
PAST LEARNINGS

Fairness, the Ultimatum Game and Public Policy

George Yarrow

No. 2

June 2024

REGULATORY POLICY INSTITUTE

The Past Learning series of papers is predicated on the twin propositions that the great bulk of valuable knowledge about the principles of political economy is to be found in past research and that much of this knowledge is forgotten or ‘written off’ in consequence of an undue preoccupation with the particular issues and thinking of the day. Whilst the search for new incremental knowledge is admirable, the amnesia is not, because it implies a shrinking base to which the increments can be added. In consequence the total value of the associated, collective human capital applied to economic policy problems progresses more slowly than it could and is also prone to significant periods of regression. The contrast with the physical sciences in this regard is stark.

The aim of the series is to provide a mild corrective to the amnesia by re-presenting some of the older wisdoms and drawing attention to their direct and immediate value in application to today’s policy challenges.

First published by the Regulatory Policy Institute
9 Royal Belgrave House, Hugh Street, London SW1V 1RR
www.rpieurope.org

June 2024

© Regulatory Policy Institute

Fairness, the Ultimatum Game and Public Policy¹

Introduction

The notion of ‘fairness’ is widely referenced in public policymaking and enforcement, but with no settled meaning. What we see is incoherent application of the notion across economic contexts, a form of policy disorder with which we are well familiar. Moreover, the disorder (‘entropy’) appears to be increasing over time.

One observable tendency is toward use of references to fairness to justify decisions that are, in reality, driven by other factors, but which may be highly challengeable if those factors were transparently exposed. The usage of the word ‘fair’ then serves to buttress Politicians’ Logic: something must be done; this is something; therefore we will do this”. It offers a source of lazy answers to questions like ‘Why must something be done?’ and ‘Why this?’ Answers: ‘because the existing situation is unfair’ and ‘it will be fairer to do this’.

Away from the rhetorical tricks, issues of fairness *are* substantive matters of relevance to governmental decisions. They therefore merit close examination.

Developing greater coherence and order in governmental decision making does not entail settling on a single meaning of fairness. Meanings vary with context. Commenting on the words fair, just and equitable in a box labelled ‘awkward synonyms’ the New Oxford Thesaurus says: “*These words express judgments springing from a belief that everyone should get what they are entitled to.*” In relation to the word ‘fair’ it goes on to say that “*It is an emotive word, expressing a deep-seated view of what is basically right, coupled with indignation if it is not achieved. ... Fair is used not only to judge people and their behaviour but also to complain about impersonal bad luck.*”

These comments point to a potential minefield into which over-casual, unthinking recourse to notions of fairness can lead us. We may *complain* about impersonal bad luck, but are we typically *indignant* about it? What precisely do people think they are entitled to, and on what bases are such expectations formed? In most sets of circumstances in which the word ‘fair’ is used, there are no unambiguous, black-letter specifications of entitlement to determine matters: different people may have very different views of their own, individual, perceived entitlements.

Rather than simply leaving the meaning of fairness to be addressed by government authorities on an ad hoc, discretionary basis, I think the challenge for decision-makers is to develop stable mappings that link/map different meanings of ‘fairness’ to types of context. In this way affected economic agents can, on the basis of their own assessments of the context, make more informed judgments about the likely conduct of the relevant authorities. This would be conducive to reductions in policy uncertainty, a major bugbear for all who have to make investment decisions, whether in physical, financial or human forms of capital.

¹ A developed text version of a presentation, ‘*Perspectives on Fairness*’, given at a Regulatory Policy Institute conference in Westminster on 3 May 2017.

It is instabilities in these mappings that cause the major problems, not the absence of some single, all-encompassing interpretation of the notion of fairness; and the major instabilities tend to arise from the clash of claimed entitlements of political partisans, the interest groups that importune government on the basis of interpretations of 'fairness' that are convenient and advantageous for their own, particular interest. The groups include politicians, civil servants, business interests, financial interests, single-issue groups, ideological partisans, and academics/intellectuals. The relative influence of each waxes and wanes over time. It's a recipe for policy disorder.

What follow are some thoughts on how to reduce the entropy.

Domains of judgment

Let's start the explorations with questions about the nature of what precisely it is that is being judged fair or unfair. i.e. about what might be called the 'domain' of the judgments at issue. Candidate domains include:

- The *outcomes/consequences* of economic processes as measured by numerical variables such as the quantities and qualities of products/services supplied and purchased, prices of product/services, relative levels of incomes, relative levels of wealth, etc. (Issues in this domain often attract the label 'distributive fairness'.)
- The *conduct* or behaviour of economic agents (and note the reference to behaviour in the Thesaurus citation above).
- The relevant socio-economic *rules* established or evolved to guide conduct/behaviour, whether they are formal (laws or regulations) or informal (e.g. social and cultural norms and conventions), including rules for the enforcement of a given set of rules and for changing those rules over time.

Read in reverse order, these categories are congruent with a familiar schema used in economic assessments of markets for both regulatory and competition law purposes, which is based on distinctions between market structure (a state variable), commercial conduct (a flow variable) and market performance (measures of outcomes/consequences).

In practice, contextual assessments of 'fairness' might involve different combinations of the identified domains. Thus, consideration of *procedural fairness* might involve assessment of both the relevant rules in play and the conduct in their implementation/enforcement. Similarly, notions of *fair play* might involve assessments of both actual conduct within a given set of rules and judgments concerning the appropriateness of the rule-set itself in the given context.

When they are not trying to evade fairness issues completely, for example by concentrating all efforts on the notion of 'efficiency' (estimated by totting up perceived costs and benefits), economists tend to have a proclivity to focus very heavily on the first of the listed domains, i.e. outcomes/consequences, with a secondary interest in the third (the rule utilitarianism examined by J.S. Mill being an example).

This is arguably a reflection of the consequentialist utilitarianism that has dominated economic thought for most of the post-classical period. Thus, utility functions representing the preferences of economic agents are generally defined over outcomes/consequences, either directly or indirectly. They are most usually specified in terms of the outcomes for the relevant

individual (quantities of goods and services consumed, time available for leisure, etc.), but inter-personal considerations can be accommodated by inclusion of the outcomes for others. Either way, it is outcomes ‘all the way down’ (to a ‘reduced form’ expression).

In contrast, when explaining the meaning of ‘fairness’ the Thesaurus citation on ‘awkward synonyms’ explicitly references judgments about behaviour (conduct), and it takes but little reflection on quotidian contexts to be convinced that conduct is a ubiquitous domain of judgment. We are constantly making judgments about the conduct of others and, in more reflective moments, about our own past and prospective conduct.

School rules

Here is one, illustrative example. It concerns the rules of a South Yorkshire school – the Percy Jackson Grammar School (PJGS) in the late 1950s and earlier parts of the 1960s.

There were twenty school rules in total, mostly covering sets of specific behavioural contexts, like running in corridors and students’ interactions with milk bottles. Rule 1, however, was a meta-rule (or principle or standard) explicitly directed at conduct across all school-relevant matters. It is almost certainly the most memorable for alumni, very many of whom will have written it out hundreds of times as punishment for some or other misdemeanour. It stated:

“In all conduct the aim should be due consideration for others.”

It’s a nice piece of rule-making. I have no information on attribution of its origins, but the foundation of the school itself, in 1939, is largely attributable to one J W ‘Jimmy’ Lane JP, an alderman of the West Riding and a check weighman at Brodsworth Colliery, whose pit-head complex lay a little to the west of the school’s grounds (and was traversed many thousands of times by more senior students on outward and return parts of the path of their cross-country runs). The school was named after another West Riding alderman, Percy Jackson, a noted advocate for education on the County Council in the 1920s and 1930s.

The ‘all’ in Rule 1 explicitly signals its comprehensive (in-all-contexts) nature, and the most interesting word is ‘due’. It calls for individual judgments by both students and teachers, the latter in their capacity as enforcers of the rule. Without it, the injunction would range from being trite² to a call for a saintliness well beyond reasonable expectations for teenage students. It embeds a proportionality principle, although (wisely) not one that is rigidly specific in form.

In relation to the Thesaurus’s references to ‘entitlement’ and ‘indignation’ in teasing out the meanings of the awkward synonyms, it can be noted that, if the rule is accepted as ‘basically right’ – or alternatively and less strongly expressed, if it is accepted as legitimate -- what is created is a collective set of expectations about entitlements. The normative injunction of Rule 1 establishes an obligation on each individual student to give due consideration to others, but, simultaneously and in consequence of the collective reach of the obligations created, it induces an expectation or ‘sense of entitlement’ on the part of each student that others will give her/him

² See the first sentence of Adam Smith’s ‘Theory of Moral Sentiments’. Today a cognitive psychologist might say that consideration for others is hard wired into the human brain, a consequence of the evolution of a social animal species. A Scottish Enlightenment expression of the same point is that consideration for others is an aspect of human nature (the study of which was a major passion among the literati of that period).

due consideration in their own conduct. In a nutshell, each and all can expect (or to some degree feel entitled to) *reciprocity*.

It can be noted that the examination of this set of rules can be extended out to encompass obligations and entitlements as between students and teachers, but, for current purposes, the main point is simply that the primary domain of judgment is not outcomes/consequences, but rather conduct. Economic analyses engaging with fairness issues solely on the basis of judgments about consequences turn a blind eye to this, as the experimental evidence on the Ultimatum Game (“UG”) illustrates.

The ultimatum game

The ultimatum game has been ranked as one of the two most studied ‘games’ in experimental economics (the other being the prisoners’ dilemma). In a nutshell, its structure is as follows:

1. Person P (the proposer) is given charge over a sum of money and asked to offer a fraction of it to person R (the responder).
2. P and R are anonymous to each other.
3. The game is played only once.
4. R is made aware of the amount over which P has charge.
5. R decides whether or not to accept the amount offered.
6. If R accepts the offer, R gets the amount offered and P gets the rest.
7. If R rejects the offer, both parties get nothing.

Thus, if P is given charge of £100, offers £25 and R accepts, the outcome is £75 to P and £25 to R. If, on the other hand, R rejects the offer, the outcome is £0 to P and R alike.

On a narrow, individualistic, utilitarian analysis of R’s decision, rejection of a 75%/25% offer appears puzzling: it suggests a preference for £0 over £25. Yet the experimental results are clear: a significant fraction of 75%/25% offers are rejected. Further, the greater the asymmetry in offers, the higher is the rejection rate.

What the ultimatum game evidence establishes is that a standard working assumption that appears systematically in the teaching of economics – that in a singular exchange transaction with an otherwise unconnected counterparty, the aim will be to seek the best feasible, individual outcome (the *homo economicus* assumption) – is not even approximately true.

Among economists of old, such falsification of a major assumption would, by and of itself, have been a fatal analytic flaw for, say, Joan Robinson (mantra: “always interrogate the assumptions”). For a positivist like Milton Friedman it would not have been, so long as the reasoning leads to a good account (prediction) of the outcomes of interest. Since the latter condition is not satisfied by the UG’s experimental results, the subversive nature of the results is the same, irrespective of the difference in economic philosophies.

When an important belief, theory, working hypothesis or assumption/presumption is challenged by disjunctions with evidence, there is a natural tendency of scholars to look first

at possible variations in auxiliary assumptions that might serve to sustain a privileged, prior conceptual framework, and so it has been with the ultimatum game. For example, the specification of the individual objective functions can be adjusted to allow the individuals involved to have preferences over outcomes for others, not just themselves. On that basis, if there is a general aversion to asymmetric outcomes, the experimental results can be explained in abstract and can be fine-tuned to the data by the introduction of some measure of the ‘degree of aversion to asymmetry’ and of parameters of its distribution among individuals.

Notwithstanding the validity of an assumption that, as social animals, humans do in fact exhibit an interest and concern about the well-being of others, which therefore has direct effects on their own feelings of well-being – Adam Smith notably asserted it in the very first sentence of the *Theory of Moral Sentiments* – I am unconvinced of the merits of this line of thinking. Auxiliary hypotheses may be adjusted to explain the outcomes in one particular context, e.g. playing the UG, but a question to ask is whether those very same adjustments work in other, related contexts? The evidence suggests not.

Consider, for example, the Dictator Game, in which the set-up is the same as for the ultimatum game, except that P has unilateral powers to determine the allocation of the money (the pie to be divided). R has no say in the matter. In a sense, P ‘owns the pie’.

If the UG results are driven solely by considerations of ‘fairness of outcomes’ it is to be expected that the distributions of offers observed in playings of the DG will be very similar to those of the UG; but the evidence indicates that they are not.

A wider perspective on the ultimatum game

Something else must be going on and it’s pretty obvious what it is: in the UG P does not ‘own the pie’. There is no unilateral power of P to determine the outcome. Indeed, the very existence of the pie itself depends on R consenting to P’s offer. That *power* (here the ability to influence final monetary outcomes) is not as asymmetrically distributed in the UG than in the DG is a plain fact, to be taken into account in seeking to understand and explain the empirical results from the UG.

P knows that, to get any payment at all, the offer made must be accepted/approved by R. It is therefore obviously advisable for P to give *due consideration* to the likelihood that an offer will be rejected. To do this, P must put herself/himself ‘in the shoes’ of R. The questions for P are: how will R ‘see’, feel about, and react to this offer?

The key point I want to make is that a ‘representative’ R will not be interested only in the money on the table. It will be a factor, but not the only factor. R will observe that the offer to be evaluated is determined by the *conduct* of P (it is a matter of choice for P, and he/she could have chosen differently). Evaluation of the offer will therefore necessarily, explicitly or implicitly, entail the forming of some judgment about the appropriateness of the conduct, and here we return simultaneously to the *Theory of Moral Sentiments* (a book devoted to the

analysis of such judgments and PJGS School Rule 1 (“*In all conduct the aim should be due consideration for others*”).

Has P given due consideration to R? In particular, has P given due consideration to the *power* of R to reject the offer – and note that possession of that power is the singular thing that P knows about R – thus denying P her/his own share of the money on the table? If not, that will be a negative factor in R’s balance sheet when making a judgment on P’s offer. In Smith’s terminology, there will be disapprobation of P’s conduct, which, to put it at its simplest, is not good for R’s feelings and consequently carries a negative value.

That does not settle the accept/reject decision of course. The balance sheet may still lead to a decision to accept the offer: the positive valuation of the money available might exceed the negative valuation attached to the disapproval of the conduct. Different Rs will make different assessments, but across samples as a whole there is a strong, negative relationship between the asymmetry of offers and the number of rejections.

The argument can also be taken one level deeper by noting that (a) in making an accept/reject decision, it is a characteristic of human nature that R may reflect on and judge her/his *own* conduct and (b) this self-judgment has a retrospective, as well as an ‘in the moment’, dimension. Questions here might be: Faced with a P behaving badly, will I later regret accepting this highly asymmetric offer? By choosing to benefit from inappropriate conduct, am I, in effect, endorsing this conduct?

Application to public policy

Public policy, including all formal systems of regulation, is an exercise in rule making and rule enforcement, and it engages with fairness questions in each of the three domains identified at the outset. Are the rules themselves fair? Is enforcement conducted fairly? Are the expected outcomes fair?

The policies promised by politicians at election times can be likened to offers made to the public, on which each individual member of the electorate can make a judgment and can make a response via a cross on a ballot paper. Indeed, some politicians refer explicitly to their “retail offers”, though what the word ‘retail’ is doing in that expression is unclear, suggesting as it does that those who use this term self-identify as middle managers of grocery stores?

The offer/response pattern is therefore akin to that in the UG, although the offers are fuzzier (they are promises only, and promises that are frequently not fulfilled); are much wider in scope (than in the UG); and it is a collective (an aggregated number of votes), not an individual, response that matters in determining outcomes.

Between elections voters are, however, constantly making judgments about the conduct of politicians, which is continuously brought to their attention in the media and which includes conduct in fulfilling promises. When the time for the ballot paper comes around, approval or

disapproval, will depend on a whole back history of conduct, not simply the electoral offers/promises made.

Though they do not affect the value of the UG ‘gestalt’ in thinking about policy determination, there are two important features of the contextual differences that it is important to note:

1. The asymmetry of power between the offerors and the *individual* responders in the democratic process is much greater than in the lab experiments on the UG;
2. Policy offers have effects on collectivities of individuals, so the judgments of government conduct will encompass the expected effects of the proposals on others, as well as on the individual voter. If the offer promises a pie, the distribution of that pie among different, identifiable groups of individuals can be expected to be a factor that influences the degree of approbation or disapprobation afforded. Particularly in political systems where ‘equality before the law’ is an embedded, normative principle (establishing a sense of justifiable entitlement), significant disparities in conduct toward different groups or types of voter are liable to trigger the Thesaurus’s ‘indignation’ among offerees. It is disparities in conduct and outcomes between comparable offerees that rankles most, right down to a highly localised level. Are the folk three doors up the street, in a very comparable socio-economic position to us, getting more from the government (or are putting less into the Exchequer) than we are? Does the state give greater consideration to their interests than to ours?

The significance of the first feature comes from another, very widely shared normative principle: greater power entails greater responsibility. Articulations of it are to be found in the Spiderman comics and movies, but also in a much earlier, widely promulgated literature (e.g. St Luke 12:48 “*For unto whomsoever much is given, of him shall be much required: and to whom men have committed much, of him they will ask the more.*”) There is a good articulation of it by Churchill in a 1906 speech in the House of Commons, in the context of Britain’s responsibilities for ensuring racial justice in the colonies (South Africa in particular): “*Where there is great power there is great responsibility, where there is less power there is less responsibility, and where there is no power there can, I think, be no responsibility.*” This brings out an underlying notion of proportionality that is expressed in the PJGS school rules by the word “due”.

In a nutshell, given the power of government, much is to be expected from it in the way it conducts its business. If expectations are not met, that is akin to a breach of PJGS Rule 1, a breach of a governing, normative principle. Inadequate consideration has been given to relevant others.

The significance of the second feature derives from what is called ‘horizontal equity’ (dealing with people in similar situations in similar ways). ‘Equality before the law’ is an example of a horizontal equity principle, as applied to a particular set of interactions between the individual and the state. The appearance of the concept in economics is marginal, largely to be found only

in discussions about taxation. The focus tends to be much more on issues of ‘vertical equity’, for example on disparities in income and wealth across the populace, on which much has been written. However, these latter disparities do not generate the same sense of strong entitlement (and hence indignation and anger at violations of entitlement) as ‘horizontal’ disparities.

Sociologists and psychologists do better with their concept of a reference group, the set of people with whom comparisons are highly salient for an individual. One ubiquitously found type of reference group comprises other people seen to be in similar socio-economic positions who are either known personally to the individual or are brought to their attention for some reason or other, for example via the media.

Policymakers ignore this dimension of ‘fairness’ at their peril. As the UG evidence has shown, simply arguing that a particular policy offer will increase the size of the pie is, even if credible, far from rejection-proof.

Immigration policy and the Brexit referendum as examples

There have been several studies of the economic effects of immigration, usually done with an eye on policy implications, but they tend to focus on estimating *aggregated* benefits and costs. Aside from the obvious difficulties of putting numbers to these things, from a policy perspective there are two deficiencies here: (a) since few would doubt that some level of immigration is beneficial to an economy, the salient policy issues concern marginal benefits and costs (is immigration too high, or not high enough?), and (b), put loosely, they focus on the size of a pie that is to be divided. In relation to (b) the empirical results from the UG serve as a warning against being riveted by ‘size-of-pieism’, and it is on these that I will focus.

A major finding of the UG research is that the offer/response curve is relatively insensitive to the size of the pies available to the participating subjects of the research. That can come as a surprise to consequentialist utilitarians who might expect that, with a pie that is doubled in size, there would be fewer rejections of, say, a 75%/25% offer, because the monetary cost of rejection by the Rs will cost them twice as much.

It is not a surprise though if, as argued, the rejection rate is largely determined by judgments made about the appropriateness of the conduct in the presence of an asymmetry in power between P and R. The *relative* power of the Ps and Rs is unchanged by the size of the pie: if the pie is made twice as large, R may suffer a doubled monetary penalty if the offer is rejected, but, by the same token, the loss inflicted on P by use of R’s power of veto will also have doubled: P also has more to lose and both parties know that.

To link this back to another practical policy issue, viewed in the light of the UG it is easy to see one of the problems with the ‘Project Fear’ campaign in the 2016 referendum on EU membership: the emphasis was on the size of the pie. Due consideration was not afforded to the kinds of factors that the UG indicates are in play when judgments about the conduct of those possessed of greater power are made (and it can be noted in passing that the same point

can be made about advocates of free trade when they also focus their rhetoric too heavily on the size of the pie).

Electors have much less power of influence than ministers, but they do have slightly more power than non-electors by virtue of their votes. They can reasonably expect that fact to be recognised and their interests taken into account to a degree that is greater than, and certainly no less than, the interests of non-electors. As Burke put it to the electors of Bristol: *“To deliver an opinion, is the right of all men; that of constituents is a weighty and respectable opinion, which a representative ought always to rejoice to hear; and which he ought always most seriously to consider.”*

Immigration had become a major issue in 2015 – it received substantial coverage in the media across Europe – and new immigrants had become a more significant reference group for UK electors. Horizontal equity issues were in play. Even if the pie would be larger in the event of remaining in the EU (itself a contested issue), the conduct of government in using its powers was a serious matter awaiting the judgment of the electorate. How, for example, would any net benefits (increments to the pie) be divided between members of the electorate and new arrivals?

Putting it delicately, there are strong grounds for suspicion that, in per capita terms, the answer to that question is that the division has been skewed rather heavily to new arrivals, even possibly to the extent of them reaping more than 100% of any net benefits. Hard evidence is scant, because the relevant questions concerning horizontal equity and the appropriate use of power have been largely ignored. That, by and of itself, is an indictment of a lack of due consideration that the Ps in this story had given to the Rs.

In a Judicial Review of administrative conduct, a unit of government that did not ask questions that were highly salient to its conduct/decisions, and/or had not troubled itself to gather evidence required to answer those questions, could be judged to have been acting ‘irrationally’ or ‘unreasonably’. Yet so divisive have immigration issues become that it is those responders who would reject the current policy approach who have been so labelled, and worse: ‘undereducated’, ‘stupid’, ‘contemptible’, and the rest. What I am suggesting is that it is the proposers, not the responders, who have been acting irrationally/unreasonably – by ignoring falsifying evidence in making an implicit *homo economicus* assumption when ‘offering’ policies to the electorate.

And once the high salience of horizontal equity issues for the public is recognised, it should be rather easier to develop immigration policies that are less divisive and more fit for purpose than current arrangements. For example, a cap-and-trade system for residency is one approach (to flattening a perceived asymmetry) which any government *“ought ... most seriously to consider.”*