Prediction is very hard, especially about the future - Yogi Berra

In this chapter, we apply the predictive model framework to explain the possibility of the wisdom of crowds. Our analysis culminates in two theorems: *The Diversity Prediction Theorem* and *The Crowds Beat Averages Law*. The first states that a crowd's collective accuracy equals their average individual accuracy minus their collective predictive diversity.[1] So for predictive tasks, the answer to the question, how much does diversity matter? The answer is just as much as ability. No less. No ore. The second states that the accuracy of the crowd's prediction cannot be worse than the average accuracy of its members. Thus, the crowd necessarily predicts more accurately than its average member. So, we are, on average, above average. Furthermore, the amount by which the crowd does out predict its average member increases as the crowd becomes more diverse.

These results exist independently of our predictive model framework. They've long been known though in different form. Thus, we spend the bulk of this chapter connecting these theorems to our predictive model framework, and showing how these theorems can be seen as showing the value of diverse predictive models.

In this chapter, we also see how crowds of people using diverse interpretations can predict more accurately than the models based on independent signals would suggest. Though convenient, an assumption of independence may understate the predictive ability of small crowds. We also see that it overstates ability of large crowds. In addition to looking at crowds alone, we compare the performance of crowds to that of experts. We provide conditions under which we would expect the crowd to predict more accurately and conditions under which we would not. We even consider crowds of experts – what we call the *crowd of models*. These crowds may predict best of all. Finally, we show how incentives can improve the accuracy of predictions.

We restrict our discussion in this chapter to crowds of people. We could also include examples of other species – ants, crows, and bison – as well as examples of machines and algorithms. The bison example provides a hint of how other species exploit diversity. Bison take different routes across a mountain range. Each bison leaves a trail. Two trails in fact. A continuous

[1]By accuracy, here we mean the distance between the model's predictions and they actual outcomes, and by predictive diversity, we mean the average distance between each model's prediction from the average predictions of other models.

foot trail and a discrete trail that we ignore. More heavily traveled routes become more beaten down and encode the collective wisdom of the bison. Norman Johnson has constructed models that demonstrate this phenomenon. The collective trails over time become efficient.[2]

Covering this substantial terrain requires moving back and forth between several types of models. For many readers, this will prove challenging for many readers. The analysis includes quite a few mathematical calculations. Thought no one calculation takes much effort: the hardest require squaring the difference of two numbers. But we do this quite a lot. The payoff at the end proves worth this effort. We get the theorems. These theorems are not political statements, but mathematical facts. To understand that fact we must roll up our sleeves and work through a few simple formulae.

Before we begin, we should keep in mind that we implicitly limit attention to challenging predictive tasks and intelligent predictors. Any easy predictive task (which will be warmer on January 8th, 2006 at noon, International Falls, Minnesota or San Diego, California?) requires neither a crowd nor an expert. Similarly, if the individual predictors do not know much of anything, the crowd may not predict well. If we ask ten thousand first graders to guess the weight of a fully loaded Boeing 747 (about 40 tons), we should expect their average guess to be well off the mark. Some might guess as low as one hundred pounds. Some might guess as high as a billion billion tons. You cannot even make a silk purse out of lots of sow's ears.

The Wise Crowd from Screening Success

To lay a foundation for the remainder of the chapter, we return to *Screening Success* and consider Ray, Marilyn, and Deborah as a crowd. We find this crowd to be wise indeed. In a few pages, we show why crowds must be wiser than the people in them and why this particular the crowd does so well. For the moment though, we revel in the mystery of how this occurs.

Recall that the task in *Screening Success* was to predict whether a given screenplay would produce a profitable movie. Ray, one of our predictors, con-

[2]See Johnson, Norman L (1998) "Collective Problem Solving: Functionality beyond the Individual." Los Alamos Working Paper LA-UR-98-2227; and Johnson, Norman L. (1999). "Diversity in Decentralized Systems: Enabling Self-Organizing Solutions." presented at the Decentralization II Conference, UCLA.

sidered the amount of sexual content. Another predictor, Marilyn, considered the level of violence. These were their interpretations of the screenplays. Our third predictor, Deborah, used a much more complicated interpretation which relied on balancing the amount of sexual content and violence. We revisit her predictive model as well as Ray and Marilyn's in the next paragraph. First, we present the mapping from screenplay attributes to whether the movie would be profitable (i.e. good, denoted by G) or unprofitable (i.e. bad, denoted by B).

The Screenplay Attribute to Movie Quality Mapping

S /V	*None*	*Low*	*Mod*	*High*
None	B	B	G	B
Low	B	B	B	G
Mod.	G	B	G	G
High	B	G	G	G

Remember that Ray predicts those screenplays with medium or high levels of sexual content will be good and that others will be bad; Marilyn predicts that those screenplays with medium or high levels of violence will be good; and Deborah predicts that those screenplays that are balanced will be good. Their predictive models can be written in tabular form as follows.

Ray's Predictive Model

S /V	*None*	*Low*	*Mod*	*High*
None	B	B	B	B
Low	B	B	B	B
Mod.	G	G	G	G
High	G	G	G	G

Marilyn's Predictive Model

S /V	*None*	*Low*	*Mod*	*High*
None	B	B	G	G
Low	B	B	G	G
Mod.	B	B	G	G
High	B	B	G	G

Deborah's Predictive Model

S /V	*None*	*Low*	*Mod*	*High*
None	B	G	G	B
Low	G	B	B	G
Mod.	G	B	B	G
High	B	G	G	B

To capture how these three predict as a crowd making a prediction, we assume that they vote based on their predictions. As each person predicts either a good or a bad outcome, we cannot have any ties. If Ray and Marilyn agree that a screenplay is either good or bad, they leave Deborah with no say in the matter. Given our assumptions, Ray and Marilyn make the same predictions on screenplays that have relatively high sexual content and violence (these fall in the four upper left boxes) and on screenplays that have relatively low sexual content and violence (these fall in the four lower right boxes). We can call this their *agreement set.* We extra terminology helps us keep track of what's what.

Ray and Marilyn's Agreement Set

S /V	*None*	*Low*	*Mod*	*High*
None	B	B		
Low	B	B		
Mod.			G	G
High			G	G

Inside their agreement set Ray and Marilyn have total say. Deborah is irrelevant. Outside of this agreement set, Deborah becomes all-powerful.

Political scientists call her pivotal – her prediction determines the crowd's prediction. Looking at the table just above, we see that Ray and Marilyn make different predictions for eight of the boxes – the boxes not in their agreement set. Filling in Deborah's predictions in those eight cases gives the following predictions from the crowd:

The Crowd's Predictions

S /V	*None*	*Low*	*Mod*	*High*
None	B	B	G	B
Low	B	B	B	G
Mod.	G	B	G	G
High	B	G	G	G

This table should look familiar. It is the original table showing that map from attributes to outcomes. The crowd predicts accurately every time. Amazing? Yes. But given that my parents are named Ray and Marilyn and my older sister is named Deb, shouldn't we have expected something like this? Clearly this example was carefully crafted, but for a purpose. The example shows how diverse predictive models can aggregate in ways far more subtle and sublime than the putting together of distinct pieces described by Aristotle. The Law of Large Numbers cannot get you to one hundred percent, neither can canceling errors.

To make sense of how the crowd can be 100% accurate, we need to compare this example to the sweater example in *The Gravity of Truth Model.* In both examples, each individual predicts correctly three fourths of the time. However, in the sweater example, the crowd predicted correctly only 84% of the time. What accounts for this difference? In the sweater example, we assumed independent individual signals. One person's reaction to the wool was independent of another's. We made no such assumption in *Screening Success.* The predictions in *Screening Success* must not be independent. They must be better than independent some how. And, they are. In those cases that Ray predicts incorrectly, Marilyn predicts correctly more than three fourths of the time. Thus, she purposefully as opposed to randomly cancels out his errors. (That's true of my parents as well.) This reduces the probability that the crowd makes an error. Statisticians cal this *negative correlation* As will become clear, the wisdom of crows resides partly in the presence of negative correlation or the lack of positive correlation.

To show negative correlation mathematically, we show that when Ray predicts correctly, Marilyn is *less* likely to predict correctly, as this implies that when Ray predicts incorrectly, Marilyn is more likely to predict correctly. So she cancels out his mistakes. To do this, we first write down the screenplays that Ray predicts correctly. We then highlight those screenplays among them that Marilyn also predicts correctly.

Correct Predictions by Ray

S /V	*None*	*Low*	*Mod*	*High*
None	**B**	**B**		B
Low	**B**	**B**	B	
Mod.	G		**G**	**G**
High		G	**G**	**G**

The table shows that Ray predicts correctly in twelve of sixteen cases. Marilyn predicts correctly in just eight of those twelve cases or 2/3 of the time. If her probability of predicting correctly had been independent of Ray's probability, then she would predict correctly 3/4 of the time, or in nine of the cases. Eight is less than nine, therefore Marilyn predicts correctly less often that would be the case if her predictions were independent of Ray's. Or, put in the formal language of statistics: the correctness of their predictions is negatively correlated.[3]

As we have already admitted, this example is contrived. However, it reveals a deeper truth. Ray and Marilyn's negatively correlated predictions provide the key insight. Notice that they look at different attributes of the same perspective. Earlier, we called these *projection interpretations*. Ray and Marilyn's projection interpretations do not contain any of the same attributes, so let's get precise and call them *non overlapping projection interpretations*.[4] In yes or no, good or bad predictive tasks such as the

[3]The eager reader can calculate whether Deborah's predictions are also negatively correlated. (Hint: they are)

[4]If screenplays had a third attribute, say level of humor, and both included it in their interpretations, then they would overlap.

one considered here, non overlapping projection interpretations *always* create negatively correlated predictions. We state this formally.[5]

The Projection Property *If two people base their predictive models on different variables from the same perspective (formally if they use non overlapping projection interpretations) then the then the correctness of their predictions is negatively correlated for binary predictions.*

Understanding the Projection Property requires careful thought. It says that if two people look at different attributes of the same perspective, i.e. different dimensions, And if the task is to predict success or failure, or any other binary outcome like good or bad, or yes or no, then when one person is correct, the other person is less likely to be correct. They're better at collectively predicting than they'd be if they got independent signals.

At first, this result might seem hard to believe, or at least unintuitive. Yet it has a simple explanation that goes as follows. We know that it must be possible for two people to make predictions so that when one is right, the other is more likely to be wrong. The obvious way to do this would be to make diverse predictions. How better to make diverse predictions than to look at different attributes? To see this in another way, if the two models looked at the same attributes, then their predictions would be the same. As a result, when one person predicts correctly, so will the other.

The projection property implies that crowds containing people who look at diverse attributes will be wise. Unfortunately, this insight cannot be leveraged as much as we might hope. The dimensionality of the perspective defines the number of non-overlapping projection interpretations. A perspective that creates a five dimensional representation of an event or situation can support at most five non-overlapping projection interpretations. A perspective that creates ten dimensions can only support ten non-overlapping projection interpretations.

To avoid positive correlation as the number of people in the crowd becomes larger, people must either use cluster interpretations or they must base their interpretations on different perspectives. Deborah's interpretation is an example of the former. Though based on the same perspective, it is not a projection interpretation. This may not be likely with lots of voters.

[5] For a precise statement of the claim see Lu Hong and Scott Page "Generated and Interpreted Signals" on my web site.

Many papers, in fact, even seminal papers in political science and economics assume infinite numbers of people getting independent signals (believe it or not, using infinity makes the math easier). If these signals come from predictive models, that assumption just doesn't have logical foundations. It's convenient though, but so is self toasting bread. In writing good models, we shouldn't confuse convenient assumptions with good ones. By assuming independent signals, these scholars assume more diversity than may exist.[6]

The Diversity Prediction Theorem

Now that we've worked through an example, we're ready to turn to the more general theorems that reveal the importance of diverse predictive models among the members of a crowd. Versions of these theorems can be found in computer science, statistics, and econometrics.[7] To describe this theorem, we need two measures. The first captures how much a collection of predictive models differs. The other captures how accurate they are. Both are based on the same accuracy measure: *squared errors*. In statistics, errors are squared so that negative errors and positive errors do not cancel one another out. If errors were added, a person who was equally likely to be overestimate or underestimate an amount on average, would make no errors (-5 + 5 = 0) . If we first square the errors, then the negative and positive errors do not

[6] The aggregation of predictive models through voting should not be confused with processes that generate *common knowledge*. Information becomes common knowledge when everyone knows that everyone knows that everyone knows something, and so on. So, for example, because Ray knows the rows and Marilyn knows the columns, they must collectively know the row and the column and between the two of them they can predict with one hundred percent accuracy. To see this, suppose that the applicant no sexual content and a medium level of violence. Under an assumption of common knowledge each would know the other person's interpretation as well as the outcome associated with each combination of attributes. They would therefore know that Marilyn's prediction is correct. See Geanakoplos, John. 1992. "Common Knowledge" *The Journal of Economic Perspectives* 6(4):53-82.

[7] The version of the theorem that we describe follows Krogh, A. and Vedelsby, J. 1995. "Neural Network Ensembles, Cross Validation, and Active Learning." In *Advances in Neural Information Processing Systems 7.* Edited by G. Tesauro, D. S. Tourtetsky, and T.K. Leen, 231-38. Cambridge, MA: MIT Press. For more general background see Leamer. E (1978) *Specification Searches - ad hoc Inference with Nonexperimental Data.* New York: John Wiley and Sons.

cancel $((-5)^2+5^2 = 25+25 = 50)$. To build the logic of the theorem, we first construct an example. Suppose that Micheala and Julianna have developed models to predict where three students Maggie, Cole, and Brody will place in an upcoming spelling bee at Rudy Giuliani Elementary. The table below shows their individual predictions, their average prediction, and the actual outcome from the bee.

Predictions of the Rudy G. Bee.

	Micheala	*Julianna*	*Average*	*Outcome*
Maggie	6	10	8	6
Cole	3	7	5	5
Brody	5	1	3	1

We first compute the squared errors of Micheala and Julianna's predictions. Michaela picks Maggie to take sixth place and she takes sixth, an error of zero. She picks Cole to take third and he takes fifth, an error of two. And she picks Brody to take fifth, but he takes first place, an error of four. Squaring these three errors gives zero, four, and sixteen. The sum of the her errors equals twenty.

Micheala's Individual Error: $(6-6)^2+(3-5)^2+(5-1)^2 = 0+4+16 = 20$

We next make the same calculation for Julianna. She misses Maggie's placement by four, she misses Cole's by two and gets Brody's place exactly right. Squaring these errors gives sixteen, four, and zero, for a total squared error of twenty.

Julianna's Individual Error: $(10-6)^2+(7-5)^2+(1-1)^2 = 16+4+0 = 20$

The sum of each of their squared errors equals twenty, so their average sum of squared errors also equals twenty. We call this the *average individual error*. Here that's easy because their errors are the same.

Average Individual Error: *Average of the individual squared errors*

$$\frac{20+20}{2} = 20$$

We next compute the error of their *collective prediction:* the average of their individual predictions. The collectively predict that Maggie will take eighth place. She takes sixth for an error of two. Their collective prediction for Cole, fifth place, is correct, and their prediction for Brody is off by two. Squaring these errors gives four, zero, and four, for a total of eight. We call this their *collective error.*

Collective Error: *Squared error of the collective prediction*

$$(8-6)^2 + (5-5)^2 + (3-1)^2 \ = \ 4 + 0 + 4 \ = 8$$

Notice that their collective prediction is more accurate than either of their individual predictions. The explanation for this can be found in the diversity of their predictions. When one of them predicts too high, the other predicts too low and their mistakes, while not canceling entirely, become less severe. To make this relationship between the diversity of their predictions and the accuracy of their collective prediction more formal, we calculate how much their predictions differ. We do this by calculating the Juliana's squared distance from the collective prediction and Micheala's squared distance from their collective prediction. We then average these two numbers. Statisticians call this the *variance* of their predictions. We call it the *prediction diversity.*

We first compute Micheala's squared distance from the collective prediction. The collective prediction for Maggie is eighth place. Micheala predicts sixth place for a difference of two. The collective prediction for Cole is fifth place, and she predicts third place for a difference of two. Finally, the collective prediction for Brody is third place, and she predicts fifth, a difference also equal to two. The squares of these differences are four, four, and four which sum to twelve.

Micheala's Squared Distance from the Average: $(6-8)^2 + (3-5)^2 + (5-5)^3 \ = \ 4 + 4 + 4 \ = 12$

As there are only two predictors in this example, Julianna's distance from the average in each case must be the same as Michaela's. That calculation can be made as follows:

Julianna's Squared Distance from the Average:$(10-8)^2 + (7-5)^2 + (1-3)^2 \;=\; 4+4+4 \;= 12$

The *prediction diversity* equals the average of these two distances, in this case, it equals twelve.

Prediction Diversity: *Average squared distance from the individual predictions to the collective prediction.*

$$\frac{12+12}{2} \;=\; 12$$

Notice the relationship between the collective error (8), the average individual error (20), and the prediction diversity (12): *Collective error equals average error minus diversity.* This equality is not an artifact of our example. It is always true. And, even better, it holds for any number of predictors, not just two predictors as in our example. Thus, we call this the *Diversity Prediction Theorem.*

The Diversity Prediction Theorem: *Given a crowd of predictive models*

Collective Error = Average Individual Error - Prediction Diversity

We have to be careful to not over or under state what this theorem means.. It doesn't say that you don't want all accurate people. If individual people predict perfectly, they cannot be diverse. If average individual error equals zero, then diversity must also equal zero. Notice also that prediction diversity equals the *average* squared distance from the collective prediction. Adding someone who predicts differently need not increase overall prediction diversity. Prediction diversity only increases if the additional person's predictions differ by more, on average, than those of other people. This implies a limit to the amount of predictive diversity we can have. If a collection of people has an average individual error of one thousand, then their prediction diversity cannot exceed one thousand. Any more diversity and the collective error would becomes negative, an impossibility.

Fine we've got some caveats. But they just reveal some of the theorem's subtleties. What's important that we keep in mind is the core insight: individual ability (the first term on the right hand side) and collective diversity (the second term) contribute *equally* to collective predictive ability. *Being different is as important as being good.* Increasing prediction diversity by a unit results in the same reduction in collective error as does increasing average ability by a unit.

Contrasting the *Diversity Trumps Ability Theorem* with this theorem about prediction reveals important differences. In making a prediction, a group of randomly selected predictors might or might not predict more accurately than a group of the best predictors. Randomly selected predictors will be more diverse to be sure, but they will also be less accurate. Those two effects work in opposite directions. So, we cannot expect that a random intelligent group will predict more accurately than the group of the best. Yet, that stronger claim holds in the problem solving context. The reason why is that poor performers fail to drag down problem solving teams. If we bring Larry, a social scientist, into our cheese making business, his lack of relevant tools won't hurt our cheese making. We just ignore him. He may cause delay or frustration, but if he only has bad ideas - peppermint cheese – those ideas won't be adopted. However, if we're predicting how much cheese to make, we won't know that he doesn't know and his prediction gets averaged along with everyone else's. And, he could make the crowd less wise.

An implication of the theorem is that a diverse crowd always predicts more accurately than the average of the individuals. This runs counter to our intuition. We can call this *The Crowd Beats The Average Law.*

The Crowd Beats The Average Law: *Given any collection of diverse predictive models, the collective prediction is more accurate than the average individual predictions*

$$\textit{Collective Prediction Error} < \textit{Average Individual Error}$$

The Crowd Beats The Average Law follows from the *Diversity Prediction Theorem.* The *Diversity Prediction Theorem* says that collective error = average individual error - prediction diversity. Prediction diversity has to be positive if the predictions differ. Therefore, the collective error must be larger than the average individual error. There's no deep math going on. But the insight is powerful nonetheless.

We now have a logic for the wisdom of crowds. In an ideal world these formal claims would replace pithy statements like "two heads are better than one" but they may not be catchy enough. We can try though. We might replace the Diversity Prediction Theorem with "the wisdom of a crowd is equal parts ability and diversity" and *The Crowd Beats The Average Law* with "the crowd predicts better than the people in it." Not memorable, but accurate.

A Crowd of Draft Experts

To cement our understanding of the logic, let's consider some real data. Die-hard theorists prefer constructed examples because they are neater and cleaner. But sometimes, even a theorist cannot help put peak out the window. If we want to look at some data, we might as well look at something important: football draft selections. The table below shows predictions for the top dozen picks in the 2005 NFL draft from eight prognosticators. The players are listed in the order that they were selected. Each predictor provides a ranking of the draftees. We use the NFL draft because it clean, integer valued data and because it can be seen as ramped up version of our earlier example that involved Julianna and Michaela and because these experts' predictions came from detailed analyses. They don't call them draft experts for nothing. These people, er men, devote long days and nights evaluating team needs, player skills, and a host of other factors.

If we look at their predictions, we see that they differ in their accuracy. The table below reveals that some do far better than others. The last column, by the way, shows the average prediction.[8]

Experts Predictions of 2005 NFL Draft

[8] The analysts are Scott Wright from *NFL Countdown*, James Alder from About.com, the Fanball Staff at Fanball.com, the Sporting News, Paul Zimmerman from *Sports Illustrated* and Clark Judge and Pete Prisco from *CBS Sportsline*.

Player /Predictor	Wright	Alder	Fanball	SNews	Zimm	Prisco	Judge	Crowd
Alex Smith	1	1	1	1	1	1	1	1.0
Ronnie Brown	2	2	4	2	2	5	2	2.7
Braylon Edwards	3	3	2	7	3	2	3	3.3
Cedric Benson	4	4	13	4	8	4	8	5.9
Carnell Williams	8	5	5	5	4	13	4	6.4
Adam Jones	16	9	6	8	6	6	9	8.1
Troy Williamson	13	14	12	12	13	7	7	9.7
Antrell Rolle	6	6	8	10	9	8	6	7.9
Carlos Rodgers	9	8	9	9	16	9	9	9.9
Mike Williams	7	7	7	6	7	12	12	8.0
Demarcus Ware	11	15	14	24	11	11	13	13.9
Shawn Merriman	12	11	3	11	12	10	11	10.1
$Error^2$	158	89	210	235	112	82	75	34.4

These data show *The Crowd Beats The Average Law* in full force. The average of the individual errors equals 137.3 The collective error, shown in the last column, equals about one fourth of that, 34.4. In this example, the crowd even predicts more accurately than its most accurate member. The Crowd Beats The Average Law makes no such claim.[9] The example also shows the power of diversity. These predictors are so diverse that they collectively predict well.

Even more amazingly note that this comparison between the crowd and its most accurate member is unfair. In selecting the best person after the fact, we stack the deck against the crowd. No one, other than perhaps Clark Judge himself, would have predicted Judge (despite his name) to be more accurate than the others. Next year, Judge may not be the best predictor. To take another example with higher stakes, successful investment funds differ from year to year. If at the beginning of the year, we could pick the fund that would do best at the year's end, investing would be fun and easy. But we cannot, so we diversify. By going with the crowd, we take on less risk. We should only go with the expert if we know that person to be far more accurate than the others and the others to make similar predictions.

[9]Dan Catlin performed a similar analysis on the 2005 NBA draft and found that Pete Prisco defeated the crowd, but he was the only one.

Points and Ranges

Up until now, we have focused on the difference between the predictions and the outcomes. In many instances, we may want to know best and worst case scenarios. We want to know the range of possibilities. In building a stock portfolio, an investor may care about the range of possible prices. How high might the stock price go? How low might it go? In predicting a potential political uprising, a policy analyst may care less about having an accurate point prediction than about knowing worst and best case scenarios. We can look at the best and worst predictions and the actual outcomes. In every case, the outcome falls within the range of predictions.

Range of Predictions of 2005 NFL Draft

Player	*Actual*	*Low*	*High*
Alex Smith	1	1	2
Ronnie Brown	2	2	5
Braylon Edwards	3	2	7
Cedric Benson	4	4	13
Carnell Williams	5	4	13
Adam Jones	6	9	16
Troy Williamson	7	7	14
Antrell Rolle	8	6	9
Carlos Rodgers	9	8	21
Mike Williams	10	6	12
Demarcus Ware	11	11	15
Shawn Merriman	12	3	13

Amazing? No, not given the diversity of the predictions.

The Madness of Crowds

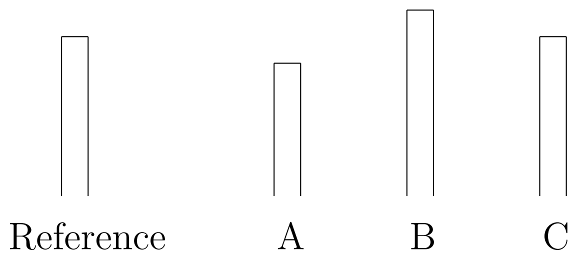
Our analysis so far has implicitly assumed that people cannot communicate. If they can, then they might become less diverse. As Socrates put it, it's much easier to go with the flow, and people often change their predictions to match those of others. And, rather than seeing wisdom emerge, we might see madness. We might see speculators buying tulips at crazy prices. We can use he *Diversity Prediction Theorem* to explain the madness of crowds. When we think of a crowd being mad, we think of a collection of people all taking an action that in retrospect doesn't make sense. The madness of crowds led people to drink the green Kool Aid. The madness of crowds leads people to burn cars and sometimes even houses after sporting events. The madness of crowds also explains stock market bubbles and stock market crashes.[10]

For a crowd to be mad, its members must systematically make the same bad decision. If people make these decisions in the heat of the moment – such as when burning a couch – we can chalk it up to the human tendency to go with the flow, a topic we return to in the epilogue. If though, people have time to construct what they believe to be reasonable predictive models, then we can often blame a lack of diversity. The *Diversity Prediction Theorem* implies that a crowd can make egregious errors if the crowd members lack both accuracy and diversity.

The theorem shows the double-edged sword of deliberation. If people communicate with one another, if they share information and criticize one another's models, they can increase the accuracy of their models. However, they can also reduce their diversity. And, it has been shown with little effort that people often choose to abandon accurate predictive models in favor on inaccurate models. In a classic experiment, Solomon Asch asked to compare the lengths of several lines. Each was given pictures with a reference line and three other lines market A, B, and C.[11]. The figure below provides an approximation of Asch's pictures

[10] A complete theory of bubbles and crashes requires a bit more complication, for though we might blame a lack of diversity for collective predictions gone awry, we can sometimes blame diversity, at least a little bit of it, for bubbles. It's easy to see that diversity in predictive models can drive price increases. For a price to rise, someone must think that the current price is too low. That person must have preferences that differ from those of other people, though they do not have to differ by much. See Sheinkman, J. and Xiong, W. (2003). "Overconfidence and speculative bubbles". *Journal of Political. Economy*, volume 111, 1183-1219.

[11] Asch, S. E. (1956). Studies of independence and conformity: A minority of one against a unanimous majority. *Psychological Monographs,* 70: 416



Subjects were asked which lines were longer than the reference line, which lines were the same length as the reference line and so on. The first subjects to answer were planted by Asch. They purposefully gave wrong answers. Asch found that others follow the majority - giving wrong answers - about a third of the time. Given that people will abandon their stated beliefs in the lengths of lines, we can hardly be surprised that they would abandon their beliefs in their predictions about the stock market, housing prices, or winning number combinations in the lottery.

More than just conformity leads to the madness of crowds. Often, in a group setting, the people will move too far in the direction of the majority opinion. So, if on average people think that prices are going to rise, then the group may work itself into a frenzy and begin to believe that because most people think prices are going up, prices are going to rise substantially.