IN THIS ISSUE

ARTICLES

Civil Society as a Complex Adaptive Phenomenon ........ 3
Mikayla Novak

Reconsidering Urban Spontaneity and Flexibility after Jane Jacobs: How do they work under different kinds of planning conditions? ............. 14
Stefano Cozzolino

Cities, Agriculture, and Economic Development: The Debate over Jane Jacobs’s ‘Cities-First Thesis’ .... 25
Sanford Ikeda

Detecting and Directing Emergent Urban Systems: a Multi-Scale Approach ......................... 32
Elena Porqueddu

Connecting the Dots: Hayek, Darwin, and Ecology ....... 51
Gus diZerega

The Role of Spontaneous Order in Video Games: A Case Study of Destiny ..................... 63
William Gordon Miller

REVIEWS

The Infidel and the Professor: David Hume, Adam Smith, and the Friendship that Shaped Modern Thought by Dennis C. Rasmussen ......................... 73
Gordon Graham

Legislating Instability: Adam Smith, Free Banking, and the Financial Crisis of 1772 by Tyler Beck Goodspeed ..................... 78
Pablo Paniagua

Pablo Paniagua

Envisioning Real Utopias by Erik Olin Wright ......... 94
David Ellerman

Cognitive Autonomy and Methodological Individualism: The Interpretative Foundations of Social Life by Francesco Di Iorio .................. 104
Gabriele Ciampini

Editorial Information ........................................ 109

http://cosmosandtaxis.org
INTRODUCTION

As observed throughout the course of human history and in all the places human beings inhabit, individuals incessantly interact with others in the pursuit of benefits and achievements. A broad-scale conception of the domain through which varying guises of human association takes place is coined “civil society,” and the objective of this paper is to provide an account of its nature and salient features.

Any attempt to provide a definitive account of civil society surely poses as an intellectually tall order. This reflects recognition that “civil society is a … necessarily contested idea” and is “not a concept that yields to easy consensus, conclusion, or generalization” (Edwards 2011, 480). Ambiguities surrounding the definition of civil society reflect scholarly disputes over the appropriate scope of human activity to be encompassed within the civil society notion, as well as contestation surrounding the normative propriety of claims made about its efficacy in explaining the trends and tendencies of actual societies. Difficulties in grasping the idea of civil society may also be attributable to shifts and permutations in the kinds of cooperative human activities undertaken therein.

Inspired by developments in complexity and evolutionary science literatures, especially their applications in social scientific contexts, I depict civil society as a complex and adaptive phenomenon. To be more specific, civil society is that combinatorial sphere of cooperative relations between individuals and groups of people spanning the economic, communal and political orders, with each order consisting of both spontaneous and non-spontaneous elements (diZerega 2014; Madison 1997).

The explanation of civil society as a structured process (embodying complexity) unfolding in novel, and often unforeseen, ways (reflecting adaptation) yields several important insights. This approach invites an inclusive, non-reductionist approach toward examining the kinds of relationships and interdependencies forged when humans act to procure gains from each other. It also draws attention to the claim economic, communal and political situations within civil society are, themselves, the manifestation of decentralised actions by the many, and not the intentional design by the few. Finally, understanding civil society as a complex, adaptive phenomenon helps us to understand and appreciate the implications of observed changes that both surround us, and in which we participate.
I now turn to the structure of the paper. First, I draw upon a range of contemporary social scientific insights to depict some generic properties of civil society as a complex, adaptive phenomenon. This is followed by an analysis of a multi-phase processes in which change is realised through the economic, communal and political orders of civil society. Entrepreneurial action sets out to discover novel modes of conduct, the entrepreneurial insight is then competitively tested against other insights for prospective support, and the most popular insights potentially become embedded in the institutional frameworks of civil society as enduring rules. Next, I draw upon the dynamic picture of civil society explicated in this paper to critically assess the claim that observed changes necessarily illustrate that civil society is in “decline.” I finish by making concluding remarks.

CIVIL SOCIETY AS A COMPLEX MEDLEY OF SPONTANEOUS ORDERS

Theorising about the nature of civil society has assumed a prominent position within the classical liberal philosophical tradition since its very inception, arguably during the seventeenth century.

Contrary to Hobbes’s suggestion that the “state of nature” of human interaction is characterised by endlessly violent frictions, necessitating nothing less than the absolute sovereignty exercised by political actors to quell it, John Locke considered that people generally tended to be naturally cooperative and peace-loving in their natural state.

Lockean harmony in the state of nature did not suggest, however, that conflict and discord would be non-existent. Therefore, Locke hypothesised that a “political society” could be legitimately instituted for the relatively limited purpose of instituting laws protecting properties justly acquired and held by individuals. For Locke, the idea of civil society was coterminous with this political order mutually agreed upon, as if culminating in the form of a “social contract,” by the members of society.

Witnessing the embryonic phase of the British Industrial Revolution, the Scottish Enlightenment figure Adam Smith suggested that voluntary exchange relationships taking place within markets served as the centrepiece of civil society. The trading of commodities between sellers and buyers, according to Smith, echoed a certain inclination within human nature “to truck, barter, and exchange one thing for another.”

For Smith the sheer scale and scope of market activities, reflecting an increasing division of labour, provided a lucrative opportunity to enhance living standards, even for those in economically impoverished positions. Furthermore, in sentiments reminiscent of the doux-commerce thesis of France’s Montesquieu, the Scotsman conceived that commercial development “contributes greatly to the extirpation of prejudice and parochialism and the cultivation of a shared sense of humanity. Commerce polishes away the rough edges of human nature and leads to the cultivation of civility and manners” (Boyd 2013, p. 444).

The nineteenth-century French aristocrat Alexis de Tocqueville studied the social and political conditions underpinning the formative American democracy, and in so doing formulated yet another version of civil society theory. Intermediate associations, both local and voluntary in their character, were elevated as the pivotal component of civil society. Tocqueville noted that diverse associations and social groupings encouraged Americans to devise their own solutions to problems, rather than depending upon government to do so purportedly on their behalf. Residing between atomistic individuals and the overbearing state, the constellation of free and intermediate associations described by Tocqueville undergirded a vibrant civil society.

The inclusion of each, or all, of the aforementioned forms of human activity in a definition of civil society has long elicited intellectual dispute, with arguments over which of these forms of interaction should be excluded. By contrast I consider there is merit in pursuing the alternative of a more inclusive stance concerning which relationships should be classified as being represented in civil society. This would enable us to more openly recognise different, and in fact the oft-intertwined, ways in which humans associate and mingle with each other, and the opportunities and challenges for peace and prosperity arising from such interdependences.

G. B. Madison described it best when he called for a “more extensive conception of civil society” to inform analysis of human association, in which the notion of civil society itself is not “an entity intermediate between the “family” and the “state” … but as society organized in a particular way” (Madison 1997, p. 36). Within this expansive, and unashamedly pluralistic, definition, Madison submitted that “[c]ivil society is composed of distinct yet overlapping orders, these orders being nothing other than the sedimented results of human agency in different spheres of life” (Ibid., p. 37).

I conceive civil society as generically composed of three nested and ordered systems in which humans interact to derive advantages and benefits conducive to their senses
of prosperity and wellbeing: the economic order; the communal order; and the political order. The three orders mentioned here can be construed as spontaneous orders, a dynamic pattern of mutually adaptive relationships which arises from a broad-scale agreement about rules rather than results:

[s]pontaneous orders are discovery processes structured by abstract procedural rules through which individual plans responding to local knowledge and personal insight are brought into greater coherence with similar plans by an infinite number of other individuals mutually unknown to each other (diZerega 2008, p. 1).

The observed regularities arising through spontaneous orders come about as individuals, and groups of individuals, follow particular rules of conduct which are independent of any specific or particular end (McCabe and Pitt 2011). Other features of (highly functional) spontaneous orders include: that the rules underpinning a given order apply to all participants in equal measure, with no discriminatory exemptions or privileges on the basis of personal identity or affiliation evident; that positive and negative feedback mechanisms exist to facilitate cooperation and encourage mutually beneficial outcomes; and that individuals can freely enter into, and exit from, arrangements which pervade the given spontaneous order and, indeed, may challenge pervading notions of “what works” by proposing novel ideas and solutions (diZerega 2008; Potts 2013).

I now turn to briefly outline the key characteristics of the economic, communal and political orders of civil society in which humans extensively interact.

The conception of markets within the economic order as a spontaneously ordered form of mutual cooperation between producing, trading and consuming strangers is arguably the central proposition expounded by scholars within the mainline tradition of political economy (Boettke 2012). For one of the figureheads of this mainline tradition, Friedrich Hayek, the “marvel of the market” rests upon its unheralded capacity to distil and coordinate dispersed economic knowledge—of how, why, where and when to produce, trade, and consume goods and services—that is possessed in fragmented, locally situated forms by multitudes of idiosyncratic individuals.

How does economic knowledge become coordinated, generating wealth within the reach of evermore members of civil society? Coordination is achieved with the aid of movements in relative prices (which, in turn, inform shifts in profitability), superimposing a rank of goods and services according to their relative value in competing uses. Relatively higher prices induce consumers to ration yet encourage suppliers to invest and produce more output; relatively lower prices induce consumers to accumulate yet discourage suppliers to invest and produce more. That said, prices are not necessarily unambiguous bits of information and so there is scope for relative price changes to be subject to interpretation and understanding, instead of necessarily reflexive action and reaction (Lavoie 1986; Ebeling 1991; Dekker and Kuchař 2017).

It is often overlooked by market critics that the unceasing mutual adjustments in market settings, which are the product of oftentimes bustling and rivalrous interactions between people to exchange valuable commodities on mutually agreeable terms, happen to foster spontaneously ordered cooperation between millions of strangers (Rubin 2014). To put it simply, “Paris gets fed.” Virgil Storr (2008) also observed the markets composing the economic order serve as a forum through which friendships and other social relationships emerge, contravening the suggestion that market-based interactions are always and everywhere anonymous and impersonal.

The idea that the economic order is a fount of mutual assistance provides an almost perfect segue to a description of what I call the communal order, which was at the heart of Tocqueville’s seminal meditation on civil society yet has seemingly vanished from much of modern classical liberal scholarship for varied reasons (Cornuelle 1993; Ealy 2005; Garnett Jr. 2011).

Richard Cornuelle was best known for his unremitting call to include what he called the non-market, non-state “independent sector” (I consider this to be a major element of the communal order) into a comprehensive theory of civil society. He referred to the communal order as a spontaneous order in the respect that “[c]ommunity is a consequence. It results when people come together to accomplish things that are important to them and succeed” (Cornuelle 1993, p. 32).

Action in the communal order is predominated by altruistic and solidarity acts between people. Oftentimes, individuals self-organise into agencies and other bodies to provide and to carry out acts of caring and enlightenment. In the United States alone “[t]here are hundreds of universities, elementary and secondary schools, thousands of hospitals, museums, symphony orchestras, and libraries, thousands of mutual aid groups like Alcoholics
Anonymous, to mention only a few of its more visible entities. Its institutions range from the giant Teachers’ Insurance and Annuity Association, … to a clearing house through which amputees can trade their useless left or right shoes or gloves” (Cornuelle 1992, p. 7).

The contributions of philanthropists and social entrepreneurs as of great import to the maintenance of non-commercial assistance and reciprocity be it in the form of grant funding or goods-in-kind to the needy or to participants involved in special causes, or through help to people which effectually build their capabilities, self-esteem and resilience.

Also situated within the communal order of civil society are social movements (for example, racial equality and queer rights activists) and interest groups (for example, business associations and labour unions) who utilise their divergent understandings of economic, social and political life to promote alternative, and prospectively improved, ways in which people can cooperate and align their prerogatives.

Several scholars suggest that feedback signals exist within the communal order to help participants “do well by doing good,” even if the feedbacks are less finely grained than relative prices and profit-and-loss mechanisms conditioning the economic order. In the communal order it is supposed that reputational mechanisms help direct donors and supporters toward credible groups and associations which demonstrate an effectiveness in the fulfillment of their missions, and provide incentives for investment (of time, money and degree of engagement) in worthwhile ventures of mutual assistance (Boettke and Coyne 2009; Chamlee-Wright and Myers 2008).

Social movements and interest groups, along with the other manifestations of associational life within the communal order, play important roles “in stimulating political participation, increasing the political efficacy and skill of democratic citizens and promoting an appreciation of the obligations as well as the rights of democratic citizenship” (diZerega 2001, p. 765). Insights into the potentially democrtising effect of people interacting within the communal order brings me to a third kind of spontaneous order situated within civil society—namely, the political order.


As is the case for the economic and communal orders, the political order does not exist, nor is it designed, to serve any particular purpose or to achieve any specific goal (Madison 1990). According to Nell (2017), the value of the democratic process is that facilitates the articulation from the bottom-up, as it were, of valuable information—including that which seems to be tacit, or otherwise un conveyed, knowledge about political demands or preferences.

The efficacy of the political order from this standpoint depends, of course, upon an abiding respect for certain norms and procedures facilitating a continuous “discussion,” in the Knightian sense, amongst all those potentially affected by policy changes:

In democracies abstract criteria for political membership and procedures for participation can be linked with any specifiable interests. Political leaders and measures are selected by balloting in which every citizen’s vote counts equally, and in which procedural and reasonably open criteria determine who runs for office. Civil liberties safeguard an indefinite and unpredictable variety of political opinions and programs. Freedom of speech can be used to support or attack political leaders and their policies. … Criteria for democratic citizenship and participation are completely divorced from citizens’ substantive views and values (diZerega 2001, p. 765).

The mark of a functional political order is the potentially indefinite and unpredictable numbers of political positions competing for the democratic affections of the general public (diZerega 1989; Madison 1990). The unfolding spontaneous order of politics embodies the fundamental notion that political decisions are the product of discovery processes orchestrated by sovereign-citizens. Buchanan also reminds us that “democracy as “government by discussion” implies that individual values can and do change in the process of decision-making. Men must be free to choose, and they must maintain an open mind if the democratic mechanism is to work at all” (Buchanan [1954] 1999, p. 99).

By no means do spontaneous orders eschew the role of purposeful action by individuals, or groups of individuals, to achieve their objectives in concert with others. Although centralised, top-down planning was construed by liberals as inimical to market functioning, associational life, and democratic deliberation, “decentralized planning by
many separate persons” (Hayek [1945] 1948, p. 79) is an entirely legitimate mode of human interaction (see also Vaughn [1982] 1994). A corollary of such insight is that “[s]ocial systems (of the human kind) are mixtures of deliberate arrangements of man-planner-designer, interacting with complex, spontaneously emerging orders, increasingly complex and important, made by no one, preconceived by no one, foreseen by no one. Both aspects—artificial and natural, man-made and spontaneous, designed and self-produced, simple and complex—should be studied in their interaction and mutual co-determination” (Zeleny 1985, p. 118).

The civil society notion entails several degrees of phenomenal complexity over and above those prevailing in any given component spontaneous order. Hayek once observed that the character of (civil) society is such that “the whole is more than the mere sum of its parts but presupposes also that these elements are related to each other in a particular manner” (Hayek 1967, p. 70). Thus, the irreducibility of civil society to given actions by individuals, with their heterogeneous values, beliefs, purposes, plans and ideas, indubitably reflects the ensemble of those interactions and relationships that people forge (Martin 2011; Lewis 2011).

The activities of those interacting within a given spontaneous order are shaped somewhat by the incentive-feedback signals emergent from within each order—that is, relative prices of the market order; reputation and esteem of the communal order; and votes of the political order. That a bewildering array of ordered interactions takes place implies that civil society itself accommodates a thick multiplicity of values in ways that the thinner economic, communal or political orders singularly cannot. This is because “feedback signals arising from different spontaneous orders merge within the more encompassing cosmos of civil society. Collectively these signals help people make the decisions needed to achieve their goals more effectively. But no single signal dominates. Each member attends to the feedback signals they wish, as much as they wish, and ignores the rest” (diZerega 2015, p. 19).

A highly functional civil society correlates well with the realisation of individual liberty since it allows for the greatest range and depth of unforced cooperation between non-intimates possible. Not only can individuals respond, or not respond, to certain feedback signals; they are free to act in a “modular” fashion, and by that I mean they can associate and dissociate with others without fear of being shunned, or punished in some other manner (Gellner [1925] 1994). There is no single focal point in civil society as individuals design, shape and direct their own lives, guided by their own interests, ideals, and passions (Ebeling 1993). The key condition, of course, is that each and all respect the equal liberty of others to engage in activities, and pursue their objectives, without rendering harm.

THE ADAPTIVITY OF CIVIL SOCIETY: A MESO PERSPECTIVE

Civil society is the epitome of the notion that, to paraphrase Friedrich Hayek, a person can live in many worlds at once. Although one can make intellectual distinctions concerning how the medley of spontaneous orders work, it should be recognised that the orders themselves are not insular or impregnable.

As highlighted by Richard Wagner (2007, 2010, 2016), civil society is exhibited by interactionist entanglements between individuals and their enterprises, organisations and ventures across the economic, communal and political orders. It is this entanglement of human relations which gives credence to the widely-held observation by social theorists that civil society possesses “structure,” and for others still, especially Hayek (1964, p. 10), “the structure of modern society has attained a degree of complexity which far exceeds that which is possible to achieve by deliberate organization.”

Although the entangled structures inhabiting civil society are of great significance, our analysis of the nature of civil society would be incomplete if it were to rest solely upon describing its structural properties. Buckley (1968, p. 497) criticised the focus of sociological (and presumably other social scientific) studies upon structural concepts, suggesting that this has led to “a rather static, overly deterministic, and elliptical view of societal workings.” Indeed, “for the sociocultural system, “structure” is only a relative stability of underlying, ongoing micro-processes” (Buckley 1968, p. 497).

Observing the structural ramparts of civil society are one thing, but comprehending “the actions and interactions of the components of an ongoing system, in which varying degrees of structuring arise, persist, dissolve, or change” (Ibid.) helps us appreciate the evolutionary dynamic of structural phenomena pervading civil society. Further still, even though the “continuous morphogenic process” (Ibid.) of civil society is not wholly reducible to individual action, classical liberals are deeply abiding methodological individualists who appreciate the pivotal role that individuals play in shaping the trajectory of orders posited within civil society.
The adaptive potential of civil society, as people act and react to novelty and other guises of change, has long been recognised in the liberal canon. The nineteenth-century French philosopher Destutt de Tracy ([1817] 2001, p. 95) observed that “society is purely and solely a continual series of exchanges. ... consequently society is an uninterrupted succession of advantages, unceasingly renewed for all its members.” Much more recently James Buchanan valorised the exchange paradigm as the *modus operandi* for human interaction in market and non-market (including political settings) (Buchanan 2005; also Alvey 2009 and Garnett Jr. 2011).

Exchange in the sense described here entails a regime of mutually beneficial cooperation involving people. That said, there needs to be a given party, or parties, that reach out to the other party, or parties, with the aspiration of initiating the given exchange. *Even before then* there must have been people sufficiently mentally alert and socially adept to opportunities that enhance their position, and that of others, in ways that fortuitously render exchanging activity within civil society. These people are what are known as “entrepreneurs,” and they are widely recognised for their pivotal role in propelling processes of economic, social and political change.

Entrepreneurial conduct need not be understood merely in its well-known, yet exclusively, economic incarnation but, in fact, as an aspect of generalised human conduct present in all societies across time and space (Boettke and Coyne 2009; Koppl 2006; Koppl and Minniti 2008). As Koppl (2006, pp. 1-2) pointedly states, “[e]ntrepreneurship is an aspect of all human action. Entrepreneurship is a human universal,” and for our purpose we can roughly map the concepts of economic entrepreneurship, social entrepreneurship and political entrepreneurship as the initial injections of novelty emergent within the economic, communal and political order trajectories, respectively.

It should be recognised that acts of entrepreneurship, prevalent in whichever order of civil society, represents an act of *dissensus* from, rather than consensus with, conventional ways of being, doing and knowing (i.e. rules). In civil society—within which a multiplex of beliefs, interests and values are subscribed to by interacting individuals—what Richard Wagner calls the “social tectonics” of discord emerge as very real features of our shared existence. In this context entrepreneurial action, born of “the unpredictable, the creative, the imaginative expressions of the human mind” (Kirzner 1982, p. 147), frustrates as much as it accommodates and discoordinates as much as it coordinates.

It is a truism that living in a civil society enables us to modularly adjust ourselves to seek greater alignment with our fellows, but the fact that entrepreneurship oftentimes arouses controversy implies that novelty often elicits a modicum of discomfort and resistance among those who bear the costs and fatigues of adjustment and readjustment. This applies to successful entrepreneurial ventures into economic, social and political unknowns. Other, less successful proposals to inject novelty endure insufficient profits, esteem or votes to gain a foothold of critical mass in civil society.

According to Martin (2011, p. 141):

> [w]ithout the possibility of individual entrepreneurs conceiving of new ways of doing things, dissenting from existing social structures, those structures would simply reproduce themselves *ad infinitum*. Interaction may give rise to novelty, but that novelty must be recognized and acted upon by individuals to have any lasting effect.

Moving from the initial phase of a given individual or agency (i.e. a socially organised rule carrier, such as a for-profit firm, non-profit charity, or political party) partaking in entrepreneurial activity we can, then, conceptualise the successive adoption of a novel economic, social or political rule by a population as the second phase of an evolutionary trajectory ramifying throughout civil society (Dopfer et. al. 2004; Dopfer and Potts 2008).

Isabel Almudi and colleagues (2017, 2017a, b; see also Markey-Towler 2016) have developed a propagation model in which people openly contribute economic resources toward supporting their preferred ideological position (or “utopia”) with respect to civil liberty, culture, environmental amenity, and the scope of market and state action. For Almudi and colleagues societal transformation, running from novel ideas to corresponding changes in social structures, is centred upon actors drawing upon their resources (effort, time, money and other resources) to convince others to accept their preferred utopia. The disparity of individual contributions toward promoting their utopia generates a contestable exchange of ideas, in which people constantly accept or reject offers of alternative organising principles for civil society.

What emanates from this complex process, given that “once an idea emerges, it can generate variations around the prior original conception” (Almudi et. al., 2017a, p. 632) and that individuals can alter their resource contributions
toward those multiple causes, is essentially a meso-level population of tenuously cohabitating ideas about how to live. According to Almudi and her colleagues their model, incidentally, seems to describe certain features of socio-economic-political change in the United States reasonably well.

We have moved from the origination phase of entrepreneurial action, embodying typically novel and sometimes useful but almost always unsettling values, to a depiction of a diffusion process whereby different ideas, interests and values are championed by varied groups and accommodated in a modus vivendi of an ongoing, albeit peaceful, struggle for prominence. Finally, I turn my attention to the final phase of change within civil society, centred upon the perpetuation, or stabilisation, of certain values into rule-systems that are commonly called “institutions” (Dopfer et al. 2004; Potts 2013).

Numerous definitions of institutions abound, though one which seems reasonably representative is as follows: “Institutions are the humanly devised constraints that structure political, economic, and social interaction” (North 1991, p. 97). This definition encompasses informal customs, mores and norms through to formal legislative edicts by the state.

An inherent value of institutions is that they provide focal points of orientation for individual planning that, incidentally, helps to ease problems of inter-subjectivity by bringing expectations of one another’s plans and intentions into better alignment, thereby reducing inconsistencies and conflicts (Martin 2011). Similarly, institutions shape the relative payoffs associated with conducting entrepreneurial activities in ways which either advantage, or disadvantage, other people.

Institutions which facilitate trust between strangers and do not unduly interfere with the free adjustment of economic mechanisms of relative prices, several property and profit-and-loss mechanisms will encourage economic entrepreneurs to strive against each other to offer up a multitude of “preferred utopias” in the form of improvements to the existing production techniques and trading possibilities.

Similarly, institutions can also facilitate cooperative social and political endeavours. Rules maintaining an ease of entry to, and exit from, existing associations within the communal order, and tolerance for the creation of new, hitherto unknown bodies and groups of like-minded persons can facilitate the peaceful co-existence of the conventional and the non-conventional. This insight was at the heart of John Stuart Mill’s defence of liberal civil society, as outlined in his famous tract On Liberty. As for political activities, political institutions (e.g., one-vote-one-value for all adults, open and contestable elections) which effectively treat individuals as intelligent, deliberating equals, and not as objects of domination and subjection, are more likely to generate outcomes perceived as reflective of the public (instead of special) interest.

Institutions come about when origination of a novel proposal leads to willing adoption by large segments of the population which then persists through time. Institutions can be so successful that they exert what Lewis (2012) calls a “downward causal influence” shaping attitudes, dispositions, interests and values in that most people routinely follow them, and indeed epitomise the enduring representations of social reality we call “culture.” By the same token, however, institutions themselves are not immune to challenge and change if there are insufficient incentives to justify their reproduction through time. In those instances we observe the processes of origination and adaptation of proposed, and actual, amendments of incumbent rule-systems at play within the various orders of civil society.

What has been offered here is a meso-centric narrative of change to the structured (and entangled) interactions between persons along economic, communal and political dimensions. It is from this perspective that we see that “social selection and relative stabilization or institutionalization of normatively interpreted role relations and value patterns occurs through the variety of processes usually studied under the headings of conflict, competition, accommodation, and such; power, authority and compliance; and “collective behavior,” from mob behavior through opinion formation processes and social movements to organized war” (Buckley 1968, p. 495).

Contestation over different ways to organise human affairs doubtlessly creates many avenues for disagreement between individuals and groups, depending upon the issues at stake. While there are always tensions within and across economic, communal and political orders of interaction, not least because of the fact that the incentive structures and value sets redounding in civil society are often incommensurable if not outright incompatible, it is nonetheless eminently possible for one to discover and achieve mutually agreeable exchanges with other people. This has long been demonstrated in the form of the “Great Enrichment” of market-tested betterment (McCloskey 2010) and other manifestations of human progress (Norberg 2016).
REFLECTIONS ON THE “CIVIL SOCIETY IN DECLINE” NARRATIVE

In recent decades, if not longer, arguments to the effect that civil society is in a state of “decline,” or gripped by some form of “decay,” have proliferated in both academic and popular literature. This has been reflected in studies pointing to falling membership rates of groups such as fraternal organisations, labour unions and religious orders, and declining voluntarism within clubs and societies (Putnam 1995, 2000). Although many of these studies do acknowledge the heightened interest in environmental concerns, and a growing affinity with feminist, racial and queer civil rights groups, they also suggest levels of commitment to these newer groups is of somewhat fleeting and ad hoc in nature (Wuthnow 1998; Skocpol 2004).

Those concerned with deteriorating civil society point out the more active involvement of people in a variety of associations in the past was not only indicative of a vibrant communal order, but tended to strengthen the level of social capital within civil society more generally. A greater social capital stock is associated with greater levels of trust and mutual respect, enabling communities to more effectively pitch together to resolve common problems without recourse to the likes of state action and other distant, top-down interventions. The observation that membership of, and participation in, numerous associations were declining therefore suggests generic deterioration in the stock of social capital, in turn threatening the functionality of civil society.

There is no doubt that the kinds of associations grafted within civil society have waxed and waned in modern times. To suggest, however, this implies civil society in its entirety is fraying or weakening is questionable.

We should remind ourselves that civil society is a complex and adaptive process accommodating an ensemble of various spontaneous orders which, in turn, facilitate interactions and preferential attachments of a mind-boggling number. Variety and change is at the core of the ways in which humans mingle, and these characteristics are not necessarily problematic from the perspective of the functionality and viability of civil society. As Buckley (1968, p. 495) stated:

[a] requisite of sociocultural systems is the development and maintenance of a significant level of non-pathological deviance manifest as a pool of alternate ideas and behaviors with respect to the traditional, institutionalized ideologies and role behaviors. Rigidification of any given institutional structure must eventually lead to disruption or dissolution of the society by way of internal upheaval or ineffectiveness against external challenge.

The classical liberal Friedrich Hayek famously suggested that people adopted and retained institutions which were more conducive to the perpetuation and flourishing of the human species, whereas those who adopted less effective institutions were cast aside into irrelevance or oblivion. Some have made the suggestion that the retention of certain traditional ways of living are necessary to ensure our survival, and indeed Hayek’s writings on cultural evolution are often interpreted in this conservative fashion.

I suggest that survival in an evolutionary sense depends not so much upon institutional persistence, which unquestionably can matter greatly under certain conditions, but upon a capacity of institutions to adapt and change, and to accommodate discoveries and challenges to prevailing dispositions, inclinations and values. Indeed, “[e]volutionary models … need to incorporate a means of change in order to be complete” (Vaughn [1982] 1994, p. 231), and, so, purposive human action should generally be viewed as both a fundamental and welcome feature of life in civil society and the orders within it.

Other accounts expressing a fear of civil-societal decline reflect concerns that the non-permeability of spontaneous orders within civil society may lead to the workings of a given order being inordinately, thus inappropriately, influenced by paradigms and values (perhaps even feedback mechanisms) widely perceived to be alien to it. From this standpoint the inherent diversity of human existence evident in civil society may dissipate as the logics of a given order dominate the logics of others.

Scholars in the communitarian tradition have expressed the view that the Gemeinschaft of intimate personal and familial ties are sorely tested by the Gesellschaft of unrepentantly commodifying markets of an impersonal nature, whereas the critics of “neo-liberalism” suggest that privatisation and deregulation reforms signal a disturbing retrenchment of the political (and communal) from human affairs. On the other hand, classical liberals view governmental activity as threatening the displacement, or “crowding out of,” those activities voluntarily undertaken by people within market and civic associational contexts.
Richard Wagner advises us that the “structured living-togetherness” that embodies civil society is exemplified by entangled human interactions which can criss-cross the economic, communal and political orders, and to some extent entanglement is unavoidable. Even so, he raises the issue that growing political influence in economic decision-making not only induces “rent seeking” behaviour, in which entrepreneurs attempt to extract special fiscal and regulatory advantages and privileges from government, but mixes private and public ordering principles in such a way as to generate calculational problems corrupting the political economy. One could extend such insights to describe, for example, the problems surrounding the displacement of polycentrically situated civil associations, such as mutual aid societies, by a massive, yet increasingly paternalistic and fiscally unaffordable, government welfare state.

A group of mid-twentieth century liberal scholars, mainly from Germany, acknowledged the fuzzy boundaries between spontaneous orders embedded within civil society but, nonetheless, sought to enumerate policies and institutional rules to ensure that the scope of entangled relations delicately remain within reasonable bounds, and without losing the dynamic propensities within civil society that yield widespread benefits (Stützel et. al. 1982; Peacock and Willgerodt 1989a, b).

The “ordoliberals,” as they were labelled, saw the maintenance of civil society as resting upon a “constitutional cultivation” of sorts—rules should maintain an economic order featuring open, competitive markets and whittle away monopolies that harm consumers’ interests; rules should encourage the development of vibrant, local communities but not concentrate social power in overbearing hierarchies; and rules should constrain political action in accordance with the rule of law, and preferably enshrine polycentric activity in line with subsidiarity principles. Dopfer and Potts (2008) offer a modern reinterpretation of key ordoliberal doctrines for maintaining fruitful interdependence between civil society’s orders:

An open market economy in which agents are not just operationally free to choose the commodities they will consume, but also generically free to choose the knowledge they will adopt, ultimately rests upon a social, political and cultural order in which variety is sustainable. Generic openness therefore fundamentally requires tolerance of novel ideas as a primary condition, and only excitement about novel ideas as an accelerant. Such tolerance may well be constitutionally embedded in, for example, freedom of speech laws or other legislation that affords and protects the right to be different. Difference is the elemental driver of economic evolution, and societies that are tolerant of different ideas and rules carried by micro agents possess a necessary condition for economic evolution (Dopfer and Potts 2008, p. 98).

By contrast, persistently hierarchical and monopolistic relationships in any of the orders of civil society has the potential to fatally compromise those senses of equality, openness and toleration upon which we all rely to have decent opportunities to strike arrangements with others for mutual benefit. Indeed, a renewed appreciation of such lessons could prove most valuable in a time in which the pressure points of market concentration, social strife and political authoritarianism appear to be looming larger across developed countries, perhaps risking our capability to live freely and peacefully with one another.

CONCLUSION

No truer words were expressed than by Gus DiZerega when he reckoned the cosmos of civil society is that realm of individual choices across a broad range of different values consistent with peace with others (DiZerega 2015, p. 20). I can be atheist, queer, a hobby astronomer, an economist and a classical liberal all at once. Similarly, I can ignore calls to prayer; disparage pleas to privilege one race, gender and nationality over another; and critique demands to nationalise the economy at the same time. Many other members of civil society may not endorse my opinions and how I live my life, and I needn’t necessarily endorse theirs, but we maintain détente as we strive to paint our unique stories onto the canvas of life.

The ability of different people to indulge in their choices—individually and among persons, as well as over time, is instrumental to human flourishing, but choice-fulfilment often requires mutual assistance. People invariably engage with others to transform resources and earn money; to give and receive love and emotional and other supports; and to persuade others to gain political credibility and influence. For many classical liberal theorists the “variety of situations” (to paraphrase Wilhelm von Humboldt) arising from the varied orders comprising civil society serves as a very foundation for the realisation of freedom itself.

Civil society represents that ingenious admixture of evolving cooperative and competitive relationships, always peaceful and abhorring those techniques which privilege certain activities (indeed, certain individuals) above others on the basis of group identity or some other status. The
twin-marvels of civil society concern its accommodation of human diversity and its evolutionary attributes. To the extent to which these qualities are diligently and vigilantly well maintained, without extirpating individual liberties, the chances are civil society will continue to serve as that indispensable arena in which we look to accommodate mutually beneficial economic, social and political changes together.

REFERENCES


1. INTRODUCTION

Jane Jacobs (1961) argues that the city is a complex system possessing emergent characters that are the result of place-based processes of self-organization. She sustains that the more the city is planned, controlled and designed from the top, the less flexible space there is for spontaneous adaptations and improvements, connected to the use of the dispersed knowledge of ordinary people (Callahan and Ikeda 2014; Cozzolino 2017; Gordon 2012; Ikeda 2017; Moroni 2016). Despite planners do not agree in the ways in which it is possible to enable spontaneity—and many different approaches are now on the table (Rauws 2017)—today, more than sixty years after Jacobs’ seminal work, one idea seems to be largely shared by experts: cities must be considered as complex self-organizing systems. In the planning field this idea has opened the door to concepts like spontaneity and flexibility which are now indicated as valuable alternatives to orthodox comprehensive planning practices. The article discusses the different ways in which spontaneity and flexibility work under different kinds of planning conditions (being material, like buildings, open spaces or infrastructures, or immaterial, like building codes or land-use plans). In particular, it recognizes the preeminent role played by the rules over the built-environment in defining the flexible space for the evolution of emergent socio-spatial configurations.

The main thesis is that the flexible space for spontaneity and emergent configurations does not depend solely on the physical dimension of the built environment (which is ob-
viously a key aspect), but also—and even mostly—on the way in which planning rules are written and provided (Ben-Joseph et al. 2005; Kim et al. 2016; Talen 2016). In other words, the degree of spontaneity and flexibility depends, in different ways, upon both social and physical conditions.

The article explores why the relationship between spontaneity and planning conditions is multi-layered and hierarchical, recognizing the preeminent role played by the rules over the built-environment. It discusses this question without having the ambition to investigate in detail many relevant side-related issues—already well-discussed by other scholars—such as indicating which planning approaches are more suitable to welcome spontaneity (Alfasi and Portugali 2007; Andersson 2014; Buhtelaar et al. 2014; Moroni 2015; Totry-Fakhoury and Alfasi 2017), analysing the effects of planning interventions in economics (Bertaund 2014; Gleaser 2011; Pennington 2002; Webster and Lai 2003), or discussing the main aesthetical issues linked to spontaneous developments (Alexander 1967; Hakim 2014; Nilufer 2004; Romano 2010). Each of these issues would deserve a specific article. This is an attempt to demonstrate, from a Jacobsian perspective, the reasons why rules matter, at least as much as the built environment, in defining the flexible space for spontaneous actions and the development of unpredictable changes from the bottom-up.

Beyond the introduction, the article is divided into four other sections: three of them answer the following questions; (section 2) “What do you mean by spontaneity and emergent configurations and why are they relevant?”; (section 3) “How does spontaneity work within the urban fabric?”; (section 4) “How does spontaneity work within the rules?”. The article ends by providing general conclusions and devices (section 5).

2. WHAT DO WE MEAN BY SPONTANEITY AND EMERGENT CONFIGURATIONS AND WHY ARE THEY RELEVANT?

Before getting to the heart of the matter, some terminological clarifications are needed. First of all, the term action (see, for instance, Carter 1999; Cody 1967) is contextualized to the field of urban planning and, therefore, it may represent certain general types of (urban) actions, like the act of building or using a certain amount of space, or other ancillary actions, helpful for pursuing them, (e.g., obtaining loans, signing agreements and covenants, etc.). Secondly, in order to differentiate between two different kinds of phenomena, the article distinguishes between spontaneous and emergent configurations. With the term spontaneous, it refers to intentional actions developed by self-determining and intentional agents, (e.g., landowners, developers, householders, etc.), whilst the term emergent refers to the unplanned social-spatial aggregation that is the (unintended) result of uncoordinated actions and interactions. In other words, the article differentiates voluntary actions, (which can be undertaken by an individual or a group of individuals acting in cooperation; see Mises 1963/1998), from evolving situations, that are the result of many, unaccountable actions interacting over time (Schelling, 1978).

The term spontaneous derives from the Latin word spontaneus which means “of one’s free will”. In general, the word describes persons and characters with a sense of “acting of one’s accord” or “occurring without external pressure”. Hence, we can think of the word “spontaneity” as a particular quality of actions (Beito et al. 2002), whereas, the concept of emergent configuration means any (socio-spatial) configuration that is the result, over time, of a countless number of actions, but not a direct consequence of a single action or design (Hayek 1960; Polanyi 1951). Configurations of this kind are detectable in any system composed of a multitude of agents who pursue different plans and separate actions. The highest expression is observable in and within the city (Holland 1995, p. 41).

The main arguments for the evolution of emergent configurations lie in the possibility of adapting, over time, the physical and social world in ways that none could predict in advance, leaving to the system the opportunity of efficiently reacting to various contextual needs, and the possibility of making more efficient use of the so-called dispersed knowledge (Hayek 1945 and 1960). Emergent configurations, however, are not synonymous with chaos or disorder, but they are a series of stable (but evolving) patterns which are not defined through central coordination. On the one hand, emergent configurations enable social interaction and the self-coordination of systems; on the other, they structure the character and peculiarities of specific places. Their evolution is open-ended (i.e., unpredictable and uncertain), but path-dependent.

However, to recognize the relevance of emergent configurations doesn’t necessarily mean to be in favour of anarchism, or laissez-faire, but to appreciate the intrinsic and instrumental values of spontaneous actions in the functioning city. All this without being against the presence of good planning rules (Moroni 2010). Firstly, spontaneity is good in itself, at the level of the individual, because it allows people to pursue their own ends by means of their knowl-
edge and creativity, also experimenting with new solutions and actions. Second, from continuous processes of trial and error, society can get benefits reaching a level of progress and innovation otherwise unachievable under a state of extensive overall control, developing self-organizing systems in which society can make more efficient use of polycentric forces (Ikeda 2004).

The city is irremediably both the product of planning interventions, which introduce artificially constructed orders, and the evolution of emergent configurations (Bertaud 2004). So, we cannot think of the city as a fully emergent configuration but, it is rather more appropriate to see it as a spatial and social configuration that may reach different degrees of overall spontaneity (Buitelaar et al. 2014). This works as a trade-off between planning conditions and the space for self-determining (unpredictable) spontaneous actions, where the former introduces a certain degree of spatial control over a world of infinite possible becoming (Ikeda 2017).

As already argued, two different types of planning conditions constrain spontaneity: the material dimension of the built environment (i.e., physical objects, such as condominiums, streets, parks and so on), and the social dimension of rules. Therefore, we may speak of two trade-offs: the first has a spatial dimension; the second has a regulative dimension. Both levels, in different ways, influence the flexible space for the evolution of emergent configurations and the ways in which spontaneity can be manifested or less.

3. HOW DOES SPONTANEITY WORK WITHIN THE URBAN FABRIC?

This section examines the issue of spontaneity within the dimension of the built environment (Bergevoet et al. 2016; Manewa et al. 2009; Roggema 2014). In particular, it examines the relationship between the design of the physical space (i.e., the intentionally created spatial order), and its level of flexibility for the evolution of emergent configurations.

Jacobs highlights that any physical element that has been deliberately designed and built always provides a certain degree of flexibility, depending upon both internal and external factors. In fact, a common feature of any work of architecture is that, with the passage of time, it (potentially) starts to differentiate itself from its initial design and it is subject to reinterpretation (Brand 2010; Easterly 2015). This process happens for two main reasons: first of all, some (unpredictable) actions occur inside; secondly, it undergoes the effects of all the actions and changes that happen outside. As regards the physical dimension of the built environment, two important factors are worth considering: the first is the scale of design and its level of detail; the second regards the passage of time (Ikeda 2017).

The scale of design and its level of detail

The first implication is easily understood: the higher the level of the detail provided by a particular project or plan, the less will be the flexible space for spontaneous actions and unexpected emergent future arrangements. Moreover, the greater the scale of design that is subjected to a unitary plan, the greater will be the amount of space, subject to the control of that specific plan. For example: imagine a unitary designed residential area which includes a hundred condominiums. Architects could face this design in different ways, with variable levels of scale and detail. To demonstrate that, we may address the issue moving from a first design, that guarantees a high degree of flexibility, to one in which the flexible space is narrowed down and nearly cancelled.

A first way to design the new residential area may be that of solely to draw the open public spaces, and then leave open, to other architects, builders and inhabitants, the possibility of defining and filling the grain of the area. Such developments could emerge according to a succession of spontaneous actions, constrained by the presence of open public spaces (for instance, roads, green areas and so on) and certain planning rules.

In the second case, it could be that the architect does not design only the open public spaces, but also the size, the footprints of all the buildings and also their physical relationships. The differences between the first two approaches are remarkable.

Let’s go further. Suppose that the work of the architect is not only limited to the design of the open public spaces and the footprints of all the buildings, but he or she also decides to determine their interior subdivisions, as well as their functions. At this point, the design will begin to assume a very rigid character. However, its level of detail could still be increased. In fact, we can imagine that, in addition to the design of open public spaces, buildings and their internal divisions and functions, the architect may also decide to design, in a unified and detailed way, the interior furnishings. At this point, only a little flexible space would remain for self-determining spontaneous actions and the use of dispersed knowledge of future inhabitants.
In brief, the implications of the scale and detail of design regards designers’ intentions to impose a particular spatial order, which, in its turn, may be more or less extended and accommodate more or less flexibility. To better comprehend it, we can think of highly planned cities like Brasilia, or the Chinese “ghost city”, or the typical neighbourhoods of the Soviet era (Hirt 2013; Ikeda 2017; Zarecor et al. 2012) as relevant examples of large scale design with a high level of detail of the built environment. On the contrary, in the case of a high degree of flexibility, for instance, we can refer to forms of organic developments diffused in the last years in the Netherlands (Cozzolino et al. 2017; Oosterma et al. 2015; Rauws & de Roo 2016).

4. HOW DOES SPONTANEITY WORK WITHIN THE RULES?

Now we look at the built environment and the flexible space for the evolution of emergent configurations deriving from the presence of planning rules. The general idea is to understand the relationship between the physical environment, (planning) rules and the flexible space for actions that can alter the pre-existing urban structure, bringing it into an adapted state of affairs.

What has been pointed out in the previous section is fundamental to the following discussion. However, further observations must be set as crucial pre-conditions for any agents’ actions. In fact, more than the physical space in itself, what primarily influences agents’ actions space (Ikeda 2007) within the built environment are the rules (Moroni 2015). To show this, we go through three comparative empirical examples which describe imaginary spatial differentiation processes, showing why the rules (more than the design of the space itself) are the prior condition for the adaptation of the physical space. This statement is not an attempt to diminish the importance of the design of the physical space (which is obviously a fundamental condition); rather, it is an attempt to distinguish two different levels of the discourse in which rules become the meta-condition, whilst the built environment is the material concretization of actions that, in their turn, are always conditioned by certain rules.

The first two examples compare two initially identical urban settlements but with different rules. The third example shows that it is not so much the initial design that determines, in absolute terms, the flexible space for the evolution of new emergent configurations, (i.e., it is not important so much to know if the city, or part of it, evolved spontaneously or not); rather, what influences the extension of the flexible space are, first of all, the rules which discipline agents’ actions in space. In fact, it could be that cities, or part of them, evolved with a high degree of spontaneity, at some point, they may have little flexible space for the evolution of emergent configurations; whilst, vice versa, highly-designed environments at some point may host a high degree of flexible space for the evolution of emergent configurations.

Example 1: two identical buildings with the same framework-rules, inside two identical neighbourhoods with different framework-rules

The following example demonstrates that, even in the case...
in which an urban element (such as a building) is designed and regulated in a very detailed manner (at the point to scale down, at the minimum level, the flexible space), the way in which its surrounding area (for instance, its neighbourhood) is regulated, inevitably influences the future adaptations of such an element.

Imagine two identical buildings: B (with the same design detail), as well as their respective neighbourhoods, X and Y, in their turn identical but with different rules. Specifically, to put it simply, these buildings could be two identical old farms now restored as restaurants that are located within an expanding urban realm.

Now imagine an incremental transformation process set in three steps: t1, t2, t3. With the passage from t1 to t2, and from t2 to t3, neighbourhoods X and Y, (not the buildings B), will be subject to different kinds of actions. These actions will alter the initial spatial configuration of both BX and BY; firstly, in BX” and BY”, and then in BX”” and BY””. Despite the two buildings B keeping their design and function unchanged, the hypothesis is that what happens during the process respectively in X and Y will lead to the buildings having substantial differences.

Example 1:

<table>
<thead>
<tr>
<th>Evolution</th>
<th>BX</th>
<th>BY</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td><img src="image1.png" alt="Diagram" /></td>
<td><img src="image2.png" alt="Diagram" /></td>
</tr>
</tbody>
</table>
| Actions | - New expressway behind the building  
- Construction of two single family houses with private garden | - Widening of the existing road  
- Construction of three multi-storey buildings along the main road  
- New local road |
| T2 | ![Diagram](image3.png) | ![Diagram](image4.png) |
| Actions | - New electric high-voltage line  
- Construction of a double family house | - Construction of multi-storey buildings  
- New neighbourhood park  
- Subway station |
Even though the buildings remain unchanged over time, in their initial function and design, we may assume that to some extent, they have been subjected to a process of differentiation, derived from the actions undergone in their respective neighbourhoods, X and Y, which were in their turn regulated and planned in different ways. For instance, we can imagine that, at the end of the transformation process undergone in both neighbourhoods, the building in neighbourhood Y has become more attractive than the building in neighbourhood X (or vice versa). In other words, the process of differentiation has been flawed from the outset by the existence of different rules that must be considered as the precondition for any changes which occurred in time T1 and T2 in both neighbourhoods. We may imagine, for instance, that at the initial state of the process, T1, X and Y possessed completely different rules as regards the list of possible land uses, the maximum FAR, and so on.\

However, with this example, we see the importance of rules but we still cannot fully understand the degree of overall spontaneity. In fact, in both cases BX and BY, despite their final differentiation, the degree of overall spontaneity may still be very low; for instance, their respective rules could be thought to be instrumental toward a predetermined constructed order. If this is the case, in both areas, the flexible space for spontaneous action provided by the rules could be considered almost absent. It might be that—we cannot exclude this chance—the rules correspond to precise planners’ will to reach a prefigured end state (for instance, to transform X into a garden city, or to concentrate in Y all the urban transformations). If this is the case, there wouldn’t be almost any degree of spontaneity since everything has been planned from scratch.

In brief: two identical buildings which are the product of the same design and regulation may be subjected to processes of differentiation if their neighbourhoods (which are in their turn identical) are regulated in different ways.

**Example 2: two identical and unitarily designed neighbourhoods with different rules**

The following example demonstrates that, if two identical urban settlements that are the result of a unitary design are regulated in different ways, they may undergo completely different developmental processes. Moreover, the example emphasizes that the initial architectural design does not necessarily coincide with the rules that discipline future actions and adaptations. The two levels

### Example 2:

<table>
<thead>
<tr>
<th>Evolution</th>
<th>W</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td><img src="image1.png" alt="Diagram" /></td>
<td><img src="image2.png" alt="Diagram" /></td>
</tr>
</tbody>
</table>
| Actions   | Ordinary maintenance | - Widening of the existing road  
- Construction of multi-storey buildings along the main road  
- Opening of new neighbourhood shops |
| T2        | ![Diagram](image3.png) | ![Diagram](image4.png) |
| Actions   | Ordinary maintenance | - Construction of multi-storey buildings with different uses and substitution of three old houses  
- Creation of two squares |
| T3        | ![Diagram](image5.png) | ![Diagram](image6.png) |
do not necessarily coincide; or, better still, the design of the physical space and the rules must be clearly distinguished. If we completely evade the question of rules, we cannot understand that the work of architecture is simply a starting point for the evolution of new emergent configurations. In fact, material artefacts are modifiable conditions, open to future adjustments.

Imagine that W and Z are two residential suburban neighbourhoods that are the product of an identical design, composed mostly of one or two-family houses with private gardens. Imagine also that the level of design detail of both W and Z is very high but they have different degrees of rules’ prescription. In W, for instance, the rules set that the initial state of affairs cannot be altered (only residential uses are permitted, and there is no possibility to have additional FAR). In contrast, in Z there are no specific prescriptions about possible uses, and owners are always allowed to expand their proprieties by purchasing and building additional FAR.

As we see with this example, although W and Z were initially identical (both are the product of unitary design and both were built all at once), with the passage of time, due to extremely different rules, they have undergone a significant process of differentiation. If W at time T3 is still the same residential neighbourhood as defined by the original master-plan, Z, on the other hand, has reached a considerably new overall urban character. Therefore, although this is an extreme case, we clearly understand that architectural design is not the only determining factor for the evolution of emergent configurations. In the first instance, planning rules may expand or restrict the range of possible actions that can bring the physical space toward unexpected adaptations. Naturally, there are architectural constructions that are more adaptable than others. However, rules remain the primary condition for unpredictable future spontaneous actions.

Example 3: an urban settlement that evolved with a high degree of overall spontaneity now subject to stringent rules, and a highly designed urban area now open to spontaneous actions

To comprehend how flexible the space is to the evolution of emergent configurations, it is not so important to study whether certain settlements evolved spontaneously or not; rather, we should firstly look at the existing rules. The following example demonstrates two main issues: firstly, it is not only the initial design of the space that matters for the evolution of emergent configurations; secondly, rules play a major role regardless of the fact that an area evolved with a high degree of spontaneity or it has been designed and built all at once. In particular, it helps to distinguish urban areas developed with a high degree of overall spontaneity (or not) from their factual existing capacity to welcome future emergent adaptations. In fact, due to the introduction of very prescriptive rules, it may be that a formerly highly emergent urban environment no longer has room for the evolution of emergent configurations. Whilst, on the contrary, highly detailed and designed settlements may turn in areas that have flexible space for spontaneous actions and the evolution of emergent configurations. This eventuality is demonstrable by considering two cases of the previous examples.

To do this, we compare the case “BY” (example 1) with the case “W” (example 2), both at time t3. The case BY may hypothetically represent a case of high overall level of spontaneity, whilst the case W represents a neighbourhood built all at once with a very low degree of spontaneity. Thus, differently from the two previous examples, we do not compare two areas with identical features, but completely the other way around.

Now imagine that, at a certain point, W and BY are subject to very different regulative measures. On the one hand, planners decide that BY has reached an intrinsic historical value, and consequently, they choose to put it into a state of absolute protection and preservation. On the other hand, at the same time, planners consider that it is necessary to densify W (for instance, to contrast the phenomenon of soil consumption) giving the opportunity to the various landowners to densify the area. If this is the case, a sudden inversion would make W more flexible than BY and consequently more open to future emergent adaptations. This eventuality is evident in the following table.
Example 3:

<table>
<thead>
<tr>
<th>Evolution</th>
<th>T3</th>
<th>W</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><img src="Example3_T3_W.png" alt="Image" /></td>
<td></td>
</tr>
<tr>
<td>Actions</td>
<td>Widening of the existing road</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Construction of multi-storey buildings along the main road</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Opening of new neighbourhood shops</td>
<td></td>
</tr>
<tr>
<td>BY</td>
<td><img src="Example3_T3_BY.png" alt="Image" /></td>
<td></td>
</tr>
<tr>
<td>Actions</td>
<td>Ordinary maintenance</td>
<td></td>
</tr>
</tbody>
</table>

According to what has just been shown, it is empirically irrelevant to distinguish areas which have evolved with a high degree of spontaneity from others that are the results of unitary design. In both cases, the flexible space for the evolution of emergent configurations first of all depends upon the existing rules which discipline agents’ action in space. For example, it may be that, as often happens, typically emergent configurations, such as the medieval settlements of European cities, become almost immutable objects, with no space for future adaptations or spontaneous actions, whilst modern unitarily designed neighbourhoods start to be adapted and reused in completely different—sometimes also innovative—ways (Bergevoet 2016; Franck et al. 2013; Urhahn Urban Design 2010).

CONCLUSION

The article maintains that the level of spontaneity in cities is influenced both by the physical dimension of the built environment and by the social dimension of rules. Although the physical configuration of cities is a relevant condition, in the end, it remains (with different degrees) always modifiable and open to future adjustments and adaptations and this is regardless of their scale of design or detail. Differently, rules can completely stifle the flexible space for spontaneous actions and the progressive evolution of emergent configurations. In other words, rules can easily decrease the possibility of adapting over time to the physical and social world in ways that none could predict in advance, preventing the system from reacting efficiently to various contextual needs. From this perspective, rules influence and filter the way in which the physical dimension of the built environment (independently from its inherent spatial flexibility) is step by step adjusted and modified.
Hence, spontaneity and flexibility ought to be analysed, both at the primary meta-level of rules and also at the level of the built environment. Moreover, even though the two levels are connected (i.e., they refer to the same spatial configuration), they work following deeply different logics.

In brief, from a planning perspective, flexibility and spontaneity can be welcomed in two ways: (i) building spaces that are adaptable and easily reinterpretable, and (ii) providing rules that enable spontaneity, unpredictable changes and innovation.

To give an example, during the twentieth century, many large-scale rigid transformations were planned and developed all at once, following the idea that detailed masterplans and comprehensive plans could have reached, in the short-run, efficient socio-spatial configuration (Akbar 1998; Bertaud 2014; Callahan et al. 2014; Romano 2010). This approach, with the passage of time, has shown clear limits; already in the ‘60s, Jane Jacobs well anticipated certain adverse effects. In particular, she maintains that most of the top-down settlements developed all at once and without too much space for spontaneity and further reinterpretations, have clear limits in promoting vibrant environments and are more inclined to stagnation and undergoing a slumming process exactly because they are not open to further reinterpretation and adjustment by their inhabitants (Jacobs 1961). On the contrary, according to the author, the rise of lively environments (lively in the long run) are more likely to emerge when the configurations of neighbourhoods facilitate intense spontaneous interactions and when the built environment can be slowly adapted and innovated by keeping alive its existing and constitutive patterns. Configurations of this kind emerge organically over time, are continuously adjusted, and cannot be planned and built all at once from scratch. In brief, the design of physical conformations must not, in other words, be such as to permit only a scant number of actions, but must instead be capable of accommodating change, both in terms of function and structure, and permit continuous improvements (see for instance, Aravena 2012; Bergevoet et al. 2016; Franck et al. 2013; Roggema 2014; Urhahn 2010).

On the other hand, if we recognize the essential role of spontaneity and flexibility in the city functioning (i.e., its creative role for innovation, the importance of localized knowledge, the need for perpetual adjustment and improvements of the built environment, and so on), we cannot regulate the city in order to obtain or preserve the specific and predetermined social-spatial configurations as we would like or desire (Alfasi et al. 2007; Buitelaar et al. 2010; Cozzolino 2017; Moroni 2011; Rauws 2017). We must provide rules that welcome flexibility and can guarantee that existing socio-spatial configurations can be adapted over time. In other words, rules cannot be concerned with the overall physical outcomes, nor even indicate precise functions or locations. Rules should not predefine specific states of affairs but they might leave the future open to a wide array of solutions, within a process of long-term transformation (Moroni and Cozzolino forthcoming). This, in general, means that we should discard the traditional, strictly instrumental idea of law embracing a nomocratic approach (Moroni 2010). Flexibility and spontaneity however, have often been interpreted as a never-ending opportunity to adapt the institutional framework any time that novelties—in terms of needs, opportunities and problems—emerge in society, leading the planning system toward the disputable idea of making rules always more unstable. This interpretation is distant from a serious approach to the law (Buchanan et al. 2008; Epstein 2009; Kaza et al. 2011). Rules require stability. If rules are determined afresh each morning by the toss of a coin, there are no rules at all. In other words, to be flexible should not be the rules but the space that they leave for the development of innovative actions, which are, by definition, unpredictable.

To conclude, flexibility has not to be sought in continuous adjustment of the institutional framework, but it has to be assumed as a principle, to write good rules that are as general and simple as possible and, at the same time, rules that are able to avoid the emergence of undesirable emergent configurations (Moroni et al., 2018).
NOTES

1 Oosterwold is a large-scale transformation of 43 km² that has not a master plan or zoning map, but only a limited number of public rules regarding the construction of plots that allows for the development of a self-organized urban realm (Cozzolino et al. 2017).

2 With the term “rule”, we refer to all the intentional social conditions, such as the land-use plan, building codes, taxation, etc., that discipline action space and are imposed by a public authority (Moroni 2015).

3 The idea of architectonic adaptability proposed by Alejandro Aravena (2012) is an interesting example. However, Aravena’s ideas would be nothing without a legal framework that leaves enough room for flexibility and spontaneous actions.

4 For instance, the allowed FAR in X was very low compared to that of Y (we put, for example, that in the case of X it was possible to build only isolated buildings in the middle of the plot, while in the case of Y was possible to build up to 15 floors with a maximum land coverage of 80%). Moreover, X and Y differ also for the list of permitted land uses: for instance, in the case of neighborhood X, the land-use plan wants to preserve an agricultural landscape, while in the case of the neighborhood Y, the land-use plan promotes high-density development.

REFERENCES


Cities, Agriculture, and Economic Development: The Debate over Jane Jacobs’s ‘Cities-First Thesis’

SANFORD IKEDA
Professor of Economics
Purchase College, SUNY

Web: https://www.purchase.edu/live/profiles/624-sanford-ikeda
Email: sanford.ikeda@purchase.edu

Abstract: Jane Jacobs (1969) argues that cities necessarily precede agriculture. Smith, Ur, & Feinman (2014) argue that archeological evidence conclusively demonstrates that Jane Jacobs’s thesis is wrong. Taylor (2014) takes issue with their definition of a city, while I point out certain limitations on their concept of agriculture. Nevertheless, even if Jacobs is wrong about the precedence of agriculture, I argue that the urban processes that she describes and that I have elaborated on elsewhere can still serve as the basis for understanding the spread and further development of agriculture as well as the central role of urban centers in economic development and innovation.

INTRODUCTION

There is a widespread consensus among urbanists of all stripes that cities have been and continue to be the main drivers of economic development (Glaeser 2011). One of the strongest proponents of this idea, Jane Jacobs, went so far as to argue (1969) that the only sensible explanation for the development of agriculture, perhaps the single most important innovation in human history, is that it is the result of, rather than the necessary precursor to, urbanization. That, of course, is a view contrary to that of most archeologists as well as the general public.

Recently, Smith, Ur, and Feinman (2014) argue that Jacobs’s “cities first” thesis is easily contradicted by the weight of modern archeological evidence and expert archeological opinion. It is that assertion and its consequences for the framework Jacobs developed to explain the role of cities in economic development, and innovation in particular, that I wish to explore. Jacobs’s analysis incorporated diversity of skill, knowledge and tastes, population density, entrepreneurial discovery, and economic development as part of a dynamic, primarily urban process, in which creativity and innovation is a natural, emergent outcome. If Smith et al. are right, how does it impact Jacobs’s thesis and her analytical framework?

JACOBS’S “CITIES-FIRST” THESIS

In the first chapter of her The Economy of Cities (1969) Jane Jacobs argues that densely populated cities based on trade and comprised of diverse, socially distant individuals must logically have preceded the appearance of agriculture. That, of course, is at odds with conventional thinking.

The dogma of agricultural primacy says: agriculture first, cities later. Behind the dogma lies the notion that in pre-Neolithic times hunting men lived only in small and economically self-sufficient groups, finding their own food, making their own weapons, tools and other manufactured goods. Not until some of these primitive groups learned to cultivate grain and raise livestock, it is thought, did settled and stable villages emerge, and not until after the villages were built did complex divisions of labor, large economic projects and intricate social organization become possible. These advances, coupled with a surplus of agricultural food, are supposed to have made cities possible (Jacobs 1969, p. 5).

So trade is the genesis of cities, and cities give rise to agriculture. Although her book is concerned mainly with setting out a controversial theory of economic development and innovation based on the provocative idea of “import replacement,” most of the controversy appears to issue from
this “cities first, rural development later” thesis (Jacobs 1969, pp. 3-48). I have summarized her theory of economic development elsewhere (Ikeda 2012) and will focus here on the aspect of that theory relevant to the cities-first thesis.

Cities drive innovation

Jacobs defines a “city” as “a settlement that consistently generates its economic growth from its own local economy” (Jacobs 1969, p. 262). For her the essence of a genuine city is the way it drives economic development through innovation, and agriculture and husbandry are the earliest examples of major economic innovations. She rejects the orthodox view that hunter-gathering gives way to small, farming settlements that eventually become wealthy enough to construct the kinds of infrastructure we conventionally associate with a city—monumental buildings, etc. (Childe 1950)—because it was so very unlikely to have happened that way.

I have found it helpful to distill her argument in the following way (Ikeda 2012, pp. 65-8).

Central to the process of innovation is experimentation. First, innovation requires repeated experiment through trial-and-error. Failed experiments in a rural settlement are costly, including starvation and death, and the challenge for the innovator is to minimize those negative consequences. Second, in order to experiment, the innovator needs exposure to novel ideas and habits of thought, but in a rural environment, where peaceful association with socially distant people is rare, inspiration for original thinking is hard to come by. Third, a successful innovation could result in a greater division of labor and specialization of knowledge, but as Adam Smith points out, the division of labor is limited by the extent of the market (Smith 1976, pp. 31-6). In order to successfully implement any new method on the supply-side, therefore, the demand for the goods the new method produces must be great enough to sustain its production, which is unlikely in a region of farming settlements unless trade among them is already extensive. Fourth, if innovation does take place it may not spread very far or very quickly, even within the community of the innovator let alone to more-distant communities, because small, rural villages may be hostile to new ideas and new ways of doing things.

Jacobs argues that it makes sense for hunter-gatherers to begin trading with distant peoples by establishing trade centers, which lowers the cost of acquiring a far greater range of resources and tools. Over time, the opportunities for trade attract a larger number of more diverse peoples and the settlement evolves into a city in Jacobs’s sense. In the initial stages of the settlement, the residents are primarily hunter-gatherers trading wild or minimally domesticated goods but as the population grows their knowledge and tastes diversify and grow as well. Here we see the beginnings of sophisticated agriculture and husbandry. Why?

First, with a greater range of work (and begging opportunities) than a rural village, the cost of failure from experimentation is far lower in a city than in a rural village (Glaeser 2012, p. 455). Second, contact with socially distant people is much more likely in a city than in a village, so exposure to different ideas and methods is greater, which is a source of stimulus and inspiration for creative thinking largely absent in a village. Moreover, in dense and diverse populations create opportunities for accidental or serendipitous combination multiply. Third, the larger population of the city-settlement can better sustain an extensive division of labor and specialization resulting from innovations. Fourth, in a dense trading environment, ideas diffuse rapidly both because of trade and because of the relatively greater social tolerance for differences in appearances, beliefs, and practices that a city must have in order to thrive. In short, a city in Jacob’s sense checks all the boxes that encourage innovation, including especially agricultural advance.

The city and agriculture as a spontaneous orders

Specifically, in her story of “New Obsidian” Jacobs explains how agriculture probably emerged in cities before it spread to rural areas (Jacobs 1969, pp. 18-31). It is a story reminiscent of Carl Menger’s theory of the origin of money (Menger 1976, pp. 257-62)—an example of unplanned, spontaneous order driven by self-interest. She imagines a variety of relatively wild goats brought to the settlement by traders, who hire a person to tend them (a new kind of work). In order to minimize his effort, when given the choice the goat tender will slaughter the goats that are the hardest to handle first and reserve the tamest for last. Given the knowledge he acquires from this specialization it would make sense for him then to further specialize by breeding the tamer goats, which in the large, diverse market of New Obsidian could be sold for profit. In this way animal husbandry takes place as an unintended consequence of trade. Jacobs also describes the same kind of process for seed hybridization, and one can see how the application of the “accidental-combined-with-self-interest” story can be applied to new specializations, in different areas including manufacturing, technology, organization, and even religion. All of these ap-
pear to be consistent with Menger’s genetic-causal explanation.

THE ARCHEOLOGICAL COUNTER-ARGUMENT

Among the supporters of Jacobs’s cities-first thesis is Peter J. Taylor, an emeritus professor of geography at Loughborough University, UK. He has developed a concept called “city-ness” which he describes as a combination of cluster/agglomeration processes within cities and network/connectivity processes between cities. These processes create unprecedented communication potentials that make cosmopolitan cities the crucibles of new ideas, innovations and inventions. This is what has made the impact of cities so extraordinary today and in the past (Taylor 2012, p. 417).

I will comment on city-ness, and its role in the debate over Jacobs’s thesis, later in the paper.

Taylor has written a number of articles extending Jacobs’s cities-first thesis, one of which (2012), published in the International Journal of Urban and Regional Research in 2014 provoked the following response from three archeologists—Michael E. Smith, Jason Ur, and Gary M. Feinman (2014, p. 7):

Agriculture preceded urbanism. They did not, however, evolve independently. Settlement and agriculture developed in tandem, often making it impossible to say whether one was a response to the other.

Smith et al.: Agriculture first, cities later

In their response, accompanied with some obvious exasperation, Smith, Ur, and Feinman (2014) report on what they regard as conclusive evidence against the cities-first thesis and by implication against Taylor and a host of others who have developed the thesis (Braudel; Reader; Ikeda).

The Smith et al. counter-argument consists of three parts: (1) a definition of what a “city” is, (2) a somewhat broad conception of what constitutes “agriculture,” and (3) the latest archeological data on the first appearances around the world of cities and of agriculture.

Regarding cities, Smith et al., unlike Jacobs, use a non-processual definition of a city and explicitly cite Louis Wirth’s (1938) essentially functional definition, which characterizes a city as a three-variable problem, with population size, density of settlement, and heterogeneity of population as arguments. In contrast, Jacobs (1961) characterizes a city, we have seen, as a problem of “organized complexity.” Smith et al. dispute Jacobs’s claims about Çatalhöyük, and its alias, “New Obsidian.”

Ultimately what is ‘urban’ is a matter of definition…, but Çatalhöyük does not meet the criteria of either of the major definitions of urbanism used in archeology and history. Louis Wirth’s (1938) influential demographic definition of urbanism requires a high population size and density, coupled with social heterogeneity. As a relatively homogenous village of 15 hectares, Çatalhöyük does not come close to qualifying as urban. The alternative functional definition (Fox 1977; Marcus 1983) requires settlements to have activities and institutions—whether economic, political or religious—that affect a hinterland. Lacking such urban functions, Çatalhöyük does not match this definition either (2014, p. 1530).

As we will see, Taylor (2015) argues that this concept has led Smith et al. to look in the wrong places for evidence of urban emergence.

Regarding agriculture, Smith et al. seem at times to define agriculture as simply “farming,” which includes farm management, irrigation, and tilling.

Long before some plants showed the physical traces of domestication, human communities were managing (i.e. cultivating) morphologically wild plants via tilling, seeding, tending, harvesting and storing (Bar-Yosef 2011, pp. 181-2; Zeder 2011, pp. 224-6). As early as 10000-8700 BCE several signs point to such management (2014, p. 1529).

Although they also include plant and animal “domestication,” what counts as “agriculture” and especially whether it aligns with what Jacobs is referring to when she uses those terms, remains contestable. While I am prepared to concede to the scientific authority of archeologists on what in the professional literature they consider “agriculture,” the Merriam-Webster (online) dictionary offers the following definition:

the science, art, or practice of cultivating the soil, producing crops, and raising livestock and in varying de-
What constitutes agriculture in the professional literature notwithstanding, let us note, with respect to the ordinary meaning of the word, that while early farmers may have cultivated the soil and raised livestock as a practice or an art, it would probably be a stretch to claim that early Neolithic settlers did so systematically and experimentally in the manner of a science. Thus, whether the archeological evidence rebuts the “agriculture-first” thesis would seem to depend crucially on whether Jacobs and the archeologists are talking about the same thing.

Regarding the archeological data, Smith et al. first report on studies that show evidence of animal and plant domestication and breeding in PPNA (8500-7600 BCE) which predates Çatalhöyük, but does this really reflect the kind of innovative agriculture Jacobs has in mind in her cities-first thesis? In any case, they assert that “current understandings are nuanced and suggest complicated processes…the overall picture is clear…”

In the Near East, unequivocal cereal domestication (identified on the basis of morphological changes resulting in reproductive reliance on humans) is clearly attested between 8400 and 7500 BCE for various plant species (Fuller et al. 2012; Willcox 2012).

To further their argument, Smith et al. claim that such changes follow “management strategies” that occurred thousands of years earlier.

It is now clear that morphological changes to plants and animals are not ‘leading-edge indicators’, but rather the result of less archeologically visible management strategies for plants and animals that began millennia earlier (around 9500 BCE for animals and before 10000 BCE for plants…. (2014, p. 1530).

They next argue that the “conventional understanding of urban origins places the first Mesopotamian city at Uruk” (4000-3200 BCE) where recent archeological work “uncovered a series of enormous and ornately decorated buildings.”

To summarize, cities in the Near East emerged over more than a millennium, with initial proto-urban agglomerations around 4400-3900 BCE, unequivocal cities in northern and southern Mesopotamia around 3900-3100 BCE and ubiquitous urbanism in the era of city-states around 2600-2000 BCE (reviewed recently in UR, 2010; 2012). At the start of the sequence, human communities were using an integrated agricultural economy that was already three millennia old (2014, p. 1531).

Thus, they argue that “integrated agricultural economies” predate the first cities by three-thousand years. And this, of course, also hinges on how they define “city.” Three-thousand years prior to the Mesopotamian city-states, for example, would place the emergence of agriculture right about the time of Çatalhöyük, i.e. circa 7000 BCE.

Interestingly, Smith et al. are confident that what they deem conclusive evidence does not detract from Jacobs’s other analytical achievements:

For the most part we limit ourselves to the narrow question of whether the earliest agriculture preceded or postdated the earliest cities within individual world regions. The fact that Jane Jacobs made a basic and elementary error on this question (a relatively minor part of her overall output) has no bearing on the validity or usefulness of her other work. “Her general legacy is not in question.” (2014, p. 1532).

They do not specify, however, what part of that legacy remains. In particular, they do not refer to her theory of economic development via urban innovation and import-replacement as part of that unquestioned legacy.

Hodder: The unintentional emergence of agriculture
In any case, how do archeologists explain the appearance of domesticated plants, for example, in a pre-urban environment? Ian Hodder, who currently directs the excavation of Çatalhöyük, posits a theory that has Jacobsian overtones and that, as I will explain, can actually be seen as consistent with Jacobs’s theory of innovation. In particular, Hodder suggests that

Settlements had increased in size during the warming at the end of the Pleistocene, but the setback of the Younger Dryas forced intensification in order to maintain settled life in agglomerated villages. The result of that intensification was that people domesticated plants and animals—perhaps unintentionally (Hodder 2006, p. 243).
Hodder explains that such an unintended consequence may occur when settlements become more densely populated or when many different settlements locate in proximity to one another.

Indeed, it seems quite possible that people who had come together largely because of the benefits (prestige, exchange, status, control over resources) that this network allowed, ended up “accidentally” domesticating plants and animals. The large agglomerations of people would have depended on a wide range of local resources which increasingly have had to be more intensively collected (just because of the large numbers of people exploiting the same landscape) (Hodder 2006, p. 244).

How might this be related to Jacobs’s overall analysis? Recall that for Jacobs, innovation takes place in the presence of trade, population density, and population diversity, all of which are characteristic of living cities. It is possible that similar conditions might arise, outside of an urban context, in “proto-urban” environments—perhaps Taylor’s “city-ness” network ideas come into play here—that mimic such a context.

THE REBUTTAL

Taylor on “city-ness”

In his response to Smith et al., Taylor (2014) concedes that agriculture preceded the first appearance of cities (a concession that I am not quite ready to make).

My first key point is that I have no disagreement with this evidence; as a non-expert on this topic I can only say that the evidence appears to be soundly based on proven methods (2014, p. 168).

But he does take issue with the concept of “city” that they use. The functional definition of Wirth used by Smith et al. (2014) places too much emphasis on “things”—monumental buildings and such—and not enough on the processes and networks that constitute living cities. He writes:

Following Jacobs (1969) and Castells (1996), I treat cities as a process (economic development) that operates through inter-city relations (networks of cities). This process of ‘city-ness’ (further elaborated in Taylor, 2013, pp. 63—87) is especially conducive to innovation and its diffusion, due to the communication potential within and between cities that totally dwarfs levels of human communication in other types of settlement (2014, pp. 169-70).

Since such networks and processes may not leave lasting traces, so that existing artifacts may be missing what is essential about living cities. He writes:

Basically what I am arguing is that inter-city relations are better described by central flow theory than central place theory. Second (and more important), it is the nature of cities that is at issue. City-ness is built upon a relational approach to understanding cities; archaeologists as represented by Smith and his colleagues use what Jacobs (2000, pp. 32—4) calls a ‘thing theory’—definition by content rather than process. Here we get the only reference to a social science view of cities (Smith, et al. 2014, p. 1530), Wirth’s (1938) three criteria of size, density and heterogeneity. Particularly in the archaeological argument, the ‘things’ emphasized are what are found in early Mesopotamian cities—see, for example, reference to other sites that ‘have not revealed any monument architecture’ (ibid. p. 7)—so that settlements not containing these things are deemed not to be cities, ipso facto Mesopotamia has a head start in claiming the first cities (2014, p. 169).

I find Taylor’s line of argument valid but weak. While it is true that processes and networks themselves may not leave measureable traces, it seems reasonable that artificial evidence of “things” traded over those networks over great distances or sophisticated and rapidly changing tools or commodities produced via those processes would point to the existence of those networks. The burden seems to be on Taylor to specify what kind of artificial evidence would, even if only in principle, corroborate his assertion. Otherwise, his defense of city-ness in this context would appear to render it immune from refutation.

Price and Bar-Yosef: Another archeological view on the origins of agriculture

Despite these reservations, I do favor Taylor’s way of conceptualizing what a city is, which is I think much more consistent with the way Jacobs looks at urban processes as incubators of ideas, experimentation, innovation, entrepreneurship, and economic development. However, as I indi-
cated earlier I will instead take issue with the way Smith et al. characterize “agriculture.”

First of all, their statements about the consensus among archeologists on the first appearance of agriculture appear to be at least somewhat overstated. In their introduction to a symposium held in 2011 on “The origins of agriculture: new data, new ideas” which appeared in *Current Anthropology*, T. Douglas Price and Ofer Bar-Yosef (2011) offer a much broader and evidently more representative view of the archeological evidence on the first-appearance of agriculture than do Smith et al.—one that reflects a great deal more caution. They observe, for example that

There is yet no single accepted theory for the origins of agriculture—rather, there is a series of ideas and suggestions that do not quite resolve the question. At the same time, of course, the evidence we have is scanty and limited. (2011, p. S168).

Elsewhere in their overview they report:

In spite of extraordinary advances in a variety of fields, many detailed at the symposium, we really know very little about the origins of agriculture.[...]We are still at the early stages of the process of identifying and understanding this transition from hunting to farming.

There is another aspect of their summary, which when combined with Hodder's observations on the conditions that may be conducive to domestication, is broadly consistent with the essence of Jacobs's cities-first thesis:

Another commonality among the cradles of agriculture is the rich environments in which farming originates. Experiments in domestication do not take place in marginal areas but amid concentrations of population and resources across the globe. It also appears that in each area where several different species are involved in the transition to agriculture, there are multiple centers of domestication within the region. A number of groups appear to be manipulating their natural world.

So, in various places around the world, plentiful resources and a high diversity of groups or settlements within a region appear to have been most conducive to the emergence of agriculture. Price & Bar-Yosef observe, moreover, that on the list of conditions the participants in their symposium analyzed, wealth accumulation ranks last.

The most important factors in the transition from the perspective of the authors presented here include, in order of suggested importance, available protodomestics, human sedentism, higher population density, resource abundance, geographic and/or social constraints, processing and harvesting technology, storage, and wealth accumulation (2011, p. S172).

This is at least consistent with Jacobs’s argument that high-levels of capital accumulation were more the result of than a precondition for cities. Nevertheless, these areas are still not cities in Jacobs’s sense, and the archeological orthodoxy Jacobs criticized appears to maintain that wealth accumulation is at least a weak precondition for cities.

But the essence of Jacobs’s analytical framework emphasizes high concentrations of people from socially distant backgrounds who are economically secure enough to (unintentionally) experiment with different forms of plants and animals, and that useful discoveries they made would be able to diffuse throughout the region at low cost.

Also, I earlier gleaned from Hodder his observation that agriculture may have first arisen “unintentionally.” In addition, Price & Bar-Yosef point out that

It is important to separate the origins of agriculture from the process of domestication and to distinguish biology from culture in the transition from hunting to farming (2011, S171).

I read this as acknowledging that agriculture and domestication are different things and that they may have arisen at different times, which further muddies efforts to date precisely their origins against the first appearance of cities.

I suggest that by combining these two observations it may be possible to tentatively conclude that, while the archeological evidence shows clearly that plant mutation occurred before the Çatalhöyük became a major settlement, nevertheless in order for the agricultural innovation and diffusion to take place on the scale that Jacobs describes still requires the urban process of a Jacobsian city.
It seems to me that there are three ways to respond to Smith et al.’s argument. The first to question, as I have just done, whether their definitions of city and agriculture agree with those of Jacobs and Taylor.

One could also investigate, if indeed Smith et al.’s critique is definitive, what part of her “general legacy is not in question.”

Thus, a second way to respond to them would be, again as I have just done, to suggest that Jacobs’s framework, in which density and diversity give rise to processes that foster innovation, might be applied to a non-urban situation. While this response abandons her cities-first thesis, it retains the essential elements of her explanation of how innovation of all kinds happens, for which Jacobs’s account of “New Obsidian” is really just a paradigm. Thus, while her broader thesis about the role of cities in innovation may not be true of the earliest forms of agriculture, it may well be valid for a wider class of innovations, e.g. literacy, numeracy, calendrical sciences, hierarchical governance, monumental architecture, and organized religion, and so on, all of which are part of V. Gordon Childe’s ten characteristics of what a city is.

Similarly, a third response might be to again abandon her specific thesis, but to argue that, while millennia of incremental change may have resulted in some domestication (perhaps a la Hodder) that long predate the appearance of cities (rightly understood) Jacobs’s framework would suggest that the existence of those cities exponentially accelerated the process of innovation of all kinds, including agriculture. [Adam Smith argued that compared to manufacturing agriculture would see comparatively little innovation.] This suggests that we look for evidence of a significantly greater variety and volume of goods consumed after true cities were established and that likewise agricultural innovation accelerated and spread quickly via trade throughout any given region.

**NOTES**

3. Emphasis added S172
4. Emphasis added S169

**REFERENCES**


Extraordinary Cities: Early ‘City-ness’ and the Origins of Agriculture and States. Also online at http://www.lboro.ac.uk/gawc/rb/rb360.html
Fostering emergent diversity: directing rather than mastering change
The recent theory of planning and urban design recognizes how healthy, vibrant cities behave as complex living systems: they are open, bottom-up, self-organizing, emergent (Allen and Sanglier 1981; Christiaanse, Rieniets and Sigler 2009; Dovey 2012; Portugali 1999; Sennett 2013). The main characteristic of an emergent system is that it cannot be predicted in advance because it is in a constant state of becoming. Such a system arises from unforeseen interactions rather than being determined by an a priori intention. It is resilient because disturbances can be accommodated by fluid adjustments: in fact, the heterogeneous components of the system retain the capacity to endlessly re-organize themselves into emergent interconnections (see section 3).

Jane Jacobs (1961) anticipated many years ago that a city is a problem of organized complexity and that diversity is an essential condition for its vitality and resilience. She highlighted how places which behave as complex living systems retain an endless capacity to encourage unforeseen exchanges between disparate people and how this concentration of people from different backgrounds and their constant interaction stimulates economic growth, innovation, and cultural emancipation (Jacobs 1961; Christianse 2009).

Despite this awareness, the theory and practice of planning and urban design often fails to elaborate strategies that can cooperate with emergent urban systems and foster city diversity. While, in planning theory, the relationship between complexity and regulation is being thoroughly investigated (Moroni 2015), the relationship between design and complex living systems remains underexplored. In fact, urban design theory and practice are still widely based on master plans, which aim at predicting and controlling the development of specific areas. Master plans tend to propose a new comprehensive vision of the future. They impose a new overall form—considered more healthy and efficient—on an existing reality, considered problematic or obsolete. Master plans usually present themselves as finished projects designed from the top-down: they show the final outcome of the transformation within a pre-defined bounded site. In their best examples, they are animated by a sincere intention to produce a regime of urban complexity able to generate diversity. Nonetheless, they remain external interventions in an otherwise spontaneous and complex urban process, thus they cannot cooperate with emergent urban systems, which are instead bottom-up and self-organizing phenomena (Porqueddu, forthcoming).

The contrast between comprehensive master planning and the city as a problem of organized complexity was similarly highlighted by Jane Jacobs (1961), who argued that...
modern top-down planning strategies were destroying diversity in American cities. Jacobs highlighted how vital cities require “an intricate and close-grained diversity of uses that gives each other constant mutual support, both economically and socially”. According to Jacobs, “the science of city planning and the art of city design, in real life for real cities, must become the science and art of catalyzing and nourishing these close-grained working relationships” (14). For Jacobs diversity cannot be entirely planned or designed because it is an emergent property of the relationship between people and their physical environment, rather than the mere consequence of a peculiar spatial layout or of a certain density of pre-defined functions (Porqueddu 2015; Ikeda 2017).

Furthermore, she highlights how top-down planning and design are not the only threats to emergent diversity and she stresses how diversity can spontaneously veer toward its self-destruction. In this respect, she advocates the need for a kind of planning and design which does not interfere with the spontaneous emergence of diversity, but at the same time, intervenes in order to prevent its emergent decline. She suggests that the role of designers and planners is not shaping the city from the top-down, but monitoring the unpredictable evolution of emergent orders within a city and understanding when an action is indispensable for inverting a negative cycle, in order to foster or maintain “city vitality (something that the planner or planners alone, and the designs and designers alone, can never achieve)”(14).

The aim of such planners and designers is not fixing the problems through external interventions, but rather becoming acquainted with the behaviour of people across space, in order to make more efficient use of their natural capacity to maintain and increase diversity. In this respect, they need to observe more and design less. This attitude requires research methods and approaches, which make it possible to question places, in order to detect—within emergent transformations—the risks for the self-destruction of diversity and to uncover latent capacities for its emergence.

In order to shape such a method of inquiry, the present paper builds upon Jacobs’s phenomenological social-spatial approach, but it extends it to a multi-scale level. In fact, a multi-scale understanding becomes indispensable in the contemporary city, where the rise of mobility and communication technology has brought about a network of exchanges between discontinuous places—opening up new possibilities for interaction and exchange beyond the traditional relationship of proximity (Harvey 1990; Massey 1994; Castells 1996; Amin and Thrift 2002; Boelens 2009).

In the domain of this new metropolis, diversity emerges according to the distinct mixture of both local and wider flows intersecting at specific points in time (Massey 1994). In this respect, the mere scale of architecture would not highlight the social-spatial dynamics beyond the boundaries of a specific site and the wide metropolitan scale would not reveal anything about the micro-space scene of perception and human interaction.

In order to shape this multi-scale approach, this paper combines:

1. the recent advance in theories of emergence, complexity, assemblages, adaptation (Allen and Saliger 1981; Portugali 1999; Dovey 2012; De Landa 2006; Gunderson and Holling 2002; Miller and Page 2007; Walker and Salt 2006), which inform an understanding of adaptive cycles across scales.
2. The current advances in mapping techniques (Corner 1999), which can visualize unforeseen relationships between emergent urban form and heterogeneous data concerning the behaviour of people across specific territories and their emergent ways of using and transforming the existing urban spaces.

By integrating Jacobs’s phenomenological social-spatial approach with this multi-scale understanding, the current paper endeavours to formulate a method of inquiry which not only helps us understand the city as an emergent system, but also enables designers and planners to detect place-specific latent capacities for emergent diversity and to uncover the risks for its self-destruction. Here this approach is considered fundamental to developing design and planning strategies which cooperate with emergent transformations rather than trying to master them. In this respect, the present paper (1) traces a strong (but non-linear) connection between place analysis and design practice; (2) shapes a method of inquiry capable of informing every design project that tries to cooperate with the unpredictable nature of cities and their adaptive cycles across scales.

The article is divided into five sections. The first section illustrates the contrast between master planning and the unpredictable emergent nature of city vitality through two examples. The second highlights how the recent advance in theories of emergence, complexity, adaptation and assemblage can inform a multi-scale approach. The third section introduces the Multi-Scale Atlas as a tool for such social-spatial multi-scale investigation. The fourth section illustrates how the Multi-Scale Atlas explores (1) the cross-scale
effects of emergent changes, (2) the place-specific conditions which foster and threaten emergent diversity across scales, and (3) how this understanding can inform the actions which increase emergent diversity and contrast its decline (this section refers to the case study where the Atlas has been developed). The last section draws general conclusions.

**The contrast between master planning and emergent urban systems**

If we fully embrace the idea that diverse vital cities behave as living systems, we must also accept their unpredictable nature: diversity can emerge and self-destruct in different places and situations far beyond the imagination of a single human being or of a group of them. In this respect, there can be no fixed cause-effect relationship between the emergence of diversity and certain specific built form or density in terms of FAR (Floor Area Ratio) or a certain a priori defined combination of uses (Porqueddu 2015). Even more in the contemporary metropolis crossed by flows of people who travel beyond the boundaries of specific sites.

If diversity cannot be the direct result of human design (Jacobs 1961), then the city proposed by New Urbanism (Duany, Plater-Zyberk and Alminana 2003), apparently based on Jacobs’s arguments, is instead linked to the modern illusion that an a priori top-down designed spatial configuration can generate diversity (Porqueddu 2015). In fact, New Urbanism elaborates a new model of the metropolis, rather than focusing on upgrading and updating the everyday experience across the urban fabric of the American suburban sprawl, which it tends to reject and replace. New Urbanism superimposes a top-down designed new city on a reality considered negative. While the aim is laudable—to solve the problems linked to urban sprawl and to upgrade the everyday experience of these places, which are in fact often desolate and lacking in diversity—this approach is still based on the illusion that the problem can be fixed by external agents, designers and planners, who apply a new and a priori designed spatial configuration onto an existing context.

Nonetheless, there are other ways of understanding urban design, animated by the intention to observe the existing reality and to cooperate with emergent orders. These approaches completely abandon the modern illusion that it is possible to fix the actual situation from the outside by replacing it with a better one designed a priori. Instead, they aim to provide a specific place with the missing ingredient which can activate its ability to evolve in positive directions and to self-produce the solutions to emergent problems, without trying to predict and control the final, formal outcome of the transformation (Porqueddu, forthcoming).

A contemporary example concerns the realization of the Metrocable network in Medellin—as part of the PUI (Proyectos Urbanos Integrados; Integrated Urban Projects), the complex programme of city transformation promoted by the Medellin municipal government and coordinated by Alejandro Echeverri (Echeverri and Orsini 2010). The Metrocable basically endeavours to reconnect and re-integrate the poor, informal sectors of Medellin with the rest of the city through selective intrusion into their social system and minimal damage to their existing structures. The design idea emerges here by detecting the spatial conditions which can foster or threaten emergent diversity. A deep knowledge of the local emergent orders is developed through the collaboration between architects, planners, community representatives, social workers, city administrators and the private sector (Davila 2013).

This detailed observation revealed that these informal settlements have an important resource: a peculiar and vital network of micro-connectivity, which supports the emergent social and economic relations on a local scale. Nonetheless, it also revealed that one of their limits—beside the well-known problems related to extreme poverty and lack of services—is that this labyrinthine structure is not inserted into the network of fast connections on a metropolitan scale and is thus cut off from potential exchanges with other places and therefore from the potential for emergent diversity. In this context, the Metrocable project is a creative solution that can provide fast connections which, in turn, can enhance rather than damage the existing slow micro-connectivity (Porqueddu, forthcoming). In actual fact, the adopted strategy enables to: (1) build fast and direct connections between these informal settlements and the wider metropolitan territory without damaging the existing slow micro-connectivity; (2) turn the steep terrain of the mountains into a potential for a stunning panoramic view over the whole city; (3) utilize the Metrocable stations as hubs to cluster a variety of cultural spaces and sport facilities, thus partially compensating for the lack of services in the area; and (4) reduce social and economic costs (it eschews the demolition of a large number of dwellings in order to build roads); (5) build the Metrocable over a relatively short period of time as it requires little in terms of land acquisition (Porqueddu, forthcoming). In this respect, the Metrocable is the missing ingredient that can incorporate these informal settlements into a new system of rela-
tionships. In this case, an effective collaboration involving planners, designers and the public administration generates an intervention which is not intended to re-design the informal settlements (by proposing a new order from the top-down). Instead, it endeavours to provide a cross-scale connection which can prevent the self-destruction of diversity across the area and increase the existing potential for its emergence. In this respect, by connecting the existing vital slow micro-network with the wider metropolitan territory, the Metrocable activates the site-specific ability for unforeseen exchanges with the rest of the metropolis, without trying to pre-define or control them.

Around the world, there are some scattered examples of this approach to design. In another forthcoming paper, I begin to collect and compare them, exploring what they have in common, in order to start developing a theory of emergence in urban design. On the one hand, they are all extremely different, operating at different scales and belonging to very distant geographic areas and heterogeneous urban situations: these strategies must necessarily be unique, being creative responses to immanent and site-specific conditions. On the other hand, they do have something in common: they are based on a deep understanding of the place-specific social-spatial dynamics across different scales and they are intended to make more efficient use of the spontaneous capacity of places to maintain and increase their potential for diversity. In this respect, although these approaches do not even refer to Jacobs's theory, they are consistent with it, being based on continuous minimum adjustments and real life observation rather than on overall comprehensive plans based on abstract principles. Jacobs affirmed that cities are themselves suggesting to the better planners and designers (the ones who are able to listen and carefully observe complex processes) “what principle of planning and what practices in rebuilding can promote social and economic vitality in cities, and what practice and principles will deaden these attributes”(4).

In this respect, these kinds of intervention highlight the urgent need for theories, methods and tools aimed at questioning places in order to understand their emergent complex orders. The multi-scale approach presented here has emerged in order to respond to this need for methods of inquiry, which are an integral and fundamental part of such a design process.

Understanding emergent urban systems: a Multi-Scale Approach

The following section illustrates how theories of complexity, emergence, self-organization, assemblage, adaptation can support a deeper understanding of cities as complex living systems and can consequently inform possible actions consistent with their emergent nature. In actual fact, complex systems have been explored in disciplines which are not directly linked to urban studies. Non-linear systems have been investigated by biologists, mathematicians, and physicists; by ecologists such as Gunderson and Holling (2002) and Walker and Salt (2006); and by social scientists such as Miller and Page (2007). Furthermore, Complex Adaptive Systems (CAS) theory has recently been applied to urban studies by Kim Dovey, who combined it with Assemblage Theory (De Landa 2006; Deleuze and Guattari 1980; Dovey 2010; 2012), in order to frame and explore places as socio-spatial territories in a continuous state of becoming and cities as emergent systems (Johnson 2001).

CAS theory and Panarchy (Gunderson and Holling 2002) are particularly relevant because they show how complex living systems evolve in time, across different scales and how it is necessary to understand their cycles in order to prevent their decline. On the one hand, these theories stress that diversity and redundancies of parts are critical to a system’s vitality and resilience: in diverse systems no single part is crucial to the success of the whole; therefore the system can endlessly adapt to unpredictable situations by re-organizing itself into emergent inter-connections. On the other hand, they also show how diversity and redundancy can self-destruct across adaptive cycles and how cross scale effects play an important role in this process. These recurring cycles consist of four phases: rapid growth, conservation, release, and reorganization (Gunderson and Holling 2002; Walker and Salt 2006).¹

Nonetheless, when we talk about adaptive cycles in cities it is easy to become too focused on the specific scale in which we are interested (a building, a street, a plaza, a block, a neighbourhood) and ignore cross-scale effects. According to CAS and Panarchy theory, the crucial point is in fact to consider that the scale in which we are interested is connected to and affected by what is happening at the scales above and below, both in time and space, and that the linkages across scales play a major role in determining how the system is behaving on another scale (Guderson and Holling 2002). In this respect, if we fully consider the city as a complex living system, we cannot successfully interact with it by focusing on only one scale.²
Long before these studies, Jacobs (1961) anticipated the tendency for outstandingly successful diversity in cities to destroy itself across time cycles and she argued that often the same forces which nourish diversity could contribute to its destruction. Her observation also revealed how linkages across scales are a key aspect in understanding diversity and resilience cycles. In fact, she argued that streets which experience the self-destruction of diversity after a successful period (late conservation phase), can quickly regenerate their diversity (from release to rapid growth), only if they are surrounded by other streets which are in a phase of flourishing diversity (rapid growth). That is to say that in this case a micro-cycle can be positively affected by wider scale processes.

Although CAS and Panarchy theories are not mentioned in Medellin project, the Metrocable strategy is based on a deep understanding of cross-scale links: the potential to retain diversity at the micro-scale of the neighbourhood was enhanced by tracing new connections with the larger scale of the city; that is to say that a widescale action was undertaken in order to retain and improve city vitality on a micro scale. In this respect, CAS and Panarchy theories offer a good approach for identifying the appropriate scale of the intervention according to a multi-scale understanding of diversity cycles.

In order to develop such a multi-scale approach, it is necessary to investigate the dynamic relationship (dance) between the physical components of urban space and the existing and emergent behaviour and activities of people across them. In fact, as a set of material components, the city is a simple system (Portugali 2013). Buildings, roads, bridges, fences, traffic lights, sidewalks, trees, shop windows, do not interact without people—they are just scattered elements of potential living systems emerging in the presence of its changing components: people, behaviour, needs, desires, tastes, actions, money, electricity, water, etc. (Porqueddu 2015; Dovey 2010).

The Multi-Scale Atlas presented in the next paragraph, combines Jacobs’s social spatial understanding with a multi-scale approach, which becomes indispensable in the domain of the contemporary metropolis, where our everyday life increasingly develops far beyond the boundaries of a specific settlement and where the emergence of diversity is increasingly related to mobility, time cycles and event geographies. In this respect, the Atlas enables us to understand heterogeneous micro-spaces as components of emergent socio-physical networks stretching out of local boundaries rather than as parts committed to local units (Porqueddu 2015), thus revealing unexpected potential for diversity to flourish.

A Multi-Scale Atlas

In order to explore the multi-scale nature of emergent urban systems, a Multi-Scale Atlas is here proposed as a tool for an empirical investigation which develops a social-spatial understanding across scales.

In this respect, the Atlas combines a multi-layer analysis with a multi-scale mapping in a matrix (Figure 1) which makes it possible to explore the links between people’s behaviour, activity rhythms and the physical layout that supports them at multiple scales. Here, the focus is not on individual maps, but on the actual and potential relations between them (Dovey 2010; Dovey and Wood 2015; Corner 1999). The Atlas faces the following questions: how can the existing and emergent spatial layouts support emergent diversity across scales? How do they foster its decline? The matrix is organized into five thematic strips, according to the following five layers (Porqueddu 2015):

1. **Intensity and rhythms**: It shows daily concentrations of people in specific places according to different time cycles, event geographies and commuter flows. This layer is important because, as previously highlighted, diversity emerges from a distinct mixture of heterogeneous people crossing in space (Massey 1994; Jacobs 1961).

2. **Activities and uses**: It shows the distribution, concentration and types of activities across the territory. The mix of heterogeneous activities and uses is considered one of the main generators of diversity.

3. **Network of connections**: It shows how the network of streets, parking lots, sidewalks, etc. supports fast and slow flows of people. Accessibility is considered one of the main factors which fosters emergent concentration of people.

4. **Boundaries and private-public interfaces**: It shows how the layout and types of physical borders encourage or discourage unforeseen interaction between heterogeneous activities and between public and private spaces. This layer is important because complex living systems require open unfixed interactions between heterogeneous components.

5. **Building footprint and density**: It shows the concentration and mix of building types combined with data on density in terms of Floor Area Ratio (FAR). This layer is relevant because density in terms of Floor Area Ratio...
Ratio and compact urban form are still central in the discussion on city diversity.  

Every layer is investigated on a wide range of scales, within the framework of:

1. 15 m and 15,000 m². This scale shows how space is perceived and how it supports face to face interaction between people. This is the scale of architecture and design.
2. 1.5 km². This scale shows whether the spatial layout—at the scale of a village or neighbourhood—supports people movements and their interaction across different activities and spaces. This is the typical scale of urban design.
3. 300 km². This scale shows the main infrastructures in relation to the existing urban fabric, natural environments and administrative boundaries, and it highlights the flows between different settlements. This is the typical scale of planning.
4. 15,000 km². This scale situates this territory within the wider scale. It is typical of regional planning and geography.

These maps enable us to study specific micro-spaces in relation to a wider network of relationships. In this respect, they can help frame the scale and type of problem by highlighting the multi-scale dynamics in which micro-spaces are immersed.

In order to frame the scale of the problem, and develop an appropriate action, it is necessary to observe existing territories from different points of view. In this respect, the Atlas also aims to become a cross-discipline tool for intersecting data collected through investigation from different disciplines such as geography, social science and anthropology, landscape urbanism, etc. In actual fact, geographers and planners collect data about people and their emergent activities, but they often show these data through widescale maps, which do not illustrate the physical urban environments (on a micro scale) where these activities and businesses take place. Architects often focus on the human scale, but they rarely observe the behaviour of people across different scales. Social scientists and anthropologists observe human behaviour and informal practices, but they rarely map them (and mapping is the only tool that can draw connections between space, people behaviour and flows, emergent mixes of uses and activities). Landscape ur-

<table>
<thead>
<tr>
<th></th>
<th>15 mt</th>
<th>15000 m²</th>
<th>1.5 km²</th>
<th>300 km²</th>
<th>15000 km²</th>
</tr>
</thead>
<tbody>
<tr>
<td>intensity - rhythms</td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
</tr>
<tr>
<td>network of connections</td>
<td><img src="image6.png" alt="Image" /></td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
<td><img src="image9.png" alt="Image" /></td>
<td><img src="image10.png" alt="Image" /></td>
</tr>
<tr>
<td>boundaries</td>
<td><img src="image11.png" alt="Image" /></td>
<td><img src="image12.png" alt="Image" /></td>
<td><img src="image13.png" alt="Image" /></td>
<td><img src="image14.png" alt="Image" /></td>
<td><img src="image15.png" alt="Image" /></td>
</tr>
<tr>
<td>activities - uses</td>
<td><img src="image16.png" alt="Image" /></td>
<td><img src="image17.png" alt="Image" /></td>
<td><img src="image18.png" alt="Image" /></td>
<td><img src="image19.png" alt="Image" /></td>
<td><img src="image20.png" alt="Image" /></td>
</tr>
<tr>
<td>FAR - building footprint</td>
<td><img src="image21.png" alt="Image" /></td>
<td><img src="image22.png" alt="Image" /></td>
<td><img src="image23.png" alt="Image" /></td>
<td><img src="image24.png" alt="Image" /></td>
<td><img src="image25.png" alt="Image" /></td>
</tr>
</tbody>
</table>

*Figure 1. Multi-Scale Atlas (Matrix example)*
banists are the masters in using multi-scale mapping. This is because they work with living materials and they need to understand eco-systems (which are emergent systems) and life-cycles, but they mainly focus on natural elements and they rarely use mapping as a tool to investigate emergent city diversity (cultural, economic, social, etc.). In this respect, the Atlas aims at cutting across all these disciplines in order to welcome heterogeneous types of data, both qualitative and quantitative.

The maps amalgamate the data collected:

1. Through interviews (with inhabitants, administrators and associations).
2. Through behavioural-photographic surveys (the fieldwork required the direct experience of the mapped areas including travelling across the territory at different speeds—driving, walking, public transport experience—and engaging with people while observing their behaviour across space).
3. Through data analysis on commuting and tourist flows and on the distribution of activities across the territory (data collected from ISTAT- Census 2001).
4. Through multi-scale mapping.

The Atlas has been developed across a low density Italian area, called the Oltrepo’ Pavese, a fine grain network of old settlements and more recent nodes, situated 50 km south of Milan (within the Province of Pavia) at the intersection of two important infrastructures connecting Milan, Genoa, Turin and Bologna. The Province of Pavia is a mosaic of 190 micro-municipalities, 172 with less than 5,000 inhabitants: 43% of the population lives in these old traditional settlements at a distance of 2-7 km from each other (Figure 2). This pattern—significant local differences notwithstanding—is prevalent throughout the whole Italian territory, where 72% of municipalities have fewer than 5,000 inhabitants and 43.7% of Italians live in such settlements (Census 2001). The Area of Observation (300 Km²) is situated at the border between the plain and the hills—agricultural and wine-growing areas—and includes eight of these municipalities (Figure 3). The major centre is a small town of approximately 40,000 inhabitants, the others being villages of 900 to 5,000 inhabitants.

Figure 2. The Province of Pavia. (a) Administrative boundaries; (b) Network of connections; (c) Built-up areas
One of the reasons for selecting the Oltrepò was that the importance of a multi-scale social-spatial approach is more evident through a low-density territory where a traditional observation would never highlight the emergent urban complexity of this territory, which is visible just through a social-spatial exploration across scales. In actual fact, one single settlement here is neither large nor dense enough to generate diversity within its boundaries. The only possibility for flourishing diversity here is related to temporary concentrations of locals and strangers—coming from discontinuous places—within the same urban fabric. In fact, local street life can be here nourished just by wider flows, developing on a metropolitan scale. Jacobs herself argued that even a small town can become a complex realm, once it “is encompassed in a metropolitan orbit with its multiplicity of choices and complexity of cross uses” (Jacobs [1961] 1992, p. 435).

Instead, by focusing also on the observation of activity rhythms and on the behaviour of people in space and time (how they use the territory in their everyday lives), the Atlas highlights how this territory is becoming far more complex than a sum of small settlements depending on a wider metropolis. In fact, data analysis, interviews and behavioural surveys reveal the presence of considerable everyday horizontal flows between small municipalities and from major centres toward smaller settlements (Figure 4). In this respect, when observed on a wide scale, this territory behaves as an emergent low-density distributed system, an open network where every single settlement has the potential to become a node connected with several other nodes. In such a system, hierarchy is always shifting: every node could become the centre according to time cycles and event geography—a “meeting place” (Massey 1994) behaving as a component of socio-physical networks stretching out of local boundaries. In this respect, the Atlas shows that even a settlement or suburb which might look neither large nor dense enough to generate diversity within its boundaries, could in fact do so across its urban fabric, according to temporary concentration of locals and strangers—coming from other settlements (Porqueddu 2015).7

Figure 3. (a) Observation perimeter (300 Km² frame): Eight municipalities; (b) Aerial photo
The specific sites for small scale investigation were selected according to the information collected during interviews and data analysis, which made it possible to identify, across the case study area, the places which own the potential to become nodes or which already behave as nodes in the network (Figure 5).

Figure 4. Major settlement: Commuter flows (Mon-Fri). (a) Outgoing commuters; (b) Incoming commuters.

Figure 5. Six sites
Detecting risks and potential of emergent transformations

The Atlas illustrates how the emergent change can generate positive and negative effects (here the transformations which foster emergent diversity are considered positive, while the transformations which foster its decline are considered negative).

On a wide scale, the Atlas highlights the presence of a new layer of distributed metropolitan attractions (such as hotels, spas, sport centres, new workshops, shopping outlets, shopping centres, important showrooms, famous restaurants or clubs) that can catalyze metropolitan flows of people. These attractions are spread across the entire territory, generating a new intensity of flows in and out of small municipalities. Nonetheless, the maps also reveal how these new activities can increase or decrease the potential for emergent diversity and how this is chiefly related to their distribution across the micro-scale spatial layout.

In actual fact, the Atlas illustrates the presence—at the micro-scale of architecture and urban design—of two different spatial layouts. The first encourages an endless exchange between metropolitan flows and place-based street life, between regional attractions and local businesses; the second discourages any unpredictable synergy between latest hubs and existing activities. In actual fact, the first tends to orient the emergent flows of people across traditional settlements, thus increasing the potential for emergent diversity and for new flourishing activities. The second tends to orient these regional flows toward top-down designed unalterable precincts, which are totally disconnected from the local network.

The potential for emergent diversity within small settlements is increased when these latest hubs are placed at a walkable distance from local activities—such as the bakery, the butcher, the newsagent, the primary school, the church, etc.—and, more important, when they co-exist and interact with them in space and time. Figures 6, 7, 8, 9, show that this occurs according to the following spatial conditions (Porqueddu 2015):

1. Metropolitan attractions are interwoven with local activities across the fast-slow network of connections (Figure 6).
2. This mix of local and metropolitan activities is situated at a walkable distance from regional arterial roads (Figure 6-7).

Figure 6. Local-metropolitan interaction
Figure 7. Multi-scale Atlas. Activities-uses
Figure 8. Multi-scale Atlas. Network of connections and parking lots
Figure 9. Multi-scale Atlas. Boundaries, public-private interfaces
3. Small and medium size parking lots are scattered at the edges of the settlement as well as across its fine-grain urban fabric; their strategic position and size encourage people coming from other places to leave their cars and walk through the settlements. This increases the pedestrian traffic in front of local shops—thus nourishing their businesses—and plugs new flows of strangers across the small settlement, thus fostering the potential for random and unpredictable face to face interactions, encounters and exchanges (Figure 8).

4. New activities are not committed to wider precincts or, if they are part of a wider precinct, then they retain a certain degree of autonomy in terms of rhythms and management (if one of these closes at 6 p.m., the others can remain open until 3 a.m.). Such precincts are provided with multiple access which integrate them into the external fast-slow network of connections and parking systems. They have diverse and non-definitive borders—single units can overlook both the inner pathways and the outer sidewalks. Accesses and private-public interfaces can change according to different time cycles and adapt to unpredictable conditions, needs and behaviour, thus fostering unforeseen synergies between heterogeneous activities (Figure 9).

The potential for emergent diversity within small settlements is decreased when the latest hubs are disconnected from local activities. Figure 10 shows that this occurs when:

1. New metropolitan attractions are enclosed within precincts and committed to them. The boundaries of these precincts are impermeable and definitive. All the activities overlook the internal part of the precinct, with no possibility for interaction with the external sidewalks or other local activities, even when they are adjacent to them. Furthermore, different activities within a single precinct do not retain any autonomy in terms of rhythms and management and in terms of spatial layout. They are strongly top-down pre-defined and controlled: there is no possibility of adaptation and no possibility for unpredictable interactions between heterogeneous businesses, also within the precinct.

2. These precincts have only one access, connected with a large private parking lot overlooking fast arterial roads, without any connection with the local network of activities, even when they are placed at a walkable distance from them. Flows of people travelling on a regional scale can directly reach these activities without walking through the settlement, which remains cut off from these new flows. In this respect, these new metropolitan flows do not increase the potential for encounter and exchange (diversity) within the traditional small settlement.

Figure 10. Local-metropolitan disconnection
On the one hand, the Atlas highlights how this territory, when observed on a wider scale, retains the potential to behave as a complex distributed metropolis, where every front door is a threshold between a potentially vibrant and diverse urban street and the peaceful slow rhythm typical of a country village (Lynch 1984). This could evolve into a low density rural metropolis where every single settlement might become a node able to combine: (1) the intensity and diversity of the city; (2) the sense of safety and calm of the village; and (3) the experience of nature in the countryside. On the other hand, it also highlights the emergent spatial conditions, which tend instead to crystallize a separation between old settlements and new hubs, thus contrasting emergent diversity. Nonetheless the Atlas shows that local businesses and metropolitan attractions can also coexist and interact within small settlements and that this interaction can enable these settlements to remain open, adaptive and diverse (Figure 6, 7, 8, 9).

In this respect, the Atlas highlights how the challenge of the local administrations, planners and urban designers consists in shaping strategies, which could (1) encourage the presence of these new activities within the urban fabric of existing villages and new hubs, thus contrasting emergent phenomena can neither be planned nor designed, these strategies should never aim at producing a top-down definition of the types of activities and their place within existing villages and across the wider network. Instead they can focus on:

1. Understanding the reasons which encourage the owners of the new businesses to place them outside the existing settlements.
2. Shaping a series of actions which can invert this trend and render the inner part of the settlement appealing for new activities and businesses. These actions should not interfere with the spontaneous tendency of metropolitan businesses to spread across this territory, but at the same time it would transform their negative effects into new potential synergies with existing local activities.
3. Revise the network of streets, in order to better integrate the slow micro-network of the villages into the regional network of fast arterials.

Some of the reasons which might induce the owners of businesses to situate new activities far from the old villages concern the spatial layouts, other the planning system. With regard to the latter, building and planning regulations play a crucial role in orienting the position of metropolitan attractions. In fact, in Italy rules tend to be more restrictive within historical centres (Moroni 2015). While the intentions of these restrictions are valid (to protect the aesthetic value of these urban areas), often they can also prevent the adaptations necessary to update the historical buildings and render them suitable for emergent activities and uses.

Furthermore, the procedures to obtain construction or renovation permits are more complicated when the project is developed in historical buildings or within the historical urban fabric. The overall bureaucratic process is usually longer and thus entails major economic costs (which are already higher due to the greater value of lands and buildings within the historical nuclei).

With regard to the spatial layout, the old nuclei might present problems of accessibility. On the one hand, they usually offer an intricate network of micro-connectivity, which can foster unforeseen synergies between heterogeneous activities and between private and public spaces, but often this intricate structure is not integrated into the network of fast arterial roads. On the other hand, the new nodes situated along them are often isolated from this micro-network, even when they are adjacent to it (Figure 10).

The question then becomes: what innovations concerning the system of rules and the spatial layout are necessary in order to facilitate the adaptation of the existing urban fabric (recent and historical) to emergent changes, while enhancing its aesthetic quality? The system of rules is not investigated in this paper, although it is crucial to managing city complexity (Moroni 2015). In regard to the spatial layout, the main challenges—for urban designers and architects—are:

1. Experimenting innovative spatial and infrastructural solutions (place-specific) that can integrate the historical urban fabric into the regional metropolitan network, without generating traffic problems or other kinds of conflicts within single settlements. This entails the revision of the current infrastructural network and of the layout and distribution of parking lots across the existing urban fabric. Here the creativity of designers is crucial in order to turn existing problems into opportunities to update and upgrade our daily experience of urban space (functionally, technically and aesthetically).

2. Experimenting innovative spatial solutions, materials and technologies, which make it possible to adapt existing buildings and urban spaces to emergent ac-
vities and uses. This type of innovation concerns soft interventions which enable fast adaptations with minimum costs. The techniques which facilitate high quality transformations with minimum structural and architectural interventions are particularly interesting. In actual fact, they can update the existing urban fabric of the historical centres, by enhancing their aesthetic quality. (Interior design and landscape architecture are the fields which better explore technique for “soft” transformation).

3. Experimenting innovative techniques which could provide existing and new buildings and precincts with responsive, reversible private-public interfaces and accesses that can easily adapt to unpredictable situations and unforeseen individual-collective needs. As previously mentioned, definitive boundaries are one of the main obstacles to emergent synergies between different activities and between private and public spaces.

While point one mainly concerns the action of public administrations and urban designers, points two and three include the work of architects and interior designers for private investors.

Learning from places: understanding and managing multi-scale adaptive cycles in the contemporary metropolis

The present paper presents the Multi-Scale Atlas as a tool for understanding complex diversity cycles across the wide territory of the contemporary metropolis and it stresses how this understanding is crucial for informing every design-planning strategy which aims at cooperating with emergent transformations. In actual fact, such design strategies are inspired from the self-ability of existing urban environments to remain dynamic, adaptive and diverse. The idea is to interpret the inventions that life brings to the city every day and to work as much as possible with (and not against) the place-based forces which shape them. By accepting the city (and its inhabitants) as co-authors of their project, such designers also agree to limit their interventions as much as possible.

The multi-scale approach presented in this paper highlights how such strategies entail a deep understanding of cross-scale effects and how the multi-scale atlas can inform this understanding. The Atlas highlights the presence of two main levels of understanding (and consequently of intervention) and makes it easier to explore the dynamic between them.

The first level concerns the micro-scale of individual initiative. This is the small scale of fast changes which derives from the will of citizens to endlessly transform existing spaces according to their emergent needs. This level also concerns the ability of architects and interior designers to form innovative solutions that can update the existing urban fabric in a short time. This first micro-scale level concerns two types of intervention:

1. “soft” interventions of design (concerning the revision of the inner distribution, new finishing, lighting design, but also the restoration or restyling of existing facades and the adaptation of public-private interfaces) able to endlessly re-interpret the existing buildings for new purposes and to adapt them to the new technological requirements, with minimum structural changes. The paper stresses how these micro-interventions can have a huge incremental effect on the metropolitan scale. Within the analyzed case study, they can become powerful agents of adaptive reuse of the existing urban fabric, thus reducing the need for urban expansion across the countryside.

2. the construction of new buildings and precincts according to emergent activities and needs. These transformations are slower than the previous ones, but also crucial for the evolution of the existing urban fabric as well as for the quality of urban growth across the territory. The Multi-Scale Atlas highlights how the new buildings and activities can foster emergent diversity or its self-destruction according to their position across the existing urban network and the design of their access and public-private interfaces.

The second level concerns the ability of the public administration to establish a frame (both spatial and normative) which can guarantee that the unforeseen individual interventions of the first level foster city diversity, without the need to over-control the development of every single project. The normative frame is not explored in this paper, although it is considered crucial to foster and manage city complexity (Moroni 2015). The spatial layout requires the ability (on the part of urban designers) to shape a basic infrastructural network (streets, transports, parking lots size and distribution, water and energy networks, parks, public buildings for activities such as schools, hospitals, etc.), whose stable structure provides the basis for balanced transformations without needing to pre-define them. Such designers must accept that their project represents the beginning of a transformation, rather than its final stage.
While a good structure (spatial and normative) minimizes the need to control every single intervention, the cross-scale effects of these incremental changes across the territory need to be constantly monitored, in order to prevent the self-destruction of diversity. The Atlas, in this respect, offers a tool to read these cross-scale effects and diversity cycles. This understanding is crucial to framing the scale of the problem, thus informing any revision of the network, which aims at inverting a negative cycle, through minimum interventions and with the minimum waste of economic resources. In Medellin the small-scale problems of specific slums were mainly dealt with by revising the wide network of connections and related public services. This strategy also made it possible to minimize economic costs also because the new metropolitan connections fostered the existing micro-economic network, thus increasing the ability of a place to regenerate through the individual initiative of its inhabitants. In the Oltrepò the Multi-Scale Atlas shows how a better integration between the slow and fast network of connections could re-orient emergent individual activities and businesses toward the existing urban fabric. This strategy would turn the spontaneous economic vitality spreading across the territory into a generator of diversity, thus inverting a negative cycle.

In this respect, the paper shows how a multi-scale understanding is necessary to shape every design and planning strategies which can channel (in a positive direction) the enormous quantity of individual (place-specific) energy that is concentrated in cities. In actual fact, the simple sum of individual performances, even if each of them is of a certain quality, does not necessarily guarantee the collective quality of the common good.

Finally, the present paper highlights the importance of understanding and detecting the multi-scale nature of emergent phenomena in space. It highlights how an understanding of the dynamic interaction between parts and whole, between micro and macro scale, between spatial details and larger phenomena, is crucial to informing every creative design strategy which is intended to cooperate with emergent place-specific transformations and foster their positive evolution. In this respect, a multi-scale approach is considered an essential part of any planning and design process which aims to channel the enormous quantity of individual initiatives that our cities contain toward the common good. This will hopefully increase the potential for city diversity across the heterogeneous landscapes of our everyday lives which are in a continuous state of becoming.

NOTES

1 These cycles describe how a social-ecological system organizes itself and how it responds to a changing world. During the Rapid Growth, the system’s components are weakly interconnected, and its internal state is weakly regulated. This is the time for innovation and growth. The transition to the conservation phase occurs because the system becomes more strongly interconnected and regulated: different ways of performing the same function (redundancy) are eliminated in favour of performing the function in just the most efficient way. The cost of efficiency is a loss in flexibility: such a system is increasingly stable, but over a decreasing range of conditions. In other words, its resilience declines. Under a small shock the system’s web breaks apart and suddenly comes undone. The release phase is brief and chaotic, but the destruction that ensues has a creative element: tightly bound capital is released, and all options are open. This phase leads quickly into a phase of reorganization and renewal. Novelty arises in the form of new inventions, creative ideas and people. The point in managing and adaptive cycles becomes then how to prevent a large collapse in the late conservation phase. The strategies elaborated by clever managers usually consist in undoing some of the constraints of the conservation phase, in order to navigate a graceful passage through the growth phase, without falling into a release phase (which is costly and unpleasant and involves the loss of capital) (Gunderson and Holling 2002; Walker and Salt 2006).

2 Managers who understand cross-scale effects often avoid a release phase at the scale of concern by generating release and reorganization phases at lower scales, thereby preventing the development of a late conservation phase at the scale of concern (Gunderson and Holling 2002; Walker and Salt 2006).

3 Jacobs (1961) describes this self-destruction cycle in six steps: (1) a diversified mixture of uses at some places becomes a popular and successful assemblage, (2) the success fosters an ardent competition for space and the locality develops, (3) few dominating uses emerge: the winners of the competition represent only a narrow segment of the many uses which together generated success, (4) visually and functionally, the place becomes monotonous and loses its appeal, (5) the local-
ity’s suitability, even for predominant use, declines, (6) the place becomes marginal.
4 Since the early studies of Jane Jacobs (1961) and Christopher Alexander (1965) it has been clear how complex urban orders emerge in the interaction between subjects and objects in space and time (Porqueddu 2015).
5 Jacobs (1961) already stressed the importance of the part-whole dynamic. She described the “ballet of good city sidewalk” as a complex order made up of movement and change, “an intricate ballet in which the individual dancers and ensembles all have distinct parts, which miraculously reinforce each other and compose an orderly whole” (50).
6 The Multi-Scale Atlas shows that density in terms of Floor Area Ratio does not necessarily foster city diversity. Across the case study, maps show how dense and compact settlements can be less diverse and vibrant than low-density settlements (Porqueddu 2015). On the other hand, it has been demonstrated that dense urban environment can be totally lacking in city diversity (Dovey and Symons 2014).
7 This approach is based on the idea that, in the contemporary metropolis, diversity emerges according to the distinct mixture of both local and wider flows intersecting at specific points in time (Massey 1994).
8 In this respect, the Atlas highlights how the problem is not to establish whether the increasing distribution of these metropolitan attractions across the territory is positive or negative. The existing local debate is often confused and polarized: on the one hand, the new hubs are considered responsible of the death for historical centres because they attract flows of people out of them. On the other hand, it is clear that these new activities witness a certain vitality and foster emergent flows across the network of small settlements. The multi-scale maps show that their emergent presence is positive, while their position and spatial layout can create problems. The question then becomes: how to turn these emergent forces into generators of diversity?
9 In this respect, the local debate is often confused and contradictory. The restrictive rules applied to historical centres are in contrast with the widespread intention to protect the beautiful local landscape and to reduce soil consumption. In actual fact, they indirectly encourage the diffusion of the new activities across the countryside outside existing settlements.
10 In Medellin the creative solution consisted of using an established technology—generally associated with ski and tourist sites—as a means of public transport.

REFERENCES

Connecting the Dots: Hayek, Darwin, and Ecology

GUS DIZEREGA

Email: gdizerega@gmail.com
Web: http://dizerega.com

Abstract: Social institutions viewed from a Hayekian perspective closely match evolutionary and ecological perspectives in biology. All rely on the same systemic relationships of variation, selection, and inheritance. What Hayek called spontaneous orders are variations of a larger range of related phenomena. Concepts developed in one such field can enrich our understanding of analogous phenomena in others. Among the most important concepts explored here are individuals, organisms, species, and ecosystems. This integration carries important implications for how human societies can exist sustainably on the earth.

Keywords: Aldo Leopold, biology, Darwin, ecology, ecosystem, emergent order, evolution, Geerat Vermeij, Hayek, individuality, invisible hand, meme, organism, organization, spontaneous order

Modern ecology is the creation of two groups... The one studies the human community almost as if it were a separate entity, and calls its findings sociology, economics, and history. The other studies the plant and animal communities...The inevitable fusion of these two lines of thought will, perhaps, constitute the outstanding advance of the present century.

Aldo Leopold, A Sand County Almanac

This essay integrates the study of social institutions from a broadly Hayekian perspective with evolutionary and ecological perspectives in biology, arguing all rely on the same systemic relationships of variation, selection, and inheritance, albeit in very different contexts. Hayek’s “spontaneous orders” are important examples of this much larger range of related phenomena. Consequently, we may learn more about any of them by exploring research of analogous processes in another.

Evolutionary theory rests on three insights that, in different words, are often applied to the social sciences of complex phenomena: variation, selection, and inheritance. When variations appear in organisms within a particular environment, some survive better than others, with the most successful leaving more descendants. As they do they pass their advantageous traits on through their offspring. This is the evolutionary process. Ecological theory utilizes these same insights from a different perspective, emphasizing relations between different species within a given network rather than the change from one species to another along a given trajectory. Although these basic principles have been greatly refined and developed since 1859, no subsequent scientific work has challenged Darwin’s core insights.

Earlier observers thought they saw similarities between the economy of nature and the economy of a household, and assumed a higher authority was responsible for both. Recognizing these same similarities, today many natural and some social scientists suggest they arise from social and natural invisible hand phenomena. Variation, selection, and inheritance are common to all, but the nature of the variation, the principle of selection, and how inheritance is passed on, are specific to the order we are examining. So how far can we go in exploring these systemic similarities?

Biologist Geerat Vermeij argues we can go quite a ways, writing “Perhaps evolutionary theory and the long record of life on earth might reveal insights that political scientists, psychologists, and historians . . . have missed” (Vermeij 2010, p. 60). I believe Vermeij is right, and this paper will seek to build on his argument from a social scientific direction.

DISCOVERY PROCESSES

Hayek described competition as a discovery process, and explored this way of looking at it from a market and a scientific perspective (1978). His insight applies to both biological and social invisible hand phenomena. Be they innovative...
products, scientific hunches, political proposals, or new species; variations on ways of existing within such systems are kinds of hypotheses that are then subjected to evaluation by the systemic contexts on which they depend. There has to be a successful fit between the “hypothesis” and the systemic constraints within which it is evaluated. Successful hypotheses acquire systemically defined resources of money, acceptance by scientists, votes, or offspring, and so flourish (diZerega 2010). Within any of these processes, at any point new hypotheses can arise to be tested in turn and possibly replace one that had long been dominant.

Successful hypotheses fall along a continuum between marginal changes and something truly new. In economic theory one end of such a continuum can be described as Kirznerian arbitrage entrepreneurship: alertness to opportunities for successful market transactions (Kirzner 1973). The other end is pure Schumpeterian pioneering entrepreneurship, the process of “creative destruction” (Schumpeter 1961). Most actual entrepreneurial projects fall somewhere along a continuum between these models, disrupting some plans while having an arbitrage dimension as well.

The same observation applies in science, where, Thomas Kuhn distinguished between “revolutionary” (Schumpeterian) and “normal” (Kirznerian) science (Kuhn 1996). Scientific knowledge usually increases at its edges as the implications of existing theories are applied to explore ever more questions. Sometimes, however, a major discovery transforms how scientists think about what had once seemed settled knowledge, as happened in physics with the discovery of relativity and quantum mechanics, and previously, with Darwin and evolution. Other discoveries can fall somewhere in the middle, transforming a particular discipline while not challenging the larger scientific consensus within which it exists. The theory of continental drift is an example. Again we are observing a continuum of more or less rather than two distinct categories of change. I have made the same distinction regarding policy innovation within democracies (diZerega 1988, pp. 454-7).

The same patterns appear in biology. Vermeij quotes Wolfgang Steerer arguing “An organism represents a hypothesis of its environment, continually tested by selection for its predictive value and modified by adaptation for a better fit” (Vermeij 2004, p. 26). Over time new developments arise based on those that have previously succeeded. Vermeij explains, “Adaptation in general is the formation, and continual testing, of hypotheses about the environment” (Vermeij 2004, p. 26).

As in social systems, successful adaptations are usually marginal, such as a species gradually developing greater speed or resistance to a pathogen. However, sometimes these changes will be deeply transformative, as when the evolution of grasses made the rise of tropical savannah ecosystems possible at the expense of forests (Stromberg 2011). Perhaps the most profound change is when some species develop new ways of organizing themselves, such as the social insects who individually are small and vulnerable and whose ancestors were largely solitary. Most individual social bees and all ants are less viable than their solitary ancestors, however as societies they are now among the most successful life forms to have ever evolved (Wilson 2012). We human beings are the most extreme earthly case of ecological transformation caused by this process.

NATURAL AND SOCIAL NETWORK EFFECTS

Because they are networks, when acting within ecologies Garret Hardin emphasized we cannot do only one thing (Hardin 1986). As with social invisible hand phenomena, everything is linked with everything else in networks of relationships. Some of these links are weak but others are strong, and we often have little idea what the stronger ones are until we have disrupted them. For example, no Japanese forester imagined cutting coastal forests would significantly harm ocean fisheries, or even harm them at all. When a marine chemist then discovered decomposing leaves from forests significantly fertilized plankton, trees were replanted in coastal places. Where this was done, fish and oyster harvests increased (Robbins 2002). These kinds of networked relationships go every which way. Scientists studying Pacific Northwest forests discovered returning salmon’s bodies were an important source of nitrogen, enhancing forest growth, which in turn enhanced the region’s suitability for sustaining salmon (Helfield 2001).

The same observation applies in society. When the city of Berkeley adopted rent control, neither critics nor advocates imagined one of its impacts would be to make Berkeley a center for “California cuisine.” The long-term influence of rent control was to reduce student populations, replacing them with young professionals with more money to spend and more interest in eating out. The result was a burgeoning of creative restaurants while the population rent control was intended to assist declined (diZerega 2000, pp. 311-14).

While many organisms’ presence or absence has little impact on the ecosystem they inhabit, if an organism’s influence is strong enough its presence can transform one
ecosystem into a different one, as happened with the evolution of grasses mentioned above. Their physical size is not important. Grasses are smaller than trees. But the networks they influence are very important. Being grazed often kills small trees, but because they grow from the bottom up, grasses are little damaged, and many even thrive on it. In ecological studies, species whose presence or absence changes the broader ecological patterns are called “keystone species.” For example, salmon are a keystone species in Pacific Northwest forests, even though they are rarely present (Grames 2012).

We see the same patterns of networks and stronger or weaker links in society. Cities are social ecologies constituted by intricate networks of relationships. Jane Jacobs famously argued the nature of city blocks had a large impact on a neighborhood’s safety and social and economic vitality. Small diverse blocks encourage varied pedestrian use because they are more pleasant places to walk. Pedestrians add to the attractiveness of a neighborhood, and provide constant observations of what happens on the streets, increasing public safety. They promote a diversity of social networks. Short blocks are more than simply an ‘inefficient’ street layout for car transportation. Increasing the speed of traffic flow by widening streets and creating longer blocks discourages these other positive urban features, degrading a city’s social and economic health (Jacobs 1961, pp. 112-40). Imposing these requirements from above is like converting a natural forest to a plantation, where a complex network is subjected to a hierarchy of priorities, in this case facilitating traffic flow above other values.

Vancouver, British Columbia, is usually described as one of the most livable cities in North America, if not the world. Vancouver’s urban policies generally reflect Jacobs’ and others’ insights about vital neighborhoods being more important for urban well-being than encouraging rapid transportation. They cultivate a vital urban ecology. The wisdom of these insights has been confirmed for many decades (diZerega and Hardwick 2011).

Short blocks are not the only factor encouraging urban vitality. Jacobs argued dense mixed use neighborhoods populated by people of varied incomes had a similar impact, again by increasing the variety of interactions. High-rise apartments surrounded by open space, no matter how park-like, inhibited the growth of social networks, reducing a region’s ability to support and nourish rich human relationships and increased crime. As in natural ecologies, complex network effects existed at every level (Scott 1998). However, due to the networks of relationships they encourage short diverse blocks could be described as a keystone feature of vital urban ecologies (Schiller 2016).

BLURRING BOUNDARIES

Most of the time we describe biological ecosystems as if they possessed clear boundaries, as I just did when describing the Pacific Coast forest ecosystem. But in fact, all ecosystems are integrated into the larger biosphere, the total ecology of life on earth. Boundaries between the ocean and land, forest and savannah, or a lake and its surroundings are porous, sometimes very much so, as those Japanese scientists discovered. But for most purposes the “ecology of life” is too complex to comprehend, and so scientists look at smaller sub-systems, such as coastal rain forest or arctic tundra, as reasonably unique in themselves and hope to discover patterns that might also exist elsewhere, or perhaps encounter something new to science.

Ecosystems’ boundaries are determined by the questions we ask. The Pacific salmon ecosystem overlaps with that of Pacific Northwest rain forests but raises different questions for research. When studying the Amazon rainforest, usually an ecologist does not need to pay attention to wind-blow dust from parts of Africa, but for other questions this more inclusive frame is important because the dust is a major source of phosphorus for earth’s largest tropical forest (Yu 2015, p. 42).

Again, the same pattern exists in the social world. The market order is the network of products and production that incorporates all that is bought and sold. Science grows in complexity as new discoveries open up additional fields for exploration. Neither the market nor science are coterminous with the social world. We can select out subsets of social ecosystems such as markets or science for study, but in reality, as with natural ecosystems, they interpenetrate. The boundaries we perceive arise from our questions. Markets, science, and other spontaneous orders in Hayek’s sense combine and interpenetrate in the larger social world we call civil society (diZerega 2014).

INDIVIDUALS IN BIOLOGY

The same kind of relativizing and blurring of what were once considered reasonably distinct ecosystem boundaries is also being discovered in the organisms within these systems. To focus initially on nature, over time the ecology of life tends to become increasingly complex (Vermeij 2004, pp. 252-4). Growing differentiation creates possibilities for
new kinds of biological integration. For example, two very different species of tree, paper birch and Douglas fir, live cooperatively, aided by yet another species, a mycorrhizal fungus, that links them at their roots. Each sends carbon to the other during stressful periods. During the summer, young firs suffer if they are in the shade, and receive aid from birches. In the winter the birch have no leaves to produce energy, and receive nutrients from firs (Frazer 2013). The result is what Nature called the “wood wide web” (Helgason 1998). There are many similar examples. We have even discovered varieties of bacteria within our guts can make good health possible, and influence our minds for the better (Schmidt 2015).

Sometimes a species will have existed so intimately with another for so long its members can no longer live independently, as with some fungi ants cultivate and the corn we cultivate. This dependency can even embrace both organisms, as with yucca moths and yuccas, each of whom needs the other to reproduce.

This blurring of once taken for granted boundaries goes still further. What we have long considered individual organisms are not individuals in the way we thought. In some cases, distinct organisms become so dependent on one another they become a separate individual in their own right, as with the partnership of algae and fungi we call a lichen. In recent years, these relationships have been discovered to be far more common than once imagined. Even the eukaryotic cells making up our bodies and the bodies of all other multicellular organisms are composed of what were once at least two distinct organisms (Margulis 1999).

Bacteria are a different kind of individual entirely. Individual ‘species’ of bacteria can trade DNA back and forth allowing resistance to antibiotics to rapidly spread through a population. The borders between individual bacteria are more open than they are in eukaryotic cells. Individuality in any strong sense does not exist.

In all these cases boundaries between organisms remain, but, depending on the issue at hand, they can be porous and sometimes only contextual. In terms of our own biology, increasingly scientists call us “ecosystems” (McFall-Ngai 2013) or “super organisms” (Milius 2014, p. 15). Our individuality is certainly real, but emerges from a node of intersecting relationships rather than existing as some kind of psychological or biological billiard ball separate from other similar balls.

This revision of biological individuality helps deepen our understanding of it in the social world.

**INDIVIDUALS IN SOCIETY**

Human beings are biological individuals and social individuals. In nature most biological individuals are not social in our sense of having a culture and division of labor and knowledge. They are two different kinds of individuality. I want to suggest in turn that not all social individuals are biological individuals. Each kind of ecosystem develops its own kinds of individuals and species, and individuals in the social world include more than what we think of as individual human beings. We are its only biological component, but not its only social component.

Species are collections of individuals with common traits that, on the whole, maintain their distinctiveness from other species. As we have seen, the boundaries between biological species are not rigid, but they are strong enough for the species distinction to be important (Stamos, 2013). Is there a social-ecosystemic equivalent to biological species?

Geerat Vermeij helps us here, arguing in economic systems an ‘occupation’ is like a ‘species.’

The key to the analogy between species and occupations lies in the roles that these entities play in the systems in which they live. The conditions of life create a regime of selection that yields adapted individuals whose phenotype [its observable features] reflects not just ancestry and the means by which some species can be told apart from one another, but also the ways of life the members are able to lead thanks to heritable adaptations. Similarly in human society, occupations have phenotypes that allow us to identify and classify individuals economically. Like species. (Vermeij, 2004, p. 45)

Building on Vermeij’s insight about occupations, organizations are like complex organisms where different species of professions symbiotically combine to accomplish what they could not achieve by themselves (diZerega, 2015). There are many biological equivalents to this kind of relationship, such as lichens, slime molds, and corals. The key to this insight is to focus on the relationships (occupations) rather than the biological individuals (human beings).

From a human point of view many organizations resemble slime molds. Most of the time slime molds exist as individual amoebae living separately on the floor of a forest. However, should conditions deteriorate, they combine together into a larger “slug” able to accomplish what they
cannot individually. In their new collective relationship individuals act differently than when they were separate from one another, including self-sacrificing behavior in service to the larger entity. Yet in many slime molds their individual cellular boundaries remain, and if conditions improve the larger ‘slug’ dissolves again into individual amoebae (Cellular Slime Molds). The relationships between slime mold amoebae shifts between one kind of individual to another. Is the amoeba an individual organism? Is the ‘slug’? Are they both?

Applying this perspective to society, occupations come together to form organizations to accomplish what they could not achieve singly. Individual human beings are expendable so long as the occupations they practice are performed by another.

Should the organization cease to exist, many occupations can still survive on their own. As with a slime mold, human beings will behave differently on their own than they otherwise might when part of a strong organization, a phenomenon behind some of our most heroic acts, and also many of our most despicable (diZerega, 2015). But what makes them a part of the organization is not their individuality but their performing essential occupational functions. Thus, the organization is an organism composed of occupational relationships shaped by the context within which they exist together.

Like occupations, organizations seek their goals in a larger selective environment that imposes constraints as well as offering opportunities to flourish. As with successful occupations, successful organizations adapt to handle their environmental constraints as well as exploiting their opportunities. Like an organism they can evolve and adapt. Organizations are complex non-biological organisms existing within the social ecology.

Civil society as ecology

The model I am developing sees civil society as an ecology. On the one hand, because we are biological beings, civil society is a part of the ecology of biological life. But it is also a network of voluntary cooperation, shaped by and in turn shaping human ends. Their strongest strands are the invisible hand phenomena of markets, science, democracy, and other spontaneous orders. Now we are seeing another dimension to this network: organizations and occupations comprise species that to some significant degree can be separated from their biological identity as human beings. Of course, they work through human beings, but they can in turn actively shape the human beings who comprise them.

Far from needing expression by biological organisms, increasingly many occupations are being performed by robots, and organizations increasingly prefer the latter.

Civil society contains many different kinds of organizations: sole proprietorships, partnerships, family owned concerns, cooperatives, research teams, churches, charities, sports teams, universities, political parties, and corporations. Each kind of organization brings different strengths and weaknesses to its task. Some are specialized to flourish in science, some in the market, some are profit oriented and some are philanthropic. Some members within an organization are tightly linked, as in a sports team, and some are loosely linked, as in a university. Because it best facilitates individuals coming together to cooperate for common ends, civil society is the richest and most diverse social ecosystem for organizational species.

If organizations in a social ecology are like organisms in a biological one, we will gain in understanding by viewing them as subject to similar evolutionary pressures leading to differentiation and adaptation. Organizations seek to succeed within an environment shared with other organizations, each shaping the environment to some degree and in turn shaped by it. There are producers who seek systemic resources and are then subject to predation and parasitism as well as scavengers who reconstitute organizational resources into more viable forms. There will also be keystone organizations as well as others that play a far less important role in shaping a social ecosystem. Like a life form, each can be understood teleologically.

Distinguishing occupations and organizations from the human beings who embody them sets the stage for a final step in grasping how civil society can be understood as the most complex ecosystem ever to have arisen on this planet. Institutions, be they organizations, occupations, or other patterns of relationship we engage in are cultural, not biological, phenomena. And cultural phenomena are maintained or changed by how we think about them.

The extended nature of the human mind

We evolved in a chain of descent probably going back at least 3.8 billion years (Rosen, 2015). We express this heritage in at least our metabolism, our senses, and our physical needs. Today scientists are learning this deep earthly rootedness even shapes our psychological well-being and the conditions for maintaining physical health (Maller, 2006, p. 45-54; Williams, 2017). We are more intimately connected to this planet than the abstractions of ‘economic man’ and ‘rational action.’ can ever grasp.
Yet all of us are aware we also differ from the rest of nature. There is no region of the earth where life has not been significantly affected by human action other than perhaps bacteria or archaea living deep underground, in undersea vents, or deep in thermal pools. Our era is increasingly called the anthropocene because in important ways we are increasingly influencing global ecosystems and even the world’s geology. Even more distinctly, we know we are doing so.

Rather smugly, we attribute our disproportionate influence over other life forms to our intelligence. But as we usually think about our intelligence, we are wrong. We are cleverer, but not in the way most of us like to imagine.

Human intelligence comes with few instincts able to orient us in the world. If necessary, a newborn fawn can soon run. Absent that necessity it knows to lie quietly to avoid predators. A human baby will normally take a year to walk, two to run and many more to run fast. If infants needed to be quiet to survive, we would be extinct. Individual intelligence is real enough, but is primarily useful on the margins, benefiting from and occasionally modifying achievements of a greater intelligence located in relationships outside our physical bodies rather than inside our heads.

Our minds consist of much more than what our individual intelligence can discover or create. They also include our cultural inheritance. As Hayek put it, “Mind is as much a product of the social environment in which it has grown up and which it has not made as something that has in turn acted upon and altered these institutions.” (Hayek, 1973, p. 17). Hayek elaborated “[M]ind can exist only as part of another independently existing distinct structure or order, although that order persists and can develop only because millions of minds constantly absorb and modify parts of it (Hayek, 1979, p. 157).

Based on work with an unusually isolated Amazonian tribe, linguist Daniel Everett’s newest work demonstrates how, rather than having an innate instinct for language, as has long been thought, language evolved along with culture. It is not an artefact of biological evolution, but rather of cultural evolution, promoted by the need to communicate effectively and quickly. Like Hayek, Everett argues culture and language are the foundations for human reason and understanding (Everett, 2017; 2012).

Tom Wolfe writes of Everett’s conclusions “language had not evolved from . . . anything. It was an artifact. Just as man had taken natural materials, namely, wood and metal, and combined them to create the axe, he had taken natural sounds and put them together in the form of codes representing objects, actions, and, ultimately, thoughts and calculations and called the codes words” (Wolfe, 2016).

Mind in this more inclusive cultural sense constitutes an emergent order of its own with its own variation, selection, and inheritance, one in important respects distinct from the orders in the biological world. For example, language enables the rapid inheritance of acquired characteristics, and that inheritance can come from any person who is able to pass on or be imitated by another, regardless of biological connections. If sex speeded up evolution and adaptability by mixing two genomes, language does so by mixing the insights of countless individual minds. The world of ideas operates in important respects in different ways than other ecological processes, but with the same abstract principles underlying them.

**Memes and institutions**

Richard Dawkins’ term “meme” describes an idea that enters into culture and can then shape as well as be shaped by it. As such memes have a kind of ‘objective’ existence in the sense Peter Berger and Thomas Luckmann used the term for culture as both a human creation and an objective reality (Berger, 1967 p. 61). They exist independently of any particular subjective mind, and need those minds to replicate. Successful memes influence behavior, sometimes powerfully.

Dawkins famously compared memes to genes, arguing like genes they adapt, flourish, and die, in their case through our success or failure in incorporating them into our lives (Dawkins, 1989). I think equating memes with genes is a most useful heuristic, and perhaps considerably more than that. As memes, ideas in the broad sense adapt, spread, die or mutate through their interaction in a social ecosystem.

From this perspective ideas are like organisms requiring mental rather than physical energy from people to flourish. Ideas compete for this mental support because we have limited attention to give them, and the most successful often form symbiotic relations with others. Without support, an idea “dies,” or perhaps goes dormant awaiting a more supportive environment.

Ideas manifest in the material world by influencing our behavior. As elements within a culture, ideas influence the world through the medium of the people guided by them, shaped by the institutional framework within which they act.

Institutions are not quite like ideas. They are patterned ways of doing things that also shape our perception of what
to do and how to do them. Institutions ‘tell’ us how to do things but not what things to do, even if their own values influence our actions somewhat independently from our intentions. They shape and coordinate our dealings with others, and so usually magnify our power to act in the world while also channeling and shaping that power. They can be supported by memes, but are not themselves ideas.

In civil society, language, science, markets, democracy, a variety of organizations and many other elements of our lives such as what constitute a family and what it means to make a contract are institutions. Some arise in an evolutionary way, the product of our actions but not our design, others are deliberately created and found useful enough to persist until they take on a reality independent of their creators. These institutions shape and coordinate our dealings with others, to some degree independently of our intentions, and as such magnify our power to act in the world while also channeling and shaping that power.

In Berger and Luckmann’s terms, like institutions, memes take on an “objective” independent existence. (Berger, 61) Initially we encounter them as a part of our taken for granted world, where what it is to be a parent is as uncritically accepted as true in the context of an institution as what it is to be a rock in the context of its hardness. Some institutions are spontaneous orders and some are the organizations acting within those orders (diZerega, 2015).

An institution is not necessarily a meme. If an institution is questioned this questioning becomes a part of the ideational ecology but need not impact how the institution operates. The market, for example, existed long before economic theory and those who support it through their actions need not have any conception of it at all. Alternatively, the influence of the market as a meme need have little impact on how it functions, which is independently of individual intent or even awareness. Ideas exist separately from institutions, but each shapes the other (diZerega, 2015).

In its immediate impact on us the ideational meme-network dwarfs biological ecologies and evolution and is at least as powerful as the institutional ecologies of language, markets, science, democracies, and civil society which shape how it manifests. If institutions shape how we go about acting and cooperating, this ideational ecology provides the environment wherein we decide what it is we will do. Institutions are not creative.

Two examples help clarify my meaning. As an institution, what we often call the “scientific method” can be applied in similar ways over long periods but what counts as scientific knowledge verified by that institution changes all the time. The rules of the market and of market oriented organizations can apply to either relatively uncreative activities, such as supplying a market niche with a product that does not change, or to innovative and inventive companies. For example, markets can coordinate building buggies, cars, and rocket ships but the ideas for what to build must come from the outside.

I distinguish between institutional and ideational ecosystems because how ideas manifest is powerfully shaped by the institutions in which they act, just as these institutions can be shaped by them. While ideas and institutions mutually influence one another, for some purposes we can treat them as different. The human relation with the rest of nature is one of them.

We can learn from many people, reconfigure what we learned, and adapt at a speed making it possible within a single generation’s lifetime to change from a society where horses were still common and cars a rarity, to one where we watched a live broadcast of the first man to walk on the moon. But when that walk occurred no one imagined the interconnected world of today with its smart phones, internet and computers at everyone’sbeck and call. The institutions shaping all this changed, but not nearly as much as the ideas manifesting within them.

The big disconnect: institutions and ideas as cultural accelerators

In the biological world one close analogy to developing social institutions would be the evolution of eusocial collective societies by organisms that were once solitary, such as ants or bees. While rooted in biological rather than social evolution, their evolved way of life completely changed the roles these organisms play in the world. Many who study them argue an ant colony in particular should now be considered a super organism, an organism distinct from the individual ants who make it up, or perhaps a physical extension of the queen, as if our fingers could take leave from our hands but still be subject to our intentions (Wilson, 2012). Eusocial insects can be said to have genetically determined institutions but so far as we know they have no memes, which is perhaps why they are relatively unchanging. We are obviously far more able to learn and change our institutions than are ants, but in the modern world an individual could be said with little exaggeration to know as much about the collective mind of its society as an ant does about its.

Human beings have developed two conceptually distinct and, in practice, deeply interrelated dimensions of non-biological invisible hand phenomena we combine together into
what we broadly term “culture.” Both institutions and the ecology of ideas are unique to human society and so differentiate it from the biological world that ultimately supports it. They are different dimensions of the larger “mind” which makes possible and empowers what we consider our individual minds. Because they are not biological, there is a basic disconnect between the cultural world in which human beings most immediately live and the biological world that ultimately supports them.

Biologically we inherit genes and to some degree the impact of previous experiences from our parents and even great grandparents (Interlandi, 2013). Institutionally we inherit systems of cooperation and coordination, systems rewarding some kinds of cooperation and penalizing those that do not fit their values. We take most of this institutional framework for granted, sometimes seeking to change some from within the perspective of the others. But for the most part institutions and memes exist independently of individual human beings.

In terms of the language I am developing here, social institutions are subject to processes of change and adaptation, but much more quickly than most such processes in biological systems. The other dimension of mind is the realm of ideas, and here the problem is even more severe. Culturally we have developed and maintain an ideational ecosystem of beliefs from what “everybody knows,” shading to more specialized beliefs. All of these meanings are rooted in their relationships with other such concepts. We have created an ecological field of meaning within which we orient ourselves. This system adapts even faster than does the institutional ecosystem. Biological evolution proceeds at the speed of successful reproduction. With some bacterial exceptions, institutional evolution is considerably faster. Ideational evolution proceeds at the speed of thought.

In both its institutional and ideational dimensions, culture creates a deep disconnect between human life and the biological world that ultimately supports it. Rapidly reproducing life forms able to exist as generalists can successfully adapt to us, but more specialized or more slowly reproducing ones are at a disadvantage. Rats and mice, flies, bacteria, and dandelions do pretty well in the modern world. Elephants, lions, whooping cranes, whales, redwoods, and increasingly, song birds, not so well. These beings need special protection or consideration if we are to continue benefiting from their presence.

The natural ecosystems that support us are more like redwoods and elephants than dandelions and mice. The long-term viability of ecosystems around the world on which we depend is being threatened by anthropogenic global warming. But global warming is only the most extreme instance of culturally based power disconnected from natural processes. Topsoil normally takes many years to form and can be wiped out in days. Coral reefs and kelp forests support important nurseries and shelter for the fish on which so many of us depend. They can be easily destroyed and recover far more slowly. Given their power over natural processes in the short run and dependence on their health in the long, cultures run the risk of destroying the natural processes that support them.

Some societies succeeded in living sustainably and others failed, but all societies face this challenge (Diamond, 2005). Viking settlements died out in Greenland whereas the Inuit succeeded. The fertile crescent is one of agriculture’s points of origin, but today land that long sustained earlier civilizations has little agricultural value. North Africa became Rome’s breadbasket after the Romans impoverished their own soil, but today North African soil is also much degraded (Montgomery, 2007, pp. 64-7). Egypt was saved from this problem because annual floods by the Nile enriched the soil, a matter more of luck than wisdom. Today due to the Aswan Dam the Nile no longer renews the soil and Egyptian farmers face the same problems less fortunately located farming civilizations had to face. Over the past 40 years the world as a whole has lost perhaps one third of its arable land, and we hope to be around for thousands of years to come (Milman, 2015).

Human culture’s disconnection from nature is amplified by our unequalled capacity to shift from one resource to another, eliminating any short-term reason for conserving the first, while avoiding the natural checks that keep other species from completely consuming their resource base. The earth’s life forms, the ecosystems that sustain them, and its minerals all constitute our resource base. When one resource is exhausted or made extinct, we find substitutes. In doing so we impact natural relationships so deeply as to transform one ecosystem into another, usually a far more impoverished one. Alternatively, we generate short term abundance at the risk of long term failure, as with the depletion of underground aquifers. Like a factory that never reinvests any of its profits in upkeep, natural operating ‘capital,’ such as soil and water, is exhausted.

Our capacity to substitute one plant or animal for another weakens any purely instrumental incentive to conserve them. But as we substitute one organism for another we usually do so based only on their utility as resources in the narrowly human world and ignore any role they play
in their ecosystems. *We act linearly in a nonlinear environment.*

**The problem of ignorance**

The more powerful a culture’s impact on its environment the more foolish this approach becomes. Its implications duplicate arguments against market interventions such as price controls: each intervention generates unexpected problems inviting another external intervention until its ability to coordinate plans becomes seriously crippled.

The Scottish Enlightenment’s thinkers criticized people thinking they could rationally reconstruct society. Within the same tradition, Hayek criticized imposing a hierarchy of goals on the market’s invisible hand phenomena, arguing efforts to override the price system’s coordinating abilities with direct interventions such as price controls over some or all market phenomena would backfire. Maintaining a market order required *not* making short run interventions to achieve particular goals if doing so interfered with the market’s feedback processes. The short-term advantages would be clear and the future damages would *always* be un-clear and so in terms of costs and benefits, *always* be under estimated. Were the results known, the interventions would usually have never have been made.

Wes Jackson made the same point with respect to ecologies: “Given that we have far more ignorance than knowledge about the workings within, the ecosystem becomes the best conceptual tool to help us understand how to get along in this world” (Jackson, 2008, p. 27). Neither Jackson nor Hayek believed nothing could be done to serve human goals better if market or ecological processes interfered with our well-being. However, they emphasized whatever we do should harmonize with the underlying dynamics of a complex social or biological system rather than imposing an external goal on it. To use an agricultural metaphor, in both societies and nature, the technique for creating positive changes should be to encourage and cultivate rather than construct and command.

**Ignorance in action**

Many economists have tried to integrate environmental theory into economics. If my argument is sound, they are getting it backwards. By ignoring the problem of ignorance and applying linear reasoning rooted in market based rationality to nonlinear issues, they end up with flawed analyses and mistaken perspectives. Consider two examples from *Free Market Environmentalism*, by Terry Anderson and Donald Leal. One deals with forests, the other with oceanic fisheries.

Michigan’s Kingston Plains once contained a large old growth white pine forest. After the forest was clear-cut and the land burned over, the forest did not regrow. Between the kind of logging performed and subsequent fires the soil is now too poor for the trees to regenerate. Nor is the land useful today for agriculture or recreation. To most people the way the Kingston Plains was logged would appear to be a bad use of the land.

Anderson and Leal argue otherwise:

> When the trees were cut, good timber stands in the Great Lakes area were selling for around $20 per acre. In order to determine whether it would have made more sense to invest in trees by forgoing the harvest, we must consider the return on other investments. Had the income from selling these trees been invested in bonds or some other form of savings at the time, it would now be worth approximately $110,000 per acre, or $2.8 billion for the forty square miles. . . . Because the land in this area is not worth anything close to this, we must infer that harvesting the trees was the correct choice. (Anderson and Leal, 2001, p. 43)

Years before Anderson and Leal wrote their book, Garret Hardin observed “*At high rates of interest the present value of the distant future effectively vanishes*” (Hardin, 1986, p. 74). The old growth redwoods we treasure today are not 500 years old, the minimum standard for old growth trees of that species. They are at least 1500 years old. As Hardin observed, “He who finds ecstasy in the wonder of today’s mature redwood forest benefits from a preservation a pre-Christian economist could not have justified” (Hardin, 1986, p. 75).

The forest that once stood on the Kingston Plains were considerably younger, but according to economic reasoning still far too old to justify investing in its maintenance. Better to cut it all down, even if it never re-grows.
Anderson and Leal describe the core market logic for using resources: “the long-run prevailing interest rate serves as a guide for determining the rate at which... resources should be harvested.” (41). While their book depicts some innovative and successful ways, particularly in which local fishing communities have managed to develop sustainable practices, they are able to do so because either the relevant ecosystem is able to be encompassed by traditional property rights owned by individuals and communities rather than corporations, or they exist within a larger network of regulations.

A traditional economist might argue that I am describing only an aesthetic preference for forests over fields. While I think this argument is incorrect, addressing it would take us far afield. Happily, (or from another perspective unhappily), a concrete example goes well beyond aesthetics...

In Northwestern forests Anderson and Leal’s market logic was applied to the Pacific Yew. Because Pacific yews were not valuable lumber trees, most that grew in logged areas were burned in slash-disposal fires (Egan, 1992; Leary 1992). By the time scientists discovered it produced taxol, the yew had become much less abundant. However, taxol is effective in treating ovarian, breast, cervical, pancreatic, and some kinds of lung cancer, as well as Kaposi’s sarcoma, and is now on the World Health organization’s list of essential medicines. Powerful struggles emerged over whether to harvest it for its medical value, even if unsustainably from seriously depleted populations, or allow less of the medicine to be available in order to preserve the tree. At the time there was no other source of taxol (Kolata, 1991). Fortunately for people and yew alike, synthetic varieties of taxol have now been developed, but one can legitimately wonder whether the drug would have ever been discovered, if market logic alone had governed forestry in the Pacific Northwest.

The problem of a disconnection between power and importance applies to more than forests. In many of the world’s oceans overall catches have begun to decline, in some cases catastrophically so. This decline happened in a matter of decades, as world fish stocks were healthy after WWII.

To justify their expense, modern capital intensive fishing trawlers are virtually forced to take every fish they can locate. Often they do not even use many of the fish they catch, because they are the wrong size or species. In 1994 a record 751 million pounds of dying edible fish were dumped back into Alaska waters alone, up from 740 million pounds in 1993 and 500 million pounds in 1992. David Germain reports 17 million pounds of halibut, 4 million pounds of herring, around 200,000 salmon, 360,000 king crabs, and 15 million tanner crabs were discarded, the equivalent of 50 million meals. Because many of these fish and crustaceans are caught deep in the sea, where they were adapted to the high pressures of the depths, raising them to the surface normally kills them (Germain, 1995, p. A-16). This practice is economically “efficient” in exactly the way Anderson and Leal described with respect to cutting and abandoning smaller trees to get at the more valuable larger ones (Anderson and Leal, 2001, pp. 39-40).

The results, if the logic is allowed to work itself out unchecked, is catastrophic. In 2008 the United Nations Environment Program reported as much as 80 percent of the world’s main fish catch species have now been “exploited beyond or close to their harvest capacity”. If fishing rates continue unabated by 2048 all of the species currently fished for food will have disappeared (Oliver, 2008).

I have explained how all societies are challenged by the deep disconnect between biological and cultural ecosystems, with the latter being dominant in the short run but the former dominant in the long run. More than one society has failed to bridge this disconnect well enough to survive. The enormous increase in market power is making a challenge once met or failed by particular cultures in particular places a world-wide problem. Market logic simply cannot think long term unless enormous profits are promised for doing so, as Anderson and Leal emphasize.

This disconnection is the fundamental tension between economic and ecological processes. It is why, if our world is to last as a prosperous and good place for us to live, economics must be subordinated to natural processes.
REFERENCES

Adelstein, Richard 1996. Language Orders. 7 Constitutional Political Economy, Fall.

NOTES

1. I would like to thank Chirag Kasbekar for alerting me to Geerat Vermeij’s fascinating work, and to a referee for helping me clarify what is a fairly complex argument.

THE IMPORTANCE OF ETHICS

When the social sciences are divorced from the biological ones, and the commonality and interweaving of different emergent systems becomes confused or invisible, the results will be misunderstanding of vital issues and destructive policy conclusions.

Given that people often urge interventions as soon as the advantages they seek appear to outweigh immediate disadvantages while the future is ‘purely theoretical,’ there is an inherent tendency to make long term mistakes. With respect to the market, Hayek considered ethical principles the only reliable check to off-set the apparent advantages of short term interventions into complex market processes at the cost of long-term sustainability (Hayek, 1973, pp. 55-71). On this important issue Hayek was one with the ecologist Aldo Leopold who also emphasized ethics as a restraint on power (Leopold, 1970, pp. 237-64). In the absence of ethical relationships the short-term advantages of intervening into complex systems will always outweigh the somewhat hypothetical long term disadvantages. As such, an appropriate ethics is a essential tool for dealing with ignorance when our power is greater than our knowledge. Perhaps the subject of environmental ethics is as important for maintaining a sustainable society as is the ethics men like Hayek long argued are foundational for maintaining a free society (diZerega, 1996).

That, however, is beyond the scope of this paper.1


The Role of Spontaneous Order in Video Games: A Case Study of Destiny

WILLIAM GORDON MILLER

Manuel H. Johnson Center for Political Economy
Sorrell College of Business
Troy University
Troy, AL 36082

Email: wmiller65275@troy.edu
Website: https://www.linkedin.com/in/gmiller1991/

Abstract: Complex and uniquely human spontaneous orders are central to the Austrian understanding of the market process and price system. Yet, these spontaneous orders are often difficult to study with traditional empirical methods. Since virtual worlds serve as extensions of our own world, video games offer alternative methods to empirically study spontaneous order. To that end, this paper presents a case study analyzing the role of spontaneous (or emergent) order in facilitating enhanced value for players. Through exploring the emergence of social institutions in Destiny, I provide evidence of the benefits of decentralized decision-making within virtual worlds. Given the theoretically limitless number of game elements from which to draw, this analysis utilizes the most basic and highly representative examples from the Destiny universe to showcase the phenomenon of spontaneous order within this context.

Keywords: Spontaneous order, emergent order, Destiny, video games, Austrian economics

JEL Codes: B53, P49, Z10

"The curious task of economics is to demonstrate to men how little they really know about what they imagine they can design."—F. A. Hayek (1988, p. 76)

I. INTRODUCTION

While paramount to the Austrian understanding of the market process and the price system, complex and uniquely human spontaneous orders are difficult to study in the experimental laboratory or with computational economics. While these methods are useful in advancing our understanding of spontaneous order, as are historical case studies and comparative institutional analysis, the rise in popularity of video game platforms provides a unique avenue for studying the formulation and effect of spontaneous order in complex environments inhabited by diverse agents with varied, and often conflicting, intentions.

Having risen immensely in popularity over the last three decades, the interactive worlds of online video games have developed into richly detailed universes that display many of the same features inherent in our own. For example, many of the triple-A games—typically those which are heavily promoted and possess the highest budgets—released today showcase lush environments, a large variety of character interactions, and complex economies, thereby making some of their worlds hard to distinguish from reality (beyond their inherently fantastical nature). Furthermore, with the advent of online multiplayer gaming, these universes have become extensions of our own in which social networks, money, and learning fluidly transfer back and forth between the virtual and real world.

Video games provide an appropriate research methodology to evaluate economic theory and the accuracy of economic observation for at least three primary reasons. First, while informative, the choices made by players in laboratory experiments, and the resulting consequences, often have
little to no direct effect on the world in which we live. Video games provide a more consequential, but admittedly less controllable, laboratory. Second, while experimental laboratories often pull from a sample of highly-educated college students, video game players, at least those in America, often come from a wide variety of diverse backgrounds (Duggan 2015). 

Due to the mass appeal of this new form of media, large sample sizes are naturally available which serve to make video games a sort of natural experiment. Finally, the social interactions that have developed in tandem with these games allow social scientists to directly observe, at a more informal level, real human interactions. In doing so, theories concerning human action become more well-informed, thereby contributing to the robustness of overall knowledge available for analysis. In comparison, computational economics lacks the uniquely social aspects of human behavior that are important for understanding spontaneous order. 

Although there is a growing body of literature focusing on applying economics to video games and using video games to advance our understanding of economics, there are few studies focusing on the economic concepts emphasized by the Austrian school of economics. This paper contributes to that literature by explicitly focusing on the Austrian concept of spontaneous order through video games. Specifically, I look to the formation of emergent order in the game world of Destiny. I hold that this order might be observed primarily through the ways in which players interact to overcome production problems, especially those related to the lack of predesigned social institutions within the game. Understanding how this order emerges should help us to understand the importance of such social institutions more broadly, particularly those related to online mediums. 

The rest of this paper proceeds as follows. Section II reviews the relevant literature on spontaneous orders, issues concerning experimental vs. field research, and the similarities to agent-based computational modeling inherent in this line of research. Section III covers the theoretical structure of the paper and connects it to the larger body of work related to the theory of the firm. Section IV provides my case study of the two levels of emergent order evident within the world of Destiny. Section V concludes.

II. LITERATURE REVIEW

Compared to mainstream economists, Austrian economists have placed far more emphasis on understanding the role of spontaneous order (Boettke 1990; D’Amico 2015). This is due to both theoretical emphasis and methodology. Austrian economists seek to understand how self-interested individuals create complex social arrangements and the division of labor by placing the science of exchange at the core of their emphasis (Boettke, Fink, and Smith 2012). They thus tend to utilize empirical methodologies such as comparative institutional analysis (Boettke, Coyne, Leeson, and Sautet 2005), analytical narratives (Boettke 2000), laboratory experiments (V.L. Smith 1994), and agent-based computational modeling (Nell 2009; Seagren 2011; Wallick 2012). In comparison, mainstream economists tend to focus on institutionally sterile environments of choice (Boettke, Fink, and Smith 2012). They thus tend to focus on developing mathematical theory, in terms of agent optimization subject to constraints, and econometric empirical methods using conventional data sources. 

The open models of Austrian economics enable the examination and understanding of the role of spontaneous order. While this limits the development of policies for engineering the economy (Wagner 2011), a downside of the Austrian approach to many mainstream economists, it does advance Hayek’s (1988, p. 76) definition of what the primary task of economists is, “…to demonstrate to men how little they really know about what they imagine they can design.” Interfering with complex social arrangements, according to Austrians, should not be taken lightly since individuals naturally develop cooperative institutions in order to solve various problems within society (Boettke and Candela 2015).

Modern research in Austrian economics has extended our understanding of the role of spontaneous order to pirate organizations (Leeson 2007), prison organizations (Skarbek 2012), and commercial law (Benson 1989). These studies and others provide evidence of the role of spontaneous order within society. The purpose of this paper is to extend such an analysis of spontaneous order to the world of video games. 

As previously suggested, video games represent a fairly realistic laboratory for analyzing human behavior. While not controllable to the extent most laboratory experiments often are, they do tend to incorporate a higher degree of consequential realism. As Falk and Heckman (2009) argue,
the laboratory methodology is widely criticized by social scientists due to its lack of “realism” and “generalizability.” However, they contend, that despite its limitations, it provides a rich narrative upon which to base an argument. Laboratory experiments and field research are comple-
ments, not substitutes, and both should be valued for the potential they have to enrich our knowledge of the social sciences.

Notably, Normann, Requate, and Waichman (2014) find that short-term lab experiments actually predict the long-
term behavior of field research pretty well under certain circumstances. Furthermore, V.L. Smith (1994) argues that experimental economics ultimately contributes to the robustness of our understanding of the social sciences by satisfying one or more of at least seven purposes. Given the aforementioned nature of video games, they serve as a natural compromise between the two methodologies of lab experiments and field research.

Interestingly, video games might also be viewed as a class-
ification of agent-based models (ABMs). Niazi and Hussain (2011, 2) provide one definition of an agent within these models as the following: “a representation of an interacting social component of a large system used to explore emergent global behavior in a simulation.” The importance of such an application to Austrian work can be found in Seagren (2011). Specifically, he argues that agent-based modeling can serve as the Austrian answer to the mainstream economists’ emphasis on mathematical models. Furthermore, he proposes that this type of modeling could serve as a more explanatory alternative to traditional neoclassical analysis, which unfortunately leads to the exclusion of so many relevant aspects of human activity.

Wolfram (2002), working from the perspective of physics and computer science, lays the foundation for such models in his exploration of how a surprising amount of complexity in computation tends to be generated even by what he refers to as relatively simple computational systems. In doing so, he provides several examples of this phenomenon related to a wide variety of systems including cellular automata, mobile automata, Turing machines, etc. To the extent this idea holds, he argues that a driving principle in science ought to consider experimentally investigating this type of complexity as a complement to traditional models of mathematical exploration and engineering principles. This idea might best be summarized in his concept of computational irreducible in which it simply remains impossible to describe this behavior in some concrete way. Rather, empirical and experimental approaches best help to inform our understanding.

Interestingly, his work also suggests that these types of simple computational programs represent a minimalistic variety of emergence in that they provide for interesting behavior even in an environment of less explicit direction/randomness. In other words, this behavior in programmatic language occurs beyond just the confines of its basic design. Additionally, making the underlying framework more complex results in a relatively insignificant amount of extra behavioral complexity. As such, much of the interesting behavior should be observable within the simpler model.

Essentially, the current paper extends this line of thought by looking at video games as relatively simplistic programs (at least along certain margins) that lead to undesigned complex behavior. Though a “designed game” exists on some level, the focus in this paper lies primarily on the more complex social institutions that emerge within this framework. By serving as a type of hybrid methodology of all the above approaches (i.e. experimental, empirical, and computational), video games provide us a window into the social aspects of interest, and as such, the succeeding case study provides a more thoroughly realistic presentation of human behavior, specifically as it applies to spontaneous order within a programmed environment.

III. THEORETICAL FOUNDATION

This section of the paper functions to provide a theoretical foundation for the succeeding case study. The first subsection relates work on the theory of the firm to the question of emergent social institutions in video games. Following this, I provide an overview of the case study methodology as well as a general framework for Section IV.

3.1 Theory of the Firm and Social Emergence

As previously hinted, the social institutions in the world of Destiny remain distinctly organic, and therefore, they represent the most relevant factor in considering the phenomenon of emergence within this realm. Why do such institutions emerge though, and what are the advantages of this emergence as opposed to the architectural design of such institutions? The answer to the first question might be gleaned through the theoretical work on the firm while the answer to the second is addressed more in the work from Hayek (1945) and A. Smith ([1776] 2003).

Essentially, the theory of the firm tries to explain both the reason why firms emerge as well as the disinguishing
characteristics between the form and functions of various firms. Coase (1937) provides one of the earliest neoclassical models for the firm in which transaction costs represent the decisive factor in whether or not individuals will contract on the market or organize more formally. If the costs of using the price mechanism are sufficiently high, then a more optimal solution lies in creating some type of firm, the size of which depends on the marginal benefits and costs of each unit produced in the firm (i.e. via the entrepreneur-coordinator) relative to simple market exchange (i.e. via the price mechanism). Other work builds on this by looking at things such as team production (Alchian and Demsetz 1972), ownership structure (Jensen and Meckling 1976), asset specificity (Williamson 1983), and agency problems in organizational competition (Fama and Jensen 1983).

Granted, the social activity seen in *Destiny*, as well as video games more generally, differs significantly from a traditional firm, but it does occur for similar reasons. For example, the transaction costs of simply exchanging goods produced in game, assuming they can somehow be priced, tend to be sufficiently high to warrant coordinated production. Such costs include the fixed costs of setting up a marketplace and the variable costs of maintenance.

In addition, for the purpose of game balance, many games, including *Destiny*, actively prohibit the peer-to-peer exchange of loot (weapons, armor, items, etc.). Though this constitutes an artificial barrier, it helps to counteract in-game equity problems that ultimately lead to efficiency problems (i.e. it helps keep the game from becoming boring due to certain players being too rapidly “overpowered”). Notably, market exchange might still occur through the buying and selling of whole accounts, but this remains largely infeasible as goods would likely have to be bundled in relatively unmarketable ways. Ultimately, this barrier does drastically end up increasing transaction costs by making it nearly impossible to facilitate any sort of workable market transaction, at least in the traditional sense.

Another way in which the theory of the firm relates to the emergence of social institutions in gaming centers around the idea of increased output from team production (Alchian and Demsetz 1972). While this does rely on appropriate mechanisms for problems of metering and moral hazard, games tend to have such mechanisms readily built into them. In regards to the former, rewards generally come from a non-rivalrous source since such gains are typically digital in nature (i.e. they are infinitely reproducible, at least in any practical sense). Hence, the marginal cost of creating the reward is effectively zero and the only real cost comes in the time it takes to acquire them. Unfortunately, this means that in regards to the latter, a player may simply shirk his responsibility in terms of allowing others to generate his rewards, and as such, an incentive exists to underproduce the desired good. However, in practice, as we shall see, this tends to not be that much of a problem due to strong social mechanisms of exclusion and reputation.

Considering all of this, it becomes apparent that low transaction costs and increased potential gains incentivize the emergence of social institutions in gaming. The question now becomes, does this lead to better outcomes compared to the ex ante creation of such insitutions? Well, the answer may not be straight-forward. Certainly, players are prone to mistakes and will make them frequently. So, for a particular case under strict assumptions, perhaps it is possible for a better outcome to result from a centrally made institution (in this case, a social platform made by the game developer). However, such decisions tend to be very case-specific, and situations can and often do arise that rarely conform to that particular case. Again, the time and place knowledge that arises in a given instance is crucial for effectively making any such decision (Hayek 1945). The question then becomes: Can institutions be architecturally designed to facilitate spontaneous order in gameplay (Beaulier, Smith, and Sutter 2012)?

The value of spontaneous order, as opposed to the constructed order, is that it allows for outcomes through the “invisible hand” that remain hard to properly design through central direction. Adam Smith ([1776]2003, 572) encapsulates the basic idea of spontaneous order in the following:

As every individual, therefore, endeavours as much as he can both to employ his capital in the support of domestic industry, and so to direct that industry that its produce may be of the greatest value; every individual necessarily labours to render the annual revenue of the society as great as he can. He generally, indeed, neither intends to promote the public interest, nor knows how much he is promoting it. By preferring the support of domestic to that of foreign industry, he intends only his own security; and by directing that industry in such a manner as its produce may be of the greatest value, he intends only his own gain, and he is in this, as in many other cases, led by an invisible hand to promote an end which was no part of his intention. Nor is it always the worse for the society that it was no part of it. By pursuing his own inter-
Nevertheless, players still end up collaborating and coordinating, as analysis is outside the scope of this paper. However, they do preclude negative outcomes from such institutions. Such an emergent order within the overall scope of the game.

As this quote relates to video games such as *Destiny*, improving societal gains comes through increasing the value of the game for everyone through individual self-interested actions. Each individual only cares about enjoying the game for him or herself, but in coordinating with others to achieve such satisfaction, each individual ultimately enhances enjoyment of the game for others as well. While such coordination or decisions may not always be optimal (and realistically speaking, they rarely are), the market (or in this case, the collective decisions of the players) constantly adjusts to approach such optimality. Allowing the relevant institutions (in the case of *Destiny*, those related to social platforms) to emerge organically, as opposed to attempts at predicitively defining them, promotes experimental competition and market discovery. As such, I argue that this tends to encourage more well-being in the long run.

### 3.2 Overview of Case Study Methodology

In order to effectively get at the question of how emergent order improves social outcomes for gamers, I am employing a case study of a few different mechanics within the *Destiny* universe. By using the more expansive term universe, I intend to imply the inclusion of factors within *Destiny* that lie outside of gameplay itself such as dedicated forums, matchmaking websites, and other user-generated materials. In addition, numerous articles have been written concerning various aspects of *Destiny*’s game world. All of these resources serve to highlight the effect and value of spontaneous order within the overall scope of the game.

In regards to specific instances of emergent social institutions within the game, I argue that *Destiny* promotes order on two levels, namely the functional and applied levels. The first deals primarily with a foundational level of order which the second then develops into well-established norms. As previously mentioned, these examples are by no means meant to be considered all-encompassing representations of how *Destiny* embraces such concepts, nor do they implicitly preclude negative outcomes from such institutions. Such an analysis is outside the scope of this paper. However, they do provide a solid starting point for considering how such institutions might enhance overall gamer welfare.

### IV. A CASE STUDY OF DESTINY

Now consisting of two games, *Destiny 1* and *Destiny 2*, as well as a number of expansions (all of these combine to form an ongoing continuation of the same game), the *Destiny* universe represents a dynamic and evolving world, especially in regards to its social institutions. Given its unique place in the repertoire as a hybrid MMORPG (massive-multiplayer online role-playing game) and first-person shooter, this game relies on robust social networks to not only survive, but to thrive. Despite this, very little infrastructure exists within the game itself to coordinate such social activity. While a formal automated matchmaking system does exist, it remains relatively limited in function compared to games similar in type. Nevertheless, players still end up coordinating, outside of any central direction, and as a result, some interesting norms within the game tend to form.

#### 4.1 The Functional Level of Order

This level of order deals primarily with the matchmaking and loot systems within *Destiny*, and it describes a foundation upon which player norms tend to develop. In regards to the former, a number of different options exist for matchmaking within *Destiny*’s gameplay. From a core perspective, the game offers a matchmaking service that differs from other multiplayer games by utilizing what has been termed mesh networking. This allows the entire game world to be populated at all times so that no one ever feels alone. This operates in stark contrast to many other games that have limited servers in which players may end up in an area completely by themselves.

In effect, an automated matchmaking service adds players to your fire team based on a number of different criteria whenever you enter a specific game mode. However, there are still some areas of gameplay that do not facilitate matchmaking through this automated system. In such cases, it becomes the player’s responsibility to find other people to join his or her fire team in order to participate in the given activity. For example, gameplay features such as the popular raids, nightfalls, Trials of Osiris, and Trials of the Nine require manually pre-made teams in order to participate. This has resulted in some controversy as some players prefer such a system while others think that automated matchmaking should be extended across all game modes.
To overcome the problem of finding players for such activities, numerous LFG services such as DestinyLFG have emerged which allow players either looking for a team or looking for more team members to coordinate on manual matchmaking. Essentially, players list themselves as either LFG (Looking for Group) or LFM (Looking for More), respectively, along with details around proposed game modes, platforms, and levels. Each group then can view listings in the other group to find individuals with whom to play and also send messages to connect. This process overcomes commitment and relationship issues associated with automated matchmaking while at the same time providing an avenue for those who do not have pre-formed teams to find individuals with whom to play. Other options for manual matchmaking exist as well such as the Bungie Forums and fan-created tracking sites. Furthermore, players are also free to create fire teams from their friend lists on PSN, XBOX Live, etc.

Through such varied approaches to matchmaking, Bungie has established a framework in which players have a large number of options for creating their fire teams. As a result, players are able to coordinate effectively with one another in order to find good players with whom to tackle the variety of tasks in Destiny’s game world. This order comes about not from centralized decision-making in which fire teams are determined directly by Bungie, but rather from a multitude of self-interested actions that lead to a wide number of beneficial outcomes for most players. Granted, as previously stated, there is a push to pressure Destiny to extend automated matchmaking to all game modes. However, given the recent spike in Destiny’s registered player base to 25 million users as well as the fact that average engagement is 3 hours per day, it is unlikely that this is a major widespread concern.

Ultimately, this structure comes about due to the lack of a more “complete” matchmaking service created by the game’s developer, Bungie. To that end, it plausibly represents a second-best alternative compared to centralized design. I argue, however, that it presents a more beneficial institution primarily since it allows for the formation of teams and the exchanging of information relevant to individual needs. Moreover, this system creates stronger incentives and accountability mechanisms to see activities through once begun as anonymity is at least partially removed. Regardless though, it still constitutes a defined social institution in which players can connect, share information, discuss strategies related to certain activities, and most importantly, socially engage in the game itself. As we shall see, this leads to some interesting norms for activities related to end-game activities and the Crucible, Destiny’s version of competitive multiplayer.

Another area in which Destiny demonstrates functional emergence lies in character customization and the procurement of loot. From the wealth of options available to create a distinctive character at the start of the game to a database of approximately 5,000+ items, many of which are gear pieces the player can equip, opportunities for creating a truly unique character abound. As such, this robust mechanism represents the other part of the foundation upon which social order emerges within the Destiny universe.

In fact, given the highly repetitive nature of this game, creating such a character and collecting this item database is likely the primary motivation behind most players’ sustained commitment to Destiny. In a video entitled “Destiny: The Hardcore Gamer’s Slot Machine” on his show The Point, Danny O’Dwyer explores what he considers to be ethical concerns around the perceived lottery structure of Bungie’s game. In essence, he argues that, given the monetary DLC component attached to the loot grind in Destiny, the game can be likened to gambling. Therefore, he suggests that it carries with it concerns around exploitation. Although such a moral question is outside the scope of this paper, the illustration he provides comparing Destiny to a slot machine remains highly apt as a description of the incentive structure in place.

Simply speaking, each activity in Destiny has a loot table, which continuously changes. From this loot table, there are set probabilities for the given activity to drop a particular item. Additionally, with various updates, there are also a few different methods of affecting these probabilities and ensuring at least some value out of every drop. Furthermore, a number of smaller, less valuable, items are guaranteed depending on the activity in which the player is participating. Notably, there are also a few items that can be acquired via micro-transactions, but these are generally limited to cosmetic items. However, some updates have included the sale of practical items as well. Regardless, the game rewards effort (via playtime) and dedication (via playtime and money allocation) through creating a system in which players constantly can work to improve upon their individual character designs and signal certain aptitudes/advantages.

As this all relates to spontaneous order, players are constantly pursuing self-interested motives in order to improve as well as diversify each of their characters, and often, these goals are in conflict with one another. After all, the piece of
gear or improvement that one player is pursuing may not be the same one that another player is pursuing. However, given different character builds, complementary abilities, and varying levels of progression, it becomes advantageous for such players to team up and work towards different loot goals together. Thankfully, the game provides plenty of opportunities for this and, most notably, not through design. Instead, it occurs in a truly spontaneous fashion coordinated through emergent social institutions. This leads to certain established norms which I explore in the next subsection.

Notably, this also facilitates greater team well-being as each player is satisfying his or her comparative advantage (based on character design, goals, and availability) through the division of labor in helping to achieve each group member’s goals. Logically, it would be highly unrealistic for a central planner to know not only which particular item a given player wishes to pursue, but also which combination of players will best help him or her acquire it. Within this game, we instead see players voluntarily forming mutually beneficial arrangements through emergent social apparatuses to accomplish various objectives without any sort of direction.

4.2 The Applied Level of Order
Building on this foundation, an applied level of order forms in which some interesting norms emerge within the Destiny universe. For instance, challenges such as the various raids provide opportunities for securing unique pieces of gear that cannot be found anywhere else. Why then might we see veteran players who have already acquired this gear continue to participate? Outside of simply enjoying the game or securing additional materials, a strong incentive exists for these individuals to help less developed players to learn and complete the raid, namely the reputational benefits and favor economy that emerge from this interaction.

Since the automated matchmaking system does not extend to the raids, players often must resort to LFG sites and forums to put together their fireteam and the most sought after players from these sites (especially for newer players) tend to be those who have already reached the level cap and completed the raid multiple times (this results from transaction costs related to the development of relevant knowledge). Such veteran players, in return, receive a reputation boost within the social community for their assistance and new connections upon which they can call to help complete other in-game tasks. Furthermore, even if the players one recruits are not veterans, the connections formed in this medium serve to develop new regular raiding groups and sub-communities with which players often continue to participate. Ultimately, these norms help build and establish lasting connections within the community that assist individual players within the game itself, and they extend to other end-game content and Crucible activities as well. Interestingly, much like other games, these connections also often go beyond the Destiny universe and develop into real friendships.

Another more concrete area in which emergent social institutions have established order can be found in the game’s customization system. This extends across both the cooperative and competitive multiplayer modes, and the sharing of information within the community leads to a set of standard norms for ability and gear selection. Examples for cooperative play include commonly understood practices such as using the weapon Gjallarhorn to quickly and efficiently take down the raid boss Crota, taking advantage of the Hunter’s invisibility ability to easily revive dead teammates, using the Titan’s bubble shield to easily generate more super energy, etc. Importantly, these norms are established both through word of mouth and the same forums/LFG sites that players use to construct fireteams.

Regarding competitive multiplayer, the Crucible can also be a fantastic demonstration of this norm. Given the strategic nature of competing against “real” players, choices surrounding weapon and armor selection can be crucial. The trick for each individual player is to decide on an equipment set that will deliver the strongest advantage in actual gameplay. This can vary depending on such factors as the type of match being played, the map that is randomly selected, and even the other players within the match itself. Regardless, certain consistent choices prove to be most advantageous in this mode, and much like its cooperative mode counterpart, these come about through the emergent social institutions at play. Examples of such norms here include the coordinated collection of both heavy ammo refills (to keep it away from the other team), use of standard sniping locations on each map, the formation of specific strategies in Control, the domination of certain weapons during particular versions of the game, etc.

These examples by no means constitute an exhaustive list of the norms that have developed within the Destiny universe. They do, however, provide a starting point for considering how order is spontaneously established within this game and the key role that emergent social institutions play in interacting with the given programmed environment. Furthermore, it provides a foundation for exploring other...
emergent phenomena within video games, and it helps us to understand how online communities naturally form and develop in a broader context.

V. CONCLUSION

In conclusion, video games serve an effective role in mirroring the economic outcomes of various institutions within society. As they essentially act as a natural experiment for examining economic inquiries with minimal consequences, they provide a solid avenue through which to explore various theories and ideas. Unfortunately, analyses concerning uniquely Austrian concepts within this medium have been woefully under-pursued. To that end, my contribution to the existing literature helps to fill this gap by exploring how emergent social institutions work to establish spontaneous order within the video game *Destiny*.

By examining two theoretical levels of order and considering previous research on the emergence of the firm, I provide several examples in which the formation of online communities helps to establish certain norms within the game’s universe. While this analysis is by no means exhaustive, it does provide an initial look into how these institutions operate. Sometimes, this takes the form of players solving various problems through innovative solutions and passing along the information through relevant networks. In other instances, it simply amounts to a naturally forming and mutually enforcing arrangement of decisions among several different players. In each respect, this order does not result from any centrally directed action by the game’s developer, but rather through multiple agents pursuing different goals in the service of self-interest.

Since many video games mimic reality (at least to some degree) through similar environments, the presence of human interaction, and robust economic systems, they likely have much to tell us about our own world, at least in regards to our online personalities. Given the naturally social element within *Destiny* as well as the fact that it is fundamentally a game driven by human players (as opposed to artificially intelligent agents), its universe is essentially an extension of our own. What applies to human interactions within this game world should also intuitively apply to reality (or again, at least the digital version of it). More research, both of the econometric and analytical varieties, should be done to demonstrate this claim. Examples might include the analysis of social relationships formed while playing these games as well as the transference of these principles to other “real-world” interactions. However, this paper provides initial support for such an argument.

In addition, there are a number of other lines of research that could be pursued from such a perspective. One notable extension of the current paper involves measuring the frequency, breadth, and regularity of matchmaking queries on *Destiny* forums and LFG sites. A few examples related to other topics include case studies of the same idea with different game genres, the interaction between spontaneous order and creative destruction in video game worlds, and the effects of entrepreneurship on the emergence of order. Other Austrian concepts could also be explored in relation to video games such as the application of the structure of production to video game worlds, the effects of government regulation on the industry, and the knowledge problem as it relates to the various decisions of players. In any case, regardless of how this stream of research evolves, video games can provide us much insight into the economic questions of both today and tomorrow.

Lehdonvirta (2005), for example, explains how the economic systems within video games, especially massive multiplayer online (MMO) games, mimic real life systems. There is also a growing body of literature in economics analyzing various aspects of video game worlds. Most of the research focuses on providing examples of, and support for, basic economic concepts (Bilir 2009; Castronova 2003, 2006; Hunter 2003; Lastowka and Hunter 2004, 2005). Salter and Stein (2016), however, provide a notable exception to this literature, using Austrian economics to understand how monetary institutions in Diablo II emerged to facilitate exchange.

Notably, however, a recent methodological work does explore the role of emergent order within virtual worlds more generally. Mildenberger (2015) explores economics as it relates to the anarchical world of virtual pirates. Although such individuals tend to be inherently conflict-loving, rules emerge to mitigate their otherwise destructive behaviors.

For a criticism of the Austrian concept of spontaneous orders, see Sandefur (2009).

Although V.L. Smith (1994, 113-115) contends there are likely more, the seven reasons he lists for economists to do experiments are to: 1) test a theory, or discriminate between theories; 2) explore the causes of a theory’s failure; 3) establish empirical regularities as a basis for new theory; 4) compare environments; 5) compare institutions; 6) evaluate policy proposals; and 7) test institutional design (V. L. Smith 1994). As it relates to the current paper, video games most fittingly serve purposes 1, 3, 4, and 5, though they could conceivably fit purpose 7 as well.

Notably, Destiny 2 attempts to address some deficiencies (as perceived by some of its player base) within the matchmaking system through the introduction of a formal clan system and the Guided Games mechanic. For more on this, see Eddie Makuch’s 2017 article “Destiny 2 Guided Games: Here’s What You Need to Know Ahead of Launch” at [https://www.gamespot.com/articles/destiny-2-guided-games-heres-what-you-need-to-know/1100-6435321/](https://www.gamespot.com/articles/destiny-2-guided-games-heres-what-you-need-to-know/1100-6435321/). While time will tell how these new ideas will influence the development of the franchise, they matter very little for the analysis of emergent social institutions within this paper. They do, however, present an interesting feedback development which begs the question, “Do game developers respond effectively to broad consumer interest or are such changes merely to appease a vocal minority within the fan base?”


See Matt Martin’s 2015 article “Destiny Players Grow to 25 Million, Putting in 3 Hours per Day” at [https://www.vg247.com/2015/11/03/destiny-20m-users-3-hours-per-day/](https://www.vg247.com/2015/11/03/destiny-20m-users-3-hours-per-day/).

A complete listing of all current loot can be found at Planet Destiny’s website, [http://db.planetdestiny.com/](http://db.planetdestiny.com/).


See the Destiny sub-reddit, [https://www.reddit.com/r/DestinyTheGame/comments/2ry8bj/can_someone_here_explain_how_the_loot_drop_system/](https://www.reddit.com/r/DestinyTheGame/comments/2ry8bj/can_someone_here_explain_how_the_loot_drop_system/), and Anthony Taormina’s 2015 article “Destiny’s New Loot System is a Huge Improvement that Rewards Player Effort” at [http://gamerant.com/destiny-review-loot-system-151/](http://gamerant.com/destiny-review-loot-system-151/).


See Brenna Hillier’s article “Destiny: How to Get a Raid Group Together Using LFG Sites” at [https://www.vg247.com/2015/02/10/destiny-how-to-get-a-raid-group-together-using-lfg-sites/](https://www.vg247.com/2015/02/10/destiny-how-to-get-a-raid-group-together-using-lfg-sites/).
REFERENCES:


This is a highly readable, meticulously researched, and superbly produced book. Turning the pages is a wonderful reminder of how poor a substitute e-books are for the best that fine paper, printing and binding have to offer. Princeton University Press is to be warmly congratulated.

To those reasonably familiar with the lives and writings of Hume and Smith, Rasmussen’s book will not have anything very novel to add, but it is undoubtedly, as one of the puffs on the back cover says, “an appealing introduction for the novice”. Yet, it presents the novice with a familiar, but highly distorted picture, in which Hume is the Enlightenment hero (with Smith his admirably faithful young, or younger, friend) who battles to bring light to contemporaries sunk in the darkness of religious bigotry and intolerance, and is abused and reviled as a result. To my mind, the book prompts this question: At what point does biography become hagiography?

As Rasmussen notes, when Hume moved from the Old Town of Edinburgh to a more spacious house in the Georgian New Town then under construction, the street in which his new house stood had not yet been given a name. Someone, perhaps a young woman to whom he was close, chalked “St David Street” on the wall. The name stuck, though almost no one who lives there nowadays knows how it came about. The appellation was something of a joke, of course, since Hume was held to be a religious skeptic, and above all a scourge of superstitious practices such as the veneration of saints. For Rasmussen, it seems, the appellation turns out not to have been a joke. Hume was a saint, if by saint we mean someone who approaches “as nearly to the idea of a perfectly wise and virtuous man, as perhaps the nature of human frailty will permit”—Adam Smith’s famous final tribute to Hume, which Rasmussen repeats several times. It is evident, I think, that Rasmussen does not merely admire Hume; he venerates him.

Is there anything wrong with this? What people commonly mean by hagiography is the embellishment of a biography by fanciful invented episodes that only the credulous would believe. Nothing of this sort appears in Rasmussen’s book. His claims are all as well grounded as the evidence will permit. And yet there are more subtle hagiographical elements. Let us suppose, though it is a matter to be returned to, that Hume was regarded in his own times as the opposite of a saint, someone who mischievously sought to undermine accepted values, and make a mockery of the simple faith of ordinary people. Rasmussen systematically inverts the meaning of the terms in which this judgment sometimes found expression. Thus ‘infidel’, the word he uses in his title, becomes a term of commendation, meaning ‘enlightened’, with the consequence that “notorious infidel” is an even greater compliment. Conversely, Hume’s critics are invariably referred to “the pious” or “the devout” in a way that makes these negative terms, synonymous with ‘bigoted’ and ‘blinkered’. Relatedly, the educated Scottish clergy with whom Hume engaged intellectually, and among whom he counted many friends, are referred to as such only occasionally. The term Rasmussen prefers is “Edinburgh literati”.

This helpfully sets them apart from their religious affiliation, though the label cannot accommodate Aberdeen clergy such as Thomas Reid and George Campbell, who thanked Hume for the intellectual stimulus he had given them, or Alexander Gerard to whom Hume (among others) awarded a prize. Nor does it do much to explain the offer of employment from Lord Hertford “noted for his religiosity”, and with whom Hume seems to have formed a most satisfactory relationship. How does this appointment fit
with “the customary clamour against Hume’s irreligion” which, Rasmussen tells us, was a key factor in the ending his employment? Rasmussen expresses some surprise that one of the Edinburgh clergy, Hugh Blair—“who was a minister, recall”—should expressly agree with Smith’s final tribute to Hume. It is worth adding that Blair was no ordinary clergyman, but Minister of the High Kirk of St Giles, and the most popular and effective preacher of his day. Similarly, Rasmussen notes the concern and solicitude about Hume’s health expressed by William Robertson in a letter to Smith, and the great loss Robertson says Hume’s death would constitute. He does not note in the same place that Robertson was Moderator of the Church of Scotland’s General Assembly for many years, which rejected an ill-conceived attempt to excommunicate Hume, as well as Principal of the University of Edinburgh where Hume, for reasons that remain obscure, had failed to get the Chair of Moral Philosophy.

This leaves us asking what we are to make of all those clergy (and others), who were tolerant of Hume, admiring of his intellect, and many of whom counted themselves personal friends. Were they, or were they not, among Rasmussen’s “pious” and “devout”? If we say they were, their tolerance and affection for Hume becomes unintelligible. If we say they were not, then their religious affiliation, it would appear, was superficial, hypocritical, self-serving and sustained only by the desire to protect their social standing. This alternative is not born out by any evidence. Their re-classification as “literati”, happily, allows us to lay the question to one side.

This is a major flaw on Rasmussen’s part, I think, because it overlooks the inconvenient fact that Hume’s attitude to religion was, and remained, ambivalent. To begin with, and contrary to what Rasmussen often implies, no more than anyone else did he rise completely above the assumptions and temper of his own times. On the contrary, what he writes about “monkish virtues”, superstition and “frivolous observances” would have been music to the ears of Protestants in general and Calvinists in particular. Equally welcome would have been the mocking tone he adopts in the Natural History when he recounts “the beliefs of our brethren the Catholics”, and his contention in the same place that the doctrine of the Real Presence of Christ is “so absurd that it eludes all force of argument”. Why would the Calvinist “pious” take offence at this? Moreover, far from concluding the Natural History with an easy or emphatic dismissal of religion, Hume muses. “Look out for a people entirely destitute of religion. If you find them at all, be assured, that they are but a few degrees removed from brutes”, and the persistence of religion, he concludes, cannot be simple foolishness. It is a “riddle”, “an inexplicable mystery”.

In the essay ‘Of Superstition and Enthusiasm’ Hume expressly refers to ‘true religion’. There has been much debate as to whether he really believed there was such a thing, and just what he meant when he identified it (elsewhere) as a “species of philosophy”. But he does say in a letter, in a spirit rather more agnostic than atheistic, that he only records doubts, and leaves the matter of convictions to others. Perhaps this serves to explain a measure of ambivalence even at the time of his death. It seems certain that Hume never exhibited the fearful, tearful penitence that some of the ‘pious’ might have wished him to do. Rasmussen, on the other hand, is determined that he faced death cheerfully, without thought of consolation. Yet in a letter to his ‘pious’ brother John, written just a few months before his death, Hume writes “Dr Black (God bless him) tells me that . . . I shall dye with much greater tranquillity in St David’s Street . . . Besides, where can I expect spiritual Assistance from Dr Blair and Dr Robertson? Not to mention a casual Exhortation from Carlylese or from you.”

Could Rasmussen’s Hume write this letter sincerely? Was he pandering to his brother’s ‘piety’? Was he mockingly blessing his attentive physician? Was he gently ridiculing his three clerical friends? The existence of any desire for consolatory spiritual assistance, it seems, is incompatible with enlightened sainthood. So what about this letter from St David?

Hume supported the idea of an established religion, giving preference (in common with most Scottish Protestants) to a religion that places ‘improving’ sermons at the heart of its observances. Why was this? Was religion something that had to be fed to the poor and ill-educated to prevent disturbances and make them content with their social subservience? I doubt if Hume held this view, but as Rasmussen observes in one or two places, unlike Smith (in the Theory of Moral Sentiments), he did not have much sense of the real role religion could play in the lives of those who were not “of the middle station” of life. Whatever Smith’s personal religious beliefs may have been, he understood how a deep belief in a justly ordered universe could elevate the lives of people whose lot was mostly endless toil, and could provide some consolation to the innocent victims of violence and
injustice. This important function, Smith thinks, “is of too much importance to the happiness of mankind, for nature to leave it dependent upon the slowness and uncertainty of philosophical researches” (TMS III.V.2). If, as Hume suggests ‘true’ religion is indeed “a species of philosophy”, it may serve to enlighten the educated and well-to-do, but it has nothing to offer the vast majority of people, living lives of a kind that Smith seems much more cognizant of than Hume.

There is no doubt that Hume differed sharply from many of his contemporaries on the matter of religion. He thought, as anyone who regarded the previous century could hardly fail to think, that religious conflict had been enormously damaging. Strictly speaking, however, it was religious factionalism, not simply religion in itself, that had caused so much suffering and disorder. Hume’s *bête noire* was faction, political no less than religious, which is why he often lumps “Whigs” and “Tories” with “Christians” in the class of the dogmatically partisan. Rasmussen misses this, I think, and ironically commandeers Hume as a partisan on the side of an anti-religious faction. This becomes clear in his treatment of Hume’s *Dialogues Concerning Natural Religion*.

There has long been an assumption that Philo is Hume’s voice in the *Dialogues*. This assumption confronts an interpretative puzzle. If Philo is Hume, what are we to make of Philo’s conclusion right at the end, that “a person seasoned with a just sense of the imperfections of natural reason will fly to revealed truth” and that “to be a philosophical sceptic is, in a man of letters, the first and most important step to being a sound believing Christian”? Why, moreover, does Hume not endorse Philo’s scepticism, but conclude the whole work with the judgment that “Philo’s principles are more probable than Demea’s, but that those of Cleanthes approach still nearer to the truth”? Is this Hume being inauthentic out of caution? Why would he see any need to do that? This was a work that he was very keen to have published, but the public, he knew, was not going to see until after his death when he would be beyond any “clamour” that it might cause.

In *The Moral Culture of the Scottish Enlightenment 1690–1805* (2014), Thomas Anhert makes a very strong case for thinking that the moderate Scottish clergy whom Hume knew well, would have subscribed to a view something like Philo’s. They had a more “just sense of the imperfections of natural reason” than the more zealous Calvinist, and this made them more inclined to look for “revealed truth”. It may even have inclined them to the view that being a philosophical sceptic is an important step in becoming a believing Christian. It is possible to read Thomas Reid in this light, and Hugh Blair described the *Dialogues* as “exceedingly elegant”. If this is correct, then the *Dialogues* leaves open the possibility of what the times might have called ‘enlightened religion’. According to Rasmussen, however, nothing of this sort is left open. “[I]n his published writings [Hume] had always refrained from marshalling all his sceptical challenges at once, thereby appearing to leave some kind of refuge for the devout. . . . The uniqueness of the *Dialogues* lies in its comprehensiveness, which leaves the pious reader no way out, no safe haven” (p. 188, my emphasis). Is this view borne out by the text of the *Dialogues*? Or is it rather, the view that Rasmussen would like to borne out by them?

In *Hume: a philosophical biography* (2016), James Harris argues persuasively that the form of the *Dialogues* is essential to them. That is to say, this is not a philosophical refutation that happens to be cast into a dialogue and might as well have been a treatise. Rather, it is an exemplification of how the discussion of contentious and difficult philosophical subjects, on which equally serious opinions differ, can avoid faction and dogmatism. It is a literary demonstration of how Hume thought ‘enlightened’ philosophy ought to proceed. The implication is that, contrary to what has been supposed almost without demur since Kemp Smith’s edition, Hume has *no voice* in the *Dialogues*. His own voice is the ironic one that composed the work. This way of construing the text does not solve every interpretative puzzle, but it does allow us to take it seriously without impugning Hume’s own integrity. Philo’s is not Hume’s *voice*. He is Hume’s *invention*, and consequently these closing statements are not stratagems to appease the wrathful ‘pious’ (*post mortem*!), but a final acknowledgement that no one position can dogmatically insist on being the last word on the subject. It did not prove easy for Hume’s critics to see the *Dialogues* in this light. Rasmussen’s treatment of the work shows that it does not prove much easier for his admirers to do so either. Thus it was predictable that (some of) the ‘pious’ would castigate Hume for them, and this may explain Smith’s hesitation in taking on responsibility for seeing them to publication. But it is no less predictable that convinced atheists would find in them a vindication, and hence a reason for Hume’s veneration. If Harris is right, both parties have missed the point.

Rasmussen correctly observes that “nearly everything Hume wrote bears on religion in one way or another”. It is a subject that seems to have had a great fascination for him. Yet Hume, in my estimation, lacked any religious sensibility, in a way that Nietzsche, for instance, arguably Christianity’s fiercest critic, did not. When it came to religion, Hume, we
might say, was tone deaf, and along with Smith, he uncritically accepted the Scottish Calvinist’s narrow conception of religion as theological doctrine plus moral injunction. Any wider conception of religion escaped him. In the period that Hume was writing, Wren’s great architectural masterpiece—St Paul’s Cathedral—had been in existence less than 50 years, Bach composed the *St Matthew Passion*, Handel composed *The Messiah*, and Charles Wesley wrote verses that came to be regarded as among the finest devotional poems ever written in the English language. None of these examples fits the category of ‘frivolous observance’, and it is absurd to suppose that everything good about them could be generated by a “species of philosophy”. This lack of sensibility did not prevent Hume from writing about religion with critical acumen, wit and occasional insight. There is nonetheless never any sense in his writings that he appreciated why religion matters to people, still less why it matters so much.

The same might be said of Rasmussen. His accounts of the ‘attacks’ on Hume by, for example, Bishops and Archbishops, can only construe them as diatribes of narrow minded intolerance. He evidently sees nothing that might be learned from such people, despite their being, as Hume acknowledged, among his most highly educated contemporaries. John Wesley has no criticisms to offer, he only “thunders” denunciation, and the “baneful and pestilential influence of false philosophy on the human heart” of which Horne, the Vice-Chancellor of Oxford warned, and about which philosophers from Plato to Nietzsche have been concerned, is not, it seems, a possibility that we ever need take seriously. Indisputably, religious adherents have their share of intolerance, prejudice and narrow mindedness. But they do not have a monopoly on it, and philosophers, it appears, can have their share.

This is an especially significant matter for Rasmussen’s book. He means it to be an account of “the friendship that shaped modern thought”. I will return to the matter of friendship, and leave aside the question of what exactly ‘shaping’ in this context means. The important point to note here is that one of the ways modern thought was shaped by the times about which Rasmussen writes, requires us to employ a conception that can find no place in the picture he paints. This is the ideal of the unity of ‘science and piety’. The pursuit of this ideal underlay the establishment, and shaped the development of college education in the American colonies, and subsequently the United States. Douglas Sloan’s celebrated book *The Scottish Enlightenment and American College Ideal* (1971) recounts this history, and a key figure in it is John Witherspoon. Witherspoon was allied to the evangelical or “Popular” party in the Church of Scotland, and so unquestionably falls within Rasmussen’s category of “the pious”. In 1768 he became President of the College of New Jersey (subsequently Princeton University). As President, and in the interests of religion, Witherspoon modernised the curriculum by introducing experimental science and the moral philosophy of Francis Hutcheson, ‘Father’ of the Scottish Enlightenment. A leading activist in the Presbyterian Church in the United States and a signatory of the Declaration of Independence, under Witherspoon Princeton provided a steady stream of leaders for the new Republic in politics, law and education. Similar curricular reforms were emulated across the country. How is this possible if Scottish enlightenment and piety were as far apart as Rasmussen depicts them?

This aspiration to a ‘unity of science and piety’ continued to be influential long into the nineteenth century, and the Scottish philosophers who were studied included Hutcheson and Reid, and later Brown and Hamilton. By contrast, after the third edition of 1804, Hume’s *Dialogues* were not reprinted until 1875. Something of the same gap occurred for almost all his works. It was the twentieth century before Hume’s philosophical work began to attract serious attention again.

Smith’s *Wealth of Nations* fared rather better. Five editions appeared in his lifetime, many editions after his death, and the work has never been out of print. Furthermore, it is possible to find reference made to it by both British and American politicians and policy makers. Rasmussen makes a case for thinking the Hume’s ideas influenced Smith very significantly, not only in the *Theory of Moral Sentiments* (about which many commentators would agree) but also in the *Wealth of Nations*. In the chapter that he devotes to it, Rasmussen finds Hume’s presence in several important places. He notes that Smith himself identified Hume as the sole writer to have noticed the wider social impact of the growth of “commerce and manufactures” and particularly on the promotion of personal liberty and security. He is even prepared to say that the “basic outlines” of “one of the most famous sections of Smith’s most famous book . . . are lifted straight from Hume [since] a similar narrative is present in both the Political Discourses (very briefly) and several volumes of the History of England.

Rasmussen is careful not to overstate his case, but he undoubtedly attributes a very significant role to Hume in the gestation of the *Wealth of Nations*. Since the work proved hugely influential, then even if Hume’s own philosophical
writings fell into the shadows for a good many decades, there is some support here for the subtitle’s claim that it was their friendship that shaped modern thought.

But what do we really know about that friendship? A perpetual source of frustration to biographers of Smith has been the lack of material relating to what we may call his inner life. He was not a great correspondent, and any papers that might have given some indication of thoughts he had not published, were probably among those his friends, on his instructions, destroyed. Even the evidence of what he did, where he travelled, whom he met is on the patchy side, and evidence for what he felt and believed is virtually non-existent. Our knowledge of what he thought is confined to what he thought fit to be published. Every biography, accordingly has of necessity to include a large measure of surmise—probably this, and possibly that. Rasmussen’s account of Smith and Hume’s friendship is no less given to surmise.

There is undoubtedly evidence of friendship, but at one point Rasmussen describes Smith as Hume’s closest friend. What are the grounds for supposing the Hume’s friendship with Smith was closer than many of the other warm relationships he had (with Drs Blair and Robertson, for instance)? It is impossible to overlook the fact that for long periods they did not meet, and for equally long periods they do not appear to have corresponded, or if they did, we no longer have a record of it. It seems that Hume often urged Smith to visit him in vain. When they were both in London for several months, Rasmussen says “we can be sure that they spent some of this time together”, but he is obliged to add, “though there is virtually no record of it”. When they overlapped in Paris, Rasmussen says that “Hume surely would have introduced Smith and the young duke to some of his Parisian friends and admirers”, only to add “though there is little record to go by”. Over the long years in which Smith retired to Kirkcaldy, his home town, to work on the Wealth of Nations, he was a relatively short distance by sea from Edinburgh, but visited very occasionally. Remarking on this period, Rasmussen says “though Hume and Smith had seen each other very little in the previous three years, they did their best to make up for lost time”. Once again he has to add, “Unfortunately, there is little surviving evidence of what they did or said during this time” (p. 207). It’s just a guess, then, that whatever they did, or did not do, was nevertheless “making up for lost time”.

Hume had many good friends. Smith was one of them. He wrote a beautiful account of Hume’s last days, and made him a most eloquent tribute. They shared a view on many matters, and there are striking parallels in some of the things they wrote. Hume’s influence can be detected not only in the Theory of Moral Sentiments, but in the Wealth of Nations. While there have been lengthy periods of neglect, their writings have generally stood the test of time, and continue to attract the serious engagement of philosophers and economists today. All these propositions are true. Even in combination, however, they fall dramatically short of supporting the claim that this was “the friendship that shaped modern thought”.
1. INTRODUCTION

Years after the recent economic and banking crisis, policy makers and economists are still discussing how we should reform our banking systems so that they can be more resilient. It has become conventional wisdom in the post–Great Recession literature to attribute to the banking sector and its fragilities the causes and length of the economic crisis and its capacity to rapidly spread to the whole economy. This emphasis on banks has obviously led governments to lay special focus on banking reform and banking restrictions, and to adopt a severely pessimistic view regarding the inherent fragility of banking systems.

Despite the fact that Goodspeed’s book does not address the current financial crisis, his historical analysis concerning the Scottish free-banking period and his novel interpretation of the Ayr (Douglas, Heron & Co.) bank crisis are a fresh contribution that has relevance to understanding the role of banks and regulation in economic crisis and where to pursue institutional reforms. Goodspeed’s book is a revision of the history of banking episodes in eighteenth-century Scotland, but it has substantial contemporary relevance and invaluable insights for policy makers in regards to the perils of banking regulation, the role of interest groups in shaping reforms of the financial sector, and the fragility or resilience of alternative banking systems. Conceivably, in order to build more resilient banking systems, we should scrutinize history to find successful banking experiences and compare them institutionally with our existent ones, so that they might illuminate the path to institutional reform (or at least illuminate the dangers of ill-conceived reform).

Approximately from 1716 to 1845, Scotland’s banking system was among the most dynamic, resilient, and competitive banking systems in Europe. The Scottish system effectively absorbed several economic shocks that affected the economies of Scotland and England and threatened the stability of financial markets. Scotland operated in a highly competitive and lightly regulated environment that had no central bank to act as a lender of last resort, no monopolist issuer of currency, no legal restrictions on entry, no ex ante binding limits on the number and size of banks, and finally no capital and reserve requirements (p. 7). Yet despite the lack of all those formal rules, governmental support, and explicit state-created institutions, the Scottish banking system was remarkably more stable and robust than the English system.

The book addresses in depth the institutional structure of Scottish banks and the wide forms of bank competition during the Scottish free-banking period, both of which helped the banks (and the system) to become more resilient than their English counterparts. It also explores in depth the regulatory environment in which banks operated, and how the ‘regulatory quality’ of that environment devolved due to interest groups and government involvement. More critically, it discusses the endogenous and collaborative institutional and contractual arrangements that banks generated (among themselves and with their shareholders) to attempt to endogenously and privately solve problems such as lack of a lender of last resort, volatile capital outflows, and banking crises without needing an outside agency. Indeed, the historical case reviewed could also be broadly interpreted as factual evidence of the real possibilities of robust institutional heterogeneity in banking, and how voluntary forms of decentralized self-governance, liability contracts, and polycentric structures are robust private means of efficiently governing banking matters without the need for centralizing banking services (Paniagua 2017).

REVIEW

Legislating Instability: Adam Smith, Free Banking, and the Financial Crisis of 1772 by Tyler Beck Goodspeed

PABLO PANIAGUA

Web: https://www.pablonpaniaguaprieto.com

Cambridge, MA: Harvard University Press, 2016, 224 pages, $39.95
Before 1765, the collapse of a bank did not entail any substantial threat to the stability of the system, mainly because no single bank acquired sufficient size, interconnectedness, and systemic importance that its failure would constitute a severe menace to counterparties and credit markets. Even if a bank had done so, the system had contractual mechanisms, such as optional clauses and unlimited liability, for mitigating systemic shocks (p. 137). The biggest threat to the Scottish system occurred during the financial crisis of 1772, in which sixteen banks failed. Goodspeed revisits this experience in detail, arguing that prior regulations and impositions from seven years earlier severely undermined banks’ resilience and flexibility to deal with the crisis. The fact that Scottish free banking and the Ayr crisis occurred during the time Adam Smith was working on *The Wealth of Nations* makes this historical episode particularly remarkable. Goodspeed revises Adam Smith’s arguments concerning free banking and regulation, concluding, contrary to Smith’s claims, that the crisis, far from revealing the weaknesses of free banking, actually supplies important lessons on ill-conceived and ill-implemented banking reform arising when banks as interest groups engage in regulatory capture.

2. THE BOOK ITSELF

The nearly 130 years of success of the Scottish free-banking period is of particular interest, not only for the history of banking, but also to understand alternative and plausible institutional banking arrangements that could be more resilient and stable than contemporary ones based on central banks. If we are indeed interested in reforming banking systems so that we could avoid or at least ameliorate the damage of economic crisis, then we should seriously analyze all the real institutional possibilities or structural liability reforms that we have available. The case of free banking in Scotland is particularly important for us today in a post–Great Recession world, especially in understanding contemporary debates on banking reform, which type of regulations to undertake, and more broadly how sound and resilient a free-banking arrangement would be compared to a centrally controlled one. I consider this case relevant for three reasons.

First, the Scottish system has been one of the closest (if not the closest) banking systems to that proposed in the theory of free banking: “Scotland from 1716 to 1845 is widely considered by economic and financial historians to have been one of the closest ever historical approximations to ‘free banking’” (p. 7). Hence understanding how the Scottish system actually functioned, and identifying its contractual and market mechanisms and how well it dealt with crisis, is a valuable way of understanding how a real (close to an ideal) free banking system could actually function.

Second, the system suffered in 1772 a financial crisis that severely threatened the system. This crisis has been considered by banking scholars an indication of how fragile a banking system would be in the absence of other institutional features such as stronger regulation, a lender of last resort, or even a central bank. Hence understanding the timing and legal context of the 1772 Ayr crisis, particularly in light of the recent drastic changes in the regulatory environment in 1765 and the political and macroeconomic context in which it occurred, would help to clarify misconceptions and doubts concerning the resilience of real free-banking systems and their vulnerabilities in the face of regulatory changes.

Finally, the third reason concerns the unique role that regulations and restrictions play—both directly and, indirectly, by interacting among themselves—in the evolution and erosion of the institutional structure of banking and finance. The Scottish experience offers a unique opportunity to identify the first- and second-order systemic and institutional effects of banking regulation, and how they can interact and unintendedly undermine other existent institutional features that are relevant to sustaining banking resilience. In addition, it sheds light on how additional and unforeseen regulatory complexity and interventions emerge from the unintended interactions among restrictions, and how that complexity is created from regulatory capture and vested interests. Particularly, it illuminates how top-down or ill-conceived and politically-led regulations can—even when well intended—*interact* among themselves, with prior existing restrictions, and with institutions, exacerbating the fragility of the entire system and bringing severe costs to society in ways that are impossible to foresee. Regulatory interventions therefore might not only have direct adverse effects, but more critically, as the case of the Ayr Bank suggests, they can have severe unexpected destabilizing interaction effects, undermining the whole institutional structure.

Thus there is a complex nonlinear relationship between politics, regulatory changes and the institutional development and fragilities that such regulations and their interactions will generate. In other words, there is a vital aspect of institutional and governance structures related to their en-
dogenous changes and their evolutions, that is significantly related to how we attempt to regulate and alter their environments and legal context; either through exogenously imposed political or legislative means. In fact, Goodspeed’s analysis actually demonstrates that ill-conceived and interest-group-captured regulation in concomitance with restrictions implemented in a nonlinear highly complex environment, might actually have the tendency to reinforce each other negatively, severely increasing both systemic banking fragility and its negative spill over effects into other markets in ways that are unforeseeable. Thus “adding complexity to complex systems—such as highly interconnected financial networks—can exacerbate the risk that adverse shocks are amplified and propagated throughout the system” (p. 22). Therefore, regulatory matters cannot be dissociated from institutional dynamics and changes, systemic risk, and complexity, making the art of politically-led regulation a very difficult and dangerous task. Recognizing those potential institutional erosions and fragility costs, alongside the dangers of enlarging the systemic fragility of banking systems that regulations could bring, should be vital theoretical aspects of economic regulation.

Goodspeed’s central argument is “that the salient [1772] financial crisis of the Scottish free banking period, the obtrusive exception to the hypothesis of greater financial stability under free banking in Scotland, was, pace Adam Smith, made more rather than less likely by precisely those regulated or “unfree’ elements of Scottish banking which the author of The Wealth of Nations promoted” (p. 8). Goodspeed notices that this conclusion should not be surprising, once we acknowledge that “the oldest, largest, and most established banks in Scotland … lobbied for … legal restrictions on banking; regulations that had the effects of raising barriers to entry, lowering competition in the provision of short-term credit, increasing the efficient scale of banking, and therefore, ultimately, amplifying the level of systemic risk” (Ibid.).

Despite these conclusions, Goodspeed clarifies that “it is not my contention that the introduction of legal restrictions into Scottish banking caused the 1772 crisis, but rather that they critically undermined the flexibility and resilience previously exhibited by Scottish finance, and thereby elevated the risk that adverse economic or financial shocks might metastasize into broader threats to financial stability” (Ibid., emphasis in original). The historical evidence suggests that the institutional and systemic resilience of the Scottish system drastically changed and was severely undermined during the regulatory changes of 1765–72. Despite those changes however, the system remained stable and responded relatively well in the face of such a severe global crisis.

Chapter 1, “A Very Melancholy Situation,” starts with a short introduction to the historical context in which the crisis occurred. Goodspeed briefly reviews the unfolding of the 1772 crisis—in particular, the liquidity problems for Scottish bills in the London discount market, and how that affected refinancing and the rollover of short-term debt of Scottish banks. The crisis affected several banks’ liquidity and soundness, but most importantly it ruined the “banking behemoth” Douglas, Heron & Co., the Ayr Bank. Goodspeed reviews how Adam Smith, during 1765 and 1772, advocated more regulation and restrictions on banks to deal with the problem. The contradiction however is that the reforms Smith advocated—mainly, a prohibition on small-denomination banknotes, a maximum legal rate of interest, and a prohibition of contingent-liability banknotes—“were already law seven years before the crisis of 1772. More curious still is the realization that these restrictions … were in reality the products of intense political lobbying” (p. 6, emphasis in original).

Chapter 1 also reviews how the Ayr crisis is a great example of what George Stigler designated the “theory of economic regulation”—specifically, the dynamics of regulatory capture. In the Scottish case, Goodspeed shows that we have a great example in which a government agency or a regulatory body designed and intended to serve the public interest “in fact serve[d] instead mainly to advance private concerns with concentrated interests in the regulated sector … The crisis itself presents a stark warning of the risks posed by regulatory capture, particularly in financial markets” (p. 21). Perhaps this is the most relevant argument Goodspeed gives as to why the 1772 crisis is relevant for us today, especially in the context of the international banking reforms such as Dodd-Frank among others, since it sheds light on how banking regulation and restrictions do not necessarily work in the public interest and to build resilience in our systems. In fact, quite the opposite could occur.

When regulations and restrictions are largely designed to satisfy concentrated interests, there is a substantial unperceived “institutional-fragility cost” that is shifted from the key institutional players into society, or onto the rest of the institutional structure. In other words, the dynamics of regulatory capture entail substantial concentrated monetary benefits for interest groups, at the expense of the transfer and dispersal of institutional fragility and its costs: the transmission of institutional fragility and financial risk.
to the rest of the system and the enlargement of systemic risks. Hence “capturing agents are essentially incentivized to produce negative externalities … [T]hese negative externalities can be especially costly when reduced competition gives rise to systematically important financial institutions” (p. 22).

The core insight here is that, institutionally speaking, regulations and restrictions are not only not positive sum, but also, quite certainly, negative sum. Banking regulation not only may not work to satisfy the public interest, but through time it can also undermine and deeply damage the public interest and their well-being, through destroying and eroding the long-term existent institutional resilience of a given system, making it more fragile and increasing its systemic risk. This is a relevant contribution to the theory of economic regulation and institutional analysis since it suggests that the social cost of regulatory capture is quite difficult to measure in quantitative terms. Its greater cost is increasing: regulatory complexity, long-term institutional erosion, and a gradual enlargement of the systemic risk and the spill over effects of banking failures into the rest of the economy, which amounts to more than a single shift in the cost-benefit equilibrium of the system. It seems the biggest welfare cost of regulatory capture resides in its capacity to undermine through time the resilience and the welfare-enhancing properties of a system as a whole.

Chapter 2, “Beggarly Bankers,” analyzes—contra Smith—the positive role that the emergence and proliferation of the so-called “beggarly bankers,” or the small-scale note issuers, played in actually bringing greater credit competition, diversification of risk, and macroeconomic stability; elements that contributed to building a remarkably resilient banking system that withstood severe adverse economic shocks prior to 1765. Goodspeed reviews here the tumultuous years between 1760 and 1765, which comprehended a severe balance-of-payments crisis in 1762, and how these events and the proliferation of small issuers affected Smith’s theoretical views on banking. Goodspeed also challenges Smith’s convictions that this period was marked by a “small note mania” attributable to the freedom of banks to issue banknotes for sums below £1 in addition to the free use of dilatory optional clauses in banks’ notes. Goodspeed finds “not only little evidence of a ‘mania,’ but also that the alleged banking offenses—namely, the issuance of notes in denominations bellow £1 and the adoption of optional clauses—were in fact rational and effective market responses to the very real challenges confronting the Scottish economy” (p. 30). Indeed, this chapter severely challenges the misapprehensions around the freedom to issue small notes and the use of optional clauses.

The Scottish economy was a small open and fast-growing emergent economy with a fixed exchange rate, open to large, volatile capital inflows and outflows that contributed to generating a balance-of-payments crisis and immense pressure on the financial system. Since the economy experienced large capital outflows due to international political events, the issuance of small notes was a rational endogenous response of the system to provide liquidity and credit in the face of an acute shortage of circulating media, particularly when big banks curtailed credit to attempt to ride out a real exchange rate depreciation. This 1760–65 period of relatively higher issuance of notes cannot be considered “manic” since it did not produce any marked difference in inflation rates in Scotland when compared to England (pp. 45–47). Furthermore, regarding the privately issued notes, a secondary and very dynamic market developed in which “there was certainly no shortage of willing bankers and agents to exchange them … [A] considerable ecosystem of quasi banking and clearing institutions formed around the issuance of small commercial banknotes” (p. 52).

The optional clauses, on the other hand, were rarely used and exclusively applied against “high-volume English speculators and arbitrageurs.” The clauses were “essentially a private application of capital controls on large ‘hot money’ outflows” (p. 30). The optional clauses therefore, were useful also to allow illiquid but solvent banks to liquidate assets without precipitating fire-sale losses and to form an orderly backstop against bank panics, and emerged endogenously from the oldest and biggest Scottish banks around 1730, which were originally trying to defend themselves from hostile note raids by rival banks. In addition, they eventually developed mainly as a useful and selective way around banks’ inability to respond to massive international capital outflows and external drains of specie, in turn because existing usury laws prohibited them to increase their deposit rates above 5 percent (p. 43). The optional clauses then were crucial to impose temporary and selective capital controls that “allowed time for the current account deficit to sort itself out through nominal adjustments in the bills of exchange market” (p. 128), while small-note issuers helped to avoid a contraction in the money supply and additional deflation.

These two features, though criticized by Adam Smith and banking scholars, were the product of competition and voluntary freedom of contact, and both contributed deeply to the resilience and flexibility of the system (pp. 45–47). They
also allowed the system to absorb and contain international financial shocks while reducing their capacity to metastasize into systemic banking panics and limiting their transmission to the real economy.

Chapter 3, "Procuring an Act," reviews the major "higher-order effects," or the unintended complexity-institutional effects that the 1765 restrictions played in the dynamics of banking and the stability of the overall Scottish system. This chapter traces the legislative history and origins of the 1765 Bank Act, reviews the details of regulatory capture in its formation and serves to contrast the drastic institutional effects that the new act, which was largely implemented at the insistence of pressure groups, had on the prior competitive system reviewed earlier. The act was intended to restrain the profusion of paper currency experienced during 1760–65, and it was therefore a legal response to the exaggerated views policy makers had at the time on the supposed small-note mania. The act "did indeed curtail the issuance of private banknotes … [T]hey did so at the cost of exacerbating the vulnerability of the Scottish financial system to adverse shocks" (p. 61). Unfortunately, the act "did nothing to resolve the fundamental problem … while at the same time it undermined some of the strengths that had previously enabled the Scottish banking system to absorb such volatiliy" (p. 129).

The act, by prohibiting both the issuance of small notes and the inclusion of optional clauses, increased the efficient scale of banking. Banks then got substantially bigger in size and lowered competition (quasi-banking or informal bank operations exited the market), having the double effect of amplifying the risk assumed by individual banks and also the system-level risk. Goodspeed notices that this result "should hardly come as a surprise … [T]he underlying purpose was to achieve precisely what it did achieve … to raise barriers to entry and limit competition" (p. 61). He specifically finds historical and statistical evidence of substantial institutional changes as direct consequences of the act, including an increase in the average bank size, substantially lower competition and fewer bank formations, drastic shifts in the composition and riskiness of banks’ assets and loan portfolios, shifts in the composition of reserves away from gold and silver toward big rivals’ bank notes, and an increase in the frequency of bank failures (from less than one per decade to more than eight) (pp. 75–84). In consequence, banks were "getting bigger [riskier, more interconnected], and there were fewer of them" (p. 78). The 1765 act therefore increased the efficient scale of banking, forced the exit of small-note issuers, elevated the level of counterparty risk, and created a rare institutional environment characterized by a “credit vacuum” (due to the exit of small-note issuers) that increased the likelihood that the institutional evolution of banking would lead to the establishment of a much larger, difficult-to-oversee, and more systemically important institution, as eventually did occur when the Ayr Bank was actually established in 1769 (p. 89).

Goodspeed exhibits here a case of regulatory capture, a substantial vested interest in the legislative process, and the extreme unintended negative consequences, or “higher-order effects,” of legal restrictions given the complexity and opacity of banking and financial markets. This case illustrates how different restrictions might actually reinforce each other unpredictably and undesirably, producing nonlinear impacts on banks’ size, counterparty and systemic risk, and institutional fragility. Here he finds that the restrictions and limitations on Scottish banking enacted before 1772 “elevated the level of systemic risk in Scottish financial markets,” concluding that “attempts to regulate specific categories of financial activity can therefore generate not simply offsetting changes in bank behavior, but changes that interact in often unpredictable ways with existing institutions, as well as with unanticipated changes in economic circumstances” (p. 138).

In chapter 4, "Prodigals and Projectors," Goodspeed addresses the crucial role that unlimited legal liability of shareholders in Scottish banks played in increasing system stability of the banking system and in securing a rapid recovery after 1772. He also convincingly argues that it was actually the unlimited liability of banks’ partners that worked as an efficient market-based lender of last resort that avoided disrupting credit markets. Despite the regulatory changes explored in chapter 3 that severely damaged the overall institutional structure and resilience inherent in the Scottish system, Goodspeed recognizes that “Scottish banking nonetheless retained much of its resilience” (p. 25). This fact is particularly relevant since the restrictions on small-note issuance and optional clauses had a significant negative effect in lowering competition in Scottish credit markets. Nevertheless, Goodspeed sheds new light on how resilient the Scottish banking system remained, mainly due to unlimited shareholder liability, appropriate rules for rapidly sequestering shareholders’ assets and providing for their equitable distribution among creditors, and robust co-partnery agreements (p. 100).

Setting aside the litigation among the Ayr Bank’s unlimitedly liable shareholders, the creditors and depositors of the bank were fully restituted between 1774 and 1786 through
the issuance of tradable eight-year, coupon-bearing bonds that were also secured by the Ayr shareholders’ personal assets and wealth (p. 113). With that mechanism, the Ayr Bank was converted into a sort of “bad bank” whose sole function was to slowly work off its toxic assets and gradually pay its creditors, while the assets and property of the bank’s owners functioned as a financial backstop. This particular market-based contractual solution to reassure, pay, and protect creditors is a valuable example of how a decentralized free-banking system could provide credible, organized, and legally enforceable private solutions to liquidations and bank failures.

This case illustrates how—even if a big bank fails in a free-banking environment, which is as plausible an outcome as in any other banking system—there are still several liability regimes and financial mechanisms to provide credibility and security to creditors and mitigate counterparty risk through endogenous and voluntary forms of contracting equity bail-ins. These forms of equity bail-ins have the positive effect of “unfreezing” and restoring calm and transparency to the credit and interbank lending systems, diminishing counterparty risk. They mitigate financial shocks that otherwise could have spread and affected the real economy, while simultaneously allowing mismanaged banks to deservedly fail, in turn discouraging long-term “too big to fail” problems and moral hazard. Furthermore, the issuance of tradable bonds backed by the Ayr Bank enabled all its counterparties to quickly resume lending and discounting bills after 1773, as these bonds were used by the remaining banks as collateral for loans (p. 112).

Hence through unlimited liability of banks’ shareholders and the issuance of bonds to creditors linked to those shareholders’ assets, the Scottish system had a built-in private contractual mechanism to restore and unfreeze credit markets whenever a banking crisis occurred. Through an expedient legal process of transparent, publicly advertised, and predictable sequestration and equitable distribution of the banks’ and owners’ liquidated assets, the system mitigated widespread bank panics and fire-sale liquidations (pp. 100–102).

The 1772 crisis was, no doubt, severe. However, comparatively speaking and unlike the Great Recession, most of the social and economic losses and distress eventually ended up being borne solely by the owners and directors of the bankrupted banks themselves, which had done, after all, a negligent job in securing sound banking practices, respecting the principles of copartnery, and enforcing managerial rules (pp. 117–21). Simultaneously and positively, by “bailing in” and making responsible and accountable the bank owners and directors for the misbegotten bank behavior, the system also protected note holders, depositors, and creditors and arrested the crisis and facilitated a rapid credit and economic recovery.

Goodspeed shows that the unlimited liability of shareholders of the bankrupt banks essentially served the homologous role of a private decentralized lender of last resort since the sequestration of shareholders’ personal assets “bailed them in” for more than their subscribed capital. Moreover, even if shareholders transferred their shares, the partners still remained indefinitely liable for all debts incurred during their ownership (p. 116). This form of voluntarily contracted lender of last resort, being market based (apolitical) and free of interest groups, might actually be a more institutionally robust alternative financial mechanism to provide stability and liquidity compared to a lender of last resort as represented by a government agency bailing out banks by socializing losses and using taxpayers’ money. A key difference between the Scottish lender of last resort and its more common form today resides in the fact that banks were indeed allowed to fail, and the cost of bankruptcy was borne by the banks themselves without inflicting losses on the rest of society. This should invite us to reflect more critically on how we allocate financial liability today (p. 139).

These properties allowed the Scottish system to recover quickly and to avoid a severe credit contraction that would have affected the rest of the economy. Indeed, the recovery of the Scottish economy after the Ayr collapse is particularly telling concerning the comparative macroeconomic recovery effects of alternative financial systems and different liability regimes. During its collapse, the Ayr Bank was allowed to issue transferable bonds to its creditors in order to fully restitute its claims. The bond issue also allowed the bank to satisfy through time the creditors while providing an orderly liquidation of the company’s assets and those of its partners. From that time (1773–74), credit markets thawed and the Scottish economy rebounded sharply (p. 111).

This historical case suggests that systems that possess sound, incentive-aligned, and market-based mechanisms for liquidating banks, in addition to “liability regimes that (unlike static, technical rules) automatically generate voluntary contracted, countercyclical equity bail-ins, may in fact challenge Reinhart and Rogoff’s conclusion that deep financial crises necessarily entail long, slow economic recovery” (p. 26, emphasis in original).
This point is particularly relevant today for diagnosing and understanding the slow and long recovery suffered by the United States. Whereas there might be other factors affecting the recovery, we have to acknowledge that a significant part of the explanation comes from the substantial damage that the banking system did to credit markets and to financial intermediation (Reinhart and Rogoff 2009). However, if we take Goodspeed’s insights seriously, the conclusions drawn by Reinhart and Rogoff (2009) in regards to the unavoidable “inherent drag” imposed on recoveries by banking crises have to be institutionally contextualized. The Scottish experience suggests that not all banking systems need to impose a severe, long, and slow recovery whenever a banking crisis occur. Institutional context, incentives for shareholders and bank managers, and the rules of banking liability matter considerably in determining the severity of and recovery from banking crisis. Painful and slow economic recoveries are not inherent, unavoidable outcomes of any banking systems failing or inherent maladies of capitalistic economies. They rather seem to be an institutionally conditioned outcome of fragile, inefficient, and overly (yet inadequately) regulated banking systems.

In chapter 5, “Upon Daedalian Wings,” Goodspeed acknowledges that the reforms and the regulatory changes lobbed for and implemented in 1765 were not simply and entirely an outcome of anticompetitive rent seeking or regulatory capture on behalf of Scotland’s biggest banks, “and ought therefore to caution us against too simple a conception of regulatory capture” (p. 23). Goodspeed acknowledges that the act was partly a genuine attempt of Scotland’s largest banks “to impose order on a highly difficult and poorly understood macroeconomic situation, an effort that, from their perspectives, was complicated by unrestricted entry of smaller banks and notes issuers” (p. 137). Nevertheless, those impulses for reform were—due to the complexity of the problem in a context of multifaceted banking regulation—ill directed. They led to a superficial interpretation and misdiagnosis of the problem since there was a general confusion in regards to the true sources of Scotland’s complicated troubles.

The 1765 legislation interacted severely and negatively with the existent usury restrictions, affecting the institutional evolution of banking and severely contributing to the Ayr Bank’s formation, massive size, risky insider-lending practices, lack of managerial accountability, and eventual demise (p. 123). The act did not solve the perennial problem of the balance of payments, but furthermore it “increased the likelihood that future financial shocks would be more disruptive” (p. 132). By increasing the average size of banks, it also undermined in part the benefits of unlimited liability in effectively monitoring management, thus making lack of managerial accountability severely problematic.

Finally, Goodspeed interprets the lessons from the 1772 crisis and the institutional evolution and demise of the Ayr Bank as ways to think more thoroughly about “the potential implications of particular types of regulatory change and institutional arrangements in financial markets” (p. 137). This case is an illustrative example of how banking practices, systemic risk, and management oversight are severely shaped, limited, and at times undermined by the constitutional context and complex legal system in which banks operate (p. 142; see also Paniagua 2017). With the Great Recession in mind, Goodspeed considers the Scottish experience as “a cautionary tale of the risks of rushing to regulate in the middle of an ongoing financial crisis and before the causes of that crisis are sufficiently understood” (p. 24, emphasis in original). It seems that today we have not taken that cautionary tale seriously, once again rushing to regulate banking without yet understanding the origins of the crisis. Hence perhaps today we are adding further complexity and planting the regulatory seeds that will make the next crisis not only more likely, but more severe.

3. CONCLUSIONS FOR AN ANTIFRAGILE RESEARCH AGENDA

As suggested, the Scottish system and free banking in general, possessed institutional and learning properties that greatly resemble systems that Nassim Taleb has designated antifragile. In other words, free banking possesses incentive and competitive structures, feedback mechanisms, and institutional properties that allow the system to become more robust from macroeconomic shocks (Paniagua 2017). It also has the discipline of markets and lacks bailout policies, which aligns the incentives of individuals in the system and incentivizes private bankers and customers to improve their behavior and develop market and contractual mechanisms based on past experiences. These incentives and virtuous institutional properties encourage the system as a whole to evolve, learn, and gain from economic stressors.

As evidenced throughout Goodspeed’s book, Scottish banking certainly resembled an antifragile system. The notion of antifragility resonates also with the concepts of institutional robustness and with polycentrity in political economy (Paniagua 2017). It would be certainly noteworthy to undertake an antifragile research agenda to develop
a framework of institutional analysis that incorporates all three concepts in order to analyze systemic risk and the macrostability properties inherent in different banking systems. Evidence suggests that free-banking regimes possess a great degree of antifragility. In contrast to Goodspeed’s historical case, contemporary societies have followed an alternative approach that, in fact, has driven us ever further away from antifragility into severe and ever-growing systemic fragility. Now that we are reflecting on the Great Recession, and looking for ways to fix our banking maladies, perhaps we should take a closer look at what banking history might tell us concerning the dangers of legislative reforms, complexity, and the alternative, robust institutional structures that we should consider more seriously.1

1 Indeed, of the sixteen banks that failed in 1772, only three failed to pay their creditors in full. However, their biggest creditor was actually the Ayr Bank, meaning that the losses were ultimately borne by the shareholders of the Ayr (ibid., 98). The other thirteen banks, in fact, transparently and in an orderly manner settled their debts in full, with stable expectations and confidence in the security and wealth of their proprietors to render them able to pay and be liable for more than enough to cover the liabilities.

2 I am grateful to Harry David for editing assistance.

REFERENCES

1. INTRODUCTION

It might seem odd today to review a book that was originally published in 1985. One could think this review arrived thirty years late, and hence it is engaging in an exercise of history of economic thought to reconsider some old ideas perhaps no longer relevant. Nevertheless, Don Lavoie’s *National Economic Planning: What Is Left?* ([1985] 2016), recently republished by the Mercatus Center, has not only retained much of its academic freshness and readability, but his ideas have preserved (if not increased) their relevance for contemporary discussions on public policy, the design of institutions, and political economy in general. This book is not only relevant but also foundational for future research on political economy and comparative institutional analysis and to enrich our understanding of existent institutions.

The reason why this book will retain its relevance resides in its explicit treatment of the inexorable role of incentives, power, and, more critically, contextual knowledge in affecting all forms of decision-making and governance structures. In addition, it examines how these elements are shaped by alternative rules (formal or informal) and the institutional framework, and ultimately how incentives and knowledge directly affect individuals’ capabilities to enact sound and wealth-enhancing policies. Since incentives and knowledge will always be crucial and pervasive in shaping human decision-making and capabilities, this book should always remain relevant.

The book critically suggests that social scientists cannot fully understand public policy’s effectiveness and the soundness of rules and governments’ decision-making—and their possible limitations and propensity to fail—if they neglect the role of incentives and contextual knowledge. Social scientists may keep discussing public policy, including the efficiency and wider functions of governments and other institutions, without fully acknowledging the role and implications of incentives, conflicts of interest, and personal knowledge. But in so doing, they are not overcoming or solving these inexorable problems (inherent in decisionmaking), which real decision makers must always face. They are instead ignoring or assuming these issues away. Therefore, they are engaging in an illegitimate and incomplete form of scientific and institutional political analysis. In assuming the aforementioned issues away while evaluating or proposing public policy and institutional solutions to be implemented in real life—and thus grounding their analysis on ideal theory and idealistic assumptions of human nature—social scientists are precariously proposing idealistic, unfeasible solutions to vital and pressing social issues. Unfortunately, such well-intended public policy proposals (based on questionable assumptions about human nature, incentives and knowledge) will always reveal themselves inherently fragile and susceptible to failure whenever confronted with real conditions of politics and human ignorance, including limited knowledge and misaligned incentives (Paniagua 2016). It is in reminding us to never forget these enduring, pressing issues and to incorporate them explicitly into our analysis of governments, public policy and institutions that Lavoie’s book will always maintain its relevance.

Lavoie’s book was originally published in 1985 as part of a two-book sequence. His other book, *Rivalry and Central Planning: The Socialist Calculation Debate Reconsidered* ([1985] 2015), which provided a novel review and account of the socialist-calculation debate, is more well known today in economics. That book could be considered an exercise in the history of economic thought, the history of market socialism, and comparative analysis. Most of the arguments contained in that book reiterate and appraise the original arguments given by both sides of the debate. Therefore, it was a vital retrospective and revisionist contribution about the history and unfolding of the debate and its arguments.
Lavoie’s National Economic Planning: What Is Left? ([1985] 2016) can be interpreted as a complement to the previous book, but in a way that looks forward rather than back on the arguments about knowledge and planning. This book builds upon the historical insights and arguments exposed on his previous book and enriches them with modern ideas concerning the philosophy of science, epistemology, and ontology. The outcome is a novel framework that incorporates past insights but enriches them with a multidisciplinary literature that makes the arguments more compelling than before and connects them to more explicit scientific foundations based in the philosophy of science. Finally, by successfully applying this framework to evaluate and assess different public policy ideas and notions of contemporary economic planning, Lavoie also provided the foundations for further applications in comparative institutional analysis. Whereas his other book helped rescue crucial past ideas in political economy previously overlooked by modern economics, the book under review instead provides a novel vision (and a tentative framework) for building a new multidisciplinary and scientifically broad political economy. Overall both books appeared as a serious challenge to the widely held view of the apparent success of market socialists in reconciling planning and markets. They also made economists more deeply aware of the technical, epistemological and scientific unfeasibility of socialism and central planning of the economy.

The next section outlines the book and calls attention to some crucial insights of Lavoie’s peppered throughout it. Section 3 indicates some of the recent explorations and far-reaching applications of Lavoie’s framework for comparative institutional analysis and his substantial contribution to robust political economy (Pennington 2011). Section 4 concludes by considering the work of Lavoie’s as providing a radical vision for a multidisciplinary research agenda.

2. THE BOOK ITSELF

Lavoie’s book was intended to challenge the new conviction that piecemeal or “noncomprehensive” planning is a much more feasible and perhaps even better alternative to comprehensive and large-scale national economic planning. In Lavoie ([1985] 2015), he shows how comprehensive economic planning fails severely to achieve market coordination and a rational economic order. He provides a powerful case for why comprehensive planning is economically irrational since it is based on unattainable and inaccurate philosophical and economic foundations and on idealistic and static assumptions concerning human knowledge. Lavoie ([1985] 2016) applies a similar analytical framework not only to comprehensive planning, but also to modern top-down contemporary forms of noncomprehensive (selective) planning. As the subtitle of the book suggests, part of its scope is to analyze what is currently left in the theory of national economic planning as a method of economic organization once it has been proven a failure as a comprehensive (total) approach. Next, it scrutinizes what is left of that theory and its implications for the current crisis of identity and of viewpoints in the intellectual left.

The book is both a theoretical development and refinement of the critiques of planning, and also a practical assessment of actual planning proposals along with evaluating contemporary planning alternatives. Lavoie sets the stage in chapters 1 and 2 by examining the inevitable “economic problem,” or how we must deal with socioeconomic coordination and scarcity. Building from Hayek, he describes economic and plan coordination as the main social problem all institutional arrangements must pay attention to in order to deal successfully with scarcity and economic survival (pp. 26–28). In so doing, Lavoie analyzes in depth the concept and positive role of “social intelligence” and the crucial role of social-relational dynamic processes in forming a better and “higher order” type of social relational intelligence. Subsequently he explores how the economic problem and social intelligence relate with each other and are dealt with by three alternative forms of social organization—tradition, markets, and planning—and how these forms of organization can actually deal with scarcity and coordination. Lavoie argues that tradition, albeit a primitive form of coordination, is able to use informal rules and taboos to successfully coordinate and organize simple and small societies. Markets instead allow for extending the complexity and dynamism of coordination to new realms of social activities and interactions, deepening and extending collaboration, the division of labor and knowledge and increasing well-being. Planning, however, according to Lavoie, attempts to replace tradition and markets with a more comprehensive and “rational” social order. It attempts to consciously and explicitly apply reason and science to coordinate and establish a path for future developments and to control the destiny of society, all grounded in a single, overarching design.

Lavoie also argues that planning suffers more critically than markets and tradition from two severe problems. The first problem is the “power problem” generated by planning, in which there is a dangerous and excessive concentration of control and economic and decisional power in the...
hands of a planning authority. The whole idea of planning is to radically replace the decentralized decision-making process of the market with a centralized entity. This institutional shift has deep repercussions for social control, the balance of power and governance in society, undermining individual choice, previously established interactions, and freedom of association. Furthermore, this top-down control of society can lead to severe unchecked abuses of power, severely hampering human dignity, morality, liberty, and well-being (pp. 20–21). The second difficulty, which according to Lavoie is the most fundamental one, is the “knowledge problem,” which he reviews in depth in chapter 3.1

Chapter 3 “contains the central argument of the book” (p. 5); hence I will put particular emphasis on it. Here the author lays out a thorough description of the market process and develops fruitful analogies of it with interesting examples of “organizational social patterns” and social processes borrowed from the fields of modern biology and scientific discoveries. Here is where he also lays out in more depth and clarity than ever what he considers the “knowledge problem.” Lavoie’s restatement of the knowledge problem does more than simply reiterate Hayek’s and Mises’s ideas concerning economic planning, and it could be considered instead as redefining and enlarging the arguments they raised. By borrowing from the growth-of-knowledge literature, sociobiology, the modern philosophy of science, and Polanyi’s sociology of scientific knowledge (Polanyi 1951), Lavoie redefines and broadens the knowledge problem. He puts more emphasis on the individual, socio-relational and tacit aspects of economic knowledge and more critically on the contextual and institutionally dependent aspects of the growth and emergence of such crucial knowledge (pp. 65–81). Ultimately, he argues that economic rationality and the efficiency of the market order are not social elements that can be fully defined and designed ex ante or predetermined with a static fully explicit plan. Rather, they are special and socially-conditioned complex emergent phenomena stemming from specific market-based social relations and market institutions. This critical point suggests that only specific types and combinations of orderly social relations and institutions can allow economic knowledge and the unique ontological market properties relevant for coordination to form at all. These crucial emergent and ontological social properties will not arise without the market institutional context and the use of money to relate (Paniagua 2018). This chapter sheds light on the indivisible relationship between institutional analysis, epistemology, and social ontology (Lewis and Lewin 2015; Paniagua 2018). The chapter also suggests that the market’s competitive process itself, in addition to the local interactions between individuals under market institutions (mediated through the use of money), is the actual foundation and unique context in which novel and more intricate “higher order” market knowledge emerges and gets communicated. In Lavoie’s words, what central planners have failed to notice, and their critical epistemic flaw, is in overlooking that the competitive market is itself the primary source of knowledge about which goods are to be produced and which production methods are feasible ... the social function performed by a particular complex of legal and market institutions makes them indispensable tools for the solution of certain unavoidable economic problems. (p. 4)

Contemporary planners and market socialists who have sought to guide and design the economy have partially recognized the power problem and the control-political dangers individuals face in relinquishing power and governance to a centralized agency. However, they have largely overlooked the epistemic and ontological implications they will face in establishing central planning over decentralized decision-making processes and market institutions. Thus they have disregarded the severe limitations—and even impossibility—they will face in producing and acquiring relevant procedural economic knowledge. This problem is actually institutionally inexorable and particularly acute whenever they attempt to replace, undermine, or govern some market-institutional relational aspects foundational for economic knowledge and coordination to emerge in the first place. Consequently, both market institutions and social relations perform unique epistemic and complexity-ontological functions that are vital to deal with scarcity and to ameliorate the economic problem while improving plans’ coordination. This novel institutional and social-procedural argument concerning how economic knowledge and collective intelligence are socially generated is crucial but still almost entirely unaddressed by the literature on planning, comparative institutional analysis, and politics.

Whereas contemporary proposals for planning have focused on retaining democratic values, incentives, “good people” in “the right place” and better computational-information processes, they have regretfully remained silent about the role of rules, institutions, and decentralized social interactions in forming knowledge itself—knowledge that planning theories simply assume to exist as available
data and thus as both already produced and already codified (and therefore of easy access). In so doing, they erroneously treat market knowledge not only as ‘already there’, but also as ontologically reducible to the pre-existent epistemic resources held by agents in isolation (regardless of their social relations); rather than focusing on how individuals, by being actually social-relationally organized through money in specific market contexts, produce a type of complex intricate knowledge arrangements that did not previously exist (pp. 65-76; Paniagua 2018). Hence planning proposals commit the fatal error of assuming that the emergent epistemic and ontological properties crucial to impart rationality and coordinate markets can be disassociated from, and can exist anterior or outside of, the actual social processes and institutions that allow them to exist. Social scientists have thus disregarded the contextual complexity of social orders, the ontological category of emergence, and their implications for knowledge in market institutions (Lewis and Lewin 2015; Paniagua 2018). Regarding market knowledge, there is a complexity and an ontological aspect of its emergence, suggesting that the “overall intellectual capacity of several interacting intelligences may be quite different from that of its constituent parts … [W]hether the whole would be greater or less than the sum of its parts … depends … crucially on the method of interaction among them” (p. 27, emphasis added).

As suggested Lavoie’s critical insight is that economic knowledge is not only dispersed, local, and hard to articulate (tacit). More importantly, it is a conditional and contextual-emergent property of the unique combination of decentralized interactions and the use of money in competitive social relations. Following Polanyi (1951), he sheds light on the complexity aspects of knowledge as a social phenomenon and the epistemic emergent ontological properties that competitive social relations in markets generate. A vast extension of economic knowledge therefore appears to be an ex post indivisible complex outcome of a system of social relations sustained by money and competition (pp. 76-86; see also Paniagua 2016; 2017). In the words of Lavoie: “The spontaneous order that emerges on the social level is the outcome of the rivalrous competition among individuals. It is a higher-level order that evolves out of a furious turmoil of lower-level disorder” (p. 69). Replacing or altering the exclusive interactions and social processes of the market will deprive the social order of the framework and context in which new and wealth-enhancing “higher order” ontological and epistemic properties can emerge.

I consider Lavoie’s reframing of the epistemic problem along the lines of sociobiology, complexity and Polanyi’s sociology of knowledge a very important contribution, one that cannot be said to simply restate anterior arguments (pp. 65–85). Furthermore, it invites us to think deeply about the ontological properties of economic knowledge and markets and the role of rules and social relations by which such emergent coordinative properties can arise (Paniagua 2018).

The indivisible relationship between the existence and growth of economic knowledge and the institutional context for the rules and social-relational procedures that constitute it has been ignored in the social sciences. Hence Lavoie suggests that “when we study a social system, we have to focus on the method of mutual coordination among the individuals” (p. 28). Overall these first three chapters can be read independently of the following ones since they provide the core theoretical background by which to evaluate any form of public policy, planning or government attempt to control aspects of the economy. Moreover, these chapters can be read as a key contribution to an emergent-social-order perspective on knowledge.

After providing the theoretical framework in chapters 1, 2, and 3 that explains in depth the power problem and the knowledge problem, Lavoie systematically proceeds to scrutinize different contemporary practical forms of non-comprehensive planning; such as the macroeconomic aggregative planning represented by Leontief’s input-output method (chapter 4), broad economic democracy, which seeks to extend democratic inclusion and participation in private economic decision-making (chapter 5), and selective reindustrialization policies known as “structuralist” strategies, which seek to modernize and revitalize particular relevant or “structural” sectors of the economy (chapter 6). Here Lavoie goes beyond the theoretical arguments and focuses on a practical analysis of what occurs when non-comprehensive forms of planning are actually established, as well the implications. Taken together, these three chapters are excellent examples of how in practice the power and knowledge problems can be applied to provide an institutional-epistemic analysis of governance’s alternatives and assessment of real public policy and their limitations. Thus they allow readers not only to understand the severe limitations of contemporary proposals, but to get a notion of how to apply the theory to actual institutions and governments’ policies in order to analyze their merits and limitations regarding authority, knowledge, and incentives.

What all these three proposals have in common, and the reason why they fail to achieve their rational economic
goals, is that all of them are variants of economic planning, which seeks to institutionally centralize control, decision-making powers, and governance in the hands of a few or a group. Moreover, they seek to replace individual sovereignty, social relations, and local knowledge with centralized group command and data aggregation. They are forms of top-down or group systems that override or dominate individuals’ local knowledge, their relations, and their decision-making powers. Therefore, they are extremely vulnerable to both the power problem and more critically to the knowledge problem, preventing them from generating and using the contextual knowledge required to make rational economic choices. By replacing decentralize interactions and decision-making in the market for democratic or political command, they simultaneously become unable to generate knowledge and social intelligence required to make accurate decisions and to impart economic rationality. It is ultimately this crucial unavoidable structural relationship and trade-off between the replacement of institutional market arrangements and social interactions, on the one hand, and the loss of both context-specific knowledge and complex ontological properties, on the other hand that lies at the core of the fragility of planning and inexorably leads planning to fail. Recognizing this inherent relationship and unavoidable trade-off undermines all arguments for non-comprehensive economic planning and provide the key insight to be highly skeptical of all sorts of public policy and government institutions that seek to guide and plan complex social orders.

Chapter 7 addresses the ideological and practical challenges the modern left faces given the failures of planning previously exposed. Lavoie argues that the left made a severe mistake in embracing both comprehensive and non-comprehensive planning as means to a rational social order. The chapter demonstrates with its historical analysis of the power problem that the ends the radical left sought to pursue are paradoxically undermined and systematically violated by national planning. There is an historical and factual abyss (and perhaps a lethal contradiction) between the laudable goals of the radical left and the real outcomes of central planning it has supported. Planning has further militarized the economy and oppressed individuals. The chapter also provides a historical overview of the intellectual development of the left, militaristic control of society, and the left’s practical experiences with planning, along the way suggesting where they went astray.

Finally, the appendix provides an invaluable account of the modern philosophy of science and its implications for epistemology, public policy, and economics. Lavoie does this to defend his arguments on tacit and emergent knowledge against the claim that they do not rest on scientific foundations. By doing so, he actually shows that the current disavowal and demise of positivism and objectivism within the philosophy of science indicates that economic planning is essentially unscientific and built upon an erroneous positivistic conception of scientific discoveries and knowledge (pp. 261–64; see also Polanyi 1951).

3. EXTENDING LAVOIE’S WORK: ROBUST POLITICAL ECONOMY

The relevance of Lavoie’s work today for political economists is not only in clarifying the insurmountable epistemic problems with central planning, but in thinking about what alternative institutional arrangements are in fact feasible and perhaps better than existent ones. The larger contribution of this book is to lead us think critically about the properties and robustness of our existent social organizations and institutions and see if there are relatively superior alternatives that we can move toward. His insights can be fruitfully extended to compare and evaluate institutions in order to understand their economic and social implications. In other words, we “must dare to imagine radical alternatives to the kinds of social institutions to which we have become so accustomed, and to explain why these new alternatives would work qualitatively better than the present ones” (p. 16). Given the pressing socioeconomic difficulties we face, we need to seriously question and evaluate how our existent institutions actually operate and to see whether they enhance human well-being. Ultimately societies benefit from continuously and thoughtfully experimenting with institutions to find radical alternatives or improve existent systems when they are found lacking.

Lavoie’s work suggests we should concentrate on real conditions and abandon the idealistic assumptions concerning politics, incentives, and knowledge that very often have plagued our theoretical analysis. In other words, it asks us to abandon ideal theory for real comparative institutional analysis—a form of inquiry that responds to Buchanan’s plea to abandon romanticized visions of politics and institutions. Hence it is an open invitation to a research agenda for comparative institutional analysis ‘without romance’. We must analyze alternative arrangements and see how they actually deal with human imperfection and whether in practice they produce beneficial or detrimental outcomes. We must think more seriously about the long-run effects of
those institutional arrangements when individuals are not perfect.

The radical agenda Lavoie gave social scientists is to constantly think critically about how all real-life institutional arrangements need to manage and ameliorate both the knowledge and the incentive-power problems—inexorable problems that decision makers will always face regardless of the institutional arrangement. The book could be seen as providing a way to extend and apply these crucial insights to a variety of proposals and institutions that have not previously been analyzed critically through a framework that takes into account Lavoie’s considerations seriously. It provides fertile ground for open-ended and interesting applications of the power and knowledge problems for scholars working in comparative institutional analysis and public policy. There is a wide range of contemporary social problems and original public policy proposals that aim for non-comprehensive planning solutions but might be amenable to critical scrutiny under Lavoie’s framework.

In fact, Lavoie’s framework has already been applied in political economy under the name of robust political economy (RPE). Thus the book has already contributed to an open-ended ongoing research program. Robust political economy means the analysis of comparative robustness of institutions and political-economic systems, or their actual ability to produce welfare-enhancing outcomes despite severe deviations from ideal conditions and problems concerning individuals’ incentives, motivations, and knowledge (Pennington 2011). Thus it relaxes the assumptions of complete knowledge, perfect rationality, and benevolence and asks whether institutions can still accomplish their objectives in the face of those defects in human nature. RPE seeks to address both the knowledge problem as laid out by Lavoie and the political, power, and incentives problem as further elaborated in the public-choice approach of Buchanan and Tullock in order to evaluate how resilient alternative arrangements are. Taken together the two approaches provide a strong—less idealistic and more conscientious—foundation for analyzing public policy and comparing institutions.

There have been interesting applications of the Lavoie-RPE approach in different subjects, expanding our understanding about the comparative robustness of alternative institutions in areas such as monetary policy and banking (Paniagua 2016; 2017), among others (see also Pennington 2011). It would be interesting to see scholars apply the Lavoie-RPE framework to institutional proposals in other fields such as economic development, immigration, and financial regulation. It would also be of interest to apply the RPE framework to new theories that frame governments’ economic involvement as an efficient and dynamic way of generating “entrepreneurial innovation.” Such explorations will help scholars engage more productively in contemporary debates in public policy about “government entrepreneurship” and state-led innovation.

4. CONCLUSIONS FOR A RADICAL RESEARCH AGENDA

Lavoie’s work shows that planning faces severe, insurmountable problems concerning power, incentives, control, and knowledge. Both models of economic planning (comprehensive and noncomprehensive) are unable to solve in a rational and efficient manner the basic economic problem of how to coordinate activities. Despite showing that economic planning is unfeasible and even contradictory to our values of freedom, dignity, and abolishing political oppression, Lavoie does not reject planning in favor of the status quo. In fact, he proposes a “more scientifically sound radicalism” (p. 1). The only viable solution to these insurmountable problems, Lavoie suggests, resides in radically decentralizing governance and decision-making, allowing individuals at the local level the liberty to use their personal knowledge and resources as they see fit. By doing so, they can rely on voluntary cooperation, relations and market institutions to generate the crucial conditions required for forming and communicating relevant knowledge. His radical proposal resides in a complete and emphatic rejection of planning and an embrace of decentralization, freedom of association and markets as the mechanisms of social organization and well-being.

Facing the crippling epistemic limitations of planning for attaining the left’s long-desired goals, Lavoie suggests a radical, scientifically rigorous, market-based, and decentralized solution to our social problems. His proposal challenges modern radicals to scrutinize further their own preferred institutional proposals, given the inexorable trade-off between planning and knowledge explored in section 2. Economic planning is epistemologically unworkable and furthermore leads inevitably to strong control and to a reactionary, militaristic type of policy-making that goes against the very same goals and values the radical left seeks to pursue. Therefore, such forms of organization and planning must be abandoned by both the radical left and by conservatives.
Lavoie’s radical vision of society should still be relevant and appealing today to everyone concerned about human dignity and freedom of association. The book asks us to think creatively about economics and political economy. I recommend it to every social scientist interested in building new scientific paradigms and in developing new avenues of research based on multidisciplinary scholarship. It will help scholars think creatively and differently about their own fields of expertise. Lavoie’s framework is particularly useful to start dealing more explicitly and thoroughly about knowledge problems and epistemic constrains and incorporating them in our analysis of comparative politics and institutions. The enduring significance of the book is to remind us to constantly push back against romanticized political theory and idealistic economic policy and to always question the unwarranted epistemological and incentive assumptions scholars often make about humans in general and relevant decision makers in particular. Finally, Lavoie’s contributions can be the foundation for novel research projects in comparative institutional analysis and public policy to understand the relative robustness or fragility of different institutional solutions to our pressing social issues. Henceforth this book does not simply clarify past ideas, but serves as a relevant and vital part of the current extended academic dialogue that is essential for the advancement of social sciences in general and for an unromanticized, rigorous, but multidisciplinary political economy in particular.

NOTES

1 Planning is here defined as “only those policy measures that involve concentrating power to shape the economy in a special government agency” (p. 2). Comprehensive planning is the complete abolition of private ownership, prices, and money and the complete substitution of market institutions and their social processes with government control. Instead, noncomprehensive planning means all variants of partial control over some areas of the economy while still retaining some market institutions.

2 “The knowledge problem is the contention that a central planning board, even if well intentioned, would lack the knowledge to combine resources in a manner that is economic enough to sustain modern technology” (p. 52). Lavoie adds emphasis on the fact that the problem of obtaining knowledge is not about gathering information and data, but about social access to and communication of a certain kind of practical knowledge that is both hard to convey (tacit) and also whose mere existence is conditioned by a set of institutions and social relations.


4 This insight does not invalidate or exclude the other aspects of the knowledge problem. In fact, Lavoie throughout the book uses all the different aspects of it. For tractability, the knowledge problem comprises three aspects: First, knowledge is fragmented and subjective and thus possesses deep local and individual properties. Second, some key knowledge is actually tacit and skill embedded and therefore unable to be conveyed accurately through language, questionnaires, or statistics (Polanyi 1951). Third, some knowledge needs to be socially generated and discovered under specific institutional contexts. Hence knowledge is also contextual and procedural. It emerges as a complex property of a competitive system of social relations which produces a type of knowledge nonexistent anterior to or outside of those rivalrous social processes that specifically constitute it (p. 6). Moreover, such “higher order” knowledge possesses ontologically and qualitatively distinct properties irreducible to the fragmented knowledge held by the agents anterior to or outside of specific market relations (Paniagua 2018). All three aspects of the knowledge problem complement and reinforce each other, making an epistemological and ontological case against central planning much stronger than previously recognized.
REFERENCES


INTRODUCTION

This review of Erik Olin Wright’s *Envisioning Real Utopias* (2010) will focus on his understanding of ‘capitalism,’ his conception of worker cooperatives, and the general issues surrounding markets, the Left, and Marxism.

THE FUNDAMENTAL MYTH OF ‘CAPITALISM’

The disagreements run very deep—even to the very designation of “the system” to be changed, i.e., “capitalism,” and the system to replace it, i.e., “socialism.”

So let’s start with Wright’s definition of “capitalism” as the target of his critique.

In capitalism, the means of production are privately owned and the use of those means of production is controlled by those owners or their surrogates. The means of production by themselves, of course, cannot produce anything; they have to be set in motion by human laboring activity of one sort or another. In capitalism, this labor is provided by workers who do not own the means of production and who, in order to acquire an income, are hired by capitalist firms to use the means of production...

Economic coordination in capitalism is accomplished primarily through mechanisms of decentralized voluntary exchange by privately contracting parties—or what is generally called “free markets”—through which the prices and quantities of the goods and services produced are determined...

The combination of these two features of capitalism—class relations defined by private ownership and propertyless workers, and coordination organized through decentralized market exchanges—generates the characteristic competitive drive for profits and capital accumulation of capitalist firms. (Wright 2010, pp. 22-3) [unless otherwise indicated, all quotes are from the June 2009 version freely available on the internet and thus unencumbered by “capitalist intellectual property rights” (p. 232)]

Wright’s article in *Jacobin (Online)* (12.2.2015) was a bit more concise.

Capitalism as a way of organizing economic activity has three critical components: private ownership of capital; production for the market for the purpose of making profits; and employment of workers who do not own the means of production.

Thus attacking “capitalism” is assumed to involve attacking: (1) “private ownership of capital” or more broadly in Marxism, the private property system, (2) “coordination organized through decentralized market exchanges”, and (3) “employment of workers who do not own the means of production.” The alternative of “democratic egalitarian socialism” would thus negate these three features.

In attacking “private ownership of capital” or “private ownership of the means of production,” it behooves the critics to at least understand those legal notions in the “capitalist” system (also a misnomer; see below). The concept is inherited straight from Marx whose notion of “private ownership of the means of production” was one of his biggest blunders. Marx understood that in the feudal system, the rights to govern the people living on the land and the right to appropriate the fruits of their labor were all part and parcel of the “dominion” of the “Lord” of the land. As Frederic Maitland put it: “ownership blends with lordship, rulership, sovereignty in the vague medieval *dominium,*...” (Maitland...
1960, p. 174) or in the words of Otto von Gierke: “Rulership and Ownership were blent” (Gierke 1958, p. 58). But then Marx blundered by carrying over that notion with capital replacing land in the modern era.

It is not because he is a leader of industry that a man is a capitalist; on the contrary, he is a leader of industry because he is a capitalist. The leadership of industry is an attribute of capital, just as in feudal times the functions of general and judge were attributes of landed property. (Marx 1977 (1867), pp. 450-1)

Hence the title of Marx’s opus magnum and the very name of the system “capitalism.” The standard form of this conceptual mistake might be called the:

Fundamental myth: that the legal rights to appropriate the product of production and the governance rights over the production process are part and parcel of (i.e., “an attribute of”) the “ownership of the means of production.”

This myth is one point of total agreement between Marxists and many “capitalist” thinkers—which is why Marxists have for some time functioned as ‘capitalist tools’ (Ellerman 2010). However, one of the most philosophically astute and careful defenders of the current system, Frank Knight, pointed out that the Marxist label is a misnomer since capital can be rented out and the legal party undertaking a production process (Knight’s “entrepreneur”) might not actually own the capital goods or “means of production.”

Karl Marx, who in so many respects is more classical than the classicals themselves, had abundant historical justification for calling, i.e., miscalling—the modern economic order “capitalism.” Ricardo and his followers certainly thought of the system as centering around the employment and control of labor by the capitalist. In theory, this is of course diametrically wrong. The entrepreneur employs and directs both labor and capital (the latter including land), and laborer and capitalist play the same passive role, over against the active one of the entrepreneur. It is true that entrepreneurship is not completely separable from the function of the capitalist, but neither is it completely separable from that of labor. The superficial observer is typically confused by the ambiguity of the concept of ownership. (Knight 1956, p. 68)

However, Wright’s view of the “capitalist” system and even his view of worker cooperatives (!) is totally imbued with the fundamental myth. For instance, he sees the private ownership of capital as including the:

rights to control the use and allocation of the surplus generated with the use of the means of production (i.e. the net income generated by the use of the means of production). (Wright 2010, p. 75)

The footnote on the same page includes:

The right to control the surplus generated through the use of means of production is very close to what economists refer to as “residual claimancy”—i.e. the right to all of the income generated in some economic process that remains after all expenses are paid. (2010, p. 75)

Or concerning the governance rights over workers, Wright also considers that as part of the rights of capital ownership:

At the core of the institution of private property is the power of owners to decide how their property is to be used. In the context of capitalist firms this is the basis for conferring authority on owners to direct the actions of their employees. (2010, p. 34)

But the basic point is that in the misnamed “capitalist” economy, capital goods are just as rentable as people. If any legal party rents the necessary capital goods (perhaps from different capital owners) and buys the other inputs, then that party will be the residual claimant, even though the owners of the capital goods are still just that. Hence the residual claimant rights are not part and parcel of the ownership of capital goods. The residual claimancy role is a contractual role, and one does not ‘own’ a contractual role in a market economy.

There is a common confusion among both critics and defenders of market economies about what is in fact owned and what is determined by the pattern of market contracts—by who rents what or whom. For instance, a conventional corporation has owners (the common stockholders) and a corporation may own capital goods, but it is
the pattern of market contracts that determines who is the ‘firm’ in the sense of the going-concern business using that capital. And market contractual-patterns are not ‘owned’. This holds even when the owner of the capital is a corporation and the capital good is a whole factory.

For instance, after the merger of the Studebaker and Packard companies in the early 1950’s, the Studebaker-Packard Corporation for a few years produced Packard car bodies in a plant leased from the Chrysler Corporation (Wikipedia). During that lease arrangement, Chrysler “owned the factory”, but it was neither the residual claimant nor the manager of the employees producing Packard car bodies in its factory due to the market contract leasing the factory to the Studebaker-Packard Corporation. The common phrase “ownership of the firm” is fundamentally confused. The Chrysler shareholders “owned the firm” in the sense of the owning the Chrysler corporation, and the corporation “owned the factory,” but Chrysler did not “own” the going-concern business using those “means of production.” These simple points of conceptual clarity should not be hard for both critics and supporters of markets to understand.¹

This misunderstanding of the ownership of capital was no small mistake in Marx. It determined the whole focus of his theoretical efforts (e.g., the name and topic of his main work) and it lead to the disastrous call for the abolition of the “private ownership system” and to the general use of “private” as a swear word on much of the Left. The real distinctive feature of the current private enterprise system is the legal contract for the hiring, employing, or renting of human beings.

Referring to the employment relation as the ‘renting’ of persons is not common usage but it is technically correct. In America, common usage is to “rent a car” and “hire a person” but in the UK, rental cars are called “hire cars.” Regardless of the language, it is the buying of the services provided by an entity as opposed to buying the entity itself. As the first American winner of the Economics Nobel Prize, Paul Samuelson, put it:

> Since slavery was abolished, human earning power is forbidden by law to be capitalized. A man is not even free to sell himself: he must rent himself at a wage. (Samuelson 1976, p. 52 (his italics))

Or as other neoclassical economists have put it:

Strictly speaking, the hourly wage is the rental payment that firms pay to hire an hour of labour. There is no asset price for the durable physical asset called a ‘worker’ because modern societies do not allow slavery, the institution by which firms actually own workers. (Begg, Fischer, & Dornbusch 1997, p. 201)

Marxism has both in theory and in practice nationalized rather than abolished the human rental system so that private employees become public employees instead of becoming the owner/members in firms such as worker cooperatives where the people who work in a firm are jointly working for themselves renting or owning the necessary capital. This brings us to Wright’s treatment of worker cooperatives.

### ON THE LEGAL STRUCTURE OF A WORKER COOPERATIVE

Wright is so imbued with the fundamental myth that he even uses it to try to characterize worker cooperatives. The workers in a worker cooperative or self-managed firm (an older variant) are said to have the residual and governance rights because they are the owners of the means of production. For instance, in the *Jacobin* article, Wright writes:

> In a worker-owned cooperative, all of the assets of the firms are jointly owned by the employees themselves, who also govern the firm in a one-person-one-vote, democratic manner. (12.2.2015)

In an essay written before (1979) workers’ self-management had collapsed in Yugoslavia and when the younger Wright was more committed to using Marxist buzz-words to signal his membership in the “ideal speech community” (Gouldner) of Marxists, he put the same idea using that jargon.

Workers’ self-management constitutes production in which the workers in a particular enterprise own the means of production and control the production process, and thus exercise rights over the disposition of the surplus produced by themselves within that enterprise. The mechanism of appropriation of the surplus labor can thus be designated “private-collective self-appropriation.” The direct producers appropriate their own surplus labor, and they do so through a collective process of management and
control over the production process. But this collective process remains essentially private in that the means of production are fully alienable and thus the surplus is appropriated by the workers in individual enterprises rather than by the working class as a whole. The mechanism of appropriation thus contains within itself both capitalist and communist elements: it is communist in that it is collective self-appropriation rather than exploitation; it is capitalist in that it is private rather than social. (1994, p. 142)

In the book under review, he expresses essentially the same idea without all the Marxist jargon.

A stand-alone fully worker-owned cooperative firm in a capitalist economy is a form of social capitalism: the egalitarian principle of one-person one-vote of all members of the business means that the power relations within the firm are based on voluntary cooperation and persuasion, not the relative economic power of different people. Jointly they control through democratic means the economic power represented by the capital in the firm. (2010, p. 91)

Or:

At the other end of the spectrum are firms characterized by two principles: they are fully owned by their employees and they are democratically governed by their members on a one-person-one vote basis. Such firms are called worker cooperatives or producer cooperatives. (2010, p. 167)

As an aside, it might be noted that in the American usage, “producer cooperatives” refers not to worker cooperatives but to agricultural marketing and processing ‘cooperatives’ (e.g., Land O’Lakes, Ocean Spray, Sunkist, etc.) which are completely conventional from the viewpoint of the employees and whose members are mostly agri-business corporate ‘producers.’

Or again the same idea:

If they [the workers] were owners of the firm, for example in the form of a worker-owned co-op, then their individual interests would be much more strongly aligned with those of the firm in which they worked, and fewer resources would have to be devoted to the tasks of social control. Since in general workers would work harder with less monitoring when they own the means of production, the heavy social control apparatus of capitalist production is a source of inefficiency. (2010, p. 41)

It is not difficult to conceptually separate being the residual claimant in a firm such as a worker cooperative from being the owner of the means of production. For instance, a family farm (as a one family worker cooperative) or an agricultural worker cooperative may well lease the land they work on, e.g., from a neighbor who can no longer work the land or from a land trust.

Moreover, it is important, particularly for Marxists wielding capitalist notions of the “ownership of the firm,” to realize that there are other types of rights afoot than property rights. The other major type of right, sometimes called “personal” or even “human” rights, are assigned to fulfilling a certain functional role—such as the voting rights attached to residing in a democratic living community. It makes no sense to treat these rights as alienable in a market transaction since the buyer may not have the qualifying functional role, and if she or he did, then they would not need to buy the rights. The litmus test to differentiate property rights from personal rights is inheritability (or bequeathability). When you die, all your personal rights are extinguished but your property rights pass to your heirs.

The members of a worker cooperative do not get their membership rights (net income and self-governance rights) as part of the property rights such as “ownership of the means of production” or as “ownership of the firm.” Instead they are personal rights attached to the functional role of working in the firm—just as one’s municipal voting rights are attached to the functional role of residing in the municipality. This has been spelled out quite clearly in the literature on worker cooperatives for at least 40 years, e.g., the pamphlet (ICA, 1978) published by the oldest organization in the country devoted to worker co-ops, the Industrial Cooperative Association (now The ICA Group). The legal structure of a worker co-op has been spelled out repeatedly over the years, e.g., in the 1982 collection of essays called Workplace Democracy and Social Change (Ellerman 1982), or in the 1984 collection of called Worker Cooperatives in America (Ellerman 1984) as well as in academic journals, e.g., (Ellerman 1984) or (Ellerman & Pitegoff 1982-3), and in a host of other papers and books. It is not clear how anyone modestly familiar with the literature on worker co-ops has to fall back on the idea that the members of a worker
co-op have the residual and governance rights because they “own the means of production.”

If Marxists have trouble grasping the concept of having personal rights attached to the role of working in a worker cooperative, then it is best not to delve into the whole matter of the worker-members’ property rights in a worker co-op in the form of internal capital accounts (or “divisible reserves”) (Ellerman 2007) which were pioneered in the Mondragon cooperatives and were completely unmentioned in Wright’s whole discussion of Mondragon. That is just too much complication for those thinking within the framing of who “owns the means of production.”

**IS WRIGHT IN FAVOR OF WORKPLACE DEMOCRACY?**

In his very definition of democracy, Wright uses what has been called the affected interests principle.

The second normative principle underlying the diagnosis and critique in this book concerns individual freedom and democracy. These two ideas are linked here because they both concern the power of people to make choices about things which affect their lives. This is the core principle: people should control as much as possible those decisions which affect their lives.

“Freedom” is the power to make choices over one’s own life; “democracy” is the power to participate in the effective control of collective choices that affect one’s life as a member of the wider society. The democratic egalitarian principle of political justice is that all people should have equal access to the powers needed to make choices over their own lives and to participate in collective choices that affect them because of the society in which they live. (2010, p. 12)

This is the core reason why Wright is actually not in favor of workplace democracy at the firm level as in a worker cooperative or democratic self-governing firm. Say, a worker co-op produces computers or cell phones that may be used all over the country or all over the world. Then all those users are “affected” by the potential decision by the worker co-op to raise prices or to change the design, so they are all “stakeholders” who need a “say” in the decision.

The term “stakeholders” is a contrast with the term “shareowners”. Share owners are the set of people with private property rights in the means of production. Stakeholders are all those with a “stake” in the means of production because their lives are affected by how those means of production are used. The idea that social ownership of specific means of production should extend to all stakeholders is the principle most consistent with the normative ideals of radical democratic egalitarianism discussed in chapter 1. Recall that the democratic egalitarian principle of political justice is that all people should have equal access to the means necessary to participate in decisions which affect their lives as individuals and as members as communities. This corresponds to the expansive notion of social ownership in which all “stakeholders” have ownership rights. (2010, p. 77)

Aside from characterizing (supposed) democratic rights as “ownership rights”, the point is the usual socialist one that “democratic control of the means of production and distribution” must involve potentially everyone or at least all the “stakeholders” in “society.” Thus, there can be no workplace democracy at the firm level—although “society” may delegate certain minor decisions to the workers in a firm in “democratic egalitarian socialism”. The expression “extending democratic rights to the workplace” has the usual socialist twist to mean not democracy in the workplace but extending the reach of social-political democracy or “democratic social power” to include otherwise private workplaces.

What is the alternative definition of democracy? The usual non-socialist definition of democracy is self-government of, by, and for those who are governed, not those who might just be affected. This uses the rather basic distinction between positive-direct-decision-making rights (e.g., within an organization) and negative-indirect-decision-constraining rights (e.g., in the marketplace). The rights of those who are only affected are usually enforced by their negative-indirect-decision-constraining rights such as their veto power in the marketplace (don’t buy the product!) or by consumer protection legislation in the overall political democracy.

But socialists are driven by their goal to largely eliminate the market so they disregard the usual indirect control rights exercised in the marketplace and thus have to extend “social democratic power” to include all workplaces. The suppliers of capital and other things to a productive firm, the buyers of the products, and the local residents are all not
governed by the management of the firm. They are all only potentially affected by the decisions of the firm and thus should have effective indirect control rights through the market or otherwise to protect their affected interests. That is the real import of the affected interests principle.

It might be noted that Robert Dahl also conceptualized democracy in terms of the affected interest principle, and he also tended to buy the fundamental myth that control (and residual) rights are part of the “ownership of economic enterprises” (Dahl 1985, p. 62)—as noted in the review (Ellerman 1985) of his book A Preface to Economic Democracy. Wright quotes Dahl approvingly.

Robert Dahl has argued, in an important book on the meaning of democracy, that there is no logical reason why rights to private ownership confer rights to dictatorial power over employees. (2010, p. 34, fn. 15)

After noting that Dahl is against the fundamental myth as if it were a legal fact, Wright goes on to paraphrase a very different argument against the real legal source of the governance rights over the workers, namely the human rental or employment contract.

Just as we have abolished slavery even in cases where a person might want to voluntarily enter into a contract to be a slave on the grounds that people should not be allowed to permanently give up their rights to autonomy (or “self-ownership” as some philosophers call it), we could prohibit people from giving up their right to autonomy within the employment contract of capitalist firms. People could still invest in firms, but this would only give them rights to a stream of earnings from the investment, not any rights to control the activities of people within the firm. [reference to (Dahl 1985)]. (Wright 2010, p. 34, fn. 15)

An argument against a voluntary contract to give up one’s “right to autonomy within the employment contract” is normally called an inalienable rights argument. Prior to Marx, Hegel clearly spelled out the inalienable rights argument against allowing a person to “voluntarily enter into a contract to be a slave” in the Philosophy of Right (Hegel 1967, p. Sec. 67) and that argument clearly also applied to the human rental contract—but it all sailed right by Marx. In fact, Marx went on to ridicule such arguments against the wage-labor contract itself by characterizing the sphere of exchange as the “very Eden of the innate rights of man” (Marx 1977 (1867), Chap. VI) and Marxists have ever since only snickered at the very mention of inalienable rights. Hence Wright does not follow up on Dahl’s promising lead and instead, at least in his earlier and more intense Marxist phase, appealed to Marx’s labor theory of value and exploitation (Wright 1994, p. 128).

Fortunately, Dahl’s use of the affected interests principle is innocent since he does not draw the usual socialist conclusion that all non-governmental firms (of any size) should fall under the political-social democratic governance. He ignores the socialist arguments when it comes to his “Sketch of an Alternative”. Instead Dahl comes out in favor of straightforward workplace democracy in the form of “worker’s cooperatives or examples of self-management or industrial democracy; but I prefer the term self-governing enterprises.” (Dahl 1985, p. 91)

A LABOR THEORY OF ‘WHAT’?

In commenting on what is in a book like Envisioning Real Utopias, one might also note “some dogs that didn’t bark.” In addition to Marx’s misunderstanding the basic structure of legal rights in a “capitalist” private property market economy, Marx spent most of his massive theoretical efforts in developing the inchoate “labor theory” inherited from Locke, Smith, and particularly Ricardo into a labor theory of value and exploitation. This theory ended up being a major train wreck in addition to being inherently superficial even if it had been a plausible theory of value. That is, as a theory of value it could at best only be a proof that wage-labor was systematically underpaid. For instance, in Marx’s discussion of overtime work, he said that even if the labor during the normal work day was “paid for at its full value”, there would still be unpaid labor extracted in overtime. In his own words:

It will be seen later that the labour expended during the so-called normal day is paid below its value, so that the overtime is simply a capitalist trick to extort more surplus labour. In any case, this would remain true of overtime even if the labour-power expended during the normal working day were paid for at its full value. (Marx 1977 (1867), p. 357; Chap. X, sec. 3)

While Marx was no doubt personally against the institution of wage-labor per se as evidenced by his “moralistic invective,” his theory (even assuming it was valid) was only a theory trying to show that wage-labor is not “paid for at
its full value.” In this respect, Marx can hold hands with neoclassical economists who have their own equally superficial “bourgeois” wages-are-too-damn-low exploitation theory that under non-competitive conditions, wage-labor is also not “paid for at its full value”, i.e., is not paid at the value of its marginal productivity (Ellerman 2016). Since the Marxian labor theory of value and exploitation was a train wreck as a theory of value and was in any case superficial—like its bourgeois brother—one wonders why so many Marxists have one of two reactions:

1. cling to the labor theory of value and exploitation as one’s ‘badge of Red courage’ to signal membership in the dwindling band of Marxists, or
2. just drop it altogether without understanding that there was another labor theory that might be valid.

John Locke, Adam Smith, and David Ricardo bequeathed to the 1800s an inchoate set of ideas which might be called “the labor theory” based somehow on the special status of human labor among the other factors of production. Those who saw their social role as finding a ‘scientific’ apology for “the system” needed to find some alternative theory where human labor would play no distinguished role. That was only accomplished towards the end of that century with marginal productivity theory as part of the whole marginalist revolution. But the critics of “the system” tried to develop the labor theory as a criticism of the system. Marx made the crucial decision to develop “the labor theory” as a labor theory of value and exploitation—and to target the system as “the capitalist system.” But that was not the only way to develop that inchoate set of ideas nor the only way to identify the system.

The alternative path was to develop it as the labor or natural rights theory of property, e.g., (Schlatter 1951), and to identify “the system” as the human rental system. That was essentially the path taken by the small band of political economists sometimes called “Ricardian socialists” although they were neither. Some of the principals in the school were Thomas Hodgskin and Pierre-Joseph Proudhon.

There are many Marxists who have taken second option of wisely dropping the labor theory of value and exploitation, and thus who have to get their Marxist membership ticket punched on the basis of some other theory like historical materialism, crisis theory, class analysis, or just a postmodernist word-cloud of the right buzz-words. What I at first found puzzling was why they didn’t cross over to the labor theory of property? That theory delivers a critique of the institution of wage-labor per se and is independent of any theory of value or wage rates. In the vernacular, Marx ‘brought a knife to a gun fight’, i.e., he brought a value theory to a property theory fight, so, of course, he lost out even if it had been a plausible value theory.

But how is a Marxist to find that alternative labor theory? It is not in Marx as some true-blue Marxists have testified.

---

**Figure 1: The fork in “the labor theory” road**

---
None of this, by the way, implies that Marx intended the labor theory of value as a theory of property rights, a la Locke or even Proudhon. (Shaikh 1977, p. 121)

Marxists might discover the alternative “labor theory” by at least reading the titles of the other labor theory books such as:

• Proudhon’s main book, What is Property? (Proudhon 1970 (1840)),
• Hodgskin’s book, The Natural and Artificial Right of Property Contrasted (Hodgskin 1973 (1832)), or
• Anton Menger’s book about the core tenet of that school, The Right to the Whole Produce of Labour: The Origin and Development of the Theory of Labour’s Claim to the Whole Product of Industry (Menger 1970 (1899)).

That is, in the labor theory of ‘what?’, the alternative ‘what’ is property.

But Marxists cannot adopt the labor theory of property because it is the normative basis for private property (“getting the fruits of your labor”). Marxists are dedicated to the “abolition of private property” (for the fallacious reasons outlined above concerning the fundamental myth)—as well as to the “social appropriation” of the product of industry. In short, Marxists are inherently against the people working in each enterprise privately and jointly appropriating the positive and negative fruits of their labor. Hence Marxists are not going to drop the labor theory of value in favor of the labor theory of property.

An interesting case of a potential crossover Marxist is the late G. A. Cohen. Cohen was one of the founding and leading members of the Analytical Marxism group which includes Wright and seems to be his main reference group; see the Prologue in (Wright 1994). At some point, Cohen had the brainstorm that the core critique actually had nothing to do with value!

And it is this fairly obvious truth which, I contend, lies at the heart of the Marxist charge of exploitation. The real basis of that charge is not that workers produce value, but that they produce what has it. (Cohen 1981, p. 219)

One would think it might occur to Cohen at this point that if the “fairly obvious truth” is not about value, then the “labor theory” might really be the “labor theory” of something else. But Cohen was not able to find that other labor theory—which, in any case, would not deliver the sine qua non of Marxist socialism, the “social” appropriation of the product of industry. Hence Cohen goes on to argue that all the inputs are “socially produced” and thus appropriation must take place at the “social” level.

It is a shame that Cohen did not understand that the direct implication of his “fairly obvious truth” is that the inputs in one enterprise are what the workers produce in a supplier enterprise. For instance, the drill presses used to produce the product in one enterprise are the products produced not by “society” but by the people working in a drill press enterprise. And the appropriate notion of “product” that the people in an enterprise produce is the whole product (i.e., the production vector of output-assets and input-liabilities) which includes not only the output-assets but also the liabilities for the inputs they use up, e.g., the liabilities for using the services of the drill presses to be satisfied by buying or leasing the drill presses.

By the labor theory of property, the people who work in one enterprise should:

• jointly own the assets they produce as outputs, and
• jointly owe the liabilities they also create by using up the inputs.

Those liabilities would be satisfied by buying the necessary inputs from the supplier firms. But that would typically be a market transaction—which is another reason why this whole train of thought is not available to someone whose pre-analytical predilection is abolishing market relationships.

The non-development of the labor theory of property and the condemnation of the “private property system” is perhaps the greatest gift of Marxism to the professional apologists for the human rental system. That is why Marxism now functions as a capitalist tool. The human rental system gives the people (“employees”) working in each firm:

• zero percent of their negative product (the liabilities for using use the inputs), and
• zero percent of the positive product (the assets produced as outputs).

That is, it gives them zero percent of the whole product (positive + negative product). Instead of the Marxist (and conventional) picture of the workers getting part of the product with the employer appropriating the "surplus prod-
uct,” the legal facts are that the employer appropriates the whole product (all the produced outputs and all the liabilities for the inputs) and the workers none. As an economic sociologist, Wright should appreciate the accurate statement of the legal facts by another sociologist in the early 20th century.

The laborer at no time owns any part of what is passing through his hands or under his eye. Never can he say, “This product, when finished, will be mine, and my rewards will depend on how successfully I can dispose of it.” There is much theoretic discussion to the “right of labor to the whole product” and much querying as to how much of the product belongs to the laborer. These questions never bother the manufacturer or his employee. They both know that, in actual fact, all of the product belongs to the capitalist, and none to the laborer. The latter has sold his labor, and has a right to the stipulated payment therefore. His claims stop there. He has no more ground for assuming a part ownership in the product than has the man who sold the raw materials, or the land on which the factory stands. (Fairchild 1916, pp. 65-6)

Instead, the employees are treated as suppliers of the commodity labor, and thus as the party to whom the labor liability (the labor costs) is owed. That liability and the other input-liabilities are paid off by the employer, who is thus the residual claimant.

By being against “the private property system” and by eschewing the whole labor theory of property as the only legitimate basis for private property, Marxism has delivered a priceless gift to the apologists for the human rental system. The apologists can pose as the defenders of “the private property system”—when in fact the human rental system is directly based on denying rented people (“employees”) the appropriation of the positive and negative fruits of their labor.

As for the dogs that didn’t bark in the case at hand, Wright in this book and elsewhere in his writings studiously eschews any of the “bourgeois” language or concepts of the labor theory of property (e.g., getting the fruits of your labor or, in juridical terms, imputing the positive and negative legal responsibility for the results of production to those who are de facto responsible for producing those results) from the whole discussion and critique of the current system. Instead Wright espouses the collective or social appropriation by “the working class as a whole” or, in this book, by a vague mélange of “social associations.” In either case, from the history of real-existing Marxist socialism in the 20th century, one does not have to wonder what institution would end up doing the actual “social” appropriation.

In conclusion, we might quote the genuinely radical proposals of the democratic classical liberal and yet Conservative MP (who was dubbed “Minister for Thought”), Lord Eustace Percy, who remarked when considering the post-WWII reconstruction:

Here is the most urgent challenge to political invention ever offered to the jurist and the statesman. The human association which in fact produces and distributes wealth, the association of workmen, managers, technicians and directors, is not an association recognised by the law. The association which the law does recognise—the association of shareholders, creditors and directors—is incapable of production and is not expected by the law to perform these functions. We have to give law to the real association, and to withdraw meaningless privilege from the imaginary one. (Percy 1944, p. 38)

NOTES

1 It is often said the Marx did not elaborate on his vision of an ideal society. But when one condemns “the system” because it has X (e.g., X = private ownership of capital or X = market relations), then one has clearly stated that one’s ideal at least does not have X.

2 This is a statement about legal rights, not about “social power” which, everyone knows, is typically in the hands of the employer/capital-owner.

3 The attempt to clarify Marxian thinking about the structure of legal rights is usually met with the Foucauldian response: “How can you be so superficial to worry about legal rights? It is all just congealed power relations.”

4 And Dahl footnotes a non-socialist source: “In clarifying my ideas on this question I have profited greatly from a number of unpublished papers by David Ellerman, cited in the bibliography, … .” (Dahl 1985, p. 91, fn. 1)

5 Albert O. Hirschman’s verdict was that Marx’s “works exhibit a simple juxtaposition of scientific apparatus and moralistic invective, wholly unversöhn” [i.e., unresolved] quoted in: (Adelman 2013, p. 570).
6 For instance, when conventional economists want to “take on the opposition,” they typically search out the nearest Marxist to lecture on the deficiencies of the labor theory of value. A perfect example of this genre is Robert Solow’s scolding of Marxist Duncan Foley in (Solow 2006).

REFERENCES


FRANCESCO DI IORIO’S *Cognitive Autonomy and Methodological Individualism: The Interpretative Foundations of Social Life* by Francesco Di Iorio

GABRIELE CIAMPINI

University of Florence/Paris-Sorbonne University

Berlin and New York: Springer, 2015, 189 pages. $129.00

Francesco Di Iorio’s *Cognitive Autonomy and Methodological Individualism* is a revised version of his doctoral thesis and investigates Hayek’s contribution to methodological individualism and the *Verstehen* paradigm. The latter conceives of individual actions to be understood as being caused by an active process of interpretation on the part of the subject. Focusing on issues related to the philosophy of action, Di Iorio strives to offer an original reading of Hayek.

According to Di Iorio, methodological individualism has often been misunderstood, having been confused with an atomistic conception of society. While Di Iorio acknowledges that an atomistic variant of methodological individualism does exist (e.g. standard economic models), he argues that methodological individualism cannot be reduced to this atomistic variant because there is a long and authoritative philosophical and sociological individualistic tradition that is non-atomistic. This non-atomistic individualistic tradition is long-standing and includes Adam Smith, Bernard Mandeville, David Hume, Montesquieu, Alexis de Tocqueville, Karl Menger, Georg Simmel, Max Weber, Herbert Spencer, Alfred Schütz, Ludwig von Mises, Karl Popper, Robert C. Merton, Friedrich Hayek, Raymond Boudon, Michel Crozier and Jon Elster.

The book discusses two fundamental assumptions concerning (non-atomistic) methodological individualism. The first, the concept of “autonomy”, as applied to the social agent. “Autonomy” is understood as “self-determination”; meaning that the cause of the action is not to be found in the external environment, but in the individual. Accordingly, this assumption, challenges various holistic paradigms in that individuals cannot be “pigeonholed” in a given environment and considered to be devoid of freedom. The second assumption is that social phenomena cannot be understood only as the planned results of human action because the consequences of individual actions are often unintentional and unpredictable.

Di Iorio explains that methodological individualism, as understood by Hayek, posits that social phenomena cannot be regarded as predetermined; they must be explained in terms of a spontaneous order. The inherent complexity of social phenomena is one of the main reasons why Hayek criticizes social planning and political constructivism. The interesting thing about Di Iorio’s thesis is that it highlights the dialectic between human actions and their unintentional consequences, focusing on how Hayek conceives of action. According to Di Iorio a detailed study of Hayek’s theory of action has often been neglected in the scholarly literature:

The history of methodological individualism is the history of the attempt to eradicate the concept of hidden determination from the study of society, and make the human being the starting point of social analysis. Methodological individualism rests on the understanding that the social order is the unintentional product of many autonomous human actions. On this reading, the conflict between sociological holism and methodological individualism may be interpreted as a conflict between a theory of heteronomy and a theory of autonomy (p. 3).

Di Iorio’s originality lies in his linking Hayek’s theory of human autonomy and methodological individualism to cognitive science, specifically to enactivism, via Hayek’s theory of the sensory order that considers the human mind to be a self-organizing complex system. The book is a good example of interdisciplinary dialogue between cognitive science and sociology, usually treated independently.
Chapter 2 deepens Hayek’s critique of the holistic approach to the social sciences. Holism denies the intentional dimension of human action, maintaining that social order is predetermined by social laws—laws that control the individuals. Hayek takes holists to task for not understanding that the social world is largely the result of unintentional purposeful human actions. Di Iorio clarifies the cognitive presuppositions of action from the standpoint of non-atomistic methodological individualism. Di Iorio illustrates the agreement between Hayek and Gadamer regarding the interpretative nature of knowledge, stating that Hayek’s theory of the sensory order is consistent with the hermeneutical theory of knowledge: “Hayek and Gadamer also agreed that, since the human being is an interpreter, he/she is hermeneutically free” (p. 12), i.e., a self-determined being.

For Di Iorio, Hayek’s originality resides in that he was one of the first thinkers who sought to establish a link between a theory of individual autonomy analogous to hermeneutics and phenomenology with cognitive psychology. Di Iorio argues that the epistemological implications of Hayek’s *The Sensory Order* (1952) have not been sufficiently analysed by social philosophers. In his view, the connections between Hayek’s theory of mind and the *Verstehen* tradition are relevant from the standpoint of the individualism-holism debate precisely because Hayek’s theory of mind supports human autonomy as understood by methodological individualism: “One of the goals is to demonstrate that Hayek’s reflections on mind include a very original argument in favor of Verstehen, an argument that has been rather neglected within the philosophy of the social sciences” (p.1). Hayek conceives of the mind as a complex dynamic system—a system that can only be explained through an “explanation of the principle” (p. 40). The logic of this system determines the existence of consciousness from a neurophysiological point of view as well as the cognitive autonomy of the agent.

Hayek’s philosophy of the mind is inconsistent with the mind-computer analogy. Those who believe that the mind works like a computer assume that every cognitive function is due to a “decision-making center,” i.e., to the Central Processing Unit (CPU) which controls every mental activity through a predetermined “protocol.” Hayek’s vision is diametrically opposed to this:

the mind is made up of billions of components—neurons—whose activity is not pre-programmed but self-determined. The neurons do not follow specific instructions, but work in a sense in an independent manner. They build up the perceptive categorizations by connecting spontaneously to each other. They create complex chains of impulses that correspond to the different kinds of “patterns” humans are able to recognize (p. 40).

Taking Hayek’s idea of mind as a self-organizing-system, Di Iorio illustrates how Hayek uses the idea to criticize some cognitive theories, such as behaviorism and, especially, methodological holism. Despite their differences, these two approaches share the basic assumption that the cause of action must be sought outside the individual: “Behaviorists consider action to be determined mechanically by physical reality, understood as a pre-given reality, while sociological holists consider action to be determined mechanically by the socio-cultural environment, which they similarly regard as a pre-given reality” (p. 55).

Both approaches deny the interpretation that social agents give their surroundings is relevant and that the mind is a self-organizing system that acts as a “cause of itself”, meaning it is self-determined and cannot be perturbed by external factors (p. 44). This is where the connection between *Verstehen* sociology and Hayek’s cognitive psychology comes into play. According to Hayek, the social context does not determine the actions of individuals, but how they interpret the context is what causes their actions. Di Iorio further explains that for Hayek human interpretative autonomy is not absolute. Social factors play a role in the cognitive process and influence individual action. However, their influence is not mechanical because there is always a dialectic between these factors and the way in which the individual interprets them.

Chapter 3 investigates in more detail the relevance of Hayek’s *The Sensory Order* to the individualism/holism debate. Hayek argued that methodological holism and behaviorism are governed by the same mechanistic paradigm, that denies the interpretative activity of the individual. Di Iorio, interestingly, connects Hayek’s theory of mind to that of Maurice Merleau-Ponty who also argued that the mind is a self-organizing system and that action is always determined by the interpretation by the subject, and on that ground criticized sociological holism. Hayek and Merleau-Ponty (a phenomenologist strongly influenced by Edmund Husserl), are commonly regarded as being very dissimilar to each other. However, Hayek claimed that Merleau-Ponty developed a perspective very similar to his own. As Di Iorio stresses, both these thinkers shared a critique of the objectivist conception of knowledge and explained the mind in terms of self-organizing system and interpretative...
device: “Merleau-Ponty studied in greater depth the issue of the impossibility of explaining consciousness as an epiphenomenon of sociality” (p. 6), i.e. in holistic terms. And “Consequently, reading Hayek in the light of Merleau-Ponty is extremely useful for understanding why the idea that mind is a complex self-organizing system and an interpretative apparatus implies a criticism of sociological holism” (Ibid.).

Chapter 4 analyses “how the interpretative autonomy of the actor is related to the systemic structure of the social world” (p. 76). Following Hayek, this “chapter attempts to clarify the theory of social systems as intended by non-atomistic methodological individualism, and the way that this approach reconciles the indeterminism of action and existence of social conditioning” (Ibid). The reductionist interpretations of (non-atomistic) methodological individualism developed by various authors (e.g., Roy Bhaskar, Harold Kinkaid, Lars Udehn) in the last few decades are carefully analysed and criticized. This chapter reconstructs the intellectual roots of both individualism and holism. Following Hayek, Di Orio targets Hegel and Comte, two very different thinkers, but who shared an organicist conception of society: “In spite of the strong and undeniable differences between Positivism and Idealism, Comte and Hegel nevertheless share some common points. In particular, both defend a theory of action based on the idea of heteronomy and connected to a deterministic conception of historical development” (p. 79). Di Iorio states that:

Comte’s and Hegel’s organicist theory of society is related to their concern for the profound social changes of their epoch. Their organicism must be considered, among other things, as a conservative reaction to the French Revolution, its subversive ideas and the de-structive egoism of modern industrial society. Both Comte and Hegel, in spite of their philosophical differences, developed a collectivist and historicist theory to provide a reassuring view and drive out the fear of chaos (p. 87).

On these grounds, society is conceived as something that transcends the individual; it is ‘emancipated’ from its individual components. It is seen as an independent phenomenon that controls individuals from outside and makes harmony and social order possible. Comte and Hegel have influenced, albeit in different ways, several social theories that came later, from Durkheim to Bronislaw Malinowski, Alfred Radcliffe-Brown and Talcott Parsons. These authors have conceived society in holistic terms and regarded action as determined by holistic macro-laws that control the individual and make the social order possible. Even structuralism is affected by this holistic approach: the idea that the societies of each historical period are characterized by the presence of ‘structures’ that determine a particular cultural and economic order has at its base the conception originally developed by Hegel and Comte.

In Chapter 5, Di Iorio stresses that methodological individualism has often been confused with a utilitarian theory of action and criticized because of its commitment to utilitarianism. According to Di Iorio, the confusion between methodological individualism and utilitarianism must be avoided because utilitarianism is only one variant of methodological individualism. Hayek and other methodological individualists such as Mises and Boudon do not agree with utilitarianism and support a broader conception of rationality, claiming that all human actions, including those carried out without a utilitarian purpose, not based on an instrumental rationality, must be considered rational. Following Hayek and the Verstehen tradition as understood by hermeneutics and interpretative sociology, Di Iorio defends a broad concept of rationality. He considers rationality to be a general feature of human action and assumes that even religious beliefs, ethical choices, and attitudes that cannot be explained in utilitarian terms and that are commonly regarded as irrational, are actually rational because they always presuppose “an intelligent process of interpretation and meaning-construction” (p. 125). Action must never be seen as an uncritical adherence to existing cultural models because it always presupposes an active role on the part of the social agent. As argued by Boudon and Gadamer, the agent’s rationality is not necessarily utilitarian and Cartesian. It can also be argumentative, fallible and characterized by a certain vagueness. This second kind of rationality is termed by Boudon as “cognitive” or “ordinary” rationality.

The final chapter of the book deals with the relationship between interpretation and explanation. It attempts to answer the following question: “Is the interpretative approach of methodological individualism compatible with the methodology of the natural sciences?” (p. 8). Following Hayek, Popper, Mises and Carl Hempel, Di Iorio argues “that the interpretative approach is not incompatible with the use of covering laws, and more generally with the methodology of the natural sciences” (Ibid.). While Hayek and Mises are sometimes ambiguous on this issue, some of their writings “imply a belief in a common methodology between the so-
social and the natural sciences” (Ibid.). Hayek’s epistemology is compared to Mises’ and Popper’s ideas. Following Barry Smith, and Hayek himself, Di Iorio criticizes “the widespread thesis that the epistemological views of these two authors are radically incompatible” (p. 8) and stresses “that Mises’ apriorism and Popper’s fallibilism are reconcilable” (p. 9). On these grounds, Di Iorio argues “that Hayek’s defense of fallibilism must not be interpreted as a radical critique of Mises” (p. 9).

Di Iorio’s book is an original and valuable contribution to the philosophy of social science that breaks new ground in our understanding of Hayek’s thought and of methodological individualism more generally. One of the great merits of the books is that it demonstrates why many biases on the account of which methodological individualism is usually criticized must be rejected as being both logically and historically unfounded.
AIMS AND SCOPE
COSMOS + TAXIS takes its name and inspiration from the Greek terms that F. A. Hayek famously invoked to connote the distinction between spontaneous orders and consciously planned orders.

COSMOS + TAXIS publishes papers on complexity broadly conceived in a manner that is accessible to a general multidisciplinary audience with particular emphasis on political economy and philosophy.

COSMOS + TAXIS publishes a wide range of content: refereed articles, unreferred though moderated discussion articles, literature surveys and reviews.

COSMOS + TAXIS invites submissions on a wide range of topics concerned with the dilemma of upholding ethical norms while also being mindful of unintended consequences.

COSMOS + TAXIS is ecumenical in approaches to, and not committed to, any particular school of thought and is certainly not a talking shop for ideologues of any stripe.

COSMOS + TAXIS is not committed to any particular school of thought but has as its central interest any discussion that falls within the classical liberal tradition.

Books for review should be sent to:
Laurent Dobuzinskis
Department of Political Science
Simon Fraser University
AQ6069—8888 University Drive
Burnaby, B.C.
Canada V5A 1S6

http://cosmosandtaxis.org

SUBMISSIONS
Submitting an article to COSMOS + TAXIS implies that it is not under consideration (and has not been accepted) for publication elsewhere. COSMOS + TAXIS will endeavor to complete the refereeing process in a timely manner (i.e. a publication decision will be made available within three months). All submissions should be in digital format, and emailed to: david.andersson@rmit.edu.vn

Papers should be double-spaced, in 12 point font, Times New Roman. Accepted papers are usually about 6,000-8,000 words long. However, we are willing to consider manuscripts as long as 12,000 words (and even more under very special circumstances). All self-identifying marks should be removed from the article itself to facilitate blind review. In addition to the article itself, an abstract should be submitted as a separate file (also devoid of author-identifying information). Submissions should be made in Word doc format.

COSMOS + TAXIS welcomes proposals for guest edited themed issues and suggestions for book reviews. Please contact the Editor-in-Chief to make a proposal: david.andersson@rmit.edu.vn

All business issues and typsetting are done under the auspices of The University of British Columbia. Inquiries should be addressed to the Managing Editor: leslie.marsh@ubc.ca

ELEMENTS OF STYLE
1. Submissions should be in English, on consecutively numbered pages. American, Canadian and UK spellings and punctuation are acceptable as long as they adhere consistently to one or the other pattern.

The reference style is as follows:
Author, J. E. and Author, B. (Eds.) Title. City: Publisher, pp. 1-10.

3. All notes should be as end notes.

4. No mathematical formulae in main text (but acceptable in notes or as an appendix).

Please consult the latest issue of COSMOS + TAXIS to see a fully detailed example of the Journal’s elements of style.