adopt self-government only through the coordination from a bishop (if $z_t = 1$). In this case Compliers are cities with a low return, which arguably are less likely to experience persistence otherwise. In this case, the instrument identifies persistence in cities with relatively low returns and consequently low persistence.

Finally, in Alesina et al. (2013) the instrument identifies the persistence of gender attitudes as affected by the exogenous adoption of the plow. It is conceivable that some locations are Always Takers, that they have adopted the plow motivated by their previous gender attitudes - independently of whether their land were suitable for plow agriculture. If these locations are characterized by stronger persistence of these attitudes, the instrument identifies persistence in location with relatively lower persistence.

5 Conclusions

In this chapter, we have argued that it is often natural for heterogeneous effects to occur in the context of Persistence studies: if different mechanisms affect the variable of interest with different intensity over locations over time, then different combinations of mechanisms that aggregate into the same value of the treatment variable generate heterogeneous effects.

Behavioral-equilibrium models can help the interpretation of the causal effects uncovered in Persistence studies when treatment effects are heterogeneous and the LATE parameter identified in quasi-experimental designs differs from the Average Treatment Effect. In this context, we have shown, even minimal models of treatment take-up may shed some light on what information the estimated effects provide, at least conditionally on assumptions that sometimes can be empirically tested. Adding structure, as recent and current research on the dynamics of institution and culture is doing, can only be of help; see Acemoglu et al. (2021, in this book); Bisin and Verdier (2021, in this book); Persson and Tabellini (2021, in this book). Importantly, these arguments are independent of the soundness of econometric techniques and of data reliability.

The interpretation of the estimated parameters is important when Persistence studies have policy implications. If multiple slow-moving mechanisms (with heterogeneous effects) underlie long-run correlations, then policies motivated by the estimated parameters, which are not necessarily the average treatment effect, may produce unexpected outcomes if they do not operate through the channels that are *locally* identified by the research design.

The correct interpretation of the estimated parameters is important not only for policy implications, but also when the researcher is interested in counterfactual historical analysis. When treatment effects are heterogeneous, the instrument operates by identifying a specific historical path whose effects may be different from those that may have been generated in a counterfactual.

Pushing the arguments of this chapter forward, this line of analysis sheds doubts on the notion of *origin*, often used in these studies. A causal relationship realized in history at time τ with effects at time t does not preclude another historical causal relationship realized prior to τ.[32] In many cases the relevant question is not what was the origin of a phenomenon, but what are the counterfactual *quantitative* effect that results from changing variables in the historical past on variables in the present. Each variable at different points in the past may affect the present differently. This paper highlights how the empirical implications of Persistence studies may be interpreted using simple models that inform how these mechanisms affect the variable of interest suggesting directions for future research.

LATE for school

The clearest applications of the distinction between LATE and ATE are in labor/education economics. In this Appendix we exploit Rosenzweig and Wolpin (2000)'s comment on Angrist and Krueger (1991)'s instrumental variable approach to estimating returns to schooling.[33]

Consider the relationship between schooling attainment and earnings. A simple regression fails to identify the causal effect of schooling because an omitted variable bias: for example, children of higher ability may earn higher wages for given schooling

[32] For instance, the origin of the Mafia in Sicily has been reduced with good arguments to the rise of socialist Peasant Fasci organizations at the end of the 19th century (Acemoglu et al., 2020); to a price shock on sulphur and lemon in the 1850's (Buonanno et al., 2015; Dimico et al., 2017); to the lack of city states in the XIV'th century - in turn a consequence of Norman domination (Guiso et al., 2016); to the Paleolithic split into nomadic pastoralism in 7th millenium B.C. (Alinei, 2007).

[33] Underlying this analysis is the debate in econometrics regarding identification power in reduced-form causal inference design: see Angrist and Krueger (1991); Imbens and Angrist (1994); Angrist and Imbens (1999); Deaton (2010, 2020); Heckman (1997, 1999); Heckman and Urzua (2010); Imbens (2010).

and also choose a higher schooling attainment. A valid instrument for schooling may be adopted to identify the causal effect. Angrist and Krueger (1991) propose to use Quarter of birth. The birth date cutoffs for school-entry age combined with minimum compulsory schooling ages induce some children born during the last months of the year to complete more years of schooling relative to children born at the beginning of the year, because they are induced to start schooling at an earlier age (this is true for example if they intend to leave school at the mandatory minimum age). The arguably random variation in date of birth provides a "natural" instrument for estimating the return to schooling.[34]

Consider the case in which the treatment effect is heterogeneous, that is, returns to schooling vary in the cross-section of students; for instance, they increase with the underlying unobservable ability. To illustrate the difference between LATE and ATE, limit the students' choice to one extra year of schooling after mandatory schooling age, and assume that high ability students attain an extra year of education regardless of their Quarter of birth, but low-ability students always intend to drop out of school at mandatory age, and can only be "forced" to undertake an extra year of education if, by being born late in the year, they start attending school one year younger.

The instrumental variable technique "works" by randomly inducing a subset of students, low-ability students, to undertake an extra year of education. The extra earnings gained by students not born in the first quarter, relative to the earnings of those born during the rest of the year, are generated only by low-ability students that attend an extra year of education. All other students have the same schooling attainment regardless of the date of birth, therefore the instrumental variable estimate only identifies the returns for low-ability students (a LATE effect), not the return of high-ability students, nor the average return of all students (the ATE).

While the estimated parameter remains of great interest, this interpretation is consequential in that it suggests that the specific IV procedure adopted, while valid, under-estimates the returns to schooling. Furthermore, it suggests caution when considering the policy implications of the estimates, because a different policy with the goal of inducing higher schooling attainment may not induce the same subset of students to comply.

[34]In fact, the validity of quarter of birth as an instrument has been questioned, see Buckles and Hungerman (2013), but for the purposes of this introduction, we assume the instrument to be valid from an econometric standpoint.

This stylized model highlights how the Quarter of birth instrument is likely to identify the returns of the low-ability students, because arguably these students are the most likely to leave schooling at the minimum compulsory age. It is both a logically reasonable assumption, and an empirically testable implication that could provide additional evidence about the interpretation of the estimated returns. It is in this spirit that we proceed, in this paper, to formalize stylized models of persistence studies.

References

Acemoglu, Daron, Giuseppe De Feo, and Giacomo Davide De Luca (2020) “Weak states: Causes and consequences of the Sicilian Mafia,” *The Review of Economic Studies*, Vol. 87, No. 2, pp. 537–581.

Acemoglu, Daron, Georgy Egorov, and Konstantis Sonin (2021) “Institutional Change and Institutional Persistence,” in Bisin, Alberto and Giovanni Federico eds. *Handbook of Historical Economics*, Amsterdam: Elsevier North Holland, Chap. 13.

Acemoglu, Daron, Simon Johnson, and James A Robinson (2001) “The colonial origins of comparative development: An empirical investigation,” *American Economic Review*, Vol. 91, No. 5, pp. 1369–1401.

——— (2002) “Reversal of fortune: Geography and institutions in the making of the modern world income distribution,” *The Quarterly journal of economics*, Vol. 117, No. 4, pp. 1231–1294.

Alesina, Alberto, Paola Giuliano, and Nathan Nunn (2013) “On the origins of gender roles: Women and the plough,” *The Quarterly Journal of Economics*, Vol. 128, No. 2, pp. 469–530.

Alinei, Mario (2007) “Origini pastorali e italiche della camorra, della mafia e della’ndrangheta: un esperimento di Archeologia Etimologica,” *Origini pastorali e italiche della camorra, della mafia e della’ndrangheta*, pp. 1000–1039.

Angrist, Joshua D and Guido W Imbens (1999) “Comment on James J. Heckman," Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations",” *Journal of Human Resources*, pp. 823–827.

Angrist, Joshua D and Alan B Krueger (1991) “Does compulsory school attendance affect schooling and earnings?” *The Quarterly Journal of Economics*, Vol. 106, No. 4, pp. 979–1014.

Angrist, Joshua D and Jörn-Steffen Pischke (2008) *Mostly harmless econometrics: An empiricist’s companion*: Princeton university press.

——— (2014) *Mastering’metrics: The path from cause to effect*: Princeton University Press.

Ashraf, Quamrul and Oded Galor (2013) "The 'Out of Africa' hypothesis, human genetic diversity, and comparative economic development," *American Economic Review*, Vol. 103, No. 1, pp. 1–46.

Ashraf, Quamrul, Ömer Özak, and Oded Galor (2010) "Isolation and development," *Journal of the European Economic Association*, Vol. 8, No. 2-3, pp. 401–412.

Bisin, Alberto and Thierry Verdier (2001) "The economics of cultural transmission and the dynamics of preferences," *Journal of Economic theory*, Vol. 97, No. 2, pp. 298–319.

——— (2011) "The economics of cultural transmission and socialization," in Benhabib, Jess, Alberto Bisin, and Matthew O. Jackson eds. *Handbook of social economics*, Vol. 1: Elsevier, pp. 339–416.

——— (2017) "On the joint evolution of culture and institutions," Working Paper 23375, National Bureau of Economic Research.

——— (2021) "Phase Diagrams in Historical Economics: Culture and Institutions," in Bisin, Alberto and Giovanni Federico eds. *Handbook of Historical Economics*, Amsterdam: Elsevier North Holland, Chap. 16.

Bottero, Margherita and Björn Wallace (2013) "Is There a Long-Term Effect of Africa's Slave Trades?" Economic History Working Paper 30, Bank of Italy.

Buckles, Kasey S. and Daniel M. Hungerman (2013) "Season of Birth and Later Outcomes: Old Questions, New Answers," *The Review of Economics and Statistics*, Vol. 95, No. 3, pp. 711–724, URL: https://ideas.repec.org/a/tpr/restat/v95y2013i3p711-724.html.

Buonanno, Paolo, Ruben Durante, Giovanni Prarolo, and Paolo Vanin (2015) "Poor institutions, rich mines: Resource curse in the origins of the sicilian mafia," *The Economic Journal*, Vol. 125, No. 586, pp. F175–F202.

Cantoni, Davide and Noam Yuchtman (2021) "Historical Contingencies, Econometric Problems: The Analysis of Natural Experiments in Economic History," working paper, National Bureau of Economic Research.

Casey, Gregory and Marc Klemp (2021) "Historical instruments and contemporary endogenous regressors," *Journal of Development Economics*, Vol.

149, p. 102586, URL: http://www.sciencedirect.com/science/article/pii/S0304387820301619, DOI: https://doi.org/10.1016/j.jdeveco.2020.102586.

Cioni, M, G Federico, and M Vasta (2021) "The two Revolutions in Economic History," in Bisin, Alberto and Giovanni Federico eds. *Handbook of Historical Economics*, Amsterdam: Elsevier North Holland, Chap. 1.

Deaton, Angus (2010) "Instruments, randomization, and learning about development," *Journal of economic literature*, Vol. 48, No. 2, pp. 424–55.

——— (2020) "Randomization in the Tropics Revisited: a Theme and Eleven Variations," Working Paper 27600, National Bureau of Economic Research.

Dimico, Arcangelo, Alessia Isopi, and Ola Olsson (2017) "Origins of the sicilian mafia: The market for lemons," *The Journal of Economic History*, Vol. 77, No. 4, pp. 1083–1115.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales (2016) "Long-term Persistence," *Journal of the European Economic Association*, Vol. 14, No. 6, pp. 1401–1436, DOI: 10.1111/jeea.12177.

Heckman, James (1997) "Instrumental variables: A study of implicit behavioral assumptions used in making program evaluations," *Journal of human resources*, pp. 441–462.

Heckman, James J (1999) "Instrumental variables: response to Angrist and Imbens," *The Journal of Human Resources*, Vol. 34, No. 4, pp. 828–837.

Heckman, James J, Lance J Lochner, and Petra E Todd (2006) "Earnings functions, rates of return and treatment effects: The Mincer equation and beyond," *Handbook of the Economics of Education*, Vol. 1, pp. 307–458.

Heckman, James J, Daniel Schmierer, and Sergio Urzua (2010) "Testing the correlated random coefficient model," *Journal of econometrics*, Vol. 158, No. 2, pp. 177–203.

Heckman, James J and Sergio Urzua (2010) "Comparing IV with structural models: What simple IV can and cannot identify," *Journal of Econometrics*, Vol. 156, No. 1, pp. 27–37.

Heckman, James J and Edward Vytlacil (2005) "Structural equations, treatment effects, and econometric policy evaluation 1," *Econometrica*, Vol. 73, No. 3, pp. 669–738.

Imbens, Guido (2019) "Potential Outcome and Directed Acyclic Graph Approaches to Causality: Relevance for Empirical Practice in Economics," Working Paper 26104, National Bureau of Economic Research.

Imbens, Guido W (2010) "Better LATE than nothing: Some comments on Deaton (2009) and Heckman and Urzua (2009)," *Journal of Economic literature*, Vol. 48, No. 2, pp. 399–423.

Imbens, Guido W. and Joshua D. Angrist (1994) "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, Vol. 62, No. 2, pp. 467–475, URL: http://www.jstor.org/stable/2951620.

Kelly, Morgan (2019) "The standard errors of persistence," Working Paper DP13783, CEPR.

Nunn, Nathan (2008) "The long-term effects of Africa's slave trades," *The Quarterly Journal of Economics*, Vol. 123, No. 1, pp. 139–176.

Nunn, Nathan and Leonard Wantchekon (2011) "The slave trade and the origins of mistrust in Africa," *American Economic Review*, Vol. 101, No. 7, pp. 3221–52.

Pearl, Judea (2009) *Causality*: Cambridge university press.

Pearl, Judea, Madelyn Glymour, and Nicholas P Jewell (2016) *Causal inference in statistics: A primer*: John Wiley & Sons.

Persson, Torsten and Guido Tabellini (2021) "Culture, Institutions and Policy," in Bisin, Alberto and Giovanni Federico eds. *Handbook of Historical Economics*, Amsterdam: Elsevier North Holland, Chap. 15.

Przeworski, Adam (2004) "The last instance: Are institutions the primary cause of economic development?" *Archives Européennes de Sociologie/European Journal of Sociology/Europäisches Archiv für Soziologie*, pp. 165–188.

Rosenzweig, Mark R and Kenneth I Wolpin (2000) "Natural "natural experiments" in economics," *Journal of Economic Literature*, Vol. 38, No. 4, pp. 827–874.

Voigtländer, Nico and Hans-Joachim Voth (2012) "Persecution Perpetuated: The Medieval Origins of Anti-Semitic Violence in Nazi Germany*," *The Quarterly Journal of Economics*, Vol. 127, No. 3, pp. 1339–1392, URL: http://dx.doi.org/10.1093/qje/qjs019, DOI: 10.1093/qje/qjs019.

Voth, Joaquim (2021) "Persistence - Myth and Mystery," in Bisin, Alberto and Giovanni Federico eds. *Handbook of Historical Economics*, Amsterdam: Elsevier North Holland, Chap. 7.