

Paola Palminiello

## A Reply to Comments

(doi: 10.2383/85815)

Sociologica (ISSN 1971-8853)

Fascicolo 3, settembre-dicembre 2016

**Ente di afferenza:**

()

Copyright © by Società editrice il Mulino, Bologna. Tutti i diritti sono riservati.  
Per altre informazioni si veda <https://www.rivisteweb.it>

### Licenza d'uso

Questo articolo è reso disponibile con licenza CC BY NC ND. Per altre informazioni si veda <https://www.rivisteweb.it/>

# A Reply to Comments

*by* Paola Palminiello

doi: 10.2383/85815

Thank you Daniel Little, Antonino Palumbo and Thomas Prosser for your comments. Although coming from different and even opposing fronts, none of them are particularly kind with my essay. Indeed some observations betray a hint of satisfaction for delivering what was thought to be a fatal blow. However, it might be not over yet. Here are my replies.

## **1. The Rational-choice Explanations of Institutional Change**

Daniel Little [2016] didn't find any of the theses I have proposed persuasive, starting from my critical analysis of the not only functional explanations of institutional change put forward by the new economic institutionalism and a number of political economy papers. Indeed, according to Little, these explanations would be fully satisfactory, exactly as, after all, those proposed by the many political and social scientists who support the rational-choice approach, some of whom Little, sometimes extensively, quotes. Perhaps I could be right – he concedes – with respect to the explanations suggested by researchers concerned with the emergence of the institutions of European cooperation, and Giavazzi and Pagano [1988] in particular, but the “broader tradition” would not make the sort of mistakes I have underlined in the paper.

I would like to point out, however, that I have not only illustrated my point by proposing other cases in addition to the ones concerning European institutions, some of which I believe are particularly clear [Palminiello 2016, 4-5], but also that Giavazzi and Pagano have made use of a mechanism (pre-commitment, self-binding or tying one's hands) which has been extensively tapped into, beginning with the Nobel prizes Kydland and Prescott [1977], and to deal with economic and social phenomena that are quite different from the introduction of European institutions. It was invoked by Shepsle himself [1989], whom Little quotes to support his position, and Coleman's explanation of cooperation norms (that again Little presents as a successful rational-choice explanation) is structured in the exact same way (and falls therefore in the same kind of contradiction). Again, it is built on the idea that an agent with the incentive to do X (defecting in a two-person Prisoner Dilemma: doing X is rational for him) can also have the incentive to prevent himself from doing X (the incentive to introduce a norm that could ban doing X) [Coleman 1990, Chapter 10]. And I would like to mention that Elster [2007, 54] too has recognized that economists run the risk of falling in a "sort of rational-choice functionalism."

Conversely, the passages and works that Little cites in order to prove the adequacy of rational-choice explanations [Little 2016, 2-4] either consist in *petitio principii* or offer theses that are too imprecise and vague to be able to fulfil the task they have been assigned by Little. A very clear example may be found in the programmatic statements used by Coleman [1990] to introduce his explanation of cooperation norms (the first of the two long passages that Little quotes), considering that the mechanisms suggested by the author only a few pages later (not considered by Little) blatantly contradict them. As the second long quotation provided by Little shows, Shepsle upholds a different approach from the one proposed by Coleman and new institutionalist economists, which engenders a different kind of problems (I allude to some of them below): according to Shepsle [2006], institutions are equilibrium ways of doing things and not constraints on agents' actions, and therefore "reflect the willingness of (nearly) everyone to engage with one another according to particular patterns and procedures." But this definition says nothing about the way in which those equilibria can be found. And the thesis that "the institutional arrangements are [...] focal [...] and may induce coordination around them" [Little 2016, 2] focuses on the positive consequences that those arrangements have for the system,<sup>1</sup> rather than on their cause. The same holds true for the "explanation" Little

<sup>1</sup> Institutions have the function to be focal when strategic interactions have many different equilibria and therefore individuals' rationality cannot help them find the game solution. This is one of the problems I have just alluded to.

attributes to Crozier and Friedberg [1980], according to which institutions would be “solutions to collective action problems [...] created, invented and established [by] relatively autonomous actors [...] with their particular resources and capacities” [Little 2016, 2-3], and that Crozier and Friedberg would have constructed without invoking neither social efficiency nor “other positive contribution to the larger system” [*Ibidem*]. But these words do not constitute an explanation. And, after all, what is the ability to solve collective action problems if not a positive contribution to the system? Maybe my adequacy criteria for the reconstruction of individuals’ actions and interactions are stricter than Little’s, but I believe that they need to be, as it is only their plausibility and consistency that we can rely on in checking the robustness of our explanations. Or does Little think we can rely on something else? And, if that’s the case, what would that be?

## 2. Explanandum External Definitions, Strong and Weak Individualism, Holism

Anyway, the main thesis of the first part of my essay is not that there are no adequate or satisfactory rational-choice explanations of institutional change. As a methodological individualist, I rather try to demonstrate, by analysing in detail a selection of works, that there is no way to construct adequate (coherent, plausible, etc.) individualistic explanations of social phenomena (institutions, for instance) that are defined from an external point of view or from above, in terms of social efficiency, above all, but clearly also in terms of something else, i.e. of other kinds of external or objective consequences, of any deep social “meaning” or “essence” [Elster 1983; 2007], or even of any supposedly objective criterion of equality or inequality. I basically try to demonstrate that these “objective” consequences or these supposed “objective” social entities<sup>2</sup> cannot have adequate (coherent, plausible, etc.) micro-foundations; hence, they have to be considered devoid of any plausibility and can only be explained using circular or tautological arguments. After all, isn’t it obvious that externally defined results cannot be explained in terms of the actions and interactions of individuals? But Little never tackles these points, he does not make any comment on my purely methodological defence of strong (or rigorously subjective) individualism against the weak (or institutional or structural) one, and he doesn’t even pay attention to the nature of my attempt to explain institutions (so much that he ends up attributing an external definition of institutions even to me: I would conceive institutions as “solutions to problems of coordination” [Little 2016, 3]). Quite the

<sup>2</sup> Or even emergences.

opposite: when he agrees to deal with the issue of institutions' supposed "functions", he does it from the substantive rather than from the methodological point of view, denying the efficiency effects of institutions but only to highlight those of inequality and power [Little, 4].

I think that this can also clarify, at least indirectly, why I did not examine, as suggested instead by Thomas Prosser in his comment [2016], approaches like the ones of Streeck and Thelen [2005] or Schmidt [2008], who also criticize the rational-choice explanations of institutional change and are among the major voices in the debate on the evolution of institutions. As I mentioned, my critique to rational-choice explanations is not only basically methodological and not substantial, as are those provided by Streeck and Thelen and Schmidt, but, as I affirmed in the previous paragraph and reiterated in my essay [Palminiello 2016, 1-3], it is also internal to methodological individualism or based on individualistic premises. My point is that those explanations, breaking with a crucial principle of individualism, forget to look at the phenomena to be explained from the point of view of the individuals who produced them and replace it with a characterization of those phenomena in terms of the external and supra-individual aim of social efficiency. Therefore, they end up revealing themselves, together with structural and other forms of weak individualism, very similar to holistic approaches and, exactly like them, destined to propose only circular or tautological explanations. In my essay, I also argue that, given the similarity between rational-choice explanations of new economic institutionalism, structural or institutional individualism and holism (which I believe is especially attributable to transcendentalism, i.e. researchers' claim to be able to locate themselves in a very privileged position outside of or above the social world), we should replace the contradistinction between individualism and holism with the one between strong or substantial individualism and holistic or individualist objectivism [Palminiello 2016, 2].

This is why I did not examine approaches like the ones of Streeck and Thelen and Schmidt: from the methodological point of view, I consider these explanations very far from ideal. Some years ago, I analysed in detail the construction of some well-known explanations of the welfare state reforms ratified in Italy during the last decades in an unpublished paper.<sup>3</sup> The critique of holistic explanations that I propose in my essay [Palminiello 2016, 9-10] is, to a certain extent, an outcome of this previous work. And I think that the same arguments, except for the third, can be applied to Streeck and Thelen's essay too. It seems to me that they also aim to know the

<sup>3</sup> In particular, the explanations proposed by, among others, Maurizio Ferrera and Elisabetta Gualmini, and Massimo Paci.

social world by looking at it from a point of view that is external to society or in a supposed “objective” way and therefore leave out the perspective of the individuals (their beliefs and intentions) involved in its production. But this approach proves again to be not only the cause of the same contradictions I mention in the essay [*Ibidem*, 9], but also misleading.

Streeck and Thelen’s essay presents a theory of gradual, incremental and endogenous institutional change – as the ongoing process of liberalization would be – as opposed to sharp, abrupt and exogenous change. The problem is that the proposed theory, instead of providing a strong and reliable support to this notion of change<sup>4</sup> and to the related interpretation of the liberalization process, seems constructed in order to accommodate it and made up of elements that are only apparently plausible. I am thinking in particular of the idea of the gaps between institutional rules and their implementation as a way out of the assumption of the constraining character of institutions [Streeck and Thelen 2005, 11-16]. Or let’s take as an example the so-called mechanism of displacement: how can it be, to address just one problem, that “a growing number of actors (end up defecting) to a new system previously deviant, aberrant, anachronistic” [*Ibidem*, 20], given that the very idea that someone can choose an institution instead of another by himself/herself is paradoxical? Or let’s look at the subsequent mechanism, called layering or differential growth: how can the “cunningly orchestrated” introduction of a private funded pension scheme as a minor pillar of the existing public system clarify also “its subsequent explosive growth” [*Ibidem*, 23] which is actually the fact to be explained? It clearly cannot.

That said, Prosser is certainly right: it can be absolutely useful to carefully examine also the non rational-choice theories of institutional change, at least those considered among the most important and influential ones.

### **3. Theory of Explanation, Rational-choice Theory and the Two-level Bargaining Game**

So, in the first part of my essay I deal with a problem of theory of explanation: I try to demonstrate that it is not possible to explain phenomena defined in terms of the external, super-individual (and considered very positive) goals of social efficiency in a consistent individualistic way, and I do that by illustrating the mistakes that the authors of such explanations make just to see their starting, mainly normative, hypotheses confirmed.

<sup>4</sup> Which is, in addition, defined in a very imprecise way: when can a change be considered “deep”? When is it “gradual”?

In his comment to my essay, Antonino Palumbo [2016] seems however to make this problem coincide with a different issue: can rational-choice theory adequately and consistently explain cooperation? Proven this theory unsatisfactory, could it be supplemented or changed? In fact, Gauthier's notion of constrained maximisation [1986] cited by Palumbo answers these questions, not mine. Unlike Giavazzi and Pagano [1988], Gauthier does not contradict himself: his criticism and non-compliance with rational-choice theory is explicit and intentional. Obviously, there is some sort of link between caring mainly about social efficiency and conceiving rationality in the way of rational-choice theory,<sup>5</sup> but the problem I deal with and the one of Gauthier's (and many others) are nonetheless distinct: for me, it is the explanandum definition that doesn't work, whereas for Gauthier it is a component of the explanans or some of its characteristics that don't work. Or I could put it this way: explaining efficiency results consistently would still be impossible even if rational-choice theory succeeded in accounting for collective action.

In my view, there are three types of rational-choice explanations of institutions seen or defined in terms of social efficiency. In my essay, I dwell mainly on two of them.

The first is the functional one or more precisely the one afflicted by rational-choice functionalism, in the essay illustrated by Giavazzi and Pagano's explanation. Here social efficiency is the goal of the social system. These kinds of explanations are characterized by a mechanism that I called "self-correction" [Palminiello 2016, 7]. After ascribing to the institution to be explained the effect of fulfilling a need of the socio-economic system or solving an inefficiency problem by changing individuals' incentives, they ascribe to the actions of the agent who has supposedly introduced such institution the purpose of correcting, through said institution, his previous rational but negative behaviour that would have caused that alleged malfunction or the inefficiency of the system. The problem is that desiring to establish an institution with the purpose of changing one's own behaviour through that institution is already desiring to adopt the behaviour that the institution is intended to induce, that means having a motivation that the agent cannot have if we want his actions to be the cause of the problem that the institution would have to solve [Elster 2007, 209-210]. It's either one or the other.

In the second type of explanations, efficiency results are the unintended consequences of actions that agents carried out with other intentions, and more precisely

<sup>5</sup> It is clear, for example, that, as any choice has to reveal what the agent prefers or attributes the greatest utility to, traditional game theory cannot but require that the agent chooses the immediately best result at every game move, even though this implies that he treats his counterpart's strategies as events or nature instead of the choices of a fully rational player [Sugden 1991].

consequences that are strongly appreciated by researchers from outside the world but ignored by the agents themselves precisely because they can be viewed only from a perspective that is extraneous to them, and in respect to which individuals are no more than means. Among the many explanations of this type are, for instance, those that conceive institutions as equilibria, as in Shepsle [2006], and those of contract theory. Here there are two possibilities: either the identification of the positive unintended consequences produced by individuals' actions does not affect the explanation of these actions and thus the efficiency or inefficiency judgment can be considered as simply not pertinent in an explanatory context, or it does affect the explanation and "sacrifices" it to its own needs. Agents' actions are explained in terms of the "objective" goals that prove to be compatible with those consequences, not in terms of the desires and beliefs that could be possibly attributed to agents after careful analysis. One of the reasons why in my essay I did not dwell upon this type of explanations is that they often are merely conjectural or speculative and, unless one conducts the empirical research that their authors avoided, it is essentially impossible to make any further comment about them.

The third type of explanations is the one I illustrated in the essay through Dyson and Featherston's [1996], Dixit's [1996], and Vreeland's [1999] essays. According to Antonino Palumbo, however, my critique is not persuasive. I will therefore try to articulate it in a different way.

In my view, the main mistake that Vreeland [1999] makes is that he ascribes conflicting interests to the various agents he considers (the small government élite, the domestic opposition, and the International Monetary Fund)<sup>6</sup> but no common interest<sup>7</sup> (the government depends on both the opposition and the IMF, but not vice versa).<sup>8</sup> Indeed, he explicitly denies that they have common interests (neither the government élite nor the domestic opposition would consider the IMF loan important for the country or their own business, a point that Palumbo appreciates), perhaps because otherwise the government élite could not freely choose the adjustment policies she intends to dictate. The problem is that, where there are no common interests or no one depends on anyone, there is nothing on which people can bargain (one can threaten someone else only if he/she controls something his/her counterpart is

<sup>6</sup> To summarize the basic premise of the game, the government élite would bring in the IMF in order to prevail over the domestic opposition and impose its adjustment policies. Owing to its minority position, it could not otherwise have these policies approved by parliament or congress, nor take any other kind of domestic action.

<sup>7</sup> In the meaning in which Schelling [1960] uses this notion: see pp. 11-12, number 5.

<sup>8</sup> Actually, as the government doesn't need any loan (see below) and simply wants that the IMF imposes the adjustment policies on the country, its dependence on the Fund is in fact superficial.



interested in),<sup>9</sup> and Vreeland is therefore forced to somehow bring them back: hence the idea that rejecting the IMF agreement can frighten the riotous domestic opposition too as it sends a negative signal to creditors and investors. But this revealing artifice cannot be effective if the IMF loan is unnecessary, not at least if the involved parties are supposed to be rational (as they are): in fact, why should the domestic opposition be afraid of creditors and investors when all it needs to do is to return the loan funds? The government can credibly threaten the domestic opposition only if the IMF loan is considered essential by both the latter, which would then agree to approve the budget policies for fear of losing it, and first of all by the government itself, because, otherwise, we could not possibly consider it to be forced to put into effect its unpopular policies (not being dependent on the IMF in any way). The same can be said about the two-level game that Palumbo applies to the EMS: the French government can use the agreement with Germany against parliament, trade unions and entrepreneurs only if parliament, trade unions and entrepreneurs consider France's participation in the EMS important; otherwise they could free themselves from the domestic policies imposed by the government by simply forcing it to exit the EMS (according to the game's assumption, it is the government that depends on them, not the other way around).<sup>10</sup>

I recognize that conspiracy explanations too can be structured in this way (here the *élite* would pursue *élitist* rather than altruistic goals).<sup>11</sup> My point is anyway that this must be certainly the structure of those explanations that tell us how an *élite* that is as small as it is benevolent (given the general selfishness hypothesis, utilitarians must be rare)<sup>12</sup> managed to put in place institutions or policies that favoured social

<sup>9</sup> "Your money or your life!"

<sup>10</sup> And this is not the only problem. According to Palumbo [2016], given the characteristics of the agreement with Germany in 1978, the centre-right French government could have chosen between two different kinds of policies: those favourable to trade unions and those favourable to the entrepreneur community. But if that was the case, then Germany would not have imposed anything on France, and one begins to wonder why the French government could not do the same without entering the EMS altogether. Considering that Germany does not impose anything on France, the answer cannot be that in that case the French government could not have shifted the responsibility for implemented policies on the international agreement, unless we inconsistently suppose the irrationality of the domestic oppositions. The agreement between France and Germany too is somehow paradoxical: as the first would have willingly given up its plan in order to find itself forced to implement the domestic policies it wanted, it would have gained something without giving something in return.

<sup>11</sup> And probably this is the version Palumbo has in mind. However, as I better state in a section below, this kind of *élite* too needs its counterparts to share common interests with it if it wants to impose something on them.

<sup>12</sup> Speaking of utilitarianism, let me reply to Palumbo's first comment. I agree with him: Hume was a supporter of rule utilitarianism. But the opposition between act and rule utilitarianism does not correspond to the one between direct and indirect utilitarianism. (This last kind of utilitarianism (or "Government House utilitarianism", as it is often called) is the focus of the above paragraph).

welfare, economic growth or national wealth in spite of the opposition by the rest of the politicians and by citizens in general, all of which interested only in their own welfare. In fact, all individuals who do not belong to the *élite* are or would be nothing but means with respect to the pursuit of these external and super-individual goals, and if they are rational they will not agree to ratify the policies required to pursue them. No one will do it, not even individuals who end up being accidentally favoured by these policies, given that they could gain more with other policies (the external and super-individual goals reduce everyone to means). If the *élite* were a little larger and therefore capable of taking some action against the self-interested politicians and citizens, then these would depend on it for the fulfilment of some of their interests, and the *élite* could threaten them and achieve (almost) what it wants (through bargaining), but unfortunately this is not the case. The authors of these explanations do not acknowledge that they are weak because they are the first supporters and defenders of the adjustment policies we find in their work and of the goals of social efficiency, growth and so on, and, as a result of wishful thinking, tend to see them fulfilled wherever they can. It is possible that there are not many explanations of this type. However, we can find the same incorrect ideas in the substantial body of normative literature produced by economists who strive to find a way to curb policy-makers' opportunism and count on the only remaining benevolent agents, i.e. economists. And a few researchers realized that [see for example O'Flaherty and Bhagwati 1996]: if society had the necessary institutions, the opportunism of policy-makers and citizens in general would not be an obstacle to social efficiency maximization, but how to introduce these institutions if all policy-makers are opportunist and the control mechanism of elections does not work? Just as any other benevolent *élite*, economists do not have much leeway.

#### 4. Why Retrieving Schelling's Approach

Palumbo believes that the main topic of Schelling's *The Strategy of Conflict* [1960] is distributive issues as opposed to efficiency ones; thus, since in my essay I would "blame rational-choice explanations of institutional change for using efficientist arguments inconsistently with their own methodological tenets," the reader "would have thought that the main reason for retrieving Schelling's approach [...] was indeed that" [Palumbo 2016, 3]. I would like to note that the opposition "distributive vs efficientist questions" cannot be superimposed on the one between "external or 'objective' point of view" and "individuals' point of view." A person who is bargaining over the purchase of a house wants to pay as little as possible and not

that the distribution of her and the seller's gains from trade is one way or another. From a subjective point of view, cooperating can mean avoiding capsizing the boat and drowning into the lake; from the efficiency perspective, it doesn't lead to a sub-optimal outcome but leads instead to an optimal one.

Therefore, no, it is not for that reason that I retrieved Schelling's approach. As I declare in my essay [Palminiello 2016, 12], I went back to it because, going against traditional game theory, it credits individuals with the ability to cooperate, and more precisely to "change the game." As I do not wish to add anything to what I wrote in the essay, I will not further elaborate on this point. I will just underline that Schelling considered strategic moves – threats and promises we would make even though they were not part of the game – precisely as proofs of this particular ability, as the focal points by which we would succeed in coordinating even in mixed-interest interactions would in fact be. Indeed, by using these kinds of moves or by meeting our partner's mind at a focal point, we would solve or stop the "familiar spiral of reciprocal expectations" [Schelling 1960, 87], allow an agreement with our partner and make the satisfaction of (part of) our conflicting interests possible thanks to the achievement of the common ones. So, I do not think that Schelling regards strategic moves primarily as means by which each bargainer tries to gain the most for himself from bargaining, as Palumbo maintains; or, better, they can be considered as such only because they are first of all the means through which an agreement may be reached.

[W]inning in a conflict does not have a strictly competitive meaning; it is not winning relative to one's adversary. It means gaining relative to one's own value system; and this may be done by bargaining, by mutual accommodation and by avoidance of mutually damaging behavior [Schelling 1960, 4-5].

Although it might not be necessary at this point, I will make an additional consideration, as an appendix to these clarifications. In Schelling's works, the meaning of self-commitment is completely different from the one we find in the works on strategic time inconsistency (of Kydland and Prescott, Giavazzi and Pagano, and many others, including Elster): it is not used to prevent oneself from choosing the immediately best option and therefore to avoid renegeing on promises and threats that one would not have, strictly speaking, the incentive to keep, but is simply used to credibly communicate one's own true intentions. Therefore, despite what Palumbo thinks [2016, 3-4], I do not run into a contradiction in the second part of my essay in supposing that agents can bind themselves to make credible promises and threats, I am simply following Schelling. I acknowledge however that the statement of the assumptions at the base of my bargaining model will be complete only when I will have clarified what goes wrong in rational-choice and traditional game theory, thus

providing the reason for setting them aside. Schelling's reasons – the overabundant formalism and abstract nature of traditional game theory, its tendency to deal with mixed-interests interactions as they were zero-sum – do not, I think, tell the whole story. It is Gauthier's problem, which however comes out in the second, and not in the first, part of my essay.

## 5. The Institutions from the Point of View of Individuals, Conflicts of Interests and Power

If I understand correctly, many of the objections Little makes to the second part of my essay (as severe as those raised to the first part) originate from the idea that, when the parties have conflicting interests, power is distributed unequally or there are authoritative “super-agents” [Little 2016, 4], institutions cannot benefit everybody, there cannot be an unanimous consent, and therefore the explanation of their introduction cannot lie in a bargaining agreement.<sup>13</sup> Hence I would have proposed a bargaining model capable of explaining only institutions of pure cooperation or devised by agents with identical interests.

It is very clear, however, that bargaining would be actually unnecessary in those cases. And in fact, as Palumbo acknowledges (but does not appreciate), the institutions I try to explain are possible solutions of mixed-interest interactions rather than of pure cooperation games. According to my model, bargaining involves parties aware of their mutual dependence for achieving some of their respective objectives,<sup>14</sup> and concerns the definition of a cooperative scheme between these parties that could ensure that each of them achieves those personal objectives (stopping the erosion of one's own cultivable plots;<sup>15</sup> enjoying the health benefits of a mutual aid society; not importing yellow fever with tea from Far East colonies vs preventing the suspension of maritime trade with the West) thanks to the achievement of the common ones (planting new trees to stop the land erosion; contributing to the common fund in employment periods; cooperating in the inspections of ships, etc.). The basic idea is that, when an agent acknowledges that for the satisfaction of some of his goals, he is dependent on someone else because of some kind of social causality, and everybody

<sup>13</sup> Jack Knight, who Little cites in support of his ideas, argues something similar: “[I]f strong actors can constrain others to choose a particular equilibrium strategy the weak will comply whether or not they want to do so. From this it follows that social actors respect these institutional rules not because they have agreed to them [...] but simply because they cannot do better than to do so” [1992, 127].

<sup>14</sup> According to these objectives either one party damages the other or could guarantee the latter an advantage.

<sup>15</sup> Elster's example [1983]: erosion on one plot can be stopped if and only if trees are planted on that plot and on both the adjoining ones.

is rational, devising cooperation schemes is the best way he has to pursue his own interests.

Nor do I suppose that cooperation must invariably benefit everyone equally. In presenting my first type of bargaining [Palminiello 2016, 14], I imagine that the contracting parties could be heterogeneous (or conceive themselves as such) rather than homogeneous, and therefore pursue different goals (i.e. advantages) by means of their cooperation. Little's pirate captain and his crew (or an entrepreneur and his workers) could, for example, devise a scheme of this kind: in fact, despite his (supposed) power, the pirate captain too depends on rank-and-file pirates, for the success of boardings and for his very survival, and this means that he is himself exposed to threats and has to grant the crew some, or maybe many, advantages (which will be directly proportional to the degree of the sailors' awareness of his dependence on them). The parties sign the agreement if they both prefer the cooperation it establishes to non-cooperation, not if they receive equal benefits.

With the third type of bargaining [Palminiello 2016, 16], I also suppose that each contracting party can try to pursue its conflicting interests also by devising a new cooperation scheme (a different way to deal with their dependence) that is more beneficial for itself than the other party's (this new scheme would allow the proponent to achieve goals it values more than the goals it would achieve by signing the scheme proposed by his/her counterpart). In that case, it may happen that a party surrenders and signs the agreement with its counterpart even if the scheme agreed upon – compared to the scheme in effect or the one proposed by such party – will worsen its situation, if in any case that scheme turns out to be the only one feasible and it appears to that party to be better than not cooperating at all. In other words, the agreement will not be Pareto-optimal in respect to those alternative schemes (the scheme in effect or the one proposed by such party). The cooperation scheme between employers and workers that has come into being during the last decades – we can define it a scheme of individualization of labour's risks – benefits employers more than the previous social democratic one, but is clearly preferred to non-cooperation by workers too. Therefore it is not true that my bargaining model leaves out interest conflicts and ignores power. Instead, I suppose that power does not consist, or at least does not consist only, in owning more resources than the others, but instead in the ability, not at all ordinary, to devise a cooperation scheme that is more advantageous for oneself and to make one's promises and threats credible in Schelling's sense.<sup>16</sup>

<sup>16</sup> Who indeed criticizes the objectivist interpretations of bargaining power, for example by noting that a country not able to control its balance of payments, to collect taxes or appeal for political unity to defend itself is more likely to get some help than a country able to control all

Yet there is a point about which Little is instead certainly right, that is that there is no way to explain “that some parties [end up being] worse-off within the proposed institution than outside [of it]” except for referring to an “artful application of various forms of power on the part of one [...] participant against the rest,” made possible by “significant inequalities of resources” [Little 2016, 4]. But this is so only because here explanandum and explanans are nothing more than products of the imagination of the researcher. Certainly they are not compatible with an approach to explanation in terms of individuals’ actions and interactions. Given the intentionality and rationality of all the other agents, no one, no matter how powerful, can cause a collective outcome by himself, for instance an outcome that is, as in this case, good for himself and very bad for everybody else. An agent who sees his situation getting worse as a result of an other agent’s action will react, if he is rational, and the relationship between the two will turn from parametric into strategic: the “strong” person will not be able to achieve the outcome he wants without coordinating with the other, in a tacit or in an explicit way (that is the same as bargaining). Of course, they will not necessarily reach an agreement. If a party’s proposal threatens to leave the other party with less than the latter could obtain with its outside options, they will not agree. But a person who accepts absolutely unfavourable conditions, for example under the threat of war, nevertheless accepts them, chooses his best option, and in an individualist approach this is the explanation we are looking for.<sup>17</sup> Therefore, it is untrue “that some parties may [end up being] worse-off within the proposed institution than outside of [it]” [Little 2016, 4]: this is merely a researcher’s moral judgment of no importance in an explanatory context. In this context, according to individualism, the only justice judgements that are worthy of attention are those of the agents. So, I reply this to Little: that, as Elster [2007] underlined, before one tries to explain something, one has to verify if it exists or existed (or did take place), and if it is defined from an external point of view it will not be there.<sup>18</sup> Or I could say

its resources [Schelling 1960, chapter 2]. Threats and promises that are based on objective elements are in essence based on incentives and therefore actually redundant [*Ibidem*, 123, number 5].

<sup>17</sup> This is an objection that can of course be made also to Knight’s thesis quoted in footnote 13. If the weak cannot do better than to comply with the institutional rules wanted by the strong, then complying is their best option. In the context of individualist explanations, the best option does not mean that it is the ideal one or what the agent - strong or weak, it does not matter - would absolutely prefer or would have preferred, but just means that it represents the best option within his opportunity set. These are the options - and not the ideal ones - one needs to explain actions. By asking oneself about agents’ degree of positive freedom, one deals with an important moral question that however is totally irrelevant from the explanatory point of view.

<sup>18</sup> With this, I do not intend to exclude brute force in general. Anyway, I would like to point out that probably resorting to brute force is rational only when dealing with the weakest counterparts, and

that, while economic explanations end up substituting an interpretation of agents' actions from the perspective of the external result of social efficiency to the agents' subjective point of view, the explanations aimed at denouncing power imbalances – the explanations of Little, Knight, Moe [2005] and many others – elect the occurrence of an event that is supposedly very negative for a social group (the weak) and positive for at least one of the others (the strong) as their explanandum, and explain it with the “strong” group's determination to pursue its interests against the “weak” one, but again to the detriment of the point of view of the agents.<sup>19</sup>

To conclude this section, I would like to add that I do not dismiss Little's, Knight's, or Moe's worries. The issue is however how to respond to them. And in this regard I believe that my analysis of the role that social efficiency has in rational-choice explanations of institutions shows very well what it is that transforms interests that are considered in principle conflicting into pieces or elements of an organic and ordered system that is organized to achieve a unique super-individual good. Because it avoids “summarizing” events and facts from a single perspective above the world and gives instead room to individuals and their ways of identifying and pursuing their interests, the rigorously subjectivist approach I propose avoids this mistake.

Two more words on the rational-choice explanations of institutional change proposed by Knight [1992], and presented by Little as both methodologically correct and capable of making room for inequalities of resource and power. I don't agree with Little's assessment. Here are a few problems besides the one already pointed out in footnote 13.

a) Knight's idea is that institutions are the intentional or unintentional result of distributive conflicts won by the strong against the weak, and that their function is to reproduce and safeguard the unfair distribution produced by those conflicts by solving an agents' strategic problem: to inform the defeated about the future actions of the winners and therefore about their own (of the defeated) equilibrium behaviour.<sup>20</sup> Knight defends this idea by setting it against the thesis of the majority of rational-choice theories of institutional change “that social institutions exist because they benefit us” or produce “collective benefits” [1992, 26]. However, what is striking

precisely with those lacking even the awareness of their own right to self-defend. With counterparts that, albeit weak, claim some rights and are able somehow to resist, the rational thing to do is probably bargaining: if the threat is effective, one can get (almost) everything one wants without resorting to violence or dealing with the violence of the counterpart (though perhaps very limited). In the first case, the strong party acts in a parametric environment, while, in the second case, in a strategic one.

<sup>19</sup> In my essay [Palminiello 2016], I make a brief reference to these kinds of explanations in footnote 43.

<sup>20</sup> See for example Knight [1992, 17, 23, 38-42, 126].

about his critical analysis is its extreme confusion. His analysis (just like Little's) betrays a total misunderstanding of mutual dependence, a misunderstanding which, as we will see below, is at the root also of the idea already mentioned in footnote 13 that "the weak [...] respect [...] institutional rules not because they have agreed to them [...] but simply because they cannot do better than to do so" [*Ibidem*, 127].

b) This is Knight's criticism: rational-choice theories would succeed in "explaining" that institutions produce collective benefits only by forcibly making self-interest coincide with collective interest. This would be proven by their inability to "provide micro-foundations for their explanations in terms of individual rational action" [*Ibidem*, 39 and more generally 27-39]. This criticism is apparently very similar to mine, but it really isn't, and certainly just not because it is not the adopted explanandum that worries Knight. It is not always clear what is Knight specifically imputing to the explanations he wants to refute: when he supposes that the explanations he is examining ascribe benevolent motivations to agents, and when he recognizes that, according to many rational-choice explanations, collective benefits can be (from the point of view of observers) the unintended consequence of the self-interested choices of the involved individuals.<sup>21</sup> Anyway, Knight refutes these explanations too. Here is how. Suppose – the example is Knight's [*Ibidem*, 33-34] – that X and Y are choosing an institutional rule, and that point A on the Pareto frontier made up of the points that maximize the collective welfare of X and Y represents a possible rule. But imagine also that X prefers point C which is inside the frontier but more generous to him than point A (C is SE of A), and that Y prefers instead point B, which is too inside the frontier and more favourable for Y than A is (i.e. B is NW of A). Against rational-choice explanations of institutional change, Knight argues that, given either point B or C (that is, if both X and Y would be either in B or in C), the favoured agent will never agree, if motivated by self-interest only, to move to A. The issue is that rational-choice theorists do not see things differently: contrary to what Knight thinks, A is not Pareto-optimal in respect to B and C (given either B or C, only the points on the frontier that are NE either of B or of C represent an improvement for both X and Y). But, as I just stated, Knight does not realize this and thus absurdly concludes that a rational agent will opt for the less socially efficient rules if these promise him a bigger share of profits (will he try to move from A to B?) [*Ibidem*, 33]. Therefore, according to Knight, an agent can produce a collective result by himself.

c) Knight's explanation of institutions as distributive mechanisms is based exactly on this same idea: these institutions would reproduce the equilibrium that is

<sup>21</sup> They are the explanations of the second type mentioned above in section 3.



most favourable for a group and unfavourable for the others that the first group succeeded in imposing to these last groups against their will. The references to bargaining [Chapter 5] must not be misleading: they do not make reference to the process through which an agreement is reached and thus a favourable, albeit perhaps unbalanced, exchange is made possible. Knight identifies bargaining mainly with the art of compelling other people to undertake certain actions by pre-committing oneself to a future behaviour, that is by making credible threats and promises [*Ibidem*, 41-42, 129-131]. (The possible ways in which we could modify other people's feasible set by pre-committing ourselves include many notions that are not at all compatible with the project of explaining social phenomena in terms of the desires and beliefs of individuals, but I will not elaborate further on this).<sup>22</sup>

d) According to Knight, only some actors can affect the alternatives available to others to get them to act in the ways they want or can enjoy the degree of credibility that is necessary to do this, and they are the strongest or richest [*Ibidem*, 131-133]. Yet, this idea – which is clearly crucial for a theory of institutions as instruments of stabilization of the unequal distribution of resources and power – raises at least two problems. 1) As Schelling makes very clear, “threats” that rely on objective facts (power and wealth, for instance) instead of pre-commitment to future actions are credible simply because they are based on incentives, that is they are not actually threats but rather warnings that will make threats unnecessary [Schelling 1960, 123-124, number 5]. Certainly, as they communicate nothing else but the incentives of the strong and the rich, they cannot change the weak's feasible set more than they have already done. Hence, that idea – that the winners in bargaining are the rich and the powerful, an idea which is in fact characteristic of traditional bargaining theory – is not compatible with the one stated at point c), which summarizes Schelling's theory.<sup>23</sup> 2) If it is the groups favoured by the asymmetries of power who win in bargaining interactions, how did the first asymmetry develop?

e) So, institutions would stabilize distributive asymmetries imposed by means of asymmetries of power. But asymmetries in comparison to what? Clearly not to a primordial condition of equality from which distributive conflicts would have later distanced us, but rather to an idea which is again clearly moral and belonging to the researcher, the idea (which I certainly support, of course – but that is not the point)

<sup>22</sup> Knight supposes, for instance, that the stronger agent A can exercise his power on B “by adding [to B's feasible set] alternatives that are in accord with B's subjective interests but are contrary to B's real interests,” or even by “manipulating B's preferences” [1992, 42].

<sup>23</sup> Knight himself sees this problem and tries to challenge the relevance of one of the most important implications of Schelling's way to conceive bargaining, that is that “precommitment allows the weak to take the strategic advantage away from the strong” [1992, 131, number 5].

of the equal moral value of all human beings. But if that is the case, what is the topic of “*Institutions and Social Conflict*” then?

## 6. Still about my Bargaining Model: Three Types of Bargaining or More?

Palumbo lingers over my bargaining model at length with questions, conjectures, critiques, emendations and content jabs. These are my replies.

*i)* As I declare in my essay, the distinction between “devising and adopting a cooperative scheme” and “choosing between alternative cooperative schemes” [Palminiello 2016, 14] has to be read in opposition to traditional bargaining theory and in particular to the idea that the bargaining situation is given (by virtue of what, in fact?) and bargainers negotiate over the division of the gains of cooperation. Drawing on Schelling, I argue – as I already stated above – that bargaining means first of all succeeding in dealing with a mutual dependence by devising a cooperation scheme that could be accepted by one’s counterpart. The idea that, given a certain situation of mutual dependence, two or more cooperation schemes could be devised and that in that case bargainers would have to choose the one to be adopted, has two premises: 1) the idea that, with the exception of exchange situations, bargainers try to foster their respective interests by devising a scheme that is more advantageous for themselves, instead of bargaining on the distribution of the gains deriving from cooperation, and 2) the thesis that, being each scheme a tool for pursuing many kinds of different goals, bargainers can disagree about the exact characteristics of the cooperation to be organized to have those interests satisfied.

*ii)* Palumbo himself admits that my first type of bargaining corresponds with alternative *c* that he believes should be added to my typology to make it analytically complete. But if it was already there, what’s the point of adding it? Anyway, unfortunately my first type of bargaining does not correspond with what Palumbo identifies as “a genuine bargaining game in which redistributive questions are all that matters.” Actually, that my first type of bargaining could correspond to this is explicitly excluded. I will repeat what I have very briefly affirmed in the previous point: my cooperation schemes are ways of reorganizing mutual dependences, whose aim is to allow all involved parties to achieve at least some of their goals over time. So, as I wrote in the essay:

[E]ach bargainer’s share of benefits is settled by the requirements of the scheme “functioning” or effectiveness with respect to its aims, and nobody has anything to gain by modifying these requisites [Palminiello 2016, 13].

If today I decrease my workers' wage under the minimal level required for the correct functioning of our scheme and the achievement of the goals that we assigned to it, tomorrow they will be undernourished and weak and, in the end, I will be the loser.

*iii)* Differently from what Palumbo writes (see his type *b*), in my third type of bargaining – parties choose between alternative cooperation schemes which pursue different goals – not only the means imagined by A for his own scheme are different from the means devised by B for his scheme, being these schemes ways to achieve different goals, but A and B disagree also about the appropriateness of the means for each other's scheme: the proposer A claims that the means proposed by B to achieve the goals of B's own scheme cannot work, and vice versa, even though the arguments they offer in this debate are not necessarily sincere [Palminiello 2016, 19]. Therefore, this third type of bargaining corresponds to the one Palumbo labels as *d* and whose introduction we would owe to him. As I do not see what could Palumbo's alternative *b* possibly add to it (my third type, named *d* by Palumbo), I consider my typology complete.

*iv)* All bargaining with threats or commitments are *Chicken Games*, not just those of my third type: pre-committing is worthwhile for A if and only if B does not do the same, and vice versa. However, as Palumbo himself points out, this is so only if the game is symmetrical, and this is unlikely to occur in real-life bargaining. Schelling himself makes this point and in fact blames traditional game theory for “treating perfect symmetry between players as the general case rather than a special one” [Schelling 1960, 119] and emphasizes the inherently empirical nature of a theory of strategic interactions hinged on players' ability to coordinate their expectations. References to the specific features of a certain interaction (agents' value systems, what they know about the world, the available commitment devices, the structure of communication, and so on) have to be considered *ad hoc* if used in building models that are meant to be deductive and purely formal, but are essential for the explanation of concrete institutions.

## **7. Repeated Cooperation Schemes and Sincere Arguing**

The thesis that sincerely describing the characteristics of the scheme is in the interest of its proposer is considered by Palumbo “not that relevant.” I hope that point *ii)* in section 6 above can help understand at least one of its premises.

Little [2016] too criticizes this thesis, but interpreting it in his own way: I would have argued that the proposer of a scheme has to communicate sincerely to his/her

counterpart what he/she will get from cooperation in order to avoid that he/she reneges on the agreement soon after (as soon as he/she becomes aware of the trick). Little denies that this veracity is necessary in the first rounds:

a pirate captain [could] have an interest in over-stating the amount of booty likely to be captured [and] the rank-and-file pirates [would] have been persuaded to contribute their labor on false pretenses [Little 2016, 5].

However he finally acknowledges that “perhaps honesty is the best policy in the long run,” but underlying again that “the typical pirate captain makes plans over short and medium term” – as if this example was mine and institutions meant to be short-lived. Yet not only the idea concerning the first rounds can hold only if the recipient of the proposal is irrational (if cheating is in the interest of the proposer, the recipient ought to be aware of this, if he is rational, and distrust him). It is also a reply to a thesis that is different from mine.

## 8. Explaining the EMS

In his comment, Palumbo declares that once I finally got to the point – explaining the EMS –, I would have fallen into the same inconsistencies I had imputed to Giavazzi and Pagano, Dyson and Featherstone, Vreeland and so on.

I have already made clear why I consider the two-level game used by Putnam and Vreeland unpersuasive and Palumbo’s explanation of the EMS unsatisfactory. So, no, it is not true that I used that kind of game in my explanation. According to the hypothesis I propose, the internal negotiation launched (and won) by entrepreneurs in each of the high-inflation countries (of the third type) is independent from the 1978 international negotiation, even though it was prompted by two consequences of the EMS: the ongoing process of European integration and the prospect of an intensification of international trade. I also wish to underline that, according to my reconstruction, the negotiations are three: the first two of the third type [Palminiello 2016, 22-23] and the last one, signed in 1987, of the first type [*Ibidem*, 25].

Naturally, my explanation can be wrong. As Prosser rightly notes, in order to propose something more than mere conjectures, one needs more data than those I gathered. Indeed, well aware of the inadequacy of my information, I was the first to set limited goals for this part of my essay [Palminiello 2016, 21], and explaining why it was Germany that won the 1978 bargaining with France is not among them.

## References

- Coleman, J.S.  
1990 *Foundation of Social Theory*. Cambridge: Cambridge University Press.
- Crozier, M. and Friedberg, E.  
1980 *Actors and Systems: the Politics of Collective Action*. Chicago: University of Chicago Press.
- Dixit, A.K.  
1996 *The Making of Economic Policy: A Transaction-Cost Politics Perspective*. Cambridge, Mass.: The Mit Press.
- Dyson K. and Featherstone K.  
1996 "Italy and the Emu as a 'Vincolo Esterno'. Empowering the Technocrats, Transforming the State." *South European Society and Politics* 1: 272-299.
- Elster, J.  
1983 *Sour Grapes*. Cambridge: Cambridge University Press.  
2007 *Explaining Social Behavior*. Cambridge: Cambridge University Press.
- Gauthier, D.  
1986 *Morals By Agreement*. Oxford: Clarendon Press.
- Giavazzi, F. and Pagano, M.  
1988 "The Advantage of Tying One's Hands: Ems Discipline and Central Bank Credibility." *European Economic Review* 32(5): 1055-1075.
- Knight, J.  
1992 *Institutions and Social Conflict*. Cambridge: Cambridge University Press.
- Kydland, F.E. and Prescott, E.C.  
1977 "Rules versus Discretions: The Inconsistency of Optimal Plans." *Journal of Political Economy* 85(3): 473-492.
- Little, D.  
2016 "Comment on Paola Palminiello/1. Calibrating the Utility of Rational Choice Institutionalism." *Sociologica* 3: [doi:10.2383/85812](https://doi.org/10.2383/85812)
- Moe, T.M.  
2005 "Power and Political Institutions." *Perspectives on Politics* 3(2): 215-233.
- O'Flaherty, B. and Bhagwati, L.J.  
1996 "Will Free Trade with Political Science Put Normative Economists Out of Work?" *Economics and Politics* 9: 207-219.
- Palminiello, P.  
2016 "An Attempt to Provide a Non-circular Explanation of Economic Reforms. Elements for a New Individualistic Theory of Institutional Change." *Sociologica* 3: [doi:10.2383/85811](https://doi.org/10.2383/85811)
- Palumbo, A.  
2016 "Comment on Paola Palminiello/2. Playing Europe. Critical Remarks on Palminiello's Non-circular Explanation of the European Monetary System." *Sociologica* 3: [doi:10.2383/85813](https://doi.org/10.2383/85813)

Prosser, T.

2016 "Comment on Paola Palminiello/3" *Sociologica* 3: [doi:10.2383/85814](https://doi.org/10.2383/85814)

Schelling, T.C.

1960 *The Strategy of Conflicts*. Cambridge, Mass.: Cambridge University Press.

Schmidt, V.A.

2008 *Bringing Ideas and Discourse Back into the Explanation of Change in Varieties of Capitalism and Welfare States*." Working Paper n. 2, University of Sussex: CGPE.

Shepsle, K.A.

1989 "Studying Institutions: Some Lessons from the Rational Choice Approach." *Journal of Theoretical Politics* 1: 131-147.

2006 "Rational Choice Institutionalism." Pp. 23-38 in S. Binder, R.A.W. Rhodes, and B. Rockman (eds.), *Oxford Handbook of Political Institutions*. Oxford: Oxford University Press.

Streeck, W. and Thelen, K.

2005 "Introduction: Institutional Change in Advanced Political Economies." Pp. in W. Streeck and K. Thelen (eds.), *Beyond Continuity: Explorations in the Dynamics of Advanced Political Economies*. Oxford: Oxford University Press.

Sugden, R.

1991 "Rational Choice: A Survey of Contributions from Economics and Philosophy." *The Economic Journal* 101: 751-785.

Vreeland, J.R.

1999 *The IMF: Lender of Last Resort or Scapegoat*. Paper presented at the Midwest Political Science Association Annual Meeting, Chicago, IL, April 15th-17th. Available online at: <http://faculty.georgetown.edu/jrv24/imf.pdf>

## A Reply to Comments

---

Abstract: The essay replies to the comments by Daniel Little, Antonino Palumbo and Thomas Prosser on the author's essay "An Attempt to Provide a Non-circular Explanation of Economic Reforms. Elements for a New Individualistic Theory of Institutional Change", lingering over rational-choice explanations of institutional change and the thesis that institutions are structures of power that can make some groups' situation worse than it was without institutions.

---

*Keywords: Rational-choice Explanations; Two-level Bargaining Game; Bargaining; Thomas Schelling; Exercise of Power.*

---

Graduate in Philosophy, **Paola Palminiello** was a temporary research assistant at the Rimes Department at the University La Sapienza of Rome, and got a PhD in Economics and Sociology. She currently collaborates with various research centres and works as a translator for several Italian publishing companies. She is now working on rational-choice theory.