To Specialize or to Innovate?  
An Internalist Account of Pluralistic Ignorance in Economics

November 23, 2013

Rogier De Langhe

Abstract

Academic and corporate research departments alike face a crucial dilemma: to exploit known frameworks or to explore new ones; to specialize or to innovate? Here I show that these two conflicting epistemic desiderata are sufficient to explain pluralistic ignorance and its boom-and-bust-like dynamics, exemplified in the collapse of the efficient markets hypothesis as a modern risk management paradigm in 2007. The internalist nature of this result, together with its robustness, suggests that pluralistic ignorance is an inherent feature rather than a threat to the rationality of epistemic communities.

Keywords: pluralistic ignorance, exploitation, exploration, Thomas Kuhn, scientific revolutions, economics

Acknowledgments

The author thanks the participants of the 4th Copenhagen Lund Workshop in Social Epistemology for comments and discussion. In particular he would like to thank its organizer Frank Zenker for many helpful comments and suggestions during the editorial process.

1 Introduction

Around the turn of the century, in economics “there was a dominant conventional wisdom that markets were always rational and self-equilibrating, that market competition by itself could ensure economic efficiency and stability, and that financial innovation and increased trading activity were therefore axiomatically beneficial.” (Turner 2010) A time of prosperity had begun that came to be called the “Great Moderation” (Bernanke 2004) in which the “central problem of depression-prevention has been solved, for all practical purposes, and has in fact been solved for many decades.” (Lucas 2003) Economics was characterized by a strong collective adherence to the efficient market hypothesis. Only in the summer of 2007 did this situation abruptly
change after two Bear Stearns hedge funds collapsed and officially started the 2007-8 financial crisis. The efficient markets hypothesis underpinned a "modern risk management paradigm [that] held sway for decades. The whole intellectual edifice, however, collapsed in the summer of last year" (Greenspan 2008). However, it was striking that in private remarks, many leading economists had already uttered longstanding disagreement with that consensus. But at the same time they felt powerless to defect from it because they were convinced that all others subscribed to it. For example another Nobel Prize winning economist, Robert Solow, wrote to a friend

"I wholeheartedly agree with the point that economics self-destructs in part because we insist on supposing that everywhere and always individuals maximize purely individualistic preferences subject only to technological, legal, and budget constraints. It is a transparently false assumption, and the brotherhood expends vast ingenuity trying to account for facts within that silly framework." (Solow 1988)

As such the faith in the efficient market hypothesis during the run-up to the financial crisis is an example of pluralistic ignorance.

Pluralistic ignorance is the collective acceptance of a norm that agents privately reject but publicly accept because they believe others accept it. The archetypical example of pluralistic ignorance is Hans-Christian Andersen’s fairy tale “The Emperor’s New Clothes”. Nobody believed that the emperor was wearing clothes, but everyone was afraid to say it and was pressured into pretending he was, because everyone believed everyone else believed that the emperor was wearing clothes. Just as with the collapse of the two hedge funds, a small event can trigger the collapse of the collective belief: a child cries out that the emperor has no clothes. Suddenly there is information available that others also know, diminishing the social pressure to pretend and possibly causing even more such information to become available. Pluralistic ignorance seems to pose a serious challenge to the scientific rationality of both groups and their members because it suggests that social pressure can cause beliefs expressed in groups to systematically deviate from what its members privately believe. Save for certain cases of reflexivity, the truth of \( x \) is independent of the number of individuals that believe \( x \). This suggests that measures should be taken to avoid pluralistic ignorance in scientific communities. In the normative debate about the epistemic status of pluralistic ignorance, this paper argues against such measures by showing that it is possible to account for pluralistic ignorance in scientific communities entirely in terms of factors internal to science. The key argument is that social pressure can be internalist because it increases scientific productivity as a result of specialization. A tension then arises between the intrinsic value of the product and the productivity by which it is achieved and truth might sometimes arise more readily from error than from confusion. By constructing an agent-based model it is shown that this tension between specialization and innovation can explain the boom-and-bust-like features typical of pluralistic ignorance.

2 Pluralistic ignorance in science

Pluralistic ignorance is the collective acceptance of a norm that agents privately reject but publicly accept because they believe others accept it. Thus my focus here lies with norms

\[ \text{See O’Gorman (1986) and Miller and McFarland (1991) for reviews of pluralistic ignorance and Bicchieri (2006) for recent work on the topic.} \]
in science, more specifically with how norms enter science in the classification of the world. This is a conventionalist perspective to classification, following Carnap (1950) and Kuhn (1970). Conventionalism is the view that there is no unique and optimal classification of the world. It became a popular doctrine after the development of non-Euclidian geometries. On a conventionalist perspective, how we “slice up” the world is essentially a social convention. How the world is sliced up matters because this affects what questions are raised and what counts as a solution to those questions. A classificatory framework gives meaning to scientific results and provides the basis for their rational evaluation. As a consequence choice for them cannot simply be governed by truth and falsity because they constitute not the answers but the questions that constitute them. The mere absence of a unique, optimal classification of the world, however, does not necessarily entail that classification is arbitrary or that individuals can change it unilaterally. For example anything can be used as money, e.g. tobacco, cows, shells... But once a framework is in place agents cannot unilaterally change or cancel the norm. So although it may not make sense to compare the intrinsic qualities of the standards, there are important differences with respect to a standard’s pragmatic value that are contingent on the (social) context, such as the fruitfulness of a classificatory framework. This can result in a pattern of collective inertia and collapse, such as the dynamics described and documented in the “Structure of Scientific Revolutions”. Kuhn (1970) On the one hand, scientists are normally reluctant to take up a new paradigm:

“Copernicanism made few converts for almost a century after Copernicus’ death. Newton’s work was not generally accepted, particularly on the Continent, for more than half a century after the Principia appeared. Priestley never accepted the oxygen theory, nor Lord Kelvin the electromagnetic theory, and so on.” (Kuhn 1970, 150)

But when adoption of the paradigm reaches a critical size, a cascade of adoption is triggered:

“At the start a new candidate for paradigm may have few supporters, and on occasions the supporters’ motives may be suspect. Nevertheless, if they are competent, they will improve it, explore its possibilities, and show what it would be like to belong to the community guided by it. And as that goes on, if the paradigm is one destined to win its fight, the number and strength of the persuasive arguments in its favor will increase. More scientists will then be converted, and the exploration of the new paradigm will go on. Gradually the number of experiments, instruments, articles, and books based upon the paradigm will multiply. Still more men, convinced of the new view’s fruitfulness, will adopt the new mode of practicing normal science, until at last only a few elderly hold-outs remain.” (Kuhn 1970, 159)

Philosophers have often resisted classificatory relativism because it makes our classification of the world relative to the circumstances. But classificatory relativism need not lead to sceptical conclusions as long as one has a model of those circumstances in which those circumstances are affected by epistemic considerations. For example the economists William Brock and Steven Durlauf (2002) model the increasing epistemic returns to specialization between scientists, but warn that the resulting path-dependence makes communities run the risk of locking in to a (potentially suboptimal) framework and fail to adjust adequately to novel empirical evidence. The model presented in this paper avoids lock-in by introducing a second factor, innovation, which stands in a dynamic tension with specialization.
In Section 2 I introduce this essential tension. In Section 3 I use the desiderata to explain the characteristic boom-and-bust-like pattern of episodes characterized by pluralistic ignorance and in Section 4 I spell out how they constitute an internalist account of pluralistic ignorance.

3 The essential tension: innovation and specialization

The value of a contribution to a classificatory framework can be broken down into an intrinsic and a social component. The intrinsic value of the contribution denotes its value regardless of the actions of others, and the social value of the contribution denotes the change in value brought about by the actions of others. The former indicates the intrinsic propensity of a rational agent, while the combination of the former and the latter will determine its public behavior. Pluralistic ignorance, then, may occur when there is a discrepancy between an individual’s intrinsic propensity and public behavior.

3.1 Intrinsic value: innovation

Just as many economic goods, the intrinsic value of a contribution can be expected to follow the law of decreasing marginal utility. On this view, the first contribution to a new framework, e.g. Einstein’s 1905 paper introducing the special theory of relativity, will have the highest utility and this utility decreases as more contributions to that classificatory framework are made. In long-established research traditions all the low-hanging fruit is gone and scientists make only very marginal advances (Kuhn’s “normal science”). If contributions have a strictly decreasing marginal utility, the utility of a contribution is always lower than all the contributions that came before. This condition can be satisfied by writing the utility of the next contribution to the classificatory framework $j$ as

$$U_{jt} + 1 = \frac{1}{(T_{jt} + 1)}$$

with $T_{jt}$ the total number of contributions to classificatory framework $j$ at time $t$. If the utility of a contribution to a classificatory framework decreases with every new contribution, there is an incentive for individual agents to be among the first to make a contribution to a new classificatory framework. As such decreasing marginal utility fosters innovation.

3.2 Social value: specialization

It was at the dawn of the Industrial Revolution that Adam Smith visited a pin factory and realized the tremendous benefits of organizing labor. Individual craftsmen had usually produced no more than a couple of pins a day. In the pin factory the production process was divided into eighteen different tasks distributed over ten workers. The factory produced 48,000 pins a day. Smith used the scale effects from specialization to explain this extraordinary leap of productivity. Division and distribution of labor allows workers to specialize in specific skills and stimulates the development of more specialized tools. This insight became the central thesis of his “Wealth of Nations” and marked the birth of modern

\[\text{See Strevens (2003) for a discussion of a model of the distribution of cognitive labor containing only decreasing marginal returns.}\]
economics. For science, Thomas Kuhn has emphasized the benefits of communities adopting the same classificatory scheme. Smith had argued that consensus on a joint standard allowed for the development of specialized tools and specialist training of workers. Using these same arguments for specialization as Smith, Kuhn describes what happens when a community of scientists no longer needs to discuss fundamental issues.

“Freed from the concern with any and all electrical phenomena, the united group of electricians could pursue selected phenomena in far more detail, designing much special equipment for the task and employing it more stubbornly and systematically than electricians had ever done before. Both fact collection and theory articulation became highly directed activities. The effectiveness and efficiency of electrical research increased accordingly…” (Kuhn 1970, 18)

The more scientists agree on a classificatory scheme, the more scientists there are to distribute scientific labor over and hence the greater will be the benefits derived from specialization. As the example of Adam Smith’s pin factory shows, the coordination of workers on the same standard for the division of labor increases their productivity dramatically. This is explained by the cumulative nature of learning effects, where every improvement not only increases productivity directly but also improves the productivity of other improvements. Such cumulative learning effects lead to increasing marginal returns to adoption of a classificatory scheme. In contrast to material effects, information effects have no natural limit. Every improvement can be combined with all existing improvements, creating a quasi-exponential increase of possibilities. This can be expressed by an exponential increase in utility. As such the utility of the next contribution to a classificatory framework $j$ can be written as

$$U_{jt} + 1 = (C_{jt} + 1)^s$$

with $C$ the current number of adopters of classification scheme $j$ and the exponent $s$ a measure of the benefits of specialization. If contributions have strictly increasing marginal

---

4See (Wray 2011, chapter 7) for an account of the crucial role of specialization in Kuhn (1970).

5To take an example, therefore, from a very trifling manufacture; but one in which the division of labour has been very often taken notice of, the trade of the pin-maker; a workman not educated to this business (which the division of labour has rendered a distinct trade), nor acquainted with the use of the machinery employed in it (to the invention of which the same division of labour has probably given occasion), could scarce, perhaps, with his utmost industry, make one pin in a day, and certainly could not make twenty. But in the way in which this business is now carried on, not only the whole work is a peculiar trade, but it is divided into a number of branches, of which the greater part are likewise peculiar trades. One man draws out the wire, another straightens it, a third cuts it, a fourth points it, a fifth grinds it at the top for receiving the head; to make the head requires two or three distinct operations; to put it on, is a peculiar business, to whiten the pins is another; it is even a trade by itself to put them into the paper; and the important business of making a pin is, in this manner, divided into about eighteen distinct operations, which, in some manufactories, are all performed by distinct hands, though in others the same man will sometimes perform two or three of them. I have seen a small manufacture of this kind where ten men only were employed, and where some of them consequently performed two or three distinct operations. But though they were very poor, and therefore but indifferently accommodated with the necessary machinery, they could, when they exerted themselves, make among them about twelve pounds of pins in a day. There are in a pound upwards of four thousand pins of a middling size. Those ten persons, therefore, could make among them upwards of forty-eight thousand pins in a day.” (Smith 2003, 8-9)

6The number of possible connections in such a network of improvements is $n(n-1)$ with $n$ the total number of improvements.

7The more interconnected a community, the more opportunities for specialization. While $C$ and $T$ indicate the current and total size of the framework, I suggest to interpret $s$ as an empirical parameter indicating the degree of interconnectedness of the social community constituting the framework.
utility with other current adopters, the utility of a contribution always increases with the number of others adopting the same paradigm. These benefits of coordination reflect the utility of specialization.

4 Innovation and Specialization

Following the previous discussion, the expected utility of every contribution to a classificatory scheme will consist of the intrinsic value of the contribution given a classificatory scheme and the value of the classificatory scheme itself. The former puts a premium on novelty by rewarding contributions to classificatory schemes to which not many contributions have so far been made. The latter puts a premium on following the tradition by rewarding contributions to the classificatory scheme to which many contributions are currently being made. Here surfaces the conflict between both tendencies. Bringing both these conflicting tendencies to bear on the utility of the next contribution to a classificatory scheme gives the following function:

$$U_{jt} + 1 = \frac{(C_{jt} + 1)^s}{(T_{jt} + 1)}$$

(3)

Innovation and specialization are both epistemic virtues. Specialization depends on the actions of others because division of labor is only possible if others adopt the same paradigm. Innovation, on the other hand, is a function only of the number of previous contributions made to the paradigm. These might have been made by the same individual and do not depend on the current actions of others. This function attributes a utility to all existing classificatory schemes j, but also sets the utility of developing an entirely new classificatory scheme to a fixed number, namely 1. Because for a new classificatory scheme there are not yet any adopters and no contributions have so far been made, both C and T are 0. Hence the utility of creating a novel classificatory scheme is always exactly 1. Since adoption is proportional to utility, and with x the sum of the utility of existing paradigms, the probability for any agent to create a new paradigm is $\frac{1}{x}$ irrespective of the value of s. Because the utility of the other classificatory schemes changes, the relative utility of creating a novel classificatory scheme will self-regulate to make the exploration of new alternatives more probable when returns from the exploitation of existing alternatives have decreased too much. As a consequence the model finds a dynamic balance between innovation and specialization.

5 An agent-based model of specialization and innovation

Consider an agent-based model that is based on the utility assignment described in the previous section. The objective is to explain the characteristic boom-and-bust-like pattern of adopting to a norm as a result of two conflicting epistemic desiderata, innovation (eq. 1) and specialization (eq. 2). Together these desiderata constitute the overall utility function given in eq. 3. In this model, the probability that an agent makes a contribution to a classificatory scheme at the next turn is directly proportional to the utility of that scheme as given in eq. 3. It is now also possible to determine the probability that a new paradigm is created. The

---

8The notions “classificatory scheme” and “paradigm” are used interchangeably for the purpose of this paper.
9This model was coded in Netlogo v.4.1.3. The code of this model can be obtained from the author.
utility of creating a novel classificatory scheme is always exactly 1. Let $x$ be the sum of the utility of all other existing paradigms. With adoption proportional to utility, the probability for any agent to create a new paradigm is then $1/x$ irrespective of the value of $s$. To convey the dynamics of the model I will start with a simple numerical example that describes the development of the utility of a contribution to a single classificatory scheme through time. Consider a model of 1,000 agents where each makes one contribution each turn. Initially the classificatory scheme to which they contribute is randomly distributed. At the outset differences in current adoption are more important than differences in the total number of contributions. Hence one of the classificatory schemes quickly gains market share and takes over the entire system.

Fig. 1 is a simple numerical simulation of various aspects of this classificatory scheme through time. Adoption (a) represents the current number of contributors to a classificatory scheme. The intrinsic value (b) is a function of the total number of contributions to a classificatory scheme and decreases exponentially as more contributions are made. The social value (c) increases exponentially, but for every system consisting of a finite number of agents its size is limited by the number of agents. Their product, the overall utility of a contribution (d), initially increases because the loss in intrinsic value is more than compensated by the gain in social value. However overall utility starts decreasing from the moment social value reaches its maximal level. From this moment the probability of creating a new paradigm starts to increase (e).

Consider now an extension of this numerical example to an infinite number of paradigms. The utility function allows paradigms to be created endogenously, that is, as a result of the very dynamics of the system itself and not as a result of a factor external to the system. Initially each agent is assigned a random paradigm. Figure 2 shows a typical run of this model to illustrate the resulting dynamics. More generally the resulting pattern of adoption can be exhaustively characterized by three phases: the revolutionary, normal and crisis phase.

1. **Revolutionary phase:** As contributions are made to a paradigm, the private/“innovative” value of those contributions decreases, but initially this is more than compensated by the high level of social utility/“specialization” made possible by the adoption of that paradigm by others. This revolutionary phase is characterized by decreasing private utility but increasing social utility and overall utility. Because overall utility increases with adoption, every contribution to that paradigm will make it even more probable that the next contribution will also be made to that paradigm. The fixed utility of creating a new paradigm decreases as a result. Soon this paradigm takes over the entire community. Let us call this the “established” paradigm. However, the benefits of specialization are limited by the size of the community. When a paradigm reaches this limit, the normal phase is triggered.

2. **Normal phase:** Once the entire community makes contributions to the established paradigm, the social utility of a contribution can no longer increase while the private utility keeps on decreasing. This point marks the start of the normal phase of the paradigm: decreasing private utility, constant social utility and decreasing overall utility. During the normal phase the probability of defection increases by a constant proportion as the total number of contributions to the established paradigm increases.

---

In this model classificatory schemes are distinguished using colors. Netlogo allows 1400 colors from which each agent draws a random color at setup.
Figure 1: Simple numerical simulation
Although overall utility is decreasing, it is still large enough that the probability of defection is low. And even when an alternative paradigm emerges, the probability of further defection is low enough that a cascade is improbable. This accounts for the periods of collective inertia Kuhn reported in the face of the emergence of alternative paradigms.

3. **Crisis phase:** During the normal phase the overall utility of the established paradigm decreases because its loss in intrinsic value is no longer offset by an increase in social value. The fixed utility of creating a new paradigm therefore becomes relatively more likely. During the normal phase this possibility of defection increases mostly linearly with the total number of contributions (Fig. 3, \( n = 0 \)). When one agent defects (\( n = 1 \); for example the moment the child shouts “the emperor has no clothes”), the crisis phase is triggered. A defector is an agent who breaks ranks by adopting a paradigm different from the paradigm held by the community during the normal phase. As a result, the probability of defection of other agents exhibits a sudden jump and is now represented as a function of time in Fig. 3, \( n = 1 \) instead of \( n = 0 \). For example at the 1000th step of the model, the probability that an agent would defect to a new paradigm was 0.6281% but once an agent does defects the probability that another agent will also defect to that paradigm rises to 1.2626%. Each new defector will increase the probability of others defecting to that paradigm. The result is a cascade that might accelerates until the defectors’ new paradigm has taken over the entire system. This accounts for the cascade of adoption described by Kuhn in section 2. Each defector increases the probability that others will defect, but decreasingly so. As Table 1 shows, the second defector will only increase the probability of others defecting to that paradigm by 50.3%, the third by 30.6%, etc. Moreover, as illustrated by Fig. 3, the proportional size of these
Table 1: The extent to which the n-th defector increases the probability of more defectors.

<table>
<thead>
<tr>
<th>n</th>
<th>P(n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>+100.40%</td>
</tr>
<tr>
<td>2</td>
<td>+50.30%</td>
</tr>
<tr>
<td>3</td>
<td>+33.60%</td>
</tr>
<tr>
<td>4</td>
<td>+25.25%</td>
</tr>
<tr>
<td>5</td>
<td>+20.24%</td>
</tr>
<tr>
<td>6</td>
<td>+16.90%</td>
</tr>
<tr>
<td>7</td>
<td>+14.52%</td>
</tr>
<tr>
<td>8</td>
<td>+12.73%</td>
</tr>
<tr>
<td>9</td>
<td>+11.34%</td>
</tr>
<tr>
<td>10</td>
<td>+10.22%</td>
</tr>
</tbody>
</table>

Figure 3: Probability defector after n other defectors as a function of total number of contributions to established paradigm.

jumps is independent of the moment at which defection occurs. If the first defection had only occurred at the 10500th step of the model, the probability that the first agent defects to a new paradigm would be 1.0521% and this would increase the probability of a second agent defecting to 2.1084%. But the proportion by which the probability increases with first defection is constant: 100.4%.

6 Pluralistic ignorance

I will now use the model introduced in the previous section to explore the conditions under which and the extent to which the interaction between specialization and innovation generates pluralistic ignorance. Pluralistic ignorance is measured as the difference between a contribution’s intrinsic value and its overall utility. The intrinsic value is the value an agent attributes to a contribution irrespective of the actions of other agents. This can be seen as the agent’s “private” opinion about the value of the contribution. In this model intrinsic value consists in its innovative value and is given by eq. [1]. The overall utility is the value this contribution acquires “in public” (eq. [3]). To represent the difference between intrinsic and

\[ \text{Pluralistic ignorance} = \text{Intrinsic value} - \text{Overall utility} \]

\[ \text{Overall utility} = \sum \text{Actions of other agents} \]

\[ \text{Intrinsic value} = \text{Innovative value} \]

\[ n = 0 \quad \cdot\quad n = 1 \quad \cdot\quad n = 2 \]

For expository purposes I restricted description of this model in the previous section to two paradigms (“established” and “defectors”) but because agents can endogenously create new paradigms there is no limit to the number of paradigms.
overall value in a meaningful way I plot the contributions’ social value as a percentage of their overall utility. Fig. 4 shows the evolution of the total amount of pluralistic ignorance in a typical run of the model with 1,000 agents and $S = 2.7$ (1,000 time steps).

We see a linear increase in the discrepancy between intrinsic value and overall utility as more contributions are made to the dominant paradigm. The larger this discrepancy, the more fragile the paradigm becomes to defectors. As long as the benefits of specialization in the dominant paradigm outweigh the benefits of innovating to another paradigm, defection will not spread. But once specialization no longer outweighs innovation, defection becomes increasingly probable until an actual defection triggers a cascade of further defection. This explains how possibly the rational expectations paradigm could so suddenly have collapsed as reported by Alan Greenspan in the introductory section. More generally I have so far shown how possibly pluralistic ignorance and its typical boom-and-bust pattern can arise as a result of individual agents weighing the conflicting desiderata of innovation and specialization. Now parameter space is explored to establish the robustness of this result.

Fig. 5 reports the cumulated amount of pluralistic ignorance after 1,000 time steps for runs of the model under variation of the population size and the $s$-parameter. It shows that pluralistic ignorance is a robust consequence of a community of agents driven only by innovation and specialization. Pluralistic ignorance tends to rise linearly with population size and exponentially with the benefits of specialization ($s$). This means that, ceteris paribus, scientific communities such as the community of economists would have longer and larger cycles of pluralistic ignorance as they get bigger and more interconnected. This paper argues that from an epistemic point of view this need not be a reason for concern. However, while tunnel vision is beneficial to scientific progress, the financial crisis has shown that there are societal dangers to prolonged periods of tunnel vision (in this case the prolonged and disproportionate adoption of the rational expectations framework) by communities of specialists. A solution could be to artificially separate academic communities into different
**Figure 5:** Sum pluralistic ignorance as a function of population size (X-axis) and $s$ (Z-axis)
ecosystems thereby linearly decreasing the length and size of the period of specialization within such an ecosystem and preserving diversity across ecosystems. This paper shows that communities balancing specialization and innovation will tend to exhibit a natural tendency toward consensus on classificatory frameworks punctuated by periods characterized by the proliferation of alternatives.

7 Conclusion

In this paper I have characterized innovation and specialization in a utility function to model the behavior of a community of agents making contributions to a paradigm. I have shown that the aggregate behavior of a community led by these conflicting epistemic desiderata can possibly explain pluralistic ignorance and the accompanying boom-and-bust-like pattern of revolution, normal science and crisis. Next I have investigated how robust these variables are against changes in their most essential determinants and found that pluralistic ignorance increases with S and the size of the community. The robustness of this result suggests that pluralistic ignorance may be an inherent feature rather than a threat to the rationality of epistemic communities.

References


