The Thesis committee for Daniel Zantedeschi

Certifies that this is the approved version of the following thesis:

Bayesian Learning with Catastrophe Risk: Information Externalities in a Large Economy

APPROVED BY

SUPERVISING COMMITTEE:

Thomas E. Wiseman, Supervisor

Paul Damien

Jan Schneider

Bayesian Learning with Catastrophe Risk: Information Externalities in a Large Economy

by

Daniel Zantedeschi, B.A; M.Eco; M.S.Stat

THESIS

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF ARTS

THE UNIVERSITY OF TEXAS AT AUSTIN

August 2011

Bayesian Learning with Catastrophe Risk: Information Externalities in a Large Economy

Daniel Zantedeschi, M.A.
The University of Texas at Austin, 2011

Supervisor: Thomas E. Wiseman

Based on a previous study by Amador and Weill (2009), I study the diffusion of dispersed private information in a large economy subject to a "catastrophe risk" state. I assume that agents learn from the actions of others through two channels: a public channel, that represents learning from prices, and a bi-dimensional private channel that represents learning from local interactions via information concerning the good state and the catastrophe probability. I show an equilibrium solution based on conditional Bayes rule, which weakens the usual condition of "slow learning" as presented in Amador and Weill and first introduced by Vives (1993). I study asymptotic convergence "to the truth" deriving that "catastrophe risk" can lead to "non-linear" adjustments that could in principle explain fluctuations of price aggregates. I finally discuss robustness issues and potential applications of this work to models of "reaching consensus", "investments under uncertainty", "market efficiency" and "prediction markets".

Table of Contents

Chapter 1

Learning with Catastrophe Risk

1.1 Introduction

Hayek (1945) first argued that markets and organizations can foster the process of aggregation of information that is owned by its participants. Information aggregation occurs through the public observation of variables that reflect other agents' actions (such as prices) or through the private observation of other agents' actions (such as rumors or gossips).

The baseline model is an extension of the continuous time world of Amador and Weill (2009, AW hereafter) and builds on the discrete-time environments of Vives (1993, 1997). I consider a continuum of agents who, at time zero, receive both public and private signals about the state of the world. After time zero, each agent takes an action at every moment until some random time when the state of the world is revealed and her payoff is realized. I assume that an agent's payoff is independent of the actions of any other agent. After receiving initial information, but before the state of the world is revealed, each agent updates her beliefs by observing multi-dimensional noisy signals about the actions of others. The first signal is public: it is the public learning channel and could represents an endogenous aggregate variable, such as a price of some good or some macroeconomic indicator. The second bi-dimensional signal is private, only observed by the agent. The way the private signal is structured provides the main line of departure with respect to the paper of (AW). I will be show that private signals about a catastrophe state leads to information dynamics that are interesting in terms of both informational efficiency and equilibrium characterization.

I solve for an equilibrium in which agents eventually learn the truth. Like in the herding literature, even in presence of the catastrophe risk, there is still a public information externality: better public information may reduce the informational content of both the public and the private channels and slows down learning. Indeed, with an increase in the accuracy of public information, each agent's action becomes more sensitive to the public information that is already known by everyone and at the same time, each agent's action becomes less sensitive to private information. However the learning dynamics is complicated by the presence of a catastrophe states that is responsible, in expectation, for a sort of under-investment problem. It is argued that this may causes a "nonlinear" mixing of opinions between the agents. This makes it harder to glean private information from the noisy observation of agents' actions, and thus generates ambiguous learning externality: information will not necessarily diffuse more slowly through all channels. In fact the probability of a catastrophe state can mitigates this information externality as agents can only infer about the state of world through the private signal.

The rest of the Thesis is organized as follows: Section III surveys the literature within two main areas that are linked in this proposal: social learning, and catastrophes. Section IV describes the model, Section V provides details about the Bayesian Learning , Equilibrium Analysis and Asymptotics, Section V provides a comparative Welfare analysis, Section VI contains discussion about possible applications, and direction for future research, Section VII concludes this work.

1.2 Related Literature

This work is related to the recent literature on the social value of public information. Key references in this areas are Morris and Shin (2002) and Angeletos and Pavan (2007) who provide evidence about the arising of information externalities. Hellwig (2005) first studied the implications for monetary policy. It is common in all this models that agents do not accumulate information over time, they do not suffer from our dynamic learning externality presented in AW. In fact, absent payoff externalities, public information always improve social welfare of the economy.

I divide the survey of the literature in the two main areas: Social Learning and Non-Stationary Economies.

1.2.1 Social Learning

The social learning literature, started by Banerjee (1992) and Bikhchandani et al. (1992), has pointed out the role information externalities in information economics. Their famous result is the possibility of informational cascades and herds: agents may choose to disregard their private information, acting solely on the basis of the public information, and take the "wrong" action with catastrophic consequences. Thus it is implicit in their framework that public information, by facilitating the emergence of herds, can reduce welfare. However, the standard herding models are sequential move games, and the appearance of cascades and herds requires technical conditions rarely met in practice. The model presented in this thesis is an extension of (AW) (2009) similar also to the seminal works of Vives (1993 and 1997), where beliefs are unbounded and actions lie in a continuous space. Vives (1993) showed that when agents learn noisy public information from others, then they learn the truth at a slow speed of $t^{1/3}$ (where t is the number of periods of market interactions). The slow convergence result is due to the informational externality explained above. The more informative the public signal is, as more periods accumulate, the less privately informed agents rely on their private signals, so that less information gets incorporated into the public signal, slowing down convergence. In their recent work, (AW) show that when agents also learn noisy private information from others, then the rate of convergence is t, however in the short term a $\log(t)$ convergence rate determines the behavior of temporary shocks.

This work is also partially related to recent studies in the area of social learning in networks. Important references are Bala and Goyal (1998), Gale and Kariv (2003), and Smith and Sørensen (2005) Banerjee and Fudenberg (2004) and DeMarzo et al. (2003). The private learning channel of the present work, as in (AW) can be thought as reduced form model of local interactions through networks as the setup presented here, although somewhat simpler can be thought as the result of random local interactions.

1.2.2 Non-Stationary Economy and Catastrophes

The main innovation provided by this paper is the inclusion of a catastrophe state. Lately, there has been a revival of the intuition of Rietz (1988) that the simple possibility of rare disasters, such as economic depressions or wars, is a major determinant of equilibrium prices. Indeed, Barro (2006) has shown that, internationally, disasters have been sufficiently frequent and large to make Rietz's proposal viable and account for the high risk premium on traded assets. An alternative line of empirical research has focuses on the so called "Peso Problem" as see Bekaert et al. (1997): investors anticipating the possibility of a catastrophe, which cannot be measured or learned ex ante, discount strongly future returns in a way that is not dynamically consistent with a classical intertemporal equilibrium modes. The work presented here maintains the basic feature of both the Rietz and Bekaert approach.

Perhaps, the only attempt similar to the one presented in this thesis is the one by Moscarini, Ottaviani and Smith (1998). They present a the social learning model, in a discrete state space setting, following the directions of Banerjee et al. (1992) in which individuals take actions sequentially after observing the history of actions taken by the predecessors and an informative private signal. It is shown that if the state of the world is stochastic, only temporary informational cascades, situations where socially valuable information is wasted can arise. Furthermore, no cascade ever arises when the environment changes in a sufficiently unpredictable way. With the simple belief of the presence of a bad state could act as stabilizer in the sense of making the individuals realize about the informational cascade.

1.3 Model Setup

Time is continuous and runs forever. The economy is populated by a $[0, 1]$-continuum of agents whose payoff depend on some unknown and stochastic state of the world $\theta \cdot x \in R^+$, with $\theta = \{0, 1\}$. This means the economy can be in two states depending on the latent variable $\theta = \{0, 1\}$. The state $\theta = 0$ is called the catastrophe state, while $\theta = 1$ the good state. Note that we have two disjoint stochastic processes: the value of the economy in the good state x, and the chance of being in a catastrophe state determined by θ.

At each time before some exponentially[1] distributed "Armageddon time" $\tau > 0$, with parameter λ, each agent takes an action $a_{it} \in R$. At time τ, the state of the world $\theta \cdot x$ is revealed and each agent receives the payoff

$$U_i = -\int_0^\tau (a_{it} - \theta \cdot x)^2 \, dt \tag{1.1}$$

This means that in expectation, at each instant t, by denoting the true prob-

[1] The choice of the exponential distribution implies a constant hazard rate λ, that cannot be inferred by the agents.

ability of $\{0, 1\}$ as $1 - \pi$ and π the first best is to play

$$a_t = E(\theta \cdot x) = \pi\theta \tag{1.2}$$

In a market setting this could be interpreted as hedging or under-investment problem due to the risk of a catastrophe similar to the "Peso Problem" detailed above. This model, albeit of simple specification, contains two important feature:

- The agents do not have time preferences; at each point in time they have to provide the best possible guess. This is driven by the quadratic utility assumption and the fact that the end of world might happen at each instant;
- The agents learn about the state of world at each instant using Bayes rule. Bayes rule in the setup has been shown to be optimal (see Vives, 1993).

1.3.1 Information Structure

Agents are endowed with a diffused common prior over x that is normally distributed with mean zero and zero precision, and a common prior p over the catastrophe state. (Without loss of generality we can assume $p = \frac{1}{2}$). At each point in time they observe:

- a public signal Z_t centered around the average action in the economy.

- a private signal z_{it} concerning the conditional value of x given $\theta = 1$;
- a second private signal concerning the distribution of θ, denoted by ε_{it} as described above.

The initial realizations of these signals are

$$Z_0 = x + \frac{W_0}{\sqrt{P_0}} \tag{1.3}$$

$$z_{i0} = x + \frac{\omega_{i0}}{\sqrt{p_0}} \tag{1.4}$$

$$\pi_{i0}(\theta) = 1/2 \tag{1.5}$$

where W_0, (ω_{i0}) with $i \in [0,1]$ are normally distributed with mean zero and variance one, pairwise independent, and independent of everything else. In equation 1.3, P_0, p_0,S_0 represent the respective precisions of the public and the private signals.

The initial signal Z_0 represents information released by a public agency. The continuum of initial private signals, $(z_{i0}, \pi_{it}(\theta))_{i \in [0,1]}$, make agents asymmetrically informed about $\theta \cdot x$, and represents dispersed information about aggregate economic conditions.

At all times after time zero the public and the private signal evolve according to the stochastic differential equations:

$$dZ_t = A_t dt + \frac{dW_t}{\sqrt{P_\varepsilon}} \tag{1.6}$$

$$dz_{it} = x_t dt + \frac{dw_{it}}{\sqrt{p_\varepsilon}} \tag{1.7}$$

$$d\varepsilon_{it} = \theta dt + dB_{it} \tag{1.8}$$

where W_t, w_{it}, B_{it} are independent realizations or of simple Brownian motions. The private learning channels capture the decentralized gathering of information. One could think, for instance, of local interaction and of private communication, such as gossips.

1.3.2 Filtrations

Given the multiplicity of signals, it is convenient to define more formally the information structure considering the following filtrations:

- $\mathcal{H}_t$: the filtration derived by observing the public signal Z_t up to time t.
- $\mathcal{I}_{it}$: the filtration derived by observing the private signal z_{it} by individual i up to time t.
- $\mathcal{L}_{it}$: the filtration derived by observing the private signal z_{it} by individual i up to time t.
- $\mathcal{M}_{it}$: the joint filtration $\mathcal{I}_{it} \vee \mathcal{L}_{it}$ by combining the information from both the private signals for individual i at time t.
- $\mathcal{G}_{it}$: the joint filtration of $\mathcal{H}_t \vee \mathcal{M}_{it}$ by combining all the information available for individual i at time t.

Each filtration is regular according to the standard conditions presented in Oksendal (1995, Ch. 1).

1.3.3 The Agents' Problem

As agents are infinitesimal, their actions do not affect the average action process A, and hence do not affect the information they receive. By combining this with their preferences in (1.1), the agents' intertemporal problem is essentially static, and together with their quadratic payoffs, it implies that optimal actions are the expectation

$$a_{it} = E\left[x \cdot \theta | G_{it}\right] \tag{1.9}$$

of the random variable x, conditional on their information filtration $\{G_{it}, t \geq 0\}$. Finally in an equilibrium, these individual actions have to generate the average

$$A_t = \int_0^1 a_{it} di \tag{1.10}$$

We summarize all the above in the following

An Equilibrium is a collection of

- a stochastic processes a_i for each agent solving the prediction problem in (1.9),
- a stochastic process A_t determining the average action in (1.10),
- information dynamics given by (1.6).

1.3.4 Timing

The timing of the model can be described as follows:

- Time 0: Nature draws x, θ, and τ.

- Time $t > 0$ with $t < \tau$ Agents receive private and public signals, and play action $a_{it} = E\left[x \cdot \theta | G_{it}\right]$ by predicting $x \cdot \theta$ given all the information they have available. Agents are Bayesian in the way of both predicting the states for each signal, and combining them into a single action.

- Time τ. $x \cdot \theta$ is revealed, payoffs are settled and the game ends.

1.4 Equilibrium: Dynamics, Analysis and Asymptotics

In this section I first detail the continuous time Bayesian learning rules as I tried to present them and then solve for the equilibrium given the notions presented in (AW, 2009).

1.4.1 The Learning problem for the Private Signal z_{it} Public signal Z_t

The dynamics of the private belief about x and public belief on the combined $x \cdot \theta$ can be determined with a direct application of one-dimensional continuous-time Kalman filtering formula (see Oksendal, 1995, pages 85-105).

Take for example the private signal z_{it} which provides information about the value of x. In state-space formulation, the filtering problem for the signal z_{it} is equivalent to the noisy observation of a constant process represented as follows:

$$\begin{aligned}
\text{(system)} \qquad dx_t &= 0 \text{ (x is constant in time)} \\
\text{(observations)} \quad dz_{it} &= x_t dt + \frac{dw_{it}}{\sqrt{p_\varepsilon}}
\end{aligned}$$

Denote with $m = \frac{1}{\sqrt{p_\varepsilon}}$ $p = \frac{1}{\sqrt{p_0}}$. The agent produced a series a predictions based on mean square errors by minimizing $S(t) = E\left(x_t - \hat{x}\right)^2$, and

$$\frac{dS}{dt} = \frac{p^2 m^2}{m^2 + p^2 t},\ t \geq 0$$

This gives the following equation for the prediction of x for the individual i at time t, denoted by $\hat{x}_{it}$

$$d\hat{x}_{it} = -\frac{p^2 m^2}{m^2 + p^2 t}\hat{x}_{it} dt + \frac{pa^2}{m^2 + p^2 t} dz_{it}.$$

In integral form we have:

$$d\left(\hat{x}_{it} \exp\left(\int_0^t \frac{p^2 m^2}{m^2 + p^2 s} ds\right)\right) = \exp\left(\int_0^t \frac{p^2 m^2}{m^2 + p^2 s} ds\right) \frac{p^2}{m^2 + p^2 t} dz_{it}$$

which can be restated as:

$$\hat{x}_{it} = \frac{m^2}{m^2 + p^2 t} x_0 + \frac{p^2}{m^2 + p^2 t} z_{it}$$

so we have a standard conjugate Bayes rule[2] that tell us to average the prediction between the historical average and the new signal z_{it} with weights

[2] Normal Prior and Normal Likelihood gives rise to a Normal distributed Predictive distribution, given that m is known.

proportional to the sum of the precisions. This implies that at instant t individual i, believes that the predictive distribution for z_{it} to be Normal with mean $\hat{x}_{it}$ and precision $m^2 + p^2 t$.

Following the same directions by defining $M = \frac{1}{\sqrt{P_\varepsilon}}$, and $P = \frac{1}{\sqrt{P_0}}$, the prediction for $Z_t = x_t \cdot \theta_t | \mathfrak{I}_t \sim N\left(M, P_t^{-1}\right)$.

1.4.2 Bayesian Learning for the Markov Process $\theta(t)$

Let $\theta = \theta(t)$ be a random variable taking either 0 or 1 value and without loss of generality assume prior information $\pi(0) = 1/2$. The random process ε_t, $t \geq 0$, with

$$
\begin{aligned}
\text{(system)} \quad \theta &= \theta(t) \sim Markov\,(0,1) && (1.11) \\
\text{(observations)}\; d\varepsilon_t &= \theta dt + \sigma dB_t && (1.12) \\
\varepsilon_0 &= 0 && (1.13)
\end{aligned}
$$

is observed. Then the a posteriori probability $\pi(t) = P(\theta = 1|\mathcal{L}_t)$ according to Lipster and Shyrayev (Ch. 9, Theorem 9.6) satisfies the following equation

$$
\begin{aligned}
d\pi(t) &= \pi(t)\,(1 - \pi(t))\,[d\varepsilon_t - \pi(t)dt] && (1.14) \\
\pi(0) &= 1/2. && (1.15)
\end{aligned}
$$

The density of the Radon-Nykodym derivative μ_1, corresponding to the process ε with $\theta = 1$ w.r.t. the measure μ_1, corresponding to the process ε with $\theta = 0$,

$$\phi(t) = \frac{d\mu_1}{d\mu_0}(t, \varepsilon) \tag{1.16}$$

then from Bayes formula it follows that

$$\pi(t) = \phi(t) / (1 + \phi(t)) \ . \tag{1.17}$$

In the case under consideration this "likelihood functional" (Lipster and Shyrayev Theorem 7.7) is defined by $\phi(t) = \exp(\varepsilon_t - t/2)$ which is a geometric Brownian motion, and therefore,

$$\begin{aligned} d\phi(t) &= \phi(t) d\varepsilon_t & (1.18) \\ &= \phi(t)\theta dt + \phi(t)\sigma dB_t & (1.19) \end{aligned}$$

It will be noted that the a posteriori probability $\pi(t)$ (or $\phi(t)$) is a sufficient statistics in the problem of testing two simple hypotheses:

$$H_0 : \theta = 0, \ H_1 : \theta = 1$$

It follows that the learning dynamics of the probability $\pi(t)$ is given by the nonlinear SDE:

$$\pi(t) = \frac{\exp\left(\left(-\frac{\sigma}{2}^2\right) t + \sigma dB_t\right)}{1 + \exp\left(\left(1 - \frac{\sigma^2}{2}\right) t + \sigma dB_t\right)}$$

which does not have closed form solution, but it is easy to verify that for $t = 0$, $\pi(1) = 1/2$ as expected, while for $t \to \infty$ by an application of the central limit theorem (see Shirayev, 1970)

$$\lim_{t \to \infty} \pi_t = \theta^{True}$$

where θ^{True} is the either zero or one depending on Nature draw.

However from theorem 9.1 in Liptser and Shirayaev, the a posteriori probability satisfies

$$\begin{aligned} d\pi\left(t\right) &= \frac{\pi_t\left(1-\pi_t\right)}{\sigma} d\bar{B}_t \\ d\bar{B}_t &= \frac{d\varepsilon_t - \pi_t dt}{\sigma} \end{aligned}$$

where $\bar{B}_t$ is the so-called innovation process of filtering theory, see Lipster and Shyrayev (theorem 7.12). In other words the change in beflies $d\pi_t$ is normally distributed with mean zero and variance $\frac{\pi_t^2(1-\pi_t)^2}{\sigma^2}$. This result about the second moment will prove useful in deriving the equilibrium in the next Section.

1.4.3 Equilibrium Characterization: Closed Form Solution

I now show that there exists an equilibrium in which agent i's action at any time is the conditionally convex combination of two forecasts of weighted states of the world: a public forecast shared with everyone in the economy, which we denote by $\hat{X}_t$, and a private forecast, denoted by $\hat{x}_{it}$ defined earlier, containing all the information observed by agent i and no one else. As mentioned earlier, Bayesian updating implies that the action taken by agent i at time t is, then,

$$a_{it} = E\left[x \cdot \theta | G_{it}\right]$$

The public forecast, is, of course the same for every agent. On the other hand the agents have to combine the two private signals. Here we adopt a Normal approximation in order to be able to deal with signals with the same distributional properties.

1.4.3.1 Combining Private Signals with a First Order Approximation

The first two moments are easily characterized:

$$
\begin{aligned}
E\left[x\cdot\theta|\mathcal{M}_{it}\right] &= \pi_{it}\hat{x}_{it} \\
V\left[x\cdot\theta|\mathcal{M}_{it}\right] &= E\left[x^2+\theta^2\right]-\pi_{it}^2\hat{x}_{it}^2 \\
&= \left(\frac{1}{p_{it}}+\hat{x}_{it}^2\right)\left(\frac{\pi_{it}^2\left(1-\pi_{it}\right)^2}{\sigma^2}+\pi_{it}^2\left(\theta\right)\right)-\pi_{it}^2\hat{x}_{it}^2 \\
&= \left(\frac{1+p_{it}\hat{x}_{it}^2}{p_{it}}\right)\left(\frac{\left(\sigma+1\right)\pi_{it}^2+\pi_{it}^4-2\pi_{it}^3}{\sigma}\right)-\pi_{it}^2\hat{x}_{it}^2 \\
&= \frac{\left(\sigma+1\right)\pi_{it}^2+\pi_{it}^4-2\pi_{it}^3+\left(\sigma+1\right)\pi_{it}^2p_{it}\hat{x}_{it}^2+p_{it}\hat{x}_{it}^2\left(\pi_{it}^4-2\pi_{it}^3\right)-p_{it}\sigma\pi_{it}^2\hat{x}_{it}^2}{p_{it}\sigma} \\
&= \frac{\pi_{it}^2\left(\sigma+1+\sigma p_{it}\hat{x}_{it}^2\right)}{p_{it}\sigma}+O\left(\pi_{it}^3\right) \\
&= \frac{\left(\sigma+1\right)}{p_{it}\sigma}\pi_{it}^2+O\left(\hat{x}_{it}^2\right)+O\left(\pi_{it}^3\right) \\
&= 1/p_{it}^{\pi}+O\left(\hat{x}_{it}^2+\pi_{it}^2\right)
\end{aligned}
$$

Also, I am dropping every squared and interaction term. The economic reasoning behind this approximation can be summarized in the following points:

- **Model Consistency**: Agents have quadratic utility, which implies just the first two moments of signals are entering the utility computation. At the same time a Normal distribution is fully characterized by the first two moments which in turns is consistent with both the preferences and the filtering rules presented earlier. Including higher order terms, such as skewness and kurtosis components, albeit important from the empirical point of view, would contradict the internal structure of the model presented here.

- **No "Scale Effects"**: in the model here scale effects are not supposed to play a role. This means that the squared level $\hat{x}_{it}^2$ should not be important; as a partial justification it is always possible to normalize x to be in the neighborhood of 1, in this case x can be thought as a log-growth rate.

- This procedure of linearization is consistent with modern macroeconomics and asset pricing theory and practice where models are usually "log-linearized" and econometric techniques such as the Delta Method and Lagrange Approximation for the Posterior distribution.

Following this derivation, the individual predicts $z_t = (x \cdot \theta | \mathcal{M}_t) = N\left(\pi_{it}(\theta)\, \hat{x}_{it}, \frac{1}{p_{it}^{\pi}}\right)$.

1.4.3.2 Combining the Derived Private Signal with the Public One: Determination of the Equilibrium

We know that the forecasts are normally distributed, independent given $x \cdot \theta$ and that the public forecast is common knowledge. Following standard Bayesian Decision Theory (Berger, 1985 Ch.2), the action taken by individual

i at time t, is given by

$$a_{it} = E\left[x \cdot \theta | G_{it}\right] = \frac{P_t}{P_t + p_{it}^{\pi}} \hat{X}_t + \frac{p_{it}^{\pi}}{P_t + p_{it}^{\pi}} \pi_{it}\left(\theta\right) \hat{x}_{it}$$

which a precision weighted combination of the public and the private forecasts. Note that the private forecasts are unbiased and based on independent private evaluations: thus their cross section must be equal to the true $x \cdot \theta$. This means that on average

$$A_t = \int_0^1 a_{it} = \frac{P_t}{P_t + p_t^{\pi}} \hat{X}_t + \frac{p_t^{\pi}}{P_t + p_t^{\pi}} \pi_t\left(\theta\right) \hat{x}_t.$$

where $\hat{x}_t = \int_0^1 \hat{x}_{it} dt$ represent the average prediction of the agents as well as $\pi_t\left(\theta\right)$. This means that if we are interested in determining the global behavior agents across time, we can aggregate by averaging, and considering the predictions from the "mean" agent the i subscripts. This is the reason why I drop all the $i-$subscripts.

1.4.4 Qualitative Analysis of the Equilibrium

Before entering into the asymptotic considerations in the following section, I here provide some qualitative comparisons with respect to the results obtained in (AW).

The presence of the catastrophe states scales down the information percolated through the public channel by decreasing the precision of the public signal. This means that agents will pay proportionally more attention to their private forecast, and it is easy to anticipate that, at least in presence of short term shock, the "slow learning" phenomenon described in Vives will be relatively weaker. Also the higher the variance σ, the lower the precision that can be derived from the market. Furthermore the rescaled public and private are informationally equivalent.

Following considerations in (AW), the precisions of the adjusted public and private forecasts, P_t and p_t, are readily characterized by a coupled system of stochastic differential equations

$$
\begin{aligned}
dP_t &= P_\varepsilon \left(\frac{P_t}{P_t + p_t^\pi} \right)^2 dt && (1.20)\\
dp_t^\pi &= p_\varepsilon \left(\frac{p_t^\pi}{P_t + p_t^\pi} \right)^2 dt && (1.21)\\
\pi_t(\theta) &= \phi(t) / (1 + \phi(t)) && (1.22)\\
d\phi(t) &= \phi(t)\theta dt + \phi(t)\sigma dB_t && (1.23)
\end{aligned}
$$

Where P_ε and p_ε are constants induced by the common knowledge assumption and the initialization. The stochasticity is induced by the presence of the θ process which determines p_t^π.

1.4.5 Learning Asymptotics and Derivation of Slow learning Properties

The literature on "slow learning" is concerned with the behavior of the private signals. In particular, as explained in Section II, it is argued that due to information externalities the agents tend to pay less attention to the private signal than the public one.

Following the system in (1.20), we have a stochasticity induced by the presence of the θ process which determines p_t^π. Note that the the first ODE can be transformed into the second by pre-multiplying by the ratio of the constants. Hence it is possible to obtain:

$$\dot{p}_t^\pi = p_\varepsilon \left(\frac{p_t^\pi}{P_t + p_t^\pi} \right)^2$$

This result suggests the form for the following Claim.

beginclaim (Precision Asymptotics). (i) The precision of the private forecast monotonically converges to infinity as long as $p_\varepsilon > 0$ (ii) the ratio $\frac{p_t}{\frac{(\sigma+1)}{\sigma}\pi_{it}^2 P_t + p_t}$ monotonically converges to $p_\varepsilon / \left(p_\varepsilon + \frac{(\sigma+1)}{\sigma}\pi^2 P_\varepsilon \right)$, and (iii) as $t \to \infty$ the total precision, grows with rate

$$\frac{\pi_{it}^2 (\sigma+1)}{\sigma} P_t + p_t \quad \propto \quad \left(\frac{p_\varepsilon}{\frac{(\sigma+1)}{\sigma}\pi_{it}^2 P_\varepsilon + p_\varepsilon} \right)^2 \left(\frac{(\sigma+1)}{\sigma}\pi_{it}^2 P_\varepsilon + p_\varepsilon \right) t + \tag{1.24}$$

$$+2\log(t) \frac{(\sigma+1)}{\sigma}\pi_{it}^2 \frac{P_\varepsilon}{p_\varepsilon} \left(p_0 - \frac{1}{4} P_0 \right) \tag{1.25}$$

Proof. Consider Appendix A in (AW) with the modified system of ODEs. □

This means that the precision of the beliefs is growing at a linear rate when $p_\varepsilon > 0$, even though no new information is being exogenously provided to the agents: that is the social learning generated from endogenous private signals, no matter how noisy, and no matter how bad the catastrophe state is perceived, is sufficient to restore the speed of convergence to the linear rate.

From the analysis above the growth rate of the sum of the precisions can be broken in two parts

$$
\begin{array}{ll}
(A) & \left(\frac{p_{\varepsilon}}{\frac{\pi_{it}^{2}(\sigma+1)}{\sigma}P_{\varepsilon}+p_{\varepsilon}}\right)^{2}\left(\frac{(\sigma+1)}{\sigma}\pi_{it}^{2}P_{\varepsilon}+p_{\varepsilon}\right)t \\
(B) & \quad 2\log(t)\,\frac{\pi_{it}^{2}(\sigma+1)}{\sigma}\frac{P_{\varepsilon}}{p_{\varepsilon}}\left(p_{0}-\frac{1}{4}P_{0}\right)
\end{array}
\tag{1.26}
$$

where (A) grows linearly and so is responsible for the long-term behavior of the learning process while (B) dominates the short term behavior.

As (A) dominates in the long term, it also determines the "slow learning property" of agents ending up paying too much attention to the private signal In contrast (B) determines the response to precision shocks in the short term.

Also in (A), any temporary shocks not affecting directly the $p_{\varepsilon}/\left(p_{\varepsilon}+\pi^{2}P_{\varepsilon}\right)$ ratio will not change the long term behavior. This means that volatility shocks such as changes in σ or in the initial conditions p_0, P_0 will only have a temporary effect due to the (B), slowing down the learning in a way that is actually diverse with respect to the (AW) case. In fact in their paper the (B) part is primarily led by a $\log(t)$ term while in our case we have $2\log(t)\,\frac{\pi_{it}^{2}(\sigma+1)}{\sigma}$.

1.5 Comparative Welfare Analysis of the Catastrophe Risk

In this last section I draw a comparison with the Welfare analysis derived in (AW). While derivations could follow the same directions of their paper, I concentrate on explaining the differences with their findings. Let the welfare criterion by the equally weighted sum of agents' expected utility. By the law of large numbers, this criterion coincides with the ex-ante utility of a

Bayesian agent,

$$W = -\lambda E\left[\int_0^{\tau} (a_{it} - x)^2 \, dt\right] = -\int_0^{\infty} \frac{e^{-\lambda t}}{P_t + p_t^{\pi}} dt$$

where the welfare intensity is normalized and the result is obtained by using the independence between the random exponential time and the fact that $E\left[(a_{it} - x)^2 | \mathcal{G}_{it}\right] = 1/\left(P_t + p_t^{\pi}\right)$.

Hence, I have shown that the aim of a representative agent, which could be thought as the market organizer, is to maximize at each point in time the total precision of the average agent. Thus, this is connected with the characterization of the asymptotic properties derived in the previous Section. (AW) claim for all initialization p_0, P_0, there is a time time $\nu > 0$, and $\nu < \tau$, such that for $t < \nu$, $\frac{\partial W}{\partial P_t} < 0$. This means a positive Public Precision shock can actually decrease welfare in the short term. This notion is weakened in the setup presented here as P_t is discounted by $\frac{\pi_{it}^2(\sigma+1)}{\sigma}$, so if the time is short the $\frac{\partial W}{\partial P_t}$ depends on the stochasticity of π_t that could in principle change the behavior of the derivative. An increase of public precision acts non-linearly on the aggregate action as I have pointed out earlier while discussing the equilibrium rule.

(AW) analyze the dynamic learning externality by studying the optimal social diffusion of information. A planner chooses an adapted action a_i in order to maximize the ex-ante utility of a randomly chosen agent, subject to learning. In setting up this problem, I follow Vives (1997) and consider actions that are convex combinations of the public and the private forecasts:

$$a_{it} = (1 - \gamma_t)\,\hat{X}_t + \gamma_t \pi_{it}\,(\theta)\,\hat{x}_{it}$$

with $0 < \gamma < 1$. It is evident that if I set $\gamma = \frac{p_t^\pi}{P_t + p_t^\pi}$, I am back to the original Bayes Rule forecasts combination. Using dynamic programming (AW) find that $\gamma_t^* > \frac{p_t^\pi}{P_t + p_t^\pi}$ and γ. This means that as expected, slow learning is not optimal, and agents should pay more attention to their private forecast with respect of what is implied by Bayes Rule. The main difference with respect to their result is that under catastrophe risk, the social planner has different ways to implement the rule:

- Telling agents the optimal weight directly γ_t^* which would determine the end of a competitive market as agents act "optimally" by Bayes rule.
- decreasing the accuracy of the public signal. so to cause some market shock. This cannot be thought as a practical option as the public signal is common knowledge, so the agent would realize the manipulation instantaneously leading to higher order expectation effects that are not considered in this model.
- spreading rumors about catastrophe probability so to cause shocks on π_t; This is actually more effective at t grows and also practical as it might sufficient to spread some rumors to a portion of the agents via the private channel, to modify the average π_t.

The last two points are usually considered to be bad features of markets, while in fact they can improve the information processing of the agents.

1.6 Discussion: Economic Predictions and Themes for Future Research

I conclude this work by suggesting some potential applications as well as economic predictions and motivations for further research. In fact even if the model presented here is quite simple, it can be adapted to work in different area of economics and finance:

- **Reaching Consensus**: The problem of reaching consensus starting from disparate expectations has been at the center of economics of information since the seminal paper of Aumann (1976). It has been shown that repeated public announcements of a stochastically monotone aggregate statistics of conditional expectations, which need not to be common knowledge, leads to consensus (McKelvey and Page (1986) and Nielsen et al (1990)). In the iterative process, individuals compute conditional expectations with the information they have available and the aggregate statistics is announced. Individuals then compute their expectations on the basis of their private information plus the new public information, and the process continues. In the basic model of (AW) repeated public announcement of a linear noisy functions of agents conditional expectations leads to consensus but slowly, due to the "slow learning" property first studies by Vives. We could argue that rephrasing Geanakoplos and

Polemarchakis (1982), that in presence of noisy public information "we cannot agree forever but can disagree for a long time". The results presented in the "slow learning" part show that in periods of time when people think there is an high chance of a catastrophe state, convergence to the truth can be made faster. Also the section on the Welfare Analysis tells us that one way to break "ties" or periods of disagreement is to start spreading rumors about catastrophes between the agents.

- **Macroeconomic Forecasting and Investments**: It is possible to consider competitive firms, producing forecasts, in presence of some macroeconomic uncertainty. The decisions could be either financial (acquire a new company, extend a bank loan, issuing new securities etc.) or operational (acquiring new technology, investing in research etc.). Macroeconomic uncertainty can be summarized in the parameter x, however there is a chance $\pi(\theta)$ that the economy will collapse. Firms invest taking into account that the profits of their accumulated investment depend on the realization of $\theta \cdot x$. Thus the investment decision is directly linked to the prediction of $\theta \cdot x$. To predict $\theta \cdot x$ each firm has access to private signals, produced by a privately hired macroeconomists, as well as public information, aggregate past investment figures compiled by a government agency. Data on aggregate corporate investment includes measurement error. Consequently at each period the noise aggregate is made public. The research question in this setup, that can be adapted very easily to the one presented in the paper, is whether repeated announcement of the aggregate investment figures reveal the state of world, and if so how fast. The take from this study is that reducing the measurement error, for instance by adopting more strict accounting standards, may not necessarily need to an improvement in the investment decisions at least in the short run.

- **The Design and Efficiency of "Prediction Markets":** These are speculative markets created for the purpose of making predictions, meaning there is not necessarily an underlying tradable asset The current market prices can then be interpreted as predictions of the probability of the event or the expected value of the parameter $\theta \cdot x$, which is normalized between 0 and 1. Some academic research has focused on potential flaws with the prediction market concept. In particular, Manski (2005) first attempted to show mathematically that under a wide range of assumptions the "predictions" of such markets do not closely correspond to the actual probability beliefs of the market participants unless the market probability is near its boundaries. Manski (2005) suggests that directly asking a group of participants to estimate probabilities may lead to better results. However, Gjerstad (2005) shows that prediction market prices are very close to the mean belief of market participants if the agents are risk averse and the distribution of beliefs is spread out (as with a normal distribution, for example). At the empirical level Wolfers and Zitzewitz (2006) have obtained similar results, and also include some analysis of prediction market data. The outcome from my research can be reconciled with this framework. In fact as the price is based in the π-*space*, using the logit transformation presented in (1.17) can be recast in the $x \cdot \theta$ state of the private signals. The structure of the updating can be thought as the same as well as some economic predictions that should not be affected by a logit transformation which is monotone. With respect the "slow learning" and "catastrophe risk" state could play a role in determining the efficiency of information aggregation in prediction markets. In light of the results in this research, I tend to agree with Manski. A public signal of 1/2 implies perfect uncertainty and in this could be a result of both slow learning and catastrophe shocks. This certainly requires further investigation.

- **Market Efficiency and Asset Pricing Methods**: the cornerstone of modern financial economics is market efficiency which in one of its forms prescribes that all information should be accounted into prices. The model presented here seems to be consistent with the definition of market efficiency but it also introduces features of "convergence to the truth" that can explain other behavioral theories assumed to invalidate it, such as informational cascades. However modern asset pricing, rarely allow for learning and rarely offers non-representative agents theories. It would be certainly interesting to embed this work into a more general market microstructural model.

- **The Role of "Dr. Doom" in Markets**: Nouriel Roubini is a famous economist that earns his nickname as "Dr. Doom" for repeated and controversial claims about incoming catastrophes. As an example in 2008, Fortune magazine wrote, "In 2005 Roubini said home prices were riding a speculative wave that would soon sink the economy. Back then the professor was called a Cassandra. Now he's a sage". From my point of view, "Dr. Doom" could actually has a positive role as he is acting as it was suggested in the Welfare Analysis Section. Repeated catastrophe shocks can improve information processing of agents avoiding slow learning which could determine herding behavior and market crashes. Thus by calling for a crash he is actually working in the other direction. From a research perspective it would be interested to extend the framework presented here to the case of "public" catastrophe rumors. I believe the structure should not change significantly.

1.7 Conclusion

I have presented a model of learning from others' actions with the possibility of a catastrophe state. My work extends the analytically tractable framework of Amador and Weill (2009) by including sources of uncertainty that can possible restore informational efficiency. In particular I have shown that catastrophe risk shocks could act as natural economic stabilizer by having individuals paying more attention to their private information. In such a way, excessive reliance on public source, and appearance of informational cascades and herding could be avoided. I have finally discussed potential applications and extensions for future research.

Bibliography

[1] Amador, M., and Weill, P. (2009). Learning from Private and Public Observations of Others' Actions, Working Paper, Stanford University.

[2] Angeletos, G., and Pavan, A. (2007). Efficient Use of Information and Social Value of Information," Econometrica, 75, 1103–1142.

[3] Araujo, L., and Shevchenko, A. (2006). Price Dispersion, Information and Learning, Journal of Monetary Economics, 53, 1197–1223.

[4] Aumann, R. (1976). Agreeing to Disagree, The Annals of Statistics, 1:1236-1239.

[5] Bala, V., and Goyal, S. (1998). Learning from Neighbours, Review of Economic Studies, 65, 595–621.

[6] Banerjee, A. and Fudenberg, D. (2004). Word-of-mouth learning, Game and Economic Behavior, 46, 1–22.

[7] Banerjee, A.V. (1992). A Simple Model of Herd Behavior, The Quarterly Journal of Economics, 107 (3), 797–817.

[8] Barro, R. (2006). Rare Disasters and Asset Markets in the Twentieth Century, Quarterly Journal of Economics, 121, 823-866.

[9] Barro, R., and Ursua, J. (2008). Macroeconomic Crises Since 1870,Brookings Papers on Economic Activity (2008), 255-335.

[10] Bergemann, Dirk and Jusso Valimaki, “Experimentation in Markets,” Review of Economic Studies, 2000, 67, 213–234.

[11] Bekaert, G., Hodrick, R.J., and Marshall, David A. (1997). ”Peso problem” explanations for term structure anomalies,” Working Paper Series, Issues in Financial Regulation WP-97-07, Federal Reserve Bank of Chicago.

[12] Bikhchandani, S., Hirshleifer, D. and, Welch, I. (1992). A Theory of Fads, Fashion, Custom, and Cultural Change in Informational Cascades, Journal of Political Economy, 100 (5), 992–1026.

[13] Berger, J.O. (1985). Statistical decision theory and Bayesian Analysis (2nd ed.). New York: Springer-Verlag.

[14] Blouin, M. R., and Serrano, R. (2001). A Decentralized Market with Common Values Uncertainty: Non-Steady States, Review of Economic Studies, 68, 323–346.

[15] Bolton, P., and Harris C. (1999). Strategic Experimentation, Econometrica, 67, 349–374.

[16] Bru, L., and, Vives, X. (2002). Informational externalities, herding and incentives. Journal of Institutional and Theoretical Economics 158:91-105.

[17] Burguet, R., and Vives, X. (2000). Social learning and costly information acquisition, Economic Theory, 15 (1), 185–205.

[18] Chamley, C. (2004). Rational Herds, Cambridge: Cambridge University Press.

[19] Decamp, JP, and Mariotti, T. (2004). Investment timing and learning externality, Journal of Economic Theory, 18, 80–102.

[20] DeMarzo, P., Vayanos, D., and Zwiebel, J.H. (2003). Persuasion Bias, Social Influence, and Uni-Dimensional Opinions, Quaterly Journal of Economics, 118, 909–968.

[21] Duffie, D., and Manso, G. (2007). Information Diffusion in Large Population, American Economic Review, Papers and Proceedings.

[22] Gale, D., and Shachar, K. (2003). Bayesian Learning in Social Networks, 2003. Working Paper, New York University.

[23] Geanakoplos, J. and Polemarchakis, H. (1982). We can't disagree forever, Journal of Economic Theory 28:192-200.

[24] S. Gjerstad (2005). Risk Aversion, Beliefs, and Prediction Market Equilibrium. Technical report, Economic Science Laboratory, University of Arizona.

[25] Golosov, M.,Tsyvinski, A., and Lorenzoni, G. (2007). Decenralized Trading with Private Information, Working Paper, MIT and Yale.

[26] Green, E.J., (1991). Eliciting Traders' Knowledge in "Frictionless" Asset Markets, Working Paper.

[27] Hayek, F. (1945). The Use of Knowledge in Society, American Economic Review, 35: 519-530.

[28] Hellwig, C., (2005). Heterogeneous Information and the Welfare Effects of Public Information Disclosures.

[29] Jovanovic, B. and Nyarko, Y. (1995). A Bayesian Learning Model Fitted to a Variety of Empirical Learning Curves, Brooking Papers: Microeconomics, pp. 247–304.

[30] Katzman, B., Kennan, J., and Wallace, N. (2003). Output and Price Level Effects of Monetary Uncertainty in a Matching Model, Journal of Economic Theory, 108, 217–255.

[31] Keller, G., Rady, S. and, Cripps, M. (2005). Strategic Experimentation with exponential bandits, 73, 39–68.

[32] Kircher, P., and Postlewaite, A. (2008). Strategic Firms and Endogenous Consumer Emulation, Quarterly Journal of Economics, 123, 621–661.

[33] Liptser, R.S., and Shiryiayev, A.N. (1977). Statistics of Random Processes I, General Theory, Springer Verlag, New York.

[34] Manski, C.F (2005). Interpreting the Predictions of Prediction Markets, NBER Working Paper No. 10359.

[35] McKelvey, R.D., and Page, T. (1986). Common Knoledge, Consensus and Aggregate Information, Econometrica 54:109-27.

[36] Morris, S., and Shin, H.S. (2002). The Social Value of Public Information, American Economic Review, 92, 1521–1534. 4, 17.

[37] Moscarini, G., Ottaviani, S. and Smith L. (1998). Social Learning in a Changing World, March 1998, 657-665.

[38] Nielsen, L., Branderburger, A., Geanakoplos, J., McKelvey, R., and Page, T., (1990). Common Knowledge of an Aggregate of Expectations. Econometrica 58:1235-39.

[39] Oksendal, B., (1985). Stochastic Differential Equations, Berlin: Springer.

[40] Rietz, T., (1998). The Equity Risk Premium: A Solution,Journal of Monetary Economics, 22, 117-131.

[41] Rosenberg, D., Solan, E., and Vieille N. (2007). Social learning in one-arm bandit problems, Econometrica, 75, 1591–1611.

[42] Smith, L. and Sørensen, P.N. (2005). Rational Social Learning by Random Sampling, Working Paper, University of Michigan.

[43] Shiryaiev, A.N., (1970). Optimal Stopping Rules, Springer

[44] Vives, X., (1993). How Fast do Rational Agents Learn?, Review of Economic Studies, 60, 329–347.

[45] Vives, X. (1997). Learning from others: a Welfare Analysis, Games and Economic Behavior, 20, 177–200.

[46] Vives, X. (2010). Information and Learning in Markets: the Impact of Market Microstructure, Princeton Academic Press.

[47] Wallace, N., (1997). Short-Run and Long-Run Effects of Changes in Money in a Random-Matching Model, Journal of Political Economy, 105, 1293–1307.

[48] Wolfers J., and Zitzewitz E. (2007). Interpreting Prediction Market Prices as Probabilities, NBER Working Paper No. 12200.

[49] Wolinsky, A. (1990). Information Revelation in a Market with Pairwise Meetings, Econometrica, 58, 1–23.