WARWICK

ECONOMICS

CRETA

Centre for Research in Economic Theory and its Applications

Discussion Paper Series

Behavioural utilitarianism and distributive justice

Giorgos Galanis and Roberto Veneziani

October 2020 No: 67

CRETA
Centre for Research in Economic Theory and its Applications
Department of Economics
University of Warwick, Coventry,
CV4 7AL, United Kingdom

warwick.ac.uk/fac/soc/economics/research/centres/creta/papers

Behavioural utilitarianism and distributive justice

Giorgos Galanis*and Roberto Veneziani†‡

Abstract

What are the distributive implications of utilitarianism? Is it compatible with a concern for equality, as many utilitarians have argued? We analyse these questions in the context of a pure allocation problem. We consider an infinitely-lived economy and, drawing on the behavioural literature, assume that individuals have reference-dependent preferences: agents' utility is a function of current consumption and a reference point which captures consumption habits, or the agents' upbringing. Assuming a history of inequalities in consumption and welfare, we show that the utilitarian allocation is *equalising*: starting from an unequal distribution, consumption and welfare inequalities decrease over time at the utilitarian optimum. However, even though agents are in a relevant sense identical, equality does not obtain at any finite time.

Keywords: utilitarianism, inequality, reference dependent preferences.

JEL Classification: D63 (Equity, Justice, Inequality, and Other Normative Criteria and Measurement); D9 (Micro-Based Behavioural Economics)

*Goldsmiths, University of London and Centre for Research in Economic Theory and its Applications, University of Warwick.

†School of Economics and Finance, Queen Mary, University of London

‡We thank Stephen Engelmann, Marc Fleurbaey, Peter Hammond, Richard Layard, Peter H. Matthews, Herakles Polemarchakis, Jonathan Riley, John Roemer, Zvi Safra, Uzi Segal, Kostas Zachariadis, and audiences in Notre Dame, Warwick and London for useful comments on an early draft of the paper. The usual disclaimer applies.

1 Introduction

Classical utilitarianism – the doctrine according to which "the just social alternative is that which maximises the sum of individual utilities" (Roemer, 1996, p.5) – is undoubtedly one of the most prominent and widely adopted approaches in normative economics and in policy analyses. Yet, as critics have long pointed out, it has potentially undesirable distributive implications. On the one hand, the utilitarian planner is definitionally indifferent between alternative allocations given a certain level of aggregate utility. Call this *the inequality indifference critique*. On the other hand, the very maximisation of total utility may require an extremely unequal allocation of both resources and utility, if agents have different preferences. Call this the *inequality generation critique*. Utilitarians have rejected these criticisms, or at least significantly deflated their relevance and have traditionally argued that utilitarianism is compatible with what they have historically considered to be one of the fundamental characteristics of a well ordered society, namely equality.

In this paper, drawing from the behavioural literature, we provide a novel perspective on this debate. We analyse the distributive implications of utilitarianism – and the compatibility of a utilitarian approach with a concern for equality – in the context of a dynamic economy where past consumption habits influence present preferences.

To see why the insights from the behavioural literature may shed significant light on these issues, consider the utilitarian response to both criticisms. The rebuttal of the inequality indifference critique focuses on the possibility of refinements of the utilitarian view requiring an egalitarian distribution of a given amount of aggregate utility. Bentham, for example, proposed that equality along with security, subsistence and abundance should be one of the objectives of a government. Sidgwick (1907) argued that utilitarianism necessarily relied on a principle of equality in order to determine the best mode of distributing a given amount of utility.[1]

[1]It is "practically important to ask whether any mode of distributing a given quantum of happiness is better than any other. Now the utilitarian formula seems to supply no answer to this question: at least we have to supplement the principle of seeking the greatest happiness of the whole by some

The utilitarian response to the inequality generation critique involves an argument that can be interpreted either as an empirical or as a normative claim: given that agents hold similar preferences, and given decreasing marginal utility (say, of income) the optimal utilitarian distribution will be roughly egalitarian. The empirical interpretation is especially prominent in classical hedonistic utilitarian approaches, in which only the agents' actual subjective preferences (whatever they are) matter. As Samuelson (1964, p.173) forcefully put it, "If people are all alike, or potentially all alike in the longer run as environmentalists like Bentham and Mill believed, total U is maximized by an equal distribution of income".

In Bentham, for example, the egalitarian implications of utilitarianism derive from the 'axioms of moral and political pathology', namely those empirical generalisations that are "expressive of the connexion between such occurrences as are continually taking place, or hable to take place, and the pleasures and pains which are respectively the results of them" (Bentham, 1838–1843b, p.224). Based on these axioms, which include claims on the marginal utility of money and wealth, Bentham (1838–1843a, p.313) concludes that "We may observe, that in a nation which prospers by agriculture, manufactures, and commerce, there is a continual progress towards equality ... This will be the result of different habits formed by opulence and poverty".

In a non-hedonistic approach, instead, the previous argument refers to possibly counterfactual, objective, or normatively relevant preferences, or preferences that agents would hold ideally or upon reflection. Then, one may argue that, once epiphenomenal and morally irrelevant characteristics are set aside, agents' true, or rational preferences are indeed largely similar, reflecting a shared human nature. While adopting a non-hedonistic approach, for example, in his classic defence of utilitarianism, J.J. Smart explicitly argues that a utilitarian criterion will not lead to extreme inequalities (Smart and Williams, 1973, pp.34ff). Similarly, in his *The Economics of control*, Abba Lerner (1944, p.29) provides an argument based on the impossibility of knowing

principle of Just or Right distribution of this happiness. The principle which most utilitarians have either tacitly or expressly adopted is that of pure equality" (Sidgwick, 1907, p.417).

people's preferences: "If it is impossible, on any division of income, to discover which of any two individuals has a higher marginal utility of income, the probable value of total satisfactions is maximized by dividing income evenly." This egalitarian implication of classical utilitarianism is endorsed by Samuelson (1964, p.175) who argues that "if the crucial equal-ignorance assumption were really acceptable to every reasonable observer, then each and every person (subject to the postulated concave utility that renders him a risk averter) would vote for a regime of equal-distribution of income, and this constitutional feature would be instituted by unanimous vote."

The foundations of both the hedonistic and the non-hedonistic approach have been subjected to intense critical scrutiny.[2] In this paper, we abstract from foundational issues and analyse the utilitarian responses to egalitarian critics on their own ground by focusing on the relation between utilitarianism and equality in the context of a pure allocation problem. Unlike in most of the literature, however, we consider this issue in a dynamic framework and assume that in every period the utilitarian planner has to allocate a good of fixed size among a given number of agents who have, in some relevant sense, the same utility function. However, in order to capture Bentham's notion of evolving "habits formed by opulence and poverty", and consistent with a large behavioural literature on reference dependent preferences, we assume that in every period, each agent's utility depends both on current consumption and on their (or their parents', as a proxy of their upbringing) past consumption habits.

Then, assuming the planner to inherit a society with a history of past inequalities, we ask, what will the distribution be at the utilitarian allocation? How will it evolve over time? Under what conditions, if any, is Bentham's conjecture correct?

Proposition 1 proves that, starting from any initial distribution, if the common utility function is concave, then at the utilitarian optimum inequalities (in consumption and utility in a given period) decrease over time and disappear in the limit.[3] While it

[2]The literature is too vast for a comprehensive list of references. We refer the reader to the classic contributions by Rawls (1971), Smart and Williams (1973), Sen (1973, 1985), Sen and Williams (1982).

[3]This can be interpreted as the formal proof of a conjecture originally formulated by Layard (1980).

is tempting to interpret this result as confirming Bentham's conjecture – as changes in habits lead to a convergence in consumption and utility, – we shall argue that this is not the only, or necessarily the most persuasive interpretation of Proposition 1. For, noting that the reference point is endogenous, and that the agents' utility functions are, in a relevant sense, identical, our main result may be read as showing that at the utilitarian allocation equality does not obtain at any finite time, *even with identical agents*. We argue that this raises some new interesting questions for utilitarians.

This paper is related to various strands of literature, in addition to debates in normative economics and social choice. The influence of habits can be understood as a special case of of reference dependent preferences in line with Prospect Theory (Kahneman and Tversky, 1979), where the reference point evolves based on previous consumption levels. Hence, this paper provides a link between the literature on distributive justice and Prospect Theory (Kahneman and Tversky, 1979; Koszegi and Rabin, 2006, 2007), and models of habit formation (Fisher, 1930; Duesenberry, 1949) used in decision theory (Hammond, 1976), macroeconomics (Ryder and Heal, 1973; Stigler and Becker, 1977; Boyer, 1978; Becker and Murphy, 1988; Carroll et al., 2000) and finance (Constantinides, 1990; Abel, 1990, 1999; Ferson and Constantinides, 1991).

However habits may also be interpreted in a more general sense to include one's history, upbringing, and family background (or socio-economic status), which determine one's norms and expectations. At a broad conceptual level, these intergenerational effects are reminiscent of the emphasis on family circumstances that is central in the modern theory of equality of opportunity (Fleurbaey, 1995; Roemer, 1998; Roemer and Veneziani, 2004; Fleurbaey, 2008; Fleurbaey and Maniquet, 2011). In this paper, how-

Considering the utilitarian allocation of a fixed amount of income in a static economy in which agents care about both actual and *expected* income, Layard (1980) proved that if expectations depend on past incomes, then "yesterday's rich should have higher net incomes than if all had the same expectations" (Layard, 1980, p.746). However, he conjectured that if one allowed for expectations to adjust over time towards the level of income actually experienced and "there were no time discounting", then the distribution of a constant amount of income "should eventually become equal" (Layard, 1980, p.746). We explicitly consider the dynamic optimisation programme and precisely characterise the conditions under which this conjecture is correct. Thus, among other things, we show that time discounting is irrelevant for the result.

ever, we shall not assume an egalitarian planner to begin with but rather inquire on the conditions (if any) that may lead to equality at the utilitarian solution, whatever the agents' initial circumstances.

Our model is conceptually analogous to those exploring economies with cake-eating consumers (Hotelling, 1931; Gale, 1967; Dasgupta and Heal, 1974; Kemp, 1976; **?**; Adler and Treich, 2017). Unlike in this literature, the resource to be allocated is renewable (a new cake appears in every period), we introduce habits, and our focus is on distribution (rather than intertemporal optimisation, optimal savings, and the optimal consumption of non-renewable resources).

Finally, the paper also speaks to the recent literature on behavioural welfare economics (Bernheim and Rangel, 2007; Fleurbaey and Schokkaert, 2013; Infante et al., 2016; Mandler, 2020, among others). However, unlike in these contributions, we consider a rather mild deviation from the standard model with rational agents – namely, habits – and we explicitly consider normative issues in a dynamic context.

2 The main result

Consider an infinitely-lived economy with $N \geq 2$ households, which can be interpreted either as N infinitely-lived individuals or as an infinite number of individuals, each living for one period and belonging to N family lines. As we are interested in the distributive properties of utilitarianism in a dynamic economy with reference dependent preferences, we will abstract from production and growth, and consider a pure allocation problem with a fixed population size.

In each period, a divisible consumer good must be shared among the N agents. Following Koszegi and Rabin (2006, 2007), we assume that individuals care about both absolute and relative consumption. To be specific, at time $t = 0, 1, 2, \ldots$, individual utilities $u^i = u^i\left(c_t^i, r_t^i\right)$, $i \in \{1, \ldots, N\}$, depend on a weighted average of current consumption, c_t^i, and consumption relative to the reference point r_t^i. We assume that

agents have the same per-period utility function, $u^i = u : \mathbb{R}^2 \to \mathbb{R}$, for all $i \in \{1, \ldots, N\}$ but their reference points depend on past history and may differ, so that individuals may have heterogeneous preferences over current consumption. We assume that u is twice differentiable, strictly increasing in current consumption, and strictly concave.

Formally, the utility of individual $i \in \{1, \ldots, N\}$ at time t, is given by

$$u[(1-\alpha)c_t^i + \alpha(c_t^i - r_t^i)] = u(c_t^i - \alpha r_t^i) \tag{1}$$

with $\alpha \in (0, 1)$. Following the habit formation literature (e.g. Carroll (2000)), the reference point is a weighted average of the history of previous consumption:

$$r_{t+1}^i = \lambda c_t^i + (1-\lambda) r_t^i, \tag{2}$$

with $\lambda \in (0, 1]$, given a vector of initial reference points $\mathbf{r_0} = \left(r_0^1, r_0^2, \ldots, r_0^N\right)$ and

$$\sum_{i=1}^{N} r_0^i = 1,$$

and $r_0^i \geq 0$, for all i. At any t, let $\mathbf{c_t} = \left(c_t^1, c_t^2, \ldots, c_t^N\right)$. The utilitarian planner solves

$$\max_{\{\mathbf{c_t}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \left[\sum_{i=1}^{N} u(c_t^i - \alpha r_t^i)\right], \tag{MP}$$

subject to (2), and

$$\sum_{i=1}^{N} c_t^i = 1, \ \forall \ t, \tag{3}$$

given r_0^i for all i, with $\beta \in (0, 1]$.

Proposition 1. *At the solution of MP, for $t \to \infty$, $c_t^i \to \frac{1}{N}$, for all $i \in \{1, \ldots, N\}$.*

Proof

1. The solution of the MP is also the solution of the following Bellman equation:

$$V(\mathbf{r_t}) = \max_{\mathbf{c_t}} \left[\sum_{i=1}^{N} u(c_t^i - \alpha r_t^i) + \beta V(\mathbf{r_{t+1}}) \right], \tag{4}$$

subject to

$$r_{t+1}^i = \lambda c_t^i + (1-\lambda) r_t^i,$$

and

$$\sum_{i=1}^{N} c_t^i = 1, \ \forall \ t.$$

The FOCs are

$$u'(c_t^i - \alpha r_t^i) - u'(c_t^N - \alpha r_t^N) = \beta\lambda[V_N(\mathbf{r_{t+1}}) - V_i(\mathbf{r_{t+1}})] \tag{5}$$

for $i = 1, \ldots, \ N-1$, where $V_i(\mathbf{r_t})$ is the partial derivative of the value function at time t, with respect to the reference point of individual i. The envelope conditions are

$$V_i(\mathbf{r_t}) = -\alpha u'(c_t^i - \alpha r_t^i) + (1-\lambda)\beta V_i(\mathbf{r_{t+1}}), \tag{6}$$

for $i = 1, \ldots, \ N$. Then, using the envelope condition of agent N, we get

$$V_N(\mathbf{r_t}) - V_i(\mathbf{r_t}) = \alpha \left[(u'(c_t^i - \alpha r_t^i) - u'(c_t^N - \alpha r_t^N) \right] + \beta(1-\lambda) \left[V_N(\mathbf{r_{t+1}}) - V_i(\mathbf{r_{t+1}}) \right]. \tag{7}$$

Then, by substituting the RHS of (5) into (7), we get

$$V_N(\mathbf{r_t}) - V_i(\mathbf{r_t}) = \beta[\alpha\lambda + 1 - \lambda] \left[V_N(\mathbf{r_{t+1}}) - V_i(\mathbf{r_{t+1}}) \right]. \tag{8}$$

Shifting (5) one period back, we get

$$u'(c_{t-1}^i - \alpha r_{t-1}^i) - u'(c_{t-1}^N - \alpha r_{t-1}^N) = \beta\lambda \left[V_N(\mathbf{r_t}) - V_i(\mathbf{r_t}) \right], \tag{9}$$

which given (8) and (5), gives the following Euler equation

$$u'(c_{t-1}^i - \alpha r_{t-1}^i) - u'(c_{t-1}^N - \alpha r_{t-1}^N) = \beta[\alpha\lambda + 1 - \lambda]\left[u'(c_t^i - \alpha r_t^i) - u'(c_t^N - \alpha r_t^N)\right]. \quad (10)$$

Note that $\alpha < 1$, so $\alpha\lambda < \lambda$ which also means that $0 < \beta[\alpha\lambda + 1 - \lambda] < 1$. Note that if

$$c_t^i - \alpha r_t^i = c_t^N - \alpha r_t^N, \quad (11)$$

for all t, then (10), is true. Let

$$F_t = u'(c_t^i - \alpha r_t^i) - u'(c_t^N - \alpha r_t^N).$$

If (11) does not hold, then given (10), F_t is strictly increasing and unbounded as for any $\delta > 0$, there exists a t for which $F_t > \delta$. Hence, the transversality conditions are not satisfied and (11) is the only possible solution.

2. The summation of (11) over i, gives

$$N(c_t^N - \alpha r_t^N) = \sum_{i=1}^{N}(c_t^i - \alpha r_t^i) = \sum_{i=1}^{N} c_t^i - \alpha \sum_{i=1}^{N} r_t^i.$$

But from the budget constraint $\sum_{i=1}^{N} c_t^i = 1$, for all t, which also means that given (2) also $\sum_{i=1}^{N} r_t^i = 1$, for all t, hence

$$c_t^N = \frac{1-\alpha}{N} + \alpha r_t^N \quad (12)$$

If we express (11) in terms of the reference point given (2), we get

$$r_{t+1}^i - (1 - \lambda + \alpha\lambda) r_t^i = r_{t+1}^N - (1 - \lambda + \alpha\lambda) r_t^N,$$

or equivalently,

$$r_{t+1}^i - r_{t+1}^N = (1 - \lambda + \alpha\lambda)(r_t^i - r_t^N). \quad (13)$$

Note that $1 - \lambda + \alpha\lambda < 1$, hence for $t \to \infty$, $r_t^i = r_t^N$ for all i. Hence, for $t \to \infty$, $r_t^i = \frac{1}{N}$. Then also given (12), for $t \to \infty$, $c_t^N = \frac{1}{N}$. Also, given (11), for $t \to \infty$, $c_t^i = \frac{1}{N}$, for all i. □

Proposition 1 shows that that consumption inequality between any two households is decreasing over time and disappears in the limit. The mechanism underlying this result is quite intuitive. From equation (11), it follows that at the utilitarian solution, in each period consumption inequality between households is lower than the difference between the respective reference points – the ratio of consumption inequality to the difference between reference points is equal to $\alpha < 1$.[4] Then, as consumption levels of one period feed into the reference points of the next, the households' consumption habits become more similar, which spurs the equalisation process further.

Proposition 1 thus seems to vindicate utilitarianism against egalitarian critics, and to confirm Bentham's conjecture. Even if the planner inherits a history of inequalities, leading to different habits, and thus to heterogeneous utility functions *over consumption*, at the utilitarian optimum inequalities in consumption (and thus habits) decrease over time and disappear in the limit. In addition to its theoretical relevance, this result would also lend empirical support to utilitarianism given the strong evidence of the pervasiveness of reference dependent preferences (Bateman et al., 1997; Bowles, 1998).

Yet, while this simple reading of Proposition 1 is legitimate, it is by no means the only plausible interpretation, and the implications of our analysis for utilitarianism are more nuanced than they may appear at first sight.

A preliminary point to note is that while inequalities in consumption (and welfare) *per period* vanish in the long run, a history of past inequalities still implies potentially large inequalities if one considers the welfare of the N infinitely lived agents, or of the N households, over their *whole lives*.

Perhaps more interestingly, two features of the model should be highlighted which

[4] It is worth noting in passing that while the parameter α plays a key role in the proof of Proposition 1, – the lower α the higher the speed of convergence, and the result does not hold if $\alpha = 1$, – the discount factor β is irrelevant to our conclusions, which continue to hold even if $\beta = 1$.

raise doubts on the simple interpretation of Proposition 1. First, although the history of past inequalities until $t = 0$ is given for the utilitarian planner, the dynamics of the agents' reference points at all $t \geq 1$ is actually endogenous. Second, at any time t, for a given set of (unequal) reference points, the agents' utility functions over consumption are indeed different. However, the basic structure of agents' utility function, as a function of *both* current consumption *and* consumption habits is exactly the same and, in this sense, one may argue that agents are structurally identical.

Indeed, given the endogeneity of the reference point at all $t \geq 1$, one may argue that agents living in each period after the first one are actually identical from the viewpoint of the utilitarian planner. To see this, consider the extreme case with $\lambda = 1$, in which the habit stock coincides with consumption in the previous period (or by the previous generation). In this case, the planner could obliterate both consumption inequalities *and* any effect of consumption habits on current preferences in a single stroke. Hence, in all periods after the first agents do have exactly the same utility function from the perspective of the utilitarian planner at $t = 0$. Yet, by Proposition 1, equality does not obtain, except at the limit, in the very long-run.

But then, perhaps counterintuitively, Proposition 1 may be interpreted as contradicting the utilitarians' counterarguments against egalitarian critics. For it may be argued that, even though agents are, in a relevant sense, identical, the utilitarian optimum does not entail equality of consumption, or utility, at any finite time. Utilitarianism thus seems inconsistent with a concern for equality even in what has long been considered the most favourable scenario.

3 Conclusions

In this paper, we have examined the distributive implications of utilitarianism in a dynamic context, when individual preferences are reference-dependent. Assuming that all agents have the same utility function – which depends on current consumption

and past consumption habits – but the utilitarian planner inherits a history of past inequalities, we have shown that at the utilitarian optimum inequality decreases over time but it does not disappear except at the limit.

This result raises an interesting conceptual issue. If agents' preferences are interpreted as being fundamentally different, due to different reference points, then our result confirms – indeed strengthens – the relation between utilitarian thought and egalitarian principles: albeit not immediately egalitarian, the utilitarian allocation is equalising even in the presence of heterogeneity arising from past inequalities.

If, however, one notes that, except for the very first period, consumption habits are endogenous and agents have utility functions with the same structure, then our result casts doubts on the relation between utilitarianism and egalitarianism highlighted by many economists and philosophers: even if agents have the same, concave utility functions, at the utilitarian allocation inequalities persist for an indefinitely long time.

Thus, the implications of our result for debates on utilitarianism and distributive justice hinge, at least partly, on a conceptual issue, namely what it means for agents to have the same utility functions in a dynamic economy with inherited inequalities and endogenous preferences. We leave this issue for further research.

References

Abel, A. B., 1990. Asset prices under habit formation and catching up with the joneses. American Economic Review 80 (2), 38–42.

Abel, A. B., 1999. Risk premia and term premia in general equilibrium. Journal of Monetary Economics 43 (1), 3–33.

Adler, M. D., Treich, M., 2017. Utilitarianism, prioritarianism, and intergenerational equity: A cake eating model. Mathematical Social Sciences 87, 94–102.

Bateman, I., Munro, A., Rhodes, B., Starmer, C., Sugden, R., 1997. A Test of the

Theory of Reference-Dependent Preferences. The Quarterly Journal of Economics 112 (2), 479–505.

Becker, G. S., Murphy, K. M., 1988. A theory of rational addiction. Journal of Political Economy 96 (4), 675–700.

Bentham, J., 1838–1843a. "Pannomial Fragments", In: The Works of Jeremy Bentham: Published under the Superintendence of His Executor, John Bowring. Volume 3. Tait, Edinburgh.

Bentham, J., 1838–1843b. "Principles of the Civil Code", In: The Works of Jeremy Bentham: Published under the Superintendence of His Executor, John Bowring. Volume 1. Tait, Edinburgh.

Bernheim, B. D., Rangel, A., 2007. Toward choice-theoretic foundations for behavioral welfare economics. American Economic Review 97 (2), 464–470.

Bowles, S., 1998. Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions. Journal of Economic Literature 36 (1), 75–111.

Boyer, M., 1978. A habit forming optimal growth model. International Economic Review 19 (3), 585–609.

Carroll, C. D., 2000. Solving consumption models with multiplicative habits. Economics Letters 68 (1), 67–77.

Carroll, C. D., Overland, J., Weil, D. N., 2000. Saving and growth with habit formation. American Economic Review 90 (3), 341–355.

Constantinides, G. M., 1990. Habit formation: A resolution of the equity premium puzzle. Journal of Political Economy 98 (3), 519–43.

Dasgupta, P., Heal, G., 1974. The optimal depletion of exhaustible resources. The Review of Economic Studies 41, 3–28.

Duesenberry, J., 1949. Income, Saving and the Theory of Consumer Behavior. Oxford University Press, New York.

Ferson, W. E., Constantinides, G. M., 1991. Habit persistence and durability in aggregate consumption: Empirical tests. Journal of Financial Economics 29 (2), 199–240.

Fisher, I., 1930. The Theory of Interest. Macmillan, New York.

Fleurbaey, M., 1995. Three solutions for the compensation problem. Journal of Economic Theory 65, 505– 521.

Fleurbaey, M., 2008. Fairness, Responsibility and Welfare. Oxford University Press.

Fleurbaey, M., Maniquet, F., 2011. A Theory of Fairness and Social Welfare. Cambridge University Press.

Fleurbaey, M., Schokkaert, E., 2013. Behavioral welfare economics and redistribution. American Economic Journal: Microeconomics 5 (3).

Gale, D., 1967. On optimal development in a multi-sector economy. The Review of Economic Studies 34, 1–18.

Hammond, P., 1976. Endogenous Tastes and Stable Long-Run Choice. The Journal of Economic Theory 13, 329–340.

Hotelling, H., 1931. The economics of exhaustible resources. Journal of Political Economy 39, 137–175.

Infante, G., Lecouteux, G., Sugden, R., 2016. Preference purification and the inner rational agent: a critique of the conventional wisdom of behavioural welfare economics. Journal of Economic Methodology 23 (1), 1–25.

Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk. Econometrica 47 (2), 263–291.

Kemp, M., 1976. How to eat a cake of unknown size. In: Kemp, M. (Ed.), Three Topics in the Theory of International Trade. North Holland, Ch. 23.

Koszegi, B., Rabin, M., 2006. A Model of Reference-Dependent Preferences. The Quarterly Journal of Economics 121 (4), 1133–1165.

Koszegi, B., Rabin, M., 2007. Reference-Dependent Risk Attitudes. American Economic Review 97 (4), 1047–1073.

Layard, R., 1980. Human Satisfactions and Public Policy. The Economic Journal 90, 737–750.

Lerner, A. P., 1944. The Economics of Control: Principles of Welfare Economics. The Macmillan Co.

Mandler, M., 2020. Distributive justice for behavioural welfare economics. The Economic Journal 130, 1114–1134.

Rawls, J., 1971. A Theory of Justice. Belknap Press of Harvard University Press.

Roemer, J. E., 1996. Theories of Distributive Justice. Harvard University Press.

Roemer, J. E., 1998. Equality of Opportunity. Harvard University Press.

Roemer, J. E., Veneziani, R., 2004. What We Owe Our Children, They Their Children,.... Journal of Public Economic Theory 6 (5), 637–654.

Ryder, H. E., Heal, G. M., 1973. Optimum growth with intertemporally dependent preferences. Review of Economic Studies 40 (1), 1–33.

Samuelson, P. A., 1964. A.P. Lerner at Sixty. The Review of Economic Studies 31 (3), pp. 169–178.

Sen, A., 1973. On Economic Inequality. Oxford University Press.

Sen, A., 1985. Commodities and Capabilities. Oxford University Press.

Sen, A., Williams, B., 1982. Utilitarianism and Beyond. Cambridge University Press.

Sidgwick, H., 1907. The Methods of Ethics. Macmillan, London.

Smart, J., Williams, B., 1973. Utilitarianism: For and Against. Cambridge University Press, Cambridge.

Stigler, G. J., Becker, G. S., 1977. De gustibus non est disputandum. American Economic Review 67 (2), 76–90.