Endogenous institutions: the importance of history

-Introduction. -I. Where do formal institutions come from?. -II. Where does culture come from?. -III. The interaction between formal and informal institutions. -Conclusions. References.

Introduction

Today few economists would dispute that institutions are the primary determinants of long-run economic performance. Most economists also broadly agree on the main mechanisms through which institutions affect economic growth. Three mechanisms figure prominently in the literature. First, institutions are considered the ultimate determinants of the costs of transaction (the cost of both detailing the conditions of a market exchange and enforcing them). Second, institutions are considered the means through which a society overcomes collective actions problems. And third, institutions are considered to define the incentives that ultimately determine whether the bulk of a society's resources are devoted to productive rather than to redistributive activities (Olson, 1982, North, 1990 and Fogel, 1997).

The growing awareness concerning the importance of institutions has made the quest for the origins of institutions and for the causes of institutional change an important area of research. This essay attempts an overview of the economic literature concerned with institutional change. Two different types of institutions are considered. The first type includes all formal rules
enforced by the state. That is, political and judicial rules, economic regulations and contracts as well as their respective mechanisms of enforcement. The second type includes social norms, codes of conduct, and conventions that are not enforced by the state but are regularly complied with.

The reader should take notice that the previous classification leaves out all mechanisms of governance—the private arrangements devised to either economize on transaction costs, or facilitate commerce or cope with coordination and cooperation problems. So neither the firm (as conceived by Coase, 1937), nor the medieval trade guilds (as conceived by Greif, 1993) nor the community enforcement mechanisms described by Ellickson (1986) are considered here. This omission is meant to incorporate North’s crucial distinction between institutions (the rules of the game) and organizations (the actions of the players).

I. Where do formal institutions come from?

As mentioned above, formal institutions comprise the laws and regulations enforced by the state. The importance of formal institutions is apparent from the fact that no developed economy lacks a sophisticated judicial system and a well-functioning central bureaucracy. The question in order is of course: where do formal institutions come from? Or, alternatively, how do institutions originate and evolve?

Two different approaches to answering these questions can be found in the literature. According to the first approach (that we may call the rational choice approach), if we are to understand how formal institutions arise, we merely have to specify the objective function of the lawmakers and the constraints they face when trying to maximize it. According to the second approach (that we may call the evolutionary approach), if we are to understand institutional change, we must pay close attention to dynamic considerations and, in particular, we must recognize the role played by initial conditions and historical accidents. The methodological tension is evident: whereas according to the rational choice approach history doesn’t matter, according to the evolutionary approach one can hardly account for the emergence of formal institutions by the type of armchair reasoning predicated by the rational choice approach.
To fix ideas, an example contrasting the two approaches may be useful. Consider the case of France and Britain in the eighteen century—a case often cited as a clear example of institutional divergence. The difference between British and French institutions at the eve of the eighteen century was dramatic, specially regarding the crown’s scope to abuse its power: while British institutions severely limited the crown discretion to both seize property and limit personal liberties, French institutions fostered an absolutist monarchy that respected neither property rights nor personal freedom (North, 1990 and Landes, 1998).

What accounts for this difference? Two distinct explanations have been advanced in the literature. First, Nye (1997), in the spirit of the rational choice approach, argues that the more despotic and inefficient French institutions were optimal responses to both France’s larger size (which hindered the crown ability to collect taxes) and France’s greater risk of foreign encroachment. Second, North and Weingast (1989), in the spirit of the evolutionary approach, argue that the more efficient British institutions stemmed from the unique balance of power brought about by the Glorious Revolution of 1688. These authors argue, in particular, that the institutional transformation of Britain that followed the Glorious Revolution was possible by the victory of the Parliament over the army of Charles I and the later victory of the Whig faction within the Parliament.¹ The key point here is the importance of contingency: Charles I could have prevailed, the Whigs could have lost and thus Britain could have ended up with a different institutional framework altogether (Fogel, 1997).

These two approaches shouldn’t be seen as irreconcilable: they both play a role in the origin and transformation of formal institutions. Consider, for example, the well-known model of political exchange in a representative democracy. In this model, the electorate (the principal) delegates to the legislators (the agents) the crucial task of designing a “social contract.”

¹ According to North (1990, p.139), the main reforms implemented following the Glorious Revolution were: “parliamentary supremacy, parliamentary control of financial matters, curtailment of crown prerogative powers, independence of the judiciary (at lest from the crown), and supremacy of common law courts.”
While we can study the interaction between legislators and their constituents using the standard tools of rational choice theory, we also have to recognize that this interaction is plagued with constraints inherited from history (i.e., legislators barely start from scratch). Bluntly, historical considerations should be an integral part of any attempt to understand the creation and modification of rules in representative democracies.

Needless to say, history affects political interactions through many channels. For example, new policies and regulations often trigger the formation of strong constituencies that perceive huge gains from their maintenance (Dixit, 1996). These constituencies will get organized and will strongly oppose any attempt to modify the policies and even other groups that will gain from a change will oppose it in fear that it may lead to further changes that ultimately will damage their interests. In short, the endogenous formation of interest groups introduces inertia to the political process. Examples abound: rent controls that benefit existing tenants, price controls that benefit existing farmers, trade restrictions that benefit existing domestic producers, and zoning laws that benefit existing landowners.

Of course, history limits institutional change in many other different ways. Bureaucrats and organizations often invest heavily on policy-specific capital, and hence will also oppose any change that will depreciate their investments (Buchanan, 1988). Policies can also modify people’s preferences and preferences may in turn affect policies (see Section IV). Policies can also have long-lasting effects by affecting which party will win in the future (Aghion and Bolton, 1994) or by changing the preferences of the incumbents (Besley and Coate, 1998). Again, history is a force to be reckoned with in the study of institutional change.²

The previous discussion suggests that well-intentioned legislators and bureaucrats face long odds when attempting to subvert old institutional arrangements. There are, however, narrow windows of opportunity through which effective reforms can be pushed forward. Exogenous shocks, for

² The importance of history is evident from the fact that many ex-colonies have retained their original civil and commerce codes in the face of superior alternatives (La Porta et al., 1996).
example, may sometimes speed up reforms. Similarly, dramatic episodes of institutional malfunctioning may help break history's tight grip. According to Stein et al. (1998) long-needed changes in Argentina and Peru's economic institutions became viable only after hyperinflation proved beyond reasonable doubt the huge cost of sticking with the old institutions.

II. Where does culture come from?

Values, attitudes and codes of conduct often override plain self-interest in human behavior. Most people obey the law even when punishment is virtually impossible and many citizens vote and participate in politics contradicting the precepts of rational choice theorists. Bluntly, self-interest and opportunism do not account for the behavior of individuals in markets and other social environments.

If we move from the level of the individual to the level of the community, we often find large differences in the extent and type of values. For example, the extent of individuals who never free ride or who obey the law as a matter of principle is very different from country to another. These differences are often considered as key to understand the differences in economic performance between nations (North, 1990 and Putnam, 1993, Landes, 1997).

Values and codes of conduct seem to matter a great deal. They, however, cannot be changed at will, a point often forgotten by historians and political commentators (Landes, 1998 and Apuleyo et al., 1994). Again, culture is not a choice variable that can be subject to normative considerations: a society chooses neither its values nor its beliefs.

The question in order is: where does culture come from? There is no shortage of hypothesis about the origin and evolution of values, norms and beliefs. A few of the most important of these hypotheses are summarized below.

---

3 Landes (1998), for example, approaches cultural change from a purely normative perspective. He argues vehemently that "If we learn anything from the history of economic development, it is that culture makes all the difference." But he offers no clue whatever as to what determines culture. His advise to backward countries seems simple enough: adopt the values of successful countries. A similar approach is used by Apuleyo et al. (1994) in "A Handbook for the Perfect Latin American Idiot." They also bring culture to the fore and they also candidly proposes a "switching-values" strategy.

Lecturas de Economía No. 54. Medellín, enero - junio 2001
Sociologists often claim that values and beliefs will ultimately reflect the main precepts that govern social interactions. Putnam (1993), for example, argues that trust and civic participation are habits formed during a long history of "horizontal associations." Putnam then goes on to argue that trust can be nurtured, as it were, by the repeated participation in formal and informal groups: church associations, trade unions, political parties, and so on. Putnam also argues that both hierarchical religions and authoritarian governments discourage the formation of trust by imposing a vertical structure on the society. The empirical evidence in this regard is mixed: whereas Knack and Keefer (1997) do not find any relationship between group membership and trust for a sample of 29 countries, La Porta et al. (1997) uncover a negative link between trust and hierarchical religions and authoritarian governments.

In the same vein, the importance of "habituation" in the formation of values and attitudes has often been emphasized by sociologists studying the "erosion of morals" in many urban areas of the United States. Wilson (1996) argues, for example, that "the more often certain behavior such as the pursuit of illegal income is manifested in a community, the greater will be the readiness of the community to find that behavior not only convenient but also morally appropriate." Although the empirical evidence in this respect is still unimpressive, some empirical studies lend credence to the idea that values are ultimately shaped by social interactions at the level of the community (see, for example, Case and Katz, 1991).

In the discussion above, values are unintended by-products of community interactions. But values can also be deliberately shaped; or, at least, this seems to be the rationale behind many social organizations that devote considerable time and effort to the inculcation of civic values. The point here is that differences in culture among regions may be partially accounted for by differences in the efficacy and prevalence of their values-molding institutions. Although the evidence in this regard remains scant, the large investments (and long speeches) devoted to the inculcation of good values suggest a widespread belief in the receptivity of such efforts.

In clear contrast to the social emphasis of the previous hypotheses, there have been various attempts to explain cultural differences on the
basis of differences in natural environments. Kaplan (1998) argues, for example, that collectivist societies are more common in arid regions where the need for water-sharing arrangements (and, more recently, for public goods to store and distribute water) is more urgent. On the contrary, individualism tends to flourish where water abounds.\(^4\) Similarly, coffee growing communities in Latin America have long been regarded as reservoirs of social capital, which perhaps reflects both the social organization of coffee production (there are no large scale economies so small holdings are common) and the importance of public goods in the production and marketing of green coffee (water, again, is essential here). External environments affect values in more subtle ways. There is, for example, evidence showing a positive connection between homeownership and civic behavior (Dipasquale and Glaeser, 1997). There is also evidence showing that a negative connection between ethnic dispersion and income inequality, on the one hand, and social trust and civic participation, on the other (Knack and Keefer, 1996).

The previous ideas notwithstanding, many social scientists remain skeptical about any attempt to change people’s morals, values, or beliefs. The question is: should policy move beyond the simple predicament of getting the incentives right? We can argue that the theories are still too fuzzy and the facts still too controversial for political activism. But we must also recognize that the potential payoffs are huge—For example, Cohen (1995) estimates that the value to society of saving a high-risk youth may be as high as $2 million.

III. The interaction between formal and informal institutions

So far we have discussed formal and informal institutions as independent entities. This is hardly the case, however. The evolution of formal institutions affects and is affected by the evolution of values, beliefs and norms. In short, values and rules co-evolve. A variety

---

\(^4\) Ostrom (1990) summarizes the diversity of social mechanisms that have been devised to solve the “commons” problem of water provision.
mechanisms through which rules affect values and vice versa have been postulated in the literature (see, for example, North, 1990, Posner, 1997, and Bowles, 1998). This section discusses a few of such mechanisms.

North (1990) sketches an important mechanism through which values and formal institutions interact. North's model combines the model of representative democracy described in Section II with a bounded-rationality argument. According to North, people don't have a complete understanding of the implications of their choices (political or otherwise) and so often rely on ideological and moral judgements when making decisions. (In Latin America, for example, the widespread support of protectionist policies during the 50s and 60s had a clear ideological component, partially originated in the teachings of the Economic Commission for Latin America and its emphasis on dependency theories). In North's model, the political process is assumed to be capable of articulating people's views and political decisions (protectionist views gave rise to protectionist policies in Latin America). The key insight of the model is that people's views, however wrong, would not always be corrected as a result of informational feedbacks (in Latin America, protectionist policies deepened the dependency on primary products, thus reinforcing not only the existing institutional framework but also people's views about the need for protection). Thus, people's perceptions "not only rationalize the society's structure but accounts for its poor performance as well" (North, 1990, p. 99). Of course, breaking this dynamic of mutual reinforcement is not easy: it takes either a big exogenous shock or decades of stagnation.

Rules and values interact in many other forms. Law scholars, for example, often argue that the law can play an important role as a shaper (Posner, 1997). According to this view, a strict enforcement of the law will ultimately raise voluntary compliance, which will in turn facilitate enforcement. This mechanism of mutual reinforcement suggests the intriguing possibility that small differences in enforcement can have huge downstream consequences. Although the empirical evidence in this regard is scant, the recent decline of crime rates in various urban areas in the United States is often presented as evidence that harsh enforcement increases voluntary compliance of the law (Bernstein, 1998). Much
uncertainty remains, however, as to the precise dynamics of this mechanism. One will wish to know, in particular, whether voluntary compliance responds quickly to changes in enforcement (as the proponents of the so-called “Broken Windows” hypothesis seems to argue) or whether it really takes a lengthy period of strict enforcement to change well-rooted criminal attitudes. Needless to say, policy implications will hinge heavily on the answers to these lingering questions.

Likewise, the degradation of values brought about by government assistance programs lies at the core of conservative attacks to the welfare state (Becker, 1995, p. 26). The point of such criticisms is simple: government assistance begets apathy and erodes self-reliance and initiative, which in turn begets a greater need for government assistance. Although direct empirical evidence of this mechanism has yet to be produced, it played a key role in the recent reform to the welfare system in the United States. A similar story underlies an often-heard explanation for the skyrocketing of divorce rates in the United States during the last two decades (Difonzo, 1997). The point is also simple: increasing social acceptability of divorce prompted the passing of non-fault divorce laws which in turn made divorce more prevalent which in turn raised the social acceptability of divorce and so on ad infinitum. The evidence in this regard is incomplete: while the effect of non-fault divorce laws on the divorce rates seems to be significant (Friedberg, 1998), the effects of increasing divorce rates on values and of values on legislation have yet to be tested (although few will deny that the stigma attach to divorce has decreased substantially in the last two decades).

Values and formal institutions may also interact in subtle and unexpected ways. Consider, for example, the following story suggested by the empirical findings of Hoxby (1998) and Poterba (1998). These authors have uncovered compelling evidence showing signs of intergenerational conflict in the United States. They show, in particular, that support

---

5 This idea has been neatly formalized by Lindneck (1995).
6 Similar stories can be told about the interaction between gay-rights laws and anti-gay sentiments, and civil rights and racial discrimination.
among the elderly for public education has decreased significantly in recent years. Why? One can argue, and Poterba (1998) himself speculates along these lines, that the decline of extended families, partially brought about by the increasing availability of Social Security and Medicare, may have eroded intergenerational altruism, which in turn may have eroded the willingness of the elderly to support social transfers to the young. This intergenerational conflict was arguably less important in the past when the trade-off between elderly care and education was resolved within the family.7

Two major conclusions can be drawn from the previous discussion. First, the coevolution of rules and values often takes the form of spirals of mutual reinforcement that are difficult to reverse. This may explain why changes in formal rules are often inconsequential in the short run; the new rules may well replace the old rules but they can hardly override old values, moral and attitudes. Second, the effects caused by the implementation of new formal rules are difficult to predict because these rules affect informal institutions in unexpected ways.

Conclusions

There is a common thread throughout this essay: history matters in the process of institutional change. The mechanisms whereby history operates are various and usually interlocked. Some of these mechanisms are driven by dynamics of mutual reinforcement between formal and informal institutions while others are driven by dynamics that are specific to each type of institution. Rules affect values and values affect rules, but also today’s rules shape future’s rules and today’s values shape future vales. Many options are possible, as the variety of mechanisms sketched above make it clear.

This is not to say that history accounts for everything and that external factors are feeble. The common disjunctive between chance and

---

7 Stiglitz (1994) has often made the same point in. In his words, "by changing the locus of caring and responsibility from the family to the government, we change society and we change ourselves." The logic is a little bit perverse in that "attempts to improve society by having the government undertake a greater role in redistribution, may ultimately--through their effects on individuals and the nature of the social contract--have more ambiguous consequences" (Stiglitz *ibid.*).
necessity (history versus geography, to put it somewhat bluntly) is not the right one in this realm. Usually, external factors dictate some initial conditions that are in turn greatly reinforced by the workings of history. Krugman (1998) puts it perfectly, "aspects of natural geography are able to matter so much not because natural features of the landscape are that crucial, but because they establish the seed around which self-reinforcing agglomerations crystallize." In short, necessity provides the original script and chance modifies along the way.

Some general policy statements can be derived from the previous discussion. First, we shouldn't take political institutions and political practices as given. The relevant question is not what kind institutions promote growth, but rather what kind of policies can create growth-enhancing institutions. Second, while deliberate attempts to change informal institutions are difficult to justify (the theoretical foundations are shaky and the ethical questions are troubling; there is a fine line indeed between inculcating good values and brainwashing), the impact of formal institutions on values, norms and perceptions shouldn't be ignored in policy discussions. And last, making miracles is out of the question here. Institutional change is not only incremental, but largely autonomous.

References


