Internal Organization in a Public Theory of the Firm: Toward a Coase-Oates Federalism Nexus

Garzarelli, Giampaolo

University Rome, La Sapienza, IPEG, Johannesburg

9 May 2018

Online at https://mpra.ub.uni-muenchen.de/86955/
MPRA Paper No. 86955, posted 31 May 2018 21:51 UTC
Internal Organization in a Public Theory of the Firm:
Toward a Coase-Oates Federalism Nexus

Giampaolo Garzarelli
DiSSE, Sapienza – Università di Roma,
Rome, Italy & IPEG, SEBS, University of
the Witwatersrand, Johannesburg,
South Africa
giampaolo.garzarelli@gmail.com

Invited contribution to International Tax Law Review
http://www.rdti.it/en/node/5592

ACKNOWLEDGEMENTS
This paper is a parsing and comparing of many economic ideas as it is envisioned as
an explorative work for a book project developing a New Institutional Economics
approach to fiscal federalism. Special thanks to Pier Paolo Patrucco for useful
feedback at a preliminary stage and to Aldo A. Sitoe for his input beyond the call of
duty. The comments of Emma Galli, Lorenzo Infantino, Yasmina Rim Limam, Pietro
Selicato and of the two referees were very useful. Last but not least, the many
conversations with the late Domenico da Empoli over many years on this and
related federalism topics were extremely precious. I am the sole module responsible
for all inadequacies and mistakes.

THIS DRAFT 9 MAY, 2018.
Internal Organization in a Public Theory of the Firm: Toward a Coase-Oates Federalism Nexus

ABSTRACT
This paper anatomizes how the theory of internal organization of the firm relates to that of internal organization of government. This broad issue is approached by narrowing matters down to a specific type of internal organization of government: fiscal federalism. The paper introduces elements for a public theory of the firm by theoretically combining organizational and federalist insights – Ronald Coase with Wallace Oates. It shows how there are vertical and horizontal transaction cost problems in both the ex ante moment of decentralized public sector organizational design and the ex post moment of organizational adaptation. These problems embed normative and positive considerations that previous organizational theories of federalism fail to consider, and that earlier theories of federalism to some extent acknowledge but fail to develop organizationally. A subsidiary point that emerges is that more effort should be directed to exploring the ex ante moment in explicit organizational design terms. To try to jump start the explorative effort, the paper also alludes to one promising set of design principles: modularity.

JEL CODES
H11, H44, H76, L95.

KEYWORDS
Coase-Oates nexus; Comparative institutional analysis; Ex ante and ex post fiscal federalism; Intergovernmental transaction costs; Modular near-decomposition.
The debate about how to distribute fiscal responsibilities within a polity – whether locally or centrally – remains topical. In Italy, for example, the so-called Monti technical government (November 16, 2011-April 28, 2013) basically demolished any hope of Lombard irredentism of the previous executive (the IV Berlusconi Cabinet, May 8, 2008-November 6, 2011) by founding a Ministry for Territorial Cohesion. Concurrently, there are supranational forces imposing national fiscal responsibilities: the various sovereign debt crises of the Euro zone led to the “Treaty on Stability, Coordination and Governance in the Economic and Monetary Union,” also known as the Fiscal Compact, which from January 1, 2013 requires a national legal rule whereby budgets of signatory states must always be balanced or in surplus. And, of course, there is the 2016 Brexit and the 2017 Catalan independence attempt, and recurring concerns about Scottish independence and Quitaly.

At the core of the debate, as I see it, is a two-sided issue that relates to distinct organizational moments of a fiscal federation, namely the ex ante moment of organizational design and the ex post moment of organizational adaptation. Though the two moments are distinct, their common thread is classic: “that every government ought to contain in itself the means of its own preservation.”¹ That is to say that, like any other organization, government faces costly trade-offs having to do with both its raison d’être and survival.

However, there is little coherent theoretical apparatus to draw upon in order to carefully think about the two-sided issue organizationally. This paper tries to remedy the situation by outlining the connection between the basic reasoning principles of fiscal federalism, especially in the work of Wallace E. Oates, and those found in the work of Ronald H. Coase. It therefore teases out some first elements for a coherent organizational theory of fiscal federalism – for a public theory of the firm – mostly by distilling the insights of the two doyens of the fields under consideration.

One of the main contributions that Coase is known for is “The Nature of the Firm” (1937). Few would deny that this contribution single-handedly spawned the

field of the theory of the firm, or, more broadly, the modern economics of organization (Gibbons and Roberts 2013). The field gravitates around three main questions:

(1) Why do firms exist?
(2) What determines firm boundaries?
(3) How are firms internally organized?

The main focus of the proposed public theory of the firm is on a variant of the third question: how does the theory of internal organization of the firm relate to that of internal organization of government? The focus on fiscal federalism narrows this broad question. The reason for the focus is straightforward: the study of the internal organization of government – of the horizontal and vertical relationships among governments – in terms of the theory of the firm is a recent, important development in federalism theory (Oates 2008; Weingast 2014).

Government, just as a firm, relies on decentralization, the nature and extent of which, we shall see, are not transaction-cost free. More specifically, viewing the internal organization of a fiscal federation through a public theory of the firm that hinges on a Coase-Oates nexus leads to normative and positive considerations about the ex ante and ex post political economy of policy rights allocation in a fiscal federation that previous organizational approaches fail to consider, and that earlier theories of fiscal federalism to some extent acknowledge but fail to develop organizationally. That is to say that ex ante and ex post organizational moments ultimately reflect fixed and variable transaction-cost compromises that must be taken into account when considering the political economy of horizontal and vertical fiscal intergovernmental relations.

A subsidiary point that emerges is that the ex ante moment is less explored than the ex post. However, both moments result to be closely connected if one continues to reason organizationally, viz. in terms of both transaction costs and property rights. The consequence is that more effort should be dedicated to directly examining the ex ante moment in terms of explicit principles of organizational design. To try to jump start the effort, I direct attention to one set of candidate principles known as modularity.
As far as I could determine, the study of federalism in terms that are analogous to those presented in the pages that follow has not been proposed before. It falls squarely within the tradition of the New Institutional Economics (NIE), which studies institutions mainly in terms of their coordinative and incentive aligning role (Schotter 1981; E. Ostrom 2007; Bednar 2009). An important aspect of the NIE consists in considering the economizing properties that different institutions have when dealing with socioeconomic problems, such as the internalization of positive and negative externalities (Coase 1960). This aspect can be more generally captured by asking at what net cost an institution creates order relative to another (Foss and Garzarelli 2007).

Analytically, it translates into the method of comparative-institutional analysis (Williamson 1991; Diermeier and Krehbiel 2003), whose ineluctable principle is that the institutions to be compared are always feasible and not ideal ones (Demsetz 1969; Hamlin and Stemplowska 2012). For example, one does not compare the incentive properties of a feasible private health insurance scheme with the incentive properties of textbook's perfectly competitive market. Rather, one compares the alternative incentive properties of different feasible types of private health insurance schemes. The method concerns cost-benefit calculation in real- (or, if you prefer, second-best) world scenarios where there are transaction (and other) fixed and variable costs that purposive economic action tries to surmount, or, at least, decrease (Williamson 1985), and where the observed institutions better manage these costs than their (non-observed) alternatives (Gibbons 2005).

We therefore might say that comparative institutional analysis injects a significant dose of realism into economics, and social science more generally, by concentrating on the cost-benefit problem-solving attributes that feasible, alternative institutions possess (Mäki 1998). This is a salutary dose that fiscal federalism also has been receiving (Hamlin 1991), especially with its recent emphasis on the modern economics of organization (Oates 2005). This paper in an

---

2 Breton and Scott (1978) is close in spirit. However, its point of entry is fiscal federalism mostly for design issues, whereas mine is the economics of organization for a federation's design as well as adaptation issues.

3 Throughout, institutions are assumed to be “systems of established and embedded social rules that structure social interactions,” and organizations to be “special institutions that involve (a) criteria to establish their boundaries and to distinguish their members from nonmembers, (b) principles of sovereignty concerning who is in charge, and (c) chains of command delineating responsibilities within the organization” (Hodgson 2006, p. 18). Thus government is an organization in the sense of being a special type of institution.
attempt to reinforce the dose by viewing the internal organization of government through the lens of a public theory of the firm: it can be interpreted as an initial step for a more coherent theory in the spirit of Vincent Ostrom's (1987, 1991) integrated federalism project that distinguishes between constitutional and day-to-day government rules (Bish 2013, p. 237). A project that also more generally connects to constitutional political economy (Buchanan 1990; Congleton, Kyriacou, and Bacaria 2003), the law and economics of the constitution (Cooter 2000), and political decentralization (Treisman 2007).

COHERENCE IN ANALOGY OF CONTEXTS

Considering government in terms of the theory of the firm can run up a red flag. For it can bring to mind innate differences between firm and government, casting doubts on the entire enterprise of a public theory of the firm. One often emphasized difference is the more complicated objective function: a firm is profit oriented; while the nature of government’s function really includes multiple, often overlapping and more vaguely-defined foci, such as fix roads, improve taxpayers' health, and reduce pollution (Wilson 1989).

Notwithstanding this and other differences, such as selection mechanisms and hardness of budgets, what allows the analogy of contexts is that government and firm are cut from the same cloth, namely both are composed of rules of behavior (Hayek 2013[1973]; Coase and Wang 2012). This implies that organizational problems are endogenous to both (Breton 2000; Dixit 2008). Simply put, issues of credible commitment, decision allocation, knowledge distribution and use, monitoring, performance shading, transaction costs, and the like are, mutatis mutandis, central organizational challenges for both government and firm.

At the same time, looking at matters from the reverse perspective, one can reasonably argue that a nontrivial slice of the Coasean scientific enterprise, as well as its legal subtext, was always concerned with the public sphere. For example, already in the 1930s Coase explored the effects of government ownership of British utilities, an exploration that turned out to be very influential for “The Federal Communications Commission” (1959) and the derived “Problem of Social Cost” (1960). Moreover, in “The Lighthouse in Economics” (1974) Coase again explores comparative institutional efficiencies by considering the public and private supply of lighthouse services (Shirley, Wang, and Ménard 2015, pp. 230-233, 236).
The ingredient shared by all these Coasean works is that government is in reality never absent as concerns defining property rights: through its ultimate authority, government prescribes a set of use rights for each resource (Merril and Smith 2011). The relevance of this shared Coasean ingredient in our context manifests through a parallel: another way to think about the suggested ex ante and ex post organizational moments of a federation is in terms of the definition of in rem and of ad personam rights. In the Coasean world where property rights matter and transaction costs are positive, the ex ante moment can be considered as the delineation of property rights in rem, that is, as the fixed costs for an organization in setting up protected spheres of “authority” (Alchian 1965) against all (think of the constitution); while the ex post moment can be considered as regarding bilateral or multilateral identifiable property rights relationships, such as the variable costs tied to changes in ownership rights to policy when intergovernmental externality internalization is needed (ad personam rights).

HOW DOES THE THEORY OF INTERNAL ORGANIZATION OF THE FIRM RELATE TO THAT OF INTERNAL ORGANIZATION OF GOVERNMENT?

The study of the internal organization of government has a long and distinguished tradition in the public economics analysis of federalism, the modern starting point of which is, of course, Oates’ Fiscal Federalism (1972). This seminal contribution by the late Oates is grounded in the Pigouvian approach found in the equally seminal works of Arrow, Musgrave, and Samuelson. Contributions within this tradition are now commonly referred to as First Generation Theory of fiscal federalism (FGT) to distinguish them from an emerging body of literature that more explicitly deals with organizational aspects, namely the Second Generation Theory of fiscal federalism (SGT) (Oates 2011).

Implicit in the study of fiscal federalism are two archetypal assumptions. We can call the assumptions, as already indicated at the outset, ex ante and ex post. The

---

4 Somewhat paradoxically, though Coase can be seen as holding a Legal Realist position as regards to property that sees government prescription in use of resources, he neither formally defined property rights (Hodgson 2015) nor thought it was possible to crisply define the firm (Coase 1937, p. 392, n. 1).

5 Allen (2015, p. 382): the “definition of transaction costs that works is fundamentally related to ‘economic property rights.’ Namely, transaction costs are the costs of establishing and maintaining economic property rights. ... Following others, economic property rights are defined as the ability to freely exercise a choice.”
assumptions cut across common concerns of both a-organizational (First Generation) and organizational (Second Generation) theories, and for this reason we heuristically can think of them as being ‘meta’ in nature. We begin by quickly introducing the gist of the assumptions for the two sets of federalist theories, and then theoretically expand on the assumptions from the perspective of a public theory of the firm.6

Two archetypal (meta) assumptions in the theory of fiscal federalism

The ex ante assumption concerns the motivation behind why one would decentralize the public sector in the first place. The closest that we have to an explicit theoretical answer in the FGT remains Tiebout (1956). In light of the nonexclusive and nonrival nature of a public good, individuals have a tendency to not reveal their preferences. The lack of preference revelation is an externality. Given the nature of a public good, the market cannot properly price it, and for this reason the market is said to fail. But if we decentralize the supply of public goods to a series of local governments, then individuals will have a propensity to reveal their preferences spontaneously by moving to the government that comes closest to satisfying their public good demand. In this way the public sector internalizes the externality from lack of preference revelation.

Thus, to Tiebout we owe the important idea of voting with the feet as an institutional alternative to the ballot, that is, consumer mobility substitutes the vote when knowledge about preference for public goods is required. Yet the idea materializes only within a given supply-side: the local governments to which individuals with different preferences can freely sort themselves into are present a priori. And this implies, albeit one can identify institutional traits in Tiebout (Garzarelli 2004a), that the organizational design problem is exogenous even though we have an explicit motivation to decentralize.

The ex post assumption acknowledges and aligns with the ex ante assumption from the Tiebout motivation. However, it regards, in line with Oates' interests, above all studying the policy measures to take when there are interjurisdictional externalities in a decentralized public sector that is already in place. That is to say that the analysis shifts to the failure of another kind of decentralized institution: we

---

6 As will become apparent, the core of the argument is in reverse order of logic (from ex post to ex ante), but merely from expository flow.
pass from a market failure to a failure of a federal government organization. Some elaboration proves useful.

All contributions dealing with the ex post assumption have as point of reference Oates’ Decentralization Theorem. The Theorem posits that “in the absence of cost-savings from the centralized provision of a good and of interjurisdictional external effects, the level of welfare will always be at least as high (and typically higher) if Pareto-efficient levels of consumption of the good are provided in each jurisdiction than if any single, uniform level of consumption is maintained across all jurisdictions. In this way the [Theorem] establishes, in the absence of other kinds of offsetting benefits from centralized control, a presumption in favor of decentralized finance” (Oates 1972, p. 54, original emphasis; see also p. 35).

It is no exaggeration to say that Oates’ entire research program in fiscal federalism genuinely concerns studying situations when the Theorem does not hold in order to institutionally work out the arrangement of fiscal responsibility within an existing decentralized public sector. Differently put, though fiscal federalism is about both normative and positive issues, the Theorem, also by subsequent, repeated admission by Oates (e.g., 1999), is ultimately about normative analysis.

The ex post FGT assumption sees government as composed of at least two levels, central and local. The central government is a benevolent social planner that deals with distribution and stabilization, and supplies national public goods (e.g., common defense, foreign affairs). The local governments are benevolent social planners as well. However, they deal, à la Tiebout, with local public good allocation in the attempt to more precisely satisfy different dispersed preferences. But local governments (which, like levels of government, must be at least two in number) could fail to coordinate in order to correct spillovers among their jurisdictional boundaries (adapt different standards, leave economies unexploited, pass conflicting laws, etc.). The central government, however, never fails, saving the day with appropriate spillover internalization through coherent, locally targeted policy (Inman and Rubinfeld 1997a, pp. 45-8). It thus corrects for interjurisdictional spillovers from local public goods through matching grants (Pigouvian unit subsidies); and the greater the extent of the spillovers, the greater the extent of direct central intervention through grants (the larger the unit subsidies), and vice versa.

Central government intervention through Pigouvian grants is the normative manifestation of the Decentralization Theorem, or, more precisely, of the failure of...
the Theorem. It is the ex post assumption operating in practice, meaning that in the FGT grants are the policy tools permitting the variation in the extent of decentralization within an existing decentralized public sector. The intuition is that centralization increases when the (benevolent) central government intervenes to internalize an interjurisdictional externality – the extent of central intervention coincides with the amount of the grant transfer.

The SGT represents the most explicit link with Coase when it comes to considering a public theory of the firm. It is motivated by the a-organizational (and benevolent) approach innate in the FGT from the Pigouvian legacy. Thus, we find the drive to employ various insights from theories of economic organization to the public sector.

The SGT has something very specific in mind when it comes to considering the public sector organizationally: incentive alignment between consumer-voter and political representative. Two approaches dominate: incomplete contract and principal-agent.

The application of contract theory is mostly employed to study the appropriate degree of decentralization in a federation in analogy to an optimal delegation problem between consumer-voters and elected officials. The election is the equivalent of the contract, which is considered incomplete because consumer-voters and government cannot verify community welfare. In the incomplete contract approach, decentralization increases the probability that the vote of a consumer-voter will decide who wins elections in a given local region, the result of which is an increase in accountability. However, since, unlike centralization, under decentralization consumer-voters of one region cannot determine who wins elections in, and thus controls, another region, locally elected officials will disregard the spillover effects of the supply of public goods on other regions. In short, decentralization fails in the presence of spillovers (Seabright 1996).

A similar result holds in the principal-agent approach. Consider, for instance, the principal-agent approach known as the common agency model. Under the common agency model, each consumer-voter offers an incomplete contract to one political agent, taking as given the contractual relation between other consumer-voters and the agent. Under centralization, the agent is the same for all consumer-voters. Under decentralization, the agent is the same only for consumer-voters of the same local region. Since each consumer-voter takes the contracts of all other consumer-voters as given, each will try to free ride on monitoring the agent. The
higher is the number of consumer-voters who contract with the agent, the higher is
the probability of free riding since the costs of detection would decrease with the
number of consumer-voters. And, since the number of consumer-voters is higher
under centralization than under decentralization, decentralization exhibits better
monitoring than centralization. Hence decentralization improves political
accountability. However, since consumer-voters in a given local region are allowed
to contract with the agent in the relevant region but not with agents of other
regions, agents will ignore the effect of spillovers to other regions. This problem is
absent or mitigated under centralization because there is only one agent with whom
all consumer-voters contract. As a result, centralization is, more than
decentralization, able to internalize spillovers (Tommasi and Weinschelbaum 2007).

The upshot is straightforward. Under both the incomplete contract and the
principal-agent SGT approaches, the choice to centralize or decentralize depends on
the trade-off between policy coordination (reducing interjurisdictional spillovers)
under centralization and accountability to local jurisdictions (improved monitoring
of political behavior) under decentralization. In different terms, there is an
organizational design trade-off between ownership rights to policy and rent-seeking
control.

Notice how both of these intriguing SGT approaches deal with the ex ante
assumption, namely with the costs and benefits of decentralization as such. They do
not deal with how the arrangement of fiscal responsibility can vary within an
existing decentralized public sector. In other words, they do not explicitly consider
the ex post assumption so dear to Oates. An assumption that effectively concerns
the policy side of the trade-off: the role of intergovernmental grants. In fact, the
neglect of the role of grants is at the core of Oates’ (2005) sweeping evaluation of
the SGT and germane theories from public choice (Brennan and Buchanan 1980)
and political economics (Lockwood 2002; Besley and Coate 2003).

**Ex post, intergovernmental grants, and transaction costs**

Others, however, have alluded to what can be considered a Coasean role of grants
before the SGT came to be. I have in mind Brennan and Pincus (1990), which is
standard public choice in the sense that it considers the public sector, and
economics more generally, as about exchange among individual interests rather than
optimizing. The exchange angle is what renders the approach congruent with
Coase’s work.
Grants are ex post policy tools for adapting to different externalities (e.g., horizontal imbalance, uneven local service delivery). As alluded to above, the grant most central to the (failure of the) Decentralization Theorem, and under consideration here (though the same basic logic applies to all types of grants), is the one for the internalization of interjurisdictional externalities: a Pigouvian unit subsidy or, if you prefer, a cost-matching formula. The idea is that the central government will match each dollar spent by the local government on a particular public good or service that generates an interjurisdictional externality (a bridge built by one province but connecting two, meningitis vaccination sponsored by one region but open to all, natural resource exploration by one state damaging the environment of another, etc.) with a specific sum of money. For example, the central government could match each dollar spent by the local government with a dollar.

There is quite a lot of what Coase calls “blackboard economics” in this normative Pigouvian reasoning. “All the information needed is assumed to be available and the teacher plays all the parts.” The teacher “fixes prices, imposes taxes, and distributes subsidies (on the blackboard) to promote the general welfare. But there is no counterpart to the teacher within the economic system. There is no one who is entrusted with the task that is performed on the blackboard.” The principal reason is that the relevant private- and public-sector knowledge about how to promote welfare is dispersed. There “is no single entity within the government which regulates economic activity in detail, carefully adjusting what is done in one place to accord with what is done elsewhere. In real life we have many different firms and government agencies, each with its own interests, policies and powers” (Coase 1988, p. 19). This is a problem that Oates – himself a pioneer of normative Pigouvian reasoning in fiscal federalism – eventually admitted to. The idealism of blackboard

---

7 I first interacted in person with Oates in 2002 in Pavia and Rome, Italy, and on several other occasions during the following years in Rome (where Wally used to travel regularly) and once in New Orleans, US. I first exposed Wally to the SGT in 2002, which eventually culminated in Oates (2005), while, at the same time, Wally persuaded me to revise a typescript (basically some old notes later collected as Garzarelli 2004b) about the SGT that ultimately became two papers (Garzarelli 2004a, 2006). Wally always emphasized in these interactions that fiscal federalism, unlike, e.g., political federalism, is really about the economic dimension of a decentralized public sector, namely the coordinative role of grants as fiscal institutions – basically the ex post assumption. While I think that I managed to raise Wally's awareness about the knowledge problems innate in FGT reasoning, I also think that I was less successful in convincing Wally about at least questioning more often the benevolence assumption of government. See also Garzarelli and Limam (2003).
economics manifests most clearly when one considers the preference revelation problem in relation to intergovernmental grants, that is, the ex ante assumption in relation to the ex post one. There are some quite strong assumptions made ... that don't seem fully consistent with one another. We assume that the central government knows the preferences of individuals for national public goods, but not for local public goods. This seems a strange dichotomy. One might justify it on the grounds that acquiring information on national public goods is worth the cost to central authorities, while ... the value of such information on local public goods is not worth its cost of acquisition. Alternatively and less formally, we might simply presume that central government information is imperfect for all public goods, but that central provision of truly national public goods is likely to produce a better outcome than one in which local jurisdictions ignore the benefits that their outputs confer on those in other jurisdictions. Pushing this point further, if central government has little knowledge of local preferences, how can it determine the correct level for Pigouvian subsidies for local outputs that generate interjurisdictional spillover benefits? The measurement of the spillover benefits themselves requires local information on the valuation of the benefits. Again, one might argue, I suppose, that the somewhat imperfect subsidies are likely to produce an outcome that is typically better than one in which the spillover benefits are ignored entirely. But there clearly are some basic information problems here (Oates 2005, p. 359).

One way to succinctly restate Oates' observations along Coasean lines is to say that *vertical* transaction costs – those case-by-case knowledge costs attached to the intergovernmental transfer of (ad personam) policy rights for externality internalization – matter for policymaking.\(^8\)

The useful Brennan-Pincus connection here is that a grant should be considered in institutional terms: as a *quid pro quo* between levels of government rather than as a windfall Pigouvian spending formula. To express it slightly differently, a grant is a fiscal agreement that is the outcome of complex intergovernmental negotiations, not an Athena-like output that springs full-blown from the head of (central government) Zeus. And if a grant is a manifestation of exchange, then it is solving a knowledge problem innate in the vertical structure of the public sector as well: the grant communicates grantor preferences about spending decisions to grantee. The backdrop is that bureaucrats and politicians operate in a public sector where

---

\(^8\) A point about the complexity of policy design earlier recognized by the literature in general, though also not explicitly couched in transaction cost terms but in terms of Mirrlees-Vickrey imperfect information (Boadway 1997).
reputation effects on political promises glue together intertemporal exchange relations tied to individual career ambitions. The incentive to fulfill grant expectations is then self-interest – to stay in good terms with political and bureaucratic colleagues for reasons of personal career ambition, re-election, keep a committee seat, etc. (Garzarelli and Keeton 2017).

The Coasean prism consequently suggests that we are in the presence of a *transaction* between levels of government: the intergovernmental transaction exchanges property rights to policy based on promises from “mutually dependent interests” (Commons 1932-1933, p. 4). The fact that the mutually dependent interests ultimately embedded in the grant are effectively parochial, means that ex post organizational adaptation is often a residual: it is the result of individual political and bureaucratic incentives that often have little to do with the explicit objective of government survival for reasons of the public good. However, my stance on the matter is positive, not normative: survival tied to extra-economic incentives is innate in the nature of the organizational beast (Weingast and Marshall 1988).

The implication is that the vertical transaction for the transfer of property rights to policy through the grant is *more costly* for reasons of “negotiation” than for reasons of “enforcement” (Cheung 1969; Allen 1991). The extra-economic incentives generate “security of expectations” (Commons 1932-1933, p. 4), rendering the grant – the institution for adaptation to contingencies – an implicit contract: it is not strategically rewarding to deviate from the terms of the contract in that the context of the game played is political, and the political game is seldom one shot (Oates 2008).⁹ That is to say that a grant represents an ad personam institution for the exchange of economic property rights to policy. This logic leads us to suggest that (variable) intergovernmental transaction costs tied to rights to policy have more of a net impact for their negotiation (or coordination) facet than for their enforcement (or incentive) facet.

A contract in Coase (1937) is an ad personam compromise for adaptation to contingencies as well, and the more a contract is incomplete, the more likely is a firm to adapt. In a firm, “the service which is being provided is expressed in general

---

⁹ This is not to mean that incentives are always aligned. For example, when in Australia a grantee (a public university) failed to take the preferences of the grantor (the government) into account, the result was punishment through a cut in grant funding (Brennan and Pincus 1990, p. 130). However, since it is not strategically rewarding to not fulfill grantor expectations, on balance I would consider misalignment to be mostly an exception (Garzarelli and Keeton 2007).
terms, the exact details being left until a later date. All that is stated in the contract is the limits to what the persons supplying the commodity or service is expected to do. The details of what the supplier is expected to do is not stated in the contract but is decided later by the purchaser.” Therefore, a firm internalizes “the direction of resources” in the effort to also capitalize on, rather than just to obviate, contingencies (Coase 1937, p. 392).

In a public theory of the firm, this Coasean view to organizational adaptation boils down to the following proposition about internal organization. Just like the extent of firm decentralization is a dependent variable of the extent of contractual incompleteness, so the extent of government decentralization is a dependent variable of grant incompleteness. By grant incompleteness I mean that a grant that is less conditional – a grant that has fewer strings attached – is more incomplete (Garzarelli 2006). Ex post this entails that, ceteris paribus, public sector adaptation is not about going back and forth between centralization and decentralization as such, but concerns the amount of decentralization as well. That is to say that public sector decentralization is not an all-or-none phenomenon: for adaptive reasons, the extent of public sector decentralization can vary in any point in time as a result of how conditional a grant is (Garzarelli and Holian 2014).

Notice how in this case the incomplete contract – the grant – is not about ex ante organizational incentives between consumer-voter and political agent as in the existing SGT literature. Rather, it is about agents who ex post transact at different levels of government for their own public choice survival, and, in the process, residually generate, at a cost, policy coordination.

At the same time, there can be ex post vertical transaction costs that are bottom-up. These costs refer to grants from local to central government. Think of the European Union’s budget that originates from grants from member states. Similarly, there also can be ex post horizontal transaction costs, which concern the internalization of externalities through grants from governments on the same level (Boadway and Keen 1996). Though acknowledged, these bottom-up and horizontal transaction costs are not in a strict sense explored by Oates’ original approach nor by Oates’ subsequent observations about the limitations of the SGT.

No matter their nature (vertical or horizontal) or vertical direction (top-down or bottom-up), however, the uniting thread of these transaction costs is that they all ultimately refer to the ex post moment. That is to say that within grant incompleteness there also lie transactions and their variable costs: the phase of
organizational “adaptive efficiency” (North 1990, p. 80) to different problems when a federation is already in place. But the public sector, as we previously learned especially from the SGT, is also to be explored in terms of ex ante organizational design, which is also not a transaction-cost free lunch.

**Ex ante, public sector organizational design, and transaction costs**

The ex ante moment can be likened to what fiscal federalism calls the assignment problem, which studies what level of government is responsible for what duty. We can think of this first moment as the economics of organization equivalent of constitutional engineering or design (Sartori 1994; Ginsburg 2012).

In the FGT, the ex ante solution to the assignment problem is straightforward Pigouvian: it is solved by default in terms of the three functions of allocation, distribution, and stabilization. Thus, the typical solution is that local governments should deal with allocation and central government with distribution and stabilization. The reason for this division of labor is equally straightforward: local government should allocate public goods in light of its greater proximity to the consumer-voter. This default solution residually implies, as pointed out earlier, that the central government internalizes externalities between lower levels of government.

The SGT motivation for ex ante decentralization, we saw, is about improving accountability tied to political delegation. Indeed, the notion that decentralization should be favored to improve political accountability is the closest statement that the SGT has to a decentralization theorem. But the SGT stops short of proposing how to structure the internal organization of government in order to achieve this accountability. And yet, it normatively aligns – by default – with the Pigouvian view when there are externalities. Namely, the spirit is that local governments fail in the presence of spillovers because the (horizontal) transaction costs for autonomously coordinating the externality internalization are always too high (among others, Inman and Rubinfeld 1997a, pp. 48-50; Lockwood 2002, p. 319); albeit there is the qualification that the centralization-decentralization trade-off hinges on externality size and not just on externality as such (Besley and Coate 2003).

At the same time, there is no indication that the solution of centralized externality internalization is necessarily less costly, to wit that vertical transaction costs of internalization are lower than horizontal ones. Indeed, there is no a priori reason as to why this should be so (Breton and Scott 1978; Inman and Rubinfeld
1997b; Luelfesmann, Kessler, and Myers 2015). This normative alignment by default is therefore rather puzzling, not in the least because the SGT is predicated on the objective to provide a theory of decentralized public sector organizational design, and not just a motivation of it (Qian and Weingast 1997).

Essentially, the SGT (like the FGT) solves the ex ante organizational puzzle of vertically and horizontally identifying what government should be assigned to what function by difference rather than endogenously. This reasoning brings to mind blackboard economics, because there is no comparative-institutional explanation about when transaction costs to perform a function are too high and when they are not. The question of a theory of decentralized public-sector design is consequently left open. In terms of our proposed public theory of the firm, the answer to the question hinges on how internal organization of government should be structured in order to deal with the problem of efficient externality internalization. This means that we face a fixed-cost issue of setting-up in rem rights over policy to different levels of government.

What organizational design principles can we draw upon to consider this complex ex ante issue? A promising set of principles is modularity (Simon 1962; Alexander 1964). The premise of modularity is simple. It is easier to tackle a complex issue by decomposing it into smaller parts, namely modules. Recall how we all learned to solve mathematical problems by breaking them down into more manageable sub-problems.

Modularity shows how when dealing with complexity the minimization of interactions among modules is key. The minimization of interactions can be achieved by letting the modules be less directly dependent on one another – by letting the modules communicate only seldom, allowing within-module communication more often. This particular process of modularization is known as near-decomposition (Simon and Ando 1961).

As an organizational design, a nearly-decomposed modularization stands between full decomposition and integrality (or non-decomposition). A fully decomposed system is not so interesting for our purposes. It refers to a

---

10 More precisely, there is no such explanation that in addition to collective decision rules also considers organizational design. See for example Inman and Rubinfeld (1997a, pp. 48-53) and Cooter and Siegel (2010) on unanimity versus majoritarian voting or even Casella (1999) on tradeable deficit permits. My point is that we need to consider the design of organization as much as that of decision rules: the relationships between the two are not trivial and can be complementary as well as substitutable. In brief, both Coase (1937) and Coase (1960) matter.
modularization where there are zero inter-modular interactions since the modules do not share a common end – each module pursues its own idiosyncratic objective without a link to an overarching organizational one. The market comes to mind (Hayek 1945). An integral system is at the polar extreme: every module must communicate with all other modules at all times, implying that each part of the system is completely dependent on every other part in order for the system to pursue its objective. It is obvious here that the magnitude of the transaction costs would be formidable given the complete graph nature of the design. To borrow Tullock’s metaphor (1969, p. 23) from a related discussion on federalism and local government scope, it is easier for restaurant patrons to mix and match menu options from a short menu rather than to face the choice of selecting from a much longer menu that already contains all possible mix and match options. The lesson is that transaction costs from coordination and communication rise faster than benefits when the addition of more organizational elements (divisions, individuals) also requires more integration among all elements.  

This leaves us with near-decomposition. Near-decomposition, like integrality, is not perfect. However, the ex ante costs of its decentralized design are often worth paying in order to minimize the ex post costs of organizational adaptation (Langlois and Garzarelli 2008). How can this result be achieved?

A near-decomposition rests on two broad sets of design principles: visible design rules and hidden design parameters (or hidden information). Visible design rules establish modules and their purpose (an architecture), how the modules interact (an interface), and the relative fit and efficiency properties of a module relative to the organization as a whole as well as to other modules (standards). The hidden design parameters instead refer to the inner structure of a module that, to avoid too many unmanageable interactions, should not be shared among modules. The hidden design parameters should not only manage the inner workings of a module. They should concurrently be in harmony with the overall organizational objective, something that is accomplished through a well-functioning interface (Baldwin and Clark 2000).

It is rewarding to establish and maintain (and usually to adjust when necessary) ex ante rights to exclude through the design principles of modular near-

11 A challenge famously faced by software engineers before modularity in programming, and that became known as Brook’s Law.
decomposition – i.e., in rem shielded domains of authority – because it is not economical to spell out ad personam rights every time a transaction is needed (Coase 1937). Or, in Smith's (2012) vivid image, near-decomposition, like an economy's structure of property rights more generally (Coase 1960), is the fixed transaction cost shortcut that is produced in order to replace those variable transaction costs that emerge when trying to map out all pairwise contractual exchange relations. The establishment of standards (e.g., a common accounting system for tracking and comparing health-care expenditure across federated states) and the exchange of commitments subject to punishment (e.g., the Stability and Growth Pact of the EU) can be considered two illustrations of this shortcut.

Take note that this is not necessarily tantamount to asserting that, in a good decentralized design, externalities are minimized, reduced in size or eliminated. A near-decomposition is a social artifact. As such, it is fallible like its designers – its property rights, no matter how well-defined, cannot be perfect. Inter-modular externalities will still crop up; and this is another instance to take to comparative-institutional analysis to see what module(s) should perform the internalization (and how) (Langlois 2002).

This is a notion that the economics of organization has long understood when answering the second of the three Coasean questions – namely that the organization-market boundary, even in public contexts (Hart, Shleifer, and Vishny 1997; Nelson 1997; Levin and Tadelis 2010), is determined according to the minimization of internality cost. In terms of our main concern, this reasoning suggests that when designing an internal organization of government what we are in effect doing is designing two sets of boundaries: those between market and public sector organization in the broadest sense, and those among the modules that compose the internal organization. The first set of boundaries is the one that we are most familiar with: rules of behavior with a common end (organization) replace more abstract ones (market) (Hayek 2013[1973]). The second set is the quintessence of near-decomposition. It refers to the demarcation line among modules, which hinge on communication as needed, rather than at all times; said another way, near-decomposition minimizes the costs of externality internalization when the costliest transactions are left to the modules. This is achieved through common interfaces, which are key to keeping a nearly-decomposable system  

12 And even if they were perfect, it does not follow that transaction costs are zero (Allen 2015).
together (e.g., public law). A well-functioning “common interface enables, but also governs and disciplines ... communication” among modules without letting the system drift into full decomposition (Langlois and Garzarelli 2008, p. 128).

A nearly-decomposed modular system is therefore a nested hierarchy. In terms of structure, it brings to mind a Russian matryoshka, a hollow wooden doll containing a hierarchy of an arbitrary number of similarly-shaped dolls that decrease in size and that are all stacked one inside the other. In terms of function, it has the characteristic that interactions among modules are less frequent than those within modules.

At this juncture, the parallelisms between the principles of modular near-decomposition and of federalism should be intuitively evident. An architecture refers to the federated states and their policy responsibilities (modules), an interface to the (public) rules of interstate governance (e.g., constitution and other comparable legislation, inter-regional law, treaties in supranational federations), and the standards to performance (ceteris paribus, the supply of the same public service should be of comparable quality across states, interjurisdictional externality internalization should have the same cost no matter what federated state performs it, etc.).

What a well-crafted, near-decomposed internal organization of government can do is incentivize a federated state to more carefully plan the use of its resources and assets. This would lead to an improved use of local knowledge and aid the division of labor among and within federated states (Tullock 1969, p. 28). Further, when the Decentralization Theorem fails one would not automatically resort to central intervention, but actually consider the (variable) vertical transaction costs of central intervention against the (variable) horizontal transaction costs of autonomous externality internalization.

Clearly, modular near-decomposition is no easy matter, and as such it is subject to trial-and-error. For example, when experimenting with configurations, one should take into account how module size and not just module number can impact overall efficiency (Shrestha and Feiock 2011). Moreover, my sense is that public contexts, with their greater susceptibility to rent-seeking captures, pressures and tussles than private ones, arguably exhibit higher probabilities of design failure. In

---

13 For a full-blown articulation of the parallelisms that also models cost-benefit implications of feasible, alternative public-sector near-decompositions, see Garzarelli and Sitoe (2018).
this sense, a clear illustration of a badly-crafted near-decomposition from too many political interests is so-called Italian-style federalism (da Empoli 2014). Ultimately, the type of modularization that a system should undergo is an empirical issue (Smith 2012). Luckily, the off-the-shelf principles of modular near-decomposition are themselves a modular system that can be tailored to existing institutional settings.

TOWARD A COASE-OATES FEDERALISM NEXUS

We have teased out the following logic from received theories of fiscal federalism.

(1) The market always fails when it comes to public good supply, that is, a public good always generates at least an externality.

(2) The normative implication is that government ought to supply public goods.

(3) However, government can also fail in its supply of public goods when allocation only stems centrally. As a result, correction of central government failure is achieved by letting local governments deal with public good supply, that is, through fiscal decentralization of government.

(4) Still, after public decentralization is in place – namely, once we have at least two levels of government and at least two local governments with some degree of fiscal autonomy – local governments fail to autonomously coordinate to internalize all local externalities they are involved in, calling for central government intervention through grants.

(4´) The reason why there is failure of autonomous externality internalization in (4) is imputed, even if not always explicitly, to excessive transaction costs, but the institutional reason behind these transaction costs is never genuinely explained. Recall the blackboard economics image.

Received theories of fiscal federalism underscore that to the three feasible institutions studied – market, central government, and local government – can correspond also three institutional failures. I do not dispute that the three institutions in question can fail. What is needed, however, is a more precise account
of the conditions under which they fail: an assessment of the relative costs and benefits of using one institution over another. This paper argues, maybe in a too desultory fashion, that we could benefit from a more careful application of the basic principles that underlie the comparative institutional calculus, from both an organizational design and an organizational adaptation perspective. In particular, it adds the following propositions to the logic of received theories.

(5) Intergovernmental transaction costs vary in magnitude; in other words, they are not necessarily always excessive.

(5’) Moreover, transaction costs vary in direction (horizontal; top-down and bottom-up vertical) and type (fixed; variable).

(6) The nature of the decentralized organization of the public sector – the in rem structure of property rights to policy – is not a matter of indifference: alternative ex ante feasible modular near-decompositions differently affect the magnitude of variable transaction costs ex post. This is federalism in its organizational design guise.

(7) Ex post coordination in a decentralized public sector – the ad personam allocation of property rights to policy – depends on the magnitude of variable transaction costs. This is federalism in its organizational adaptation guise.

(7’) The horizontal buy (or price control) choice through grants (alternatively taxes) should be considered alongside the vertical one when assessing costs and benefits of coordination.

The substantive implication is that the demarcation line among the three functions of allocation, distribution, and stabilization would be endogenously determined. Except perhaps for a handful of more strategic policy rights that one could assign centrally – monetary policy and sovereign transactions come to mind – assignment in general (i.e., not just tied to externality internalization) would be the output of an institutional comparative efficiency criterion. So differentiation among functions would be less stringent, more fluid. After all, the assignment of

---

14 Sovereign transactions are those public-sector transactions that ought to be performed centrally rather than locally not necessarily for cost-saving or efficiency reasons, but for reasons of security (e.g., foreign affairs, foreign intelligence, military) (Wilson 1989; Williamson 1999).
duties already does not always directly translate into a crisp central-local and local-local government differentiation (e.g., tax and expenditure decisions designed to achieve optimal allocation of resources affect income redistribution and the level of employment and price stability).

In my view, the organizational economics of fiscal federalism remains an open research agenda. The incorporation of transaction costs and property rights into the analysis of the internal organization of federalism – the Coase-Oates federalism nexus – represents a first module to building a public theory of the firm.

REFERENCES


