secret_hitler

July 18, 2021

1 Secret Hitler

Nick Holden

Secret Hitler is a popular social deduction game. At the start of the game, the players are split (in secret) into liberals and fascists, where one of the fascists specially marked "Hitler."

The liberals have just enough players to have a voting majority. If there are n players, if n is even there are $\frac{n}{2}+1$ liberals, $\frac{n}{2}-1$ fascists, one of whom is Hitler. If n is odd, then there are simply $\frac{n}{2}$ rounded down fasicsts including Hitler, and $\frac{n}{2}$ rounded up liberals.

1.1 Game objective

- The liberals win if they enact 5 liberal policies, or Hitler is assinated.
- The facists win if either they enact 6 fascist policies, or Hitler is elected chancelor after the 3rd fascist policy has been enacted.

1.2 Gameplay

Secret Hitler is played in rounds. Each round has an election to form a government, then a legislative session to enact a new policy.

- **Presidency**: At the begining of a new round, the person sitting to the left of the previous President becomes the acting President for that round.
- **Chancellorship**: The President then nominates a Chancellor candidate out of the remaining players. After the nomination, all players then vote simultainously and publicly "Da" (yay) or "Nein" (nay) on the proposed Chancellor candidate. If half or more players vote no, the election fails and the round ends. Otherwise, the Chancellor candidate is elected Chancellor for the round.
- **Policy**: After a Chancellor is elected, the government moves into a legislative session. The President draws 3 policy tiles from the top of the policy deck. The deck is made up of **11 Fascist tiles, and 6 Liberal Tiles**. The President then discards one of the 3 tiles and passes the remaining two to the Chancellor. The Chancellor then publically'enacts' one of the two policies. Neither of the discarded tiles are revealed, nor can the President and Chancellor speak until after a legislative session is complete and a policy is enacted.

1.3 Discussion

There are several more rules and to this game, for example the acting president gets special powers if a certain number of facist policies are enacted. However, my goal for this blog post is to delve into the mathematics and probability related to the game, hopefully to glean some strategic edge.

The resulting enacted tile can result from one of or a combination of three factors:

Presidential action: For example if the President draws two fascist tiles and one liberal, she can choose to discard a fascist tile, giving the Chancellor a choice between a fascist and liberal policy, or discard the liberal tile, forcing the Chancellor to play a fascist tile.

Chancellor action: The chancellor chooses from the remaining two tiles. They may be given a two different tiles, they may be given two of the same tiles. The Chancellor may or may not have agency depending on what he is given.

Probabilistic action: The initial three tiles that President draws may contain 0,1,2, or 3 facist tiles with differing probability, let us consider the odds of these 4 events:

- **0 Fascist, 3 liberal:** There is only one combination of this set of tiles. So the odds of randomly drawing 3 consecutive liberal policies is $(\frac{6}{17}) * (\frac{5}{16}) * (\frac{4}{15}) = \frac{6*5*4}{17*16*15} = \frac{1}{34} \approx 0.029$ or 2.9%
- **1 Fascist, 2 liberal:** There are $\binom{3}{1}$ possilbe combinations of this set of tiles. So the odds of this set are given by: $\binom{3}{1} * (\frac{11}{17}) * (\frac{6}{16}) * (\frac{5}{15}) = \frac{3*11*6*5}{17*16*15} = \frac{33}{136} \approx 0.242$ or 24.2%
- **2 Fascist, 1 liberal:** There are $\binom{3}{2}$ possilbe combinations of this set of tiles. So the odds of this set are given by: $\binom{3}{2} * (\frac{11}{17}) * (\frac{10}{16}) * (\frac{6}{15}) = \frac{3*11*10*6}{17*16*15} = \frac{33}{68} \approx 0.485$ or 48.5% This is the most common draw.
- **3 Fascist, 0 liberal:** There is 1 possilbe combination of this set of tiles. So the odds of this set are given by: $(\frac{11}{17}) * (\frac{10}{16}) * (\frac{9}{15}) = \frac{11*10*9}{17*16*15} = \frac{33}{136} \approx 0.242$ or 24.2%. The same as the 1 fascist and 2 liberal draw.

These odds provide useful heuristics while playing. For example, lets say we are elected Chancellor and the President hands us two fascist tiles. What are the odds that the President is fascist? We cannot say exactly because we don't know how a fascist president will balance disguising themself as a liberal by passing some liberal policies, and progressing their fascist agenda. Liberals, on the otherhand, have the voting majority and tend to play much more honestly.

But lets calculate the odds if the president is 100% openly fascist, ie they will always pass as many fascist policies they can to the Chancellor.

Let A be the event that the president is fascist, and B be the event that two fascist tiles were passed. We want to find $P(A|B)$ Which is given by Bayes Thoerem: $P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{P(B|A)P(A)}{P(B|!A)P(!A)+P(B|A)P(A)}$ (!A means not A). So we are left to plug in values.

- $P(B|!A)$: the probability that a liberal president would pass two fascist tiles. This is exactly equal to the probability that the President drew 3 fascist tiles, since otherwise a liberal president will always be able to pass one liberal tile. $P(B|!A) = 24.2\%$

- $P(B|A)$: the probability that a fascist president would pass two fasicst tiles. This value depends on how openly fascist our president is. But since we are assuming that our president is not balancing their play, $P(B|A)$ = P(2 fasist, 1 liberal) + P(3 fascist). So $P(B|A) = 48.5\% + 24.2\% = 72.7\%$
- $P(A)$: the probability any random player is fascist, or the proportion of fascists in the player pool. $P(A) = \frac{\#fascists}{\#players} < 0.5$
- $P(!A)$: the proportion of liberals. $P(!A) = 1 - P(A) = \frac{\#liberals}{\#players} > 0.5$

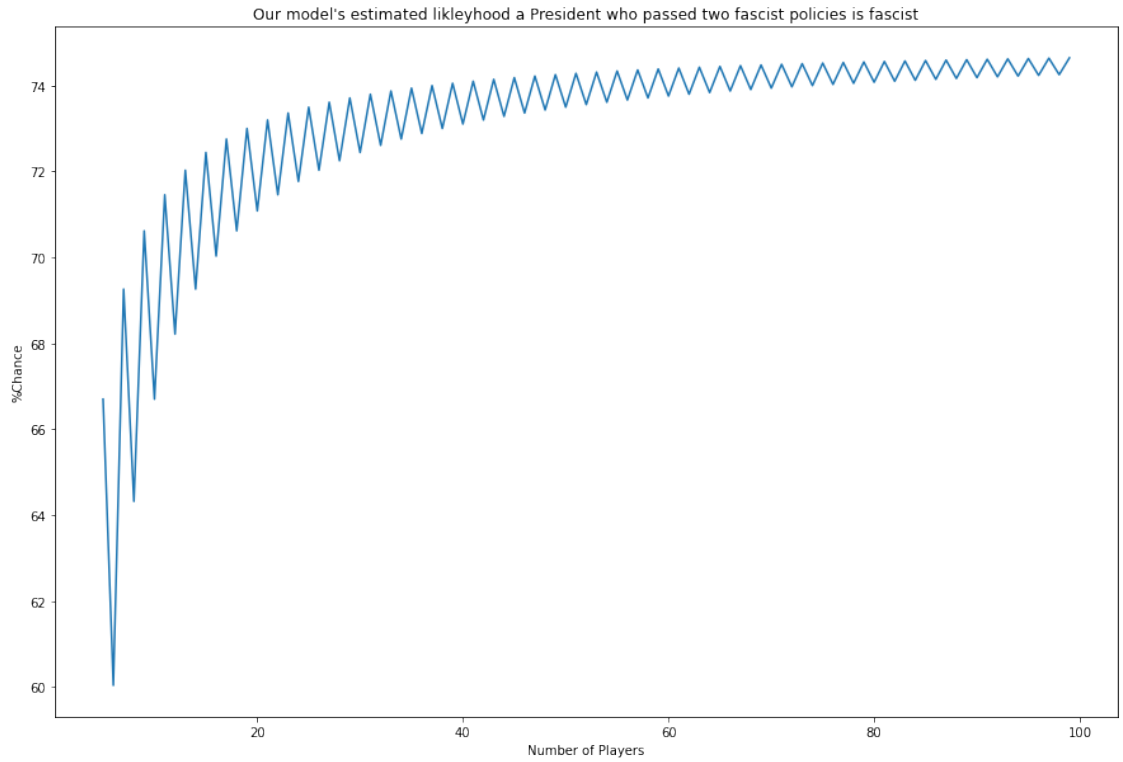
So plugging in, we find $P(A|B) = \frac{P(B|A)P(A)}{P(B|!A)P(!A)+P(B|A)P(A)} = \frac{0.727*P(A)}{0.242*(1-P(A)+0.727*P(A)} = \frac{0.727*P(A)}{0.242+0.485P(A)}$

As we can see, the probability that the president is fascist, $P(A|B)$, is dependant on the ratio of fascists, $P(A)$, which is a function of number of players. So we plot $P(A|B)$ as a function of the number of players

[49]:
```
import numpy as np
import pandas as pd
from numpy.linalg import matrix_power
from numpy.linalg import inv
import math
import matplotlib.pyplot as plt
```

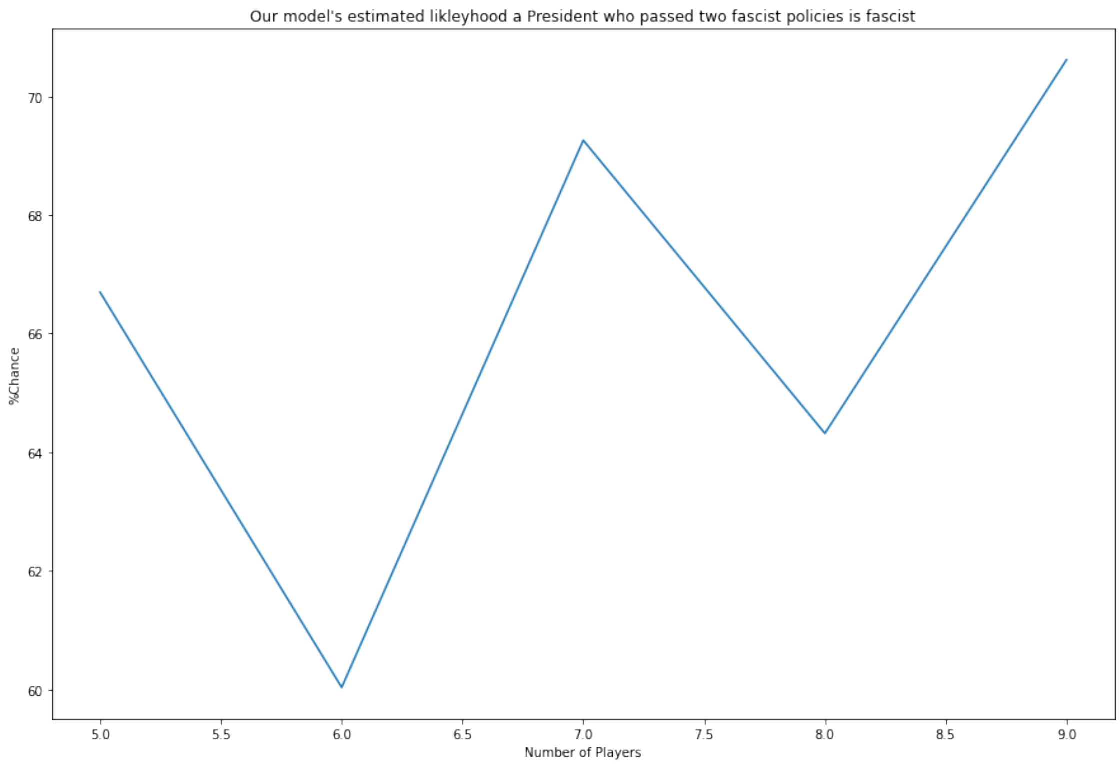
[76]:
```
def plot_2fascist(n):
    p_as = []
    for i in range (5,n):
        if(i % 2 == 0):
            p_a = (int(i/2) -1)/i
        else:
            p_a = (int(i/2))/i

        p_as.append(0.727 * p_a / (0.242 + 0.485 * p_a))
    plt.figure(figsize=(15,10))

    pas = pd.Series(p_as) * 100
    df = pd.DataFrame()
    df['P(A)'] = pas
    df.index = index = range(5,n)
    plt.plot(df)
    plt.title("Our model's estimated likleyhood a President who passed two␣
 ↪fascist policies is fascist")
    plt.xlabel("Number of Players")
    plt.ylabel("%Chance")

    plt.show()
```

[77]:
```
plot_2fascist(100)
```



We see a clear, expected, difference between games with odd and even numbers of players. We also see that passing fascist policies is less telling with less players. This is because fascists make up a smaller portion of the player pool, and therefore its less likley that our president is fascist. Since Secret Hitler is typically played with 5 - 9 players, we plot the probability an unbalanced president is fascist given that he passed the Chancelor two fascist tiles, for that range of players.

[78]:
```
plot_2fascist(10)
```



As we can see, we need to seriously take into consideration the number of players. Both as president when deciding whether or not we can pass two fascist tiles, and as Chancellor. A president passing two fascist tiles is 10% more suspicious with 9 players than with 6. So fascists should play more agressivly with a small, even number of players than in games with a larger odd number of players.

1.3.1 Strategic adjustments

In practice, because the fascists win if Hitler is elected chancelor, Hitler typically plays like a liberal to hide his identity. We also made the assumption that fascists will play honestly 100% of the time. We add these assumptions to the model.

- **Hitler Assumption:** If we assume that Hitler, one of the fascists, plays like a liberal then we have to change $P(B|A)$, the probability a fascist president would pass two tiles. Let $P(B|A)\prime$ be the updated odds a fascist president would pass two tiles given that Hitler will play like a liberal. Like before, there only two cases in which a president is able to pass two fascist policies: if he draws 2 fascist tiles and one liberal(48.5%), or if he draws 3 fascist tiles (24.2%). So, we can only change Hitler's pass rate in the first case. So $P(B|A)' = 0.242 + 0.485 * \frac{P(A)-\frac{1}{n}}{P(A)}$ where n is the number of players. So we update $P(A|B)$ to $P(B|A)' = \frac{P(B|A)\prime P(A)}{P(B|!A)P(!A)+P(B|A)'P(A)} = \frac{0.242*P(A)+P(A)*0.485*\frac{P(A)-\frac{1}{n}}{P(A)}}{0.242+P(A)*0.485*\frac{P(A)-\frac{1}{n}}{P(A)}}$
- **Strategic Balancing** Earlier we assumed that fascists will always pass the most fascist policies possible. But sometimes, especially earlier in the game, fascists might want to keep their

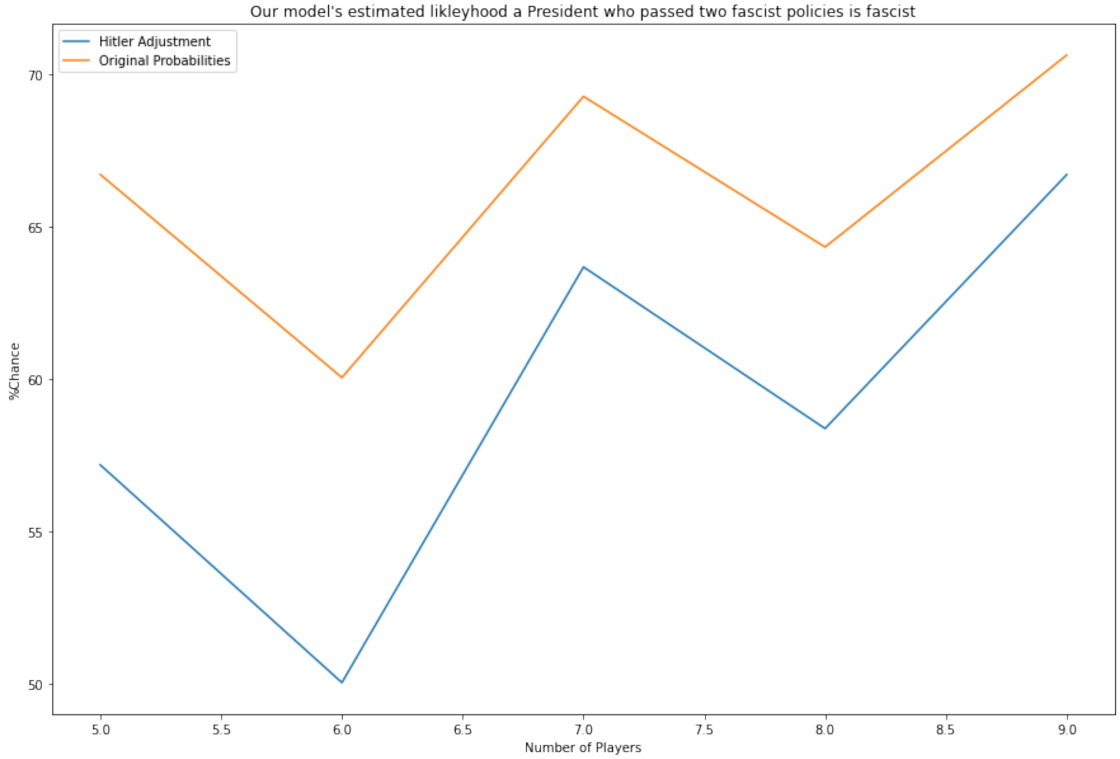
identity hidden by sometimes passing some liberal policies. Let α be the probability that a fascist president will act like a fascist. We update our original model to include α: $P(B|A)_\alpha = 0.242 + 0.485 * \alpha \implies P(A|B)_\alpha = \frac{P(A)(0.242+0.485*\alpha)}{0.242(1-P(A))+P(A)(0.242+0.485*\alpha)} = \frac{P(A)(0.242+0.485*\alpha)}{0.242+P(A)(0.485*\alpha)}$ Just as a sanity check, if we plug in $\alpha = 0$ we get $P(A|B)_\alpha = P(A)$ which makes sense if fascists and liberals play identical, the probability is just the probability of any random player being fascist.

- **Combined Adjustment** If we combine strategic balancing with the Hitler assumption, we get $P(B|A)'_\alpha = 0.242 + 0.485 * \alpha * \frac{P(A)-\frac{1}{n}}{P(A)} \implies P(A|B)'_\alpha = \frac{P(A)(0.242+0.485*\alpha*\frac{P(A)-\frac{1}{n}}{P(A)})}{0.242+P(A)(0.485*\alpha*\frac{P(A)-\frac{1}{n}}{P(A)})}$

[113]:
```python
def plot_2fascist_hitler_adj(n):
    hitler_pas = []
    base_pas = []
    for i in range (5,n):
        if(i % 2 == 0):
            p_a = (int(i/2) -1)/i
        else:
            p_a = (int(i/2))/i

        pab_prime = 0.242 + 0.485 * (p_a - 1/i)/p_a
        hitler_pas.append((pab_prime * p_a)/(0.242 - 0.242 *p_a + p_a *␣
↪pab_prime))
        base_pas.append(0.727 * p_a / (0.242 + 0.485 * p_a))

    plt.figure(figsize=(15,10))

    base_pas = pd.Series(base_pas) * 100
    hitler_pas = pd.Series(hitler_pas) * 100
    df = pd.DataFrame()
    df["hitler_adj"] = hitler_pas
    df["Original"] = base_pas

    df.index = index = range(5,n)
    plt.plot(df)
    plt.title("Our model's estimated likleyhood a President who passed two␣
↪fascist policies is fascist")
    plt.xlabel("Number of Players")
    plt.ylabel("%Chance")
    plt.legend(['Hitler Adjustment',"Original Probabilities"])
    plt.show()

def plot_2fascist_adj(n,alphas, Hitler = False):
    plt.figure(figsize=(15,10))

    df = pd.DataFrame()
```

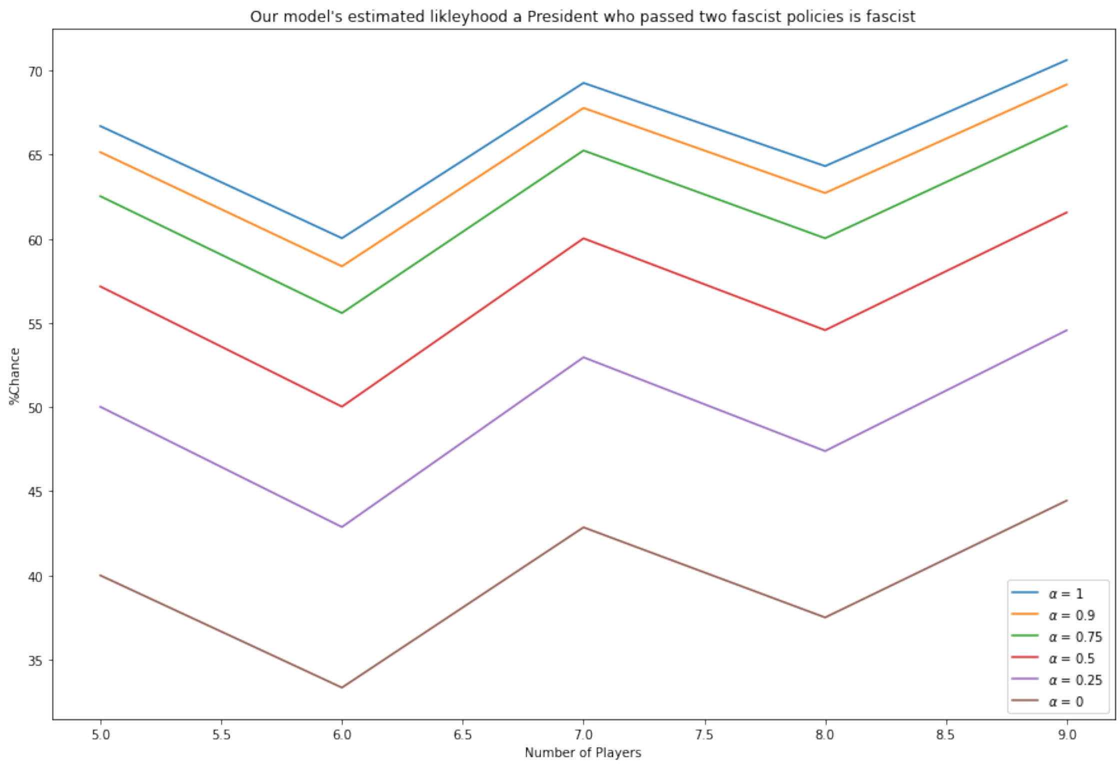
```
    for a in alphas:
        hitler_pas = []
        base_pas = []

        for i in range (5,n):
            if(i % 2 == 0):
                p_a = (int(i/2) -1)/i
            else:
                p_a = (int(i/2))/i

            pba_prime = 0.242 + 0.485 * a *  (p_a - 1/i)/p_a
            hitler_pas.append((pba_prime * p_a)/(0.242 - 0.242 *p_a + p_a *␣
↪pba_prime))
            base_pas.append( (p_a * (0.242 + 0.485 * a)) / (0.242 + p_a * (0.
↪485 * a)))

        if(not Hitler):
            df["alpha = "+ str(a)] = pd.Series(base_pas) * 100
        elif(Hitler):
            df["alpha = "+ str(a)] = pd.Series(hitler_pas) * 100

    df.index = index = range(5,n)
    plt.plot(df)
    asm_str = ''
    if(Hitler):
        asm_str = ' + (Hitler Assumption)'
    plt.title("Our model's estimated likleyhood a President who passed two␣
↪fascist policies is fascist" + asm_str)
    plt.xlabel("Number of Players")
    plt.ylabel("%Chance")
    legend = []
    for a in alphas:
        legend.append(r"$\alpha$ = "+ str(a))
    plt.legend(legend)
    plt.show()
```

[114]:
```
plot_2fascist_hitler_adj(10)
```

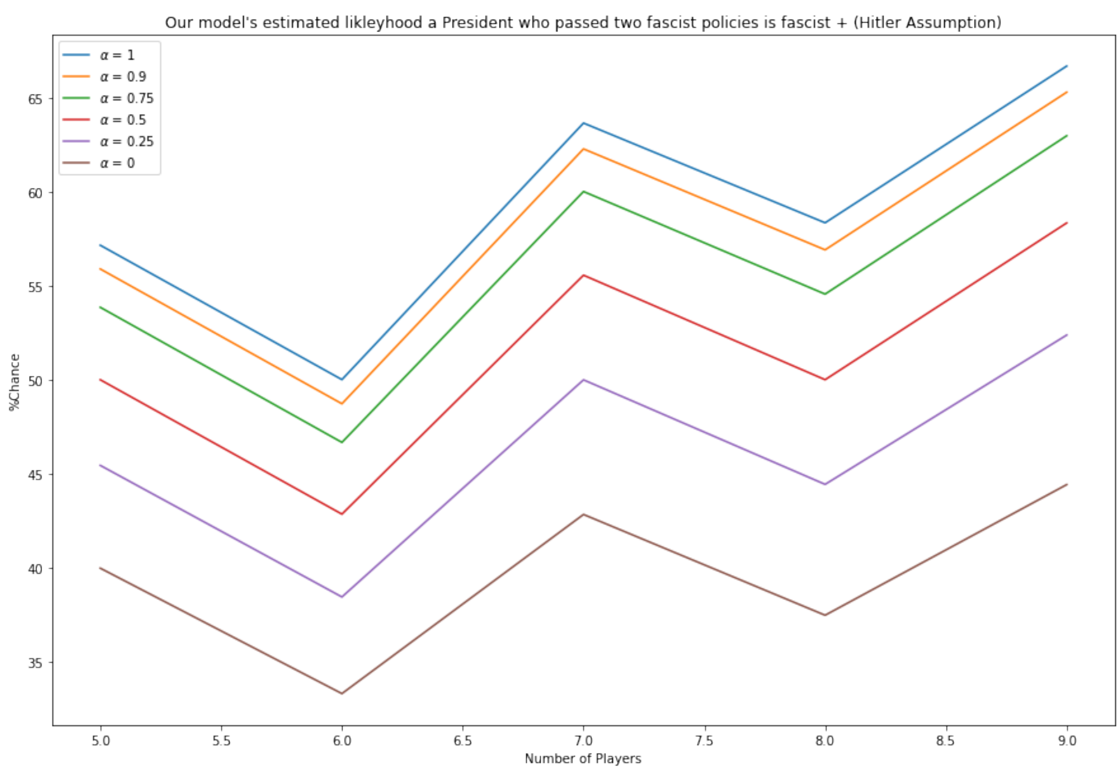


As expected, the Hitler adjustment reduces the percentages, and is more impactful with a smaller player pool as Hitler represents a larger portion of the total players.

[111]:
```
plot_2fascist_adj(10,[1,0.9,0.75,0.5,0.25,0], Hitler = False)
```




```
[112]: plot_2fascist_adj(10,[1,0.9,0.75,0.5,0.25,0], Hitler = True)
```



```
[ ]:
```