



# Beyond Pigou: externalities and civil society in the supply–demand framework

Casey B. Mulligan<sup>1</sup>

Received: 4 April 2023 / Accepted: 20 April 2023 / Published online: 2 May 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

## Abstract

The extent of voluntary cooperation in the presence of externalities is shown as an equilibrium outcome in the supply and demand framework. The analysis uses familiar ingredients to provide a new way of understanding the results of the extensive literature beginning with Buchanan, Coase, Ostrom, Shapley, Telser, Tullock, and Williamson showing that a Pigouvian tax is not the only alternative to independently acting individuals who are coordinated merely through distorted market prices. Voluntary cooperation transforms the character of the costs resulting from externalities and may have a far different incidence than Pigouvian taxes and subsidies do. The paper discusses applications including forest management, volume discounts, residential associations, energy policy, the scope of planning of household activities, and the role of workplaces in preventing infectious disease.

**Keywords** Externalities · Voluntary cooperation · Pigouvian taxes

## 1 Introduction

Meaningful externalities are present in many industries, including but not limited to agriculture, education, energy, healthcare, housing, insurance, innovation, technology, mining, and transportation. Externalities are opportunities for cooperative institutions to enhance outcomes compared to purely individualized action. However, these institutions do not necessarily have to be governmental. Focusing on natural resources, Elinor Ostrom (1990) observed that “neither the state nor the market is uniformly successful in enabling individuals to sustain long-term, productive use of natural resource systems” especially because

---

I appreciate comments from Pete Leeson, Ennio Piano, an anonymous referee, seminar participants at the University of Chicago and Tel Aviv University and financial support from the University of Chicago’s Initiative on Enabling Choice and Competition in Healthcare. This paper is part of a project of adding new chapters to Chicago Price Theory, which contains Kevin M. Murphy’s Economics 301 lectures. Those lectures (such as <https://www.youtube.com/playlist?list=PLp2AOdiHSxGezS07261P6w9akqQpAZ6g->) and further recommendations from Kevin Murphy and Robert Minton strongly influenced the preparation of this paper

---

✉ Casey B. Mulligan  
c-mulligan@uchicago.edu

<sup>1</sup> Department of Economics, University of Chicago, Chicago, IL, USA

effective cooperation “requires reliable information about time and place variables.” Instead, voluntary institutions emerge to facilitate cooperation. Echoing Coase (1937), Buchanan & Tullock (1962) cite the “business firm or enterprise [as] the best single example of an institutional arrangement or device that has as its purpose the internalization of external effects.”

Buchanan & Tullock famously concluded that a useful economic equilibrium concept must be “defined to include the modification of private institutions” as market participants seek to improve upon “purely individualized action.” Although both cooperative and non-cooperative game theory have been deployed for this purpose, these lessons are often forgotten in applied policy analysis where it is presumed that state action is the only viable alternative to independently acting individuals.<sup>1</sup> The purpose of this paper is to facilitate consideration of voluntary cooperation by providing a new analytical framework assembled from the familiar ingredients of supply and demand.

Many previous studies have results regarding voluntary associations, such as the theories of the firm like Coase (1937) and Alchian & Demsetz (1972); theories of local-externality management (“clubs”) such as Buchanan & Tullock (1962) and Buchanan (1965); and cooperative game theory in general (Shapley, 1953; Telser, 1994). Williamson (1975) also models firms as responses to market failures, emphasizing the role of corporate governance structures in determining outcomes.

Ostrom (1990) examines various case studies where voluntary associations successfully managed common-pool resources and indicates factors that made the cooperation successes possible. To name one, she describes forests near Swiss villages that have been sustained for centuries, despite being treated as a communal resource, via communal limits on the harvest season and agreements as to the trees to be cut. While Ostrom and various game theory texts such as Gintis (2009) analyze such outcomes using noncooperative games, this paper offers a price-theoretic alternative to understanding both positive and normative aspects of externalities. Having alternative modeling approaches can be a step toward public policy analysis that more frequently recognizes voluntary cooperation rather than ignoring it.

Section 2 of this paper briefly provides definitions and introduces some basic notation. Optimal Pigouvian taxes are discussed in Sect. 3, where it is shown why the magnitude of redistribution likely exceeds the potential efficiency gains by an order of magnitude. This by itself suggests that voluntary cooperation would be of a fundamentally different character than the usual proposals for state action. Sect. 4 considers Coase bargaining and the cooperation costs that are widely recognized to prevent or at least limit its applicability. My approach is to augment the usual marginal costs of production with cooperation costs, emphasizing that an important driver of such costs is that cooperation itself creates individual incentives to profit from defection—in effect buying low and selling high at the expense of others. This formalization allows the amount of cooperation to be an equilibrium outcome reflecting the tradeoff at the margin between cooperation costs and the benefits of encouraging positive externalities and discouraging negative ones. All of this can be presented in the Marshallian quantity-price diagram. Section 5 examines group size for the special case of public goods. Section 6 concludes. Throughout the paper, real-world examples are provided including forest management, volume discounts, residential associations,

---

<sup>1</sup> For example, equilibrium analysis of infectious disease beginning with Philipson and Posner (1993) through the COVID-19 pandemic (Brodeur, et al., 2021) contrast uncoordinated individual action with public policy solutions. Leeson & Rouanet (2021) is an exception.

energy policy, the scope of planning of household activities, and the role of workplaces in preventing infectious disease.

## 2 Definitions

An externality is a cost or benefit experienced by a third party due to economic activity, not represented in the market price of that activity. Though missing prices may indicate deeper market imperfections like ill-defined property rights, this paper's results are best interpreted from the price perspective. This paper begins by examining excise taxes and subsidies as proposed “solutions” to missing prices, relating the scale of redistribution to the value of efficiency gains. I then use the supply and demand framework to show how markets may respond to externalities with voluntary cooperation among some of the affected parties.

Within this framework, cooperation costs increase marginal cost. The amount of cooperation is an equilibrium outcome reflecting the tradeoff at the margin between cooperation costs and the benefits of encouraging positive externalities and discouraging negative ones. The cooperation often occurs in families, schools, workplaces, clubs, insurance plans, homeowner associations, churches and other institutions of civil society in which membership is voluntary but comes with behavioral restrictions.<sup>2</sup>

Figure 1 illustrates a production externality, modeling the preferences and private production costs in a standard supply–demand framework. Two horizontal marginal costs curves are shown, with one reflecting the private marginal cost, normalized to one, and the other marginal social cost  $1 - \delta > 0$ . A negative externality  $\delta < 0$  is the case shown in Fig. 1 because the marginal social cost is the greater of the two. The relative position of the curves is reversed for a positive externality, as shown later in this paper.

The efficient quantity  $q_e = D(1 - \delta)$  is where the demand curve crosses social marginal cost.<sup>3</sup> Often  $q_e$  is strictly positive, in which case it is efficient to have some production occurring even though every unit of production is affecting third parties. The efficient amount of pollution, for example, is not zero when the marginal value to consumers of the first unit of production exceeds the social marginal cost.

The “unregulated” or “unorganized” quantity  $q_u = D(1)$  is an equilibrium outcome when consumers independently choose the quantity at the price (one) that does not reflect any of the external effect on third parties. At that equilibrium, production imposes total costs of  $-\delta q_u > 0$  on the third parties.

## 3 Efficiency and redistribution with Pigouvian taxes

It is often suggested that either production or consumption be subject to an excise tax at rate  $t > 0$ , sometimes known as a Pigouvian tax. The excise tax would reduce consumption below  $q_u$ . If set at rate  $t = -\delta$ , sometimes known as the optimal Pigouvian rate, the excise tax would (with an important caveat noted below), reduce consumption to the efficient

<sup>2</sup> Civil society is sometimes defined as non-profit non-governmental organizations. I define it more broadly to be all organizations or associations of multiple individuals, each of whom is free to exit the group and forgo whatever benefits the group offers its members and avoid its membership costs.

<sup>3</sup> I neglect income effects on the demand for  $q$ . Defined over  $q$  and all other goods  $x$ , consumer preferences are  $x + \int_0^q D^{-1}(z) dz$ .

level.<sup>4</sup> Consumers make their decision based on the price  $1+t$  inclusive of tax, which by construction reflects the costs  $-\delta > 0$  of production imposed on third parties as well as the private marginal cost (one).

The social value of Pigouvian taxation comes from changing consumption. That value is the average gap between marginal social cost and marginal value on the units that are not consumed because of the tax:

$$\int_{D(1+t)}^{D(1)} [1 - \delta - D^{-1}(q)]dq \approx [D(1+t) - D(1)] \left[ \delta + \frac{t}{2} \right] \quad (1)$$

where  $D(q)$  is the demand curve and  $D^{-1}(q)$  is the marginal value curve, sometimes known as the inverse demand curve. Geometrically, that area is approximately the area of a trapezoid with base equal to the consumption change and heights equal to the marginal cost-value gap at each endpoint. This trapezoid reduces to a triangle at the optimal Pigouvian rate, which is the case shown in Fig. 1 as the area above the demand curve and below social marginal cost.

In contrast, we know from the cost function that the cost to the consumer of Pigouvian taxation is in the interval  $(tD(1+t), tD(1))$ , the lower bound of which is Fig. 1's revenue rectangle. At the middle of this interval is the area of a trapezoid with height  $t$  and base equal to the average of  $D(1+t)$  and  $D(1)$ . The trapezoid representation of consumer cost has height equal to the tax rate, which at the Pigouvian optimum is twice the height of the social benefit triangle. It also has greater width, which depends on the level of consumption rather than the change.

I begin with the case in which the consumers whose demand generates the external costs are distinct from those who experience most of the harm from the industry's production as well as those who receive or control the disbursement of the excise tax revenues. In this case, Pigouvian taxation redistributes much more than it adds social value.<sup>5</sup> The revenue rectangle is a transfer from consumers to those who receive or control the revenue disbursement. The rectangle bounded horizontally by  $q_e$  and  $q_u$  and vertically by the two cost curves represents the benefit to those who experience the harm from the industry's output. The upper right part of that rectangle is a net social benefit while the lower left part is a consumer loss beyond the tax paid that offsets part of the gain to those harmed.

### 3.1 Example: climate change and carbon taxes

Figure 1 illustrates some of the economics of climate change by interpreting its quantity to be the quantity of fossil fuels combusted and emitting carbon dioxide ( $\text{CO}_2$ ), which has been classified as a pollutant by the U.S. Environmental Protection Agency. If the quantity is measured in units of  $\text{CO}_2$ , the parameter  $-\delta > 0$  denotes the "cost of carbon," which is

<sup>4</sup> Excise, income, and consumption taxes used for revenue purposes are not explicitly treated here. One approach would be to put in the baseline all taxes that are for revenue purposes. The baseline equilibrium would thereby have fiscal externalities in that movement of consumer expenditure and factor costs out of the industry (a reduction in  $q$ ) would affect aggregate tax revenue. Under this interpretation, our  $\delta$  denotes the combination of fiscal externalities and the type of externalities emphasized in this chapter. The optimal Pigouvian excise tax would adjust the marginal fiscal externalities so that they exactly offset the other externalities, which in our notation is  $t + \delta = 0$ .

<sup>5</sup> See also Sallee (2019). Buchanan & Tullock (1975) emphasize the different incidence of regulatory versus tax responses to externalities.

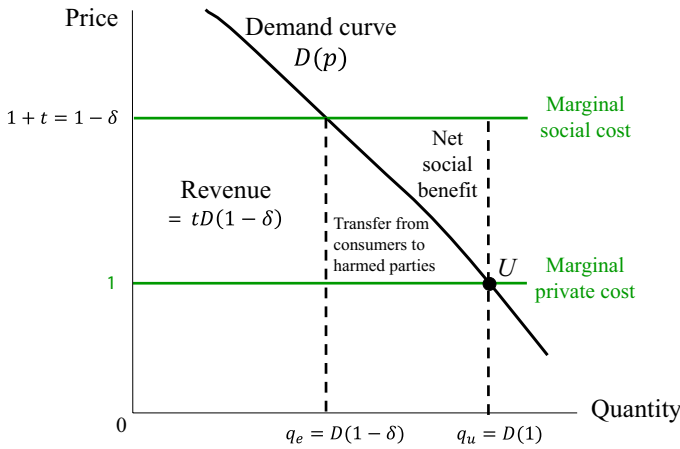


Fig. 1 A negative production externality

the net cost to third parties of each unit of CO<sub>2</sub> emissions such as harm to health and infrastructure from CO<sub>2</sub>-induced climate change (“the cost of manmade climate change”). The area between the two marginal cost curves and up to  $D(1)$  denotes the monetary amount of that cost, relative to what it would be with zero emissions.

Unless the highest-value use of fossil fuels has less value than the cost of carbon, the optimal Pigouvian tax would reduce but not fully eliminate CO<sub>2</sub> emissions. Even if it did, and Pigouvian taxes had no enforcement costs, the net value of implementing the optimal Pigou tax would be less than the cost of the manmade climate change that would occur at the quantity  $q_u$  and less than the harm to consumers from imposing the tax.

### 3.2 Gains from trade illustrated in a Marshallian diagram

Redistribution can have adverse side effects such as unintended behavioral changes in response to tax obligations. Consumers benefit from undermining the tax regime or at least from reducing the rate of taxation even when it is below the Pigouvian optimum. Those who experience the harm benefit from high rates, even those well above the Pigouvian optimum. Those who control or receive the excise revenue benefit from moving the rate in the direction of more revenue, regardless of whether the revenue-maximizing rate is above or below the Pigouvian optimum. With such divergent interests, determining and implementing the optimal Pigouvian rate will prove to be as much a political challenge as an econometric one.<sup>6</sup>

Figure 2 adds consumer indifference curves, shown in red, to Fig. 1. The indifference curves must be increasing and concave in  $q$  everywhere below the demand curve.<sup>7</sup> The

<sup>6</sup> Ostrom (1990, p. 15) notes that the disadvantages of state solutions to (common pool) externalities include “dependen[cy] on the accuracy of information obtained by a distant government official...”. She adds that “it is difficult for a central authority to have sufficient time-and-place information to estimate accurately ... the appropriate fines to induce cooperative behavior.”

<sup>7</sup> Because a consumer with income  $y$ , consuming  $q$ , and paying average price  $p$  consumes  $y-pq$  of all other goods, the price-quantity combinations that yield utility  $U$  satisfy  $p = \frac{y-U + \int_0^q D(z)dz}{q}$ .

same curves are decreasing everywhere above the demand curve. The combined surplus for third parties and the receivers of the revenue associated with any gap between the consumer price and the private marginal cost is  $(\delta + p - 1)q$ . Level curves of this function (“iso-other”) are shown in blue in Fig. 2, with more surplus in the direction of less quantity. Each iso-other curve must also be increasing and concave in  $q$  everywhere below the social marginal cost curve.

The contract curve in this market is a vertical line at the efficient quantity, where the consumer indifference curves are tangent to a corresponding iso-other curve. The contract curve includes the optimal Pigouvian allocation  $E_1$ , although that point leaves consumers much worse off than they are without taxation (allocation  $U$ ). The allocation  $E_2$  also has the efficient quantity while leaving consumers no worse off than at  $U$ . Allocations in the shaded area are each a Pareto improvement over the allocation  $U$  because they are below the consumer indifference curve through  $U$  but above the iso-other curve through  $U$ .  $E_3$  is the efficient allocation  $\left\{ D(1 - \delta), 1 - \left( \frac{D(1)}{D(1-\delta)} - 1 \right) (-\delta) \right\}$  that maximizes consumer surplus subject to the constraint that the others are no worse off.<sup>8</sup>

Aside from the single point  $U$ , the entire shaded area in Fig. 2 involves average prices for consumers that are below the private marginal cost (one). Pareto-improving allocations involve marginal prices above one, but also the others help pay for some of the private cost of production  $q$ .

The harm to consumers of Pigouvian taxes is often acknowledged with a suggestion that the excise revenue be rebated to consumers in a way that is independent of their consumption choices. However, tax rebates are not enough to make consumers whole, because of the excess burden of the tax. Rebated Pigouvian taxes in Fig. 2 are those allocations at price 1 and quantities less than  $q_u$ . They are outside the set of Pareto improving allocations. If the price elasticity of demand were close to zero, the excess burden would also be close to zero and perhaps the rebate scheme comes sufficiently close to a Pareto improvement. However, inelastic demand is also the case in which Pigouvian taxation does little to help the third parties harmed by production.

## 4 Voluntary regulation

### 4.1 Coase bargaining

Unlike the optimal Pigouvian allocation  $E_1$ ,  $E_2$  and the rest of the shaded area (other than  $U$  itself) is off the demand curve. Therefore, allocations off the demand curve are not an equilibrium if consumers are free to purchase any quantity at a single price. Consumers must be induced to purchase less than  $q_u$  by facing a marginal price  $m > 1$  while at the same time paying an average price  $p < 1$ . This could be achieved with a nonlinear pricing scheme or by having consumers agree to and comply with quantity restrictions. Ronald Coase’s famous paper “The Problem of Social Cost” predicted that affected parties would reach such agreements at least if the cost of agreement and enforcement were low enough. I refer to the allocations in the shaded area with quantity  $q_e$  as the “Coase allocations.”

<sup>8</sup> The vertical line at  $q_e$  is analogous to the contract curve in an Edgeworth box, which would have  $q$  and all other goods as the two axes rather than  $q$  and  $p$ . The indifference curves and iso-other curves correspond with indifference curves in the Edgeworth box.

Coase's (1960) solution brings our attention to an abundant phenomenon in human societies: organizations with voluntary membership that impose rules on their members for the good of the overall membership. Much activity occurs in families, schools, workplaces, clubs, insurance plans, homeowner associations, churches and other institutions of civil society. Each of these organizations restricts the choices of its members. Members who are harmed by those restrictions can leave the organization. An apartment building may require tenants to be quiet during nighttime hours, place time limits on the use of common areas, or prohibit pets. In the spirit of Coase, these organizations can be interpreted as voluntary agreements to alleviate negative externalities and encourage positive externalities among the members.

Applied cooperative game theory, especially its equilibrium concept of "the core," emphasizes the role of endogenous voluntary organizations in shaping individual and market behavior (Shapley, 1953; Telser, 1994).<sup>9</sup> Our approach reaches similar conclusions, and some new ones, by extending the price-theoretic framework to include cooperation and especially emphasizing market-level outcomes (Murphy, Snyder & Topel, 2014).

Aside from a propensity to redistribute, another shortcoming of the usual Pigouvian tax recommendations is that they ignore the voluntary regulation that may occur instead of the Pigouvian tax, or simultaneous with it. Take the case of making noise at night in an apartment building. Suppose for the moment that the noise can only be heard by those living in the same building. Each building has rules, enforced with a fine or some other punishment, about noise at night. If the city, which contains a great many apartment buildings, were to impose an excise tax on noise at night, the buildings would still (we predict) maintain their rules for noise at night. Under the tax scheme, a resident caught making noise at night would deliver revenue to the city coffers, which is no compensation for the residents in his building aside from their tiny share in the city's overall population.

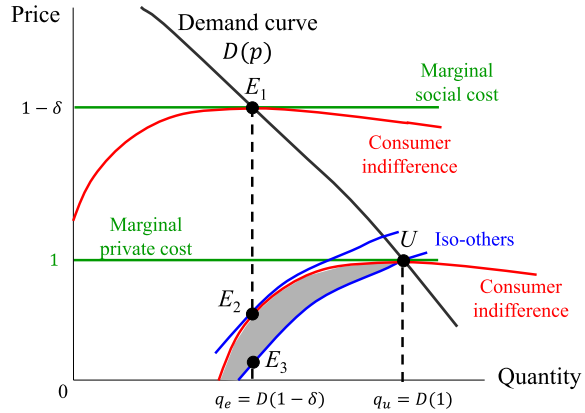
In effect, the excise tax increases both the private cost curve in Fig. 2 as well as the social cost curve, interpreted as the cost to the building, by the amount of the tax. Formally, let's say that the apartment building had set its rules to put the noise level at  $D(1 - \delta)$  absent the city tax. A city tax changes a tenant's marginal cost from 1 to  $1 + t$  and the building's marginal cost from  $1 - \delta$  to  $1 + t - \delta$ . From noise level  $D(1 + t)$  that would be chosen absent building rules, there is still incentive for residents of the building to mutually agree to make less noise. The gains from trade among the tenants are not fully realized until the noise level is reduced to  $D(1 + t - \delta) < q_e$ . Any positive excise tax by the city is overkill.<sup>10</sup>

An alternative case to consider is one in which a tenant's noise can also be heard outside his own building. Whether that creates a role for a city excise tax depends on the extent that buildings negotiate with their neighbors. Absent interbuilding cooperation, a positive Pigouvian rate can increase aggregate surplus, although the optimal rate would account for voluntary restrictions already in place.

<sup>9</sup> The core is typically defined in endowment economies as a feasible allocation for the entire economy that cannot be improved upon by any subset of agents (including a single person) that delivers greater utility to each of its members by removing each member's endowment from the aggregate and reallocating them among themselves. In this way, the core emphasizes competition among groups in which members can choose their group and groups choose their members.

<sup>10</sup> Buchanan & Stubblebine (1962) is an early paper that accounts for Coase bargains under an excise-tax regime.

**Fig. 2** Pareto improving externality management (negative externality, without cooperation costs)



### 4.2 Cooperation costs

Any nontrivial response to an externality involves inducing a marginal value for the consumer that is different from private marginal cost. The greater this gap, the stronger the individual incentive for consumers to undermine the tax or group agreement and the more likely that resources will be needed to monitor and enforce compliance. In practice, demand and costs vary over time and across individuals, which requires either resources to measure and update the tax or agreement parameters or experiencing costs of deviating from the optimal plan, or both. I refer to these costs as cooperation or management costs.

Another classic paper by Ronald Coase (1937) observed that there is an optimal scope of cooperation or “planning.” Hayek (1945) also poses the question of “whether planning is to be done centrally, by one authority for the whole economic system, or is to be divided among many individuals.” Centrally planned countries are known to perform poorly and be plagued by corruption even, or especially, in terms of managing environmental and other externalities. Coase hypothesized that the size of a voluntary organization and the scope of activities it controls expands to the point at which the marginal cooperation cost equals the marginal benefit of pushing members to act in the group interest in ways that they would not when acting independently. Williamson (1975) and Ostrom (1990) also analyze how voluntary organizations are limited by cooperation costs.

Putting excise taxation aside, I formally represent Coase’s idea with a cooperation cost that is a convex function  $c(m - 1)$  of the gap between marginal value and marginal private cost.<sup>11</sup> The marginal cooperation cost has the opposite sign of the gap  $m - 1$ . Under these conditions and regardless of the sign of the externality, the optimal voluntary regulation solves:

$$\max_m \{u(D(m)) - (1 - \delta)D(m) - c(m - 1)\} \tag{2}$$

where  $u(q)$  denotes the area under the marginal value curve  $D^{-1}(q)$  and  $m$  is the marginal value  $u'(q) = D^{-1}(q)$  induced by the cooperation agreement. If the optimum (2) is less than

<sup>11</sup> Ostrom (1990) also emphasizes cooperative institutions “resembling neither the state nor the market” and their maintenance costs, which she formally models as a parameter in a noncooperative game.



the value  $u(D(1)) - (1 - \delta)D(1)$  of not cooperating, then cooperation does not occur and the externality is not voluntarily regulated.

The first-order condition is (3):

$$m = 1 - \delta + \frac{c'(m-1)}{D'(m)} \quad (3)$$

which expands the concept of marginal social cost to include the marginal cooperation cost of increasing quantity. The optimum involves pushing the marginal value in the direction of social marginal cost but not all the way. The Pigouvian optimum  $m = 1 - \delta$  does not optimize (2) because at this point the gap between marginal value and marginal social cost is zero whereas pushing  $m$  toward 1 offers a first-order savings on the cooperation costs.

Although their analyses often ignore enforcement costs, a similar logic is relevant for Pigouvian taxes. Marginally reducing the tax rate from the Pigouvian optimum presumably economizes on enforcement costs that depend on the magnitude of the tax rate.

As in Fig. 2, it is useful to decompose the social surplus into two components: the variable part of consumer surplus, represented by  $u(q) - pq$ , and the surplus of others, represented by  $pq - (1 - \delta)q - c(D^{-1}(q) - 1)$ . The consumer indifference curves are the same as in Fig. 2. At the quantity  $D(1)$ , indifference curves for others have slopes that are independent of cooperation costs, at least under the nudge hypothesis that  $c'(0) = 0$ .<sup>12</sup> For positive externalities, and prices between private and social marginal cost, the slopes must be negative. The efficient quantity is  $D(m)$ , where  $m$  solves (3) and the iso-other and consumer-indifference curves are tangent to each other. At quantities between  $D(1)$  and  $D(m)$ , the iso-other curves slope in the same direction as they do at the endpoints.

Figure 3 depicts the cooperation-cost tradeoffs in the case of a positive externality. The upper horizontal schedule is the private marginal cost curve. The lower one is the  $1 - \delta$  component of the social marginal cost. Because the condition (3) also specifies a cooperation cost component, marginal cooperation costs are included in the figure as the upward-sloping curve that departs from the horizontal component of social marginal cost at  $q_u$ .<sup>13</sup> As drawn, I assume that  $c'(0) = 0$ , indicating that there is zero marginal cooperation cost beginning from the zero-cooperation quantity  $q = D(1)$ . The equilibrium  $m$  and  $q$  are further up the marginal cooperation cost curve, where marginal cooperation costs are strictly positive. In other words, the organization's equilibrium constraints on its members provide a more significant push than a mere nudge, creating non-negligible costs for monitoring and enforcing compliance.

The intersection of demand and marginal social cost in Fig. 3, expressed algebraically as Eq. (3), specifies the equilibrium quantity  $q$  and marginal price  $m$ . The equilibrium average price is somewhere between allocations  $E_2$  and  $E_3$ .

Notice from Fig. 3 how much better off the others are between  $E_2$  and  $E_3$ —or even at  $U$ —than they would be at the Pigouvian optimum  $E_1$ .<sup>14</sup> While the equilibrium without cooperation or taxes has too little production from the perspective of the third parties, at least they benefit from production  $D(1) > 0$  without having to pay anything to

<sup>12</sup> The nudge hypothesis, which is an implication of the standard consumer theory, is that consumers are easily moved a small distance off their demand curve in the quantity dimension (Chapter 5 of Jaffe et al., 2019).

<sup>13</sup> Expressed as a function of quantity, the right-hand-side of (3) is  $1 - \delta + c'(D^{-1}(q) - 1)/D(D^{-1}(q))$ .

<sup>14</sup> The results in this paragraph require  $D(1) > 0$ .

consumers or producers. The Pigouvian subsidy requires third parties to pay the subsidy rate on all production, even the production that would occur without the subsidy.

We see from the condition (3) and Fig. 3 that the model (2) does not go as far as “the Coase Theorem” to say that voluntary cooperation fully eliminates the externality. The equilibrium marginal benefit  $m$  exceeds the sum of the private marginal production cost (one) and the (negative) external cost of production. At the same time, the voluntary equilibrium is significantly different from  $U$ , which is the baseline case often considered for Pigouvian tax analysis. First, even if the cooperation only moves quantity half way from  $D(1)$  to  $D(1 - \delta)$ , it has eliminated essentially three-fourths of the externality.<sup>15</sup> Second, cooperation translates some of the external cost into a management cost. As drawn in Fig. 3, the equilibrium management cost is the area of the larger of the two shaded triangles. The equilibrium externality is the area of the smaller triangle. Third, a state-imposed excise subsidy may push quantity beyond  $D(1 - \delta)$  because a subsidy reduces both the individual’s marginal cost and the group’s. Voluntary cooperation can also complicate the assessment of even the sign of the equilibrium externality, as discussed in the disease example that follows.

### 4.3 Example: managing the forest commons

Consider a forest with villagers living nearby who have a demand for timber. The forest is not anyone’s private property. In terms of Fig. 2, the labor cost of cutting and hauling each tree is a private marginal cost. The social marginal cost also includes the cost of planting and growing the tree (perhaps via forgone logging so that trees reproduce themselves). Absent cooperation, the forest might quickly become depleted, a result known in economics as “the tragedy of commons.”

Ostrom (1990) describes forests near Swiss villages that have been sustained for centuries, despite being treated as a communal resource. The villages governed use of the forest themselves. Individuals were appointed to mark specific trees for harvesting. Harvesting was limited to a specific time of year. Individuals breaking the rules were punished, although in practice significant punishments were rarely needed. In our notation, these restrictions help push the quantity of logging during a year below  $D(1)$  and nearer to  $D(1 - \delta)$ . The costs of meeting, marking, monitoring, and adapting to new conditions are the cooperation costs  $c(m - 1)$ . The cooperation costs also include opportunity costs—perhaps the harvesting labor would be more economically spread through the year but limiting the harvest season has the social benefit of helping to prevent overuse.<sup>16</sup> An excise tax on logging proved to be unnecessary for sustaining these forests.

### 4.4 Example: volume discounts and buyers’ clubs

Monopolies and oligopolies are usually modeled as firms that present consumers with a single price that exceeds marginal cost. As such, the industry resembles Fig. 3, even if monopoly or oligopoly pricing are rarely described as externalities per se. The upper of

<sup>15</sup> Exactly three-fourths is eliminated if demand is linear in the interval  $(1 - \delta, 1)$ . More (less) than three-fourths is eliminated if demand is convex (concave) in that range, respectively.

<sup>16</sup> Wildlife stocks in the United States are commonly managed by, among other things, limiting the duration of the hunting season.

the two horizontal cost curves is the price that consumers would pay if they were charged a single price. Consumers would respond by choosing quantity  $D(1)$ . The lower horizontal curve is marginal cost, which crosses marginal value at  $D(1 - \delta)$ . The consumer indifference curves are as shown in Fig. 3, although the iso-other (iso-profit) curve is slightly different than shown in Chapter 13 of Jaffe et al (2019). Most relevant for this paper is that the allocation  $U$  leaves unrealized gains between consumers and producers. The Pareto improvements involve trades that are to the right of the consumer demand curve.

Monopolies and oligopolies may use advertising to nudge consumers off their demand curves. They may offer consumers volume discounts: the more they buy, the less they pay relative to list price (which we normalize to one). Volume discounts not only reduce the average price below list, but also present the consumer with a marginal price below the average price because an additional purchase reduces the price paid on the inframarginal units too. That configuration of list price (one), average price (somewhere between the prices at  $E_2$  and  $E_3$ ), and marginal price ( $m$ ) is exactly what is shown in Fig. 3.<sup>17</sup>

In sales contracts between pharmaceutical manufacturers and insurance plans, the discount is known as a rebate because the plan pays list price up front followed by a rebate payment from manufacturer to plan based on the plan's achievement of contractual quantity targets.<sup>18</sup> One way that the plans demonstrate to manufacturers that consumers have been moved off their demand curve is to place the product favorably in the plan's benefit structure. Especially, the plan covers a greater share of the drug cost (patients would pay a lesser share when they pick up their prescription) for manufacturers that pay more rebate rather than less.

Commentators have noticed that the contractual sales targets present insurance plans with a marginal price that is less than even the average price, pejoratively describing the situation as a "rebate trap." From the perspective of Fig. 3, the low marginal price is evidence that the sales quantity has been moved away from  $D(1)$  in the direction of  $D(1 - \delta)$ . Indeed, studies have observed that, in pharmaceutical markets where insurance plans are major purchasers, the quantity of prescription sales is essentially the same for monopoly drugs as for drugs with many sellers, even though monopoly list and average prices are quite high (Lakdawalla & Philipson, 2012; Lakdawalla & Sood, 2013). That is, the studies have not been able to detect underutilization as a result of pharmaceutical monopolies, although monopoly may create noticeable management costs (Mulligan, 2022).

Buyers can also take action to realize gains from trade. As discussed in Chapter 13 of Jaffe et al (2019), they sometimes form buyers' clubs that restrict their members' choices in order to obtain manufacturer discounts. Health insurance plans are buyers' clubs in healthcare markets.<sup>19</sup> The resource and opportunity costs of managing and participating in buyers' clubs are represented in my model as  $c(m - 1)$ .

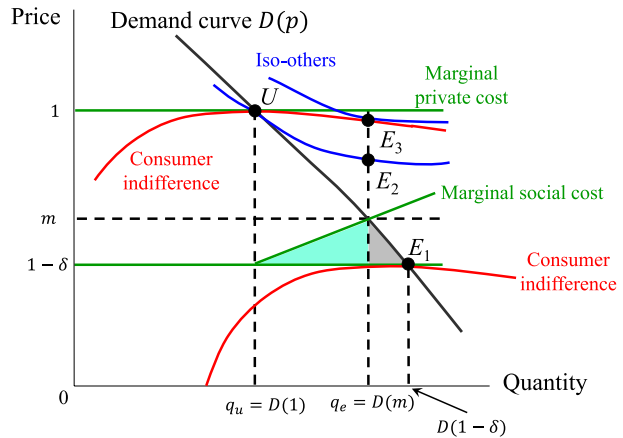
The example of volume discounts clearly points to management costs. A seller offering one customer a volume discount must prevent that customer from purchasing on behalf of other customers. In the case of prescription drugs, sellers are assisted by government

<sup>17</sup> Notice how the Fig. 3 has four distinct values that might be called prices: the list price (one), the average price (expenditure per unit quantity, sometimes known as "net price"), the marginal price (equal to social marginal cost inclusive of cooperation costs), and the part of social marginal cost without cooperation ( $1 - \delta$ ).

<sup>18</sup> Manufacturer rebates also occur in other industries.

<sup>19</sup> Indeed, the insurance plans themselves cooperate via clubs, known as group purchasing organizations and pharmacy benefit managers, in making purchases from healthcare providers (Burns 2022). Insurance plans even cooperate in small clubs for the purpose of purchasing the services of larger clubs.

**Fig. 3** Equilibrium cooperation (positive externality, with cooperation costs)



regulations, often strictly enforced, against the resale of prescriptions among consumers. Resale is more much difficult to prevent in other industries such as petroleum.

#### 4.5 Example: price discrimination by OPEC

Some industries such as petroleum production have management costs  $c(m-1)$  that are too prohibitive to support an equilibrium allocation much different from  $U$ . The Organization of Petroleum Exporting Countries (OPEC) is a group of countries that collectively attempt, with varying degrees of success, to restrain oil production in order to increase world oil prices. They do not offer significant volume discounts in part because their customers are scattered around the world, where it is difficult to monitor resale by the high-volume customers who would be getting a discounted marginal price. Moreover, some of the OPEC members have high marginal costs that leave little surplus to be gained from volume discounting.

Petroleum-product sales to consumers in Saudi Arabia are exceptions to both of these conditions. Saudi Arabia is the leading OPEC country and often takes the largest production cuts. That is, they are the country likely to have the greatest gap between world oil price and marginal cost of production. The Saudi government can also better monitor and prevent international resale by its own citizens than it can its international customers.<sup>20</sup> Therefore Fig. 3 suggests that the marginal price of petroleum products paid by consumers in Saudi Arabia should be substantially less than what consumers pay elsewhere in the world, even without regards to any political benefits of having cheap domestic oil. In fact, Saudi Arabia has the world's largest, or close to largest, domestic subsidies for petroleum products.<sup>21</sup>

<sup>20</sup> Petroleum exports by domestic consumers can also, in principle, occur indirectly by exporting manufactured products that are intensive in petroleum.

<sup>21</sup> Alyousef & Stevens (2011) and Fattouh & El-Katiri (2013).

## 5 Cooperation and group size

The special case of externality known as a non-rival public good highlights the possible role of group size in determining the gap between individual and group-level marginal cost. It is a symmetric version of the positive production externality in which the entire population benefits from the production. The magnitude of an individual's benefit is independent of his own behavior, including whether he takes any action to contribute to the production of the good. In our notation, the quantity  $q$  of the public good in a population of  $N$  individuals is the sum of each individual's contribution:

$$q = \sum_{i=1}^N x_i \quad (4)$$

Absent cooperation, an individual  $i$  taking as given the others' contributions faces a price of one. Increasing his contribution by  $dx$  increases  $q$  by  $dx$ , exactly as it does along the upper horizontal line in Fig. 3. The group, by contrast, faces a price of no greater than  $1/N$  because it can finance  $dq=1$  by requiring every member to increment his contribution by  $dx=1/N$ . This equal contribution case is represented in Fig. 3 by setting  $\delta=1-1/N$ . The Pigouvian optimum in this economy is allocation  $E_1$ , which equates the average individual's willingness to pay to the social marginal cost  $1/N$ .

Although the entire population benefits from the production of a public good, it may be supplied by a subset  $M < N$  of the population. The optimal supply from the perspective of the group of  $M$  is the point where their average willingness to pay equals  $1/M$ . Furthermore, the group supplying the public good may attempt to prevent nonmembers from benefitting in order to attract and retain group members despite the contribution costs that come with membership. This famously occurred with the Chicago Cubs when the Major League Baseball team erected screens to obstruct the rooftop views of its Wrigley Field from neighboring buildings not owned by the team.<sup>22</sup> The team ultimately reached an agreement with the rooftop owners to remove the screens in exchange for sharing rooftop revenue. In this and many other cases, at least partial exclusion is part of the technology (that is, the public good is "local" or "impure") because the quality of the view of a baseball game decreases with the distance from the game.

Allocation  $E_1$  is equivalently described as the point where the sum of willingness to pay across individuals is equal to a marginal cost of one.<sup>23</sup> This is known as the Samuelson condition for a collective consumption good. The larger the group, the greater is the number of individuals in the sum. That is, under cooperation the willingness to pay for a public good increases proportionally with the size of the group.

At the same time, even small-scale cooperation can go a long way toward reducing the opportunity costs of falling short of the efficient amount of the public good. Suppose that only two cooperate with negligible cooperation costs, and that all  $N-2$  others who benefit from the public good contribute nothing. Their cooperation results in aggregate contributions of  $D(1/2)$ , which with linear demand is more than three-fourths of the way to the aggregate-efficient point  $D(1/N)$  in terms of aggregate surplus and more than half way in the quantity dimension (Lindsay & Dougan, 2013).

<sup>22</sup> Associated Press (2002); Strazzante (2004).

<sup>23</sup> Samuelson (1954).

Per capita management costs may also fall with group size, over a range, to the extent that managing cooperation involves some fixed costs. Eventually per-capita management costs increase with group size as monitoring large, heterogenous, and geographically dispersed groups becomes increasingly challenging. That is, there is an equilibrium group size for cooperation for many of the same reasons that there is an equilibrium firm size (Viner, 1932). If the entire population of  $N$  wants to be part of a group whose optimal size is  $M$ , the equilibrium involves  $N/M$  distinct groups. Having multiple groups also has competitive benefits that help ensure that groups ultimately serve the interests of members because each can choose to exit one group and join another.

The fixed element of cooperation costs suggests that a group formed to manage one externality may take on the management of secondary externalities that affect the group's members (Mulligan & Shleifer, 2005; Murphy & Shleifer, 2004). For example, employers that are large for business reasons are more likely to provide their employees health insurance, retirement pensions, and other fringe benefits (SHRM, 2020). We suggest that a fixed element of cooperation costs helps account for this pattern. If the fixed elements of cooperation costs were significant enough, the equilibrium group size could be the entire population ( $M=N$ ), which might be interpreted as a government policy response to an externality rather than cooperation among members of smaller groups.<sup>24</sup>

## 5.1 Example: families and the supply of household goods

The framework in this paper provides a simple model of effects of policy on cooperation between mothers and fathers for the purpose of investing in children. Here,  $q$  denotes the quantity of human capital investments, with price normalized to one. Each parent perceives a value of  $u(q)$ , regardless of how the investment is financed between the parents. That is,  $q$  is a household-level public good. If each parent acted independently, taking as given the resources the other parent puts toward investment, the equilibrium  $q$  would be  $q_u=D(1)$  (the same notation as above). If instead each expected the other to make matching contributions, the equilibrium  $q$  would be  $D(1/2)$ . With cooperation costs of the form (2), the equilibrium  $q$  would be between these two values. Figure 3 shows why much, but not all (see the right-shaded triangle), of the externality would be eliminated. It also shows that cooperation costs are incurred (the left-shaded triangle) in the process of eliminating that amount of externality.

Now consider perturbing Fig. 3's equilibrium by introducing a  $q$  subsidy at rate  $s$ , financed from outside the family. The two quantities bracketing the equilibrium quantity would now be  $D(1-s)$  and  $D\left(\frac{1-s}{2}\right)$ . The equilibrium cooperation condition (5) is the analog to (3):

$$m = \frac{1-s}{2} + \frac{c'(m - (1-s);s)}{D'(m)} \quad (5)$$

where the notation in (5) allows for the public policy to shift the cooperation cost function in addition to shifting the two horizontal cost schedules in Fig. 3. For example, the subsidy may be administered in such a way to discourage mother and father from living together and thereby increase their costs of cooperating. If the demand for the household-level

<sup>24</sup> See also Buchanan & Tullock (1962, Chapter 5).

public good is not too elastic, the policy would increase total coordination costs even if it also increases  $q$ . Other cooperative investments, not eligible for the subsidy, may also be discouraged.

Not only is cooperation relevant for household investments, but it also affects the size and structure of households. The prevalence of nuclear families in the modern world suggests that cooperation on many household activities may best occur with just a handful of individuals, but larger household models have also been tried. Voluntary communes have been formed, and some still persist, that not only collectively manage household activities such as childcare and food preparation but also employment and capital investment. The kibbutz system in Israel has persisted for several generations, although eventually the kibbutzim narrowed the scope of their household planning (Abramitzky, 2018). These are questions of “who plans” of the same character as those posed by Coase.

## 5.2 Example: civil society and the spread of infectious disease

Infectious diseases often spread directly from person to person or indirectly through environmental intermediaries (“disease vectors”). From an economic point of view, individual actions taken to avoid infection or prevent transmission are contributions to a local public good. If the cost of infection were significant, as they often were with COVID-19, the economic approach predicts that voluntary organizations with sufficiently low cooperation costs would meaningfully encourage their members to contribute to slowing the spread of the disease. The Samuelson condition suggests that the per capita demand for actions that slow the spread would be greater when individuals participate in large groups than when they act alone or in small groups such as families. To an order of magnitude, the increase in willingness to pay that comes from cooperating in a group is proportional to the size of the group.

The supply of disease prevention could also be greater in a group setting due to economies of scale. Flu transmission can be prevented by administering flu shots, which workplaces have sponsored enough that they are one of the top locations where workers receive the shot. Large employers have been more likely to offer flu shots on site than smaller employers do. While workplaces were organized for business purposes rather than flu-shot administration, my approach suggests that the organizational capital acquired for the first purpose may also prove to be useful for the second purpose.

During the COVID-19 pandemic, large employers took some remarkable steps to reduce transmission that could not conceivably be organized by individuals or households on their own. To name one, the University of Illinois developed its own rapid COVID-19 testing operation for its campuses (Deliso & Bhatt, 2020).<sup>25</sup>

Because individuals travel or are members of multiple organizations, an infected member of an organization can spread it beyond the membership. If an individual’s prevention actions taken as part of complying with her member obligations are not sufficiently reflected in the price of activities taken outside the group, then organizations that encourage prevention have a positive externality on the wider population. In other words, supply and demand factors by themselves suggest that participation in a large organization has a

---

<sup>25</sup> Law prohibited the university from selling testing services in the marketplace until authorized by the U.S. Food and Drug Administration, which was not obtained (and then only on an emergency basis) until February 24, 2021.



positive externality, especially when that organization is suited to encourage cooperation among its members.

Large organizations create more opportunities for healthy members of the group to potentially be infected by a sick member, public health officials said during the pandemic. All else constant, this suggests that the disease would spread more slowly if the population refrained from in person participation in large organizations like schools and workplaces. From this epidemiological perspective, in person participation would seem to have a negative externality during a pandemic. It is therefore an empirical question whether the supply and demand factors driving more prevention in large groups is more than offset by the epidemiological factor of more contacts. This empirical work is still ongoing, but some early results suggest that in person schools and workplaces did not noticeably accelerate the spread of COVID-19.<sup>26</sup> Policies that encouraged healthy people to stay away from school and work may have been discouraging a positive health externality.

## 6 Conclusions

The Marshallian supply–demand equilibrium framework is extended to allow for voluntary cooperation in the presence of externalities. This price theoretic approach offers a different perspective on the obstacles to cooperation by emphasizing the equilibrium gap between marginal price and private marginal cost. Such a gap must be created by any state or voluntary cooperative scheme that internalizes an externality. At the same time, the gap presents individuals with incentives to undermine cooperation that are increasingly costly for a group to supervise and counteract. The approach also shows how voluntary cooperation transforms the character of the costs resulting from externalities and may have a far different incidence than Pigouvian taxes and subsidies do.

As such, this paper’s framework might be described as a novel presentation of an old idea that was previously analyzed with noncooperative game theory and cooperative game theory. As an alternative analytical approach, it has the potential to enrich public policy discussions by broadening the scope of policies in which voluntary cooperation is acknowledged rather than overlooked. Furthermore, price theory has advantages over game theory when it comes to predicting the consequences for cooperation and other market outcomes of external shocks related to tastes, the structure of costs, and public policies. On the other hand, this paper’s price theoretic approach provides little practical advice as to the strategic considerations involved in designing cooperative institutions, which are important in practice and emphasized in game theory.

## References

- Abramitzky, R. (2018). *The mystery of the Kibbutz: Egalitarian principles in a capitalist world*. Stanford University Press.
- Alchian, A. A., & Demsetz, H. (1972). Production, information costs, and economic organization. *The American Economic Review*, 62, 777–795.

---

<sup>26</sup> Mulligan (2021). Even in economics, few studies have so far raised the possibility that schools, workplaces, and other institutions of civil society might be more effective at alleviating local externalities such as infectious disease. Leeson & Rouanet (2021) is one exception.



- Alyousef, Y., & Stevens, P. (2011). The cost of domestic energy prices to Saudi Arabia. *Energy Policy*, 39, 6900–6905.
- Associated Press. (2002). Cubs, rooftop fans at odds again. *ESPN*. <http://a.espncdn.com/mlb/news/2002/0405/1363787.html>.
- Brodeur, A., Gray, D., Islam, A., & Bhuiyan, S. (2021). A literature review of the economics of COVID-19. *Journal of Economic Surveys*, 35, 1007–1044.
- Buchanan, J. M. (1965). An economic theory of clubs. *Economica*, 32, 1–14.
- Buchanan, J. M., & Stubblebine, W. C. (1962). Externality. *Economica*, 29, 371–384.
- Buchanan, J. M., & Tullock, G. (1962). *The calculus of consent*. University of Michigan Press.
- Buchanan, J. M., & Tullock, G. (1975). Polluters' profits and political response: Direct controls versus taxes. *The American Economic Review*, 65, 139–147.
- Burns, L. R. (2022). *The healthcare value chain: Demystifying the role of GPOs and PBMs*. Springer.
- Coase, R. H. (1937). The theory of the firm. *Economica*, 4, 386–405.
- Coase, R. H. (1960). The problem of social cost. *Journal of Law and Economics*, 3, 1.
- Deliso, M., Bhatt, J. (2020). Inside University of Illinois' massive COVID-19 testing operation. *ABC News*. <https://abcnews.go.com/US/inside-university-illinois-massive-covid-19-testing-operation/story?id=72686799>.
- Fattouh, B., & El-Katiri, L. (2013). Energy subsidies in the middle East and North Africa. *Energy Strategy Reviews*, 2, 108–115.
- Gintis, H. (2009). *Game theory evolving: A problem-centered introduction to modeling strategic behavior*. Princeton University Press.
- Hayek, F. A. (1945). The use of knowledge in society. *The American Economic Review*, 35(4), 519–530.
- Jaffe, S., Minton, R., Mulligan, C. B., & Murphy, K. M. (2019). *Chicago price theory*. Princeton University Press.
- Lakdawalla, D., & Philipson, T. (2012). Does intellectual property restrict output? An analysis of pharmaceutical markets. *The Journal of Law and Economics*, 55, 151–187.
- Lakdawalla, D., & Sood, N. (2013). Health insurance as a two-part pricing contract. *Journal of Public Economics*, 102, 1–12.
- Leeson, P. T., & Rouanet, L. (2021). Externality and COVID-19. *Southern Economic Journal*, 87, 1107–1118.
- Lindsay, C. M., & Dougan, W. R. (2013). Efficiency in the provision of pure public goods by private citizens. *Public Choice*, 156, 31–43.
- Mulligan, C. B., & Shleifer, A. (2005). The extent of the market and the supply of regulation. *Quarterly Journal of Economics*, 120(4), 1445–1473.
- Mulligan, C. B. (2021). *The backward art of slowing the spread? Congregation efficiencies during COVID-19*. NBER working paper, no. 28737.
- Mulligan, C. B. (2022). *The value of pharmacy benefit management*. NBER working paper, no. 30231.
- Murphy, K. M., Snyder, E. A., & Topel, R. H. (2014). Competitive discounts and antitrust policy. In *The Oxford handbook of international antitrust economics* (Vol. 2, pp. 89–119).
- Murphy, K. M., & Shleifer, A. (2004). Persuasion in politics. *The American Economic Review*, 94, 435–439.
- Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*. Cambridge University Press.
- Philipson, T. J., & Posner, R. A. (1993). *Private choices and public health: An economic interpretation of the AIDS epidemic*. Harvard University Press.
- Sallee, J. M. (2019). *Pigou creates losers: On the implausibility of achieving Pareto improvements from efficiency-enhancing policies*. NBER working paper, no. 25831.
- Samuelson, P. A. (1954). The pure theory of public expenditure. *The Review of Economics and Statistics*, 36, 387–389.
- Shapley, L. S. (1953). A value for n-person games. In *Contributions to the theory of games* (Vol. 2, pp. 307–317).
- SHRM. (2020). *2020 Employee benefits*. <https://www.shrm.org/hr-today/trends-and-forecasting/research-and-surveys/pages/benefits-report.aspx>.
- Strazzante, S. (2004). Cubs, rooftop owners make their deal official. *Chicago Tribune*. <https://www.chicagotribune.com/news/ct-xpm-2004-01-30-0401310016-story.html>.
- Telser, L. G. (1994). The usefulness of core theory in economics. *Journal of Economic Perspectives*, 8, 151–164.
- Viner, J. (1932). Cost curves and supply curves. *Zeitschrift Für Nationalökonomie*, 3(1), 23–46.
- Williamson, O. (1975). *Markets and hierarchies: Analysis and antitrust implications*. Free Press.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.