

Diagnostic Expectations and Credit Cycles

Pedro Bordalo, Nicola Gennaioli, Andrei Shleifer¹

November 2015

Abstract

We present a model of credit cycles arising from diagnostic expectations – a belief formation mechanism based on Kahneman and Tversky’s (1972) representativeness heuristic. In this formulation, when revising their beliefs agents overweight future outcomes that have become more likely in light of incoming data. Diagnostic expectations are forward looking, and as such are immune to the Lucas critique and nest rational expectations as a special case. Diagnostic expectations exhibit excess volatility, over-reaction to news, and systematic reversals. These dynamics can account for several features of credit cycles and macroeconomic volatility.

¹ Royal Holloway, Universita Bocconi and IGER, and Harvard University, respectively. Gennaioli thanks the European Research Council and Shleifer thanks the FHB initiative at Harvard University for financial support of this research. We are also grateful to Nicholas Barberis, Sam Hanson, Arvind Krishnamurthy, Yueran Ma, Matteo Maggiori, Sendhil Mullainathan, Josh Schwartzstein, Alp Simsek, Jeremy Stein, Amir Sufi, and Luigi Zingales for helpful comments.

1. Introduction

The financial crisis of 2008-2009 revived economists' and policymakers' interest in the relationship between credit expansion and subsequent financial and economic busts. According to an old argument (e.g., Minsky 1977), investor optimism brings about the expansion of credit and investment, and leads to a crisis when disappointing news arrive. Stein (2014) echoes this view by arguing that policy-makers should be mindful of credit market frothiness and consider countering it through policy.

Recent empirical research provides considerable support for this perspective. Schularick and Taylor (2012) demonstrate, using a sample of 14 developed countries between 1870 and 2008, that rapid credit expansions forecast declines in real activity. Jorda, Schularick, and Taylor (2013) further find that more credit-intensive expansions are followed by deeper recessions. Mian, Sufi, and Verner (2015) show that growth of household debt predicts future economic slowdowns. And Baron and Xiong (2014) establish in a sample of 20 developed countries that bank credit expansion predicts increased crash risk in both bank stocks and equity markets more broadly.

Parallel findings emerge from the examination of credit spreads – differences in yields between safe and risky debt. The benchmark recent study by Greenwood and Hanson (2013) shows that credit quality of corporate debt issuers deteriorates during credit booms, and that low credit spreads forecast low, and even negative, excess corporate bond returns. In addition, during credit expansions the share of credit going to risky firms rises, and this risky share, rather than credit growth per se, predicts poor economic growth. Gilchrist and Zakrajsek (2012) and Krishnamurthy and Muir (2015) relatedly establish that eventual credit tightening correctly anticipates the coming recession. Lopez-Salido, Stein, and Zakrajsek (hereafter LSZ 2015) pull a lot of this

evidence together, and show that low credit spreads predict both a rise in credit spreads and low economic growth afterwards. They stress predictable mean reversion in credit market conditions². Both Greenwood and Hanson (2013) and LSZ (2015) interpret the evidence as inconsistent with rational expectations.

The prevailing approach to understanding the link between financial markets and the real economy is financial frictions, which focus on the transmission of an adverse shock through a leveraged economy (Bernanke and Gertler 1989, Kiyotaki and Moore 1997, Lorenzoni 2008, Brunnermeier and Sannikov 2014). In some instances, financial frictions are supplemented by Keynesian elements, such as the zero lower bound on interest rates or aggregate demand effects (Eggertson and Krugman 2012, Farhi and Werning 2015, Guerrieri and Lorenzoni 2015, Korinek and Simsek 2014, Rognlie et al. 2015). The adverse shock in these models is either a drop in fundamentals, or a “financial shock” consisting of the tightening of collateral constraints or an increase in required returns. These models typically do not explain why a financial boom is systematically followed by a bust. In particular, there is no attempt to explain what causes financial conditions to deteriorate suddenly.

Perhaps more fundamentally, because they rely on the assumption of rational expectations, these models do not explain predictable negative or low abnormal returns on debt in over-heated markets (Greenwood and Hanson 2013, LSZ 2015). Nor do they come to grips with the survey evidence on investor and manager expectations that points to predictable expectations errors linked to extrapolation of the past (Greenwood and Shleifer 2014, Gennaioli, Ma and Shleifer 2015). To account for the evidence more completely, one may need to abandon rational expectations.

² An older literature on financial asset prices and economic activity includes Bernanke (1990), Friedman and Kuttner (1992), and Stock and Watson (2003), among others.

In this paper, we follow this path. We propose a psychological model of investor confidence and credit cycles that accounts for much of the evidence described above, and articulates in a fully dynamic setup the phenomenon of credit market overheating. It implies that in a boom investors are overly optimistic and will systematically become more pessimistic in the future, leading to crises even without deteriorating fundamentals. The model allows us to discuss in a unified framework such phenomena as adaptive expectations and extrapolation (e.g., Cagan 1956, Greenwood and Shleifer 2014, Barberis et al. 2015a, b), as well as neglect of risk (Gennaioli, Shleifer, and Vishny 2012). Critically, households in our model are forward looking, and recognize policy shifts. As a consequence, the model is not vulnerable to the Lucas critique, which has plagued an earlier generation of behavioral models. Indeed, for any data generating process, rational expectations emerge a special case of our model.

Our principal contribution is to write down a psychologically-founded model of beliefs and their evolution in light of new data.³ Importantly, the model we propose is taken from a very different context and adapted to macroeconomic problems, rather than just designed to match credit cycle facts. It is portable in the sense of Rabin (2013). Our model of belief evolution is based on Gennaioli and Shleifer's (2010) formalization of Kahneman and Tversky's (KT 1972, 1983) representativeness heuristic describing how people judge probabilities. According to KT, people estimate types with a given attribute to be more common in a population than they really are when that attribute is representative of diagnostic for these types, meaning that it occurs more frequently among these types than in the relevant reference class. For instance, beliefs about the Irish exaggerate the share of red haired people among them because red hair

³ Many models of beliefs in finance are motivated by psychological evidence, but use specifications specialized to financial markets (e.g., Barberis, Shleifer, and Vishny 1998, Fuster, Laibson, and Mendel 2010, Hirshleifer et al 2015, Greenwood and Hanson 2015, Barberis et al. 2015a,b). Fuster et al (2010) review evidence from lab and field settings documenting deviations from rational expectations.

is much more common among the Irish than in the average national group, even though the true share of red-haired Irish is small. Similarly, after seeing a patient test positive on a medical test for a disease doctors overestimate the likelihood that he has it because being sick is representative of testing positive, even when it remains unlikely despite a positive test (Casscells et al. 1978). This formalization of representativeness can account for several well-documented judgment biases, such as the conjunction and disjunction fallacies, as well as base rate neglect. It also provides a description of stereotypes consistent with a good deal of empirical evidence (Bordalo, Coffman, Gennaioli, and Shleifer (BCGS 2015)).

We show in this paper that this formalization of representativeness can be naturally applied to modeling expectations, where agents form beliefs about the future, and update them in light of new data. Analogously to the medical test example, agents focus on, and thus overweight in their beliefs, the future states whose likelihood increases the most in light of current news *relative to what they know already*. Just as doctors overestimate the probability of sickness after a positive test result, agents overestimate the probability of a future state when the news point in the direction of that state. Following TK (1983)'s description of the representativeness heuristic as focusing on diagnostic information, we refer to such beliefs as diagnostic expectations.

This approach has significant implications. For example, a rising path of news will lead to excess optimism, a declining path to excess pessimism, even when these paths lead to the same fundamentals. There is a kernel of truth in assessments (after a rising path the investor revises upwards, after a falling path he revises downwards) but revisions are excessive. When news stabilize, the agent no longer extrapolates change, but such a systematic cooling off of expectations itself leads to a reversal. Excessively volatile expectations drive cyclical fluctuations in both financial and economic activity.

We construct a neoclassical macroeconomic model in which the only non-standard feature is expectations. In particular, we do not include financial or any other frictions. The model accounts for many empirical findings, some of which also obtain under rational expectations, but some do not. In our model:

- 1) In response to good news about the economy, credit spreads decline, credit expands, the share of high risk debt rises, and investment and output grow.
- 2) Following this period of narrow credit spreads, these spreads predictably rise on average, credit and share of high risk debt decline, while investment and output decline as well. Larger spikes in spreads predict lower GDP growth.
- 3) Credit spreads are too volatile relative to fundamentals and their changes are predictable in a way that parallels the cycles described in points 1) and 2).
- 4) There are predictable forecast errors in investor beliefs, and thus systematic abnormal bond returns, that parallel the cycles described in points 1) and 2).

Prediction 1) can obtain under rational expectations, and the same is true about prediction 2) provided fundamentals are mean reverting. Predictions 3) and 4), in contrast, centrally depend on our model of diagnostic expectations.

In the next section, we present our basic macroeconomic model. Section 3 focuses on our specification of the expectations mechanism, and introduces diagnostic expectations. It also describes how expectations evolve in our model, and relates our formulation to extrapolation and neglected risk. Section 4 examines credit markets in this model, and in particular focuses on credit spreads, total credit, and credit share accruing to risky firms. Section 5 turns to the predictions of the model for aggregate macroeconomic volatility. Section 6 describes some alternative ways to specify diagnostic expectations, and considers their implications. Section 7 concludes.

2. The model

2.1 Production

Time is discrete $t = 0, 1, \dots$. The state of the economy at t is captured by a random variable $\Omega_t \in \mathbb{R}$, whose realization is denoted by ω_t . This random variable follows a Markov process whose conditional distribution on Ω_{t-1} is normal, as in the AR(1) case $(\omega_t - \bar{\omega}) = b(\omega_{t-1} - \bar{\omega}) + \epsilon_t$, with $\epsilon_t \sim N(0, \sigma^2)$ and $\bar{\omega} \in \mathbb{R}$, $b \in [0, 1]$.

A measure 1 of atomistic firms uses capital to produce output. The productivity of firms at t depends on the state ω_t , but to a different extent for different firms. Each firm is identified by its risk $\rho \in \mathbb{R}$. Firms with higher ρ are less likely to be productive in any state ω_t . Formally, if at t a firm of type ρ has invested capital k , its output is given by:

$$y(k|\omega_t, \rho) = \begin{cases} k^\alpha & \text{if } \omega_t \geq \rho \\ 0 & \text{if } \omega_t < \rho \end{cases} \quad (1)$$

where $\alpha \in (0, 1)$. The firm produces only if it is sufficiently safe, $\rho < \omega_t$. Safe firms, for which $\rho = -\infty$, produce k^α in every state of the world. The higher is ρ , the better the state ω_t of the economy needs to be for the firm to pay off. At the same capital k , two firms produce the same output if they are both active, namely if $\omega_t \geq \rho$ for both firms.

A firm's riskiness is common knowledge and ρ is distributed across firms with density $f(\rho)$. Capital for production at $t + 1$ must be installed at t , before ω_{t+1} is known. Capital fully depreciates after usage. At time t each firm ρ demands funds $D_{t+1}^f(\rho)$ from a competitive financial market to finance its capital investment, namely $D_{t+1}^f(\rho) = k_{t+1}(\rho)$. The firm issues risky debt that promises a contractual interest rate $r_{t+1}(\rho)$. Debt is repaid only if the realized state of the economy allows the firm to be productive. If at t the firm borrows $D_{t+1}^f(\rho)$ at the interest rate $r_{t+1}(\rho)$, next period it produces and repays $r_{t+1}(\rho)D_{t+1}^f(\rho)$ provided $\omega_{t+1} \geq \rho$, and defaults otherwise.

Because there are no agency problems and each firm's output has a binary outcome, the model does not distinguish between debt and equity issued by the firm. Both contracts are contingent on the same outcome and promise the same rate of return. For concreteness, we refer to the totality of capital invested as debt.

2.2 Households

A risk neutral, infinitely lived, representative household discounts the future by a factor $\beta < 1$. At each time t , the household allocates its current income between consumption and investment by maximizing its expectation of the utility function:

$$\sum_{s=t}^{+\infty} \beta^{s-t} c_s.$$

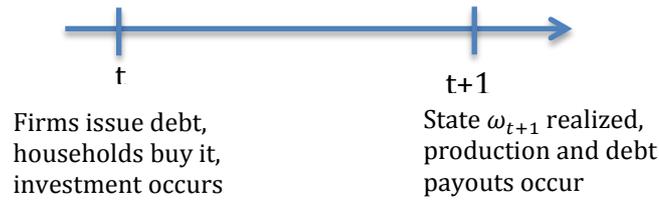
The household's investment consists in buying the claims issued by firms – which then pay out or default in the next period – while its income consists of the payout of debt bought in the previous period, the profits of firms (which are owned by the household), and a fixed endowment w . Thus, for each time s and state ω_s its budget constraint is:

$$c_s + \int_{-\infty}^{+\infty} D_{s+1}^h(\rho) f(\rho) d\rho = w + \int_{-\infty}^{+\infty} I(\rho, \omega_s) [r_s(\rho) D_s^h(\rho) + \pi_s(\rho)] f(\rho) d\rho,$$

where c_s is consumption, $D_{s+1}^h(\rho)$ is capital supplied to firm ρ , $I(\rho, \omega_s)$ is an indicator function equal to one when firm ρ repays, namely when $\omega_s \geq \rho$, and $\pi_s(\rho)$ is the profit of firm ρ when active. The household's income depends, via debt repayments, on the state of the economy: the worse is the current state (the lower is ω_s), the higher is the fraction of firms that default and thus the lower is the household's income.⁴

The timeline of an investment cycle in the model is illustrated below.

⁴ As we show later, the endowment w ensures that household income is high enough that the equilibrium expected return (and the expected marginal product of capital) is equal to β^{-1} . We could alternatively assume that there is a riskless, fixed size, technology that guarantees such income to the household.



Investment decisions by households and firms depend on the *perceived* probability with which each firm type ρ repays its debt in the next period. When the current state is ω_t , the objective probability of repayment is given by:

$$\mu(\rho|\omega_t) = \Pr(\omega_{t+1} \geq \rho|\omega_t) = \int_{\rho}^{+\infty} h(\Omega_{t+1} = \omega|\Omega_t = \omega_t) d\omega, \quad (2)$$

where $h(\Omega_{t+1} = \omega|\Omega_t = \omega_t)$ is the probability density of next period's state conditional on the current state. The firm defaults with complementary probability $1 - \mu(\rho|\omega_t)$.

Under rational expectations, the repayment probability expected by households and firms is given by the objective measure $\mu(\rho|\omega_t)$. This is not the case when representativeness shapes beliefs. We now introduce a model of diagnostic expectations based on the representativeness heuristic.

3. Diagnostic Expectations

3.1 A Formal Model of Representativeness

We build our model of expectations from the ground up, starting with research on heuristics and biases in human decision making. One of Kahneman and Tversky's most universal decision heuristics is representativeness, which they define as follows: "an attribute is representative of a class if it is very diagnostic; that is, the relative frequency of this attribute is much higher in that class than in the relevant reference class (TK 1983)." Kahneman and Tversky argue that individuals often assess likelihood by representativeness, thus estimating types or attributes as being likely when they are

instead representative, and present a great deal of experimental evidence to support this claim. Gennaioli and Shleifer (2010) build a model in which judgment biases arise because decision makers overweight events that are representative precisely in the sense of KT's definition. In this section, we summarize some of this work as well as a closely related application to stereotype formation by Bordalo et al. (BCGS 2015) to motivate our model of diagnostic expectations.

A decision maker judges the distribution of a trait T in a group G . The true distribution of the trait is $h(T = t|G)$, but – due to limited working memory – the decision maker disproportionately attends to representative traits, which are easier to recall. In GS (2010), the representativeness of the trait $T = t$ for group G is defined as:

$$\frac{h(T = t|G)}{h(T = t| -G)}$$

where $-G$ is a relevant comparison group. As in KT's quote, a trait is more representative if it is *relatively* more frequent in G than in $-G$. The agent then forms his assessment of the distribution h by overweighing highly representative traits.

To illustrate, consider an individual assessing hair color among the Irish. The trait T he must predict is hair color, the conditioning group G is the Irish. The comparison group $-G$ is the world at large. The true distributions of hair color are:⁵

	$T = red$	$T = blond/light brown$	$T = dark$
$G \equiv Irish$	10%	40%	50%
$-G \equiv World$	1%	14%	85%

The most representative hair color for the Irish is red because it is associated

⁵ See http://www.eupedia.com/genetics/origins_of_red_hair.shtml and http://www.eupedia.com/europe/genetic_maps_of_europe.shtml. Shares of red hair are based on the sampled distribution of the small number of genetic variants that cause this phenotype. As such, these shares are relatively accurate. The definitions, and shares, of brown and dark hair are approximate.

with the highest likelihood ratio among hair colors:

$$\frac{Pr(\text{red hair}|\text{Irish})}{Pr(\text{red hair}|\text{World})} = \frac{10\%}{1\%} = 10.$$

Our model of judgment by representativeness thus predicts that assessments exaggerate the frequency of red haired Irish, which they do. Decision makers here consider the possibility of blond or dark-haired Irish, but underestimate the frequency of these more likely but less representative traits. Intuitively, after hearing the news “Irish”, the agent inflates in his beliefs the likelihood of the hair color “red”, whose probability has increased the most relative to the baseline but remains low. BCGS (2015) show that this model explains many empirical features of stereotypes, including the finding that they contain a “kernel of truth” (Judd and Park 1993).

As shown in GS (2010) and BCGS (2015), representativeness also sheds light on fundamental biases in information processing, such as base rate neglect. Suppose that a doctor must assess the health of a patient in light of the result of a medical test. In this case, $T = \{\text{healthy}, \text{sick}\}$, G is the tested patients and includes patients who test positive $G = +$, and patients who test negative, $G = -$. The comparison group $-G$ is untested patients. Suppose that the test comes out positive, namely we are looking at $G = +$, and want to assess the probability that the patient is actually sick. In this case, the sick patient is representative of $G = +$ provided:

$$\frac{Pr(T = \text{sick}|G = +)}{Pr(T = \text{sick}|-G)} > \frac{Pr(T = \text{healthy}|G = +)}{Pr(T = \text{healthy}|-G)}$$

This condition is always satisfied provided the test is minimally informative, in the sense that a positive result raises the likelihood of having the disease relative to no information. By overweighting the representative scenario, the doctor inflates the probability that the patient is sick. The psychological mechanism here is the same as the red-haired Irish stereotype: after learning that the test is positive, the doctor inflates

the probability of the outcome “sick” because it has increased. Diagnostic information is overweighed. In doing so, the doctor might neglect the base rate information that the disease in question is very rare, and its likelihood even after a positive test is very low. This analysis is consistent with the evidence in Casscells et al. (1978) on base rate neglect by physicians.⁶ A significant literature in psychology explores this finding, and the mechanism above captures TK’s (1974) verbal account of base rate neglect.

In GS (2010) and BCGS (2015) we show that this approach to the representativeness heuristic provides a unified account of several widely documented judgment biases, including base rate neglect, conjunction and disjunction fallacy, over- and under-reaction to data, as well as the confirmation bias. It also explains empirical evidence on beliefs about different genders, races, and ethnic groups on a variety of dimensions. We next show that the same logic of representativeness can be naturally applied to analyzing the evolution of beliefs in a macroeconomic context.

3.2 Diagnostic Expectations

Our model of representativeness is portable to dynamic environments such as the Markov process, and in particular the AR(1) process, described in Section 2. Here, the agent seeks to represent the distribution of a future state, say Ω_{t+1} , entailed by current conditions $\Omega_t = \omega_t$. The model is easily generalized to longer term predictions Ω_{t+T} and to richer AR(N) processes. As in the medical test example, where the doctor assesses the health of the patient conditional on a positive test outcome, here the agent

⁶ Casscells et al (1978) asked physicians: “If a test to detect a disease whose prevalence is 1/1000 has a false positive rate of 5 per cent, what is the chance that a person found to have a positive result actually has the disease, assuming that you know nothing about the person’s symptoms or signs?” While the correct answer is $\Pr(G + |T = sick) \frac{\Pr(T=sick)}{\Pr(G+)} \cong 2\%$, they found an average response of 56%, while the modal answer was $\Pr(G + |T = sick) = 95\%$.

assesses the distribution of Ω_{t+1} conditional on the current state $\Omega_t = \omega_t$. We refer to $G \equiv \{\Omega_t = \omega_t\}$ as the “group” of all possible future states conditional on $\Omega_t = \omega_t$.

A rational agent assesses Ω_{t+1} using the true conditional distribution $h(\Omega_{t+1} = \omega_{t+1} | \Omega_t = \omega_t)$. The agent judging by representativeness also knows the true distribution, but he overweighs the probability of future states ω_{t+1} that are representative or diagnostic of $G \equiv \{\Omega_t = \omega_t\}$ relative to a comparison group $-G$.

In the medical example, after a positive test result the representative health status is one whose objective frequency goes up the most relative to not observing the test outcome. One natural way to extend this logic to a dynamic setting is to assume that, after seeing the *realized* current state $G \equiv \{\Omega_t = \omega_t\}$, the most representative future state ω_{t+1} is the one whose objective probability has increased the most relative to what the agent knew in the past. In this case, the comparison group is the *true* distribution in the absence of new information. This pins down $-G \equiv \{\Omega_t = \mathbb{E}_{t-1}(\omega_t)\}$, where $\mathbb{E}_{t-1}(\omega_t)$ denotes the past rational expectation for today.

There are other ways of specifying the comparison group $-G$.⁷ For instance, $-G$ could be slow moving, and include more remote recollections. Alternatively, the reference $-G$ may be specified in terms of diagnostic rather than rational expectations. In Section 6 we analyze these cases in detail.

The representativeness of state ω_{t+1} is then measured by the likelihood ratio:

$$\frac{h(\Omega_{t+1} = \omega_{t+1} | \Omega_t = \omega_t)}{h(\Omega_{t+1} = \omega_{t+1} | \Omega_t = \mathbb{E}_{t-1}(\omega_t))}, \quad (3)$$

⁷ One could almost equivalently define $-G \equiv \{\Omega_{t-1} = \omega_{t-1}\}$, which uses the same information set as the above. Now the comparison distribution is $h(\Omega_{t+1} = \omega_{t+1} | \Omega_{t-1} = \omega_{t-1})$. The analysis is more cumbersome (because the target and the comparison distributions have different variances), but the results are qualitatively the same. Expectations are still represented by Equation (4) but the coefficient on the incoming news becomes equal to $\frac{\theta r_\sigma^2}{1 + \theta(1 - r_\sigma^2)}$. In this expression, $r_\sigma^2 = \sigma^2 / \hat{\sigma}^2 < 1$, where σ^2 is the variance of $h(\Omega_{t+1} = \omega_{t+1} | \Omega_t = \mathbb{E}_{t-1}(\omega_t))$ and $\hat{\sigma}^2$ is the variance of $h(\Omega_{t+1} = \omega_{t+1} | \Omega_{t-1} = \omega_{t-1})$.

which is indeed higher for states whose objective likelihood increases the most in light of recent news $\omega_t - \mathbb{E}_{t-1}(\omega_t)$. We formalize overweighting of representative states by assuming that the agent attaches to future state ω_{t+1} the distorted probability density:

$$h_t^\theta(\omega_{t+1}) = h(\Omega_{t+1} = \omega_{t+1} | \Omega_t = \omega_t) \cdot \left[\frac{h(\Omega_{t+1} = \omega_{t+1} | \Omega_t = \omega_t)}{h(\Omega_{t+1} = \omega_{t+1} | \Omega_t = \mathbb{E}_{t-1}(\omega_t))} \right]^\theta \frac{1}{Z},$$

where Z is a normalizing constant ensuring that $h_t^\theta(\omega_{t+1})$ integrates to one, and $\theta \in (0, +\infty)$ measures the severity of thinking through representativeness. When $\theta = 0$, the agent holds rational expectations, which are thus a special case of our model. When $\theta > 0$, the distribution $h_t^\theta(\omega_{t+1})$ inflates the likelihood of representative states and deflates the likelihood of non-representative ones. Because they overweight the most representative, or diagnostic, future outcomes, we call the expectations formed in light of $h_t^\theta(\omega_{t+1})$ diagnostic.

In this formulation, news do not just alter the objective likelihood of certain states. They also change, through representativeness, the extent to which the agent focuses on particular states. An event that increases the likelihood of a future state ω_{t+1} also makes it more representative, so $h_t^\theta(\omega_{t+1})$ overshoots. The reverse occurs when the likelihood of ω_{t+1} decreases. As a consequence, if the likelihood ratio in (2') is monotone increasing, "rationally" good news $\omega_t > \mathbb{E}_{t-1}(\omega_t)$ cause the household to overweight high future states, and to underweight low future states (and conversely if news are bad). In this sense, good news cause neglect of downside risk.

Proposition 1 *When the process for ω_t is AR(1) with normal $(0, \sigma^2)$ shocks, the diagnostic distribution $h_t^\theta(\omega_{t+1})$ is also normal, with variance σ^2 and mean:*

$$\mathbb{E}_t^\theta(\omega_{t+1}) = \mathbb{E}_t(\omega_{t+1}) + \theta[\mathbb{E}_t(\omega_{t+1}) - \mathbb{E}_{t-1}(\omega_{t+1})]. \quad (4)$$

Diagnostic expectations can be represented as a linear combination of the rational expectation of the same variable ω_{t+1} held at dates t and $t - 1$. The interpretation is not that the decision-maker computes rational expectations and combines them according to Proposition 1. Rather, overweighting the *true* probability of representative future states yields expectations equivalent to the linear combination (4). When adjusting beliefs about future states that have become objectively more likely, agents focus on the news that drive the updating of beliefs, and overreact to these news. This feature reflects a “kernel of truth” logic: diagnostic expectations differ from rational expectations by a shift in the direction of the information received at t , given by $[\mathbb{E}_t(\omega_{t+1}) - \mathbb{E}_{t-1}(\omega_{t+1})]$.⁸

Figure 1 illustrates the entailed neglect of risk. After good news, the diagnostic distribution of ω_{t+1} is a right shift of the objective distribution. Due to the monotone increasing and unbounded likelihood ratio of normal densities, good news cause underestimation of probabilities in the left tail (the shaded area).

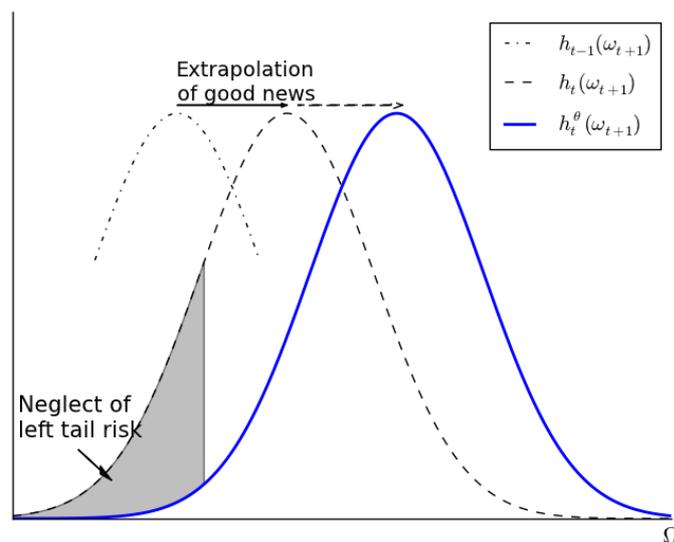


Figure 1

⁸The representation in (4) – and the tractability it entails – extends to the exponential class of distributions, including lognormal, exponential, and others.

This logic also provides a psychological foundation for extrapolative expectations. Writing the AR(1) process as $(\omega_t - \bar{\omega}) = b(\omega_{t-1} - \bar{\omega}) + \epsilon_t$, with persistence parameter b , Equation (4) becomes:

$$\mathbb{E}_t^\theta(\omega_{t+1}) - \omega_t = [\mathbb{E}_t(\omega_{t+1}) - \omega_t] + b\theta[\omega_t - \mathbb{E}_{t-1}(\omega_t)], \quad (5)$$

Diagnostic expectations ($\theta > 0$) extrapolate the current shock $\omega_t - \mathbb{E}_{t-1}(\omega_t)$ into the future. This comes from overreaction to information contained in $\omega_t - \mathbb{E}_{t-1}(\omega_t)$ (and only if the data are serially correlated, $b > 0$). Diagnostic expectations exaggerate the role of surprising new information, consistent with Kahneman's (2011) view that "our mind has a useful capability to focus spontaneously on whatever is odd, different, or unusual." This property differentiates our model from backward looking ones such as adaptive expectations. We return to this point in Section 6.

As indicated in the introduction, recent research in finance points to the prevalence of extrapolative expectations among economic agents (e.g., Greenwood and Shleifer 2014, Gennaioli, Ma, and Shleifer 2015). A parallel line of work finds evidence that agents in financial markets occasionally neglect tail risks (Gennaioli, Shleifer, and Vishny 2012, Coval, Pan, and Stafford 2014, Arnold, Schuette, and Wagner 2015). As Figure 1 illustrates, in our model neglect of risk and extrapolation are connected by the same psychological mechanism. Good news render good future states representative. The agent thus overweighs good future states, effectively extrapolating the news into the future and downplaying the probability of future bad events.

3.3 Properties of Diagnostic Expectations

To contrast our model with rational expectations, we show first that diagnostic expectations: i) exhibit too much volatility, and ii) are systematically wrong.

Proposition 2. *Equation (4) implies that:*

i) At $t - 1$, the variance of future diagnostic expectations $\mathbb{E}_t^\theta(\omega_{t+1})$ increases in θ :

$$\text{Var}[\mathbb{E}_t^\theta(\omega_{t+1})|\omega_{t-1}] = (1 + \theta)^2 \text{Var}[\mathbb{E}(\omega_{t+1}|\omega_t)|\omega_{t-1}].$$

ii) At t , diagnostic expectations lead to a predictable forecast error:

$$\mathbb{E}[\omega_{t+1} - \mathbb{E}_t^\theta(\omega_{t+1})|\omega_t] = -\theta[\mathbb{E}_t(\omega_{t+1}) - \mathbb{E}_{t-1}(\omega_{t+1})].$$

Diagnostic expectations display excess volatility because they are distorted in the direction of realized news. With positive (negative) news, expectations are too optimistic (pessimistic). Excess volatility also causes predictable forecasting errors. After good news, the error is negative, the more so the better the current state $\mathbb{E}_t(\omega_{t+1})$, and conversely after bad news. These predictions are illustrated in Figure 2 below. Panel A simulates a path of an AR(1) process and shows excess volatility of diagnostic expectations $\mathbb{E}_t^\theta(\omega_{t+1})$ (solid red line) relative to rational expectations $\mathbb{E}_t(\omega_{t+1})$ (dashed blue line).⁹ For the same simulation, Panel B documents the negative correlation between forecast errors $\omega_{t+1} - \mathbb{E}_t^\theta(\omega_{t+1})$ and current conditions ω_t .

Gennaioli, Ma, and Shleifer (2015) analyze quarterly data on the expectations of earning growth reported by CFOs of large U.S. corporations during the period 2005-2012. They find that these expectations are too volatile relative to fundamentals and that they have an extrapolative structure: the error in forecasting earnings growth is negatively related to past earnings. This evidence is inconsistent both with Rational Expectations and with models where decision makers underreact to news (such as adaptive expectations for persistent processes). Diagnostic expectations offer a psychologically founded way to account for these findings, as shown in Figure 2.

⁹ The simulated process is $\omega_t = 0.7\omega_{t-1} + \varepsilon_t$, with shocks $\varepsilon_t \sim \mathcal{N}(0,1)$ i.i.d. across time. The simulation started at $\omega_t = 0$ (the long-term mean of the process), and was run for 150 periods. The stereotypical thinking parameter was set at $\theta = 1$, and the technology scaling parameter at $\alpha = 0.8$.

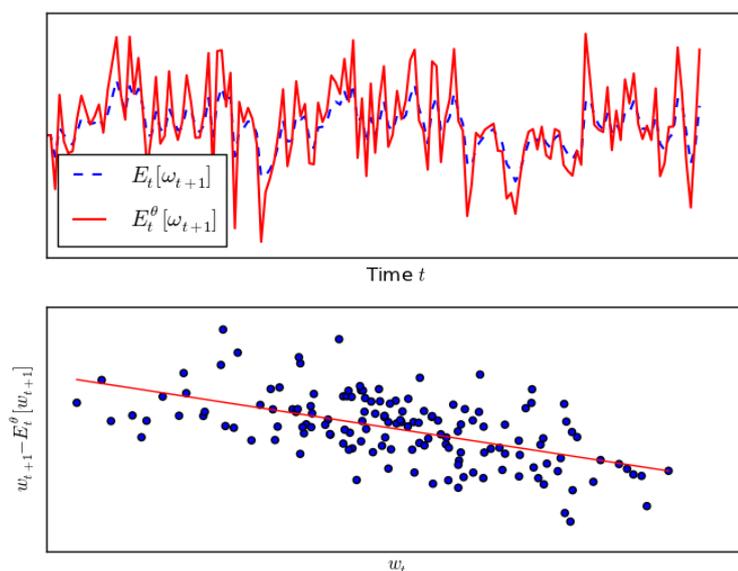


Figure 2

Diagnostic expectations can yield over- as well as under-reaction to news. Equation (4) generates over-reaction to accelerating repeated news in the same direction, such as when a small positive news is reinforced by bigger positive news. When alternatively the news fail to maintain momentum, as when a large positive shock is followed by a small positive shock, Equation (4) entails under-reaction.¹⁰

This prediction is illustrated in Figure 3, which shows two alternative paths of a random walk ($b = 1$) between the same end-points. Rational Expectations always coincide with the current state of the world (dotted line, $\theta = 0$), but diagnostic expectations do not (solid line, $\theta = 1$). The top panel shows that accelerating news cause diagnostic expectations to be too optimistic; conversely, a strong deceleration at the end of this streak causes a major reversal in expectations (back to rationality). The bottom panel shows that decelerating good news can, through cooling off of previous

¹⁰ Formally, $|\mathbb{E}_t^\theta(\omega_{t+1}) - \mathbb{E}_{t-1}^\theta(\omega_{t+1})| > |\mathbb{E}_t(\omega_{t+1}) - \mathbb{E}_{t-1}(\omega_{t+1})|$ if and only if the following condition holds: $|\mathbb{E}_t(\omega_{t+1}) - \mathbb{E}_{t-1}(\omega_{t+1})| > |\mathbb{E}_{t-1}(\omega_{t+1}) - \mathbb{E}_{t-2}(\omega_{t+1})|$.

overreaction, cause expectations to underreact, and even move in the opposite direction of news. However, in this case major reversals are avoided.

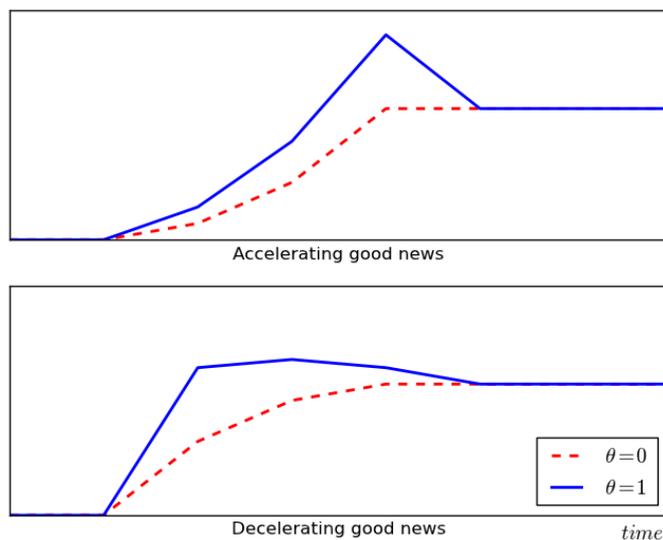


Figure 3

More generally, diagnostic expectations explicitly address the Lucas critique. This critique holds that mechanical models of expectations can be used for policy evaluation only under the very restrictive assumption that the expectations formation process is invariant to changes in policy. The validity of this assumption was challenged after empirical estimates of adaptive expectations processes revealed large parameter instability to policy change. This instability led researchers to prefer rational expectations, which account for changes in expectations due to regime shifts. But, importantly, responding to the Lucas critique does not necessitate rational expectations; it merely requires that expectations be at least in part forward looking.

Diagnostic expectations share this important feature with rational expectations. They are immune to the Lucas critique in the precise sense that expectations depend on the true data generating process. Because diagnostic expectations distort the true distribution $h_t(\omega_{t+1})$, they respond to policy shifts that affect $h_t(\omega_{t+1})$. The diagnostic

distribution $h_t^\theta(\omega_{t+1})$ incorporates changes in the objective frequency (as under rational expectations) but also changes in representativeness. Thus, if the government commits to inflating the economy, inflation expectations will also react upwards.

In his path-breaking paper introducing Rational Expectations, Muth (1961) discusses the possibility that expectations over or under-react to news. To capture these phenomena while keeping model consistency, he proposes a generalization of rational expectations that allows for systematic errors in expectations. Muth's formula is precisely of the linear form of Equation (4): relative to rationality, expectations can distort the effect of recent news. Our model shows that Muth's formulation is not ad-hoc: it follows from a natural formalization of the psychology of representativeness.

We now return to the model and show that the psychology of representativeness generates excess volatility in beliefs, over-heating and over-cooling of credit markets, as well as predictable shifts in market sentiment, credit spreads, and economic activity that are consistent with many observed features of credit cycles.

4. Equilibrium under Diagnostic Expectations

4.1 Capital Market Equilibrium and Credit Spreads

At time t firm ρ demands capital $k_{t+1}(\rho)$ at the market contractual interest rate $r_{t+1}(\rho)$ so as to maximize its expected profit at $t + 1$:

$$\max_{k_{t+1}(\rho)} (k_{t+1}(\rho)^\alpha - k_{t+1}(\rho) \cdot r_{t+1}(\rho)) \cdot \mu_t^\theta(\rho), \quad (6)$$

where investment is financed with debt issuance $k_{t+1}(\rho) = D_{t+1}^f(\rho)$. The function $\mu_t^\theta(\rho)$ denotes the firm's believed probability at time t that it is productive at $t + 1$. It is obtained by computing Equation (2) under the diagnostic distribution:

$$\mu_t^\theta(\rho) = \int_\rho^{+\infty} h_t^\theta(\omega_{t+1}) d\omega_{t+1}. \quad (7)$$

The first order condition for the profit maximization problem is given by:

$$k_{t+1}(\rho) = \left[\frac{\alpha}{r_{t+1}(\rho)} \right]^{\frac{1}{1-\alpha}}, \quad (8)$$

which is the usual downward sloping demand for capital.

Households are willing to supply any amount of capital to firm ρ at the interest rate $r_{t+1}(\rho)$ at which the expected repayment of the firm is just sufficient for the household to be willing to postpone its consumption:

$$r_{t+1}(\rho)\mu_{t+1}^\theta(\rho) = \beta^{-1} \Leftrightarrow r_{t+1}(\rho) = \frac{1}{\beta\mu_{t+1}^\theta(\rho)}. \quad (9)$$

In equilibrium, this condition must hold for all firms ρ . First, debt of all firms must yield the same expected return, which cannot be below β^{-1} . Otherwise, the household would not supply capital to some or all firms. There would thus exist investment opportunities yielding an infinite expected return, which cannot occur in equilibrium. Second, if the expected return from investment rose above β^{-1} , the household would save the totality of its income. As previously noted, the endowment w is large enough that the marginal product of capital would fall below β^{-1} , which – as argued above – cannot occur in equilibrium. Formally, we assume:

A.1 $w \geq (\alpha\beta)^{\frac{1}{1-\alpha}}$.

This condition implies that even if all firms are believed to repay for sure ($\mu_{t+1}^\theta(\rho) = 1$ for all ρ), the endowment is large enough that, if invested entirely, would drive the marginal return on capital below β^{-1} . We assume A.1 throughout.

By combining Equations (8) and (9) we obtain the volume of debt that firm ρ issues at t and households purchase, as well as the firm's installed capital stock at $t + 1$:

$$k_{t+1}(\rho) = [\alpha\beta\mu_t^\theta(\rho)]^{\frac{1}{1-\alpha}}. \quad (10)$$

Using firm level investment, we can compute other key real variables. Aggregate investment at t , and thus capital installed at $t + 1$, is given by:

$$K_{t+1} = \int_{-\infty}^{+\infty} [\alpha\beta\mu_t^\theta(\rho)]^{\frac{1}{1-\alpha}} f(\rho) d\rho. \quad (11)$$

Aggregate output at $t + 1$ conditional on state ω_{t+1} :

$$Y_{t+1}(\omega_{t+1}) = \int_{-\infty}^{\omega_{t+1}} [\alpha\beta\mu_t^\theta(\rho)]^{\frac{\alpha}{1-\alpha}} f(\rho) d\rho. \quad (12)$$

Financial markets and production are shaped by the perceived creditworthiness of firms, $\mu_t^\theta(\rho)$. When times are good, households are optimistic about the future state of the economy. The perceived creditworthiness of firms is high, households supply more capital, the interest rate falls, firms issue more debt and invest more, and future output rises. When times turn sour, households cut lending, firms issue less debt and cut investment, and the economy contracts.

4.2 Credit Spreads and the short run response to shocks

We now analyze the role of diagnostic expectations in determining perceived creditworthiness of firms and, through that, the behavior of the economy. We focus on the short term link between credit spreads, debt issuance, and investment. We analyze the economic cycle in Section 5.

Under the assumed AR(1) process, the perceived creditworthiness of firm ρ at time t (i.e. its assessed probability of repayment at $t + 1$) is given by:

$$\mu_t^\theta(\rho) = \frac{1}{\sigma\sqrt{2\pi}} \int_{\rho}^{+\infty} e^{-\frac{(\omega_{t+1} - \mathbb{E}_t^\theta(\omega_{t+1}))^2}{2\sigma^2}} d\omega_{t+1}. \quad (13)$$

A perfectly safe firm $\rho \rightarrow -\infty$ never defaults, since $\lim_{\rho \rightarrow -\infty} \mu_{t+1}^\theta(\rho) = 1$. By Equation (8), then, it promises the safe interest rate $\lim_{\rho \rightarrow -\infty} r_{t+1}(\rho) = \beta^{-1}$.

Riskier firms have to compensate debt holders for bearing their default risk by promising contractual interest rates above β^{-1} . The spread obtained on their debt relative to the debt of safe firms is given by:

$$S_t^\theta(\rho) \equiv r_{t+1}(\rho) - \beta^{-1} = \left(\frac{1}{\mu_t^\theta(\rho)} - 1 \right) \beta^{-1}. \quad (14)$$

The credit spread increases for riskier firms, namely those with lower $\mu_t^\theta(\rho)$, so riskier firms borrow and invest less in equilibrium:

$$k_{t+1}(\rho) = \left(\frac{\alpha\beta}{1 + \beta S_t^\theta(\rho)} \right)^{\frac{1}{1-\alpha}}.$$

To proceed, consider how a firm's perceived creditworthiness depends on the current state ω_t . After some algebra, we can show:

$$\frac{\partial \ln \mu_t^\theta(\rho)}{\partial \omega_t} = \left[\mathbb{E}_t^\theta(\omega_{t+1} | \omega_{t+1} \geq \rho) - \mathbb{E}_t^\theta(\omega_{t+1}) \right] \frac{b(1+\theta)}{\sigma^2} > 0. \quad (15)$$

Better news boost households' perception of creditworthiness. The effect is proportionally stronger for riskier firms, because $\partial^2 \ln \mu_t^\theta(\rho) / \partial \omega_t \partial \rho > 0$. Riskier firms are more exposed to the aggregate state, so they benefit relatively more when economic conditions improve. Improvements in perceived creditworthiness are also stronger when θ is higher: diagnostic expectations cause a stronger reaction of perceived creditworthiness to better news.

We then have:

Proposition 3. *As current aggregate conditions ω_t improve:*

i) spreads drop and become compressed:

$$\frac{\partial S_t^\theta(\rho)}{\partial \omega_t} = - \frac{1}{\beta \sigma^2 \mu_t^\theta(\rho)} \frac{\partial \ln \mu_t^\theta(\rho)}{\partial \omega_t} < 0, \quad \frac{\partial^2 S_t^\theta(\rho)}{\partial \omega_t \partial \rho} < 0.$$

ii) debt issuance and investment increase, disproportionately so for riskier firms:

$$\frac{\partial K_{t+1}}{\partial \omega_t} = \left(\frac{1}{1-\alpha} \right) \int_{-\infty}^{+\infty} \frac{\partial \ln \mu_t^\theta(\rho)}{\partial \omega_t} k_{t+1}(\rho) f(\rho) d\rho > 0,$$

$$\frac{\partial}{\partial \omega_t} \frac{k_{t+1}(\rho_1)}{k_{t+1}(\rho_2)} \propto \frac{\partial \ln \mu_t^\theta(\rho_1)}{\partial \omega_t} - \frac{\partial \ln \mu_t^\theta(\rho_2)}{\partial \omega_t} > 0 \text{ for all } \rho_1 > \rho_2.$$

As current and thus expected conditions improve, firms' perceived creditworthiness improves as well. As a consequence, interest rate spreads charged to risky firms fall. This decline in the cost of borrowing is greater for riskier firms, so that credit spreads become compressed. The decline in the borrowing costs in turn stimulates debt issuance and aggregate investment. Once again, because spreads fall more for riskier firms, the increase in debt issuance is greater for those firms.

These predictions of the model are consistent with the evidence of Greenwood and Hanson (2013). They document that when the BBB-credit spread falls, bond issuance increases and the effect is particularly strong for firms characterized by higher expected default rates. As a consequence, the share of non-investment grade debt over total debt (the "junk share") increases, as has also been documented by LSZ (2015).

Our model also accounts for the behavior of the junk share. Consider the share of debt issued by firms riskier than an arbitrary threshold $\hat{\rho}$:

$$\frac{\int_{\hat{\rho}}^{+\infty} k_{t+1}(\rho) f(\rho) d\rho}{K_{t+1}}.$$

This quantity increases as spreads become compressed (for any $\hat{\rho}$). The opposite effects arise when economic conditions deteriorate. Credit spreads increase, investment falls and there is a flight to safety, so that the junk share falls as well.

In our model, this relation between spreads and debt issuance is triggered by the asymmetric exposure of different firms to changes in fundamentals assumed in (1). More generally, the qualitative effects described in Proposition 3 do not rely on diagnostic expectations and obtain even if households are fully rational, since, as

illustrated in Equation (15), positive news always leads to an upward revision of firms' creditworthiness, and the more so for riskier firms.

However, diagnostic expectations complement the effect of fundamentals and allow the model to match additional pieces of evidence that the rational expectations assumption cannot accommodate. In particular, the excess volatility and systematic errors in diagnostic expectations described in Proposition 2 immediately translate into excess volatility and abnormal returns in credit markets:

Proposition 4. *Suppose that at time t new information $[\mathbb{E}(\omega_{t+1}|\omega_t) - \mathbb{E}(\omega_{t+1}|\omega_{t-1})]$ arrives. Then, under diagnostic expectations $\theta > 0$ we have that:*

i) Credit spreads overreact:

$$\frac{\partial S_t^\theta(\rho)}{\partial \theta} = \frac{\partial S_t^\theta(\rho)}{\partial \omega_t} \frac{[\mathbb{E}_t(\omega_{t+1}) - \mathbb{E}_{t-1}(\omega_{t+1})]}{b(1 + \theta)}.$$

Moreover, spreads exhibit excess volatility. To a second order approximation in θ :

$$\text{Var}[S_t^\theta(\rho)|\omega_t] \approx (1 + \theta)^2 \text{Var}[S_t^0(\rho)|\omega_t],$$

where $S_t^0(\rho)$ denotes the spread in the rational expectations benchmark ($\theta = 0$). As a consequence, aggregate investment also displays excess volatility.

ii) Positive news compresses credit spreads $S_t(\rho)$ at time t and predicts low average debt returns $\frac{\mu_{t+1}(\rho)}{\mu_{t+1}^\theta(\rho)} \beta^{-1} < \beta^{-1}$ at time $t + 1$. With negative news, the reverse happens.

As shown in Section 3, diagnostic expectations exaggerate the reaction of beliefs to new information. For a given ω_{t-1} , households become too optimistic when new information at t is positive and too pessimistic when it is negative. Diagnostic expectations lead to an excessive spread compression when economic times are improving, and conversely to an excessive widening in spreads when economic times are deteriorating. These effects are stronger for riskier firms.

As a consequence of overreaction, diagnostic expectations exhibit excess volatility with respect to the true underlying volatility of fundamentals. As we have seen in Proposition 2, excess volatility is due to the fact that beliefs do not just depend on the *level* of the current fundamentals ω_t (as would be the case under rational expectations). They also depend on the magnitude of the recently observed news, which corresponds roughly speaking to the *change* in fundamentals. Reaching a given level of fundamentals through large positive news causes more optimism than reaching the same level via a small upgrade. This path-dependence introduces excess volatility as well as predictability of expectations and spreads that helps account for several empirical findings.

First, several papers document that credit spreads appear too volatile relative to what could be explained by the volatility in default rates or fundamentals (Collin-Dufresne et al. 2001, Gilchrist and Zakrasjek 2012). For instance, Collin-Dufresne et al. (2001) find that credit spreads display excess volatility relative to measures of fundamentals such as realized default rates, liquidity, or business conditions. They argue this excess volatility can be explained by a common factor that captures aggregate shocks in credit supply and demand. Our model suggests that investors' excessive reaction to changing news can offer an account of these shocks.

Excess volatility in reaction to news yields another key implication of diagnostic expectations: predictable forecast errors. In a financial market context, this implies the existence of predictable anomalous returns. As Proposition 4 shows, when perceived creditworthiness is too high, $\mu_{t+1}^\theta(\rho) > \mu_{t+1}(\rho)$, credit spreads are too low. Going forward, the average return on debt is anomalously low, in our model lower than the investor's required return β^{-1} . The reverse is the case when perceived

creditworthiness is too low, $\mu_{t+1}^{\theta}(\rho) < \mu_{t+1}(\rho)$: credit spreads are now too high and there are positive abnormal returns going forward.

Greenwood and Hanson (2013) document the pattern of return predictability in Proposition 4. They find that high levels of the junk share predict anomalously low, and even negative, excess returns (point ii), and that this occurs precisely after good news, measured by drops in expected default rates (point i).¹¹ They consider conventional explanations for this finding, such as time varying risk aversion and financial frictions, but conclude that the evidence (particularly negative returns) is more consistent with the hypothesis that the junk share is a proxy for investor sentiment and extrapolation. Diagnostic expectations offer a psychological foundation for this account.

Proposition 4 describes how excess volatility in financial markets translates into excess volatility in the real economy, as measured by real investment and the economic return on this investment. Gennaioli, Ma and Shleifer (2015) find that CFOs with more optimistic earnings expectations invest more. Greenwood and Hanson (2015) study empirically investment cycles in the ship industry. Consistent with our model, they find that returns to investing in dry bulk ships are predictable and tightly linked to boom-bust cycles in industry investment. High current ship earnings are associated with higher ship prices and higher industry investment, but predict low future returns on capital.

In sum, diagnostic expectations lead to short-term extrapolative behavior, which is in line with a large set of recent empirical findings on both financial markets and production, including: i) excess volatility of spreads relative to measures of fundamentals, ii) excessive spread compression in good times and excessive spread

¹¹ One intuitive way to see this is to note (see Equation (16)) that the credit terms obtained by riskier firms are more sensitive to the biases caused by diagnostic expectations than those obtained by safer firms. Periods of excess optimism witness an abnormal increase in the junk share and disappointing subsequent returns. Periods of excess pessimism see the reverse pattern.

widening in bad times, iii) a similar pattern in the junk share, which expands excessively in good times and contracts excessively in bad times, iv) excessively volatile investment and output, and finally, v) good times predicting abnormally low returns.

Having assessed the implication of diagnostic expectations for the immediate reaction of the economy to a shock, we now study their implications for *cycles*.

5. Credit and Economic Cycles

We now analyze the implications of our model for the link between credit markets and economic activity. To address existing empirical work, we perform two exercises. First, we explore how a tightening in credit spreads at t affects output at $t + 1$. Second, we examine credit cycles by analyzing the link between credit spreads at $t - 1$, credit spreads at t , and output at $t + 1$.

Krishnamurthy and Muir (2015) document that a tightening of credit spreads at t induces an output contraction in period $t + 1$. Our model yields this pattern. Suppose that bad news at time t cause expectations of future fundamentals $\mathbb{E}_t^\theta(\omega_{t+1})$ to drop by $\Delta\mathbb{E}_t^\theta(\omega_{t+1}) < 0$. Proposition 3 and Equation (12) together lead to the following result:

Corollary 1. *For any $\theta \geq 0$, an adverse shock $\Delta\mathbb{E}_t^\theta(\omega_{t+1}) < 0$ to expectations at t causes a predictable change in total output at $t + 1$ given by:*

$$\mathbb{E}_t \left[\frac{1}{(1 - \alpha)\beta} \cdot \int_{-\infty}^{\omega_{t+1}} [\alpha\beta\mu_t^\theta(\rho)]^{1-\alpha} \frac{\partial \ln \mu_t^\theta(\rho)}{\partial \mathbb{E}_t^\theta(\omega_{t+1})} f(\rho) d\rho \right] \Delta\mathbb{E}_t^\theta(\omega_{t+1}) < 0$$

The expression above illustrates the link through which household pessimism at t predictably reduces output at $t + 1$. As confidence drops, perceived creditworthiness at t also declines. This effect, captured by the term $\partial \ln \mu_t^\theta(\rho) / \partial \mathbb{E}_t^\theta(\omega_{t+1})$ inside the integral, increases spreads at t , reducing debt issuance and investment and leading to a

decline in aggregate output at $t + 1$. The magnitude of the output drop is tempered by the fact that the firms cutting investment the most, the riskier ones, invest less to begin with. Still, because all firms cut investment (except for the safest ones $\rho = -\infty$), output declines. This decline is larger the more pronounced is the hike in credit spreads.

This result, which obtains when bad news cause households to downgrade their perception of firms' creditworthiness, becomes stronger with diagnostic expectations. When $\theta > 0$, bad news cause greater pessimism about future conditions. As a result, for given bad news and controlling for fundamentals, the term $\Delta E_t^\theta(\omega_{t+1})$ in Corollary 1 is on average more negative, the tightening of credit spreads is more pronounced, and the output decline is larger.

LSZ (2015) confirm that a tightening in credit markets at t is associated with a drop in output at $t + 1$, and show further that the current tightening can be predicted by credit conditions in the previous period. In particular, LSZ document that low credit spreads at $t - 1$ systematically predict high credit spreads at t and then a drop in output at $t + 1$. This evidence bears directly on the second issue we address, namely the possibility for our model to generate full-fledged credit cycles. There is growing evidence of systematic reversion in credit conditions and of subsequent output drops. Jorda, Schularick and Taylor (2012) document that strong growth of bank loans forecasts future financial crises and output drops. Baron and Xiong (2014) show that credit booms are followed by stock market declines.

When predicting credit spreads, LSZ (2015) do not try to tease out whether the cycle in credit spreads is due to fundamentals (e.g., mean reversion in the state of the economy) or to fluctuations in investor sentiment. According to the sentiment account, which they seem to favor, a period of excessive investor optimism is followed by a period of cooling off, which they refer to as "unwinding of investor sentiment". This

reversal contributes to a recession over and above the effect of changes in fundamentals. Baron and Xiong (2014) document that in good times banks expand their loans, and this expansion predicts future negative returns on bank equity. The negative returns to equity might reflect the unwinding of initial investor optimism, or might be caused by abnormally low realized performance on the bank's credit decisions (as per Proposition 4). Either way, excess optimism in good times seems needed to account for this evidence.

Diagnostic expectations yield these dynamics, and in particular can reconcile excess optimism with unwinding of investor sentiment. To see why, consider the dynamics of the investors' assessment of future economic conditions. Using Equation (4), we can compute the predictable change from $t - 1$ to t in investors' assessment of the future state of the economy:

$$\begin{aligned} & \mathbb{E}[\mathbb{E}_t^\theta(\omega_{t+1}) - \mathbb{E}_{t-1}^\theta(\omega_t)] \\ &= \mathbb{E}_{t-1}[\mathbb{E}_t(\omega_{t+1}) - \mathbb{E}_{t-1}(\omega_t)] \\ &+ \theta \mathbb{E}_{t-1}[\mathbb{E}_t(\omega_{t+1}) - \mathbb{E}_{t-1}(\omega_{t+1}) - \mathbb{E}_{t-1}(\omega_t) + \mathbb{E}_{t-2}(\omega_t)]. \end{aligned}$$

Under the assumed AR(1) process, this quantity is equal to:

$$\mathbb{E}_{t-1}[\mathbb{E}_t^\theta(\omega_{t+1}) - \mathbb{E}_{t-1}^\theta(\omega_t)] = (\bar{\omega} - \omega_{t-1})(1 - b)b - \theta[\mathbb{E}_{t-1}(\omega_t) - \mathbb{E}_{t-2}(\omega_t)]. \quad (16)$$

There are two terms in expression (16). The first term is mean reversion: conditions at t can be predicted to deteriorate if the current state is below the long run value $\bar{\omega} > \omega_{t-1}$. The second term instead captures reversals of past sentiment, which is a function of past news $[\mathbb{E}_{t-1}(\omega_t) - \mathbb{E}_{t-2}(\omega_t)]$.

Mean reversion can only generate a systematic cooling off in optimism if the current fundamental ω_{t-1} is above its long run mean $\bar{\omega}$. Diagnostic expectations, in contrast, generate predictable pessimism whenever the arrival of good news at $t - 1$ caused investors to be too optimistic to begin with, $[\mathbb{E}_{t-1}(\omega_t) - \mathbb{E}_{t-2}(\omega_t)] > 0$. Indeed,

when this is the case, investor beliefs on average revert (to rationality) next period. The intuition is simple: diagnostic expectations are too optimistic when agents' attention is caught by positive trends or good news. Because there is no systematic news going forward, any current optimism cools off on average. We view this cycle of beliefs as capturing what LSZ (2015) refer to as "unwinding of sentiment".

In a market equilibrium context, Equation (16) offers a way to think about predictable spread reversals. Such reversals can be created either by mean reverting fundamentals or by diagnostic expectations. The testable implication of the latter is that mean reversion is predictable in light of the level of fundamentals whereas unwinding of sentiment is predictable using past news, as Proposition 5 shows.

Proposition 5. *Suppose that expectations are diagnostic, $\theta > 0$, and at $t - 1$ perceived creditworthiness is too high due to good news, $\mathbb{E}_{t-1}(\omega_t) > \mathbb{E}_{t-2}(\omega_t)$. Then:*

- i) Controlling for fundamentals at $t - 1$, perceived creditworthiness predictably falls at t , namely $\mathbb{E}[\mu_t^\theta(\rho)|\omega_{t-1}] < \mu_{t-1}^\theta(\rho)$, and credit spreads predictably rise.*
- ii) Controlling for fundamentals at $t - 1$, there is a predictable drop in aggregate investment at t and in aggregate production at $t + 1$.*

Because on average investor optimism at $t - 1$ predictably reverts at t , there is, controlling for fundamentals, a predictable credit tightening at t which in turn depresses investment in the same period and thus output next period. This cycle around fundamentals is entirely due to diagnostic expectations: over-reaction to good news causes credit markets and the economy to overshoot at $t - 1$. The subsequent reversal of such over-reaction causes a drop in credit and economic activity that is more abrupt than what could be accounted for by mean reversion in fundamentals.

This reasoning implies that investor psychology can itself be a cause of volatility in credit and investment, and thus of business cycles, even in the absence of mean reversion in fundamentals. Even if the process for aggregate productivity ω_t is a random walk, namely when $b = 1$, and economy buffeted by shocks systematically experiences boom-bust episodes because investors react to news by becoming excessively optimistic or pessimistic, but then such excess pessimism or optimism mean reverts on average, in the absence of contrary news.

6. Relation to the Literature and Open Issues

6.1 Literature on Representativeness and Extrapolation

In this section we compare our model with several others in the literature that focus on the representativeness heuristic. Barberis, Shleifer and Vishny (1998) consider a decision maker who, in light of incoming news, learns whether the economy is in a mean reverting or a trending regime. Both regimes he believes are possible are wrong (the data generating process is a random walk) and predictions erroneously focus on the most likely (representative) model. After a series of good news, the agent extrapolates the trending model into the future.

Rabin and Vayanos (2010) also build a model of distorted learning. Representativeness is modeled as the agent's erroneous belief that, controlling for fundamentals, good shocks must be followed by bad shocks even though the true process for shocks is i.i.d. This model also yields extrapolative beliefs about fundamentals after a series of positive shocks.

We depart from prior work in two main ways. First, we offer a "literal" formalization of representativeness in a general judgment problem, following KT (1983). Our model can (and has been) applied to diverse domains ranging from lab

experiments on probability judgments, to social stereotyping, to market expectations. Second, previous models assume Bayesian learning about wrong models. Our approach, in contrast, involves distorted updating of true models. This mechanism creates an “unwinding of optimism” in the absence of bad news that cannot be obtained in these models of news-based learning. Furthermore, in our setting expectation formation endogenously depends on the true data process. As a consequence, drastic changes in the environment cause drastic changes in expectations, in line with the Lucas critique.

6.2 Open Issues

Our formal analysis of diagnostic expectations left three open questions: 1) what do diagnostic expectations look like if the reference $-G$ is defined in terms of diagnostic rather than rational expectations, i.e. $-G \equiv \{\Omega_t = \mathbb{E}_{t-1}^\theta(\omega_t)\}$? 2) what is the link between adaptive and diagnostic expectations?, and 3) why does our model of diagnostic expectations generate over-reaction to mixed news (e.g., positive news coming after negative news), in contrast to under-reaction in earlier models (Barberis, Shleifer and Vishny 1998, Rabin and Vayanos 2010)?

We now formally consider these questions. Our aim is not to close them, but to highlight how they might be addressed in our model, and to stress that they are related. In particular, introducing “backward looking” elements (in a sense to be specified below) in our formulation of representativeness, as per point 1), clarifies the connection of our model to adaptive expectations (point 2), and to under-reaction (point 3).

6.2.1 Lagged Diagnostic Expectations as Reference

Consider the possibility of specifying $-G$ in terms of diagnostic (rather than rational) expectation $\mathbb{E}_{t-1}^\theta(\omega_t)$. Intuitively, this implies that the DM focuses on future

outcomes along which the true distribution $h(\Omega_{t+1} = \omega_{t+1} | \Omega_t = \omega_t)$ is most different from his prior diagnostic expectation.¹² One simple way to apply this logic is to reverse-engineer the comparison group $-G$ as the current state Ω_t such that the true average future state Ω_{t+1} is equal to the lagged diagnostic expectation $\mathbb{E}_{t-1}^\theta(\omega_{t+1})$. Setting for simplicity $\bar{\omega} = 0$ in our AR(1) process, this state is $\Omega_t = \mathbb{E}_{t-1}^\theta(\omega_{t+1})/b$, so that $-G \equiv \{\Omega_t = \mathbb{E}_{t-1}^\theta(\omega_{t+1})/b\}$. Diagnostic expectations at time t are then:

$$\mathbb{E}_t^\theta(\omega_{t+1}) = \mathbb{E}_t(\omega_{t+1}) + \theta[\mathbb{E}_t(\omega_{t+1}) - \mathbb{E}_{t-1}^\theta(\omega_{t+1})]. \quad (17)$$

The agent is overly optimistic when news point to an outcome that is above his past expectations, $\mathbb{E}_t(\omega_{t+1}) > \mathbb{E}_{t-1}^\theta(\omega_{t+1})$, and overly pessimistic otherwise. In a sense, past expectations drive excess optimism or pessimism by acting as a reference point.

By iterating Equation (17) backwards, for $\theta < 1$ we obtain:

$$\mathbb{E}_t^\theta(\omega_{t+1}) = (1 + \theta) \sum_{j \geq 0} (-\theta)^j \mathbb{E}_{t-j}(\omega_{t+1}). \quad (18)$$

Expectations about the future are a weighted average of current and past rational expectations, where the weights are pinned down by the parameter θ capturing the importance of representativeness. When $\theta = 0$, expectations are rational. As θ rises, past rational expectations obtain a larger weight in shaping forecasts about the future. Critically, in Equation (18) the signs on rational expectations obtained in odd and even past periods alternate. This is an intuitive consequence of Equation (17). If one period ago the agent received good news so that his past expectation $\mathbb{E}_{t-1}^\theta(\omega_{t+1})$ is high, he is also more likely to be pessimistic today. In turn, this reduces the current

¹² This specification represents a large departure from the psychology of representativeness formalized in GS (2010). That paper views representativeness as an intuitive, effortless mechanism that selectively retrieves truly distinctive data – which shapes beliefs – from a long-term memory database which contains all potentially available data. When thinking about the Irish, we immediately think about red hair not because our working memory only contains instances of red haired Irish, but rather because our long-term memory contains the *knowledge* that red hair is more common among the Irish. The mistake, then, is not that we have false information about the Irish, but rather that we pay *too much attention* to the hair color that immediately comes to mind. If primed to pay attention to other hair colors, we adjust our estimates downward, because we know that the base rate of red hair is low.

distorted expectations $\mathbb{E}_t^\theta(\omega_{t+1})$, making it more likely that the agent is optimistic tomorrow. A certain piece of news thus has a non-monotonic effect on future expectations. The agent over-reacts on impact, but displays excess reversal in the next period, causing a further bout of optimism two periods from now.

Specifying $-G$ in terms of diagnostic (rather than rational) expectations preserves the two key properties of our basic model: expectations display excess volatility (and over-reaction) to news relative to rational expectations, and reversal of sentiment in the future, as captured by the cycle. We prefer the formulation that uses true distribution because it is both simpler and more consistent with our conceptual approach.

6.2.2 Adaptive Expectations

Define $\mathbb{E}_t^a(\omega_{t+1})$ to be the agent's current adaptive expectation of the future state. Cagan (1956) and Nerlove (1958) model adaptive expectations as a distributed lag of past realized states, namely:

$$\mathbb{E}_t^a(\omega_{t+1}) = (1 - \lambda) \sum_{j \geq 0} \lambda^j \omega_{t-j}. \quad (19)$$

Adaptive expectations are purely backward looking: the expectation for next period is independent of the future distribution. This implies not only that the agent makes systematic mistakes even in predicting deterministic cycles, but also that current news about the future leave the agent's expectations completely unaffected. As these properties seemed hard to justify, adaptive expectations have fallen in disuse.

From equation (18) one can see that when $-G$ is shaped by diagnostic expectations, our model shares similarities with adaptive expectations. This is evident

once we exploit the AR(1) process. Setting for simplicity the long run mean to zero $\bar{\omega} = 0$, given that $\mathbb{E}_{t-j}(\omega_{t+1}) = b^{j+1}\omega_{t-j}$, equation (18) becomes:

$$\mathbb{E}_t^\theta(\omega_{t+1}) = (1 + \theta)b \sum_{j \geq 0} (-\theta b)^j \omega_{t-j}. \quad (20)$$

The diagnostic expectation of the future is also a distributed lag of past realizations. There is, however, a key difference between adaptive expectations and (20). Under adaptive expectations all weights are positive, under diagnostic expectations formula some weights are negative. This occurs because under diagnostic expectations agents pay disproportionate attention to news. New information causes expectations to overshoot and subsequently revert back, generating cycles. Under adaptive expectations, in contrast, agents pay insufficient attention to news, thus under-reacting to data. As a result, the effect of news persists into the future.

To see this clearly, suppose that we allow coefficient θ to be negative, $\theta = -\lambda < 0$. Then, Equation (20) becomes:

$$\mathbb{E}_t^\theta(\omega_{t+1}) = (1 - \lambda)b \sum_{j \geq 0} (\lambda b)^j \omega_{t-j}, \quad (21)$$

which is very similar to the adaptive expectations formula in (20).¹³

In sum, our model is consistent with adaptive expectations provided individuals discount (rather than inflate) the highly representative states, and inflate the less representative states. Psychologically, this suggests that adaptive-type processes may occur in situations in which the agent emphasizes similarities, rather than differences, between different time periods. Perhaps adaptability and diagnosticity of expectations can be obtained within a unified model of attention in which, depending on the data observed, the agent either stresses differences from or similarities with the past.

¹³ The only difference is that in (21) the distributed lag coefficients depend on the persistence parameter b of the AR(1) process. If the process is i.i.d, $b = 0$, there is no extrapolation. This further underscores the fact that in our model expectations are forward looking and biases depend on the data generating process.

6.2.3 Slow Moving – G and Underreaction

We now briefly discuss slow moving specifications of the comparison group – G . Such specifications can allow our model to generate under-reaction to mixed data without introducing learning. Indeed, when – G is slow moving, an agent exposed to repeated good news remains focused on good outcomes for a while. He will then under-react to a single negative observation that is not drastic enough to draw his attention towards future downsides.

One way to allow for more rigid frames is to specify representativeness as a combination of the current and past likelihood ratio:

$$\left(\frac{h(\Omega_{t+1} = \omega_{t+1} | \Omega_t = \omega_t)}{h(\Omega_{t+1} = \omega_{t+1} | \Omega_t = \mathbb{E}_{t-1}(\omega_t))} \right)^{\alpha_1} \left(\frac{h(\Omega_{t+1} = \omega_{t+1} | \Omega_t = \mathbb{E}_{t-1}(\omega_t))}{h(\Omega_{t+1} = \omega_{t+1} | \Omega_t = \mathbb{E}_{t-2}(\omega_t))} \right)^{\alpha_2},$$

where $\alpha_1 \geq 0$ and $\alpha_2 \geq 0$ capture the weights attached to present and past representativeness, respectively. The coefficients α_i capture limited memory (past news are completely forgotten if $\alpha_1 = \alpha_2 = 0$), and $\alpha_1 > \alpha_2$ captures recency effect whereby recent news are more easily remembered. In this case we have that:

$$\begin{aligned} \mathbb{E}_t^\theta(\omega_{t+1}) = & \mathbb{E}_t(\omega_{t+1}) + \theta\alpha_1[\mathbb{E}_t(\omega_{t+1}) - \mathbb{E}_{t-1}(\omega_{t+1})] + \\ & + \theta\alpha_2[\mathbb{E}_{t-1}(\omega_{t+1}) - \mathbb{E}_{t-2}(\omega_{t+1})]. \end{aligned} \quad (22)$$

In this formulation, the agent remains overly optimistic after bad news at time t if:

$$\alpha_1[\mathbb{E}_t(\omega_{t+1}) - \mathbb{E}_{t-1}(\omega_{t+1})] + \theta\alpha_2[\mathbb{E}_{t-1}(\omega_{t+1}) - \mathbb{E}_{t-2}(\omega_{t+1})] > 0,$$

which is more likely to be true if there was good news in the past. Unless the current bad news is major, the agent inherits past optimism and under-reacts.

Along these lines, Gennaioli, Shleifer and Vishny (2015) present a model of beliefs in which the reference group is updated only every 2 periods and in which $\alpha_1 = \alpha_2 = 1$ (there is no recency effect), so that Equation (22) takes the form:

$$\mathbb{E}_t^\theta(\omega_{t+1}) = \mathbb{E}_t(\omega_{t+1}) + \theta[\mathbb{E}_t(\omega_{t+1}) - \mathbb{E}_{t-2}(\omega_{t+1})],$$

In this case, expectations in the intermediate period $\mathbb{E}_{t-1}(\omega_{t+1})$ do not shape beliefs, because the agent over-reacts in the direction of $\mathbb{E}_t(\omega_{t+1}) - \mathbb{E}_{t-2}(\omega_{t+1})$. Gennaioli et al. (2015) use this form of under-reaction to contrary news to explain why, during periods of market euphoria, early warnings are neglected until a large shock arrives.

This discussion stresses the idea that the agent's reaction to news depends on the frame or comparison group through which he interprets them.¹⁴ The structure of $-G$ can be investigated empirically by comparing model predictions with data on expectations elicited in the field or under experimental conditions. For example, by comparing the lag structures empirically estimated for an economic variable and for an agent's expectations of it, one might recover the weights attached to past data in $-G$. Under rational expectations the two estimated lag structures should be equivalent, so that the agent makes no predictable errors. According to our model, in contrast, the estimated lag structures should be different, and their differences identify the distortion parameters (i.e. $\theta, \alpha_1, \alpha_2$). This is an important avenue for future work.

7. Conclusion

We have presented a new approach to modeling beliefs in economic models, diagnostic expectations, based on Kahneman and Tversky's representativeness heuristic. Our model of expectations is portable in Rabin's sense, meaning that the same framework accounts for many experimental findings, the phenomenon of stereotyping, but also critical features of beliefs in financial markets such as over-reaction to data. Diagnostic expectations are also forward-looking and responsive to new data, which means that they are invulnerable to the Lucas critique of mechanical backward looking

¹⁴ As stressed by Kahneman (2003), this is a key feature of human perception "Perception is reference-dependent: the perceived attributes of a focal stimulus reflect the contrast between that stimulus and a context of prior and concurrent stimuli."

models of beliefs. We applied diagnostic expectations to a straightforward macroeconomic model of investment, and found that it can account for several empirical findings regarding credit cycles without resort to financial frictions.

Two aspects of our research most obviously require further investigation. First, we have assumed away financial frictions. Indeed, in our model debt is indistinguishable from equity, in that there are no costs of financial distress and no differential legal rights of alternative financial claims, and in particular no collateral constraints. Furthermore, investors are risk neutral, so that debt does not have a special role in meeting the needs of risk averse investors (see Gennaioli, Shleifer, and Vishny 2012). The absence of financial frictions and of risk averse investors leads to completely symmetric effects of positive and negative news. Introducing a more realistic conception of debt might be extremely useful, particularly in the context of analyzing financial crises. In particular, diagnostic expectations may interact with collateral constraints to give rise to additional consequences of tightening credit.

When the economy is hit by a series of good news, investors holding diagnostic expectations become excessively optimistic, fueling as in the current model excessive credit expansion. During such a credit expansion households would pay insufficient attention to the possibility of a bust. As fundamentals stabilize, the initial excess optimism unwinds, bringing this possibility to investors' minds. The economy would appear to be hit by a "financial shock": a sudden, seemingly unjustified, increase in credit spreads. Agents would appear to have magically become more risk averse: they now take into account the crash risk they previously neglected.

In the presence of financial frictions, the economy will not go back to its normal course. When excessive leverage is revealed, debt investors try to shed the excessive risk they have taken on, depressing debt prices and market liquidity, particularly if they

are very risk averse as in Gennaioli, Shleifer and Vishny (2012, 2015). The tightening of debt constraints causes fire-sales and corporate investment cuts, leaving good investment opportunities unfunded. Such a crisis does not occur because of deteriorating fundamentals, but because the initial excess optimism burst. Years of bonanza plant the seeds for a financial crisis. A combination of diagnostic expectations and financial frictions could thus lead to models of financial crises that match both the expectations data and the reality of severe economic contractions.

The second set of open questions relates to the diagnostic expectations themselves. On the one hand, the formulation of diagnostic expectations relies on context $-G$, which comes from memory and is therefore fundamentally unobservable. Since there is no reason to believe that $-G$ is the same across people or situations, understanding what $-G$ is in different contexts remains a big open question. We suggest, in conclusion, that our model puts a lot of structure on this question.

To begin, expectations data are themselves observable, so one way to get at $-G$ is to match model predictions of expectations with expectations data. Informally, this is what we have done in this paper, arguing that rational expectations model do not match expectations data, which strongly point to extrapolation.

Perhaps as important an advantage of our approach is that expectations are not delinked from news, but rather follow a distorted true process of the data, what we have referred to as the “kernel of truth” hypothesis. As a consequence, the model imposes tight restrictions on the processes governing actual economic variables and their expectations, and is therefore testable. Such tests, as well, may enable us to recover $-G$, the background context for the formation of diagnostic expectations.

References:

- Arnold, Marc, Dustin Schuette, and Alexander Wagner. 2015. "Pay Attention or Pay Extra: Evidence on the Compensation of Investors for the Implicit Credit Risk of Structured Products." Swiss Finance Institute Research Paper No. 14-24.
- Barberis, Nicholas, Andrei Shleifer, and Robert Vishny. 1998. "A Model of Investor Sentiment." *Journal of Financial Economics* 49(3): 307-343
- Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer. 2015a. "X-CAPM: an Extrapolative Capital Asset Pricing Model." *Journal of Financial Economics* 115 (1): 1 - 24.
- Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer. 2015b. "Extrapolation and Bubbles." Yale University Mimeo.
- Baron, Matthew, and Wei Xiong. 2014. "Credit Expansion and Neglected Crash Risk." Working Paper, Princeton University.
- Bernanke, Ben and Mark Gertler. 1989. "Agency Costs, Net Worth, and Business Fluctuations." *American Economic Review* 79(1): 14-31.
- Bernanke, Ben. 1990. "On the Predictive Power of Interest Rates and Interest Rate Spreads." *New England Economic Review* (Nov): 51-68.
- Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer. 2014. "Stereotypes." NBER Working Paper 20106.
- Brunnermeier, Markus and Yuliy Sannikov. 2014. "A Macroeconomic Model with a Financial Sector." *American Economic Review* 104(2): 379 - 421.
- Cagan, Philip. 1956. "The Monetary Dynamics of Hyperinflation," in M. Friedman (ed.), *Studies in the Quantity Theory of Money*. Chicago: University of Chicago Press.
- Casscells, Ward, Arno Schoenberger, and Thomas Graboys. 1978. "Interpretation by Physicians of Clinical Laboratory Results." *New England Journal of Medicine* 299:999-1001.
- Collin-Dufresne, Pierre, Robert Goldstein and Spencer Martin. 2001. "The Determinants of Credit Spread Changes." *Journal of Finance* 56(6): 2177-2207.
- Coval, Joshua, Kevin Pan, and Erik Stafford. 2014. "How Markets Learn." Harvard Business School Mimeo.

Cutler, David, James Poterba, and Lawrence Summers. 1990. "Speculative Dynamics and the Role of Feedback Traders." *American Economic Review Papers & Proceedings* 80(2): 62-68.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam. 1998. "Investor Psychology and Security Market Under- and Overreactions." *Journal of Finance* 53(6): 1839 – 1885.

DeLong, Bradford, Andrei Shleifer, Lawrence Summers, and Robert Waldmann. 1990. "Positive Feedback Investment Strategies and Destabilizing Rational Speculation." *Journal of Finance* 45(2): 379-395.

Eggertsson, Gauti and Paul Krugman. 2012. "Debt, Deleveraging, and the Liquidity Trap: A Fisher-Minsky-Koo approach." *Quarterly Journal of Economics* 127(3): 1469-1513.

Farhi, Emmanuel and Ivan Werning. 2015. "A Theory of Macroprudential Policies in the Presence of Nominal Rigidities." Working Paper.

Friedman, Benjamin, and Kenneth Kuttner. 1992. "Money, Income, Prices, and Interest Rates." *American Economic Review* 82(3): 472-492.

Fuster, Andreas, David Laibson, and Brock Mendel. 2010. "Natural Expectations and Macroeconomic Fluctuations." *Journal of Economic Perspectives* 24(4): 67 - 84.

Gennaioli, Nicola, Andrei Shleifer, and Robert Vishny. 2012. "Neglected Risks, Financial Innovation, and Financial Fragility." *Journal of Financial Economics* 104(3): 452-468.

Gennaioli, Nicola, Andrei Shleifer, and Robert Vishny. 2013. "A Model of Shadow Banking." *Journal of Finance* 68 (4): 1331-1363.

Gennaioli, Nicola, Andrei Shleifer, and Robert Vishny. 2015. "Neglected Risks: The Psychology of Financial Crises." *American Economic Review, Papers and Proceedings* 105 (5): 310-314.

Gennaioli, Nicola, Yueran Ma, and Andrei Shleifer. 2015. "Expectations and Investment." Forthcoming *NBER Macroeconomics Annual*.

Giesecke, Kay, Francis Longstaff, Stephen Schaefer, Ilya Strebulaev. 2011. "Corporate Bond Default Risk: A 150-year Perspective" *Journal of Financial Economics* 102 (2): 233-250.

Gilchrist, Simon and Egon Zakrajšek. 2012. "Credit Spreads and Business Cycle Fluctuations." *American Economic Review* 102(4): 1692-1720.

Greenwood, Robin and Samuel Hanson. 2013. "Issuer Quality and Corporate Bond

- Returns." *Review of Financial Studies* 26(6), 1483–1525.
- Greenwood, Robin, and Samuel Hanson. 2015. "Waves in Ship Prices and Investment." *Quarterly Journal of Economics* 130(1): 55-109.
- Greenwood, Robin, and Andrei Shleifer. 2014. "Expectations of Returns and Expected Returns." *Review of Financial Studies* 27(3): 714-746.
- Guerrieri, Veronica and Guido Lorenzoni. 2015. "Credit Crises, Precautionary Savings and the Liquidity Trap." Working Paper.
- Hilton, James, and William von Hippel. 1996. "Stereotypes." *Annual Review of Psychology* 47 (1): 237 – 271.
- Hirshleifer, David, Jun Li, and Jianfeng Yu. 2015. "Asset Pricing in Production Economies with Extrapolative Expectations." *Journal of Monetary Economics*, forthcoming.
- Jorda, Oscar, Moritz Schularick, and Alan Taylor. 2013. "When Credit Bites Back: Leverage, Business Cycles, and Crises." *Journal of Money, Credit and Banking* 45(s2): 3-28.
- Judd, Charles, and Bernardette Park. 1993. "Definition and Assessment of Accuracy in Social Stereotypes." *Psychological Review* 100: 109 – 128.
- Kahneman, Daniel, and Amos Tversky. 1972. "Subjective Probability: A Judgment of Representativeness." *Cognitive Psychology* 3 (3): 430 – 454.
- Kahneman, Daniel. 2003. "A Perspective on Judgment and Choice: Mapping Bounded Rationality." *American Psychologist* 58 (9): 697 – 720.
- Kahneman, Daniel. 2011. *Thinking: Fast and Slow*. New York: Farrar, Straus, and Giroux.
- Kiyotaki, Nobuhiro, and John Moore. 1997. "Credit Cycles." *Journal of Political Economy* 105(2): 211-248.
- Korinek, Anton, and Alp Simsek. 2014. "Liquidity trap and excessive leverage." Working Paper, MIT.
- Krishnamurthy, Arvind, and Tyler Muir. 2015. "Credit Spreads and the Severity of Financial Crises." Working Paper, Stanford University.

Lopez-Salido, David, Jeremy Stein, and Egon Zakrajsek. 2015. "Credit-Market Sentiment and the Business Cycle." Working Paper, Harvard University.

Lorenzoni, Guido. 2008. "Inefficient Credit Booms." *Review of Economic Studies* 75(3): 809-833.

Mian, Atif, Amir Sufi, and Emil Verner. 2015. "Household Debt and Business Cycle Worldwide." Working Paper.

Minsky, Hyman. 1977. "The Financial Instability Hypothesis: An interpretation of Keynes and an Alternative to 'Standard' Theory." *Nebraska Journal of Economics and Business* 16(1): 5-16.

Muth, John. 1961. "Rational Expectations and the Theory of Price Movements." *Econometrica*, 29 (3): 315 – 335.

Nerlove, Marc. 1958. "Adaptive Expectations and Cobweb Phenomena." *Quarterly Journal of Economics* 72(2): 227 – 240.

Rabin, Matthew. 2013. "Incorporating Limited Rationality into Economics." *Journal of Economic Literature*, 51(2): 528-43.

Rabin, Matthew, and Dimitri Vayanos. 2010. "The Gambler's and Hot Hand Fallacies: Theory and Applications." *Review of Economic Studies* 77(2): 730 – 778.

Rognlie, Matthew, Andrei Shleifer, and Alp Simsek. 2015. "Investment Hangover and the Great Recession." Working Paper, Harvard University and MIT.

Schularick, Moritz and Alan Taylor. 2012. "Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870–2008." *American Economic Review* 102(2): 1029-1061.

Stein, Jeremy. 2014. "Incorporating Financial Stability Considerations into a Monetary Policy Framework." Speech at the International Research Forum on Monetary Policy, Federal Reserve Board of Governors, Washington, D.C., March 21, 2014.

Stock, James, and Mark Watson. 2003. "Forecasting Output and Inflation: The Role of Asset Prices." *Journal of Economic Literature* 41: 788–829.

Tversky, Amos, and Daniel Kahneman. 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science* 27 (185): 1124 – 1131.

Tversky, Amos, and Daniel Kahneman. 1983. "Extensional versus Intuitive Reasoning: the Conjunction Fallacy in Probability Judgment." *Psychological Review* 90 (4): 293-315.

Proofs

Proposition 1. Let ω_t be an AR(1) process, $\omega_{t+1} = a + b\omega_t + \varepsilon_t$, with i.i.d. normal $(0, \sigma^2)$ shocks ε_t . At t , the true distribution of ω_{t+1} is

$$h_t(\omega_{t+1}) = h(\Omega_{t+1} = \omega_{t+1} | \Omega_t = \omega_t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\omega_{t+1} - \mathbb{E}_t(\omega_{t+1}))^2}{2\sigma^2}}$$

where $\mathbb{E}_t(\omega_{t+1}) = a + b\omega_t$. The comparison distribution is instead

$$h_{t-1}(\omega_{t+1}) = h(\Omega_{t+1} = \omega_{t+1} | \Omega_t = \mathbb{E}_{t-1}(\omega_t)) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\omega_{t+1} - \mathbb{E}_{t-1}(\omega_{t+1}))^2}{2\sigma^2}}$$

where we used the law of iterated expectations, $\mathbb{E}_{t-1}(\mathbb{E}_t(\omega_{t+1})) = \mathbb{E}_{t-1}(\omega_{t+1})$. From Equation (*), the diagnostic distribution is then:

$$\begin{aligned} h_t^\theta(\omega_{t+1}) &= \frac{1}{Z} e^{-\frac{1}{2\sigma^2}[(1+\theta)(\omega_{t+1} - \mathbb{E}_t(\omega_{t+1}))^2 - \theta(\omega_{t+1} - \mathbb{E}_{t-1}(\omega_{t+1}))^2]} \\ &= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}[(\omega_{t+1} - \mathbb{E}_t^\theta(\omega_{t+1}))^2]} \end{aligned}$$

where Z is a normalizing constant, and:

$$\mathbb{E}_t^\theta(\omega_{t+1}) = \mathbb{E}_t(\omega_{t+1}) + \theta[\mathbb{E}_t(\omega_{t+1}) - \mathbb{E}_{t-1}(\omega_{t+1})].$$

■

Proposition 2. The variance of future expectations $\mathbb{E}_t^\theta(\omega_{t+1})$, computed in period $t - 1$, is given by

$$\text{Var}_{t-1}[\mathbb{E}_t^\theta(\omega_{t+1})] = \text{Var}_{t-1}[(1 + \theta)\mathbb{E}_t(\omega_{t+1}) - \mathbb{E}_{t-1}(\omega_{t+1})],$$

where the variance is computed over all possible realizations of ω_t . The second term is constant and its variance is zero. The expression thus becomes:

$$\text{Var}_{t-1}[\mathbb{E}_t^\theta(\omega_{t+1})] = (1 + \theta)^2 \text{Var}_{t-1}[\mathbb{E}_t(\omega_{t+1})].$$

The expected forecast error at t is given by $\mathbb{E}_t[\omega_{t+1} - \mathbb{E}_t^\theta(\omega_{t+1})]$. Replacing $\mathbb{E}_t^\theta(\omega_{t+1})$ with Equation (4), we find

$$\mathbb{E}_t[\omega_{t+1} - \mathbb{E}_t^\theta(\omega_{t+1})] = -\theta[\mathbb{E}_t(\omega_{t+1}) - \mathbb{E}_{t-1}(\omega_{t+1})].$$

■

Proposition 3. We start by proving Equation (15). We have:

$$\frac{\partial \ln \mu_t^\theta(\rho)}{\partial \omega_t} = \frac{1}{\mu_t^\theta(\rho)} \frac{\partial \mu_t^\theta(\rho)}{\partial \mathbb{E}_t^\theta(\omega_{t+1})} \cdot \frac{\partial \mathbb{E}_t^\theta(\omega_{t+1})}{\partial \omega_t}$$

The first term reads

$$\begin{aligned} \frac{1}{\mu_t^\theta(\rho)} \frac{\partial \mu_t^\theta(\rho)}{\partial \mathbb{E}_t^\theta(\omega_{t+1})} &= \frac{1}{\sigma^2} \int_{\rho}^{+\infty} \omega e^{-\frac{1}{2\sigma^2}[(\omega - \mathbb{E}_t^\theta(\omega_{t+1}))^2]} \frac{d\omega}{\mu_t^\theta(\rho)} - \frac{\mathbb{E}_t^\theta(\omega_{t+1})}{\sigma^2} \\ &= [\mathbb{E}_t^\theta(\omega_{t+1} | \omega_{t+1} \geq \rho) - \mathbb{E}_t^\theta(\omega_{t+1})] \frac{1}{\sigma^2} \end{aligned}$$

This term is strictly positive for $\rho > -\infty$. Moreover, given Equation (4), the second term reads $\frac{\partial \mathbb{E}_t^\theta(\omega_{t+1})}{\partial \omega_t} = b(1 + \theta) > 0$. In particular, we have $\frac{\partial \ln \mu_t^\theta(\rho)}{\partial \omega_t} > 0$.

Consider now the proposition's point i). From the definition (14) of credit spreads, we have

$$\frac{\partial S_t^\theta(\rho)}{\partial \omega_t} = -\frac{1}{\beta \sigma^2 \mu_t^\theta(\rho)} \frac{\partial \ln \mu_t^\theta(\rho)}{\partial \omega_t}$$

which is negative, i.e. spreads drop as conditions ω_t improve. Moreover, spreads also become compressed. Formally,

$$\frac{\partial^2 S_t^\theta(\rho)}{\partial \omega_t \partial \rho} \propto -\frac{\partial}{\partial \rho} [\mathbb{E}_t^\theta(\omega_{t+1} | \omega_{t+1} \geq \rho) - \mathbb{E}_t^\theta(\omega_{t+1})] < 0$$

Turning to the proposition's point ii), recall from Equation (11) that equilibrium debt issuance for firm is $k_{t+1}(\rho) = [\alpha \beta \mu_t^\theta(\rho)]^{\frac{1}{1-\alpha}}$, while total debt issuance is given by

$K_{t+1} = \int_{-\infty}^{+\infty} [\alpha \beta \mu_t^\theta(\rho)]^{\frac{1}{1-\alpha}} f(\rho) d\rho$. Thus, firms' debt grows as:

$$\frac{\partial k_{t+1}(\rho)}{\partial \omega_t} = \left(\frac{1}{1-\alpha} \right) \frac{\partial \ln \mu_t^\theta(\rho)}{\partial \omega_t} k_{t+1}(\rho)$$

implying that total investment grows:

$$\frac{\partial K_{t+1}}{\partial \omega_t} = \left(\frac{1}{1-\alpha} \right) \int_{-\infty}^{+\infty} \frac{\partial \ln \mu_t^\theta(\rho)}{\partial \omega_t} k_{t+1}(\rho) f(\rho) d\rho > 0.$$

and that it grows disproportionately for riskier firms:

$$\frac{\partial k_{t+1}(\rho_1)}{\partial \omega_t k_{t+1}(\rho_2)} = \left(\frac{1}{1-\alpha} \right) \left[\frac{\partial \ln \mu_t^\theta(\rho_1)}{\partial \omega_t} - \frac{\partial \ln \mu_t^\theta(\rho_2)}{\partial \omega_t} \right] > 0 \text{ for all } \rho_1 > \rho_2.$$

■

Proposition 4. Suppose that no information arrived at time $t-1$, $\mathbb{E}_{t-1}(\omega_t) = \mathbb{E}_{t-2}(\omega_t)$, and that new information $\mathbb{E}_t(\omega_{t+1}) - \mathbb{E}_{t-1}(\omega_{t+1}) \neq 0$ arrives at time t . The impact on credit spreads depends on the degree θ of stereotypical thinking, which modulates how investors' expectations respond to information. We have:

$$\frac{\partial S_t^\theta(\rho)}{\partial \theta} = \frac{\partial S_t^\theta(\rho)}{\partial \mathbb{E}_t^\theta(\omega_{t+1})} \cdot \frac{\partial \mathbb{E}_t^\theta(\omega_{t+1})}{\partial \theta} = \frac{\partial S_t^\theta(\rho)}{\partial \mathbb{E}_t^\theta(\omega_{t+1})} \cdot [\mathbb{E}_t(\omega_{t+1}) - \mathbb{E}_{t-1}(\omega_{t+1})]$$

Using $\frac{\partial S_t^\theta(\rho)}{\partial \omega_t} = \frac{\partial S_t^\theta(\rho)}{\partial \mathbb{E}_t^\theta(\omega_{t+1})} \frac{\partial \mathbb{E}_t^\theta(\omega_{t+1})}{\partial \omega_t} = \frac{\partial S_t^\theta(\rho)}{\partial \mathbb{E}_t^\theta(\omega_{t+1})} \cdot b(1+\theta)$, we obtain the result. Therefore,

$\frac{\partial S_t^\theta(\rho)}{\partial \theta}$ has the same sign as $\frac{\partial S_t^\theta(\rho)}{\partial \omega_t}$ (and opposite sign as the information), so credit spreads overreact to information.

Turning to volatility, recall that expanding the variance of a function $f(X)$ of a random variable X in a Taylor series around we get $\text{Var}[f(X)] \approx \left(f'(\mathbb{E}(X)) \right)^2 \text{Var}[X]$, provided f is twice differentiable and $\mathbb{E}(X)$, $\text{Var}[X]$ are finite. Therefore,

$$\text{Var}_{t-1}[S_t^\theta(\rho)] \approx \left(\frac{\partial S_t^\theta(\rho)}{\partial \omega_t} \left(\mathbb{E}_{t-1} \left(\mathbb{E}_t^\theta(\omega_{t+1}) \right) \right) \right)^2 \text{Var}_{t-1}[\mathbb{E}_t^\theta(\omega_{t+1})]$$

and similarly $\text{Var}_{t-1}[S_t^0(\rho)] \approx \left(\frac{\partial S_t^0(\rho)}{\partial \omega_t} \left(\mathbb{E}_{t-1}(\mathbb{E}_t(\omega_{t+1})) \right) \right)^2 \text{Var}_{t-1}[\mathbb{E}_t(\omega_{t+1})]$, where $S_t^0(\rho)$ denotes the spread in the rational expectations benchmark ($\theta = 0$). Since no information arrived at $t - 1$, we have $\mathbb{E}_{t-1}(\mathbb{E}_t(\omega_{t+1})) = \mathbb{E}_{t-1}(\mathbb{E}_t^\theta(\omega_{t+1}))$, and so the squared-terms in the two expressions are equal. Thus, to second order approximation in θ , Proposition 2 implies:

$$\frac{\text{Var}_{t-1}[S_t^\theta(\rho)]}{\text{Var}_{t-1}[S_t^0(\rho)]} \approx \frac{\text{Var}_{t-1}[\mathbb{E}_t^\theta(\omega_{t+1})]}{\text{Var}_{t-1}[\mathbb{E}_t(\omega_{t+1})]} = (1 + \theta)^2$$

Credit spreads display excess volatility, for any firm type $\rho > 0$. Likewise, any function of expectations, such as aggregate investment, also displays excess volatility.

We have seen above that positive information at time t compresses credit spreads $S_t(\rho)$ at time t . Because no information arrived at time $t - 1$, good news at t cause spreads to compress too much, since by Equation (15) we have $\mu_t^\theta(\rho) > \mu_t(\rho)$. Average debt returns at time $t + 1$ are then given by $\frac{\mu_t(\rho)}{\mu_t^\theta(\rho)} \beta^{-1}$ which are abnormally low (below β^{-1}). With negative information, the reverse happens.

■

Corollary 1. Consider an adverse shock $\Delta \mathbb{E}_t^\theta(\omega_{t+1}) < 0$ to investor confidence at t . This causes a drop in perceived creditworthiness $\mu_t^\theta(\rho)$, and investment $k_{t+1}(\rho)$, for all firms $\rho > 0$. Expected total output at $t + 1$ is given by:

$$\mathbb{E}_t[Y_{t+1}(\omega_{t+1})] = \mathbb{E}_t \left[\int_{-\infty}^{\omega_{t+1}} [\alpha \beta \mu_t^\theta(\rho)]^{1-\alpha} f(\rho) d\rho \right].$$

A drop in confidence about ω_{t+1} thus translates into $\mathbb{E}_t \left[\frac{\partial Y_{t+1}(\omega_{t+1})}{\partial \mathbb{E}_t^\theta(\omega_{t+1})} \right] \Delta \mathbb{E}_t^\theta(\omega_{t+1})$.

Differentiating the expression above yields:

$$\mathbb{E} \left[\frac{1}{(1-\alpha)\beta} \cdot \int_{-\infty}^{\omega_{t+1}} [\alpha\beta\mu_t^\theta(\rho)]^{\frac{\alpha}{1-\alpha}} \frac{\partial \ln \mu_t^\theta(\rho)}{\partial \mathbb{E}_t^\theta(\omega_{t+1})} f(\rho) d\rho \mid \omega_t \right] \Delta \mathbb{E}_t^\theta(\omega_{t+1}) < 0$$

■

Proposition 5. We first derive Equation (16). We expand the identity

$$\begin{aligned} & \mathbb{E}_{t-1} [\mathbb{E}_t^\theta(\omega_{t+1}) - \mathbb{E}_{t-1}^\theta(\omega_t)] \\ &= \mathbb{E}_{t-1} [\mathbb{E}_t(\omega_{t+1}) - \mathbb{E}_{t-1}(\omega_t)] \\ &+ \theta \mathbb{E}_{t-1} [\mathbb{E}_t(\omega_{t+1}) - \mathbb{E}_{t-1}(\omega_{t+1}) - \mathbb{E}_{t-1}(\omega_t) + \mathbb{E}_{t-2}(\omega_t)]. \end{aligned}$$

under the AR(1) process $\omega_t - \bar{\omega} = b(\omega_t - \bar{\omega}) + \varepsilon_t$. We have:

$$\mathbb{E}_{t-1} [\mathbb{E}_t(\omega_{t+1})] = \mathbb{E}_{t-1} [\mathbb{E}_{t-1}(\omega_{t+1})] = b\mathbb{E}_{t-1}(\omega_t) + \bar{\omega}(1-b)$$

where the first equality follows from the law of iterated expectations. Putting it all together, we find

$$\mathbb{E}_{t-1} [\mathbb{E}_t^\theta(\omega_{t+1}) - \mathbb{E}_{t-1}^\theta(\omega_t)] = (\bar{\omega} - \mathbb{E}_{t-1}(\omega_t))(1-b)b - \theta[\mathbb{E}_{t-1}(\omega_t) - \mathbb{E}_{t-2}(\omega_t)].$$

By assumption of the Proposition, $\mathbb{E}_{t-1}(\omega_t) > \mathbb{E}_{t-2}(\omega_t)$, so the second term on the right is negative. Thus, controlling for fundamentals (the first term), expectations about the future of the economy predictably (i.e. on average) fall at time t , as a consequence of good news at time $t - 1$.

Because creditworthiness $\mu_t^\theta(\rho)$ is a strictly increasing function of expectations $\mathbb{E}_t^\theta(\omega_{t+1})$, for every firm $\rho > 0$, it follows that (controlling for fundamentals), creditworthiness predictably falls at t after good news at $t - 1$ (point i). It then immediately follows from the definitions that spreads rise, and investment falls, on average at time t , and output predictably falls on average at time $t + 1$.

■