Social Complexity, Learning and International Political Economy,

With Illustrations from the Political Economy of Trade Policy

Douglas R. Nelson

Acknowledgments

This paper was prepared for the CREDIT 10th Anniversary conference at the University of Nottingham, 17-18 September 1998. This is an appropriate time and place to thank CREDIT for continuing interest in and support of my work. This paper, in particular, has benefitted from the encouragement of David Greenaway and Oliver Morrissey. Without implication, I would also like to thank Matthew Slaughter for discussion that was particularly helpful in writing the conclusions to this paper. Comments from John Black, Kevin Grier, and participants at the CREDIT 10th anniversary conference and the 23rd annual International Economics Study Group (IESG) at St. Antony's College, Oxford, are also greatly appreciated.

Abstract

Social Complexity, Learning and International Political Economy, With Illustrations from the Political Economy of Trade Policy

Complexity is an essential element of the political economic environment. This fact implies that ignorance and learning will affect the dynamics of that environment. This paper surveys recent work from decision theory, game theory, and economics that seeks to formalize learning in a social environment. To illustrate the gains from this sort of analysis, the main body of the paper presents an approach based on Bayesian learning from the behavior of others in some detail. Throughout, the analysis is motivated by three examples from the political economy of trade policy.

Social Complexity, Learning and International Political Economy,

With Illustrations from the Political Economy of Trade Policy

Reality is hideously, uncompromisingly, and (at least for social scientists) inconveniently complex. In many contexts this complexity is essentially irrelevant. It is, loosely speaking, far enough away from the domain of decision that simple local approximations do a very good job. A butterfly in the Pacific *might* cause a tornado to appear when one wants to go shopping, but this is unlikely to be particularly important in the purchase of a tube of toothpaste. In such (locally) simple cases, the environment is sufficiently well-defined and stable that standard full-information models will be good approximations for positive and normative analysis. However, in at least equally many cases, complexity is an essential element of the decision context.[1] By "complexity" I don't have in mind any particular formalization, but rather the common sense notion that there is too much potential information and not enough theory to make sense of the situation–I will refer to this as *agent complexity*.[2] It is in precisely these situations that we might expect agents to condition their behavior on the actions of others, not for strategic reasons, but because those who have already taken actions might reasonably be expected to possess knowledge one does not

[1]It is not the purpose of this paper to discuss complexity as a problem for social scientific analysis. See Nelson (1999) for such a discussion applied to the case of trade policy reform.

[2]By this I mean that the environment is complex from the perspective of the agent. Somewhat ironically, complexity in the sense of, say, Prigogine (1984) or Kauffman (1993) is often not complex in this common sense way. This is not a criticism. That work is after something else: generating complex aggregate behavior and, specifically, phenomena like self-organization. Interestingly, agent complexity is not necessary for aggregate complexity. In fact, one of the most popular classes of bridge between agent complexity and aggregate complexity involves reducing the rationality of agents (e.g. evolutionary models, finite automata, etc.).

possess. This attempt to learn from the behavior of others is called *social learning*.[3]

The politics of economic policy are obviously characterized by agent complexity. With the exception of very rare, very local, cases (e.g. regulation of local utilities), economic policy involves making judgements about consequences even the experts (e.g. economists) cannot agree upon.[4] The result is that citizens and policy-makers attempt to derive information about aggregate consequences from directly observable facts–among them, the actions of other citizens. The problem with such social learning, as we shall see, is that it can lead to what I shall call *footloose policy preferences*. That is, aggregate opinion which is seen to strongly support one position can shift to support the opposite (or at least an opposing) position with equal conviction, without any obvious change in the fundamentals of the decision situation. While political economy modeling has resisted analyses based on agent ignorance, on the compelling grounds that we can explain virtually anything under appropriately constructed ignorance, I will argue that certain phenomena cannot be understood any other way. I will ultimately conclude that the incorporation of ignorance has interesting implications for us as policy intellectuals. First, and most attractively, social learning identifies a direct and important role for providers of expertise.[5] However, secondly, this implies that, as a profession, we should learn to communicate more

---

[3]This usage follows Gale (1996). By contrast, Vives (1996) uses the phrase to refer to the study of information aggregation at the level of the economy as a whole.

[4]What is, perhaps, worse, economists often cannot even agree on how to talk about such issues.

[5]This conclusion is very much contrary to the usual implications of positive political economy without individual complexity and social learning. Though see O'Flaherty and Bhagwati (1997) and the comments on that paper collected in a special issue of *Economics and Politics* edited by T.N. Srinivasan.

effectively with civil society.

Some notion of common beliefs has played a central role in much of the development of contemporary research on international political economy. Functionalist integration theory saw the emergence of a transnational civil society from the increased intensity of everyday interactions among citizens of integrating nations (Haas, 1958). Current research on epistemic communities seeks to understand the way common beliefs constrain the operation of power and permit cooperation among nations even in the absence of a hegemonic power (Ruggie, 1982; Haas, 1992). Even regime theory found a role for "ideas"–though in this case the role was primarily the auxiliary role of accounting for predictive failures of the main body of theory (Goldstein, 1988). Much of this work is primarily concerned with the consequences of the presence or absence of such common understandings, not with their dynamics–i.e. the issues of creation, evolution and collapse of such understandings. Dynamic issues have been addressed in the substantial literature that has studied the evolution of shared beliefs in specific cases (Goldstein and Keohane, 1993; Rueschemeyer and Skocpol, 1996).[6] However, because this latter body of research lacks a systematic framework, it is difficult to draw systematic lessons or generalize in ways that permits cumulation.

Over the last, say, decade and a half an intensely interdisciplinary effort–linking mathematicians, philosophers, computer scientists, decision theorists, and game theorists–has sought to develop theoretically well-specified models of individual and, especially, collective

[6]See Levy (1994) for a useful survey of research on learning in the international strategic context. It is notable that this is the one area where a substantial body of research addresses issues of learning. For example, the papers recently collected as Aumann and Maschler (1995) were originally produced in the late-1960s with funding from the Arms Control and Disarmament Agency.

knowledge and learning. This paper has two primary goals: to provide a brief overview of this exceptionally interesting body of research; and to develop some simple examples of the implications of this research for international political economy. This second goal will be pursued by developing one strand of this research in a bit more formal detail with illustrations from three aspects of the political economy of international trade. In the next section I motivate the analysis by briefly discussing three questions of trade policy that seem hard to address convincingly without taking social learning and footloose preferences into account: shifting public preferences on NAFTA; optimal speed of trade policy reform; and the rapid shift from import-substituting industrialization (ISI) policy to export-led growth (XLG) policy among both development professionals and policy-makers. Next I provide a simple typology of the broader literature on learning and social learning in political economy to provide a context for the following discussion. The main discussion is a survey of results from the class of social learning model that has grown from the fundamental contribution of Bikhchandani, Hirshleifer, and Welch (BHW, 1992). In the final section I conclude with a discussion of the implications of these models as suggested in the previous paragraph.

Footloose Preferences and Problems in the Political Economy of Trade Policy

From the point of view of the citizen, or policy-maker, trade policy is extremely complex. Trade policy rarely comes in the form of a single, discrete act of protection.[7] Instead, trade policy is embedded in legislative acts made up of complicated bundles of changes in the law regulating

[7]Even in the days of the classic tariff system–say, 1870-1932–the political action revolved around tariff acts with hundreds of line items. In addition, as research on 19th century voting suggests, the social meaning of the tariff was highly variable across local electorates.

trade that even experts have a hard time evaluating. So many things are changing in these situations that, even if we could give a clear account of the objective function we are working with, the construction of the problem will still be hopelessly contested.

NAFTA is a very interesting example. It was the first instance of highly public trade politics since the heyday of the classic tariff system and, as such, it may tell us something about the politics of trade in years to come. It was in the news, it was dinner table and cocktail party conversation among non-economists, every politician had (and had to have) a public opinion (sometimes different from their private opinion). Trade economists approached minor celebrity status, and the unanimity with which trade economists concluded that NAFTA was economically insignificant for the US was truly stunning. With similar unanimity economists heaped well-earned scorn on dishonest claims about the "giant sucking sound".[8] But NAFTA is a very complicated agreement: NAFTA itself is a document of over 300 pages (not including the national tariff schedules and various other lists–with these the text runs over 2000 pages) covering trade in goods, technical barriers to trade, government procurement, investment, services, intellectual property, and the administrative and institutional conventions needed for implementation; the agreement on environmental cooperation is another 20 pages; as is the agreement on labor cooperation; all of these need to be implemented with specific national regulations in all three signatory countries. Some of these regulations will affect trade in goods, others will affect trade in factors of production; some will liberalize this trade, others will restrict it. In addition, while NAFTA officially became law in 1994, full implementation will not occur in sensitive sectors for

[8]It is probably useful, given the recent vogue among trade economists for attacking free trade areas, to recall that during the NAFTA debate support for NAFTA was seen as a litmus test of one's status as a serious economist.

periods of 10 to 25 years. Reasonable people might well be excused for harboring doubts about any claims made, whether by economists, Ross Perot, environmentalists, or unionists.

A recent paper by Philip Levy (1998) argues that public opinion shifts around NAFTA are suggestive of some kind of learning. Specifically, he refers to a Gallup poll from 4 April 1991 in which:

> "72 percent of Americans thought [NAFTA] would be 'mostly good for the country' and 15 percent thought it would be 'mostly bad for the country'."

But in a Wall Street Journal/NBC News poll of 5 March 1996:

> "21 percent of Americans thought NAFTA had 'had more of a positive impact' and 48 percent thought it had had 'more of a negative impact'."

Note that this shift in opinion occurs despite that fact that, although NAFTA had been officially "implemented" in January 1994, virtually nothing of economic substance had occurred, or was scheduled to occur for several years. In addition, Mexico's economy had experienced severe political and economic problems that could in no way be attributed to NAFTA. Perhaps more surprisingly, aggregate economic performance was stronger in 1996 than in 1991. For example, unemployment was higher in 1991 (6.8% and rising) than it was in 1996 (5.4% and falling). While the current account balance was less negative in 1991 than in 1996, the administrative protection mechanism was far more active in the early 1990s than in the mid-1990s, reflecting much greater concern with import competition on the part of import competing firms.

Clearly some form of "learning" had occurred, since expressed preferences had shifted fairly dramatically. What is completely unclear is what triggered this reassessment. There was virtually no NAFTA-specific information and the aggregate facts of the economy suggest the sort of environment in which trade is viewed relatively benignly. This would seem to be a virtually archetypal example of footloose preferences. Given the weakened domestic political institutions supporting trade liberalization in the U.S., and the concomitant likelihood of increased prominence for trade as a public political issue, such footloose-ness of trade preferences could play a significant role in the future politics of trade.[9] As proponents of liberal trading relations, it behooves us to seek an understanding of such preferences.

Similar issues arise, with increased intensity, in the case of major trade liberalization, especially when adopted as part of large-scale macroeconomic reform. In the case we have just considered, the overall economic environment and the political institutions are relatively stable, if difficult to understand. Large scale policy reforms generally occur under conditions of generalized social instability. That is, trade policy is only part of, often a small part of, what is at stake. In particular, because broader macroeconomic and social policies, as well as political institutions, are at stake; often with complex links between different sets of policies, a clear understanding is surely orders of magnitude more difficult than a case like NAFTA. Thus, we should not be surprised by the spectacle of sequences of adoption and reversal of reforms, driven by little but highly unpredictable public support. In fact, this phenomenon is so common that there is now a substantial, and sophisticated, literature which seeks to provide a framework to assist in understanding and responding to such policy shifts.[10]

In the previous two cases, we sought to identify cases of footloose citizen preferences,

---

[9]See Nelson (1989) provides a discussion of the role and transformation of US trade policy institutions for the support of trade liberalism. Nelson (1995) discusses the collapse of those institutions in the early 1970s.

[10]Nelson (1998) discusses the literature on the political economy of policy reform with a focus on social complexity.

under the plausible presumption that this phenomenon plays a significant role in the determination of policy. The third case we consider has to do with herding by policy-makers that is not driven by citizen herding. Complexity of the decision context is at least as severe a problem for public decision-makers as it is for citizens. I mention two examples. First, at some point between the debates on the Hawley-Smoot tariff and contemporary debates, the presumption among policy-makers shifted from generalized support for protection to generalized support for liberalization.[11] What is striking about this is that it has occurred without any evidence of such a shift in presumption among citizens. In fact, the median voter in the US is now, and has apparently always been, essentially protectionist.[12] Similarly, just as US political elites seem to have dramatically shifted beliefs with respect to trade in the early-ish post-World War II era, political elites in the developing world seem to have shifted from a strong preference for import substitution in the 1950s and 1960s, to an equally strong presumption in favor of export orientation in the 1980s and 1990s. This last example is particularly interesting because there has also been an equally significant shift of exactly the same sort, at more-or-less the same time, on the part of development economists. This case seems particularly fruitful for studying the relationship between policy intellectuals and policy-makers in complex decision environments.[13]

---

[11]By this I mean that, during the Classic tariff era, supporters of liberalization at the margin needed to make statements of the form, "Of course I support the tariff, but..."; while contemporary proponents of protection must assert, "Of course I support trade liberalization, but ...".

[12]Nelson (1989) demonstrates that the success of the trade liberalization program in the US rested on the creation of institutions that shielded trade policy from domestic politics.

[13]Anne Krueger (1997) devoted her presidential address to the American Economic Association to precisely this question.

Research on Learning in Economics: A Typology[14]

As we noted in the introduction, ignorance would appear to be a dubious foundation for political economy (or most other kinds of) modeling. In the previous section, however, we noted that ignorance, at least arguably, looms large in public decision making on trade policy. Trade policy is not unique in this regard. Thus it is hardly surprising that there are several large bodies of research focusing directly on the characterization of ignorance and the identification of rules for learning. In this section we very briefly characterize these bodies of research to provide a context for our focus on one of these.[15] Table 1 sorts the economics literature on learning with respect to: the characterization of rationality; and the source of information from which agents learn.

–Table 1 about here–

Following more-or-less standard usage for situations of risk/uncertainty, we take Bayesian rationality as a standard baseline and treat, in addition to the case of full Bayesian rationality, the cases of bounded (i.e. less than full) rationality and practical reason (i.e. more than full Bayesian

---

[14]The limitation to research in economics (under which label I include decision theory and game theory) is a genuine limitation. Psychology in particular has a distinctive tradition of research on learning that is strongly complementary to research in economics. Patrick Suppes has produced a number of very fine surveys of psychological learning theory (Suppes 1977, 1989; Suppes, Pavel, Falmagne 1994). Not surprisingly, political science has addressed a number of learning-related issues. For applications to foreign policy learning by state actors, with a particular emphasis on strategic contexts, see the very nice survey by Levy (1994). More closely related to the considerations of this paper are attempts to account for response instability in surveys (see Zaller, 1992).

[15]Because the main purpose of this paper is not to survey learning theory in general, we will refer primarily to existing survey papers or classic references. A larger, but still selective, bibliography on learning can be found at http://www.tulane.edu/~dnelson.

rationality).[16] Since agents in all three cases are possessed of standard utility functions and are desirous of achieving maximum individual welfare relative to them, the distinction is based on the nature of the process by which they derive inferences from the data they observe. The case of Bayesian rationality is easy: in addition to their utility functions, agents possess well specified prior distributions on the relevant state space; and update the prior using observed data and Bayes' rule. By bounded rationality we will mean that agents adjust their behavior to observed data by using some relatively fixed decision rule ("rule of thumb"). Finally, by practical reason, we refer to the various attempts to endow decision-makers with something more than the capacity for first-order optimization, e.g.: the capacity to recognize patterns; the possession of higher-order logics; or more sophisticated rules for adjusting to new information.[17] [18]

As we have already noted, we take social learning to refer to learning by direct observation of other agents, while non-social learning involves making inferences from data other than the observed behavior of others. Within the category of non-social learning, we consider:

---

[16]Characterizing the third case as Bayesian rationality *plus* may be misleading since, in many of these cases, it is really *satisficing plus*. However, for our purposes here, the interesting attribute of this research is that decision-makers are endowed with higher level cognitive and/or information processing capacities than is usually assumed in standard decision-theoretic modeling. There is a strong family resemblance between this programme and that in normative decision theory of endowing individuals with higher-order moral judgements (e.g. Frankfurt, 1971; Jeffrey, 1974).

[17]Unless otherwise specified, we will be assuming that utility functions have the expected utility property.

[18]It is important to note that there are representations of standard decision-theoretic models in terms of these more general models–that is, of course, what it means to be more general. In fact, much of the literature by decision theorists and economists attempts to find the minimally restrictive set of assumptions that underwrite standard decision-theoretic results. See Dekel and Gul (1997) and Battigalli and Bonanno (1998).

single-agent learning from privately observed data; single-agent learning from publicly observed data (e.g. aggregate summary data, say, price); and learning from an aggregate public signal in strategic situations.[19] Within the category of social learning we consider non-strategic and strategic cases.

The first column contains (among other things) what must be the oldest, and most completely formalized, body of research on learning: statistical decision theory. One of the main goals of the pioneers of modern statistical analysis (R.A. Fisher, Jerzy Neyman, E.S. Pearson, Abraham Wald, L.J. Savage, *et al.*) was to formalize the process of inference (i.e. learning) based on statistical analysis. Neyman, Pearson, and Wald all worked toward a sophisticated framework of inference based on fixed, quantitative decision rules (see e.g. Wald 1950). Wald (1947), in particular, was interested in learning from sequential experiments. While highly operational in practice, these rules do not emerge as the optimal answer to a question about inference.[20] The usual defense of these rules is that they are sturdy and easy to apply, very much like general decision processes studied under the rubric of optimal rules of thumb.[21] The Bayesian approach, represented particularly clearly in Savage (1954), attempted to derive optimal decision rules from an explicit, choice theoretic foundation. This model of learning has formed the basis of a large

---

[19]In this third case agents are trying to learn *about* other agents, not *from* them. That is, all agents involved in a strategic interaction observe a common signal, based on which they will try to infer something about each other.

[20]The same criticism applies to proposals to use minimax rules as, for example, in Blackwell and Girschik (1954).

[21]In addition to the classic papers by Simon (1982, sections VII and VIII), see the recent survey by Conlisk (1996) and the very interesting lectures by Rubinstein (1998).

proportion of research on learning.[22] Applications include: consumer learning with unknown product quality, monopoly pricing with unknown demand, investment with unknown demand, and technology choice/adoption.

In recent years a substantial body of research has developed which seeks to endow decision-makers with more sophisticated inference processes than those possessed by Bayesian decision-makers. In all of these cases, the research involves both a characterization of knowledge (or belief) which is richer than that in the standard Bayesian case, and the rules for revising that knowledge. One prominent approach characterizes knowledge in terms of some version of epistemic modal logic and models revision in terms of the adjustments that maintain consistency among statements making up what one knows (believes).[23] Recently, there have been substantial successes in linking this approach to generalized Bayesian approaches based on non-additive probability theory (Mongin, 1994). An alternative approach models dynamics in terms of pattern recognition (Holland, Holyoak, Nisbet and Thagard, 1987; Holland, 1996; Gilboa and Schmeidler, 1995). There are, unsurprisingly given the complexity of the issue, many other approaches. What is common across all of them is the attempt to develop models that engage in something like practical reason, not simply mechanical updating.

---

[22]See Kiefer (1988/9), Kiefer and Nyarko (1995), and Marimon (1997) for surveys of Bayesian learning theory with application to economic decision making.

[23] Strictly speaking, epistemic logic is the logic of knowledge, while doxastic logic is the logic of belief. Since both knowledge and belief act syntactically and semantically like necessity, modal logic is a natural tool for analyzing both. Battigalli and Bonanno (1998, section 2) is an exceptionally clear short introduction to the logical analysis of knowledge and belief, while Fagin, Halpern, Moses, and Vardi (1995, chapters 2 and 3) gives a more detailed development. Chellas (1980) and Hughes and Cresswell (1996) are standard textbook treatments of modal logic. Gardenfors (1988) is a detailed development of one of the most prominent research programmes dealing with the dynamics of knowledge revision.

In the next column, the inference process becomes somewhat more social as each agent acts essentially like decision-makers in the first column, but all agents are attempting to condition their behavior on a common, public signal. For example, there is an extensive literature that asks whether or not learning of this sort will lead to a rational expectations equilibrium (REE).[24] A large body of literature asks whether a REE can be achieved if all agents are learning from public information (e.g. price) based on no more information than that possessed by an econometrician–called *least squares learning* (Sargent, 1993; Evans and Honkapohja, 1995).[25] An alternative is to endow all agents with the capacity for full Bayesian rationality in evaluating an aggregate signal like price (Frydman, 1982; Blume and Easley, 1984; Feldman 1987; Bray and Kreps, 1988).[26] It is an interesting result of this literature that, even in quite simple models and under either bounded or full Bayesian learning, aggregate behavior may be quite complex. The last of the columns is yet more social, and involves agents that know they are involved in a strategic interaction, but who must draw conclusions about the behavior of other agents based not on direct observation of their behavior, but on the basis of a public signal (e.g. price, as in the REE literature)–in this case agents are attempting to learn *about* each other rather than *from* each other. As with the literature on convergence to REE, game learning from a public signal can

[24]The earlier literature which had proved the existence of a REE had left open the question of whether it could be achieved. General overviews of this literature can be found in Radner (1982) and Jordan and Radner (1982).

[25]Least squares learning is only one of many possible ways of characterizing bounded rationality, and others have been studied. However, it has been a particularly central case and seems to have been more extensively studied than any of the other possibilities.

[26]There is a closely related literature which asks whether a group of Bayesian agents can achieve a common poseterior, and common knowledge of that fact, from observing a common signal. The key reference here is McKelvey and Page (1996).

involve either adaptive learning or full Bayesian rationality.

Whether we are interested in speech acts or international policy coordination, it is clear that all forms of sociality rely on a non-trivial degree of common understanding. While this fact has been known, and commented on, for as long as there has been social theory, formal analysis has, until recently been rare.[27] The philosophy of language, in the Frege-Russell-Wittgenstein tradition, from its beginning married axiomatic method with a fundamental interest in shared understanding. Thus it is probably not surprising that the earliest formal work on this issue appears to come from the application of epistemic logic to the analysis of shared meanings in Lewis (1962) and Schiffer (1972).[28] The work of Lewis and Schiffer was very much continuous with previous work in the field and would surely have generated continuing work, but the huge increase in research in this mode appears to have been driven by the, more-or-less simultaneous, recognition in computer science (artificial intelligence) and game theory that work of this sort offered the possibility of solutions for problems with which these fields were faced. While most research of this sort seeks to characterize common knowledge (belief) and to analyze its implications, a considerable amount of work has studied whether common beliefs emerge under various learning rules and the relationship between these beliefs and knowledge (justified true beliefs). Thus, considerable effort has been given to iterated direct communication of posteriors

---

[27]Informal analysis, such as that in Giddens (1986) has been quite common. But, as the example of Giddens suggests, the lack of a clear analytical structure has generally meant that there is little in the way of cumulation.

[28]There is a closely related body of work in statistical theory on merging of opinions with increasing information, but, until the emergence of research on common knowledge and common belief, there does not appear to have been any recognition of the broader implications of that research for social theory.

in a Bayesian environment. In the next section we consider the sizable literature on observation of the acts of other agents in a Bayesian environment. Except for some very preliminary work in computer science (Kfir-Dahav, Noa and Moshe Tennenholtz, 1996), there does not appear to be any application of more sophisticated belief revision in the multi-agent case.

Shifting to the column on the far right, there is a recent, but very rapidly growing literature on learning from others in strategic settings.[29] The primary issues in this literature are whether some learning process: 1) must lead to a Nash (or some other) equilibrium; and 2) can refine the set of equilibria further than standard static refinements.[30] There is a substantial literature on bounded rationality and learning involving adaptive programming (Marimon and McGratten, 1995; Matsushima, 1997); finite automata (Rubinstein, 1987; Kalai, 1990; Lipman, 1995); and/or evolutionary games (Kandori, 1997; Mailath, 1998; Young, 1998). Given the strong rationality assumptions imposed in game theory it is unsurprising that full Bayesian rationality is well-represented in this literature (Blume and Easley, 1995). Finally, there is an interesting case of not-quite-full (or naive) Bayesian learning that goes under the label of fictitious play. As with the case of individual rationality, epistemic logic-based approaches have been extensively applied to the evaluation of various solution concepts though, as with the case of non-strategic learning, this has focused primarily on characterization, not on the application of more sophisticated learning procedures.

---

[29]Eichberger (1992), Battigalli, Gilli, and Molinari (1992), Fudenberg and Levine (1998), and Walliser (1998) are useful overviews of this literature.

[30]This is seen as providing foundations for solution concepts. Loosely speaking, since there are a wide range of results in this literature, it seems to be the case that the stronger the rationality attributed to agents, the weaker the solution concepts that are supported by the analysis (e.g. Walliser, 1998).

To illustrate the gains from systematic analysis of learning, we now turn to a more detailed discussion of one particular class of non-strategic social learning, that characterized by learning from the actions of others.

Non-Strategic Social Learning, Information Cascades, and Herd Behavior

In both formal and informal political economy research, it is standard to assume that agents (however defined) seek to choose policy, or respond to changes in the policy environment, as best they can. However, as we noted in the introduction, situations of large scale policy change will generally imply a high degree of uncertainty among agents. One way of reducing that uncertainty is to seek information from others, either directly (i.e. by asking them) or indirectly by observing their actions. In this section we develop a simple model, due originally to Bikhchandani, Hirshleifer, and Welch (BHW, 1992), of indirect learning, focussing on the interesting implication that, even with fixed preferences over final outcomes, policy preferences may be footloose. This is related to what BHW call an information cascade.

In an environment where individuals can learn about the environment from both private information and the behavior of others, an *information cascade* occurs when agents ignore their private information and follow the behavior of others--agents engage in belief herding.[31] Where

[31]Herding occurs whenever all agents focus on a single behavior, with particular reference to cases in which there are multiple plausible candidate behaviors. The phenomenon of herd behavior is common enough that it has been used as the basis for a wide variety of economic analyses based on such things as demand interdependence (Leibenstein, 1950; Schelling, 1978; Becker, 1991) and network externalities (Dybvig and Spatt, 1983; David, 1985; Farrell and Saloner, 1985; Katz and Shapiro, 1985). As noted by BHW (1992), other social sciences develop explanations in terms of factors such as internalized social norms and sanctions on deviants. Both of these are discussed in detail in Coleman (1990) and Elster (1989), in ways that are congenial to economists. Also see the very interesting essay by Bowles (1998) for perspective on these and

the behavior of others is not perfectly informative with respect to their private information, an information cascade effectively traps socially useful information, thus permitting socially suboptimal outcomes. Thus, the essential elements of the information cascade model are coarse public signals and private signals of bounded accuracy. Without the first assumption, the law of large numbers suggests that, with a sufficiently large number of observations, the true state of the world is revealed (almost surely). In most information cascade models this assumption takes the form that agents observe the actions of other agents but not their signals, and that the actions are imperfectly informative with respect to signals. Without the second assumption, individuals might receive fully informative signals, allowing them to take actions that would break the cascade. In addition, BHW make a number of additional assumptions that permit a very simple expository model.[32]

Specifically, suppose that each of a countable number of identical, Bayesian rational agents ($n \in N$) faces a decision problem characterized by a binary state space, $\Omega = \{\omega_0, \omega_1\}$, and a binary action set, $A = \{a_0, a_1\}$. In the political economy context, the state refers to the effect of a policy (e.g. NAFTA, policy reform, export orientation) on the performance of the economy. Similarly, we might think of the actions as publicly supporting or opposing the policy. Payoffs are

---

related issues.

[32]The Appendix to this paper presents the analytics of the simplest model. Banerjee (1992) is another early reference. But for reasons that we will discuss later, this paper has not had the impact on later development of research in this area that BHW (1992) did. The basic model has been applied to the analysis initial public offerings by Welch (1992) and Truemann (1994) examines herding by market forecasters in a similar framework. In related work, agents *seeking* to herd on a common behavior end up impounding socially useful information in way similar to that in information cascade models–Froot, Scharfstein, and Stein (1992) consider herding by short-term investors, while Choi (1997) considers technology adoption with informational externalities.

given in terms of a utility function, common to all agents: $V(a_k, \omega_i) \equiv V_{ki} = \{0,1\}$, where we assume that $V_{00} = V_{11} = 1$ and $V_{01} = V_{10} = 0$. That is, the payoff to a_0 is higher than that to a_1 if the state is ω_0, and similarly for a_1 if the state is ω_1. All agents have the same period 0 prior on Ω: $\pi(\omega_0) = \pi(\omega_1) = ½$.

Each individual observes a private, conditionally independent, binary signal: $X = \{x_0, x_1\}$. These signals are, conditional on the state of the world, identically distributed for all agents such that: if the true state of the world is ω_0, agents observe x_0 with probability $p > ½$ and they observe x_1 with probability $1 - p$; and if the true state of the world is ω_1, agents observe x_1 with probability $p > ½$ and they observe x_0 with probability $1 - p$.[33] This symmetry means that the likelihood functions for the signals are: $f\left(x_0 \middle| \omega_0\right) = f\left(x_1 \middle| \omega_1\right) = p;$ and

$f\left(x_1 \middle| \omega_0\right) = f\left(x_0 \middle| \omega_1\right) = 1 - p.$ As rational Bayesians, if the agents act simultaneously, they would all choose their optimal acts, after observing their signal, by updating according to Bayes' rule (see the first equation in the appendix).

However, in the basic information cascade model, agents take their actions in a fixed sequence, and the *actions* of all previous agents are observable. Thus, except for the first agent, all agents possess the public information contained in the actions of the agents preceding them, as well as their private information. Again, it is important to the model that this information is contained in the *actions* of the previous agents; the *signals* of the previous agents are

[33]Thus, if actions are fully informative with respect to signals, after a sufficiently large number of agents have taken actions, all later agents can determine the state of the world (a.s.) by identifying which signal has occurred with greater frequency.

unobservable. We now need to characterize this public information and the way it is incorporated in the decision strategies of the agents. First, we need to represent the *history* prior to any agent's decision. Let $n \in N$ index agents (we will sometimes refer to this as *periods*, i.e. one agent acts in each "period"). For $n \geq 2$, denote by $H^n \equiv \{a_0, a_1\}^{n\text{-}1}$ the space of all possible period n histories of actions chosen by the n - 1 agents that have moved so far. $h^n \in H^n$ will denote a specific history as of period n. Under common knowledge of the game's structure and Bayesian rationality, agent's can compute a *public belief*–i.e. a posterior given $h^n = (a^1, a^2, \ldots, a^{n\text{-}1}\}$. We assume for the base case that all agents share the same period-0 prior, and we make our lives easy by assuming that it is a neutral belief, $q^0 = ½$. Each successive agent can update this public prior by applying Bayes rule to q^0 and the information revealed by the action of later agents–i.e. the h^n. Then, each agent applies Bayes rule one more time using the public prior, q^n, and the private signal, x^n, to derive the posterior on which they act.

Within this simple framework, and solving for a perfect Bayesian equilibrium, it is easy to display an information cascade. We do so in the appendix to this paper. The essential fact is that, while still noisy, the public information contained in the actions of previous agents swamps later agents' private information. BHW (1992) actually carry out their analysis on a multiple state × multiple signal × 2-act model. In the context of that model they prove two basic results:

> 1) Information cascades almost always occur as *N* gets large (Proposition 1, pg. 1001); and
>
> 2) Bad cascades are always possible.[34]

---

[34] In a very interesting application of experimental methods, Anderson and Holt (1997) implement a model of the BHW type and find strong support for both the proposition that cascades occur and that bad cascades occur. As predicted by the model, bad cascades occur with lower probability than good cascades.

More interesting than the simple fact of cascades, BHW stress the fragility of cascades. Because cascades may contain little information, and suppress much, a new information release can shatter a cascade, leading to a dramatic change in aggregate beliefs and, thus, behavior. Specifically, BHW prove (pg. 1005-6):

> *Result 3*: The release of a small amount of public information can shatter a long-lasting cascade, where "small amount" refers to a signal less informative than the private signal of a single individual. (pg. 1005)

Given my interest in footloose preferences, this is a particularly interesting fact about social learning in cascade models.

Consider the NAFTA example. Suppose that we start from an equilibrium and, thus, from a cascade involving all citizens. Specifically, suppose we start from a cascade in which all citizens conclude that trade liberalization (whether NAFTA or multilateral) would be harmful to the economic interests of the US.[35] That is, we assume that citizens as a whole believe that trade liberalization is harmful, but this opposition to increased liberalization is highly conditional and subject to large shifts. Now recall that in the run up to the NAFTA vote, virtually all respectable economists and the political leadership of both parties argued very publicly that, essentially, NAFTA was no big deal economically for the US, but that it was important politically (by being important economically for Mexico). The result, as we have already noted, was strong public

[35]Both assumptions here strike me as plausible approximations for the purposes of this example. What poll data exist with respect to trade policy generally strongly suggests that a considerable majority believes further liberalization will be harmful to the US economy–though there is little evidence of support for general increases in protection (though support for sectoral increases is often strong). Furthermore, there is very little evidence that citizens make much of a distinction between preferential and multilateral liberalization. This makes sense. The difference in complexity, from the point of view of a citizen, between NAFTA and, say, a GATT agreement, is trivial.

support for NAFTA. However, the public had not *learned* that liberal trade was good, if by "learned" we mean "identified the true state of the world", but had simply shifted to another weakly held prior. Once NAFTA was passed by Congress, economists in general (and trade economists in particular) not only lost interest in NAFTA, but began to argue that maybe NAFTA wasn't such a great idea after all. It isn't really relevant that the free trade fundamentalist critique of NAFTA was that it distracted political attention from broader trade liberalization. The public listens for the conclusion, not the argument–especially when the argument is, at best, arcane.[36] The result, in the face of continued aggressive public relations against NAFTA, as we have already noted, is that public opinion shifted back to opposition to NAFTA.

Since the fundamental contribution of BHW, the basic information cascade model has been extended in a number of interesting ways. Perhaps the most useful of these from the perspective of political economy modeling involves the introduction of endogenous ordering of moves. Chamley and Gale (1994) develop a model of investment in which the payoff from an investment is an increasing function of the total number of investments, in which the option to invest is private information. Thus, the state of nature (the number of agents with an investment option) can only be revealed by the behavior of other agents.[37] Chamley and Gale are primarily

[36]Abstracting from details of trade creation and trade diversion, which are characterized by complexity considerations of the sort central to this paper, it is notable that virtually all of the recent arguments by economists against regionalism, whatever their validity, are arguments that only an economist would love. Unlike the simple models used to illustrate powerful, but difficult, notions of comparative advantage and gains from trade, which are based on assumptions that isolate the key causal relation generating gains from trade, the political economy arguments used to argue against free trade areas are based on assumptions that seem unrelated to the core processes involved. Their purpose seems more to be to stiffen the spine of the profession in its support of multilateralism than to persuade citizens or their representatives.

[37]Note that this is a pure informational problem, there is no technological externality.

interested in the fact that investment will be delayed as firms wait for others to invest.[38] Depending on the (common) discount rates of the firms, all holders of options invest immediately, no one invests, or some finite delay occurs. There is no social learning in either of the first two cases, but the third is very much like an information cascade. Especially since total investment collapse is possible (i.e. no one ever invests), even though investment is profitable, this is just a (very) bad cascade. In the delay case there is also a cascade that is eventually broken by sufficient impatience and the rapid onset of a good cascade.

Zhang (1997) presents a model of investment cascading that is more closely related to BHW. As in BHW (1992), Zhang presents a 2-state × 2-act model in which payoffs to *ex ante* identical agents are essentially as above [$V_{00} = V_{11} = 2$ and $V_{01} = V_{10} = 0$].[39] In addition to signals, x^j, conditionally distributed as above, every agent's signal is associated with a *signal quality* which is also private information. While all signals are of bounded quality, in the sense that no signal is perfectly informative with respect to $\overline{\omega}$ (the true state of the world) and no signal is perfectly uninformative, some signals are more informative than others. Because all firms have strictly positive discount rates, firms face a tradeoff between acting immediately on their private information and waiting to observe the information revealed by the actions of others. Zhang

[38]Gul and Lundholm (1996) also discuss endogenous ordering of moves, but are primarily interested in the effect of continuous action sets, which we consider below. A recent paper by Aoyagi (1998) synthesizes and extends both Chalmey/Gale and Gul/Lundholm. There are a number of other related papers that are directly related to the question of frenzies (i.e. rapid onset of behavior). Romer (1993), Bulow and Klemperer (1994) and Caplin and Leahy (1994) focus in financial markets, while Lohmann (1993a, b; 1994a, b; 1995) presents an analysis of costly political action seeking to induce a political frenzy.

[39]For technical reasons, Zhang needs a finite cost of investment to prove existence of an equilibrium.

demonstrates that, in equilibrium, the agent with the highest quality signal moves first. However, since all signals are of bounded quality, there is a finite period with no investment. Once some player moves, all other players replicate the action of that player.[40] The higher the common discount rate (i.e. the more costly is waiting), the sooner the cascade starts; but, the more firms there are, the longer each firm waits.[41] Note that, just as in BHW, cascades always start and bad cascades are always possible. Furthermore, since signals are of bounded accuracy, the most accurate signal may still be not very accurate, and the cascade still impounds a large amount of private information, so cascades will still be fragile.

As an example of the application of cascade models with endogenous ordering of actions, we consider the political economy of trade policy reform. vanWijnbergen (1992a, b) develops an argument in favor of rapid policy reform in terms of a model with public information and simultaneous action by all agents. By comparison with the models we have been developing to this point, vanWijnbergen's model involves simultaneous, rather than sequential, moves and a public signal. Specifically, although in vanWijnbergen's development the agents are characterized by *ex ante* heterogeneous priors on the success of the policy reform, it is equivalent to follow the practice (common in game theory, and like BHW) of assuming that agents share a common prior but receive differing signals. In modeling the underlying economy, vanWijnbergen focuses on the important relationship between supply response and sustainability of reform. In a two-period model of price reform in which goods produced in period one can be stored for intertemporal speculation, it is shown that (with a given probability of the reform collapsing) larger reforms

---

[40]Proposition 1, page 198.

[41]Proposition 2, page 199.

produce lower hoarding and larger supply response. As a result, there will be fewer problems with shortages. At the same time, citizens, each of whom has a prior evaluation of the likelihood of program collapse and updates that evaluation on observing the period one supply response, will vote either for the incumbent (who will continue the reform to its completion in period 2) or the opponent (who will reverse the reform) prior to the realization of the period 2 price. All citizens observe the actual supply response (a common public signal) and update according to Bayes rule, so vanWijnbergen is able to characterize the conditions under which the median voter will shift toward (or away from) continuation of the policy.[42] Since gradualism leads to hoarding, and thus to low supply response, citizens' posterior evaluations reflect a greater expectation of policy failure. vanWijnbergen uses this framework to argue in favor of "cold bath" reform rather than gradual reform.

Now consider a model like that of vanWijnbergen (1992a,b), except that instead of a public information release, following Zhang, every agent receives a private signal of varying, but bounded, intensity (but common accuracy). Furthermore, we assume that the policy reform succeeds (i.e. is welfare increasing) with positive probability, but not certainty. For example: some agents get raises; some get reductions in pay; some move to better jobs; some get fired; and so on. Then, prior to the election, every agent has the opportunity to take a public action opposed to, or in favor of, the policy reform. One might plausibly argue that the probability of engaging in protest is increasing in the magnitude of the private signal, and that in the short-term

[42]Note that the signal, while fully informative about the state of the world following the policy reform, is a noisy signal with respect to the likelihood of success of the reform because supply response is chosen strategically by firms. In fact, the policy, if completed, is welfare improving by construction.

there are both more losers and that (at least some of) those losers experience larger negative signals than under a gradual reform.[43] If this is the case, even with heterogeneity, public protest could generate an informational cascade against the cold bath reform leading to reversal where a gradual reform would be sustainable.[44]

The (almost surely empirically important) role of *ex ante* heterogeneous preferences, among a variety of other generalizations, is considered in detail in an important paper by Smith and Sørensen (1998).[45] In this case the distinction, developed by the authors, of convergence in beliefs and convergence in behavior (herding) is essential. The authors are able to show that, not only are cascades (good and bad) possible outcomes, but what they term *confounded learning* must be added to the set of outcomes: "it may well be impossible to draw any clear inference from history even while it continues to accumulate privately-informed decisions". In a standard endogenous policy model based on, say, two goods and two factors of production, with common weakly held priors, it is straightforward to produce a generalization of the NAFTA analysis suggested above. That is, while agents will eventually converge on common beliefs, their behavior (e.g. responses to poll questions about their preferences over policy) will differ due to

---

[43]Such an asymmetry need not be a statistical fact. It could emerge from an asymmetric perception of the form that agents take positive outcomes to reflect personal superiority/ability, while negative outcomes reflect unfair outcomes of the political process.

[44]This is logic is closely related to that presented in Lohmann's (1994a) analysis of the Leipzig demonstrations of 1989-1991 and the collapse of East German communism, though as with the theoretical work on which this paper is based, her work is primarily concerned with costly actions undertaken to induce a cascade (1993a, 1994b, and 1995). Similarly, in Lohmann (1993b) the actions are taken to influence voting behavior prior in an upcoming election.

[45]This paper presents the most detailed investigation into the information cascade model since the original presentation of BHW (1992).

the distributional impact of the policy.

Another important consideration, developed in Moscarini, Ottaviani, and Smith (1998; also see Moscarini and Ottaviani, 1997), is explicit incorporation of a stochastically changing environment. When the environment is known to change stochastically over time, it is therefor known that any state of knowledge is subject to change–that is, that knowledge depreciates in value. As a result, only temporary cascades can be supported. Without new information, the depreciation of old information results in the collapse of cascades. It is easy to imagine this phenomenon playing a large role in the case of large-scale policy reforms, though distinguishing this from the case of multiple public information releases and a single true state would be virtually impossible.

One extension that has been widely noted, while important for understanding the limits of the model, strikes me as of fundamentally less interest to social learning in general, and political economy modeling in particular–*continuous action spaces*. Lee (1993) develops the extension of the BHW model to the case where agents can choose from, say, a continuum rather than a discrete set of actions.[46] Because agents can fine tune their actions conditional on their private information and the state of the public prior, their actions fully reveal the value of their private signal. Thus, by the strong law of large numbers, as the number of agents taking actions becomes large, the true state of the world is revealed with probability 1. Vives (1996) seems to take this,

[46]Specifically, there are still a finite number of states and two signals. Strictly speaking, all Lee requires for his demonstration is that the optimal set of actions for all states be connected. Banerjee (1992) displays a non-fully revealing (i.e. "bad") cascade with a continuum action set, but this result rests on degenerate payoffs: positive payoff obtains only if the agent identifies exactly the correct action. In Lee, agent payoff is increasing in proximity to the correct action. Huck and Oechssler (1998) have recently produced a cascade in a model with two signals in which the optimal action is uniformly distributed on a circle.

and the assumption that signals are of bounded precision, to be a strong criticism of the information cascade model, at least as applied to markets.[47] Gul and Lundholm (1995) make essentially the same point with respect to the endogenous order of actions model. However, it seems to me that the essential element of social learning is precisely that behavior provides coarse evidence of any agent's private information. This is, perhaps, especially true in the political context where observable actions are few (e.g. {participate, abstain}). Similarly, the assumption that signals are of bounded accuracy is simply a trope representing the complexity of the decision situations we have been considering.

To this point we have been considering what Gul and Lundholm (1995) usefully term *statistical cascades*. That is, the cascade emerges simply as a function of the Bayesian rationality of agents in a sequential decision problem.[48] We finish our review by considering what Gul and Lundholm call *reputational cascades*, i.e. cascades driven by an agency relation embedded in the sequential decision problem. The central reference here is Scharfstein and Stein (1990), who consider a model of investment by managers in an environment characterized by common prediction error in their decision to make one or another investment.[49] This means that owners, in

[47]However, in related work, Vives (1997) considers a model of social learning with a continuous action set in which the public information is observed with noise. In the context of this model, although completely accurate learning eventually occurs, learning is slow.

[48]In fact, Smith and Sørensen (1997) develop an identity between the structure of the information cascade model, and the model of sequential experimentation in statistical decision theory.

[49]See Zwiebel (1995), Jeon (1998), and Khanna (1998) for related work. Jeon (1998) and Khanna (1998) are closer to Scharfstein and Stein, and both are interested in outcomes when owners adopt optimal contracts to align manager incentives more closely with their own. Zwiebel (1995) considers a different information structure, but still generates a kind of herd behavior.

evaluating the performance of managers, consider both outcomes and whether a given manager did the same thing that other managers did. This creates an incentive for herding, even if there is no convergence in beliefs. As in BHW, in Scharfstein and Stein's model, the manager that acts second will follow the behavior of the first manager independently of his private signal, thus impounding socially useful information in a cascade.[50]

Reputational cascades would seem to be a particularly useful tool in the attempt to understand rapid onset of adoption of substantial changes in policy regimes. Consider the case of the fall of import-substituting industrialization (ISI), the rise of export-led growth (XLG), and the possible fall of XLG. The widespread adoption of ISI policies in the immediate post-War/post-Independence period, and the subsequent shift to XLG by many of the same governments is a standard part of the lore of development economics.[51] Clearly one could tell a statistical cascade story consistent with these facts, but the active role of public and private international financial institutions suggests a substantial payoff from examining reputational aspects in this case. That is, suppose that these international institutions, in making decisions about the allocation of their resources, evaluate the performance of governments not only in terms of fundamentals, but relative to the performance of other, say, LDC governments. As suggested in the models of Scharfstein and Stein, and others, this could create incentives for governments to ignore private

[50]In fact, the cascade happens faster, and more certainly, in Scharfstein and Stein than in BHW. In the latter case a cascade cannot happen until the third decision maker at the earliest, and could take an arbitrarily long time (i.e. this would require a long sequence of strictly alternating signals with appropriate coin flip outcomes). It should also be noted that reputational cascades are not fragile, at least not in the same straightforward sense, as they are in BHW.

[51]See Little (1982) for a nice, chronological presentation of this story. Bruton (1998) is also very useful in this respect.

information and herd on policies. The incentives to herd might be particularly strong if this reputational cascade were contained in a statistical cascade among economists. That is, suppose that the international institutions, in evaluating country performance, consider not only absolute performance and performance relative to other clients, but also degree of deviation from the consensus of professional development and trade economists (where this last factor is weighted by the degree of that consensus). Even if there is only weak information relative to the performance of even widely different policy regimes, a herd among economists could impose quite tight reputational constraints on policy and the opportunity for potentially useful experimentation.[52]

Conclusion

It is important to be clear that, while the class of models considered here is positive, they are not predictive over the domain of final political-economic outcomes. These models do not avoid, in fact they rest on (or, more accurately, provide a formal representation of) the fact that, with appropriately chosen priors, we can reproduce virtually any final outcome.[53] Thus, any predictions of these models with respect to final outcomes are vacuous. Nonetheless, I hope that

---

[52]There certainly is considerable casual evidence supporting the notion that economists, as a result of their rigorous and highly uniform training, are more prone to behavioral herding than other social scientists. The experiments of Marwell and Ames (1981) suggest some slightly more systematic evidence for this, though see the survey by Ledyard (1995) for qualification. There is certainly sufficient evidence of herding in both research topics and policy advice to render this a plausible area of concern. However, both Little (1986) and Krueger (1997) make compelling cases for the argument that learning by economists, and governments, has more in common with learning the true state of the world than an informational cascade.

[53]It is the *re*production of outcomes that is most worrying in social scientific analysis. Thin predictions (unless strictly unfalsifiable) will not last, but compelling post-dictions ("stylized facts") can sustain empirically weak theoretical analyses for long periods.

this paper, and the examples it contains, suggest that ignorance and learning in a social context are issues of first rate importance, both as empirical phenomena and as potential determinants of trade policy outcomes. Furthermore, there are, I think, several implications of these models (and of ignorance/learning more generally) for us as policy analysts. I will consider two: one with respect to evaluating the predictive content of our positive models; and the second with respect to the role of economists in the public discourse over trade policy.

There is no substitute for basing predictive political-economy models on political and economic fundamentals. We have good reason to expect such fundamentals to play a central role in determining trade policy, and we have equally good reason to predict the direction of the effects. However, if learning effects also play a, largely unpredictable, role, we also need to expect prediction errors that are occasionally large. That is, the right kind of ignorance can yield wildly different outcomes from those predicted by models. An excellent example, away from the trade policy focus of this paper, is the poor performance of macro political-economy models in the 1992 Bush-Clinton election (Haynes and Stone, 1994). While the data seem to show recovery of the economy, and thus success for the incumbent, there was widespread perception that the economy was still in an economic crisis ("It's the economy stupid"). The result, as they say, is history.

As noted in the introduction and suggested by the discussion of the NAFTA case, one of the most interesting implications of learning models with information cascades as a prediction is the suggestion of a major role for policy analysts. Some of the recent attempts to justify an active policy role for economists turn on difficult philosophical issues of freedom of choice that seem rather removed from the actual practice of participation by economists in the public policy discourse. The advice on how and whom to advise that emerges from this kind of argument seems of limited use.[54] The problem seems to emerge from taking our models seriously where we should not. I have just argued that these models serve a very useful positive purpose in understanding and predicting public policy. However, when we abstract from complexity and uncertainty in the interest of building parsimonious models, we have abstracted from the most obvious warrant for an active advisory role. When agents, whether citizens or policy-makers, are highly uncertain about the workings of the economy (i.e. most of the time), expert advice can have a substantial effect on final outcomes via precisely the channels identified in learning models.[55] It makes perfectly good sense for citizens and politicians to listen to, and even to seek out, the advice of economists because that advice is better informed than much of the policy advice that is given during a political process–though note that this need not be even vaguely perfect information. A public signal of strong agreement among economists, especially when supported by compelling evidence, during a political process can have the effect of a public information release in the models discussed above. Even if policy-makers, or citizens, believe that this information is less informative than any individual privately observed signal, such a public release can have the effect of reopening the public discussion and dramatically shifting the

---

[54]This is the entering wedge for Dixit's (1997) comment on O'Flaherty and Bhagwati (1997).

[55]Just as trade and political economy models abstract from informational issues to focus on the causal forces of most immediate interest, learning models of the sort developed in this paper abstract from a variety of complexities to highlight the effects of ignorance and learning in a social context. In particular, these models abstract from the important, and complementary, forces that make advisors participants (in a game theoretic sense) in the political process. Dixit (1997) provides a very nice sketch, with appropriate references, of models which highlight the role of advice giving in a strategic environment with asymmetric information.

structure of governmental or public opinion.[56]

It seems to me that a focus on this non-strategic, informational role of economic advice has useful implications. To the extent that the warrant for advice-giving is uncertainty, as much about the working of the economy as about simple facts, it seems particularly fruitless to give advice based on the presumption that those receiving it are well-trained economists.[57] We need to convince our auditors that a consensus on fundamental issues related to, say, trade policy, exists, and we need to do so in ways that are clear to relatively engaged, relatively intelligent non-economists. This clearly means that complicated arguments, requiring many closely argued steps, and knowledge of economic theory, are likely to be unsuccessful. However, as Matthew Slaughter (1998) argues, while outright lies may be successful in the short-run (e.g. "NAFTA will create thousands of jobs"), sooner-or-later they are likely to backfire. The most successful of our public representatives–e.g. Milton Friedman, Alan Blinder, Paul Krugman–seem to identify simple but compelling metaphors, which are mixed with a small number of striking facts, to argue for a single clear policy point. Finally, it should be noted that, if social learning does not produce knowledge of the intertemporally sturdy type, but rather of the type suggested by informational cascade models, we need to be prepared to stay engaged in the public discourse beyond the

---

[56]Although we are presumably better informed, as we noted above, economists are at least as prone to being trapped in cascades as any other rational agents engaged in learning about the world, possibly more so. Because economists have very similar understandings of the workings of the economy–due to strong socialization–we may behave more like the agents in the BHW models than any other group. Recall that Smith and Sørensen (1998) find much more complex aggregate behavior characterizing groups with heterogeneous preferences than in groups with homogeneous preferences.

[57]While admirable, the attempt to transform citizens and policy-makers into economists is almost certainly doomed to failure.

passage of any particular piece of legislation.

Table 1: Typology of Learning Models in Economics

	Individual Learning			Social Learning	
	Private Signal	Public Signal	Strategic	Non-Strategic	Strategic
Bounded Rationality	Rules of Thumb & Classical Decision Theory	Convergence to REE with Rule of Thumb Learning	Adaptive learning in oligopoly	Rule of Thumb Learning from Others	Fixed Rules: Evolutionary Games, Finite Automata
Bayesian Rationality	Bayesian Decision Theory	Convergence to REE with Bayesian Learning	Repeated Games with Imperfect Observability	Getting to Common Knowledge and Common Belief	Fictitious Play & Full Bayesian Rationality
Practical Reason	Analytical Epistemology			Foundations of Common Knowledge	Behavioral Foundations

Appendix:
A 2-State × 2-Signal Example

In this appendix we illustrate the information cascade in a model with 2 states, $\Omega = \{\omega_0, \omega_1\}$, and 2 signals, $X = \{x_0, x_1\}$--drawn essentially from BHW (1992).[58] We start with the single agent case, retaining notation and assumptions from the main text of the paper. Suppose that the private signal is x_j. From a prior belief that the state is i, $\pi(\omega_i)$, where $f(x_j|\omega_i)$ is the probability of observing x_j conditional on the actual state being ω_i, and where $f(x_j)$ is the unconditional probability of observing x_j, Bayes' rule gives the posterior probability that the actual state is ω_i conditional on having observed x_j as:

$$\pi(\omega_i | x_j) = \frac{\pi(\omega_i) f(x_j|\omega_i)}{f(x_j)} = \frac{\pi(\omega_i) f(x_j|\omega_i)}{\sum_{k\in\Omega} \pi(\omega_k) f(x_j|\omega_k)}. \qquad \text{(Bayes' Rule)}$$

For, say, $\pi(\omega_1 | x_1)$, and recalling the flat prior $\left[\pi(\omega_1) = \pi(\omega_0) = \tfrac{1}{2}\right]$ and symmetric likelihoods [i.e. $f(x_j|\omega_i) = p > ½$ for $i = j$ and $f(x_j|\omega_i) = (1 - p)$ for $i \neq j$]:

$$\pi(\omega_1|x_1) = \frac{\pi(\omega_1)p}{\pi(\omega_1)p + \pi(\omega_0)(1- p)} = \frac{p}{p + (1-p)} = p.$$

Then agents select the optimal action with respect to the posterior distribution $\pi(\omega| x)$.

---

[58]Section 1 of Moscarini and Ottaviani (1997) also presents the 2 × 2 case, and I have made use of that exposition as well.

However, in the basic information cascade model, agents take their actions in a fixed sequence, and the *actions* of all previous agents are observable. Thus, except for the first agent, all agents possess the public information of the agents preceding them, as well as their private information. This information is contained in the *actions* of the previous agents, the *signals* of the previous agents are unobservable. We now need to characterize this public information and the way it is incorporated in the decision strategies of the agents. First, we need to represent the *history* prior to any agent's decision. Let $n \in N$ index agents (we will sometimes refer to this as *periods*, i.e. one agent acts in each "period")–so that we refer to agent n as DM^n. For $n \geq 2$, denote by $H^n \equiv \{a_0, a_1\}^{n\text{-}1}$ the space of all possible period n histories of actions chosen by the n - 1 agents that have moved so far. $h^n \in H^n$ will denote a specific history. Under common knowledge of the game's structure and Bayesian rationality, agent's can compute a *public belief*–i.e. a posterior given $h^n = (a^1, a^2, \ldots, a^{n\text{-}1}\}$. We assume for the base case that all agents share the same prior, and we make our lives easy by assuming that it is a neutral time-0 belief, p = ½. That is, DM^1 updates according to Bayes' rule. DM^2 updates the flat public prior to

$$q_i^2 \equiv q(\omega_i | h^2) = \frac{p(h^1|\omega_i)}{p(h^1|\omega_0) + p(h^1|\omega_1)} \qquad \textbf{(1)}$$

Successive *DM*s update the flat public prior to:

$$q_i^n \equiv q(\omega_i | h^n) = \frac{p(h^n|\omega_i)}{p(h^n|\omega_0) + p(h^n|\omega_1)} \qquad \textbf{(2)}$$

A final application of Bayes' rule, given the private signal, x^n of agent $n > 1$, gives agent n's posterior. That is, $q(\omega_i|\ h^n)$ becomes the prior, $q(\omega_i)$, and we update after observing x^n according to

$$\pi_{ij}^n \equiv \pi\left(\omega_i \middle| h^n, x_j\right) = \frac{q_i^n\ f\left(x_j \middle| \omega_i\right)}{q_0^n\ f\left(x_j \middle| \omega_0\right) + q_1^n\ f\left(x_j \middle| \omega_1\right)}. \quad \textbf{(3)}$$

With these posteriors in hand, we now need to characterize the procedure by which *DM*s select their optimal strategies–$a^n = \alpha^n(x^n, h^n)$.

Recall that we have assumed a common utility function, $V(a_k, \omega_i) \equiv V_{ki}$, with the properties that: $V_{00} = V_{11} = 1$; and $V_{01} = V_{10} = 0$. Thus, we seek to maximize posterior expected utility. Here, we exploit the fact that, for a fixed x, $q(\omega|x) \propto q(\omega)\ f\ (x|\omega)$. If $\alpha(x) = a_0$, the expected posterior utility is proportional to

$$q_0\ f\left(x \middle| \omega_0\right) V_{00} + q_1\ f\left(x \middle| \omega_1\right) V_{01}$$

If $\alpha(x) = a_1$, the expected posterior utility is proportional to

$$q_0\ f\left(x \middle| \omega_0\right) V_{10} + q_1\ f\left(x \middle| \omega_1\right) V_{11}$$

Then $\alpha(x) = a_0$ only if

$$q_0\ f\left(x \middle| \omega_0\right) V_{00} + q_1\ f\left(x \middle| \omega_1\right) V_{01} > q_0\ f\left(x \middle| \omega_0\right) V_{10} + q_1\ f\left(x \middle| \omega_1\right) V_{11}$$

Rearranging gives

$$q_0\ f\left(x \middle| \omega_0\right)\left[V_{00} - V_{10}\right] > q_1\ f\left(x \middle| \omega_1\right)\left[V_{11} - V_{01}\right].$$

Since we have assumed that $V_{00} > V_{10}$ and $V_{11} > V_{01}$, this implies that $\alpha(x) = a_0$ when:

$$\frac{f(x|\omega_1)}{f(x|\omega_0)} < \frac{q_0[V_{00} - V_{10}]}{q_1[V_{11} - V_{01}]} = \frac{q_0}{q_1}.$$

The equality follows from our assumption that $V_{00} = V_{11} = 1$ and $V_{01} = V_{10} = 0$. Of course, $\alpha(x) = a_1$ when the inequality is reversed. Thus, we end up with a likelihood ratio test. Define:

$k^n \equiv \dfrac{q_0^n [V_{00} - V_{10}]}{q_1^n [V_{11} - V_{01}]} = \dfrac{q_0^n}{q_1^n}$. Then, given signal x_j, DM^n selects $\alpha(x_j) = a_0$ if:

$$\left\{ x^n : \frac{f(x^n|\omega_1)}{f(x^n|\omega_0)} < k^n \right\}. \qquad (\boldsymbol{LR}^n)$$

Within this simple framework, it is easy to display an information cascade. Suppose that DM^1 receives the signal $x = x_0$. DM^1 seeks to update using Bayes' rule

$$\pi(\omega_i|x_j) = \frac{\pi(\omega_i) f(x_j|\omega_i)}{f(x_j)} = \frac{\pi(\omega_i)\, p}{\pi(\omega_0)\, p + \pi(\omega_1)(1 - p)}.$$

We have assumed that $\pi_1 = \pi_2 = ½$., so, $\pi(\omega_0|x_0) = \dfrac{p}{p + (1 - p)} = p > \dfrac{1}{2}$. DM^1 now applies

LR^1 to choose her strategy: $k^1 = \pi_1/\pi_2 = 1$; and $\dfrac{f\left(x_0 \middle| \omega_1\right)}{f\left(x_0 \middle| \omega_0\right)} = \dfrac{1-p}{p} < 1.$ So, by LR^1 DM^1 selects a_0^1. Now, in addition to his signal, x^2, DM^2 observed that $\alpha^1(x) = a^1$ (thus, $h^2 = \{a_0^1\}$). Thus, DM^2 knows that DM^1 observed x_0, so $q_0^2 = p$ and $q_1^2 = (1 - p)$. DM^2 applies LR^2 to select $\alpha^2(x^2, h^2)$: $k^2 = \dfrac{q_0^2}{q_1^2} = \dfrac{p}{1-p} > 1.$ Now there are two possibilities. Suppose that $x^2 = x_0$. Then, $\dfrac{f\left(x_0^2 \middle| \omega_1\right)}{f\left(x_0^2 \middle| \omega_0\right)} = \dfrac{1-p}{p} < 1 < k^2.$ Thus, DM^2 also selects $\alpha(x^2, h^2) = a_0$. Suppose, instead, that $x^2 = x_1$. Now, $\dfrac{f\left(x_1^2 \middle| \omega_1\right)}{f\left(x_1^2 \middle| \omega_0\right)} = \dfrac{p}{1-p} = k^2.$ We now need a tie-breaking rule. Suppose we assume that in cases where $LR^n = k^n$, DM^n flips a fair coin.

DM^3 observes a signal, x^3, and one of two possible histories: $h^3 = [\{a_0^1, a_0^2\}$ or $\{a_0^1, a_1^2\}]$. Suppose that $h^3 = \{a_0^1, a_1^2\}$. To find k^3, we need to update the public prior

$$q^3 = \pi\left(\left\{a_0^1, a_1^2\right\} \middle| \omega_0\right) = \frac{p(1-p)}{p(1-p) + (1-p)p}.$$

Since this is equal to ½, we are back to the flat prior, and DM^3 acts exactly like DM^1, that is she concludes that the state of the world is the one given in her signal and acts accordingly. Suppose instead that $h^3 = \{a_0^1, a_0^2\}$. Again, we need the public prior

$$q_0^3 = \pi\left(\left\{a_0^1, a_0^2\right\}\middle|\omega_0\right) = \frac{pp}{pp + (1-p)(1-p)}.$$

Thus, we can calculate k^3 as

$$k^3 \equiv \frac{q_0^3}{q_1^3} = \frac{q_0^3}{1-q_0^3} = \frac{p^2}{(p-1)^2}.$$

Now we need to consider two cases. Suppose the private signal is $x^3 = x_0$, so

$\dfrac{f\left(x_0^3\middle|\omega_1\right)}{f\left(x_0^3\middle|\omega_0\right)} = \dfrac{1-p}{p} < 1 < k^3$. Thus, DM^3 also selects $\alpha(x^3, h^3) = a_0$. Suppose instead that the private signal is $x^3 = x_1$, $\dfrac{f\left(x_1^2\middle|\omega_1\right)}{f\left(x_1^2\middle|\omega_0\right)} = \dfrac{p}{1-p} < k^3$. Note that we no longer need a tie-breaking rule, $\alpha(x^3, h^3) = a_0$ for this case also. This is the information cascade. If the first two *DM*s take the same action, everyone else does also, independently of their private signals. Thus, the information of everyone in the cascade is essentially lost.

Note the way this works: as soon as two *DM*s in succession take the same action, all following *DM*s also take that action, independently of the x^n. Since, for a sufficiently long string of *DM*s, two of the same actions in succession will occur almost surely, cascades occur almost surely in this model. Since the $f(x_i \mid \omega_j) > 0$ for $i \neq j$, bad cascades are always possible. In fact, under the coin flip tie-breaking rule, the probability of a bad cascade is greater than that implied by $f(x_i \mid \omega_j) > 0$ for $i \neq j$ since, in addition to the probability of two false signals in succession, there is also the probability of two different signals in succession times ½ (i.e. the coin flip).

References

Aoyagi, Masaki (1998). "Equilibrium Delay in a Simple Dynamic Model of Investment". *Economic Theory*; V. 12-#1, pp. 123-146.

Anderson, Lisa and Charles Holt (1997). "Information Cascades in the Laboratory". *American Economic Review*; V.87-#5, pp. 847-862.

Aumann, Robert and Michael Maschler (1995). *Repeated Games with Incomplete Information*. Cambridge: MIT Press.

Banerjee, Abhijit (1992). "A Simple Model of Herd Behavior". *Quarterly Journal of Economics*; V.107-#3, pp. 797-817.

Battigalli, Pierpaolo and Giacomo Bonanno (1998). "Recent Results on Belief, Knowledge and the Epistemic Foundations of Game Theory". ms: UC Davis, Department of Economics.

Battigalli, Pierpaolo, Mario Gilli, and M. Cristina Molinari (1992). "Learning and Convergence to Equilibrium in Repeated Strategic Interactions: An Introductory Survey". *Richerche Economiche*; V.46-#3/4, pp. 335-378.

Becker, Gary (1991). "A Note on Restaurant Pricing and Other Examples of Social Influence on Price". *Journal of Political Economy*; V.99-#5, pp. 1109-1116.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch (1992). "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades". *Journal of Political Economy*; V.100-#5, pp.992-1026.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch (1998). "Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades". *Journal of Economic Perspectives*; V.12-#3, pp. 151-170.

Blackwell, David, and M.A. Girshick (1954). *Theory of Games and Statistical Decisions*. New York: John Wiley and Sons.

Blume, Lawrence and David Easley (1984). "Rational Expectations Equilibrium: An Alternative Approach". *Journal of Economic Theory*; V.34-#1, pp. 116-129.

Blume, Lawrence and David Easley (1995). "What has the Rational Learning Literature Taught Us?". in Alan Kirman and Mark Salmon, eds. *Learning and Rationality in Economics*. Oxford: Blackwell, pp. 12-39.

Bowles, Samuel (1998). "Endogenous Preferences: The Cultural Consequences of Markets and other Economic Institutions”. *Journal of Economic Literature*; V.36-#1, pp. 75-111.

Bruton, Henry (1990). “A Reconsideration of Import Substitution”. *Journal of Economic Literature*; V.36-#2, pp. 903-936.

Bulow, Jeremy and Paul Klemperer (1994). “Rational Frenzies and Crashes”. *Journal of Political Economy*; V.102-#1, pp. 1-23.

Bray, Margaret and David Kreps (1988). “Rational Learning and Rational Expectations”. in G. Feiwel, ed. *Arrow and the Ascent of Modern Economic Theory*. London: Macmillan, pp. 597-625.

Caplin, Andrew and John Leahy (1994). “Business as Usual, Market Crashes, and Wisdom After the Fact”. *American Economic Review*; V.84-#3, pp. 548-565.

Chamley, Christophe and Douglas Gale (1994). “Information Revelation and Strategic Delay in a Model of Investment”. *Econometrica*; V.62-#5, pp. 1065-1085.

Chellas, Brian F. (1980). *Modal Logic: An Introduction*. Cambridge: Cambridge University Press.

Choi, Jay Pil (1997). “Herd Behavior, the ‘Penguin Effect’, and Suppression of Informational Diffusion: An Analysis of Informational Externalities and Payoff Interdependency”. *RAND Journal of Economics*; V.28-#3, pp. 407-425.

Coleman, James S. (1990). *Foundations of Social Theory*. Cambridge: Belknap Press of Harvard University Press.

Conlisk, John (1996). “Why Bounded Rationality?”. *Journal of Economic Literature*; V.34-#2, pp. 669-700.

David, Paul A. (1985). “Clio and the Economics of QWERTY”. *American Economic Review*; V.75-#2, pp. 332-37.

Dekel, Eddie and Faurk Gul (1997). “Rationality and Knowledge in Game Theory”. in David Kreps and Kenneth Wallis, eds. *Advances in Economics and Econometrics: Theory and Applications, Vol. 1*. Cambridge: Cambridge University Press, pp. 87-172.

Dixit, Avinash (1997). “Economists as Advisers to Politicians and Society”. *Economics and Politics*; V.9-#3, pp. 225-230.

Dybvig, Philip and Chester Spatt (1983). “Adoption Externalities as Public Goods”. *Journal of Public Economics*; V.20-#2, pp. 231-247.

Eichberger, Jürgen (1992). “Nash Equilibrium and Learning”. in J. Borland, J. Creedy, and J. Eichberger, eds. *Recent Developments in Game Theory*. Brookfield, Vt.: Gower.

Evans, George and Seppo Honkapohja (1995). “Adaptive Learning and Expectational Stability”. in Alan Kirman and Mark Salmon, eds. *Learning and Rationality in Economics*. Oxford: Blackwell, pp. 102-126.

Elster, Jon (1989). *The Cement of Society*. Cambridge: CUP.

Fagin, Ronald, Joseph Halpern, Yoram Moses, and Moshe Vardi (1995). *Reasoning About Knowledge*. Cambridge: MIT Press.

Farrell, Joseph and Garth Saloner (1985). “Standardization, Compatibility, and Innovation”. *Rand Journal of Economics*; V.16-#1, pp. 70-83.

Feldman, Mark (1987). “Bayesian Learning and Convergence to Rational Expectations”. *Journal of Mathematical Economics*; V.16-#3, pp. 297-313.

Frankfurt, Harry (1971). “Freedom of Will and the Concept of a Person”. *Journal of Philosophy*; V.68-#1, pp. 5-20.

Froot, Kenneth, David Scharfstein, and Jeremy Stein (1992). “Herd on the Street: Informational Inefficiencies in a Market with Short-Term Speculation”. *Journal of Finance*; V.47-#4, pp. 1461-1484.

Frydman, Roman (1982). “Towards an Understanding of Market Process: Individual Expectations and Convergence to a Rational Expectations Equilibrium”. *American Economic Review*; V.72-#4, pp. 652-668.

Fudenberg, Drew and David Levine (1998). *The Theory of Learning in Games*. Cambridge: MIT.

Gale, Douglas (1996). “What Have we Learned from Social Learning?”. *European Economic Review*; V.40-#3/5, pp. 617-628.

Gärdenfors, Peter (1988). *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. Cambridge: MIT Press.

Gilboa, Itzhak and David Schmeidler (1995). “Case-Based Decision Theory”. *Quarterly Journal of Economics*; V.110-#?, pp. 605-639.

Giddens, Anthony (1986). *The Constitution of Society: Outline of the Theory of Structuration*. Berkeley: University of California Press.

Goldstein, Judith (1988). “Ideas, Institutions and American Trade Policy”. *International Organization*; V.42-#1, pp. 179-217.

Goldstein, Judith and Robert Keohane, eds. (1993). *Ideas and Foreign Policy: Beliefs, Institutions, and Political Change*. Ithaca: Cornell University Press.

Gul, Faruk and Russell Lundholm (1995). “Endogenous Timing and the Clustering of Agents’ Decisions”. *Journal of Political Economy*; V.103-#5, pp. 1039-1066.

Haas, Ernst (1958). *The Uniting of Europe: Political Social and Economic Forces, 1950-1957*. Stanford: Stanford University Press.

Haas, Peter (1992). “Introduction: Epistemic Communities and International Policy Coordination”. *International Organization*; V.46-#1, pp. 1-36.

Haynes, Stephen and Joe Stone (1994). “Why Did Economic Models Falsely Predict a Bush Landslide in 1992”. *Contemporary Economic Policy*; V.12-#2, pp. 123130.

Hirshleifer, David (1995). “The Blind Leading the Blind: Social Influence, Fads, and Informational Cascades”. in Mariano Tommasi and Kathryn Ierulli, eds. *The New Economics of Human Behavior*. Cambridge: CUP, pp. 188-215.

Holland, John, Keith Holyoak, Richard Nisbet and P.R. Thagard (1987). *Induction: Process of Inference, Learning and Discovery*. Cambridge: MIT Press.

Holland, John (1996). “The Rationality of Adaptive Agents”. in Kenneth Arrow, *et al*. eds., *The Rational Foundations of Economic Behavior*. London: Macmillan, pp. 281-297.

Huck, Steffen and Jörg Oechssler (1998). “Informational Cascades with Continuous Action Spaces”. *Economics Letters*; V.60-#2, pp. 163-166.

Hughes, G.E. and M.J. Cresswell (1996). *A New Introduction to Modal Logic*. London: Routledge.

Jeffrey, Richard (1974). “Preferences Among Preferences”. *Journal of Philosophy*; V.71-#13; pp. 377-391.

Jeon, Seonghoon (1998). “Reputational Concerns and Managerial Incentives in Investment Decisions”. *European Economic Review*; V. 42-#7, pp. 1203-1219.

Jordan, James and Roy Radner (1982). “Rational Expectations in Microeconomic Models: An Overview”. *Journal of Economic Theory*; V.26-#2, pp. 201-223.

Kalai, Ehud (1990). “Bounded Rationality and Strategic Complexity in Repeated Games”. in Tatsuro Ichiishi, Abraham Neyman and Yai Tauman, eds. *Game Theory and Applications*. San Diego: Academic Press, pp. 131-157.

Katz, Michael and Carl Shapiro (1985). “Network Externalities, Competition and Compatibility”. *American Economic Review*; V.75-#3, pp. 424-440.

Kauffman, Stuart (1993). *The Origins of Order: Self-Organization and Selection in Evolution*. New York: Oxford University Press.

Kfir-Dahav, Noa and Moshe Tennenholtz (1996). “Multi-Agent Belief Revision”. in Yoav Shoham, ed. *Theoretical Aspects of Reasoning About Knowledge– Proceedings of the Sixth Conference*. Los Altos: Morgan Kaufmann, pp. 175-194.

Khanna, Naveen (1998). “Optimal Contracting with Moral Hazard and Cascading”. *Review of Financial Studies*; V.11-#3, pp. 557-594.

Kiefer, Nicholas M. (1988/9). “Optimal Collection of Information by Partially Informed Agents”. *Econometric Reviews*; V.7-#2 1988-89, pp. 113-48.

Kiefer, Nicholas and Yaw Nyarko (1995). “Savage-Bayesian Models of Economics”. in Alan Kirman and Mark Salmon, eds. *Learning and Rationality in Economics*. Oxford: Blackwell, pp. 40-62.

Krueger, Anne (1997). “Trade Policy and Economic Development: How We Learn”. *American Economic Review*; V.87-#1, pp. 1-22.

Ledyard, John (1995). “Public Goods: A Survey of Experimental Research”. in John Kagel and Alvin Roth, eds. *The Handbook of Experimental Economics*. Princeton: PUP, pp. 111-194.

Lee, In Ho (1993). “On the Convergence of Informational Cascades”. *Journal of Economic Theory*; V.61-#2, pp. 395-411.

Leibenstein, Harvey (1950). “Bandwagon, Snob and Veblen Effects in the Theory of Consumer Demand”. *Quarterly Journal of Economics*; V.64-#2, pp. 183-207.

Levy, Jack (1994). “Learning and Foreign Policy: Sweeping a Conceptual Minefield”. *International Organization*; V.48-#2, pp. 279-312.

Levy, Philip (1998). “Learning from Trade Agreements”. ms: Department of Economics, Yale University.

Lewis, David (1966). *Convention: A Philosophical Study*. Cambridge: Harvard University Press.

Little, Ian M.D. (1982). *Economic Development: Theory, Policy, and International Relations*. New York: Basic Books/Twentieth Century Fund.

Lipman, Barton L. (1995). “Information Processing and Bounded Rationality: A Survey”. *Canadian Journal of Economics*; V.28-#1, pp. 42-67.

Lohmann, Susanne (1993a). “A Signaling Model of Informative and Manipulative Political Action”. *American Political Science Review*; V.87-#2, pp. 319-333.

Lohmann, Susanne (1993b). “A Welfare Analysis of Political Action”. in William Barnett, Melvin Hinich, and Norman Schofield, eds. *Political Economy: Institutions, Competition and Representation*. Cambridge: CUP, pp. 437-461.

Lohmann, Susanne (1994a). “The Dynamics of Informational Cascades: The Monday Demonstrations in Leipzig, East Germany, 1989-1991”. *World Politics*; V.47-#1, pp. 42-101.

Lohmann, Susanne (1994b). “Information Aggregation Through Costly Political Action”. *American Economic Review*; V.84-#3, pp. 518-530.

Lohmann, Susanne (1995). “A Signaling Model of Competitive Political Pressures”. *Economics and Politics*; V.7-#3, pp. 181-206.

Mailath, George (1998). “Do People Play Nash Equilibrium? Lessons from Evolutionary Game Theory”. *Journal of Economic Literature*; V.36-#3, pp. 1347-1374.

Marimon, Ramon (1997). “Learning from Learning in Economics”. in David Kreps and Kenneth Wallis, eds. *Advances in Economics and Econometrics: Theory and Application–V.I*. Cambridge: CUP, pp. 278-315.

Marimon, Ramon and Ellen McGratten (1995). “On Adaptive Learning in Strategic Games”. in Alan Kirman and Mark Salmon, eds. *Learning and Rationality in Economics*. Oxford: Blackwell, pp. 63-101.

Marwell, Gerald and Ruth Ames (1981). “Economists Free Ride, Does Anyone Else?: Experiments on the Provision of Public Goods, IV”. *Journal of Public Economics*; V.15-#3, pp. 295-310.

Matsushima, Hitoshi (1997). “Bounded Rationality in Economics: A Game Theorist’s View”. *Japanese Economic Review*; V.48-#3, pp. 293-306.

McKelvey, Richard and Talbot Page (1986). “Common Knowledge, Consensus, and Aggregate Information”. *Econometrica*; V.54-#1, pp. 109-127.

Mongin, Philippe (1994). “Some Connections between Epistemic Logic and the Theory of Non-Additive Probability”. In P. Humphreys, ed. *Patrick Suppes: Scientific Philosopher–Vol 1*. Boston: Kluwer; pp. 135-171.

Moscarini, Giuseppe and Marco Ottaviani (1997). “Economic Models of Social Learning”. in Peirpaolo Battigalli, Aldo Montesano, and Fausto Panunzi, eds. *Decisions, Games and Markets*. Boston: Kluwer, pp.265-298.

Moscarini, Giuseppe, Marco Ottaviani, and Lones Smith (1998). “Social Learning in a Changing World”. *Economic Theory*; V.11-#?, pp. 657-665.

Nelson, Douglas (1989) "The Domestic Political Preconditions of US Trade Policy: Liberal Structure and Protectionist Dynamics". *Journal of Public Policy*. V.9-#1, pp. 83-108.

Nelson, Douglas (1995). “Political Economy of Protection for the US Automobile Industry”. in A. Krueger, ed. *Political Economy of American Trade Policy*. Chicago: University of Chicago Press/NBER, pp. 133-191.

Nelson, Douglas (1999). “Problems in the Political Economy of Trade Policy Reform”. *Journal of International Trade and Economic Development*; forthcoming.

O’Flaherty, Brendan and Jagdish Bhagwati (1997). “Will Free Trade with Political Science put Normative Economists out of Work?”. *Economics and Politics*; V9-#3, pp. 207-219.

Prigogine, Ilya and Isabelle Stengers (1984). *Order Out of Chaos: Man’s New Dialogue with Nature*. New York: Bantam Books.

Radner, Roy (1982). “Equilibrium under Uncertainty”. in K. Arrow and M. Intriligator, eds. *Handbook of Mathematical Economics–Vol. II*. Amsterdam: North-Holland, pp. 923-1006.

Romer, David (1993). “Rational Asset Price Movement without News”. *American Economic Review*; V.83-#5, pp.1112-1130.

Rubinstein, Ariel (1987). “The Complexity of Strategies and the Resolution of Conflict: An Introduction”. in R. Bryant and R. Portes, eds. *Global Macroeconomics: Policy Conflict and Cooperation*. New York: St. Martins/CEPR, pp. 17-32.

Rubinstein, Ariel (1998). *Modeling Bounded Rationality*. Cambridge: MIT Press.

Rueschemeyer, Dietrich and Theda Skocpol, eds. (1996). *States, Social Knowledge, and the Origins of Modern Social Policies*. Princeton: Princeton University Press.

Ruggie, John (1982). “International Regimes, Transactions, and Change: Embedded Liberalism in the Post-War Order". *International Organization*; V.36-#2, pp. 379-415.

Sargent, Thomas (1993). *Bounded Rationality and Macroeconomics*. Oxford: OUP.

Savage, L.J. (1954). *The Foundations of Statistics*. New York: Wiley.

Scharfstein, David and Jeremy Stein (1990). "Herd Behavior and Investment". *American Economic Review*; V.80-#3, pp. 465-479.

Schelling, Thomas (1978). *Micromotives and Macrobehavior*. New York: Norton.

Schiffer, Stephen (1972). *Meaning*. Oxford: Oxford University Press.

Simon, Herbert (1982). *Models of Bounded Rationality, Vol. 2: Behavioral Economics and Business Organization*. Cambridge: MIT Press.

Slaughter, Matthew (1998). "Globalisation and Jobs: Will There be a Protectionist Backlash?". World Economy Lecture, University of Nottingham, 17 September 1998.

Smith, Lones and Peter Sørensen (1997). "Informational Herding and Optimal Experimentation". ms: MIT.

Smith, Lones and Peter Sørensen (1998). "Pathological Outcomes of Observational Learning". ms: MIT.

Suppes, Patrick (1977). "A Survey of Contemporary Learning Theories". in Butts and Hintikka, eds. *Foundational Problems in the Social Sciences*. Dordrecht: Reidel, pp. 217-239.

Suppes, Patrick (1989). "Current Directions in Mathematical Learning Theory". in E.E. Roskam, ed. *Mathematical Psychology in Progress*. Berlin: Springer-Verlag, pp. 3-28.

Suppes, Patrick, M. Pavel, and J.-Cl. Falmagne (1994). "Representations and Models in Psychology". *Annual Review of Psychology–1994*, pp. 517-544.

Trueman, Brett (1994). "Analyst Forecasts and Herding Behavior". *Review of Financial Studies*; V.7-#1, pp. 97-124.

van Wijnbergen, Sweder (1992a). "Intertemporal Speculation, Shortages and the Political Economy of Price Reform". *Economic Journal*; V.102-#415, pp. 1395-1406.

van Wijnbergen, Sweder (1992b). "Should Prices Be Decontrolled Gradually or in a Big Bang?". in Alex Cukierman, Zvi Hercowitz, and Leonardo Leiderman, eds. *Political Economy, Growth and Business Cycles*. Cambridge: MIT Press, pp. 262-280.

Vives, Xavier (1996). “Social Learning and Rational Expectations”. *European Economic Review*; V.40-#3/5, pp. 589-601.

Vives, Xavier (1997). “Learning from Others: A Welfare Analysis”. *Games and Economic Behavior*; V.20-#2, pp. 177-200.

Wald, Abraham (1947). *Sequential Analysis*. New York: Wiley.

Wald, Abraham (1950). *Statistical Decision Functions*. New York: John Wiley.

Walliser, Bernard (1998). “A Spectrum of Equilibration Processes in Game Theory”. *Journal of Evolutionary Economics*; V.8-#1, pp. 67-87.

Welch, Ivo (1992). “Sequential Sales, Learning, and Cascades”. *Journal of Finance*; V.47-#2, pp. 695-732.

Zaller, John (1992). *The Nature and Origins of Mass Opinions*. Cambridge: Cambridge University Press.

Zhang, Jianbo (1997). “Strategic Delay and the Onset of Investment Cascades”. *RAND Journal of Economics*; V.28-#1, pp. 188-205.

Zwiebel, Jeffrey (1995). “Corporate Conservatism and Relative Compensation”. *Journal of Political Economy*; V.103-#1, pp. 1-25.