6 THE NASH-LINDAHL THEORY

1 THE GENERAL PROBLEM

We consider a **group** of people (*individual* agents) bound together by one or more **common non-rival shared interests**, whose satisfaction requires some kind of **cooperation/collective action**. Cooperation means **sharing the costs** of the collective action. As for the benefits, since the interests are of the common = non rival type, there is no problem of sharing them (the problem of benefit sharing arises if cooperation concerned rival interests, as in the case of unions and cartels). We shall use the variable G to denote both 1) the amount of collective action required for satisfying those common interests, and 2) the amount-quantity of the **group goods** satisfying them, *non rival* and *non excludable* among the group's members. The notation G refers to *public goods*, i.e., in our terminology, to a particular category of shared interests, namely the *public/collective interests of a political community*, but strictly the formal arguments of this Chapter are independent of this further restriction.

A more detailed discussion of cooperation/collective action is given in Chapter 1, Section 1, §§ from *Collective action/cooperation* to *Incentives to cooperate*.

For an advance insight into the topic of this Chapter we present a short statement of what we shall call the **Nash-Lindahl theorem,** NLT:

*Let G be the amount-quality of the goods satisfying the common interests shared by the group's members, and suppose the subjects are in a social state of **full freedom of action**, which means that there is no social coercive power capable of forcing them to do or not do something, i.e. **no social organization**. Then:*

*1) there will always be a **non-Pareto Nash equilibrium** level of G, **lower** than the **Pareto non-equilibrium Lindahl** level, and*

2) as the size of the group increases the Nash equilibrium level of G decreases, and tends to zero as the size of the group tends to infinity.

The Nash-Lindahl theorem is therefore the theorem stating ***the failure of free cooperation in the public economy***.

The Chapter is organized as follows.

First, we begin with a **graphical study** of the **individual incentives to engage,** or **not engage, in the collective action**, focusing on two qualitatively different stylized situations. Assuming for simplicity a group of **equal (representative)** agents, we consider in F6.1 a **small group**, SG, of 2 subjects and in F6.2 a **large group,** LG, of N subjects

Second, we present the *classical Lindahl equilibrium* argument.

Third, we present and *prove graphically* the NLT for *two agents*.

Fourth, we use the Cesi-Gorini diagram to generalize the NLT to any number of subjects.

2 A SMALL GROUP OF TWO: THE INCENTIVES (PAY-OFFS)

In spite of certain similarities, the case of the **Prisoners' game,** dealt with in Chapter 4 is different from the SG case of *2 subjects* discussed here in F6.1. The PG is a case of **strategic coordination failure** where the nature of the group and of the actions and interests involved are of no importance, whereas in the small and large group discussed here the nature of the actions and interests involved are part of the model: **the interests shared by the agents are of the common type**, and the amount/quality of the activity serving its satisfaction, denoted by G, is not fixed but **variable**. When the shared interests are of the common type, **cooperation/collective action** is the only possible type of coordination because the conflicting modality of the market is physically impossible. We shall therefore specialize our terminology by speaking of **cooperation failures in the public economy,** i.e. **in the pursuit of shared public interests** by taxpayers, voters, citizens.

F6.1 shows the case of **2 equal agents** with a common shared interest G. It shows how to use areas to represent the individual *payoffs* of different choices. The **possible individual choices** are many, but for our purpose it is enough to consider **only 4** of them: 1) unilateral action, 2) free riding, 3) equal cooperation, 4) no action.

1) **Unilateral action**. The *thick blue line* **b-k**, denoted $MB(G)$, is the *individual marginal benefit schedule*, assumed to be the same for all agents. The *thin blue line* **a-k**, denoted $SMB(G) = 2 \times MB(G)$, is the *social marginal benefit schedule*. The *thick red line*, denoted $SMC(G)$, is the *social (=total) marginal cost* of providing G. If point **b** ($MB(0)$), lies above point **c** ($SMC(0)$), as in the Figure, each agent has some incentive to *act unilaterally*, i.e. to provide some G bearing its full cost, of course only up to some level G_U where his individual marginal benefit equals the social marginal cost, $MB(G) = SMC(G)$. The **yellow area** measures his *incentive* (= to his *payoff* or *net benefit*) to act unilaterally.

2) **Free riding**. If one agent contributes (up to G_U), the other has an incentive (payoff, net benefit) to not contribute at all, exlploiting the amount G_U paid for by the former. His incentive to free ride is the area *bordered in grey* $\overline{beG_UO}$ (notice that part of it is covered by the other areas!).

3) **Equal cooperation**. Equal cooperation by equal agents means sharing in half the social marginal cost. If the two cooperated equally they would carry G up to the *efficient* level G^* where the individual marginal benefit is equal to the individual marginal cost and the social marginal benefit equals the social marginal cost:

$$SMB(G) = 2 \times MB(G) = SMC(G) \rightarrow MB(G) = \frac{1}{2}SMC(G)$$

The **yellow + orange** area measures the payoff (net benefit) of each agent.

4) **No action**. With $G = 0$ the payoff obtained by each agent is zero.

F6.1 does not tell us - in itself - what would be the strategic equilibria of this game. It only tells us something about the nature and size of the **incentives** (payoffs, net benefits) associated to the different strategies. It tells us, for instance, that the incentive to equal cooperation is higher than that of unilateral action.

It tells us also something about one particular frequent outcome under conditions of full freedom of action: if one agent, say **A**, *refuses to cooperate*, i.e to contribute, then agent **B** has no other choice than *unilateral action.* In this strategy **A**'s payoff is the grey area and **B**'s payoff is the yellow area. It is quite likely that *such a strategy would be Pareto-inferior relative to the cooperative one*. Outcomes of this sort are frequent in *non-organized groups*, bearing witness to the *weakness of the incentive to cooperate* relative to the *strength of the incentive to engage in the distributional conflict* (see **Chapter 5** on the *Cesi-Gorini game*).

As for F6.1bis, it is only added here to remind the reader of the simple yet fundamental property of public economics: the difference between public goods, requiring *vertical summation* of individual *MBs*, and private (market) goods, requiring *horizontal summation* of individual demands.

3 A LARGE GROUP OF N: THE COLLAPS OF FREE COOPERATION

F6.2 extends to an arbitrary group size the previous reasoning (maintaining where possible the same letters). We assume N **equal agents**, with the same individual marginal benefit schedule **b-k**, denoted $MB(G)$, and a given social marginal cost schedule **c-g**, denoted $SMC(G)$.

The agents are assumed to face, as before, *only 4 possible choices*:

1) **unilateral action**, up to the level of G where the individual marginal benefit equals the social marginal cost: $MB(G) = SMC(G)$. Since $c > b$ this choice yields $G = 0$*,*

2) **free riding**, exlploiting whatever level of G is being provided by the others,

3) **equal cooperation of all**, dividing the social marginal cost by N and carrying G up to the *efficient* level G^* where the individual marginal benefit is equal to the individual marginal cost, and the social marginal benefit equals the social marginal cost

$$SMB(G) = N \times MB(G) = SMC(G) \rightarrow MB(G) = \frac{1}{N}SMC(G) \rightarrow G^*$$

4) **no action**, yielding again $G = 0$.

Inspection of F6.2 shows us that in this LG case we can say a number of things about **incentives/payoffs**, and that, unlike in the SG case of F6.1, we can say something also about **equilibrium strategies**.

Given the new marginal benefit and marginal cost schedules, and assuming the new payoffs to be as indicated, the incentive situation is as follows:

1) **No agent has an incentive/positive payoff to act unilaterally**, because the social marginal cost is too high compared to the individual marginal benefit (willingness to pay): $SMC(G) > MB(G)$ for all G. For the purpose of giving numerical weights to payoffs, we assume the individual payoff of contributing alone up to G^* to be negative, and equal to **-80**.

2) **Each agent has an incentive to cooperate (equally)** up to G^* equal to **10**.

3) However, if the **number** N **is very large** each agent rightly assumes that the impact on the final outcome (any level of $G \geqslant 0$) of his own choice to contribute or not contribute is **negligible**. It follows that for any level $G > 0$ his incentive to contribute will **always, by construction, be weaker** than that of not contributing. In particular, if the final outcome were G^* the incentive to contribute is **10** while that of not contributing is **10+10=20**.

4) Since every agent, taken in isolation, makes the same assumption of a negligible impact of his own choice on the final outcome, under **free collective action** the unique **equilibrium outcome is** $G = 0$.

F6.2 offers a simple intuitive proof that **in a sufficiently large group, under conditions of free collective action the incentive to contribute always falls short of the incentive to free-ride and the satisfaction of the shared public interests collapses to zero:** $G = 0$ **is the unique non-Pareto Nash strategy** ■.

The emergence of the coercive power of government

At this stage of our resoning we can see that the Figure tells more than this. True, each agent regards his choice, taken in isolation, as having no impact on the final outcome, but agents are no fools: every one realizes that also every one else regards the impact of his own defection as negligible, and that therefore every one has an incentive to defect, leaving the shared interests totally neglected. If such interests are perceived as non-essential this may be the **end of the story**: no collective action will be undertaken. But if they were **perceived as essential** for the survival and well-being of the polity, then its members realize that, in order to have them satisfied to some degree, they must accept the transformation of the group from a *fully free community* into an **organized political community** endowed with the **coercive power** to convert cooperation from a *free choice* into a **public obligation**. Needless to add that the emergence of **governments and their power to tax** in human societies is a complex social phenomenon investigated by different branches of the social sciences, and that we must be aware that our present argument is *only one* of the many contributions to its understanding.

For a more extensive discussion of this point see the **Cesi-Gorini** paper **2014**, and for a general in-depth discussion of the meaning and role of the State in human society see **Bourdieu's** treatise **1992**.

4 THE CLASSICAL LINDAHL EQUILIBRIUM ARGUMENT

To place the NLT into perspective we first use F6.3 to present what is sometimes reported in the literature as **Lindahl**'s theory that an *efficient allocation* of *public goods* and *cost share distribution*, called a **Lindahl allocation**, can in principle be treated as an **equilibrium** because it can be negotiated by people *acting in condition of full freedom (**no political organization**)*. The argument is cast here in terms of only two agents **A** and **B** and one public good G, but in principle it can be generalized to any number of agents and public goods.

On the horizontal axis is the *quantity/quality of the public good*, on the vertical one are the *cost shares* of **A**, s^A, and of **B**, $s^B = 1 - s^A$. In the *background* are the respective *gross income/endowments* of the two. The **downward sloping blue curve** is **A**'s 'demand' for G as a function of his cost share, and the **upward sloping red curve** is the same for **B** (whose share is measured moving downwards along the vertical s^A axis). The two agents start with some *arbitrary cost share distribution*, a point on the vertical s^A axis, and see what are the respective 'demands' for G. If they coincide the two are at the *equilibrium allocation*. If they diverge, say if $s^A > s_{LS}$, they would begin to change the cost share distribution until they reach the equilibrium. Clearly, the 'demand' curves *depend also on the agents' endowment*, and they would change when such endowment changes. According to some authors this was regarded by **Lindahl** as the **end of the story**. But when **Nash** came along it became possible to prove rigorously that while the Lindahl allocation is indeed *efficient*, it *cannot be an equilibrium*. We have labeled the result as the NLT, and will explain and prove it in detail below.

5 THE NASH-LINDAHL THEOREM WITH TWO AGENTS

We shall provide a formal proof of the NLT in graphical (non-mathematical) terms, using F6.5, but we first need to explain how F6.5 is derived from F6.4.

Demand for *G* as a function of cost share and endowment

F6.4 shows graphically how to derive the individual 'demand' for G as a function of his cost share and income endowment $G^D(s, y)$. The function relates the **desired ('demanded')** level of G by an agent to his **cost share** s and his **gross income** y, on the assumption that the cost share is not freely negotiated among the group members, but is **fixed by an outside authority**. In F6.4a on the horizontal axis we measure G and on the vertical axis individual consumption c and gross income y. The downward sloping curves $c = y - sh(G)$ are **the taxpayer's budget lines** corresponding to different values of his cost share s.

The budget lines are not straight but convex towards the top-right simply because we assume a macro cost function $h(G)$ with increasing slope (increasing marginal cost).

In F6.4b the downward sloping **thick blue curve** represents the inverted 'demand' $G^D(s)$ for a given gross income y. For each cost share, the desired level of G is obtained where the taxpayer's indifference curve is tangent to the budget line.

Demand for *G* as a function of tax-price and endowment

For completeness we add F6.4bis. Here again the downward sloping line represents the individual's inverted demand for G, but in analogy with *ordinary demands for market-priced private goods*, on the vertical axis instead of measuring the **cost share** s, as in F6.4, we measure $sh'(G)$. This is the **tax-price**, which in turn is equal to the *Marginal Rate of Substitution* (MRS) *or Marginal Benefit* (MB)

$$p = sh'(G) = \underset{c(G)}{MRS}(G, c) = MB(G)$$

The downward sloping curves in the two Figures are thus both individual demands for G, but with different types of values being measured on the vertical axis.

The concept of tax-price is discussed in detail in Chapter 7, *Public consumption goods*.

0.0.1 The proof

F6.5 provides a geometrical proof of the NLT in the case of 2 *equal* agents (same preferences and same endowment). The reader must keep in mind that the assumption of *equal (representative) agents* and the use of *straight lines* are made only to simplify the the geometry. There is *no loss of generality* and the same result holds perfectly also with different agents and differently shaped lines. For the reader's convenience we provide a picture of such a more general case is in F6.5bis.

The proof given here is a restatement - with some graphical changes and a more detailed development of the argument - of the diagram in **Schotter 2009**, p. 596.

We measure G on the horizontal axis, and the cost share s^A of the *blue* agent **A** moving upwards on the vertical axis. The cost share s_B of the *red* agent **B** is the complement to unity $s^B = 1 - s^A$ measured on the same vertical axis moving downwards. The downward sloping **thick blue line** $D^A_{LS}(G)$ is the same thing as the downward sloping line in F6.4b: it is the blue agent's 'demand' for G as a function of s^A for a given y^A. The upward sloping **thick red line** $D^B_{LS}(G)$ represents the same for the red agent **B** for the same $y^B = y^A$ (keep in mind the assumption of identical agents). By reasoning on the two diagrams - F6.5 and F6.4a - the reader can see how the blue cap-shaped indifference curves of F6.5 are related to those of F6.4a: the points **a0-0-a0** along the indifference curve u_0 in F6.4a correspond to the same points **a0-0-a0** along **A**'s indifference curve u_0 in F6.5. From this rule of conversion between the two diagrams follow the properties of **A**'s indifference curves in F6.5 (all curves relative to **A** are in **blue**, while the corresponding curves relative to **B** are in **red**):

1) they are **cap-shaped** because, starting from a tangency point like **0** with its associated G level, maintaining the same indifference u_0 when s^A decreases (the budget curve rotates anticlockwise) requires a decrease or increase in G relative to its optimal level (F6.4a),

2) lower indifference curves, being associated to lower cost shares, correspond to higher welfare levels,

3) the downward sloping *thick blue line* $D^A_{LS}(G)$ (drawn straight only for geometrical ease) joins the *top points* of these cap-shaped indifference curves.

B's indifference curves and the *thick red line* $D^B_{LS}(G)$ are obtained in the same way, after turning the picture upside-down.

To prove the theorem we start with the meaning of points such as **B0**, **B1**, **B2**. If **A**'s preference for G were represented by the blue 'demand' curve $D^A_{LS}(G)$ we ask: what would be **B**'s best choice of his cost share s^B? It would be at point **B0** where **B** reaches his *highest (red) indifference curve tangent to* ***A****'s demand curve* $D^A_{LS}(G)$ and **B**'s cost share is $s^B = 1 - s_3$. In other words, if **B** chooses the cost share $s^B = 1 - s_3$ the cost share born by **A** would be s_3, **A** would demand G_{B0}, and at the combination $(G_{B0}, s^B = 1 - s_3)$ **B** would reach his *maximum attainable welfare* given **A**'s preference $D^A_{LS}(G)$. If we apply the same reasoning to find the best choice of **A**'s cost share if **B**'s preference for G were represented by by the red 'demand' curve $D^B_{LS}(G)$ we would find point **A0**, where **A** attains his highest (blue) indifference curve tangent *to* ***B****'s demand curve* $D^B_{LS}(G)$. In short, if $D^A_{LS}(G)$ and $D^B_{LS}(G)$ were the preferences for G of **A** and **B** respectively, the best choice of **B**'s cost share would be $s^B = 1 - s_3$ and the best choice of **A**'s cost share would be $s = s_4$. This shows that the best cost share choices of the two subjects would *not coincide*, and this means that neither **A0** nor **B0** are *Nash equilibrium strategies*.

So far we've identified the best cost share choices of the two agents if each one thought that the other's preference were $D^A_{LS}(G)$ and $D^B_{LS}(G)$ respectively. Let us now turn again to agent **B**. While he knows his own *true preference* for G (represented $D^B_{LS}(G)$), he doesn't know for sure what is **A**'s true preference. He may ask him, but in general he can't trust **A** to tell him the truth (we will see that it is in A's interest to under-report his preference), nor can he trust **A** to freely stick to any agreement the two might reach.

Notice, for *later discussion*, that if the group is **small**, say 2, there is the possibility - though not the certainty - that the agents might actually *know* each other's preference, and also have enough *trust* in each other's honesty for *asking* and *negotiating* a *free (binding = cooperative) agreement*. But if the group is **large**, say thousands or millions, no one may ever be able to know and ask, let alone trust, everybody else!.

Under the general condition of *ignorance of the others' preferences* and of *trust-impossibility* the *only rational behaviour* of **B** is to act according to the logic of the **Nash best reply function**, BRF. Since he doesn't know **A**'s actual preferences, his 'rational' behaviour is to consider *all possible* ***A****'s preferences*, and then see what would be his best cost share choice for each one of them. **A**'s possible preferences are represented by **A**'s **possible 'demand' curves** for G, drawn in the Figure as the *many downward sloping thin blue lines* to the right and to the left of $D^A_{LS}(G)$. Applying the previous reasoning we find the other **B**'s best choices, such as points **B1** and **B2**. By joining all such points we obtain the **thick upwardsloping dark red line** labeled $D^B_{NS}(G)$, which represents ***B****'s* BRF, in the face of all possible **A**'s preferences for G.

Now we repeat the exercise inverting subjects. We consider what would be **A**'s best cost shares vis à vis all possible 'demand' curves by **B**. The result is the **thick downwardsloping dark blue line** labeled $D^A_{NS}(G)$ representing ***A****'s* BRF. Since such thick lines are the BRFs of **A** and of **B** in this game, their intersection - the point labeled NS corresponding to the pair (G_{NS}, s^*_{NS}) - is the **Nash equilibrium strategy**, while the intersection between the thick lines labeled $D^A_{LS}(G)$ and $D^B_{LS}(G)$ - the point labeled LS corresponding to the pair (G_{LS}, s^*_{LS}) - is the **Lindahl (non-equilibrium) strategy**. The Figure shows in clear fashion that:

1) point NS is **inefficient (non-Pareto)** because the two agents' **indifference curves intersect**.

2) the Figure shows more than just **inefficiency**, it shows that in this framework inefficiency takes the form of **insufficiency**. Because of the way in which the best reply functions are constructed, point NS must **necessarily lie to the left** of point LS. The *inefficient* Nash equilibrium level (amount/quality) of G is necessarily less **less than** the efficient level.

3) point LS is **efficient (Pareto)** because the **indifference curves** are **tangent** to each other.

4) a move from NS to LS improves the **welfare of both agents**: the Pareto strategy LS is not only Pareto, it is - more strongly - **Pareto superior** to the non-Pareto strategy NS.

5) the previous points 1-4 are **universal properties of this Nash-Lindahl result**: they do not depend on the assumption of equal agents (made here for geometric ease), as they **hold for any pair of different players**, whatever their preferences and incomes/endowments. Notice that this applies in **particular to point 4 above**: since at point NS the two indifference curves intersect, there clearly must exist some **rightward path**, starting from NS and yielding a *continuous* improvement of the welfare of *both players*, up to the point where it reaches LS, where by definition such joint improvements are no longer possible.

6) with our assumptions of equal agents and equal incomes/endowments the NS share s^*_{NS} coincides with the LS share s^*_{LS}. But this is strictly dependent on those assumptions. In the more general case of non-equal agents and endowments the two shares would be different, as shown in F6.5bis.

For completeness we have drawn in the Figure also the 'equivalent' of the **contract line** in the **Edgeworth diagram**: it is the *thin black curve* going through LS and joining all points of *indifference tangency* of the two subjects.

Notice that in this 'shared interests' = public goods context this curve 'resembles' the Edgeworth contract line because of the indifference tangency and the associated property of Pareto efficiency, but it *cannot act as a proper contract line* because public goods being *non-rival* cannot be exchanged. All we can say is that at points on the curve lying above or below LS the corresponding shares s_A and $s_B = 1 - s_A$ are such that one agent woul want more of G and one would want less, but since G is *not distributed* among them, its amount can still be efficient even with *non-unanimity cost shares*. In terms of F6.7 below the government could provide the efficient amount G^* even if the cost shares were distributed equally according to $\frac{1}{2}SMC(G)$ and and the two subjects wanted the different amounts $G^B < G^A$.

A different graphical proof

In F6.6 we construct the same result of F6.5 using a *different graphical technique*.

The Figure is a simplified version of F2.7 in the **Cesi-Gorini** paper **2014**, last Section.

1) The downward sloping **thick blue line** is the *actual* 'demand' curve (drawn here not in terms of the cost share s but in terms of the marginal benefit or tax-price, as previously explained) of the two identical individuals.

2) The red curve SMC is the social marginal cost, assumed for simplicity to be constant.

3) We start by assuming the cost to be divided into equal shares, and that the two agents are at point LS, where *the individual marginal benefit is equal to the individual marginal cost*. Starting with agent **B** we ask: is he happy to stay where he is? We easily see that he is not. If **B** decides to reduce his share, **A**'s share must increase by the same amount by which **B**'s share decreases. It follows that the amount of G that **A** is prepared to pay for decreases, say to the point **g** - to the left of LS - where his MB is equal to his new and *higher* MC. Thus when **B** reduces his share he gets a *benefit* because his MC curve shifts downwards, but at the same time he suffers a *loss* because G decreases. At first the balance for **B** is a net gain, because the *horizontal distance* **f-g** = **k-o** (the *marginal gain*) is larger than the *vertical distance* **f-k** = **g-o** (the *marginal loss*). Therefore he will move away from LS. However, if he keeps reducing his share eventually a situation is reached where the balance '*marginal gain minus marginal loss*' becomes *zero*, and if he goes further the balance becomes *negative*. In the diagram the situation of *zero balance* is reached at point **B0** where the marginal cost curve for **B** is $MC_{B0}(G)$ and that for **A** is $MC_{A0}(G)$, because at this point the *horizontal distance* **r-B0** (the *marginal gain*) is equal to the *vertical distance* **e-B0** (the *marginal loss*). Thus **B**'s share in the SMC which brings his marginal cost down to $MC_{B0}(G)$ is the one that *maximizes his net gain* if the **A**'s demand' were the *blue curve* **c-q**. But if **A**'s 'demand' were the *lower thin blue curve* going through NS the new situation of *zero balance* (corresponding to the previous one) would be at point NS where the

horizontal distance **j-**NS (the *marginal gain*) is equal to the vertical distance **d-**NS (the *marginal loss*). Points **B0** and NS in this Figure have the same meaning of points **B0** and NS in F6.5. If the reader repeats the argument for as many other 'demand' curves of **A** as he likes, he would find *other points of zero balance*, and the line joining them would turn out to be **the thick black broken line c-NS-B0-B1-q**. In the Figure this black broken line represents the BRF (best reply function) of **B** corresponding to the BRF of **B** in F6.5. Since we are dealing with *identical agents*, the Nash best reply function of **A** is the same as that of **B**, and therefore we can't think here in terms of their '*intersection*'. The **Nash equilibrium** is obtained where the *uniform* BRF intersects the *uniform* individual *MC* curve, because that is where *every agent has no incentive to choose a different cost share*.

Notice that in the highly schematic picture of F6.6 point **B1**, which lies on the horizontal axis vertically below point **h**, is supposed to correspond to point **B1** in F6.5. But the fact that it lies on the horizontal axis doesn't match the position of **B1** in F6.5: in F6.6 point **B1** indicates that **B**'s best choice is to *not contribute at all*, while in F6.5 his best choice is a *positive contribution.* Though the two Figures F6.5 and F6.6 are intended to represent the same facts, they are *drawn freely*, with no attempt to produce a perfect graphical match!

6 GENERALIZING THE NASH-LINDAHL THEOREM

The different graphical technique of F6.6 relative to F6.5 allows to **generalize the result of F6.5 to any number of players**. Figure F6.6bis shows, under the assumption of equal agents (which facilitates the proof *with no loss of generality*) that, as the number of agents increases, **the NS level of G keeps decreasing, and tends to zero when the number tends to infinity**. We may regard this result as a *rigorous counterpart* of what we have anticipated in the *informal* explanation of F6.2. In large communities, under full freedom of action the incentive to *cooperate* for the satisfaction of *essential public shared interests* vanishes, and only the emergence of an **organized political community**, endowed with a **coercive power** called the **State** or the **government**,can prevent it from collapsing.

Figure F6.6bis is the first step towards the generalization of the result, the subsequent steps being an automatic extension to higher numbers. It shows that when the **number of agents** increases from **two** to **three**, the *thick black broken line* representing the BRF of each agent changes its shape in the following simple way: point **B1** shifts leftwards along the G-axis, and the line **c-B1** becomes *steeper*. While with *only two agents* (F6.6) the *black line* **c-B1** is steeper than the *blue line* **c-q** in such a way that at every point the *horizontal distance* from the vertical axis is *equal* to the *vertical distance* to the blue line (see points NS, **B0**, etc.), with *three agents* at every point the *vertical distance* to the blue line is the **double** of the *horizontal distance* to the vertical axis. The reason is that when agent **C** reduces his cost share by a certain amount, the increase in the *individual cost share* of the other two agents **A** and **B**, required to cover G's cost, is no longer equal to that same amount, but to **one half** of that amount (remember we are assuming for simplicity that cost shares are divided equally among all contributing agents). It follows that **C** reaches his point of *zero balance* when the *horizontal distance* is equal to *half the vertical distance.*

The point is that as the number of people in the community increases, **any given reduction in the contribution by a single agent** translates into an **equal** increase in the **total contribution,** but this **must then be divided among all other** $N-1$ **agents**, so that the impact of a single agent's cost share reduction on each other agent's cost share increase *becomes ever weaker*, until it becomes *negligible*, and when *the group size becomes very large* the single agent has an *incentive to bring his contribution down towards zero.* In F6.6bis as the number of people increases the *black line* **c-B1** rotates towards the vertical axis and point NS keeps *shifting leftwards* towards it. In a large community without authority everyone would have an incentive to contribute little or nothing to the community, unless he could expect everyone else to contribute spontaneously his due share, an expectation which in a large community without authority is of course totally implausible.

7 COOPERATION FAILURE: INDIVIDUAL FREEDOM PERSONAL HONESTY, PUBLIC COERCION

[versione provisoria nov 2022]

We now resume the concept of **strategic coordination failure** introduced in the previous Chapters 4, *Elementary games* & 5, *The Cesi-Gorini game*, as well as in the present Chapter, in order to probe more deeply into the nature and causes of this particular type of *coordination failure.*

Why don't people move from Nash to Lindahl? .

First stage: binding agreements, preference revelation, power We proceed in stages. The *first stage* is to go back to F6.5 and repeat **the critical question** raised in the *Prisoners' game* of Chapter 4: why don't the two agents **A** and **B** move from NS to LS, where *both would improve their welfare*? We already know the answer. Under conditions of **full freedom of action** (no authority) each agent cannot be confident that once they have moved to the LS strategy the other agent would stick to the agreement, because if he moved to a different strategy he would get a higher payoff.

1) **Binding agreements and group size**. Suppose the two negotiate to get to LS. They can do so, but then **B** has no incentive to stay in LS. He has an incentive to move to **B0**, i.e. to reduce his contributing share from $1-s_{NS}$ to $1-s_3$, where G decreases but his welfare (indifference curve) increases, while **A**'s welfare decreases. According to a large **literature on so-called cooperative**

games, the possibility of binding agreements would eliminate this obstacle, but although we are not particularly familiar with that literature we are in principle distrustful of it: **court enforcement of contractual agreements** is expensive, time-consuming, highly uncertain in its outcome. Judicial procedures are everywhere the quintessence of inefficiency, and their outcome is everywhere extremely uncertain and dependent on the partners' financial capacity of employing powerful law firms. In short, court enforcement is most obviously **no substitute for a selfimposed behavioural code of personal honesty**.

Moreover, **contractual arrangements** may be made between a **limited number of partners**, usually two or a few more. They cannot be made between thousands of people (unless people organize themselves in well-organized subgroups, such as workers' and employers' unions in the commercial economy). Therefore we are back to our main contention: **when the group is large**, under conditions of *full freedom of action* the *non-cooperative logic* of the **Nash** BRFs is the only realistic one.

2) **False preference revelation**. **A** knows that if he confessed his true preferences, then **B** would choose something like **B0** increasing **A**'s cost share. Therefore, in order to pay less he could pretend to have the *lower preference* represented by $D^A_{NS}(G)$, because by doing so he might expect **B** to choose something like NS, and he (**A**) would be better-off than in **B0**. Notice that if **A** pretended to have an even lower preference, such as $D^A_3(G)$, then he might expect **B** to choose something like **B3** where he (**A**) would now be *worse-off* than in NS.

3) **Power**. Single **powerful agents**, or subgroups of agents, may have the **power to force weak agents** or subgroups **to accept their preferred choice of contributing shares** and/or **amounts of** G. Here we must stress that both the **Lindahl allocation** of F6.3 and the **Nash allocation** of F6.5 assume a scenario of **powerless individual agents** (or subgroups), exactly as in the competitive 'market' scenario of the **Edgeword box** with a price auctioneer (Chapter 1, *Microtheory of private rent in the private economy: market power*). For instance, if agent **A** had the *power to impose* onto **B** his own choice of *contributing shares*, but *not the power to also fix the amount of* G, then *both the Lindahl and Nash scenarios would vanish.* The new *'power' equilibrium* would be at **A0**, because that is where **A** would reach his *highest attainable indifference curve*, given the **actual preferences** of **B** represented by the *light-red line*. Similarily, if agent **B** had that power the new 'power' equilibrium would be **B0**.

In Chapter 2, *Microtheory of private rent in the public economy: distortion of public choice by private interests*, we did consider situations of a similar type, but with a *significant difference*: we were considering *organized political communities*, already endowed with the *coercive power of government*, and the power under discussion was the power of private agents to distort government choices to their private advantage.

4) **Trust on agreement compliance and preference revelation**. If the two agents could *trust* each other on the matter of *agreement compliance* and

true preference revelation, and if neither of them had any *power* to force his own preferred contributing shares on the other, then they would always choose the Lindahl strategy LS, which would then become the **strategic equilibrium** in place of NS.

As already noticed such a convenient state of things can only be found in **extremely small communities (Ostrom E. 1990)**, like family, close friends, small villages or tribes, and even then not always, as we know all too well from the frequency with which fights erupt also in those 'favourable' circumstances! **Large communities** are inevitably dominated by **depersonalized relationships**. Under conditions of **full freedom of action** the **non-cooperative Nash behaviour** becomes the **only rational one**, even in a fantasy world where all individuals obeyed a **selfimposed moral code of personal honesty**.

Second stage: Legal versus moral honesty Now to a s*econd stage*. We introduce a *cultural hierarchy* concerning the nature of *personal honesty*.

1) **Legal honesty**. This descends from a **selfimposed practical behavioural code of formal legality** (or legality tout court). Such behavioural code has two meanings, a strong and a weak one.

In its **strong meaning** it consists in choosing to **respect the law, including lawful contractual obligations, because it is the law**, i.e. because compliance with the law is the only way to ensure a viable social order for oneself (and by extension for all), and to prevent society to disintegrate into a jungle of violence and exploitation (of the weak by the powerful, the poor by the rich, the servant by the master, the meek by the greedy).

In its **weak meaning** it consists in choosing to respect the law, including lawful contractual obligations, **because of the (high or low) probability of being forced to do so by a court of law**, and/or of being subjected to some kind of social **sanction/punishment** for violating it.

2) **Moral honesty**. This descends from a selfimposed **moral behavioural code** which precedes, and is independent from, formal legality. It is a behavioural code consisting in the **respect of the other fellow human beings, of their freedom-independence and their interests and values as having in principle the same value in society as one's own**. In so far as such respect is a **matter of principle**, descending not from the practical (utilitarian) need of avoiding social collapse, but from one's **view of the world and of life**, it acquires the status of a universal moral value, namely **the recognition of the identity and freedom-independence of the human being as the universal secular moral value par excellence**.

Third stage: the trade-off between individual freedom and social/public coercion Next the *third stage*. We introduce a *hierarchy of social arrangements,* envisaging on one side, say the left, *decreasing* levels of **individual freedom of action,** or **social freedom,** IFA, and on the other, say the right, *increasing* levels of **scope, role and functions of government power,**

associated to **different visions of the state** and **dfferent types of state organization**.

We mention here a *strictly philosophical point* concerning the concept of *individual liberty/freedom*, namely the distinction between **social freedom** and **moral freedom**: 'social' freedom stands for *freedom of action*, while 'moral' freedom stands for *individual consciousness and self-consciousness.* Among the many philosophers who have placed the **idea of freedom** at the centre of their intellectual investigation, the two who in my view have produced the clearest answers are **Isaiah Berlin (1969)** (on the *negative* and *positive* concept of social freedom) and **Benedetto Croce (1943 [1988])** (on the concept of moral freedom).

1) **No state**, NS, or **maximal individual freedom of action**, 100%IFA. Start by assuming a community (group) of people acting under conditions of a maximal IFA. By a maximal IFA we mean a - purely imaginary - social setting where there is **no outside authority** having the power to force individuals to do or not do something. Everyone is free to act as he pleases, including the use of physical or moral violence over others. This social setting may also be labeled Law of the jungle.

This NS = 100%IFA setting is an *abstract benchmark* because, at the level of *individual agents*, in most human societies of the past and present *it doesn't exist.* It may however be a *not-so-abstract benchmark* if we consider the **international community**, whose agents are the single sovereign state-like political entities. The international community has never known, so far, a superior authority, and sovereign states have always negotiated agreements and waged wars against each other precisely because the international community is ruled by the 100%IFA = The law of the jungle.

2) **Night-watchman state,** NWS. Next we descend to the social setting characterized by the immediately inferior IFA level: the existence of a state with the *only* function and power to forbid and sanction all interpersonal uses of **physical** or **moral violence**, over people or things.

3) **Nozick minimal state**, NMS. The next lower level of IFA is the **full Nozick minimal state**: a state whose only scope and power is to enforce not only the no-violence rule, but also the compliance with *any contractual obligation* freely negotiated between agents. This minimal state has been theorized and sustained, with exceptional clarity, by the philosopher and political scientist **Robert Nozick** in his classic work *Anarchy, State, and Utopia* (**1974**).

Nozick's vision of the state is a benchmark. Alternative visions of the role of the state in public and private life with special reference to economics are found, among others, in the wide-ranging dialogue **Buchanan-Musgrave 1999** and in **Stiglitz & others 1989**. The general nature and role of the state in human history and society is dealt with in **Bourdieu's 1992** encyclopaedic treatise. The meaning and role of moral honesty and civic/secular morality in economic life is investigated in **Gorini 2009, 2015, 2018, 2021**, and **Castellucci-Gorini 2014**, on the basis of selected economic and philosophical literature.

[18/11/19]

4) **Ordine sociale liberale-democratico** (laico), **stato socialdemocratico**, con il compito di provvedere al soddisfacimento degli **interessi pubblici**: Croce, Berlin, Aron, Bobbio, Samuelson, Bourdieu, Olson, Judt, Chomsky, Habermas, Stiglitz, Dasgupta e altri

5) Stato-**istituzioni del potere**: Gentile "Aut Cesar aut nihil"

6) **Autocrazia, dittatura, tirannia**: Olson

Strategic versus moral behaviour

The crucial concept lying at the heart of the process of negotiating a contractual agreement concerns the *reciprocal confidence* that, once the agreement has been secured, the negotiating partners will respect it. In the social condition of number **1)** above - the NS = 100%IFA - we can say that in general the reciprocal confidence would be at its lowest level. It may be there (in very small communities), but with no objective support of any kind. In the condition of number **2)** above, the NWS, the reciprocal confidence may be slightly higher, but still with no objective support of any kind. In the condition of number **3)** above, NMS, the reciprocal confidence is definitely higher. Each partner can count that the other will respect negotiated agreements, because of the strong or weak behavioural code of legality.

From the point of view of the possibility of *negotiating binding agreements*, we may deepen our understanding of the concept of *reciprocal trust* by comparing two distinct social conditions. In one social condition, the NMS, with **no selfimposed moral behavioural codes**, people may negotiate agreements because they know that they will be forced by law to abide by them. In the other, the 100%IFA under a *'cultural' state of widespread moral honesty*, people may negotiate agreements because they know that they feel morally bound to keep their word out of a belief in the **moral value of respecting each other**. From the **purely formal** point of view of the possibility of negotiating agreements these two social conditions are **similar**. Indeed, some may believe them to be **objectively identical**. But on the contrary, they are **substantially and objectively different**, because a **legally constrained behaviour** is not the same as a **morally constrained one**. Moral constraint means moral behaviour, and **moral behaviour is incompatible with strategic behaviour**. It means that a person acts in a certain way not because he expects certain choices on the part of others, but because he regards that behaviour as **the right one**. Moral constraint is therefore much **deeper** and much **stronger** than legal constraint, and of course it covers an **enormously larger area of potential agreements**. It is on the basis of this fundamental difference that we distrust the very concept of **cooperative game**. Strategic (= game-based) behaviour is by definition **non-cooperative**. In so far as the literature about cooperative games rests on the the concept of **binding agreements** based on the **legal enforcement of contracts** rather than **civic morality**, it doesn't allow to carry out the thinking about **cooperation** to its full social-political implications. Introducing **moral honesty** means moving away from strategic (game–based) behaviour towards a **social behaviour that is no longer strategic in nature**.

8 ENFORCED COOPERATION. THE POWER TO TAX AND THE PRINCIPLES OF TAX BURDEN DISTRIBUTION: BENEFIT VERSUS ABILITY TO PAY

F6.7 is the graphical representation of **two main concepts of Public economics**: that of a **Lindahl allocation** (point G^*) and that of the *cost share tax burden distribution* according to the **Ability to pay** (such as the black line $\frac{1}{2}SMC(G)$) and the **Benefit** principles (the blue and red lines $s^A SMC(G)$ and $s^B SMC(G)$). The Figure repeats F7.6 of Chapter 7, *Public consumption goods: the Samuelson-Lindahl economy*. The Chapter contains a comprehensive treatment of the topic to which we refer the reader.