Math 401

Settlers of Catan Analysis

Peter Keep

12/9/2010

Settlers of Catan Analysis

Settlers of Catan is a board game developed by Klaus Teber, a German board game designer. *Settlers* is described by Mayfair Games, the English publisher of the game:

> "Players are recent immigrants to the newly populated island of *Catan*. Expand your colony through the building of settlements, roads, and villages by harvesting commodities from the land around you. Trade sheep, lumber, bricks and grain for a settlement, bricks and wood for a road, or try to complete other combinations for more advanced buildings, services and specials.
>
> Trade with other players, or at local seaports to get resources you might lack. The first player to achieve 10 points from a combination of roads, settlements, and special cards wins."[1]

The basic goal is to gain points by "settling" the fictional island of *Catan*. *Catan* is an island made up of 5 different resources – brick, wood, sheep, wheat, and ore – that are represented by hexagonal tiles. Players build settlements and cities along the borders of these hexes to gain access to the resources. Resources are given based on rolls – each tile is assigned a number that, when rolled, doles out that particular resource. The game is based on probabilities and economics, as rolls of the dice determine resource income, but supply and demand determine resource value. Other factors that make developing a solid strategy difficult include the multiple ways of scoring points; the fact that the board is assembled randomly, making it different every game; and the interference by other players caused by playing on the same board, and essentially competing for space. Through an analysis of the game, and also a statistical analysis of different strategies, an optimal strategy, or "game theory" can be found. strategy towards gaining points, and a sort of weighing system to judge the optimal significance of each of these.

In *Settlers of Catan*, there are multiple ways to get points:

- Building Settlements
- Upgrading to Cities
- Buying Victory Point Cards
- Building the Longest Road
- Assembling the Largest Army

[1] Mayfair Games. The Settlers of Catan. 09 December 2010. <http://www.mayfairgames.com/game.php?stock=MFG3061>.

Finding the optimal strategy for where to build Settlements and where (and when) to upgrade to Cities requires a great deal of foundational strategy, only some of which is purely math related. There is a great deal of economic theory involved in this strategy, which will not be investigated as much – it will be an assumed knowledge. Before an investigation of Settlement or City strategy can take place, an analysis of the actual *Settlers* board must be done.

Board Analysis

The board is made up of 18 hexagonal tiles, forming the set of resources R={brick, brick, brick, wood, wood, wood, wood, sheep, sheep, sheep, sheep, wheat, wheat, wheat, wheat, ore, ore, ore}. Each tile is randomly paired with a number corresponding to the roll of 2 dice, forming the set of numbers N={2,3,3,4,4,5,5,6,6,8,8,9,9,10,10,11,11,12} (7's are not included, as they correspond with the Robber, another facet of the game). Judging the value of each position requires a combination of the value of the specific resource and the value of the corresponding number.

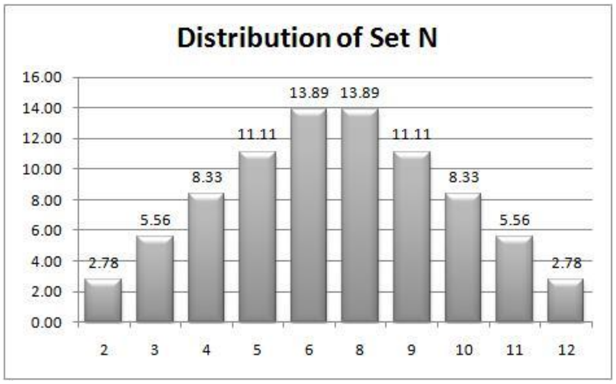

Fig - 1

Set N follows the normal distribution, as seen in Fig – 1. These probabilities essentially indicate some semblance of each number's worth. Also, each number chit (the small circular number tile) represents these probabilities by dots on each chit – the number of combinations for each roll of the 36 total combinations. The worth of numbers can be compared to each other, showing relative worth. For example, a 6 has a worth of 13.89, a 5 has a worth of 11.11, and a 3 has a worth of 5.56 compared to 2's worth of 2.78. So, a 6 is 2.78 more value points than a 5, and 3 is 2.78 more value points than 2, but the relative difference is much smaller between the 6's and 5's. A 6 is only a 25% increase on a 5, while a 3 is a 100% increase on a 2. This comparison will come in extremely useful when an analysis of resources is added.

Resource Analysis

Set R does not follow the normal distribution, and therefore is much harder to find values for. Instead, a more thorough analysis of each resource is required.

> *Settlements*: The most settlements that can possibly be built by a single player in *Settlers* is 5 – upgrading the original 2 settlements to cities and building 5 settlements gives 9 points, and no more settlements before a city us built, in

which case 10 points is reached. So, for settlement building, the most that we need is:

5 wood, 5 brick, 5 sheep, 5 wheat

Roads: With 5 settlements being the maximum number of settlements, there are 9 roads bought, assuming one connects them all to one of the original settlements, and all settlements are spaced 2 spaces apart. If one was forced towards a space of 3 roads twice in our route, there is a required 11 roads. While this is by no means an efficient way of playing, it gives a sufficient upper bound on road spending:

11 wood, 11 brick

Cities: The most cities that can possibly be built in a game are 4 – each player is only provided with 4 cities. So for city building, the upper bound is:

8 wheat, 12 ore

Development Cards: This is the first snag of a spending analysis. Technically, one could buy all of the development cards and not win. This is obviously not a great strategy. A more thorough analysis of development card spending will follow later, but for this purpose, a crude upper bound is appropriate. With a purchase of 8 Development Cards (rarely does a player buy more than 8), the maximum spending is:

8 sheep, 8 wheat, 8 ore

Total: Using this as a guide, an upper bound on resource spending can be acquired. If all of these upper bounds are used (which is impossible give the game limit of 10 points), the total spending for a given player would be:

16 wood, 16 brick, 13 sheep, 21 wheat, 20 ore

These numbers, while most definitely being crudely deciphered limits, give an insight as to importance of resources. This is where the analytical resource falls short, and statistical analysis must be used. The statistical analysis comes in the investigation of 5 elements of game-play: primary resources, secondary resources, city resources, starting empty resources, and ending empty resources.

Primary Resources: Any resource pertaining to the original 2 settlements of a player, placed before regular game-play begins.
Secondary Resources: Any resource pertaining to any settlement built by a player after the original 2 settlements have been places (i.e. any settlement built during game-play).
City Resources: Any resources pertaining to any cities built by a player.

Starting Empty Resources: The complement set of Primary Resources taken before game-play begins (i.e. any resource that is not a Primary Resource at the beginning of the game).
Ending Empty Resources: The complement set of the union of Primary Resources and Secondary Resources (i.e. any resource that is not a Primary Resource or a Secondary Resource).

These categories have been tracked statistically, and can provide insight to what worth is assumed for each resource.

Primary Resources

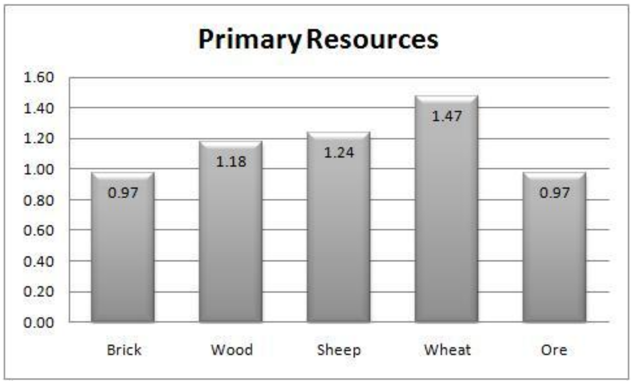

Fig - 2

Fig – 2 shows the tendency for the winning player to place their first 2 settlements on any given resource, where the y-axis represents the number of settlements on the given resource. Using this graph directly would lead us to believe that the order of importance of Primary Resources would be: (1) wheat, (2) wood, (3) sheep, (4) ore, (5) brick. The only flaw in this ordering is the fact that the supply of brick and ore are less than the supply of wood, wheat, and sheep. So, the average number on each resource can be averaged per hex of that given resource, to show the average number of settlements on each hex. As mentioned earlier, set R={ brick, brick, brick, wood, wood, wood, wood, sheep, sheep, sheep, sheep, wheat, wheat, wheat, wheat, ore, ore, ore}, giving the following values to be used as resource worth as well as the following order of importance:

Brick = 0.3233
Wood = 0.2950
Sheep = 0.3100
Wheat = 0.3675
Ore = 0.3233

1. Wheat
2. Ore, Brick
3. Sheep
4. Wood

Secondary Resources

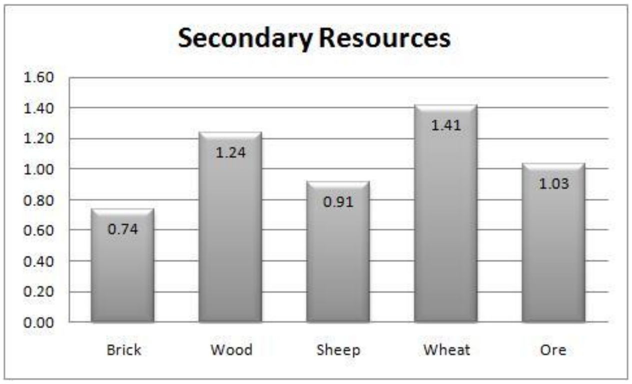

Fig - 3

Fig – 3 shows the tendency of winning players to expand and build settle on another specific resource. This data varies from the Primary Resources data significantly in areas, proving that resource worth is dependent on the demand for that resource. Demand for resources can change drastically. For example, after 2 secondary settlements have been built, players undoubtedly (should) begin attempting to upgrade a settlement to a city, or look into development cards, as there is a maximum of 5 settlements available to be built before upgrades. So, in the secondary resources phase, the resource worth can be averaged per hex again, giving the following values and order of importance:

Brick = 0.2467
Wood = 0.3100
Sheep = 0.2275
Wheat = 0.3525
Ore = 0.3433

1. Wheat
2. Ore
3. Wood
4. Brick
5. Sheep

City Resources

The last phase of the game involving collection of resources involves the upgrade to cities. Fig – 4 shows the tendency of a player to upgrade a settlement to a city on a specific resource. Again, there is a noticeable difference in the value of resources in this phase, as demand for resources are different. Calculating the average number of cities per hex of any given resource presents the following values and order of importance:

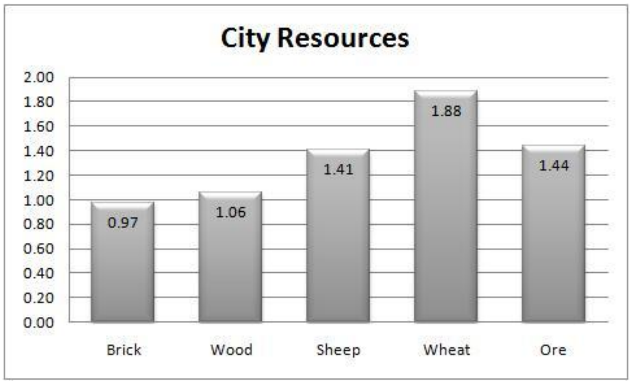

Fig - 4

Brick = 0.3233
Wood = 0.2650
Sheep = 0.3525
Wheat = 0.4700
Ore = 0.4800

1. Wheat
2. Ore
3. Sheep
4. Brick
5. Wood

Starting Empty Resources

Empty Resources show how long and how often a winner can play *Settlers* whilst ignoring certain resources. Fig – 5 shows, then, that 27% of the time, a winner will not begin the game on any brick. Essentially, this shows the worthlessness of each resource.

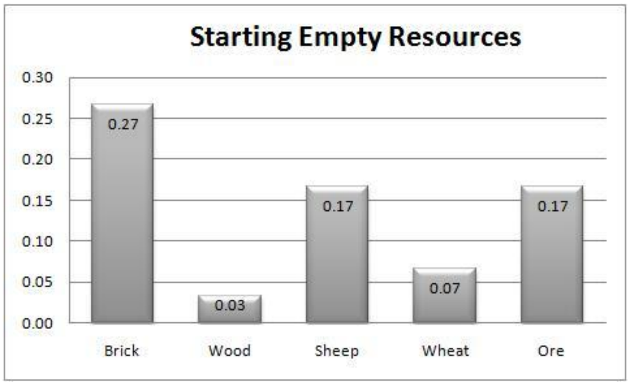

Fig - 5

For these, an average of emptiness per hex would not be appropriate, as a lesser amount of hexes contributes to a more likely probability of being empty by natural causes and not disregard or lack of worth. In this case, then, the percentage of starting emptiness for each resource needs to be corrected by some other factor – ore and sheep have the same percentage of starting emptiness, but this should correspond to sheep being ranked more worthless than ore, as there are more sheep to leave empty. Instead of dividing each value by the total number of hexes, a simple multiplication of each value by the number of available hexes to ignore will give a useful indication of the uselessness of each resource at the starting phase. When these simple calculations are carried out, the following values of worthlessness and order (with most worthless ranked higher) are given:

Brick = 0.81
Wood = 0.12
Sheep = 0.68
Wheat = 0.28
Ore = 0.51

Brick
2. Sheep
3. Ore
4. Wheat
5. Wood

Ending Empty Resources

Fig – 6 shows the percentage of time that a player makes it all the way through the game and wins without ever settling a given resource. So, for example, 13% of the time, a player who wins will never settle onto brick. This gives a great deal of information on the worthlessness of resources – if a player can win the game without ever picking up a resource, is that resource worth settling ever?

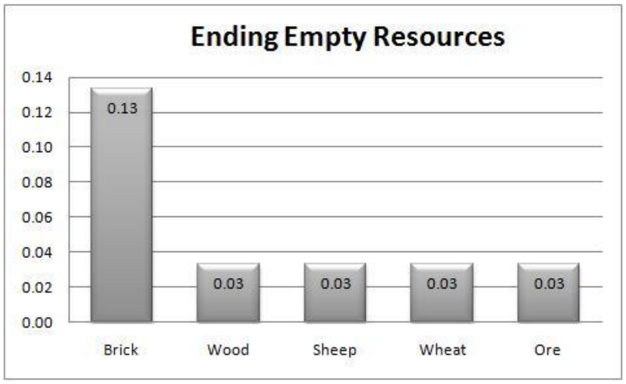

Fig - 6

When calculations similar to those in Starting Empty Resources are made, the following values of worthlessness and order of worthlessness are given:

Brick = 0.39
Wood = 0.12
Sheep = 0.12
Wheat = 0.12
Ore = 0.09

1. Brick
2. Wood, Sheep, Wheat
3. Ore

Total Value Per Resource

The intuitive way of finding total value is to consider the value at each phase of the game a positive value, and the worthlessness at each phase a negative value. The combination of all of these values will give the overall value of a given resource.

Value(Total) = [Value(Primary Resources) + Value(Secondary Resource) + Value(Cities)] – [Value(Starting Empty Resources) + Value(Ending Empty Resources)]

Using this formula, total value of each resources and rankings from highest net worth to lowest net worth are found below:

Brick = (0.3233+02467+0.3233)-(0.81+0.39)
= **-0.3067**
Wood = (0.2950+0.31+0.2650)-(0.12+0.12)
= **0.63**
Sheep = (0.31+0.2275+0.3525)-(0.68+0.12)
= **0.09**
Wheat = (0.3675+0.3525+0.47)-(0.28+0.12)
= **0.79**
Ore = (0.3233+0.3433+0.48)-(0.51+0.09)
= **0.5466**

1. Wheat
2. Wood
3. Ore
4. Sheep
5. Brick

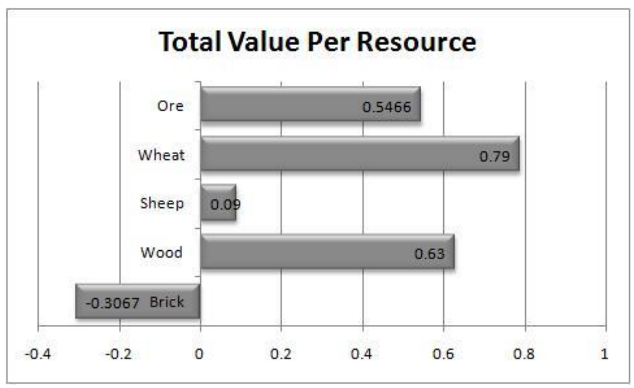

Fig - 7

Fig – 7 shows this information in graphical form. Most useful, though, is the analysis of value and worthlessness at each phase, allowing a player's strategy to adapt to the chronological state of the game. Knowledge of the importance of resources at each point of the game allows a player to manipulate and maximize the productivity of each resource.

Number Analysis

The discovery of the value of each resource throughout the game begs the question: "What about the value of rolls? Do they change throughout the game, or do they remain constantly normally distributed?" A method similar to that used to derive the value of resources can be used to analyze the numbers of *Catan*. This will include 2 categories of data: Primary Numbers, and Secondary Numbers.

Primary Numbers: Any number pertaining to the original 2 settlements of a player, placed before regular game-play begins.
Secondary Numbers: Any number pertaining to any settlement built by a player after the original 2 settlements have been places (i.e. any settlement built during game-play).

Primary Numbers

Fig – 8 shows the tendency for a player to place his primary settlements on a given number. To analyze the worth of numbers at this stage of the game, Fig – 8a can be compared to the Distribution of Set N (Fig – 1). The relative difference is deceiving, as there are only one of each of 2's and 12's available, while there are 2 of each of the rest of the numbers available. Not unlike the process completed for the resources, each number's likelihood to be settled on must be divided by its occurrence on the board. Fig – 8b lays out the distribution of numbers per chit. In theory, the worth of numbers can be found by analyzing the apparent distribution of Fig – 8b. If, for instance, it is normal, or even nearly normal, then it could be concluded that numbers are more important than resources, as the normal distribution would dictate the worth of each number. If, though, Fig – 8b more resembled a uniform distribution, it could be concluded that numbers have no importance – they are randomly paired with a resource, whose value completely dictates the value of the number. In this case, there is obviously some semblance of both the normal distribution and uniform distribution interacting. The curvature of the edges of the graph suggest that the value (or lack of value) of lower numbers act as a deterrent from those hexes, while the flat, middle section implies that those numbers have a uniform value. 2's, 3's, 11's, and 12's are inherently worth less than 4-10, and numbers 4-10 are about the same value.

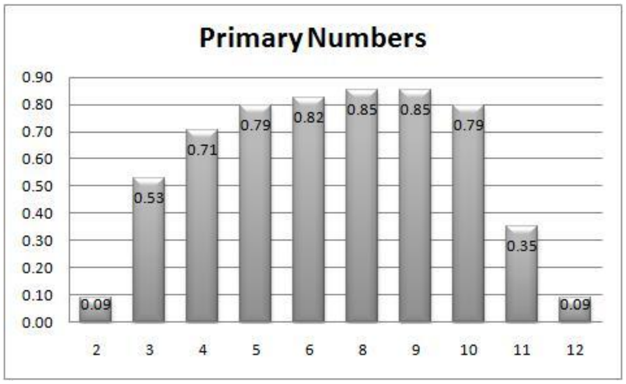

Fig - 8a

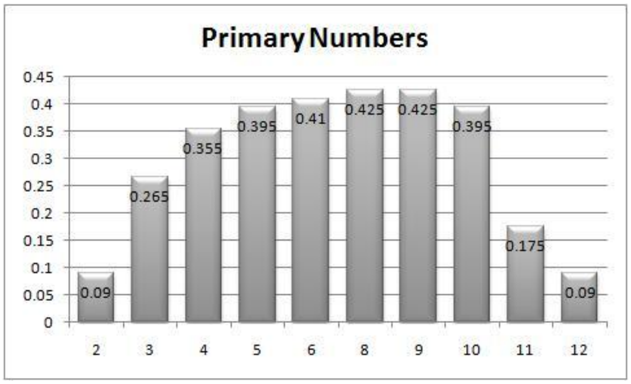

Fig - 8b

Secondary Numbers

Because of the Primary Numbers seemingly remaining fairly uniform in worth at the beginning of the game, and the knowledge that the worth of resources changes significantly throughout the game, the assumption of the distribution of Secondary Numbers remaining fairly uniform is an appropriate assumption. Fig – 9 displays the distribution of Secondary Numbers after the equality calculations have been performed. The distribution seems erratic if not uniform. This shows that the worth of each number fades as the game progresses – by the time a player builds secondary resources, they should be concerned solely with their strategy concerning resources. With knowledge of the importance of each resource, any secondary settlement can be judged by the combination of resources it gives, as well as how well represented those resources are in the players other settlements. For example, if a player has no settlements a wood hex, any secondary spot with wood as an included resource is more valuable. This uniform distribution is seen on all aspects of the game, with the exception of the Primary Numbers. While placing the first 2 settlements, it is advantageous to stay away from 2's, 3's, 11's, and 12's. Using this analysis of resources and numbers well leads to the development of a good strategy of settlements, or *placing strategy*.

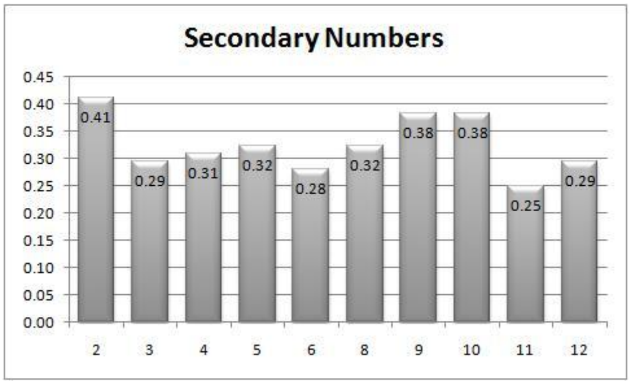

Fig - 9

Development Cards

Development Cards are the next main items to spend resources on. Value of Development Cards must be determined, but in order for this to happen, there needs to be a small analysis of each card and the distribution of each. There are a total of 25 Development Cards in *Settlers*, with the following distribution:

- 14 Knights, with a probability of 56%
- 5 Victory Points, with a probability of 20%
- 2 Road Building Cards, with a probability of 8%
- 2 Monopoly Cards, with a probability of 8%
- 2 Year of Plenty Cards, with a probability of 8%

Knights obviously have some value – aside from warding the robber off of a player's hex or onto an opponent's hex, the Largest Army is worth 2 points, and awarded to the player with the most Knights. Victory Points are of even more obvious importance, as they directly increase a player's score. Road Building, Monopoly, and Year of Plenty are principally glorified trades, as they are an exchange of the resources needed to buy the card for roads or a choice of resource cards. So, the only Development Cards with significant positive value are Knights and Victory Points. Buying Development Cards should be based on the probabilities of attaining Knights and Victory Points. For example, if obtaining the Largest Army is a priority, 6 Development Cards should be bought. Of the 3 Development Cards that will statistically not be Knights, 1 or 2 will be Victory Points (1.4 Victory Points actually), and the rest will have little or no worth. If, though, 4 Development Cards are bought, there is still a chance that the Largest Army and Victory Point will be received. Most likely, though, 2 Development Cards will be Knights, 1 will be a Victory Point, and 1 will chosen from the remaining 3 cards, failing to give the player the Largest Army.

Statistically, a winner will buy 3 Development Cards per game, with 0.8 being Victory Points. The distribution of 3 Development Cards is 1.68 Knights, 0.6 Victory Points, and 0.72 remaining cards. The difference in Victory Point cards could be a slight anomaly, which would move towards the 0.6 target with more games played, but more likely it is a luck factor – if two players buy 3 Development Cards each, and all other things equal, the player who buys a Victory Point by luck will be more likely to win (as they have more points). The lack of Development Card purchasing has heavy implications for the pursuit of the Largest Army.

Largest Army and Longest Road

The utility of the Largest Army and Longest Road can be compared relatively easily. Both give 2 points, and so there is a common reference point. It could be said that since the 2 point output remains the same, the more valuable one is the cheapest one. An evaluation of cost is also easy to perform: the cost of roads and Development Cards are fixed. The Longest Road costs 5 roads to get. This is a resource cost of 5 brick and 5 wood resource cards. The cost of keeping the Longest Road is relatively cheap. It costs 2 roads to steal Longest Road back once a player has lost it (1 road to tie, and another to surpass the opponent). The cost of obtaining the Largest Army is slightly different. Since there is a 56% chance of purchasing a Knight, it takes about 6 Development Cards to buy the necessary 3 Knights. This is a cost of 6 wheat, 6 sheep, and 6 ore. Maintaining the Largest Army is more expensive, as it takes 4 Development Cards to buy the 2 Knights necessary for surpassing the opponent, costing 4 wheat, 4 sheep, and 4 ore. The cost of obtaining and the cost of maintaining can be compared, using the values of resources determined previously. This comparison will determine relative price, and therefore the cheaper priced item will be worth more.

Obtaining Longest Road = 5(0.63) + 5(-0.3067) = **1.617**
Obtaining Largest Army = 6(0.09) + 6(0.79) + 6(0.5466) = **8.560**

Maintaining Longest Road = 2(0.63) + 2(-0.3067) = **0.6466**
Maintaining Largest Army = 4(0.09) + 4(0.79) + 4(0.5466) = **5.7064**

So, the cost of building and maintaining the Longest Road is much less than that of the Largest Army, especially when considering the total resource worth. This is evident when looking at the statistics as well, as a winner will end the game with the Longest Road 52.9% of the time, and end the game with the Largest Army only 23.5% of the time.

There are many other aspects to *Settlers of Catan* that influence the strategies held, most of which are based in some fundamental economic theory. A statistical analysis such as this is a tool that can be used and manipulated to give certain advantages, but there is no guarantee of a perfect strategy with a 100% chance of winning. Through this "game theory", though, the game of *Settlers* can be brought to life a little bit more – looking at it through a lens of mathematics brings out aspects of the total game strategy that could not be found otherwise.