Class on Elster

1. Some Main Themes

- *Norms and Rationality*: In understanding individual action—including, but not confined to, individual contributions to cooperative activities that produce common benefits—we need more than the notion of rational decision making. Rational decision-making is a matter of choosing the action that has the best expected results, as specified by an agent's beliefs and rankings of outcomes (assuming that the beliefs and rankings of outcomes meet certain minimal conditions of coherence and consistency). Main idea is that individuals also follow norms, where norms are understood as standards that serve as guides to conduct, guides that provide agents with more or less determinate instructions about how to act when situations present certain features; norms are not simply patterns of behavior that conduct instantiates. It is a norm, in the relevant sense, that people take hats off when they are in church and men cover their heads in temple; it is not a norm in the relevant sense that people put up umbrellas when it rains, even if that is as much a regularity of conduct as the previously mentioned cases. Main feature of a norm is that it provides a standard for assessing and guiding action, without reference to the results of the action. There is an interesting question about why we take ourselves to have reasons for acting in a certain way because the action itself has certain characteristics, and irrespective of its consequences. But whatever the explanation, norms seem to be important elements of practical thought.

- *Norms and Explanation*: In emphasizing the importance of norms—of norm-guided conduct—Elster offers micro-foundations of the arguments that we find in Bull on international politics and Allison/Zelikow on the role of organizations in foreign policy, as well as in Rawls on the idea of a well-ordered society, as a society whose members have a sense of justice founded on fairness—a society whose members are, as Rawls puts it, reasonable, and not only rational. Each of these views shares the idea that norms are not simply regularities of conduct or standards for assessing conduct as good or bad, right or wrong—they are not simply important from the point of view of a judge—but that they are practically important: agents internalize norms, and as such those norms help (in some measure) to explain the conduct of agents: their actions, emotions (guilt, shame for violating norms, pride for compliance), their judgments (both internal and expressed about the conduct of others).

- *Not Cheery, Naïve Story*: From the fact—assuming it is one—that norms are among the factors that explain conduct, no conclusions can be drawn about whether people are—because they follow norms—more admirable

or better at solving problems or more cooperative than they would otherwise be. Norm-guided conduct is a matter of acting in accordance with rules or standards that provide directives about how to act in view of certain features of a situation and an action, rather than in virtue of likely consequences of the conduct. But whether norm-compliance is admirable or in some way desirable or good for cooperation depends on the content of the norms, and not simply on the fact that conduct is norm-guided. Consider the norm of repaying insult with unforgiving retaliation: this norm may make insult less likely (because of the costs that will be imposed in a population in which the norm is widespread), but it will also make cooperation impossible, once we have the first insult; and it also may have the effect of escalating the first insult: as long as I am going to face unforgiving retaliation, I might as well insult as strongly as possible. On the other hand, a norm of forgiveness may invite (enable?) uncooperative behavior, though it may also have the effect of reestablishing cooperation. The point is that we need to distinguish the thesis that norm have real influence on conduct—that conduct can be explained by reference to norms—from the thesis that people are morally decent or naturally cooperative. Nothing in the idea that norms have real and independent influence—on the conduct of individuals or on the conduct of groups or states—need suggest a naively optimistic view about the prospects for cooperation or justice. It is worth recalling in this connection that Bull's Grotian view of international politics includes the claim that states embedded in an international society featuring a sense of shared interests and norms are less likely to be ruthless, but that's because of the overarching norm of justifying conduct on grounds that other states also accept, not simply because states are assumed to be norm-guided.

2. Some Main Ideas in Elster

- **Order**: two issues of order, predictability of conduct and cooperation. Norms have some connection with both:

(i) predictability: (a) it may be impossible to predict what someone will do, knowing that they are rational. If rationality predicts equilibrium and there are multiple equilibria, then we need some way of selecting among the equilibria, and it may be that some equilibrium is supported by shared norms: a norm of deference in Battle of the Sexes, for example; (b) puzzles about use of backwards induction in understanding how rational agent will act. On the one hand, we use backwards induction in predicting how a rational agent will act, but not clear if we can legitimately use it, because there appears to be a tension between the conclusion of using it (that rationality leads an agent to act in a certain way) and the actual use of it, which requires that we make sense of what happens when someone has acted irrationally;

(ii)   cooperation: here, we have order when people act in ways that are required for making everyone better off. But there are concerns about whether that cooperation will be forthcoming from rational agents, in two kinds of cases: (a) cases—like voting, tax compliance, organizational membership, organized resistance—in which the costs to the agent outweigh the benefits to the agent, but the agent's conduct confers benefits on others (externalities and helping). In these cases, the magnitude of the benefit should not matter, because an individual agent typically will not be able to make any real difference to the likelihood that the benefit will be provided; (b) problems of bargaining, where there are lots of ways of distributing the results of cooperation, disagreement about which is best, we need some decision on that in order to get cooperation, and no determinate answer to the question of what rational bargainers will agree to. Once more, norms *may* help in solving the bargaining problem, and thus in fostering cooperation. So consider the argument Rawls gives from the original position: he claims that rational parties would agree to his principles, but the plausibility of this claim—to the extent that it is plausible—depends on the normative notions of fairness expressed in his veil of ignorance. Moreover, I say that appeal to norms "may help" because disagreement extends to norms as well as interests, and the appeal to norms may deepen disagreement rather than resolving it (e.g., what is the relevance of status quo in bargaining?).

- **Social Norms**:

  (i)   Examples of Norms: (a) Rules of Thumb: norms of individual choice, such as: look before you leap; a bird in the hand is worth two in the bush; don't look a gift horse in the mouth; and don't judge a book by its cover; (b) organizational codes, SOPs; (c) Norms of Good Character, of decent and honorable behavior, which need no be specifically moral: don't be crude; maintain your dignity; show kindness to strangers; don't get down in the gutter, even if others do; stand your ground; act with integrity; (d) norms of etiquette; (e) moral norms: don't cheat, steal, lie, or make false promises; help others who are in need, at least when the burdens are not overwhelming; do your fair share in producing benefits, at least when others are doing so, by restraining their own conduct to contribute to a social good (fairness); act as you would wish everyone to act because of the results that would ensue, regardless of how they are acting (everyday Kantianism).

  (ii)   Now this last one may seem consequential, but it is not. So consider a Kantian example: You notice a person who could use

your help. Though it wouldn't cost you much in time or resources, you are nevertheless inclined not to help. So in effect you have the maxim or rule for conduct: "each person for him or herself, God for us all." Being rational you want to be sure that not helping out is prudent—will anyone see you being a bastard, with bad effects on your own future interests? But suppose you are confident that you will be no worse off for failing to help. Still, you are troubled by the question: is this the right thing to do? And Kant proposes that we answer this question by asking: Could I will the maxim of my action ("Each for him/herself, God for us all.") as a universal law: a world in which, as a matter of law, everyone acts on my maxim? If not, then I ought not to act on my maxim, and I ought to provide the help. Now is there a problem willing a world in which everyone acts on my maxim, helping others only when it is to their own advantage. Kant thinks so. For in willing such a world, I will a system in which I cannot expect to get help that I need, unless others find it prudent to help. So on the one hand, I want to get help when I really need it (not simply when others find it convenient); but at the same time, I will that no one help others except when prudence requires it, i.e. will a world in which the very assistance that I want cannot be counted on. Now this may seem consequential, but it is not: you are not going to provide the help because you think you are more likely to get it if you provide it, but because you think it is wrong to make a special exception for yourself. Because this is the nature of Kantian thought, it is hard to see why it rests on magical thinking: on a confusion of causal and diagnostic thinking. The Kantian does not think that his/her acting rightly will cause others to act in the same way, but that acting as I wish others to act is the right way to act.

(iii) Characteristics of Norms: (a) norms in each of the categories are not consequentialist (or instrumental) in their *formulation*. That is, the norms themselves do not, on their face, direct us to actions that will produce (or that make it reasonable to expect) desirable outcomes. Instead, they simply state that there are certain kinds of actions that we ought to take. In characterizing norms as being characteristically facially non-consequential, I do not mean to imply anything about either the characteristic motivations for complying with them or the kinds of justification that might be offered in support of such compliance. Thus it may be that the motives that lead me to comply with a norm that has a non-consequential structure are motives that look to consequences (I comply with norms of honesty because I expect greater monetary gain, or fear getting caught). As Elster

says, they are not outcome-oriented: but it is a mistake to say that they need to be guided by the past rather than future expectations (norms against making false promises are not guided by the past), and they need not—as Elster proposes—involve obligations or interdictions. For example, standards of good character and rules of thumb do not, and nor do the norms of etiquette; (b) practical, in the sense that they provide standards for guiding conduct and for guiding judgment of conduct of others; (c) plurality of norms: in any case of conduct, there are likely to be a variety of norms that bear on a course of conduct and each of the norms may not be fully determinate. So I may think that I ought to be honest but also that I ought not to insult a stranger who asks if I like her hat; or that I ought to let bygones be bygones, but that I also ought to uphold the honor of my family.

(iv) Social Norms are norms with two characteristics: (i) shared among members of a group; and (ii) sustained by reactions of approval and disapproval among members of the group. These conditions seem both puzzling and irrelevant. They are puzzling because the norms that Elster appeals to in his account of collective action—the norms of fairness (reciprocity) and everyday Kantianism—are not social norms. Consider everyday Kantianism: this is not a widespread norm, though it may be important to solving problems of collective action that some people are prepared to initiate by contributing irrespective of the contributions of others, just because they think the results of everyone's contributing would be very good. Moreover, it is not sustained, if it is sustained at all, social approval and disapproval: indeed it may be that the everyday Kantian is someone who is seen as fanatical and in need of restraint by the group. It is true that there may be group benefit in having some everyday Kantians around, who are prepared to do what they think everyone ought to be doing, even if they think that their own morally exemplary conduct

3. Reality and Autonomy of (Social) Norms

- *Status.* Two questions about what might roughly be called the "status" of norms: (1) Reality: are people really motivated to comply with norms by self-interest or by some other concern about consequences, or is there some sort of independent grip on the mind that norms have? (2) are norms explained by their beneficial consequences for individuals or groups? Can we account for the presence of a norm in a population by references to these consequences, even though the norms themselves are on their face non-consequential? Focus here on the reality of norms.

- Indeterminacy of Norms? Norms are manipulable in at least two ways, and by manipulating them may be able to act in own best interests: (1) individual norms do not have fully determinate content. So if I think that I ought to help others in need just in case the burdens on me are not excessive, I may be able to find some way to present the burdens as always very great; (2) because there are a plurality of norms, may be able to identify a relevant norm that instructs me to act in the way that serves my interests. On these views, norms are not at all constraining, not because people violate the norms but because you can always find a way to comply with norms (given their indeterminacy) without constraining the pursuit of your interests.

- Reasons Against Norm Skepticism: (1) in the case of some norms, it is hard to see how self-interest could be at work: political participation norms, where there is no plausible argument from self-interest for participating, no matter how large the stakes; (2) if no one cared about norms, then there would be no reason to present oneself as complying with them. Not clear how compelling this point is: might mistrust people who do not know how to put on a show of honoring norms (shows general incompetence); (3) need for consistency: cannot manipulate norms by appealing to them on some occasions and not on others. The observation seems right, but not relevant to the reality of norms, because it shows instead why people may be under pressure to comply, even though the norms do not motivate them; (4) limited repertoire of norms, which means that you may not be able to find one that conforms to self-interest. Assume that this is right: that the norm skeptic is wrong about the extent of the indeterminacy of norms. Still the point is of uncertain relevance. What it may show is why we might expect some norm compliance from individuals who are not motivated by the norms, but need to present themselves to others as so motivated for one of the reasons noted earlier.