University of Pennsylvania
ScholarlyCommons

Operations, Information and Decisions Papers

Wharton Faculty Research

2-16-2003

Understanding Under - and Over - Reaction

Cade Massey
University of Pennsylvania

George Wu

Follow this and additional works at: http://repository.upenn.edu/oid_papers

Part of the Other Economics Commons, and the Other Social and Behavioral Sciences Commons

Understanding Under - and Over - Reaction

Disciplines

Other Economics | Other Social and Behavioral Sciences

Understanding under- and over-reaction

Cade Massey
George Wu

February 16, 2003

Fuqua School of Business
Duke University
Durham, NC 27708
cade.massey@duke.edu

Graduate School of Business
University of Chicago
Chicago, IL 60637
george.wu@gsb.uchicago.edu

I. INTRODUCTION

The ability to detect change accurately is vital for economic as well as social success. Consider four very different examples. A baseball manager must decide whether his pitcher has "lost his stuff". A doctor must determine whether a patient's health has taken a turn for the worst. A production manager is unsure of whether a manufacturing line is out of control. A central bank needs to determine whether an economy has slipped into recession. In each case, the state of the world may have changed from one "regime" to another. These examples illustrate some of the difficulties associated with detecting change. First, the "signal" in each instance (an errant pitch, a sudden drop in blood pressure, or a sharp increase in defective goods, a precipitous drop in the stock market) is noisy, and thus there is a need to separate the "signal" from the "noise". Second, a decision maker must balance between making one of two mistakes. She can *over-react* (*i.e.*, act as if the world has changed, when it in fact has not) or *under-react* (*i.e.*, act as if the world has remained the same, when it in fact has changed).

The general problem of detecting change has received considerable recent attention in both the academic and popular press. Financial economics has documented both over- and under-reaction in financial markets to earnings announcements and other news (De Bondt & Thaler, 1985; Brav & Heaton, 2002). Even more recently, the run-up in stock prices in the late 1990's led to considerable debate among pundits over whether the historically high stock prices were a bubble and a short-term anomaly, or whether the market valuations were warranted because of the advent of a "new economy" (Browning, 1998; Gasparino, 1998). Finally, the management literature has emphasized how critical it is for managers to be able to detect the onset of a new regime (Grove, 1999).

This research takes a step toward understanding how good individuals are in detecting regime shifts. In particular, we are motivated by the following questions: Do individuals over-react or under-react to indications of change? When do individuals over-react, and when do they under-react? What psychological processes explain the pattern of over- and under-reaction? Two experimental studies provide some answers to these questions. Individuals both over-react and under-react. However, the pattern of over-reaction and under-reaction is systematic, not random. Under-reaction is most common in unstable environments in which signals are precise, and over-reaction is most likely to occur in stable environments in which signals are noisy. The psychological story we use to explain these results, the *system neglect hypothesis*, posits that individuals primarily react to signals of change, and secondarily to the system that generated the signal.

Our chapter is organized as follows. We begin by describing our experimental setup. This permits us to establish some terminology, and also helps us develop and explain our psychological hypothesis of system neglect. We then show how the system neglect hypothesis predicts over-reaction in some environments, and under-reaction in other environments. Next, we present the results of two studies. The two studies involve two very different tasks, judgment and choice, but both show the predicted pattern of over- and under-reaction. We conclude by reviewing the questions raised in the introduction, and posing some new questions.

Detailed descriptions of the studies, as well as much more extensive analyses, are found in Massey & Wu (2002).

EXPERIMENTAL SETUP

An Informal Description

We begin by describing the setup used in our experimental studies. Subjects in our experiments observe signals and use these signals to infer whether there has been a shift from one regime to a second regime. Figure 1 depicts our basic experimental setup.
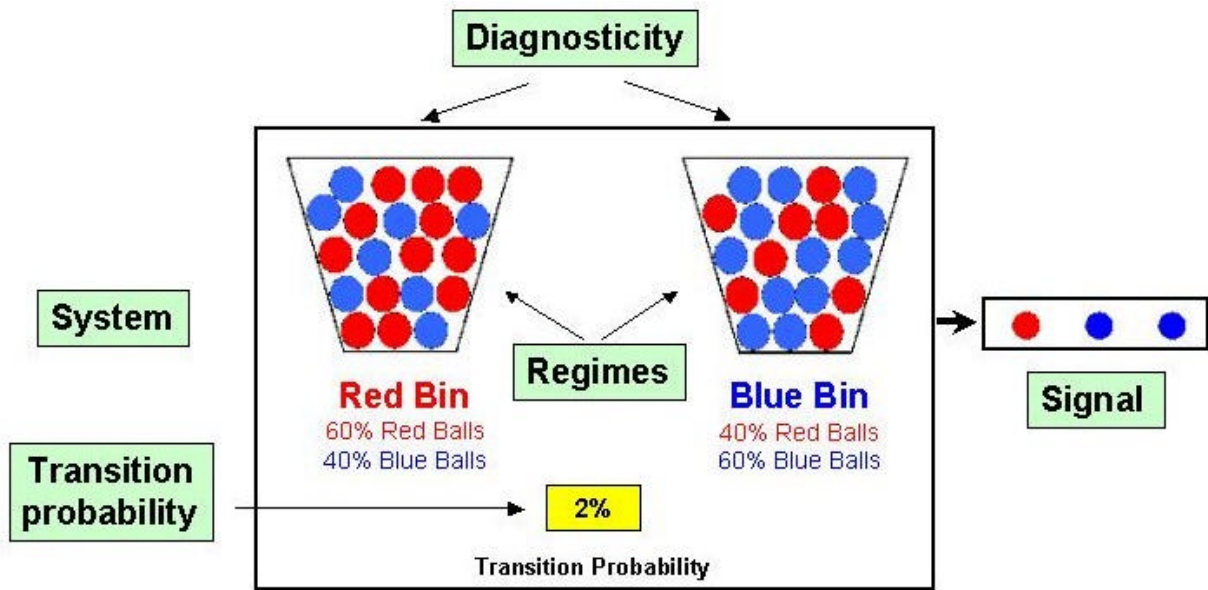

Figure 1: Experimental Setup (Low Diagnosticity, Low Transition Probability System)

We call everything inside the box in Figure 1 a *system*. A system has several important characteristics. First, a system involves two *regimes*, an *incumbent regime* (the "red bin") and a *transition regime* (the "blue bin"). The process begins with the red bin but may switch to the blue bin at any period. The blue bin is an absorbing state: if there is a switch to the blue bin at any time, the process will continue with that regime.

Subjects observe *signals* and must use these signal to infer whether there has been a shift from the red bin to the blue bin. In the example in Figure 1, a subject has observed a signal consisting of a red ball drawn in the first period and blue balls drawn in the second and third

4

periods. In judging whether the process has switched to the blue bin, two aspects of the system are critical. A subject must consider the *diagnosticity* of the signal. In Figure 1, the red bin does not differ that much from the blue bin: the red bin consists of 60% red balls and 40% red balls, while the blue bin has 60% blue balls and 40% red balls. Thus, the signals are relatively uninformative. Contrast this to the system in Figure 2 in which the red bin has 90% red balls, and the blue bin has 90% blue balls. Thus, the signal in Figure 2 is much more diagnostic than the signal in Figure 1.
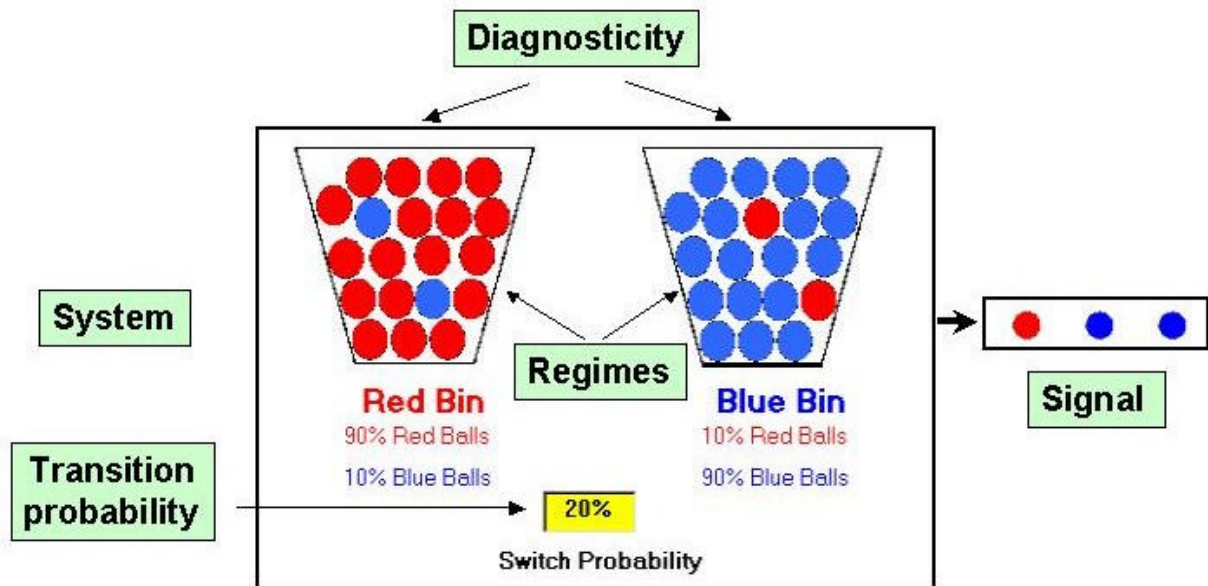

Figure 2: Experimental Setup (High Diagnosticity, High Transition Probability System)

Finally, the subjects must consider the stability of the system, as measured by the *transition probability*, or chance the system will switch from the red bin to the blue bin in any period. The system in Figure 1 is a very stable system, with a transition probability of .02. In contrast, the system in Figure 2 is considerably less stable with a transition probability of .20.

To review, a system consists of two regimes, a red bin and a blue bin, and two system parameters, diagnosticity (the relative composition of red balls in the red bin to red balls in the

blue bin), and transition probability. It is important to note that whereas the system *generates* a

signal, the signal is outside the system.

Although our experimental paradigm necessarily simplifies most real world change

processes, we can nevertheless map our introductory examples to this conceptualization.

Consider the manufacturing process example. A process begins "in control" (the incumbent

regime) but may drift and become "out of control" (the transition regime) (Shewhart & Deming,

1986). The diagnosticity of a signal, such as the defect rate, is captured by variability of the

process, while the transition probability is captured by the historical "hazard rate".

A Formal Description

More formally, we let R_t indicate that the process is in the red bin in period t, and B_t

indicate that the process is in the blue bin in period t. Each bin produces one of two possible

signals: a red ball ($r_t = 1$) or a blue ball ($r_t = 0$), where r_t is an indicator variable for a red signal

in period t. Let $p_R = \Pr(r_t = 1 | R_t)$ and $p_B = \Pr(r_t = 1 | B_t)$ denote the proportion of red balls in

the red and blue bins, respectively. In our experiments, the red bin and blue bin are symmetric:

$p_R = 1 - p_B$.   Then $d = p_R / p_B$ measures the diagnosticity of the signal. We denote

$q = \Pr(B_{t+1} | R_t)$ the transition probability, or probability that the process changes from the red bin

to the blue bin at any given period. In addition, the process begins with the red bin at $t = 0$,

$\Pr(R_0) = 1$, and the blue bin is an absorbing state, $\Pr(B_{t+1} | B_t) = 1$. Finally, denote the history of

signals from period 1 to t by $H_t = (r_1, ..., r_t)$. In Figure 1, $p_R = .6$, $p_B = .4$, $d = 1.5$, and $q = .02$,

while $p_R = .9$, $p_B = .1$, $d = 9$, and $q = .20$ in Figure 2.

It is straightforward to derive the Bayesian posterior odds given H_t:

$$\frac{\Pr(B_t \mid H_t)}{\Pr(R_t \mid H_t)} = \left(\frac{1-(1-q)^t}{(1-q)^t} \right) \sum_{j=1}^{t} \frac{q(1-q)^{j-1}}{1-(1-q)^t} d^{t+1-j-\left(2\sum_{k=j}^{t} r_k\right)}. \tag{1}$$

The derivation is found in Appendix II of Massey & Wu (2002). In our experiments, we compare empirical behavior to this normative expression.

THE SYSTEM NEGLECT HYPOTHESIS

Consider the signal in Figure 1: a red ball in period 1, and blue balls in periods 2 and 3. Clearly, the signal is suggestive of a change to the blue bin. However, a Bayesian agent must also consider the diagnosticity of the signal and the stability of the system (*i.e.*, the transition probability). The system neglect hypothesis suggests that individuals will attend primarily to the signal, and secondarily to the system that generated the signal.

The system neglect hypothesis draws on and extends research in static judgment. Edwards' (1968) research on conservatism found that subjects updated their judgments too slowly when presented with samples of evidence. Kahneman & Tversky's (1973) work on representativeness showed the opposite: that individuals were too willing to extrapolate from small samples. Griffin & Tversky (1992) resolved this apparent paradox by noting that research in the conservatism literature typically involved large samples, while representativeness research involved small samples. More generally, Griffin & Tversky proposed that individuals attend first to the "strength of evidence" and then secondarily to the "weight of evidence". Roughly speaking, the strength of evidence is its magnitude or extremity, while the weight of evidence is its reliability or validity. Suppose, for example, you need to determine whether a coin is biased towards heads or tails. The strength of evidence corresponds to the proportion of heads in a sample, while the weight of evidence corresponds to the sample size. An individual who

considers the strength of evidence but ignores the weight of evidence would find 4 heads out of 5 more convincing than 20 heads out of 30, even though the opposite is true according to Bayes Rule.

In our setup, we suggest that signals provide the strength of evidence and the system parameters (diagnosticity and transition probability) provide the weight of evidence. Thus, system neglect predicts that subjects will act approximately the same to the signal in Figure 1 and Figure 2, since the signal in each case is identical, even though the systems are substantially different.

The signal is also likely to be overweighted because it is more *salient* than the system. In virtually every real world decision problem, the system parameters are not known, and perhaps cannot be known. Furthermore, since the signal changes over time, unlike the underlying system parameters, it is natural to attend to the signal, not the system. Put differently, the signal is in the foreground, while the system is in the background. Thus, the system neglect hypothesis is similar in spirit to the correspondence bias or fundamental attribution error (Jones & Harris, 1967), and also to the cognitive psychological notion that decision makers often attend to surface features and ignore deep structure.
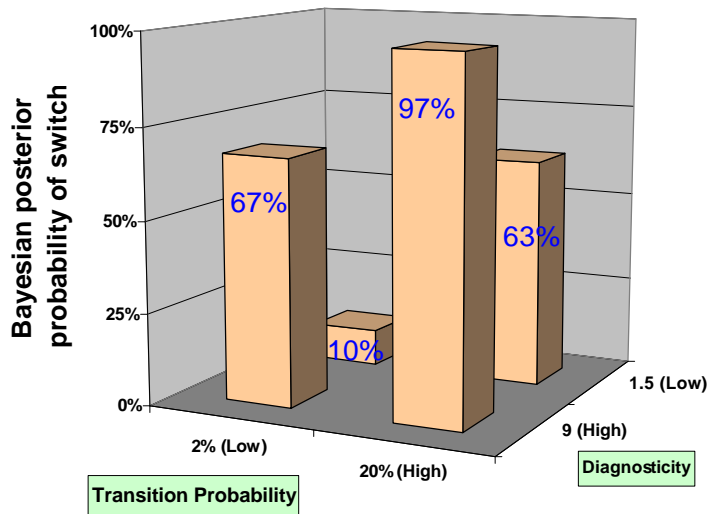
Figure 3: Posterior probabilities of a change for the sequence: red, blue, blue

The system neglect hypothesis implies a particular pattern of over- and under-reaction. We illustrate this pattern with the following numerical example. Consider four systems, corresponding to low ($d = 1.5$) and high diagnosticities ($d = 9$), and low ($q = .02$) and high ($q = .20$) transition probabilities. Suppose that a subject has observed the signal in Figures 1 and 2: a red ball in period 1, a blue ball in period 2, and a blue ball in period 3. The posterior probability that the last ball was drawn from the blue bin is given in Figure 3. The posterior probability in the high diagnosticity, high transition probability cell is almost one (.97), while the posterior probability in the low diagnosticity, low transition probability cell is close to zero (.10).

Suppose that a subject neglects the system completely and gives the same posterior in each case. Such complete neglect gives rise to a pattern in which the greatest tendency for under-reaction is in the low diagnosticity, low transition probability cell (the southeast cell), and the greatest tendency for over-reaction is in the high diagnosticity, high transition probability cell (the northwest cell). For example, if a subject's posterior probability in each case is .50, she will

under-react in 3 of the 4 cells, with the greatest under-reaction in the southeast cell, and the only over-reaction in the northwest cell (Figure 4). Note that the system neglect hypothesis only makes a prediction about the *relative* pattern of over- and under-reaction. We may see all over-reaction (for example, if the subject's posteriors are .98), all under-reaction (for example, if the subject's posteriors are .05), or the mixed pattern we see in Figure 4.

We first test for the predicted pattern of over- and under-reaction in a judgment and choice study.

EXPERIMENT 1: JUDGMENT STUDY

Methodology

We conducted two computer-based studies based on the experimental setup described above. In the first study, subjects ($n = 40$) saw a sequence of signals, and after each signal, provided a probability that the process had changed. Details of both Experiments 1 and 2 are found in Massey & Wu (2002).

The program consisted of several screens with explanations of the statistical process and the payment scheme. We then conducted 18 trials of 10 periods (signals) per trial. The trials involved different systems, constructed by crossing 4 transition probability levels and 3 diagnosticity levels. Diagnosticity levels ($d = p_R / p_B$) were 1.5, 3, and 9, and transition probabilities were .02, .05, .10, and .20. We generated 5 random sequences per condition for a total of 60 sequences. Each subject was given 1 or 2 sequences from each condition. Thus, 12 subjects provided judgments for each of the 60 sequences.
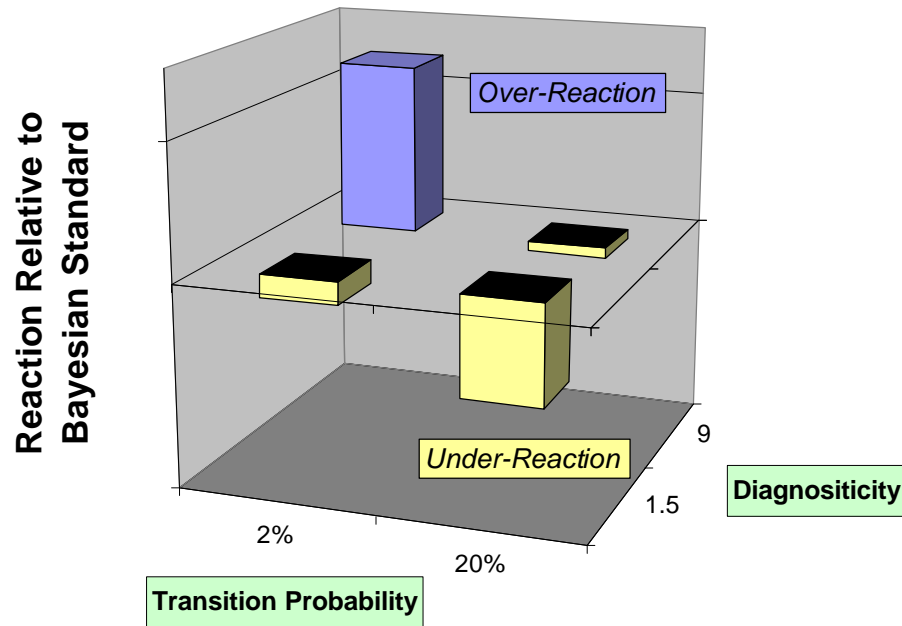
Figure 4: Hypothesized pattern of under- and over-reaction under system neglect

At the beginning of each trial, subjects were shown the parameters, p_R, p_B, and q, governing that trial. These parameters were displayed continuously throughout each trial. Subjects were then shown a series of red and blue balls. After each signal, subjects indicated the probability that the "computer has switched to the blue bin". Subjects were not permitted to change their probabilities once they were entered.

We rewarded subjects according to a quadratic scoring system that paid a maximum of $0.10 per judgment and a minimum of -$0.10 per judgment. Subjects were paid the maximum, for example, if they indicated with certainty that the process was red, and the process indeed was red. The scoring system is proper and truth-revealing for risk-neutral subjects (Brier, 1950).

At the end of each trial, subjects were given feedback as to if and when the process switched from the red bin to the blue bin. The computer also indicated how much they won or lost on that particular trial.

Results

Payments ranged from $6.60 to $14.67 (mean=$11.61). A Bayesian agent would have made $14.23 on average. Thus, an average subject made 18% less than a Bayesian agent.

We are interested in how subjects respond to indications of change, relative to the Bayesian standard. Therefore, we consider *changes* in probability judgments after a subject has observed a blue signal. Figure 5 depicts the Bayesian response (change in probability) to a blue ball for each of the 12 conditions averaged across the trials. The pattern depicted in Figure 5 is identical to that shown in Figure 3: the most reaction is required for $d = 9$ and $q = .2$ (average Bayesian change of .31), and the least reaction is called for when $d = 1.5$ and $q = .02$ (average Bayesian change of .05).

We contrast the Bayesian response with the empirical response shown in Figure 6. Although the responses were not identical across the 12 conditions, they were considerably more compressed than demanded by Bayes Rule, and reveal much less of a gradient sloping upward toward the southeast than found in Figure 5.
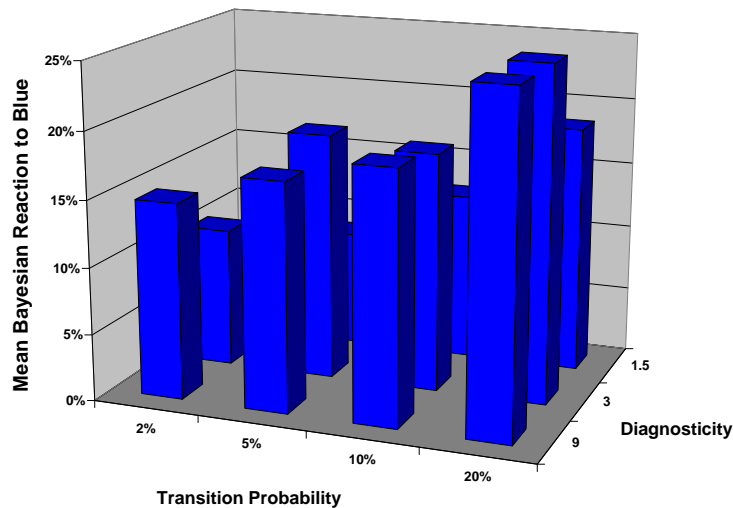

Figure 5: Bayesian change in posterior probabilities after observing a blue ball (Experiment 1)
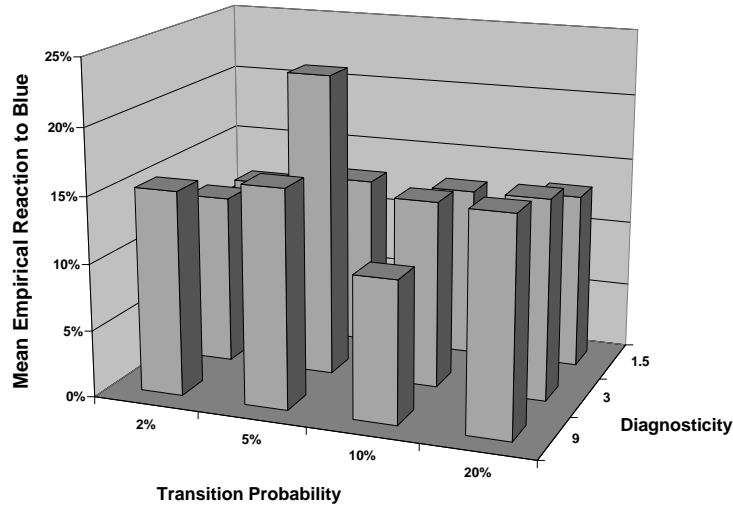
Figure 6: Empirical change in probability judgments after observing a blue ball (Experiment 1)

Our measure of over- and under-reaction is obtained by considering the difference between the empirical and Bayesian change. The plot of over- and under-reaction is shown in Figure 7. We find over-reaction in 6 conditions, and under-reaction in 6 conditions. As predicted, the greatest over-reaction occurs in the most northwest cell (lowest diagnosticity, lowest transition probability), and the greatest under-reaction is found in the most southeast cell (highest diagnosticity, highest transition probability).
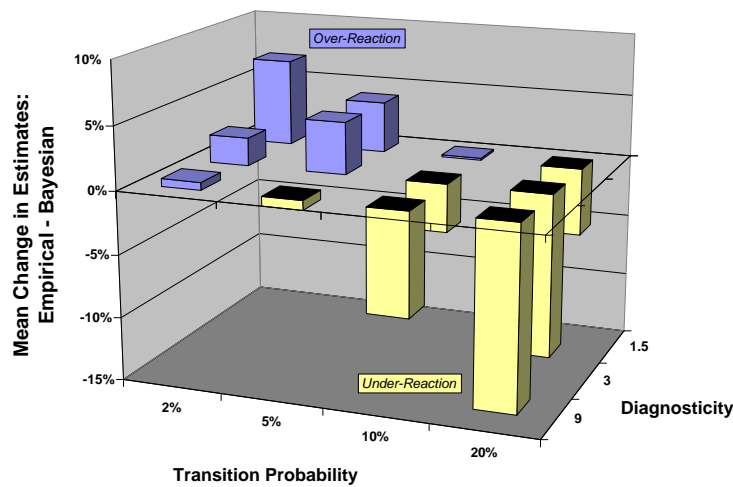
Figure 7: Pattern of under- and over-reaction (Experiment 1)

In Massey & Wu (2002), we provide more formal support for the system neglect hypothesis by fitting a series of quasi-Bayesian models (in the spirit of Edwards, 1968). These models show that subjects are sensitive to the underlying system parameters, but insufficiently so relative to a Bayesian agent.

EXPERIMENT 2: CHOICE STUDY

Methodology

We employed a similar methodology to the one used in Experiment 1. The major difference was the response mode: instead of providing a probability judgment, subjects predicted the color of the next ball.

We recruited 50 subjects for an experiment that consisted of 18 trials of 10 periods. Before each trial, subjects were shown the system parameters governing that trial. Subjects were asked to predict the color of the next ball for each period, including the first period. They were then shown the signal for that period. The task continued until the trial was completed.

Our design used 3 diagnosticity levels ($d = 1.5$, 3, and 9) and 4 transition probabilities ($q = .025$, .05, .10, and .20) for a total of 12 experimental conditions. We randomly generated 3 unique sequences for each of the 12 conditions. Each subject was thus given 50% of the 36 total sequences, and 1 or 2 of the 3 sequences from each condition. Subjects were paid 9 cents for each correct prediction (maximum of 180).

Normative Model

We compare our subjects' behavior with the normative model. The Bayesian posterior that a switch has occurred by $t-1$ is $p_{t-1}^b = \Pr(B_{t-1} \mid H_{t-1})$, and determined by manipulating (1). It

is easy to show that a subject should predict a blue ball if $\Pr(B_t \mid H_{t-1}) = p_{t-1}^b + (1 - p_{t-1}^b)q > .5$ (see Massey & Wu, 2002). Put differently, a subject should predict a blue ball if she believes that it is more likely that the process has switched to the blue bin.

Results

As in Experiment 1, the normative model outperformed the average subject. The normative model made correct predictions 69% of the time, while the average subject was correct 64% of the time (range 59% to 73%). Payments ranged from $9.63 to $11.79 (mean=$10.62).

There are several ways to measure over- and under-reaction. First, we consider when predictions at t are different from prediction at $t-1$. We take switches of this sort as a measure of belief revision, and compare the proportion of trials that a Bayesian agent would revise her predictions (10.8% of the trials) with the proportion of trials that our subjects revise their predictions (16.1% of the trials). By this measure, subjects show a tendency to revise their predictions 50% more than a Bayesian agent.

The system neglect hypothesis predicts that the difference between empirical and Bayesian belief revisions will be most extreme in the southeast and northwest cells. Figure 8 plots the difference between empirical and Bayesian switches for all 12 conditions. We interpret negative values as under-reaction, and positive values as over-reaction. We see the same general pattern as observed in Experiment 1: the most under-reaction occurs in the highest diagnosticity, highest transition probability condition, and the most over-reaction occurs in the lowest diagnosticity, lowest transition probability condition.
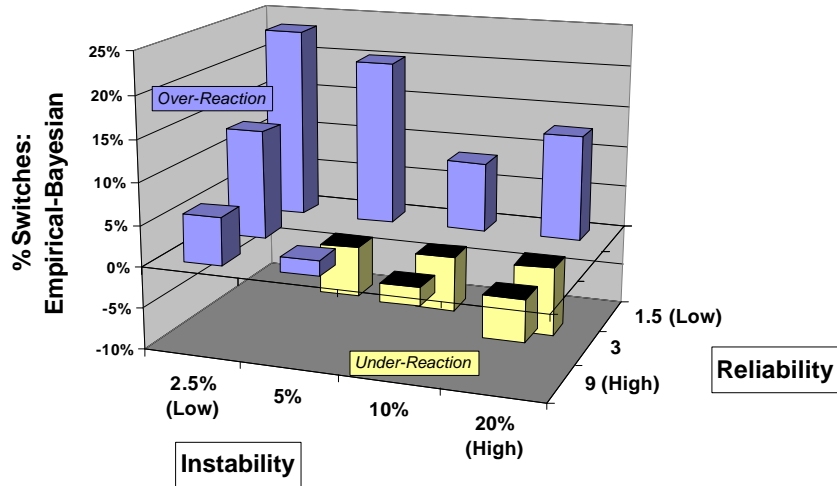
Figure 8: Pattern of under- and over-reaction measured by switches (Experiment 2)

We next consider a second measure of over- and under-reaction: the amount of evidence required to make the first blue ball prediction. Normatively, the threshold should be the same across all conditions, *i.e.*, the first t for which $\Pr(B_t \mid H_{t-1}) = p_{t-1}^b + (1 - p_{t-1}^b)q > .5$. However, the system neglect hypothesis predicts that the "standard of proof" will be different across conditions. Substantial evidence will be required in the high diagnosticity, high transition probability conditions (consistent with under-reaction), and very little evidence will be required in the low diagnosticity, low transition probability conditions.

Figure 9 plots the average Bayesian prior at the time of the first blue prediction. Consistent with the system neglect hypothesis, the threshold varied across conditions. In the low diagnosticity, low transition probability conditions, the average posterior probability for the first blue prediction was .11, considerably lower than the .50 that the normative model demands. By comparison, in the high diagnosticity, high transition probability condition, subjects required evidence indicating a .88 chance of a change to make a blue prediction. Note also that over-reaction appears to be more pronounced in Experiment 2 than Experiment 1.
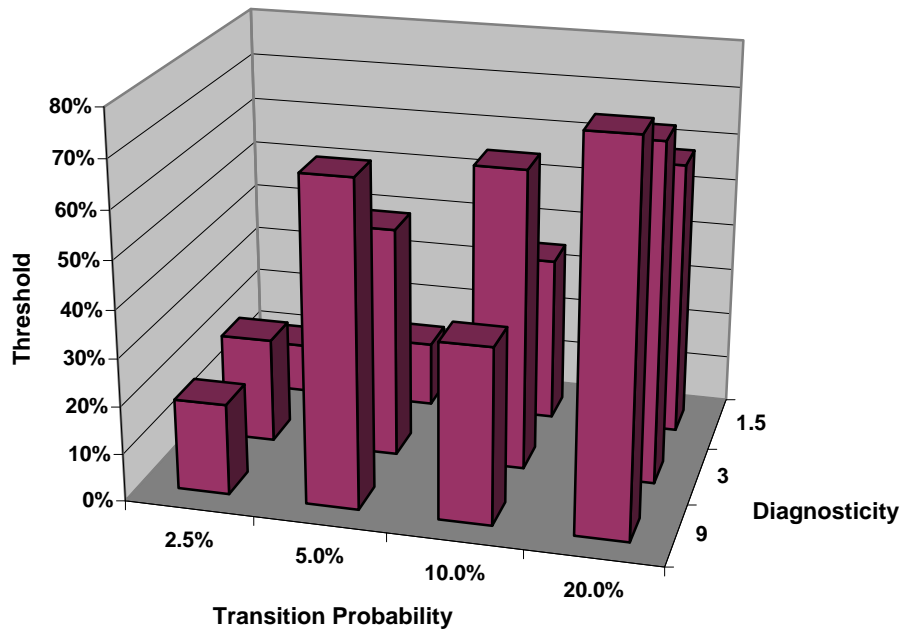
16

Figure 9: Amount of evidence required for first blue prediction. Evidence is measured as the Bayesian posterior at time of first blue prediction (Experiment 2)

We tested for system neglect formally in Massey & Wu (2002) by fitting a quasi-Bayesian model to Experiment 2's data, using a stochastic choice functional. This exercise provides formal support for our hypothesis, and also shows that over-reaction is indeed more pronounced in Experiment 2 than Experiment 1.

SUMMARY

We end by returning to the questions we posed in the introduction, then present some additional questions that offer fruitful avenues for future research.

Old Questions

Question 1. *Do individuals over-react to indications of change? Do individuals under-react to indications of change?*

Our two studies show that individuals both over-react and under-react to indications of change. We show over- and under-reaction, graphically, as well as formally in our estimates of quasi-Bayesian models (see Massey & Wu, 2002).

Question 2. *When do individuals over-react and when do they under-react?*

The pattern of over- and under-reaction is quite systematic. In both studies, there is the most under-reaction in unstable systems with clear signals, and the most over-reaction in stable systems with noisy signals. There also is more over-reaction in the choice task than the judgment task. These two tasks are quite different psychologically, but may facilitate anchoring and adjustment processes of different sorts. In the judgment task, it is natural to anchor on one's judgment in the previous period. In contrast, in the choice task, it is natural to anchor on the signal itself. Anchoring and insufficient adjustment, applied very differently due to the difference in the nature of the tasks, could then explain the differential level of over-reaction in the choice study.

Question 3. *What psychological processes explain the pattern of over- and under-reaction?*

We put forth the system neglect hypothesis: individuals attend primarily to the signal and secondarily to the system that generates the system. This hypothesis draws on psychology that has been used to explain behavior in both cognitive and social psychology. The main implication of this hypothesis is a pattern of over- and under-reaction, with the greatest tendency for under-reaction in unstable systems with precise signals, and the most over-reaction in stable systems with noisy signals. Two studies find exactly this pattern of over- and under-reaction. The system neglect hypothesis is further supported by estimation of quasi-Bayesian models.

New Questions

We close by offering three new questions and presenting some preliminary answers to these questions.

Question 4. *Does the pattern of over- and under-reaction diminish over time?*

Learning has received considerable recent attention in experimental economics, both in terms of experimental study and formal models (cf. Camerer & Ho, 1999). Do subjects learn to detect change better when given experience? In Experiment 1, we looked at the first 4 periods and the remaining 14 periods separately. The gradient shown in Figure 5 was most pronounced in the first 4 periods, but the pattern was still significant in the remaining 14 periods.

We also conducted a study in which subjects were given 20 trials from 1 of 6 different systems. System neglect was still apparent at the end of the 20 trials, but was significantly less pronounced than at the beginning of the experiment. However, most of the learning occurred in systems with high diagnosticity. A more complete discussion of the results is found in Wu & Massey (2003).

In summary, subjects do appear to learn with repetition, but most of the learning happens very quickly and is probably due as much to increased comprehension of the task as real learning. Beyond that, learning is very slow, and appears to occur reliably only in conditions highly suited for learning—conditions with high diagnosticity, where reinforcement is very strong.

Question 5. *What other psychological influences lead to over- and under-reaction?*

In this chapter, we have highlighted the role of the system neglect hypothesis in predicting patterns of over- and under-reaction. However, we believe that there are other psychological factors that may influence the ability to detect change, particularly in real world detection tasks. First, in most real world situations, individuals often have a vested interest in

one of the two regimes. For example, a "bricks and mortars" company will not be eager to see the advent of the "new economy", while those heavily invested in the technology sector will be looking for signs of change. In addition, decision makers may have invested time or money in the incumbent regime. Thus, factors such as sunk costs, commitment escalation and motivated reasoning may make under-reaction more pronounced in many real world decisions (*e.g.*, Staw, 1981; Kunda, 1990). Understanding how these and other psychological factors interact with system neglect should be a fruitful area for future research.

Question 6. *What prescriptions should be offered for this bias?*

Unfortunately, most behavioral biases do not have trivial remedies (Russo & Schoemaker, 2002). This undoubtedly holds for the behavior documented here. However, we offer some preliminary thoughts.

Much popular management literature has called for organizations to be nimble, reactive, and flexible (Bhide, 2000; Hamel, 2000; Schoemaker, 1995). What often goes unsaid is that organizations must be appropriately reactive. Clearly there is a trade-off between flexibility and perseverance. This research suggests there are environments in which managers are prone to stay the course when they should be react, as well as environments in which they react when they should stay the course. Managing this tension is a major challenge for firms operating in dynamic environments.

Our research shows that individuals in dynamic environments emphasize indications of change over the system providing those indications. For managers, this means a tendency to respond to events rather than to the environment in which those events take place. Our research suggests that organizations should devote more resources to analyzing and interpreting the environment. Firms need to understand whether their environment (the "system") is more likely

to lead to errors of under-reaction or errors of over-reaction. For example, a firm in the relatively stable electrical utility industry is likely to overreact to rumors about a change in government regulation, whereas a firm in the rapidly changing technology sector is likely to underreact to the newest market data. Understanding the environment is crucial to correctly interpreting the events affecting an organization and hence managing the tension between under- and over-reaction.

REFERENCES

Brav, Alon and J. B. Heaton (2002). "Competing Theories of Financial Anomalies." <u>Review of Financial Studies</u> 15: 575-606.

Bhide, Amar (2000). <u>The Origin and Evolution of New Businesses</u>. New York: Oxford University Press.

Brier, Glenn W. (1950). "Verification of forecasts expressed in terms of probability." <u>Monthly Weather Review</u> 78: 1-3.

Browning, E.S. (1998). "Bulls use Convoluted Measures to Justify View of the Market." <u>Wall Street Journal</u> (April 20), C1.

Camerer, Colin and Teck-Hua Ho (1999). "Experience-weighted attraction learning in normal form games." <u>Econometrica</u> 67: 827-874.

De Bondt, Werner F. M. and Richard H. Thaler (1985). "Does the stock market overreact?" <u>Journal of Finance</u> 40, 793-805.

Edwards, Ward (1968). "Conservatism in human information processing." <u>Formal representation of human judgment</u>. Benjamin Kleinmuntz (Ed.). New York: Wiley, 17-52.

Gasparino, Charles. (1998). A Top Fund Manager is Humbled When He Makes Ill-Timed Moves. <u>Wall Street Journal</u>, (October 15), A1.

Griffin, Dale and Amos Tversky (1992). "The Weighing of Evidence and the Determinants of Confidence." <u>Cognitive Psychology</u> 24: 411-435.

Grove, Andrew S. (1999). <u>Only the Paranoid Survive: How to Exploit the Crisis Points that Challenge Every Company</u>. New York: Doubleday.

Hamel, Gary. (2000). <u>Leading the Revolution</u>. Boston, MA: Harvard Business School Press.

Jones, Edward E., & Harris, Victor A. (1967). "The attribution of attitudes." <u>Journal of Experimental Social Psychology</u> 3: 1-24.

Kahneman, Daniel and Amos Tversky (1973). "Subjective probability: A judgment of representativeness." <u>Cognitive Psychology</u> 3: 430-454.

Kunda, Ziva (1990). "The case for motivated reasoning." <u>Psychological Bulletin</u> 108: 636-647.

Massey, Cade and George Wu (2002). "Detecting Regime Shifts: The Psychology of Over- and Under-reaction", working paper.

Russo, J. Edward, and Paul J.H. Schoemaker (2001). <u>Winning Decisions</u>. New York: Doubleday.

Schoemaker, Paul J.H. (1995). "Scenario Planning: A Tool for Strategic Thinking." <u>Sloan Management Review</u> (Winter): 25-40.

Shewhart, Walter A. and W. Edwards Deming (1986). <u>Statistical Method from the Viewpoint of Quality Control</u>. Dover Press.

Staw, Barry M. (1981). "The escalation of commitment to a course of action." <u>Academy of Management Journal</u> 6: 577-587.

Wu, George and Cade Massey (2003). "Learning to Detect Change: The Role of Feedback and Experience." Manuscript in preparation.