

RESEARCH ARTICLE

Incentives and taxation in Rawls's institutional theory of justice

Maxime Desmarais-Tremblay 

BETA, Université de Lorraine, France

Email: Maxime.Desmarais-Tremblay@univ-lorraine.fr

(Received 4 August 2024; revised 15 February 2026; accepted 16 February 2026)

Abstract

John Rawls proposed a theory of justice for the basic structure of society. Surprisingly, his suggestions for tax institutions were not well articulated. Rawls's principles of justice do not prescribe a unique set of tax recommendations, but his remarks on tax matters reflect his vision of society as a cooperative venture in which everyone must work. This paper makes two contributions. First, it offers a chronological, systematic, and contextual analysis of what Rawls wrote on taxation. Rawls's comments on taxation reveal his lifelong concern for preserving market incentives and his rejection of ability-to-pay as a principle of taxation. Second, the paper argues that some of Rawls's tax proposals belong to nonideal theory because they depend on a conception of individuals in tension with the conception of moral persons developed in his theory.

Keywords: incentives; institutions; ideal theory; John Rawls; nonideal theory; optimal tax theory; Richard A. Musgrave; taxation

'Economic man,' is a feat of the scientific imagination, and is not intended as a competent expression of fact. . . . But, as happens in such cases, having once been accepted and assimilated as real, though perhaps not as actual, it becomes an effective constituent in the inquirer's habits of thought, and goes to shape his knowledge of facts. It comes to serve as a norm of substantiality or legitimacy; and facts in some degree fall under its constraint, as is exemplified by many allegations regarding the 'tendency' of things. (Veblen, 1899: 422)

Introduction

When looking for a comprehensive standpoint to provide general guidance on the right tax system, John Rawls's *A Theory of Justice*, with its focus on the 'basic structure of society', is an obvious place to start. By basic structure, Rawls meant the institutions that 'distribute fundamental rights and duties and determine the division of advantages from social cooperation'. (Rawls, 1971: §2, 7). His *Theory* proposed 'standards by which to assess economic arrangements and policies' (Rawls, 1971: §41, 259). As Jon Elster (1992: 187) once remarked, 'To implement a theory of justice, we need to know what taxes, transfers, and subsidies it implies; and sometimes we do not need to know much more'. While the tax system is certainly part of the basic structure of society to which Rawls's *Theory* could apply, he never provided a systematic characterisation of a just tax system (O'Neill and Orr, 2018; O'Neill and

Williamson, 2015b; Sugin, 2004). Taxation poses an unresolved problem for Rawls's ideal theory of justice. In the midst of a discussion on whether progressive or proportional taxation is better to strike a balance between concerns of justice and efficiency, Rawls tells his readers that 'these are questions of political judgement and not part of a theory of justice' (Rawls, 1971, §43: 279). Over the years, Rawls kept coming back to the same 'rough and intuitive' arguments in support of his 'highly tentative' suggestions concerning tax institutions (Rawls, 2001: 135–136). Despite his 'political turn' in the 1980s, he did not change his position concerning tax matters.¹

This paper makes two contributions. First, it offers a chronological, systematic, and contextual analysis of what Rawls wrote on taxation. Many commentators did not feel constrained by what Rawls himself proposed and argued that Rawls provided justifications for limiting the level of taxation (Buchanan, 1984), while others instead argued that a Rawlsian tax system should include a progressive income and wealth tax (Chandler, 2023; Turmel, 2024). As I show in the next section, Rawls's comments on taxation reveal his lifelong concern for preserving market incentives and his rejection of ability-to-pay as a principle of taxation. Getting the history right matters because an incorrect genealogical use of Rawls's authority can serve to justify tax policies that should call for a fuller assessment of the institutional context. Beliefs about the history of ideas shape public policy debates.

Second, the paper shows that Rawls's comments on taxation belong to nonideal theory. To do so, in the third section, I revisit the criticisms put to Rawls by Richard A. Musgrave and G.A. Cohen in light of the distinction between ideal and nonideal theory. In contemporary debates, the distinction was introduced by Rawls and has since been used to discuss the relationship between political philosophy and concrete policy problems.² Some of Rawls's tentative tax proposals belong to nonideal theory because they rely implicitly on a conception of individuals that neglects the role of institutions in shaping their preferences. In particular, Rawls rejects progressive income tax because of a concern for incentives, which might not arise if individuals behave as moral persons as we would expect them to in his ideal theory of justice. Rawls followed Rousseau in taking 'people as they are (by the laws of nature)' (Rawls, 1999: 13). Against this naturalist stance, I take an institutionalist position with which Rawls should agree, arguing that labour supply response to taxation is endogenous to the economic culture and legal institutions in place. In other words, the paper contributes to institutional economics by highlighting that incentive-based tax arguments are not part of an idealised scheme but contingent upon the system of rules and the values and habits of thought held by individuals.

A genealogy of Rawls's comments on taxation

Before a theory of justice

Surveying Rawls's intellectual trajectory is essential to understanding his social vision from which his comments on taxation derive. John Rawls grew up in an Episcopalian family on the East Coast of the United States. Although his combat experience in the Pacific during WWII would lead him to doubt his faith, in 1942, as a young undergraduate at Princeton, he had been wrestling with the *Meaning of Faith and Sin*. In his undergraduate thesis, Rawls argued that separate but equal persons recognised themselves as forming a community of faith (Bok, 2017). Faith manifests itself by 'integration in the community', which leads to Christian love, an opposing force to egotism (Rawls, 2009: 243, 245). At the time, Rawls was suspicious of economists who offered, like Hobbes, a vision of society as a contract based on egoism (Rawls, 2009: 217). But, as his faith declined, Rawls had to find another explanation for what brought people together. He turned to developmental psychology and to economics. Moral sentiments formed the basis of a moral sense, a condition for individuals to accept the principles of a

¹On Rawls's political turn, see Rawls (1993: xvii–xxxii) and Weithman (2010).

²Ideal theory provides guidance for systems in which individuals respect the rules of the game, while non-ideal theory deals with systems where individuals do not always respect the rules, transitional systems, or systems in which policy is restricted to second-best (or worse) instruments. See Simmons (2010); Schmitz (2011); Valentini (2012); Stemplowska (2017); Gajević Sayegh (2017). The distinction has not yet been applied to Rawls's remarks on taxation.

just voluntary association. Rawls also adopted the model of the social contract in the 1960s to conceptualise a social union between individuals who have different conceptions of the good life (Gališanka, 2019).

In the meantime, after finishing his PhD at Princeton, Rawls engaged extensively with economic ideas. In the early 1950s, he read widely in economic theory from Léon Walras to Paul Samuelson, John Hicks, John von Neumann and Oskar Morgenstern, and Frank Knight, as well as contemporary works in welfare economics such as William Baumol, whose Princeton seminar he attended.³ From Knight and from game theory, Rawls borrowed the idea that society was like a game, with cooperation and competition mechanisms. Individuals could occupy different positions, provided that the rules of the game were fair. Echoing the minimax strategy discussed by Luce and Raiffa (1957), Rawls suggested that self-interested individuals would choose rules as if they were ‘designing a practice in which his enemy were to assign him his place’ (Rawls, 1958: 171). By the end of the 1950s, Rawls had assembled the structure of a theory of justice as fairness between rational and mutually self-interested persons (Forrester, 2019: 20). Provided equal liberties for all, inequalities would be acceptable if they worked ‘as incentives to draw out better efforts’ (Rawls, 1958: 173).

In such an ideal society, reasonable individuals would recognise the benefits of cooperation and respect the ‘duty of fair play’ and pay their taxes rather than act as ‘free riders’ (Rawls, 1958: 180). Between two tax policies, even if one finds the first tax policy closer to his moral principles, if the second one receives majority support under a just constitution, one has an obligation to abide by it (Rawls, 1964). These remarks about tax compliance appear to be the only ones Rawls published on tax matters before the end of the 1960s.

In 1967, Rawls sketched the institutional implications of his principles of justice for a society conceived as a ‘co-operative venture for mutual advantage’. In an economy like that of the United States, markets allocated most goods, but the government still had an important role to play in maintaining the competitiveness of markets and the full employment of resources, as well as to check large inequalities of wealth, guarantee a social minimum, and provide equality of opportunity (Rawls, 1967: 69). Rawls argued that such a mixed economy could satisfy his difference principle if remaining inequalities would ‘work out to the advantage of the most unfortunate representative man’ (ibid.: 67).⁴ To conceptualise the various functions of government, Rawls drew from the main reference in the field, Richard Musgrave’s *Theory of Public Finance* (1959). Musgrave rationalised the public budget into three branches: the allocation branch, the distribution branch, and the stabilisation branch. The allocation branch would be responsible for providing social/public goods to correct inefficiencies in allocation. Through taxes and transfers, the distribution branch would redistribute income and wealth to match the ‘widely accepted mores’ in the society as elicited by the political process (Musgrave, 1959: 18). Finally, the stabilisation branch would compensate swings in prices and employment by maintaining a high level of resource utilisation and stabilising the value of money. Each branch was assumed to work under the assumption that the other two were fulfilling their role. For instance, a given public good could be provided to all and paid for by taxes proportional to income if the income elasticity of demand for the good were unitary, under the assumption that the distribution branch had already secured the socially acceptable level of distribution of income. Importantly, each branch had its own sub-budget, collecting its taxes according to its objective. The fiscal balance of the stabilisation branch would fluctuate over the business cycle, but it would be distributionally neutral. Musgrave insisted that the branches were interdependent, and each could affect the development of the economy. For instance, ‘desired changes in the state of distribution may increase the total cake to be divided, or

³Rawls’s engagement with contemporary economic ideas continued at least until the mid-1970s. The secondary literature on economists and Rawls is vast. See, for instance, D’Aspremont (1984), Little (2014), Hawi (2016; 2024), O’Neill (2017), Coker (2021), Jackson and Stemplowska (2021), Igersheim (2022), and Walraevens (2024).

⁴Compare with F.A. Hayek’s normative use of a Smithian argument to justify inequalities in a ‘progressive society’: ‘Even the poorest today owe their relative material well-being to the results of past inequality’ (Hayek, 1960: 98). Rawls (1967) cites the *Constitution of Liberty* elsewhere in the paper but dropped that reference in *A Theory of Justice*. We can speculate that, thinking more systematically on constitutional questions at the turn of the 1970s, Rawls distanced himself from Hayek’s ideas.

they may decrease it. If there is a conflict, the case is not necessarily in favour of the greater income and against the preferred state of distribution' (Musgrave, 1959: 20). Moreover, some public goods were provided by the allocation branch, but they served mostly redistributive purposes. Transfers in kind, like housing for the poor or school meals, satisfied socially important 'merit wants' and were not meant to correct a technical market failure caused by non-excludability, like national defence (Desmarais-Tremblay, 2017). Merit goods posed a problem for the separation between the allocation and the distributional branches that might explain why Rawls either misunderstood Musgrave's framework or consciously decided to change it for his own purpose.

Rawls envisaged four branches: the allocation branch, the stabilisation branch, the transfer branch, and the distribution branch. The allocation branch was meant to 'prevent the formation of unreasonable market power' and restore market efficiency by imposing taxes and subsidies so that social benefits match the social costs. Rawls thus combined insights from the 'regulative' function of taxes theorised by American institutionalists in the interwar period with the more recent neoclassical theory of market failure (Commons, 1934, 820; Marciano and Medema, 2015). He didn't expand on the second branch, but it seems to fulfil a similar role as Musgrave's, that is, maintaining 'reasonably full employment' (Rawls, 1967: 70). Rawls suggested that a 'transfer branch' would be responsible for providing a 'social minimum' according to the criteria of 'need'. A fourth 'distribution branch' was meant first to secure a 'just distribution of income and wealth over time' and second to raise revenue to cover the costs of public goods and make transfer payments (ibid.: 71). By proposing two parts to the branch, Rawls departed from Musgrave's framework in which each branch collected its own revenues. In Musgrave's *theory*, provision for the social minimum, conceptualised as merit goods, was the responsibility of the allocation branch, which also collected its own revenues to fund those alongside other social goods – this was not a responsibility of the distribution branch.

In terms of instruments, Rawls was not very specific. He suggested that the social minimum could be attained by family allowances, unemployment benefits, or a negative income tax (Rawls, 1967: 69). The first part of the distribution branch could achieve its goal by taxing inheritance and gifts. The rationale, which Rawls developed further in *A Theory of Justice* (1971), echoed republican views such as Rousseau's: large inequalities of wealth would erode equality between citizens. Interwar economists like Henry Simons (1934: 13, 26), Knight (1935: 291), and Commons (1934: 87, 831) also defended progressive income taxation to prevent or break up the concentration of wealth that would develop under laissez-faire capitalism. Yet, in the postwar period, this was no longer the rationale economists used to justify the progressive income tax system in the United States. In the age of mass taxation, progressive corporate and individual income taxes were needed to fund the substantial level of public expenditure without putting too high a burden on the poorest, who had little financial ability to contribute to this collective sacrifice (Brownlee, 2000; Desmarais-Tremblay and Johnson, 2026).

The most surprising suggestion made by Rawls concerned the second part of the distribution branch. He argued that 'proportional expenditure taxes' could be part of an 'ideally just arrangement' to fund the costs of public goods (Rawls, 1967: 71; Fried, 1999: 185).⁵ Rawls offered three arguments in support of his proposal. The most convincing one, but the least noticed, relied on his division between the various branches. If inheritances and gifts were highly and progressively taxed to maintain equal liberty and equality of opportunity by the first part of the distribution branch, intergenerational inequality would be minimised, and the overall level of inequality would be much lower (than in any existing society). Thus, the need for progressive taxes on income would be reduced.

Rawls also argued that consumption was a preferable base to income because it aligned better with 'common-sense precepts of justice' (Rawls, 1967: 72). On those grounds, reference was made to the controversial monograph of British economist Nicholas Kaldor, *An Expenditure Tax* (1955). Consumption as a basis made sense under the benefit principle, according to which taxation is a payment for the benefits of the social union. Kaldor attributed this view to Hobbes in the epigraph of

⁵A proportional tax with a personal exemption, as advocated by Rawls and John Stuart Mill (1848: 808) before him, would result in a (mildly) progressive distribution of the tax burden.

his book. Having adopted a contractualist stance, Rawls paraphrased Hobbes by claiming that ‘a levy according to how much a man takes out of the common store of goods and not according to how much he contributes’ was a commonsensical precept of justice (Rawls, 1967: 72; Turmel, 2024). The metaphor of the ‘common store of goods’ probably references Hobbes’ often-used expression of ‘common stock’. More importantly, Hobbes claimed that to pay for the protection of one’s life provided by the state, equity called for a tax burden proportional to the ‘riches’ consumed or ‘used’: ‘For what reason is there, that he which laboureth much, and sparing the fruits of his labour, consumeth little should be more charged, than he that living idly, getteth little, and spendeth all he gets; seeing the one hath no more protection from the Common-wealth, then the other?’ (Hobbes, 1651: 239). Like a burning flame, in the early modern period, consumption was associated with waste and destruction (Trentmann, 2017). Consumption as a tax basis makes sense if we enlarge the understanding of the benefits of the social contract to encompass everything produced by society as a cooperative venture for mutual advantage (Turmel, 2024). In this sense, taxation would be proportional to the overall benefits enjoyed as a member of the social union. We can see why Rawls would support Hobbes’ view that taxation should not penalise those who contribute more through their work but rather should weigh more heavily on those who consume more, on the idle ones.⁶ It reflects his normative commitment to a vision of social cooperation based on work.

Rawls further argued that an expenditure tax would ‘interfere less with incentives’ than an income tax. It’s on this argument that Rawls referenced Kaldor’s proposal for an expenditure tax, which is technically equivalent to a tax on all income minus savings. Such an idea was proposed by the US Treasury in 1942, but since it was too radical a departure from the existing system, it was forcefully rejected by all members of the Senate Finance Committee (Kaldor, 1955: 16; Paul, 1954). For American liberals, there was a serious political risk that a proposed expenditure tax would not be as progressive as the current income tax. In Britain, Kaldor was particularly concerned with the rich men who financed their luxury lifestyle through dissaving without paying their fair share of the tax effort (Kaldor, 1955: 14). Echoing Hobbes, he argued that ‘an Expenditure base would tax people according to the amount which they take out of the common pool, and not according to what they put into it’ (ibid.: 53). His case rested mostly on the inequity of the narrow base of the existing income tax, but the efficiency argument loomed large. For Kaldor, incentives mattered. At the end of the day, any ultimate criteria of ‘fairness’ must take into account the social consequences of individual behaviour. Neoclassical economists distinguish between the income effect of a tax change and the substitution effect on the agent’s choice between working for money and enjoying leisure. A decline in wages (or a rise in tax) would render the agent poorer (the income effect), which could encourage them to want to work more, while it would also decrease the opportunity cost of leisure (the substitution effect), incentivising them to work less. Although he acknowledged that the extent of the ‘disincentive effects’ was an empirical question on which economists knew very little, Kaldor argued that an increase in a progressive income tax (where the marginal rate was higher than the average rate) would probably lead to a larger substitution effect and thus a decline in the labour supply (ibid.: 130–131). In contrast, ‘a tax assessed on expenditure does not discriminate against either saving or risk bearing; and it alleviates, even if it does not remove, the disincentive effects of progressive taxation on work’. (ibid.: 14). For Musgrave, who was like many of his American peers unconvinced by the idea of an expenditure tax, Kaldor’s plea was motivated by a desire to restore Britain’s international competitiveness through increased capital formation, a goal that a shift of the tax base from income to consumption could help attain by encouraging private saving (Musgrave, 1957).

⁶Likewise, Commons argued that ‘the man who gets his wealth by mere rise in site values should pay proportionately higher taxes than the one who gets his wealth by industry or agriculture. In the one case he extracts wealth from the commonwealth without adding to it. In the other case he contributes directly to an increase in both private wealth and commonwealth’. As Rawls, Commons was concerned with the ‘economic results’ of taxation. From the perspective of the ‘police power of taxation’, the important question is ‘What will be the best inducements to individuals to increase the commonwealth by increasing their own wealth?’ (Commons, 1934: 819).

Rawls was at pains to demonstrate that the institutional system of the mixed economy could be made fairer by satisfying his principles of justice and doing so ‘with the smallest loss in efficiency’ (Rawls, 1967: 72). But the status of the tax recommendations he made was ambiguous. He claimed that the proportional expenditure tax was ‘a question of political judgment’, yet he also argued that it was ‘part of an idealised scheme which we are describing’ (ibid.).

A theory of justice

When *A Theory of Justice* appeared in 1971, its ideas were already familiar to many philosophers and economists (Forrester, 2019: 104). Rawls aimed to offer a social-economic doctrine that would match the scope of that of David Hume, Adam Smith, Jeremy Bentham, John Stuart Mill, and Henry Sidgwick (Rawls, 1971: §5, vii). Besides constituting a comprehensive social-economic vision, utilitarianism also provided an answer to any kind of ethical and political question. But it rested on a flawed conception of morality, which Rawls tried to supplant with his *Theory*. The old welfare economics relying on interpersonal utility comparisons created a problem for 20th-century economists, one that Rawls’s theory avoided by assuming that comparisons were made on objective standards of ‘primary goods’.⁷ A procedural theory of distributive justice, as he proposed, didn’t presuppose the ‘utilitarian’s standard assumptions about individual utilities’ (Rawls, 1971: §43, 280). Rawls recognised the usefulness of the new welfare Pareto principle in comparing the efficiency of different institutional arrangements, but he was aware that it only offered a partial ranking, an indeterminacy that the difference principle was meant to solve (§13, 71, 75). But Rawls did not treat justice merely as an allocative question. First, the element of ‘pure procedural justice’ introduced by the principle of equality of opportunity meant that his theory didn’t have to ‘keep track of the endless variety of circumstances and the changing relative positions of particular persons’ (ibid.: 87). To the extent that a fair procedure was respected, one didn’t need to judge the fairness of any particular result.⁸ Second, justice as fairness was not merely an allocative conception because it recognised that goods had to be produced before they would be distributed (ibid.: §14, 88). Hence, it was important for Rawls to secure the cooperation of all members of society.

In terms of the ‘background institutions’, Rawls (1971) repeated almost word-for-word his 4-branch framework from the 1967 paper. The separation of his distribution branch into two parts received a new argument: the first part, checking intergenerational inequalities through progressive inheritance and gift taxation and adjustments to property rights to ‘correct the distribution of wealth and to prevent concentrations of power detrimental to the fair value of political liberty and fair equality of opportunity’, was meant to implement the first principle of justice (Rawls, 1971, §43, 277). The second part of the distributive branch would collect proportional expenditure taxes to fund public goods, to fund the transfer branch, to secure equality of opportunity, and to realise the difference principle (ibid.: 278). In support of the first part, Rawls added a reference to British economist James Meade’s short book on *Efficiency, Equality, and the Ownership of Property* (1964). Meade was discussing state involvement in the economy, in particular, the redistribution of property.⁹ Meade (1964: 53) cautioned his readers that using progressive income tax to redistribute from the rich to the poor ‘may have adverse effects upon incentives to work, enterprise, and save’, a concern which Rawls took on board. Rawls didn’t

⁷See Rawls (1971: §11, §13, §46). Rawls later gave a stabilised list of primary goods: basic liberties; freedom of movement and choice of occupation; powers and prerogatives of offices and positions of authority and responsibility; income and wealth; the social bases of self-respect. To apply the difference principle, Rawls assumed a ‘simplest form’ which boiled down to income and wealth (Rawls, 1982: 162). On the construction of the index, see Ferretti (2022).

⁸In contrast to nonideal comparative theories like that of Amartya Sen (2009).

⁹Rawls’s (1971: 78; 2001: 135 ff.) adaptation of Meade’s alternative social model of ‘property-owning democracy’ has attracted a lot of attention over the last two decades, especially by left-leaning thinkers who see it as a realistic but progressive alternative to capitalism (see O’Neill and Williamson, 2012; O’Neill, 2020). Ironically, the idea of ‘property-owning democracy’ was first employed by British Conservatives in the interwar period, then again by Conservatives in the postwar period before gaining political traction with the rise of neoliberalism in the 1980s (Jackson, 2012).

assume men to be selfish, but they were self-interested: 'They are conceived as not taking an interest in one another's interests' (Rawls, 1971: §3, 13). This meant that his theory of justice couldn't rely much on altruistic motivations, even if it was hoped that the principles of justice would sustain reciprocity in a well-ordered society (Rawls, 1971: §43, 281). When in doubt about the level of redistribution needed to maximise the position of the worst-off member of society, Rawls suggested erring on the side of efficiency: 'A society should try to avoid the region where the marginal contributions of those better off are negative, since, other things equal, this seems a greater fault than falling short of the best scheme when these contributions are positive' (ibid.).

Rawls entertained the idea that his principles would lead to 'a definite system of government activities and legal definitions of property together with a schedule of taxes'. Under these assumptions, the distribution of income and wealth would be just, and nothing would prevent citizens from providing themselves with extra public goods on a voluntary and consensual basis. Such collective actions would be undertaken by a fifth branch, the exchange branch (Rawls, 1971: §43, 282). The reference here was Knut Wicksell's (1896) principle of just taxation and especially James Buchanan's interpretation of it. To the extent that a package of public expenditure and the taxes to fund it received unanimous consent, the action of this 'exchange branch' would respect Pareto efficiency. In this idealised setting that assumes no discrepancies in bargaining power, nor any income effect, the fifth branch would follow the market logic, and citizens would not be coerced to pay for goods they wouldn't perceive a direct benefit from. Rawls further argued that the criterion for the exchange branch was the benefit principle and not his principles of justice (Rawls, 1971: 283). On this point, two comments are in order.

First, although the other branches are regulated by the two principles of justice, they are justified by the social contract thought experiment, which was historically tied to a comprehensive reading of the benefit principle (Musgrave, 1985: 16). In terms of the opposition between the benefit approach and the ability-to-pay approach introduced by Seligman (1894) and refined by Musgrave (1959), Rawls was much more sympathetic to the benefit principle. Certainly, his approach was more foundational than that of public finance theorists. For Rawls, every institution should respect the principles of justice, which could easily require a reassignment of property rights. In that respect, traditional tax principles 'have no independent force, however appropriate they may be in certain delimited cases' (Rawls, 1971: 280).¹⁰ Yet, it is striking how Rawls rejected the ability-to-pay principle, which had become the implicit rationale for progressive taxation among US tax experts since the turn of the twentieth century (Mehrotra, 2013).

Second, since the fifth branch was 'only a trading arrangement', representatives deciding on the public goods to provide themselves collectively did not need to abstract from their economic interests. Rawls famously argued that the principles of justice would be selected by individuals behind a veil of ignorance, where they would be unaware of their natural talent, their particular social position, and their conception of the good life (Rawls, 1971: §24). This original position constituted an objective moral standpoint to judge principles of justice; otherwise, 'if a man knew that he was wealthy, he might find it rational to advance the principle that various taxes for welfare measures be counted unjust; if he knew that he was poor, he would most likely propose the contrary principle' (Rawls, 1971: §4, 18). Inspired by Arrow's (1951: 90) rationalisation of the process of collective decision-making in the US, Rawls proposed a four-stage sequence: first, the principles are decided in the original position; second, a constitutional convention is set up. At that stage, participants are assumed to be ignorant about their own particular position, but they are knowledgeable about the general facts of their society, its history, its level of development, etc. (Rawls, 1971: §31, 197). The first principle of justice guides deliberations at this second stage. The second principle is implemented at the legislative stage, where social and economic policies are decided behind a thinner veil (ibid.: 199). Judges and administrators enter at the fourth stage and apply and interpret rules in full knowledge of the particular facts. Rawls claimed that tax matters were decided at the legislative stage, except for the decisions taken by the 'fifth branch' that

¹⁰This is the central thesis of Murphy and Nagel (2002).

are taken outside of the four-stage sequence (ibid.: §43, 284). This framework is not without its ambiguity, notably since the establishment of the income tax in the United States required the ratification of the Sixteenth Amendment to the Constitution in 1913, as Rawls (1993: 239) himself later remarked.

The status of tax institutions in Rawls's *Theory* showed the same hesitation as in the 1967 paper. He repeated that 'these are questions of political judgment and not part of a theory of justice' while claiming that a proportional expenditure tax was part of an 'ideal scheme for a well-ordered society' (Rawls, 1971: §43, 279). Rawls reiterated his proposal for a proportional tax in later writing, but he did not return to the branches of government, which led O'Neill and Williamson (2015a) to question whether this functional approach to the institutions of the basic structure retained its importance for Rawls in the later part of his career.

After a theory of justice

Rawls responded with interest to economists' attention to his *Theory of Justice*, even if they mostly focused on some very specific points and interpreted them through their own lens. Rawls (1967; 1971) himself had suggested that his difference principle was analogous to the maximin strategy of game theory applied to the context of choice under uncertainty. Economists seized this idea and understood Rawls's difference principle as a social welfare function with maximum risk aversion, according to which the government would maximise the *utility* of the worst-off agent in society. Such a social welfare function offered a tractable alternative to the utilitarian social welfare function in the emerging literature on optimal taxation (Walraevens, 2023). British economist James Mirrlees (1971) calculated the optimal rate of income taxation for a utilitarian social welfare planner faced with utility-maximising agents who were distinguished only by their ability (to convert hours worked into wages). Each agent derived utility from consuming goods and enjoying leisure (time not working). Taxing labour resulted in an efficiency loss due to the substitution effect. At the collective level, a higher level of redistribution required a higher tax rate, generating a bigger efficiency loss. Thus, Mirrlees captured better than anyone else before the now-famous tradeoff between equity and efficiency in tax theory (Desmarais-Tremblay *et al.*, 2023).

With his enthusiasm for contractarian approaches to institutional justice, James Buchanan admired Rawls's intellectual enterprise.¹¹ He invited Rawls to participate in a session on 'redistribution theory' at the American Economic Association (AEA) annual meeting in 1973.¹² Rawls spoke about distributional equity and the maximin criterion. In a draft of his presentation, he remarked that before optimal income taxation, the traditional utilitarian view of Edgeworth and Pigou based on equalising the sacrifice from taxation did not 'tak[e] into account in any systematic way the effects on incentive of taxes on income'.¹³ Even if he couldn't always follow the maths, Rawls was interested in how economists explored the implications of his maximin principle. In correspondence with Edmund Phelps, he remarked, 'How [the maximin criterion] applies to economic questions like taxation is not a matter of mere application. One is testing the viability of the conception of distributive justice itself. . . . The kind of exploration you present is necessary if we are to determine whether the criterion is really reasonable' (cited by Phelps, 1973: 337). In the published version of his AEA paper, Rawls remarked that 'the application of the maximin criterion to optimal income taxation is, then, perfectly in order, since an income tax is part of the basic structure' (Rawls, 1974b: 142). Part of that statement is surprising because, as we saw, the income tax does *not* feature prominently in his *theory*, but also because the economists' take on the difference principle was doubly utilitarian. It conceptualised individual choice by a utility function and collective choice with an individualistic social welfare function, disregarding

¹¹Although he was critical of certain elements of the theory. See Buchanan (1972) and Colin-Jaeger *et al.* (2022).

¹²Buchanan to Rawls, 2 January 1973. Rawls to Buchanan, 6 February 1973. Folder 7: TJ Correspondence, 1973, Box 19, Papers of John Rawls, Harvard University Archives.

¹³'Some remarks about distributive equity', undated, Box 24, Folder 12, Papers of John Rawls.

Rawls's rejection of utilitarianism and his insistence on grounding interpersonal comparisons on an index of primary goods. But Rawls shared the economists' overarching vision of social cooperation as a productive endeavour. The maximin criterion recognises that social cooperation 'can be made more productive for everyone by allowing certain differences to develop'.¹⁴ Anthony Atkinson's (1973) discussion of the impact of varying the elasticity of the optimal tax schedule led Rawls to realise explicitly that his maximin relied on a 'supply of work and effort schedule'.¹⁵ Yet, as Arrow remarked, Rawls was 'inexplicit about the incentive effects', which explained why he could not derive concrete tax proposals (Arrow, 1973: 259; see also D'Aspremont, 1984; Kolm, 2003: 30).

To the extent that one is concerned solely with the negative effect of income taxation on labour supply, the logical solution would be lump sum taxes, that is, taxes that do not impact behaviour and so create no substitution effect and thus no efficiency loss. That is the case of poll taxes, or head taxes, which affect everyone equally based on their existence, as well as what came to be known as endowment taxation, a direct tax on one's natural ability to earn. The obvious problem of such a lump sum tax is that natural abilities are not directly observable (Atkinson and Stiglitz, 1976). Despite his concern for the efficiency of taxation, Rawls rejected lump sum taxes:

Now the maximin criterion would conform to the precept 'from each according to his abilities, to each according to his needs' if society were to impose a head tax on natural assets.¹⁶ In this way, income inequality could be greatly reduced if not eliminated. Of course, there are enormous practical difficulties in such a scheme; ability may be impossible to measure and individuals would have every incentive to conceal their talents. But another difficulty is the interference with liberty; greater natural talents are not a collective asset in the sense that society should compel those who have them to put them to work for the less favoured. This would be a drastic infringement upon freedom. But society can say that the better endowed may improve their situations only on terms that help others (Rawls 1974b, 145).

Rawls's first argument against lump sum taxes was probably a response to one of Arrow's criticisms: as Rawls assumes that 'individuals are supposed to act justly, at least in certain contexts', would they not have an obligation to reveal their natural ability truthfully to earn income (Arrow, 1973: 260)? In *Justice as Fairness*, Rawls pressed his point further: since talents must be trained and are conditioned by social attitudes, 'a usable measure of native endowments' might not even be possible, 'even in theory' (Rawls, 2001: 158). Certainly, in the real world, the practical difficulties are indeed substantial, although proxies could be found (see Mankiw and Weinzierl, 2010).

Rawls's second argument against taxing natural abilities deserves more scrutiny. In *A Theory of Justice*, Rawls famously attacked the idea of desert. He argued that individuals with higher talents do not deserve a higher income since their talents are the result of the natural lottery. Cognitive abilities, physical strength, and one's willingness to work hard are matters of luck: they depend on the genetic lottery and the chances of being born into a good family that encourages children to work hard and develop their talents (Rawls, 1971: §17, §48). Only within a just society can talented people have reasonable expectations to receive a higher wage, and not because they morally deserve it, but because it is to the benefit of all, and, in particular, of the poorest. Rawls regarded the 'distribution of natural talents as a common asset' (ibid.: 101). In his trenchant libertarian critique, Robert Nozick (1974: 229 note) argued that his Harvard colleague's formulation suggested he would support a 'head tax on assets and abilities'. For Nozick, Rawls wanted to 'harness' people's abilities to work for the common good. As

¹⁴Among the public rules, the division of labour is at the heart of the 'scheme of cooperation' that assigns roles and positions, with varying wages and salaries to encourage greater returns for the benefit of all (Rawls 2001: 63).

¹⁵Some remarks about distributive equity', op. cit. Rawls also remarked that maximin models of optimal taxation such as Atkinson's did not lead to marginal rates of income tax above 45%. However, he noted that it might be higher if the model 'took into account [the] effect of inequality on the equal liberties, particularly political liberty' ('Comments on Atkinson', undated, Box 24, Folder 12, Papers of John Rawls). See also Hawi (2024, 598).

¹⁶Rawls referred to head taxes, but he seemed to be talking about natural endowments, or direct ability tax.

he put it very clearly: ‘The notion of liberty needs elaboration which is to exclude a head tax and yet allow the other taxation schemes’ (ibid.). Rawls responded to Nozick that ‘redistributing’ natural assets was ‘incompatible with the integrity of the person’. What the difference principle suggested was to ‘arrange the system of entitlements that encourages and rewards productive efforts’, but ‘we have a right to our natural abilities’, argued Rawls (1975: 263; see also Pogge, 1989: 68 ff.).

Invoking the priority of liberty over the difference principle, Rawls argued that taxing natural endowments constituted an unacceptable violation of the individual’s freedom to choose their occupation. Indeed, in *Justice as Fairness*, Rawls argued once more that the difference principle should not ‘penalise the more able for being fortunately endowed’. Such a tax scheme might result in people ‘not be[ing] able to afford to enter low-paying, though worthy, vocations and occupations’ (Rawls, 2001: 158). However, one could argue that while an endowment tax would make such ‘worthy occupations’ less appealing to the most talented, market signals already discourage young people from taking them in the first place. Likewise, under a differentiated university fee regime, those who build higher debt for their more expensive degree are incentivised to take up the highest-paying jobs to recoup their investment. Thus, it seems that Rawls did not satisfactorily respond to Nozick’s objection: the difference between the incentive effect of a tax on endowments and that of an income tax, or differentiated costs of education, is a matter of degree.¹⁷ One can build examples where an income tax (or a consumption tax) would influence the occupational choice of an individual more than a tax on endowment tax: ‘which tax is more intrusive is empirically contingent’ (Olson, 2010: 255). It depends on the whole set of institutions in place, including the labour market, educational system, etc.

In sum, Rawls was sympathetic to the optimal taxation models based on the maximin at the heart of which lay an individual labour-leisure tradeoff, but he forcefully rejected endowment taxation proposals that would have avoided the efficiency cost of taxing labour income. Rawls still held to his vague positions on taxation in the restatement of his theory published at the end of his life: ‘It is possible that there need be no progressive income taxation at all’ (Rawls, 2001: 161). He briefly mentioned a ‘proportional income tax’ (ibid.: 162), but then suggested that ‘income taxation might be avoided altogether’ if proportional taxation of expenditures with a personal exemption was instituted (ibid.: 161). He added that ‘the difference principle might, then, roughly be satisfied by raising and lowering this minimum [exemption] and adjusting the constant marginal rate of taxation’ (ibid.).

Incentives do not belong in ideal theory

Rawls proposed a theory of justice for a well-ordered society that had achieved a high level of economic development. The theory assumed that reasonable individuals would comply with the institutions when the basic structure was just. For instance, we saw that Rawls (1958) assumed that reasonable citizens would pay their fair share of tax and not freeride. He argued that fleshing out such a theory was an essential prerequisite to addressing the duties and obligations of persons in nonideal situations when some individuals were not fully compliant or when institutions did not guarantee equal liberties to all (Rawls, 1971: §39, 245–246). Ideal theory would give us a goal to attain, and nonideal theory would point out the steps needed to achieve it (Rawls, 1999: 90). Rawls believed that his social vision could guide nonideal theory, clarify reforms, and even help us to prioritise the injustices to correct (Rawls, 2001: 13). In this section, I argue that Rawls’s willingness to propose a ‘realist utopia’ (Rawls, 1999: 11) creates a tension in his ideal theory. Rawls’s concern for incentives suggests an interpretation of the difference principle that makes some of his tax suggestions nonideal, as they rely on behavioural assumptions that should not arise in his ideal theory.

Shortly after Rawls’s presentation at the AEA, Musgrave published his criticism of Rawls’s *A Theory of Justice* in the *Quarterly Journal of Economics*, which he was editing. Musgrave took issue with the

¹⁷An expenditure tax like that favoured by Kaldor and Rawls would have an income effect that would lead individuals to work more to afford the same level of expenditure. It would also have a substitution effect by reducing the real value of work (decreasing the opportunity cost of leisure) by making consumption goods relatively more expensive.

maximin rule and especially its use as a social welfare function within the optimal income tax model of Mirrlees (1971).¹⁸ Considering that early results of optimal tax theory had ‘dubious’ practical implications because ‘too little [was] know[n] about the magnitude of the incentive effects, particularly in the upper brackets’, Arrow had already shared his disappointment that Rawls did not explain better why he assumed that individuals would reduce their work effort if they were taxed for the purpose of redistributing income (Arrow, 1973: 258–9). Musgrave went further: ‘Throughout this discussion (including Rawls’s own work as well as discussion in economic journals) it is assumed, rather strangely, that the income tax reduces work effort. This may or may not be the case, depending on the magnitude of income and substitution effects for both taxpayers and transfer recipients’ (Musgrave, 1974: 629). To the extent that a person’s welfare depends on their consumption of leisure and goods, but that redistribution can only proceed in goods or income, the taxation of income discriminates in favour of those with a higher preference for leisure: ‘It is to the advantage of recluses, saints, and (nonconsulting) scholars who earn but little and hence will not have to contribute greatly to redistribution’ (Ibid.: 632). One of Musgrave’s points was whether an ideal theory of justice like Rawls’s should allow individuals to practise ‘defensive leisure substitution’ (Ibid.: 631). He also argued that maximin should be derived directly from the values defended by Rawls (equal worth, self-esteem, stability, harmony) rather than the unrealistic assumption that individuals would choose as if they had maximum risk aversion (Ibid.: 628). Overall, Musgrave believed that while these were ‘fun and games for the economist’, it might not belong in a philosophical theory of justice (see also Musgrave, 1992).

In his response, Rawls suggested that part of Musgrave’s objection could be addressed by including leisure in the index of primary goods. How that would work was not clear at first, but Rawls clarified his suggestion in a discussion with Philippe van Parijs in 1987. Van Parijs (1991) proposed a ‘Rawlsian’ case for universal basic income, to which Rawls objected. He argued that if Rawls was truly committed to promoting equal freedom for all, irrespective of their conception of the good life, Rawls should support the idea of providing everyone with an unconditional cash grant. Rawls (1974a) responded that the index of primary goods could include a certain number of hours of leisure equivalent to 24 hours minus a normal workday. That is, ‘those who surf all day off Malibu must find a way to support themselves’ – they cannot claim support from the state (Rawls, 1988: 257 n.; see also Rawls, 2001: 179; Freeman, 2007: 229). In contrast to the libertarian post-Fordist spirit of the 1980s in which van Parijs’ proposal emerged, Rawls consistently defended his vision of society as a cooperative venture in which everyone (who can) must work: ‘What men want is meaningful work in free association with others’ (Rawls, 1971: §44, 290). The difference principle was meant to capture this necessity of working. For Rawls, the distribution of primary goods was attached to positions that must be filled for the economy to function. As Rawls put it in a draft of his AEA paper, if we had lived in a ‘manna economy’ where the social product was fixed, the maximin would equalise income and wealth.¹⁹

G.A. Cohen famously argued that the difference principle was not a principle of justice for a well-ordered society populated by moral persons who treat each other as equals, such as the one envisaged by Rawls in his *Theory of Justice*. Cohen doesn’t put it in such terms, but his demonstration implies that Rawls’s tax recommendations, such as a proportional expenditure tax, do not belong to ideal theory. As Cohen put it, it can be a principle of public policy for an unjust society (non-ideal), but its effectiveness depends on factual claims that fail the test of public interpersonal justification (Cohen, 2008: 85, 40). At first, Cohen doesn’t challenge the normative premise that ‘economic inequalities are justified when they make the worse-off people materially better off’, but he thinks Rawlsian persons could not defend the factual premise required by the difference principle, namely that if we don’t allow these inequalities of income, the rich would ‘go on strike’ (ibid.: 33–34). In other words, in a well-ordered society characterised by the ideal theory of justice, a more talented and richer person could not face a poor person

¹⁸Besides Atkinson’s 1973 survey, Rawls was familiar with Cooter and Helpman (1974)’s paper that Cooter sent him in the spring of 1973 (Cooter to Rawls, dated 10 April 1972 [likely 1973]; Rawls to Cooter, 7 May 1973; Rawls Papers, Box 19, Folder 7). Cooter’s 1975 dissertation in economics was jointly supervised by Musgrave, Jerry Green, and Rawls.

¹⁹Some remarks about distributive equity’: 6, Box 24, Folder 12, Papers of John Rawls.

in the eyes and tell them that if they are taxed at a higher marginal rate, they are not going to work as hard as if they were taxed at the same rate as the poor. The incentive argument can be defended by humble third parties as a precautionary principle, but a rich person could not make it the public motivation of his action in front of a poor person. By defending his incentive, the rich man would demonstrate a lack of community with the poor. Couched in the naturalising third-party language of economics, the incentive premise is plausible, but it contradicts the importance that Rawls gives to 'fraternity', that is, 'equality of social esteem' among citizens of a democratic society (Rawls, 1971: §17). It also contradicts the Kantian interpretation defended by Rawls: in the ideal theory, individuals are assumed to 'knowingly act on the principles of justice in the ordinary course of events' (Rawls, 1971: §40, 253). If they had 'a highest-order desire' to 'act from the principles of justice' that they have themselves chosen as free and autonomous agents, individuals could not at the same time act from the incentive argument (Cohen, 2008: 74, Rawls, 1980). Thus, for Cohen, the difference principle does not belong in an ideal theory of justice, even if it can be a reasonable political compromise for an unjust capitalist society. A fortiori, Rawls's tax proposals, which rely on the incentive argument, such as his defence of proportional taxation (of consumption) and his opposition to progressive taxation of income, belong to the realm of nonideal theory.

Surely, the difference principle can be defended on the grounds that it is a conditional principle. It does not prescribe a certain level of inequality, nor does it recommend a given structure of incentive payment or put a priori constraints on the taxation of labour income. It only states that inequalities are acceptable if they are to the advantage of the worst-off groups. It might be that institutions established to secure the fair value of equal liberties and fair equality of opportunities for all reduce inequalities to the point of rendering the difference principle redundant. Or it might also be that some inequalities would remain after the implementation of the first and the second part of the second principle, but those inequalities would not be to the advantage of the worst off, which means they would have to be reduced in virtue of the difference principle. In the first scenario, a progressive income tax might not be necessary, and in the second one, it might be an acceptable form of redistribution to the extent that the income differential is not to the advantage of the worst-off representative member of a social/occupational group. However, Rawls's opposition to progressive income tax, his defence of proportional income tax, and his opposition to universal basic income suggest he envisaged a different scenario, even for an ideal theory.

I should acknowledge that Rawls argued that inequalities would not be too large in a society guided by the principles of justice: 'Although in theory the difference principle permits indefinitely large inequalities in return for small gains to the less favored, the spread of income and wealth should not be excessive in practice, given requisite background institutions'. (Rawls, 1971: §81, 536). Indeed, large inequalities would damage the self-esteem and self-respect of those at the bottom, making them excusably envious (Rawls, 1971: §80, 534). For Joshua Cohen (2001), Rawls's difference principle would be acceptable to egalitarians if we think that large inequalities would not arise. G.A. Cohen (2003) goes further by pointing out that the plausibility of the difference principle for egalitarian contracting parties hinges on a factual premise. If inequalities were large, the difference principle should not be acceptable. I infer that the factors that would make inequalities small in the first place would also reduce demands for incentive compensation and reduce opposition to progressive taxation of income.

Rawls doesn't seem to have drawn all the implications of a fair basic structure of society on individual motivations. He did recognise that 'the basic structure as a social and economic regime is not only an arrangement that satisfies given desires and aspirations but also an arrangement that arouses further desires and aspirations' (Rawls, 2001: §16.2, 56). The basic structure impacts 'culture' and shapes the 'character and interests of individuals themselves' (J. Cohen, 2001, 381; Freeman, 2014; Rawls, 1993: 269). These claims contradict Rawls's other claim to take 'people as they are (by the laws of nature)' (Rawls, 1999: 13).

One can build an institutionalist and normative theory of fiscal policy based on the assumption that individuals are strongly self-interested. Rules would be selected assuming the worst from others, as Rawls had suggested in 1958. An approach that is also followed in constitutional political economy in the wake of Buchanan's work (see Delmotte, 2020; Hamlin, 2018: 29; Vanberg, 2005).

However, in an ideal theory where behaviour is determined by a just basic structure as analysed by Rawls, why should we assume that individuals have a strong preference for leisure or that they will reduce their work effort when taxed at a higher marginal rate? Even in our unjust world, empirical evidence does not unanimously support what Rawls seems to take as a fact of life: estimates of the labour supply elasticity are generally small and vary widely across empirical studies (Engen *et al.*, 1997). Rawls might have responded that ‘moral philosophy must be free to use contingent assumptions and general facts as it pleases’ (Rawls, 1971: §9, 51). Still, my point is that there isn’t much empirical content in the concept of human nature that is not shaped by the institutional structure. Expectations about career prospects, remuneration, and the distribution of the tax burden are all shaped by institutions; there are no exogenous preferences.²⁰

I am not arguing that in a just society employment would be decided by a central authority and everyone would receive the same pay. Nor am I arguing that citizens would have an individual duty to do everything they can to help the worst-off groups in society. Rawls is clear that the principles of justice apply to the basic structure of society and only indirectly to individuals (Freeman, 2014). In a liberal society such as that defended by Rawls, individuals preserve considerable freedom in choosing their career path. It is in the interest of all, including the worst off, that people with specific talents are encouraged to develop them (Daniels, 2002). This justifies some differences in remuneration. However, in a just society, to the extent that individuals have a duty to promote egalitarian social relationships based on mutual respect, it is likely that democratic governance mechanisms might emerge in the workplace (O’Neill, 2020), thereby further reducing expectations of significant pay differentials.

One could assume that individuals would refuse to work if taxed at a higher marginal rate, but we could also assume otherwise, and doing so would be more in line with how Rawls characterises equal persons who live by the principles of justice and support a system of social cooperation based on reciprocity, realising the ideal of fraternity. To the extent that tax recommendations can be made in ideal theory, they would need to be compatible with first principles (free and equal citizens) and plausible ‘facts’ or behavioural assumptions fitting with the rest of the ideal institutional system proposed by Rawls. Thus, some of Rawls’s suggestions on tax matters might be valuable compromises for the nonideal world, but they are certainly not part of an ‘ideal scheme for a well-ordered society’ as he sometimes claimed (Rawls, 1967; 1971: §43, 279).

Conclusion

Amartya Sen aptly remarked that ‘the insights that people have obtained from Rawls’s analysis of ‘justice as fairness’ seem to have gone greatly beyond the limits he has himself imposed, and it is not clear to me that these insights have been all misderived and misdirected’ (Sen, 1992: 79). Political philosophy exists today in the shadow cast by the monument erected by Rawls (Forrester, 2019). The same could probably be said of many questions in normative economics. Economists and philosophers invoke Rawls in support of their proposals for institutional reform, which gives grounds for investigating what Rawls wrote and why he wrote it, as well as what he didn’t write.

Over the years, Rawls borrowed ideas from the right, left, and centre, which made his work so interesting to scholars interested in institutions from a wide range of disciplines. A humble man, he also acknowledged and cited generously. Cohen remarked that there was a tension in Rawls’s framework between a ‘bargaining conception’ and a ‘community conception of social relationships’ (Cohen, 2008: 82). Indeed, Rawls’s theory tried to reconcile the ‘kingdom of ends’ (1971: 252; 1975: 264) with a Smithian commercial world in which weak ties based on self-interest hold society together.

In terms of tax proposals, Rawls drew mostly from J.S. Mill, N. Kaldor, R.A. Musgrave, and J. Meade. While he did not propose a fully fledged and coherent theory of taxation, two aspects of his remarks on

²⁰For my purpose, the distinction between social ethos and institutions made by G.A. Cohen (2008) and J. Cohen (2001) is not needed. Both are endogenous to the institutional system.

taxation stand out. One is his rejection of the ability-to-pay principle, the basis on which most progressive tax proposals by economists in the 20th century were anchored. Rawls rejected ability-to-pay in all its forms: equality of sacrifice, vertical equity, and the taxation of natural endowments. A second, and related, is his unfaltering commitment to preserving work incentives. Rawls took welfare economics seriously. He searched for a normative criterion (and a framework, the lexicographic ordering of principles) that would resolve the indeterminateness of the new welfare economics: the maximin criterion would pick the point ‘closest to equality consistent with Pareto efficiency’.²¹ But Rawls wasn’t concerned with allocative efficiency. He envisaged a cooperative society in which people had to be incentivised to voluntarily come together to produce.

Are there unambiguous recommendations that follow from Rawls’s ideal theory? Sugin (2004) argues that while Rawls’s theory does not prescribe a unique tax system, it can rule out tax proposals that violate the principles of justice, such as a discriminatory tax based on skin colour or a highly regressive tax that would worsen the position of the poorest. Rawls claimed that justice as fairness singles out ‘with greater sharpness the graver wrongs a society should avoid’ than alternative conceptions of justice (Rawls, 1971: §31, 201). Even if it were feasible, Rawls rejected the taxation of natural endowments of talents because it conflicted with individual freedom to choose their occupation. The other policy that follows within the realm of ideal theory is the taxation of gifts and bequests. In that, Rawls belonged to a long tradition of political thinkers from Rousseau to Saez and Zucman (2019: 158), passing through J.R. Commons and many others. Referring to Meade, Rawls proposed a progressive taxation of inherited income to prevent the concentration of wealth that would confer political power and damage the fair value of political liberty and the fair value of equality of opportunity.

Beyond those suggestions, it seems that not much tax policy can be recommended within ideal theory (see also Hamlin: 2018, 18). Rawls’s discussion of economic institutions is often oversimplified (O’Neill, 2020). As Buchanan noted long ago, Rawls ‘doesn’t discuss the critically important bridge between such an idealized setting and that within which any discussion of basic structural rearrangement might, in fact, take place’ (Buchanan, 1975: 175). One cannot derive a unique set of tax recommendations from *A Theory of Justice* for the non-ideal world. As Rawls suggested, ‘in practice we must usually choose between several unjust, or second best, arrangements; and then we look to nonideal theory to find the least unjust scheme’, keeping in mind the ranking of the principles (Rawls, 1971: §43, 279, 303).

The principles of justice do not provide incontestable guidance in choosing, for instance, between a progressive income tax and a proportional expenditure tax to prevent the rapid accumulation of wealth within a lifetime and to pay for the important public goods. Based on the difference principle and his concern for incentivising work, Rawls favoured a proportional expenditure tax, but I have argued that this suggestion relies on a behavioural assumption that conflicts with his account of how institutions shape behaviour and social cooperation.

Considering the underdetermination of the tax system by *A Theory of Justice*, non-ideal tax policy must rely on ‘society’s historical circumstances, to its traditions of political thought and practice, and much else’ (Rawls, 2001: §42, 139). It needs to rely on a realistic account of individual values and habits of thought. In the real world, ‘given the injustice of existing institutions, even steeply progressive income taxes [might be] justified when all things are considered’ (Rawls, 1971: §43, 279). In this context of institutional compromise, it would be worthwhile to consider how the Rawlsian framework might be supplemented by ‘domain-specific’ principles of tax equity, such as vertical and horizontal equity.²²

²¹‘Some remarks about distributive equity’, *op. cit.*

²²Horizontal and vertical equity are ‘domain-specific’ principles of justice that can help articulate the fairness of a tax reform taking for granted other institutions. For a genealogy of the concepts, see Desmarais-Tremblay (2021) and Desmarais-Tremblay et al. (2023). For a philosophical defence of those principles, see Lindsay (2016) and Verbeek (2025).

Acknowledgements. A first draft of this paper was presented at the 27th annual ESHET conference at the University of Graz and at the 7th Economic Philosophy Conference at the University of Reims Champagne-Ardenne in 2024. I thank Herrade Igersheim, Eva Jacob, Cyril Hédoin, Nathanaël Colin-Jaeger, and Thomas Ferretti for their very useful feedback. I am also grateful to seven anonymous referees for this journal, whose comments greatly helped me improve the paper. I am fully responsible for all remaining errors.

References

- Arrow K.J. (1951). *Social Choice and Individual Values*. New Haven: Yale University Press.
- Arrow K.J. (1973). Some ordinalist-utilitarian notes on Rawls's theory of justice. *The Journal of Philosophy* 70(9), 245–263.
- Atkinson A.B. (1973). How progressive should income tax be?. In Parkin, M., Nobay, A.R. (eds), *Essays in Modern Economics. The Proceedings of the Association of University Teachers of Economics: Aberystwyth 1972*. London: Longman, pp. 90–109.
- Atkinson A.B. and Stiglitz J.E. (1976). The design of tax structure: Direct versus indirect taxation. *Journal of Public Economics* 6(1/2), 55–75.
- Atkinson A.B. and Stiglitz J.E. (1980). *Lectures on Public Economics*. New York: McGraw Hill.
- Bok P.M. (2017). To the mountaintop again: The early Rawls and post-protestant ethics in postwar America. *Modern Intellectual History* 14(1), 153–185.
- Brownlee W.E. (2000). Historical perspective on U.S. tax policy toward the rich. In Slemrod J.E. (ed), *Does Atlas Shrug? The Economic Consequences of Taxing the Rich*. Cambridge, MA: Harvard University Press, pp. 29–73.
- Buchanan J.M. (1972). Rawls on justice as fairness. *Public Choice* 13, 123–128.
- Buchanan J.M. (1975). *The Limits of Liberty: Between Anarchy and Leviathan*. Chicago: The University of Chicago Press.
- Buchanan J.M. (1984). The ethical limits of taxation. *The Scandinavian Journal of Economics* 86(2), 102–14.
- Chandler D. (2023). *Free and Equal: What Would a Fair Society Look Like?*. Dublin: Allen Lane.
- Cohen G.A. (2003). Facts and principles. *Philosophy & Public Affairs* 31(3), 211–245.
- Cohen G.A. (2008). *Rescuing Justice and Equality*. Cambridge, MA: Harvard University Press.
- Cohen J. (2001). Taking people as they are?. *Philosophy & Public Affairs* 30(4), 363–386.
- Coker D.C. 2021. *Economics to Social Philosophy: The Foundation of John Rawls's Contribution*. PhD thesis. Fairfax, VA: George Mason University.
- Colin-Jaeger N., Dold M. and Gascoin A. 2022. Realism and deliberation in normative economics. In Badei S. and Grivaux A. (eds), *The Positive and the Normative in Economic Thought*. London: Routledge, pp. 108–28.
- Commons J.R. (1934). *Institutional Economics: Its Place in Political Economy*. Vol. 2. 1959 reedition. Madison: University of Wisconsin Press.
- Cooter R.D. and Helpman E. (1974). Optimal income taxation for transfer payments under different social welfare criteria. *The Quarterly Journal of Economics* 88(4), 656–670.
- D'Aspremont C. (1984). Rawls et les économistes. In Ladrière J. and Van Parijs P. (eds), *Fondements d'une théorie de la justice: Essais critiques sur la philosophie politique de John Rawls*. Louvain-la-Neuve: Éditions de l'Institut supérieur de philosophie, pp. 83–103.
- Daniels N. (2002). Democratic equality: Rawls's complex egalitarianism. In Freeman S. (ed), *The Cambridge Companion to Rawls*. Cambridge: Cambridge University Press, pp. 241–276.
- Delmotte C. (2020). Tax uniformity as a requirement of justice. *Canadian Journal of Law & Jurisprudence* 33(1), 59–83.
- Desmarais-Tremblay M. (2017). A genealogy of the concept of merit wants. *The European Journal of the History of Economic Thought* 24(3), 409–440.
- Desmarais-Tremblay M. (2021). Généalogie du principe d'équité horizontale. Une contribution à l'histoire de la normativité en théorie des finances publiques. *Revue de philosophie économique* 22(2), 149–176.
- Desmarais-Tremblay M. and Johnson M. (2026). Distributing the tax burden in postwar America: Empirical estimates and normative principles. *History of Political Economy*. Forthcoming.
- Desmarais-Tremblay M., Johnson M. and Sturn R. (2023). From public finance to public economics. *The European Journal of the History of Economic Thought* 30(5), 934–964.
- Elster J. (1992). *Local Justice: How Institutions Allocate Scarce Goods and Necessary Burdens*. New York: Russell Sage.
- Engen E., Gravelle J. and Smetters K. (1997). Dynamic tax models: Why they do the things they do. *National Tax Journal* 50(3), 657–682.
- Ferretti, T. (2022). Measuring freedom: Towards a solution to John Rawls' indexing problem. *Erasmus Journal for Philosophy and Economics* 15(1), 1–31.
- Forrester K. (2019). *In the Shadow of Justice: Postwar Liberalism and the Making of Political Philosophy*. Princeton, NJ: Princeton University Press.
- Freeman S. (2007). *Rawls*. London: Routledge.
- Freeman S. (2014). The basic structure of society as the primary subject of justice. In Mandle J. and Reidy D.A. (eds), *A Companion to Rawls*. Chichester: Wiley-Blackwell, pp. 88–111.
- Fried B.H. (1999). The puzzling case for proportionate taxation. *Chapman Law Review* 2, 157–195.

- Gajevic Sayegh A. (2017). A case of non-ideal guidance: Tackling tax competition. *Moral Philosophy and Politics* 4(1), 141–171.
- Gališanka A. (2019). *John Rawls: The Path to A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Hamlin A. (2018). What political philosophy should learn from economics about taxation. In O'Neill M. and Orr S. (eds), *Taxation: Philosophical Perspectives*. Oxford: Oxford University Press, pp. 17–36.
- Hawi R. (2016). *John Rawls – Itinéraire d'un libéral américain vers l'égalité sociale*. Paris: Classiques Garnier.
- Hawi R. (2024). Les coulisses de la pensée économique de John Rawls. *Economia* 14(3), 573–608.
- Hayek F.A. (1960). *The Constitution of Liberty*. Chicago: The University of Chicago Press.
- Hobbes T. (1651). *Leviathan*. Edited by R. Tuck. 1991 Edn. Cambridge: Cambridge University Press.
- Igersheim H. (2022). Rawls and the economists: The (im)possible dialogue. *Revue économique* 73(6), 1013–1037.
- Jackson B. (2012). Property-owning democracy: A short history. In O'Neill M. and Williamson T. (eds), *Property-Owning Democracy*. Malden, MA: Wiley-Blackwell.
- Jackson B. and Stemplowska Z. (2021). 'A quite similar enterprise . . . interpreted quite differently'? James Buchanan, John Rawls and the politics of the social contract. *Modern Intellectual History* 18(4), 1010–1033.
- Kaldor N. (1955). *An Expenditure Tax*. London: George Allen & Unwin.
- Knight F.H. (1935). Economic theory and nationalism. In *The Ethics of Competition*. London: George Allen & Unwin, pp. 277–359.
- Kolm S.-C. (2003). Quelques souvenirs de John Rawls. *Revue de philosophie économique* 7, 21–31.
- Lindsay I.K. (2016). Tax fairness by convention: A defense of horizontal equity. *Florida Tax Review* 19(2), 79–119.
- Little D. (2014). Rawls and economics. In Mandle J. and Reidy D.A. (eds), *A Companion to Rawls*. Chichester: Wiley-Blackwell, pp. 514–521.
- Luce R.D. and Raiffa H. (1957). *Games and Decisions: Introduction and Critical Survey*. New York: Wiley.
- Mankiw N.G. and Weinzierl M. (2010). The optimal taxation of height: A case study of utilitarian income redistribution. *American Economic Journal: Economic Policy* 2(1), 155–176.
- Marciano A. and Medema S.G. (2015). Market failure in context: Introduction. *History of Political Economy* 47(Supplement), 1–19.
- Meade J.E. (1964). *Efficiency, Equality and the Ownership of Property*. London: George Allen & Unwin.
- Mehrotra A.K. (2013). *Making the Modern American Fiscal State: Law, Politics, and the Rise of Progressive Taxation, 1877–1929*. Cambridge: Cambridge University Press.
- Mill J.S. (1848). *Principles of Political Economy, vol. 3 of the Collected Works of John Stuart Mill*. 1965 Edn. London: University of Toronto Press and Routledge & Kegan Paul.
- Mirrlees J.A. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies* 38(2), 175–208.
- Murphy L. and Nagel T. (2002). *The Myth of Ownership: Taxes and Justice*. New York: Oxford University Press.
- Musgrave R.A. (1957). An expenditure tax. By Nicholas Kaldor (book review). *American Economic Review* 47(1), 200–205.
- Musgrave R.A. (1959). *The Theory of Public Finance: A Study in Public Economy*. New York: McGraw-Hill.
- Musgrave R.A. (1974). Maximin, uncertainty, and the leisure trade-off. *Quarterly Journal of Economics* 88(4), 625–632.
- Musgrave R.A. (1985). A brief history of fiscal doctrine. In Auerbach A.J. and Feldstein M. (eds), *Handbook of Public Economics*, vol. 1. Amsterdam: Elsevier, pp. 1–59.
- Musgrave R.A. (1992). Social contract, taxation and the standing of deadweight loss. *Journal of Public Economics* 49(3), 369–381.
- Nozick R. (1974). *Anarchy, State, and Utopia*. New York: Basic Books.
- O'Neill M. (2017). Philosophy and public policy after Piketty. *Journal of Political Philosophy* 25(3), 343–375.
- O'Neill M. (2020). Social justice and economic systems: On Rawls, democratic socialism, and alternatives to capitalism. *Philosophical Topics* 48(2), 159–201.
- O'Neill M. and Orr S. (2018). Introduction. In O'Neill M. and Orr S. (eds), *Taxation: Philosophical Perspectives*. Oxford: Oxford University Press, pp. 1–14.
- O'Neill M. and Williamson T. (eds) (2012). *Property-Owning Democracy: Rawls and Beyond*. Malden, MA: Wiley-Blackwell.
- O'Neill M. and Williamson T. (2015a). Branches of government. In Mandle J. and Reidy D. A. (eds), *The Cambridge Rawls Lexicon*. Cambridge: Cambridge University Press, pp.65–68.
- O'Neill M. and Williamson T. (2015b). Taxation. In Mandle, J. and Reidy, D.A. (eds), *The Cambridge Rawls Lexicon*. Cambridge: Cambridge University Press, pp. 825–827.
- Olson K.A. (2010). The endowment tax puzzle. *Philosophy & Public Affairs* 38(3), 240–271.
- Paul R.E. (1954). *Taxation in the United States*. Boston: Little, Brown.
- Phelps E.S. (1973). Taxation of wage income for economic justice. *Quarterly Journal of Economics* 87(3), 331–354.
- Pogge T.W. (1989). *Realizing Rawls*. Ithaca, NY: Cornell University Press.
- Rawls J. (1958). Justice as fairness. *Philosophical Review* 67(2), 164–194.
- Rawls J. (1964). Legal obligation and the duty of fair play. In Hook, S. (ed), *Law and Philosophy*. New York: New York University Press, pp. 3–18.

- Rawls J. (1967). Distributive justice. In Laslett, P. and Runciman, W.G. (eds), *Philosophy, Politics and Society*, third series. Oxford: Basil Blackwell, pp. 58–82.
- Rawls J. (1971). *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Rawls J. (1974a). Reply to Alexander and Musgrave. *Quarterly Journal of Economics* 88(4), 633–655.
- Rawls J. (1974b). Some reasons for the maximin criterion. *American Economic Review* 64(2), 141–146.
- Rawls J. (1975). A Kantian conception of equality. In Freeman, S. (ed), *Collected Papers. John Rawls*. 1999 Edn. Cambridge, MA: Harvard University Press, pp. 254–266.
- Rawls J. (1980). Kantian constructivism in moral theory. *Journal of Philosophy*, 77(9), 515–572.
- Rawls J. (1982). Social unity and primary goods. In Sen A. and Williams B. (eds), *Utilitarianism and Beyond*. Cambridge: Cambridge University Press, pp. 159–185.
- Rawls J. (1988). The priority of right and ideas of the good. *Philosophy & Public Affairs* 17(4), 251–276.
- Rawls J. (1993). *Political Liberalism*. New York: Columbia University Press.
- Rawls J. (1999). *The Law of Peoples*. Cambridge, MA: Harvard University Press.
- Rawls J. (2001). *Justice as Fairness: A Restatement*. Edited by Kelly, E. Cambridge, MA: Harvard University Press.
- Rawls J. (2009). *A Brief Inquiry into the Meaning of Sin and Faith*. Edited by Nagel, T. Cambridge, MA: Harvard University Press.
- Saez E. and Zucman G. (2019). *The Triumph of Injustice: How the Rich Dodge Taxes and How to Make them Pay*. New York: W. W. Norton.
- Schmidtz D. (2011). Nonideal theory: What it is and what it needs to be. *Ethics* 121(4), 772–796.
- Seligman E.R.A. (1894). Progressive taxation in theory and practice. *Publications of the American Economic Association* 9(1–2), 7–222.
- Sen, A.K. (1992). *Inequality Reexamined*. Oxford: Russell Sage Foundation and Oxford University Press.
- Sen A.K. (2009). *The Idea of Justice*. Cambridge, MA: Harvard University Press.
- Simmons A.J. (2010). Ideal and nonideal theory. *Philosophy & Public Affairs*, 38(1), 5–36.
- Simons H.C. (1934). *A Positive Program for Laissez Faire*. Chicago: The University of Chicago Press.
- Stemplowska Z. (2017). Non-ideal theory. In Lippert-Rasmussen, K. Brownlee, K. and Coady, D. (eds), *A Companion to Applied Philosophy*. Chichester: Wiley-Blackwell, pp. 284–296.
- Sugin L. (2004). Theories of distributive justice and limitations on taxation: What Rawls demands from tax systems. *Fordham Law Review* 72(5), 1991–2014.
- Trentmann F. (2017). *Empire of Things*. London: Penguin.
- Turmel P. (2024). Rawls et l'impôt progressif: deux arguments. *Economia* 14(3), 547–571.
- Valentini L. (2012). Ideal vs. non-ideal theory: A conceptual map. *Philosophy Compass* 7(9), 654–664.
- Van Parijs P. (1991). Why surfers should be fed: The liberal case for an unconditional basic income. *Philosophy & Public Affairs* 20(2), 101–131.
- Vanberg V.J. (2005). Market and state: The perspective of constitutional political economy. *Journal of Institutional Economics* 1(1), 23–49.
- Veblen T. (1899). The preconceptions of economic science II. *Quarterly Journal of Economics* 13(4), 396–426.
- Verbeek B. (2025). Isolationism, instrumentalism and fiscal policy. *Economics & Philosophy* 41(1), 79–97.
- Walraevens B. (2023). Rawls's maximin and optimal taxation theory. *European Journal of the History of Economic Thought* 30(5), 860–882.
- Walraevens B. (2024). Des inégalités économiques à la justice sociale: l'influence de John Rawls sur Thomas Piketty. *Economia* 14(3), 483–518.
- Weithman P. (2010). *Why Political Liberalism? On John Rawls's Political Turn*. Oxford: Oxford University Press.
- Wicksell K. (1896). A new principle of just taxation. In Musgrave, R.A. and Peacock, A. (eds), *Classics in the Theory of Public Finance*. 1958. London: Macmillan, pp. 72–118.