Exploring Blackjack Strategies & Variations

Benson Ngai

April 2025

1 Introduction

Blackjack is one of the most popular and commonly played casino games, where an individual player plays against a dealer, representing the "casino" or "house". The player's goal is to try and obtain a set of cards (by drawing cards from the deck) where their "hand total", the total value of the set of cards, is as close to 21 as possible without going over ("busting"), and trying to beat the dealer's hand total without busting. Although this seems like a game of random chance, Blackjack contains significant decision-making components that allow for optimal strategies to reduce the house's edge (the casino's profit off of you). These strategies can be derived through mathematical and experimental simulation, and are designed to minimize your losses in the long run. When you're playing these types of casino games, the goal is always to minimize loss, as the casino always has some edge / advantage over you for them to profit.

Motivation-wise, Blackjack an interesting problem at the intersection of probability, combinatorics, and decision-making. As I don't have much experience playing casino games, Blackjack is a simple enough game where where:

1. We know there exist **good** and **bad** strategies out there. It's not mathematically impossible to find an optimal strategy.
2. It can be easily simulated using Monte Carlo simulation to try and maximize reward/minimize loss in order to find optimal strategies.

Many people find 'chance' games like Blackjack purely random, when in fact there are mathematical ways to show otherwise. Beyond the game itself, the decision-making and statistical analysis explored in Blackjack extend to other domains such as risk-management and algorithmic trading. Although market trading operates in a dynamic environment with countless influencing factors and no fixed rules, unlike blackjack—a fixed, closed system with well-defined probabilities—both scenarios share a common principle. In both cases, decision-makers adjust their bet sizes based on the win/lose probabilities, thereby maximizing long-run growth and minimizing risk.

In this project, we focus on exploring what this "optimal playing strategy" is through computational Monte Carlo simulations. By simulating many Blackjack hands under various situations ("game states"), we can determine which strategic decisions in specific game scenarios yield the highest win rates. We aim to compare the effectiveness of this optimal strategy against other common-player strategies. In addition, our goal is to evaluate the robustness of our strategy under various variations of blackjack (since different casinos follow different Blackjack rules), such as modifications in deck composition and game structure (simplified action set vs complex action set).

2 Background

2.1 Rules

We first overview the basic rules of Blackjack, which are critical for our game simulation in this project. The player first receives two cards face-up drawn from a randomly shuffled deck, and the dealer receives two cards as well: one face-up (the "up card") and one face-down (the "hole card"). In Blackjack, card suits don't matter; only the value plays a role. Number cards 2-10 hold their respective numeric value, face cards (J, Q, K) represent 10, and Aces can either represent 1 or 11, depending on what is most beneficial for the hand total. A hand that contains an Ace can either be "hard" or "soft". A soft hand is such that the Ace represents an 11; however, if more cards are drawn ("hit") and busts, the Ace can now become a 1 to avoid the hand from busting. As a result, the soft hand has now become hard, since the Ace(s) in the hand only represent 1.

The player must first play by making decisions based off of their initial hand of two cards and the dealer's visible card. This is why the house has an edge; since the player has to make the first move, they can bust and lose immediately before the dealer even has to play. Further, as the dealer's second "hole" card is not visible, the player doesn't have complete information of how the dealer's hand total compares to their hand total, and therefore have to make decisions with partially known information.

The list of possible actions include:

1. **Hit**: Draw another card from the deck, which increases hand total.
2. **Stand**: Keep your current hand total (you're satisfied with it) and end your turn for the dealer to play.
3. **Double-down**: Double the initial bet, receive exactly one more card, and then stand.
4. **Split**: If the first two cards are of equal value, split them into two separate hands, each with its own bet. Play both hands.

5. **Surrender**: Forfeit half of your initial bet and end the hand immediately.

The dealer doesn't get to make their own decisions like you do; they must follow a set predetermined rules as determined by the casino. The universal rule is for dealers to continue to hit until reaching a hand total of at least 17. A common rule variation is for the dealer to either hit or stand on a "soft 17", where the hand contains an Ace representing an 11 and other cards that sum up to 6.

A particular "hand", or round, of the game starts by you placing your initial bet and the cards being dealt. If your initially dealt cards forms 21, this is a **natural blackjack**. If the dealer doesn't also have a natural blackjack, you profit your $[bet amount] * some multiplier, which is typically 1.5. If the dealer has a natural blackjack, you tie, which is called a "push", and your profit here is $0. If you don't have a natural blackjack by the dealer does (they can check their hole-card if their up-card is an Ace or 10), you lose your bet amount. If no natural blackjacks occur, you make your moves first. If you bust, you lose your bet amount. If you don't, the dealer then plays their move. If they bust, you profit your bet amount. If neither busts, the final hand totals are calculated, in which you can either win and profit by your bet amount, push (tie) and profit nothing, or lose and lose your bet amount.

2.2 Problem Formulation

Our primary goal involves finding the optimal strategy to play Blackjack by simulating many hands through Monte Carlo simulations in systemized approach. With the rules established in the previous section, we can first see that the entire game of Blackjack can be discretized into a series of "states" (**state-space**) S, where a **state** s is characterized by the player's current hand total (hard or soft) and the dealer's face-up card. More concretely, $s = (\text{player hand total}, \text{dealer up-card})$. Note that no matter what actions a player makes, while they have not busted, they are always somewhere in S. Next, for a particular state, the player can make an **action** a from the set of all possible actions A (**action-space**). **For the sake of simplicity, we first start by considering the set of two actions that a player can take in any** s**: 1) hit or 2) stand.** Intuitively, a strategy simply means: what action should I take at a particular situation (state)? As such, naturally, a **strategy** is defined as a mapping from each state to an optimal action (hit or stand), aiming to optimize our objective. Visually, each strategy can be represented as a state-action table. Strategies are represented in two separate tables for clarity, distinguishing between hard and soft hand totals, as shown in Figures 1 and 2. In both tables, the each (row, column) pair represents a particular state s in S. Note that J, Q, and K are missing columns on purpose as they are equivalent to 10. Further, 21 is a missing column as the only action you should do is "stand"; you can't get a better hand and will likely win!

	2	3	4	5	6	7	8	9	10	A
4										
5										
6										
7										
8										
9										
10										
11										
12										
13										
14										
15										
16										
17										
18										
19										
20										

Figure 1: Strategy (State-Action) Table for Hard Hands

Figure 1 shows the strategy table for hard hands. To clarify, a hard hand refers to a hand that either contains no Ace, or contains an Ace counted as 1, because counting it as 11 would result in a bust. A hand can become hard if it initially contained an Ace valued at 11 (making it soft), but after hitting, the Ace's value must be changed to 1 to prevent the hand from busting. Specifically, consider these scenarios:

1. **Case 1: Initially Hard Hand**. Your initial hand does not contain an Ace (otherwise, by default, it would default to 11 to benefit you in getting a higher hand total). Think of being dealt any 2 cards in the set of cards valued 2-10, with each card having 4 copies in the deck.

2. **Case 2: Soft-to-Hard Transition**. Your initial hand is soft. However, as you "hit", your hand busted, so your Ace went from an 11 to a 1, converting your hand from soft to hard.

Conversely, Figure 2 shows the strategy table for soft hands, which always contain an Ace counted as 11. Consider these soft-hand scenarios:

1. **Case 1: Initially Soft Hand**. You are dealt an Ace (representing an 11) and some other card, which could also happen to be an Ace. If you had two Aces, this would produce 11 + 1, a soft 12. The highest hand total possible, disregarding a natural blackjack here, would be 11 + 9 = soft 20.

2. **Case 2: Hard-to-Soft Transition**. Your initial hand is hard, but after hitting, you draw an Ace that can be valued at 11 to help benefit your hand

total without busting. This converts your hand from hard to soft. Note that this scenario only occurs with initial hard hands totaling between 4 and 9, as adding an Ace would elevate your hand to [15, 20], resulting in the soft strategy table. Any higher ranges of a hard hand total would either result in a 21 (stand), or a bust!

	2	3	4	5	6	7	8	9	10	A
12										
13										
14										
15										
16										
17										
18										
19										
20										

Figure 2: Strategy (State-Action) Table for Soft Hands

From the cases above, we can see a transition from some state $s \in S_{soft}$ to some other state $s' \in S_{hard}$, or even a transition from some state $s \in S_{soft}$ to $s' \in S_{hard}$, where our complete state space $S = S_{hard} \bigcup S_{hard}$. These observations made here are extremely important, and will be referred to later to create the simulation and determine our optimal strategy.

With these definitions, our goal is to find the optimal strategy, which in simpler terms means: how can we fill out each of these cells (each representing a state s) with a corresponding action $a \in A$ to optimize our objective? To answer this, we need to define what exactly our objective is. Our **objective** is the maximization of expected profit over a large number of simulated hands, mathematically expressed as

$$\max_{\pi} \frac{1}{N} \sum_{i=1}^{N} p_i$$

where N represents the total number of simulated hands, p_i represents the profit or loss obtained in hand i, and $\pi : S \to A$ is a strategy mapping states to actions. The goal of our Monte Carlo simulation is to find the optimal strategy $\pi_{optimal}$ that achieves this maximum. To do so, we aim to simulate many hands under each of these possible states in both the hard and soft tables, applying each action $a \in A$ many times to find the optimal action for that cell. Through this method, we can iteratively build our strategy π that performs best in expectation and minimizes our losses the least.

3 Simulation Setup

Using Python, we constructed a Blackjack simulator that mirrors all of the aforementioned Blackjack rules, ensuring that the game's rules are accurately reflected to guarantee that the results from our simulations are correct. As previously mentioned, we started off simple by enforcing the following conditions:

1. **Deck Composition:** We implement only 1) single decks and 2) infinite decks. With a single deck, a hand is played using the limited 52 cards in a deck. With an infinite deck, a hand is played by drawing cards with replacement, where each card is drawn with a $\frac{1}{52}$ probability regardless of which cards are already on the table. We reshuffle the deck after every hand as we are not interested in card-counting mechanisms and keeping track of the high and low card counts remaining in the deck.

2. **Action Space:** At any time, the player is only allowed to either **hit** or **stand**. These are the two most basic Blackjack actions, so we aim to analyze strategies first using the reduced action space. Introducing surrender, double-down, and split involves more state-action tables (such as with split) and complicates things, so we hope to tackle this with time permitting.

With these conditions in mind, to simulate a hand, the simulator initializes the game by shuffling the deck, and dealing two cards face-up to the player and two cards (one face-up, one hole) to the dealer. The card-drawing process accounts for both single-deck and infinite deck scenarios as described above. By default, we first use 1 deck only, and use this setup for the rest of this paper unless otherwise specified.

The player's actions are determined based on the current hand total and the dealer's visible card and the player's strategy, referring to previously imaged state-action tables via a lookup-table. The dealer's actions are determined by the default dealer behavior. For a particular hand, a winner is decided, and the next hand is played. Thousands of hands are played using a particular strategy, which we cover in subsequent sections. Throughout the simulation, detailed metrics are tracked for each hand, such as: sequence of actions taken, player/dealer hands throughout the hand, final outcomes, and profits or losses incurred for that hand. We use these metrics to later analyze performance metrics systematically and compare strategies.

4 Optimal Strategy

4.1 Algorithm & Analysis

With our game simulation setup, we need to actually figure out what our optimal strategy π is! Precisely, how do we systematically fill our tables in Figures

1 and 2 with actions ("hit" or "stand") that maximize our objective O to maximize our profits through Monte Carlo simulations?

The idea is to tackle this in a backwards approach, starting from the terminal states of Blackjack, as we know what the best actions are for these states. Trivially, the clearest terminal state is when your hand total is 21. No matter whether you were dealt a natural blackjack, or you hit your way up towards a 21, hitting any more strictly increases your hand total, resulting in a bust. As such, the only other action available is to stand. Standing on 21 is a strictly non-negative action: 1) if the dealer doesn't have 21, you profit by your bet amount (+\$bet); 2) if the dealer has 21, this results in a push (tie), and you profit 0 (+\$0). Therefore, we've filled an entire row of our state-action table: no matter the dealer's up card, if your hand total is 21, stand.

Now, the idea of "working backwards" inductively is such that if you know the most winning action for a hand total of 21, you can now find the most winning action for a hand total of 20. This is because when you either hit or stand, your hand total can only either increase or remain the same, so when you have a hand total of 20, your next state s' is either 21 (by hitting an Ace to represent 1), busting (hitting anything > 1), or 20 (standing). Since we have filled in the future optimal state we could transition to, we have all the information we need to simulate hitting and standing thousands of times for your state $s = (20, \textit{dealer up card})$ and test which action results in a higher objective O, which is the expected profit calculated by summing all the profits and dividing by the number of trials. The action corresponding to a higher O will be chosen, and we can now move onto the next lower hand total, 19, and so forth.

However, finding the optimal actions for the remainder of our strategy state-action table is trickier than expected due to the complexity of the game in which the order of events in which different states appear, particular with both soft and hard hands, does not work in a highest to lowest hand total fashion. We refer to the aforementioned cases in Section 2.2, where we previously observed possible transitions of states between hard and soft tables, where $s \in S_{hard} \rightarrow s' \in S_{soft}$ and vice versa. Naively, we first attempted to fill out S_{hard}, the state-action table for hard hand totals; we did this in a reverse chronological fashion from highest hand total to lowest hand total, and for each state $s = (\textit{hand total, dealer up card})$, would simulating hitting /standing and playing out the remainder of that hand to determine the winner for thousands of iterations to determine the optimal action for that state. However, a counterexample arose: **hard hands totaling 4-9 resulted in looking up unfilled states** $s \in S_{soft}$. With any hard hand with total $\in [4, 9]$, drawing an Ace (11) resulted a hand $\in [15, 20] \in S_{soft}$, resulting as a soft hand.

The next natural approach was to fill in the soft state-action table in a similar reverse order, starting from the highest possible soft hand total of 20 (A, 9) down to the lowest possible soft hand total of 12 (A, A). From our previous observations, one can see that drawing sufficiently large cards would cause your hand to bust, turning these soft hands into hard hands, which are unfilled (optimal actions are unknown). More specifically, enumerating the possible transitions:

- *AA (Soft 12)* $\rightarrow$ *hit 10* $\rightarrow$ *Hard 12*
- *A2 (Soft 13)* $\rightarrow$ *hit 9,10* $\rightarrow$ *Hard 12, Hard 13*
- *A3 (Soft 14)* $\rightarrow$ *hit 8,9,10* $\rightarrow$ *Hard 12, Hard 13, Hard 14*
- ...
- *A9 (Soft 20)* $\rightarrow$ *hit 2,3,...9,10* $\rightarrow$ *Hard 12, Hard 13, ... Hard 19, Hard 20*

The cases above show that **for any soft hand pair in the form** $(A, X) = Sum$**, where** A **represents Ace,** X **represents some card, and** Sum **represents the hand total, if it becomes hard after drawing a sufficiently high card, the new hard hand total will be** $\in [12, Sum]$. With these observations, we can then construct an ordering to fill our hard and soft tables such that when running a simulation to determine whether hit or stand is optimal from state s, any possible sequence of hit events will transition your state to some other state s' that already has an optimal action determined (has been visited by our search algorithm). Specifically, the order is:

State-Action Table Traversal Algorithm:

1. Determine optimal action for terminal states, which is "Stand" for any hand total of 21 regardless of dealer face-up card.
2. Determine optimal action for hard hand totals $[20, 19, 18, ..., 10]$ in descending order. These totals can only either bust, hit 21, or transition to a higher value in this set.
3. Determine optimal action for soft hand totals $[20, 19, 18, ..., 12]$ in descending order, which could transition to step 2 if cards $\in [2, ..., 10]$ are hit.
4. Determine optimal action for remaining hard hand totals $[9, 8, 7, ..., 4]$ in descendeing order, which can either transition to step 2 if cards $\in [2, ..., 10]$ are hit, or step 3 if an Ace is hit.

The order specified above populates the strategy tables in a specific, deliberate order to avoid referencing uninitialized state-action pairs which we have not determined an optimal action for. To find our optimal strategy, running a sufficient number of **trials** to determine the optimal action (hit vs stand) for a particular state was paramount. Results from various amounts of trials are shown below.

4.2 10 Trials

Hard Totals State–Action Table

	2	3	4	5	6	7	8	9	10	A
4	hit	stand	hit	stand	hit	stand	hit	stand	stand	hit
5	stand	hit	hit	stand	stand	hit	stand	stand	stand	hit
6	stand	hit	hit	stand	hit	hit	stand	stand	hit	hit
7	stand	hit	hit	stand	stand	hit	stand	hit	stand	hit
8	hit	hit	stand	stand	hit	hit	hit	hit	stand	stand
9	hit	hit	hit	stand	hit	hit	hit	hit	hit	stand
10	hit	hit	stand	hit	hit	hit	hit	hit	hit	hit
11	stand	hit	hit	hit	hit	hit	hit	hit	hit	hit
12	stand	stand	stand	stand	hit	hit	hit	hit	stand	hit
13	stand	hit	hit	hit	stand	hit	hit	hit	hit	hit
14	stand	hit	stand	stand	stand	hit	hit	hit	hit	hit
15	hit	stand	stand	hit	hit	hit	stand	hit	hit	hit
16	stand	stand	stand	stand	hit	stand	stand	stand	stand	hit
17	stand	stand	hit	stand	stand	stand	hit	stand	stand	hit
18	stand	stand	stand	stand	stand	stand	stand	stand	stand	stand
19	stand	stand	stand	stand	stand	stand	stand	stand	stand	stand
20	stand	stand	stand	stand	stand	stand	stand	stand	stand	stand

Soft Totals State–Action Table

	2	3	4	5	6	7	8	9	10	A
12	hit	hit	stand	hit	hit	hit	hit	hit	hit	hit
13	stand	hit	hit	hit	hit	hit	stand	hit	stand	hit
14	stand	hit	hit	hit	hit	hit	hit	hit	hit	hit
15	hit	hit	stand	hit	hit	hit	hit	hit	hit	hit
16	hit	hit	hit	hit	hit	hit	hit	hit	stand	hit
17	hit	stand	stand	hit	hit	stand	hit	hit	hit	stand
18	hit	stand	stand	hit	stand	hit	stand	hit	hit	hit
19	stand	hit	hit	hit	stand	stand	stand	stand	hit	stand
20	stand	stand	stand	stand	stand	stand	hit	stand	stand	stand

Figure 3: Optimal Strategy Table showing optimal action to hit or stand, for any given game state, derived from a Monte Carlo simulation with 10 trials per game state. There are noticeable inconsistencies (”hit” vs ”stand”) between adjacent, nearly-the-same game states due to high variance from the limited number of trials.

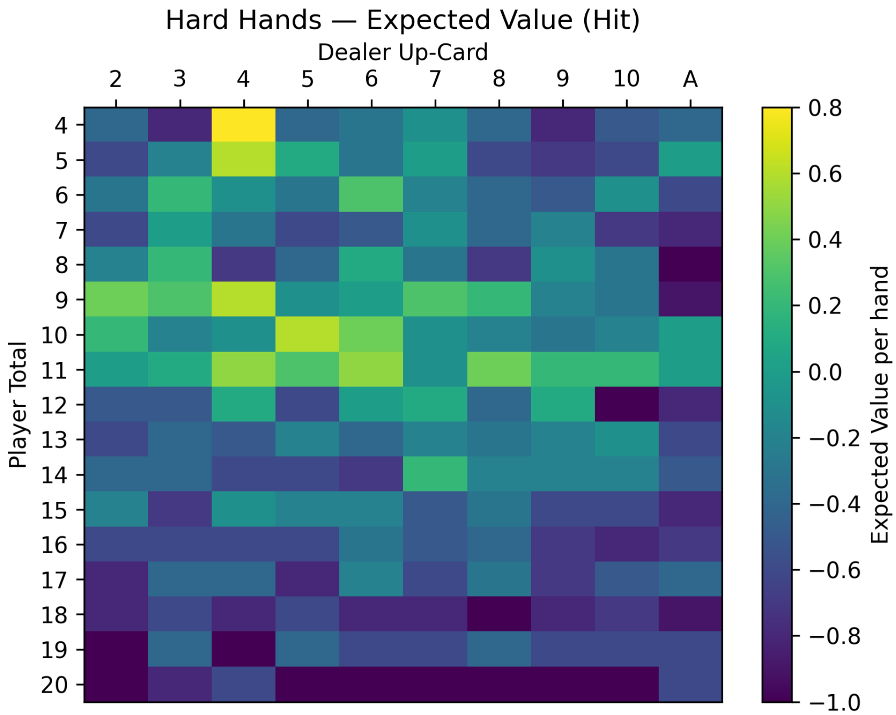

Figure 4: Hard Hands – Expected Profit (Hit). This heatmap displays the expected profit when hitting on hard hands based on 10 simulated trials per state.

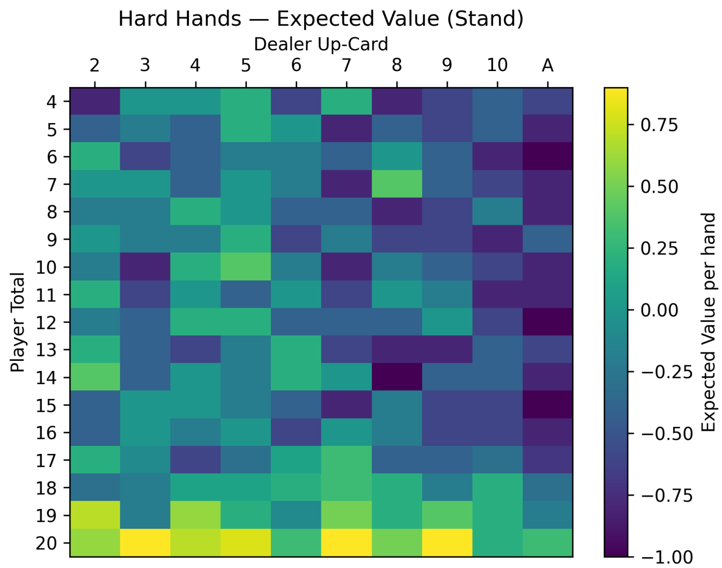

Figure 5: Hard Hands – Expected Profit (Stand). This heatmap displays the expected profit when standing on hard hands, based on 10 simulated trials per state.

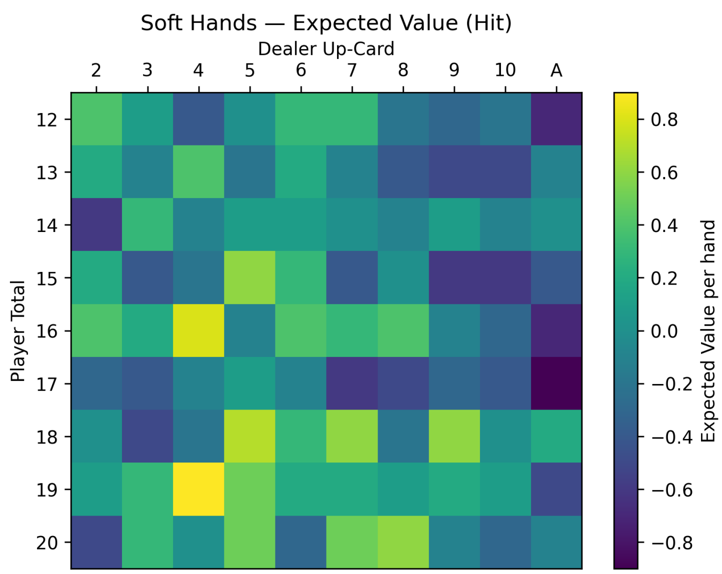

Figure 6: Soft Hands – Expected Profit (Hit). This heatmap displays the expected profit when hitting on soft hands, based on 10 simulated trials per state.

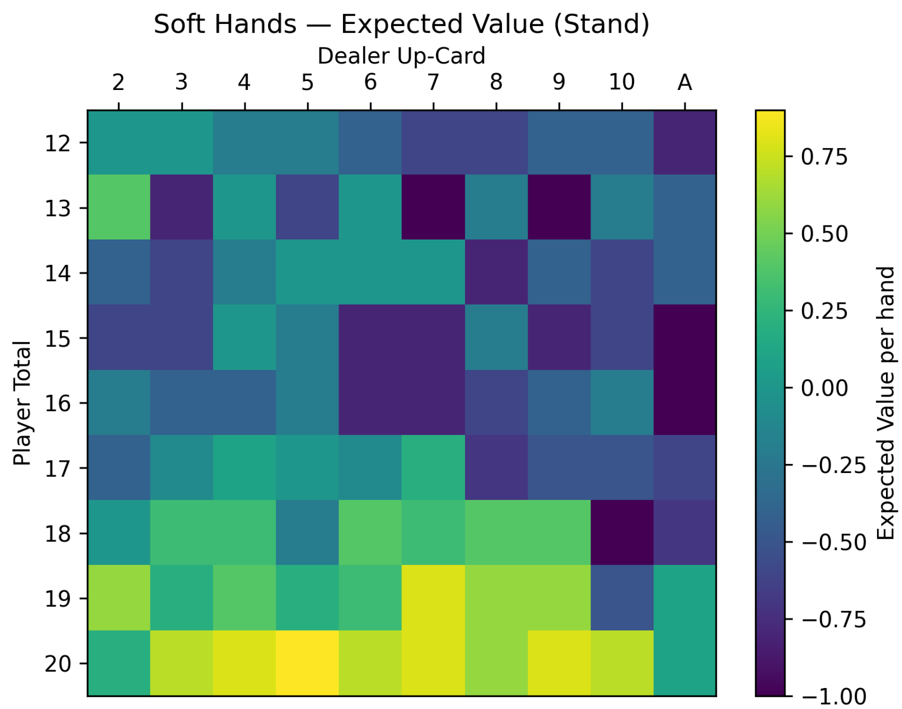

Figure 7: Soft Hands – Expected Profit (Stand). This heatmap displays the expected profit when standing on soft hands, based on 10 simulated trials per state.

4.3 100 Trials

Hard Totals State-Action Table

	2	3	4	5	6	7	8	9	10	A
4	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
5	hit	stand	hit	hit	hit	hit	hit	hit	hit	hit
6	hit	hit	hit	stand	stand	hit	hit	hit	hit	hit
7	stand	stand	hit	hit	hit	hit	hit	hit	hit	hit
8	hit	hit	hit	hit	stand	hit	hit	hit	hit	hit
9	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
10	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
11	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
12	hit	hit	stand	stand	stand	hit	hit	hit	hit	hit
13	hit	stand	stand	stand	stand	hit	hit	hit	hit	stand
14	hit	hit	stand	stand	stand	hit	hit	hit	hit	hit
15	stand	stand	stand	stand	stand	hit	stand	hit	hit	hit
16	stand	stand	stand	stand	stand	hit	stand	hit	stand	hit
17	stand	stand	stand	stand	stand	stand	stand	stand	stand	hit
18	stand	stand	stand	stand	stand	stand	stand	stand	stand	stand
19	stand	stand	stand	stand	stand	stand	stand	stand	stand	stand
20	stand	stand	stand	stand	stand	stand	stand	stand	stand	stand

Soft Totals State-Action Table

	2	3	4	5	6	7	8	9	10	A
12	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
13	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
14	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
15	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
16	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
17	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
18	stand	hit	stand	hit	hit	stand	stand	stand	stand	stand
19	stand	stand	stand	stand	stand	stand	stand	stand	stand	stand
20	stand	stand	stand	stand	stand	stand	stand	stand	stand	stand

Figure 8: Optimal Strategy Table showing optimal action to hit or stand, for any given game state, derived from a Monte Carlo simulation with 100 trials per game state. We begin to see a more smoothed out strategy table in the upper half of both the hard hand and soft hand tables, resulting in significantly less "stand"s than "hit"s with a lower hard/soft hand total. There appears to be a boundary transition from hit to stand in the middle of the hard hand total table, with more optimal actions classified as hit when the dealer face-up card is higher.

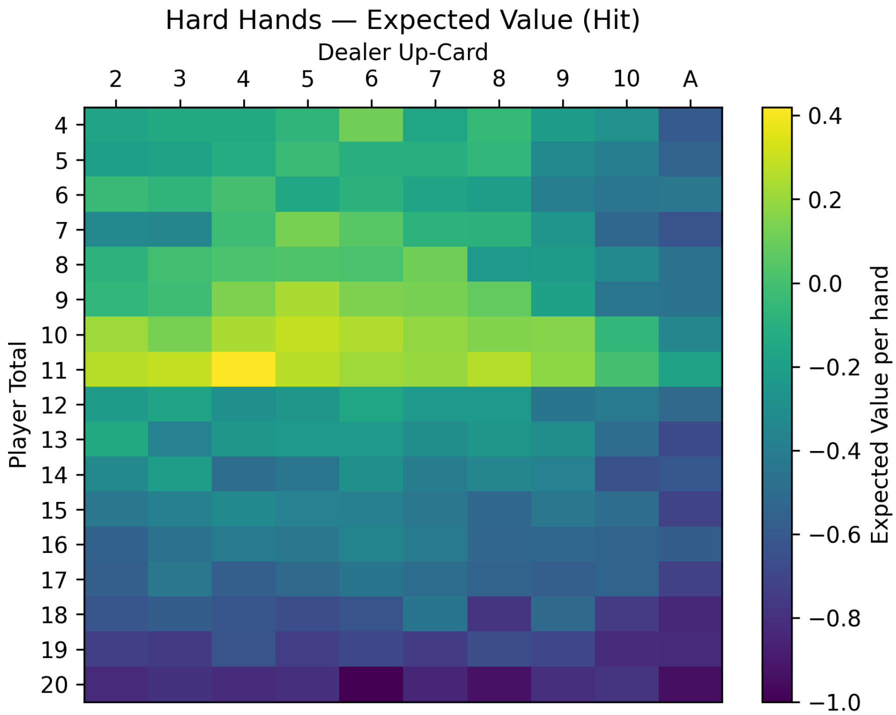

Figure 9: Hard Hands – Expected Profit (Hit). This heatmap displays the expected profit when hitting on hard hands based on 100 simulated trials per state. Compared to 10 trials, we a gradient developing on both the upper and bottom halves of the table. The upper half indicates a higher, positive expected profit when hitting on a low player total, with decreasing profits as the dealer up-card increases. The lower half is associated with a negative expected value per hand, as hitting on higher player total results in busting the closer you get to 21. There exists a clear boundary between making and losing money at the player hand total of 11/12.

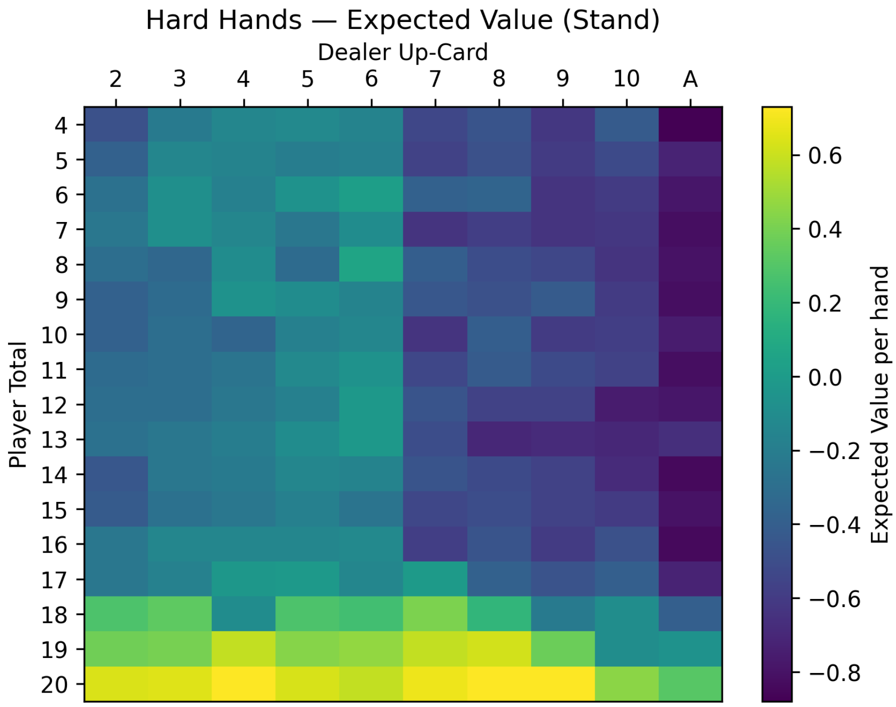

Figure 10: Hard Hands – Expected Profit (Stand). This heatmap displays the expected profit when standing on hard hands based on 100 simulated trials per state. Compared to 10 trials, the heatmap is similar such that the highest expected profits are around the player total of 19 to 20, which is clearly expected as hitting otherwise is extremely risky. We begin to see a vertical gradient at the dealer face-up card between 6 and 7, where cells to the left of this border suggest a slightly higher expected value than to the right. Standing when the dealer's face-up card is lower suggests that they are more likely to bust when they are dealt a lower face-up card.

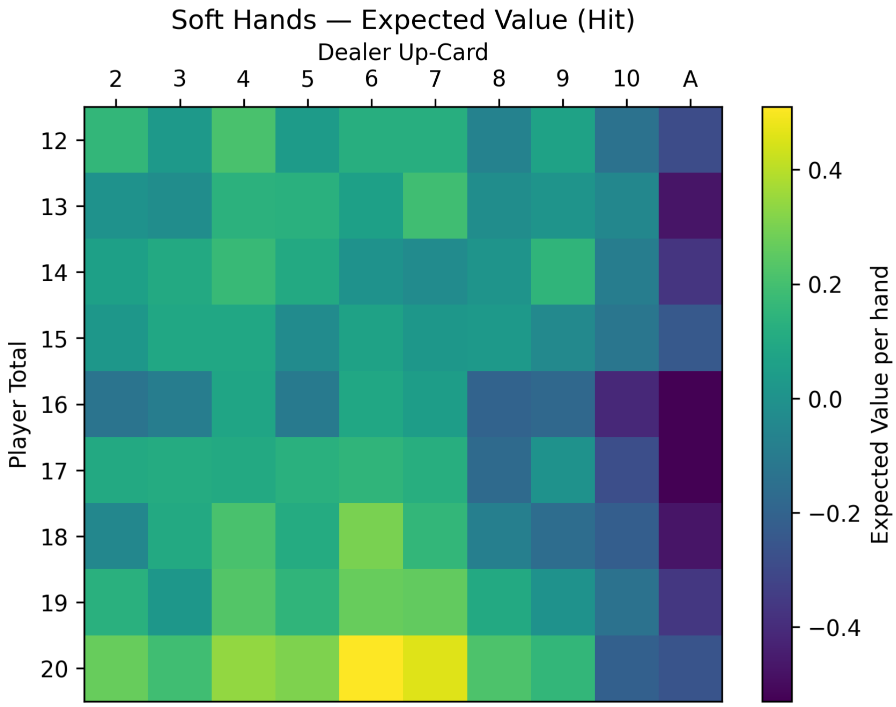

Figure 11: Soft Hands – Expected Profit (Hit). This heatmap displays the expected profit when hitting on soft hands based on 100 simulated trials per state. Compared to 10 trials, this heatmap is significantly more smooth, whereas Figure 6 looks like each cell contained a random expected profit. Here, we see a smooth horizontal gradient where you are more likely to make money (with the exception of player total 16), for any dealer face-up card from 2 to 8/9. However, at the dealer face-up card value of 8-10, we see a gradual transition towards breaking even or losing money.

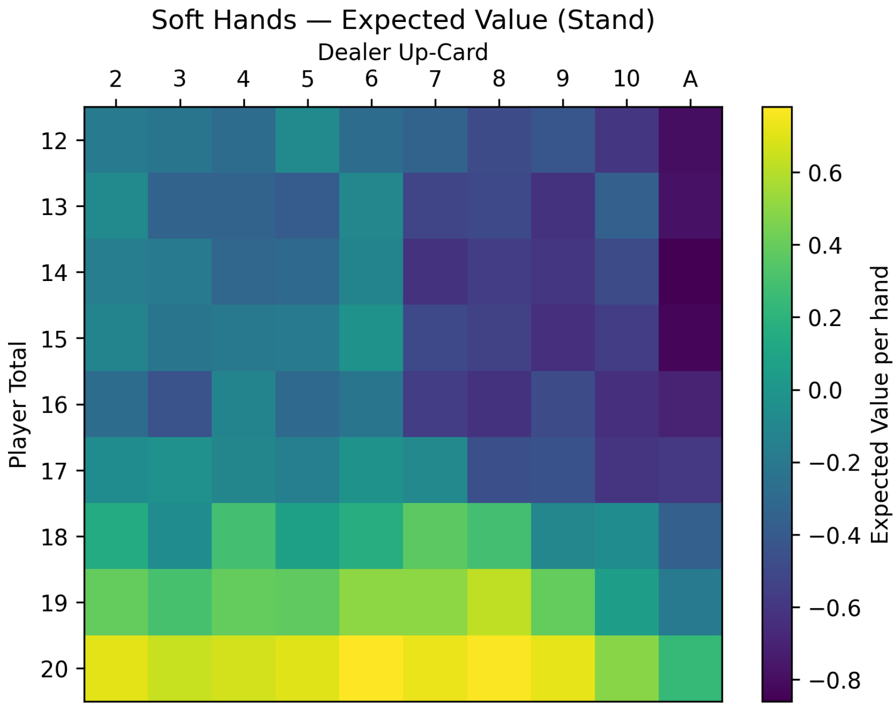

Figure 12: Soft Hands – Expected Profit (Stand). This heatmap displays the expected profit when standing on soft hands based on 100 simulated trials per state, and once again produces a smoother gradient than than the 10 trials heatmap. Similar to optimal behavior for hard hands, standing on a high player total of 18-20 yields the highest expected value. Standing on lower player hand totals results in mostly losing money regardless of the dealer's face-up card, suggesting that in these scenarios, the most feasible way to win is to have a high soft hand total and stand.

4.4 10000 Trials

Hard Totals State-Action Table

	2	3	4	5	6	7	8	9	10	A
4	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
5	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
6	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
7	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
8	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
9	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
10	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
11	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
12	hit	hit	hit	stand	hit	hit	hit	hit	hit	hit
13	stand	stand	stand	stand	stand	hit	hit	hit	hit	hit
14	stand	stand	stand	stand	stand	hit	hit	hit	hit	hit
15	stand	stand	stand	stand	stand	hit	hit	hit	hit	hit
16	stand	stand	stand	stand	stand	hit	hit	hit	hit	hit
17	stand	stand	stand	stand	stand	stand	stand	stand	stand	stand
18	stand	stand	stand	stand	stand	stand	stand	stand	stand	stand
19	stand	stand	stand	stand	stand	stand	stand	stand	stand	stand
20	stand	stand	stand	stand	stand	stand	stand	stand	stand	stand

Soft Totals State-Action Table

	2	3	4	5	6	7	8	9	10	A
12	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
13	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
14	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
15	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
16	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
17	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
18	stand	stand	stand	stand	stand	stand	stand	hit	hit	hit
19	stand	stand	stand	stand	stand	stand	stand	stand	stand	stand
20	stand	stand	stand	stand	stand	stand	stand	stand	stand	stand

Figure 13: Optimal Strategy Table showing optimal action to hit or stand, for any given game state, derived from a Monte Carlo simulation with 10000 trials per game state. Compared to the optimal strategy table for 100 trials, we see clear linear boundaries between hitting and standing. For hard hands, the general rule is to hit when your hand total is 12 or lower, or when it is 16 or lower and the dealer's face-up card is at least 7. Otherwise, standing is the optimal choice. For soft hands, the optimal move is to hit when your hand total is 17 or lower, or when it is 18 or lower and the dealer's face-up card is at least 9. We see interesting behavior where the player is encouraged to hit with a higher hand when the dealer also has a high-value up-card, suggesting that there is greater likelihood in getting closer to 21 and the dealer busting in these scenarios.

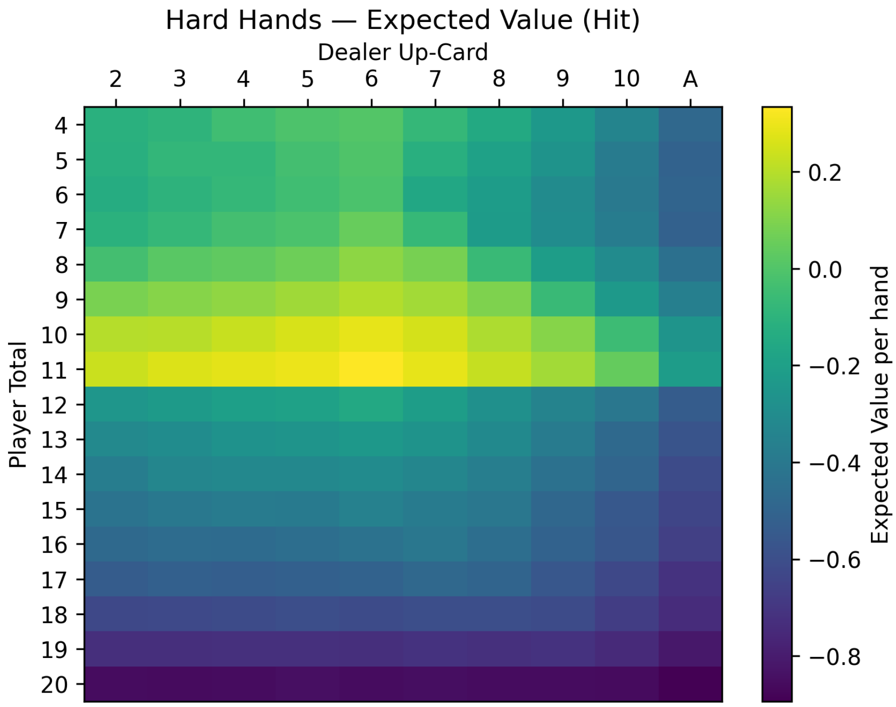

Figure 14: Hard Hands – Expected Profit (Hit). This heatmap displays the expected profit when standing on hard hands based on 10000 simulated trials per state, showing a significantly more smoothed our gradient than for 100 trials. Specifically, the upper half and lower half appear to be somewhat symmetric, where the highest expected profit for each half is at the middle of the heatmap, when your hand total is around 10-14 and the dealer's up-car is around 5-8. We observe a semi-circle like gradient for each half of the heatmap, where expected profit decreases as you traverse towards the sides and edges of the map (the extreme game states). There is a clear indication to hit when the player hard hand total is lower than 12, and the dealer's up card is less than 9.

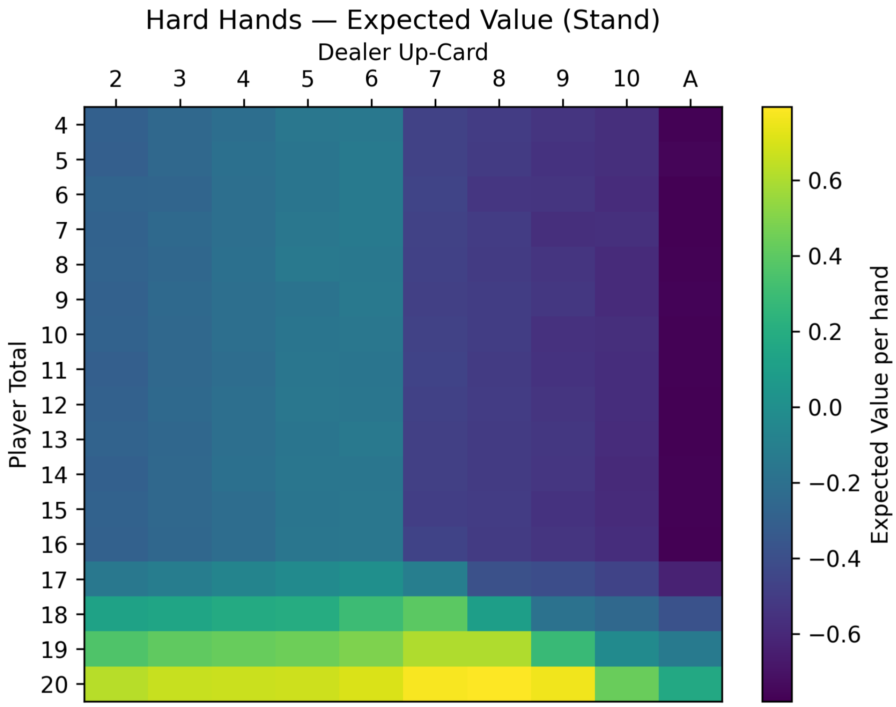

Figure 15: Hard Hands – Expected Profit (Stand). This heatmap displays the expected profit when standing on hard hands based on 10000 simulated trials per state, which indicates a profit when your hand total is generally greater than 17. We see a stair-step like behavior in the bottom right corner of the heatmap, where you lose less money the higher the dealer's face-up card is. For the most of the heatmap, standing on anything 17 or lower results in a negative expected profit.

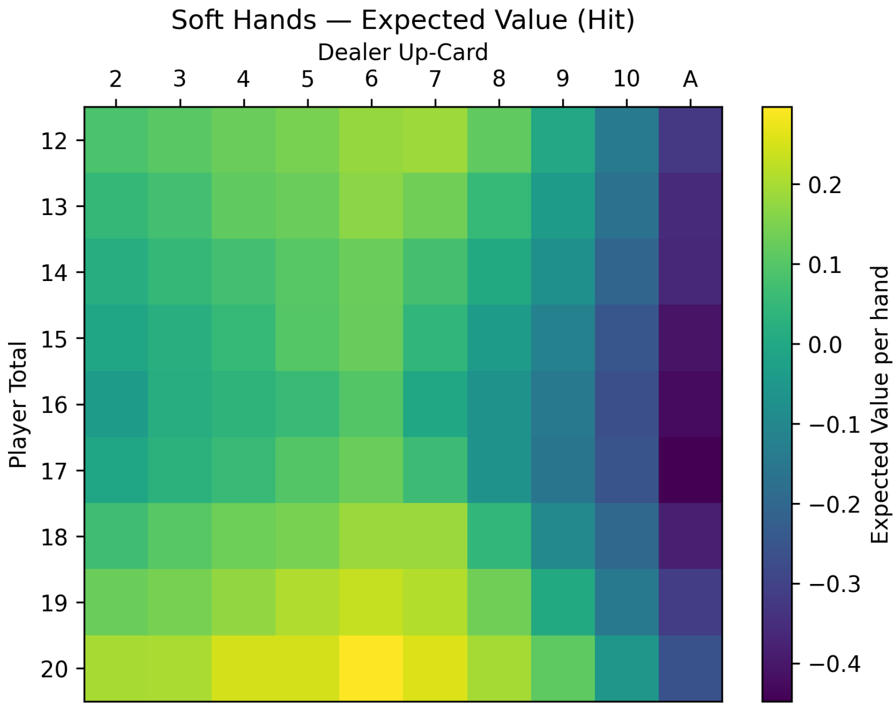

Figure 16: Soft Hands – Expected Profit (Hit). This heatmap displays the expected profit when hitting on a soft hand based on 10000 simulated trials per state. Compared to 100 trials, this heatmap is significantly smoothed out, displaying on an odd "X"-like pattern that splits the heatmap down the diagonals. Expected profits are positive generally for any player total where the dealer upcard is less than 8, and are higher in the yellow regions at the top and bottom portions of the map. We almost always lose money when the dealer face-up card is a 9, 10, or A.

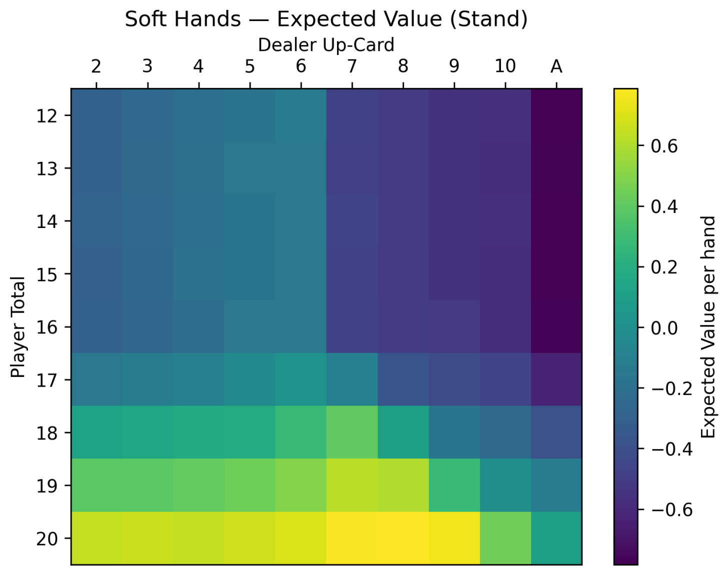

Figure 17: Soft Hands – Expected Profit (Stand). This heatmap displays the expected profit when standing on a soft hand based on 10000 simulated trials per state. Compared to 100 trials, we generally observe the same, but more smoothed out, pattern where the highest expected profits are standing at a high soft player hand total, and lose money everywhere else (losing more when the dealer's face-up card is higher).

4.5 Analysis Across k Trials

Different amounts of trials were tested for finding the optimal strategy to play Blackjack using our algorithm. Specifically, the number of trials refers to how many hands we simulated from a particular state (cell in the table). Through these trials, the profit for each these hands were recorded, and the expected value is calculated by summing our profits and dividing the number of hands played (trials). Examining the results above for 10, 100, and 10000 trials, we can clearly see differences in our resulting "optimal strategy" π in both the found state-action tables and the expected value heatmaps.

As the number of simulation trials increases, the strategy tables and corresponding heatmaps show progressively clearer convergence towards a less noisy, more interpretable pattern with more distinct boundaries. For example, in the 10-trial strategy, the state-action table as appears in Figure 3 appears noisy and inconsistent with some decisions varying unpredictably across similar hand totals. For example, standing on 11 when the dealer's up-card is a 2 versus hitting when the dealer's up-card is a 3, which in reality, should not make much of a difference. This can be attributed to the higher variance in our expected value profit estimates due to the small sample size of 10 trials being run for a particular cell. The corresponding heatmaps also reflect the noise and instability shown in this strategy table. For example, in Figure 4, even though the top half does appear to be higher than the bottom half, there is noise across the columns within each half. As our optimal action for earlier states depend on later states that can be transitioned to, this noise certainly contributed to and had a large, negative effect on the earlier states of the game.

When increasing $k = 100$, we begin to see a bit more structure in each of the heatmaps as well as the table. It appears that no matter the dealer's face-up card, hitting while your hard hand total is below 12 is generally a good idea. If the dealer's face-up card is higher (7 or above), we see that you should actually hit even at higher hard hand totals, such as at 17 or 18. For soft hands, we see that hitting anything below 18 is a good idea. In the heatmaps, we certainly see a significant reduction of noise as compared to $k = 10$ trials. In Figure 4, we start to see a clear separation at a hand total of 11 and 12, with peaks of expected value around the center of the bottom of the top half of the map.

Finally, we tested $k = 10,000$ trials to decrease the variance of our expected profits as much as possible. In Figure 13, we see a very stable and refined optimal strategy with clear boundaries between standing and hitting. Specifically, the table tells us to essentially always hit when your hard hand total ≤ 12, with the exception of one cell at $s = (12, 5)$. We see that if your hard hand total is ≤ 16, you should hit when the dealer's face-up card is ≥ 7. In regards to the soft hand, you should always hit when your soft hand total ≤ 17, with the exception of hitting at soft 18 when the dealer's face up card is either a 9, 10, or an Ace. The heatmaps, as compared to the lower k values, are certainly the most smooth and look much more like a continuous gradient. From this data, we can see that hitting is more favorable when the dealer shows strong face-up cards, while standing becomes more preferable when the dealer has a lower up-card. From our experiments, we can see the important of sufficient sample size when using Monte Carlo simulations to estimate our optimal strategy, as running more trials builds more statistical confidence and reduces variance of our objective O.

5 Suboptimal Strategies

Evaluating the effectiveness of our optimal strategy is the fun part – we wanted to be able to test how much more profit we could make (in expectation) compared to other strategies. To do so, we created a few simple strategies to represent typical ”non-expert” player strategies when playing Blackjack. Creating a strategy just involved populating that strategy's state-action table with the desired values that would match the behavior of someone using such a strategy. Some strategies we created were:

1. **Random Strategy:** At each state $s = (\textit{hand total, dealer up card})$, the player randomly chooses to either hit or stand with equal probability.

2. **Dealer Strategy:** Mimic the dealer. Always hit when hand total < 17 and stand when ≥ 17. Follow the dealer's strategy on hitting on soft 17, which in this case, they did not.

3. **Conservative Strategy:** Mimic a risk-averse player who doesn't want to bust. Stand when your hand total is ≥ 12 regardless of what the dealer's face-up card is.

4. **Aggressive Strategy:** Mimic a risky player. Hit until your hand total is ≥ 18.

We used the optimal strategy found from the $k = 10,000$ trials table to evaluate against these other strategies. Each strategy started with \$500, and all bets placed were \$1. Winning against the dealer made you a profit of +\$1, winning on a natural blackjack was +\$1.5, pushing (tying) was +\$0, and losing was −\$1. To evaluate each strategy, we simulated 10,000 hands per strategy, keeping track of the results/profits of each hand. These tracked metrics were helpful in computing some final evaluation results, such as:

1. **Expected Value (EV)**: the expected profit or loss per hand. Computed using O.

2. **House Edge**: Simply the negative expected value, which represents the casino's profit as a percentage of the player's original bet.

3. **Profit**: Overall profit or loss after the 10,000 hands played.

4. **Win Rate**: percentage of hands won.

5. **Loss Rate**: percentage of hands lost.

6. **Push Rate**: percentage of hands tied.

7. **Natural Blackjack Rate**: Percentage of hands where you won via a natural blackjack.

8. **Risk of Ruin**: whether the player when bankrupt (ran out of their initial \$500).

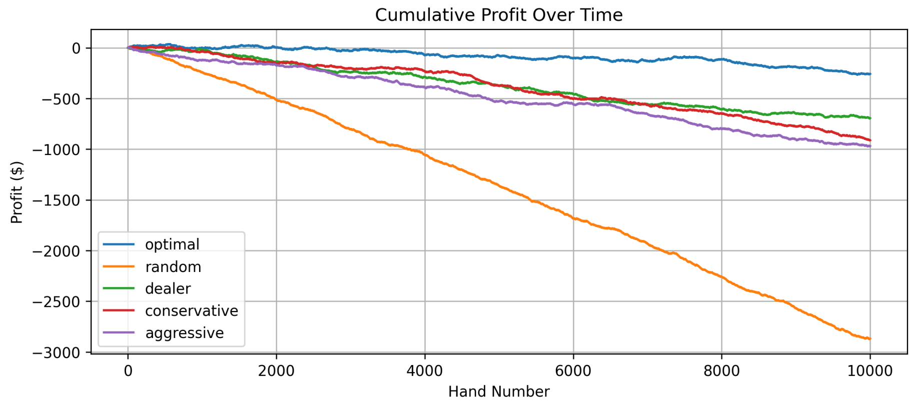

Figure 18: Cumulative Profit/Loss Over $10,000$ Hands Across Different Blackjack Strategies for Single Deck Blackjack

The cumulative profit of each of the strategies is shown in Figure 18. The first clear observation is that **no strategy was profitable**. Every strategy resulted in less money than they started with, supporting our hypothesis such that the casino has a significant edge over you based on the game setup of Blackjack (player has to play first, player doesn't have information on the dealer's hole card). From the graph, we can see that the random strategy, without surprises, loses significantly more money than the other strategies. Conservative and aggressive strategies yield relatively similar results, without sufficient evidence to prove one being better than the other. The dealer strategy performs the second best, while **our optimal strategy performs the best**, and is the only strategy that did not go bankrupt (only lost $\approx$ \$300).

Strategy	**EV**	**Win %**	**Loss %**	**Push %**	**Natural BJ %**	**Ruin @ \$500**
Optimal	-0.0258	43.2%	48.1%	8.7%	4.58%	No
Random	-0.2869	32.5%	63.3%	4.2%	4.37%	Yes
Dealer	-0.0695	40.6%	49.9%	9.6%	4.65%	Yes
Conservative	-0.0912	41.3%	52.5%	6.1%	4.21%	Yes
Aggressive	-0.0970	40.0%	52.0%	8.1%	4.75%	Yes

Table 1: Performance Comparison Across Different Blackjack Strategies for Single Deck Blackjack

From Table 1, we see that our optimal strategy results in a small expected loss per hand (-0.0258), meaning for every hand (\$1 bet), I would lose around 2.58 cents. For every 100 hands I played (\$100), I would lose \$2.58. The house edge here would be $-0.0258 * 100 = 2.58\%$, meaning the casino would profit roughly 2.58% of your initial bet per hand you play. This is an order of magnitude lower than random play, and still several times lower than the other strategies. We can see that our win percentage is only slightly percents higher for optimal strategy in comparison to dealer, conservative, and aggressive strategies, but clearly plays a large role in the long term as we lose significantly less money, as seen in Figure 18. If we mimic the dealer, this results in the highest rate of tying with them, which intuitively makes sense as our behavior is the same as theirs. As expected, no matter what strategy you play, the chance of you getting a natural blackjack is the same at roughly $4-5\%$.

6 Variations

Next, we explore how our optimal strategy performs under different variations of Blackjack when modifying deck composition and extending our simplified game structure to accommodate complex actions.

6.1 Infinite Decks

The previous results involve playing Blackjack with a single deck, where every new hand played is reshuffled. This allows us to simulate Blackjack without considering card counting strategies, as they would be ineffective in this continuous shuffling scenario, and the house edge remains constant since each hand is independent of previous hands. Now, we analyze playing Blackjack with infinite decks, where each card drawn is drawn with replacement. In this scenario, it is possible to get more than four copies of the same value card on same hand. We aim to analyze and compare this deck composition with the single-deck composition. We follow the same procedure as before to run 10000 trial Monte Carlo simulations to find the optimal strategy with infinite decks, and then simulate 10000 Blackjack hands using our strategy.

Hard Totals State-Action Table

	2	3	4	5	6	7	8	9	10	A
4	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
5	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
6	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
7	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
8	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
9	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
10	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
11	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
12	hit	hit	stand	stand	stand	hit	hit	hit	hit	hit
13	stand	stand	stand	stand	stand	hit	hit	hit	hit	hit
14	stand	stand	stand	stand	stand	hit	hit	hit	hit	hit
15	stand	stand	stand	stand	stand	hit	hit	hit	hit	hit
16	stand	stand	stand	stand	stand	hit	hit	hit	hit	hit
17	stand	stand	stand	stand	stand	stand	stand	stand	stand	stand
18	stand	stand	stand	stand	stand	stand	stand	stand	stand	stand
19	stand	stand	stand	stand	stand	stand	stand	stand	stand	stand
20	stand	stand	stand	stand	stand	stand	stand	stand	stand	stand

Soft Totals State-Action Table

	2	3	4	5	6	7	8	9	10	A
12	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
13	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
14	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
15	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
16	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
17	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
18	stand	stand	stand	stand	stand	stand	stand	hit	hit	hit
19	stand	stand	stand	stand	stand	stand	stand	stand	stand	stand
20	stand	stand	stand	stand	stand	stand	stand	stand	stand	stand

Figure 19: Optimal Strategy Table for Infinite Deck Composition, derived from Monte Carlo simulation with 10000 trials per game state.

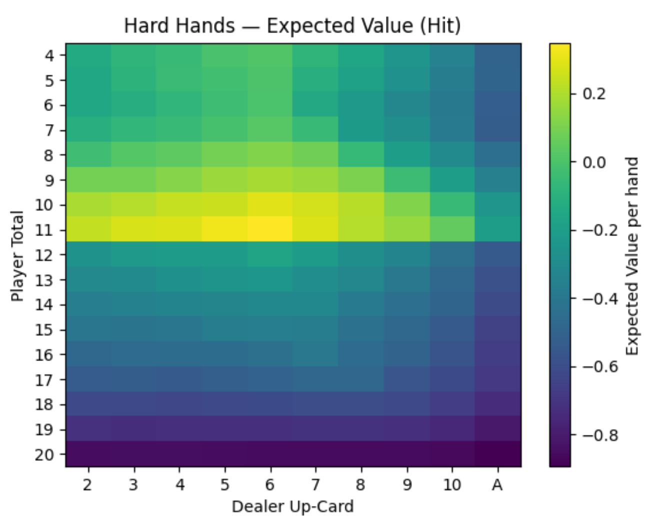

Figure 20: Hard Hands - Expected Profit (Hit) in Infinite Deck Blackjack. Highest profitability around player totals of 10-12 and a dealer up-card range of 4-6.

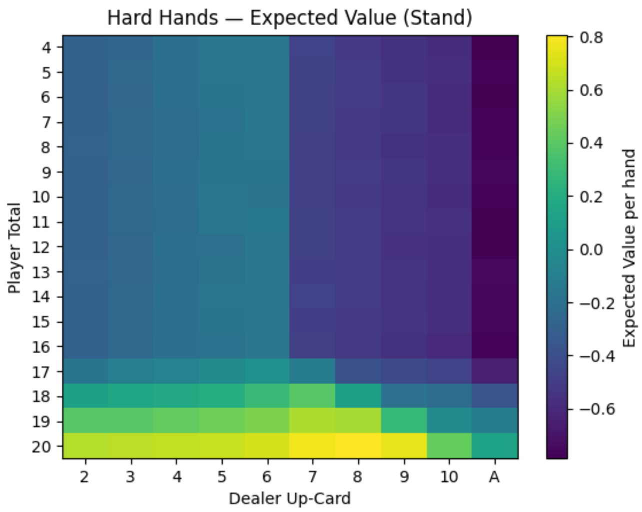

Figure 21: Hard Hands - Expected Profit (Stand) in Infinite Deck Blackjack, indicating profitability for player totals above 17, similar to single-deck Blackjack.

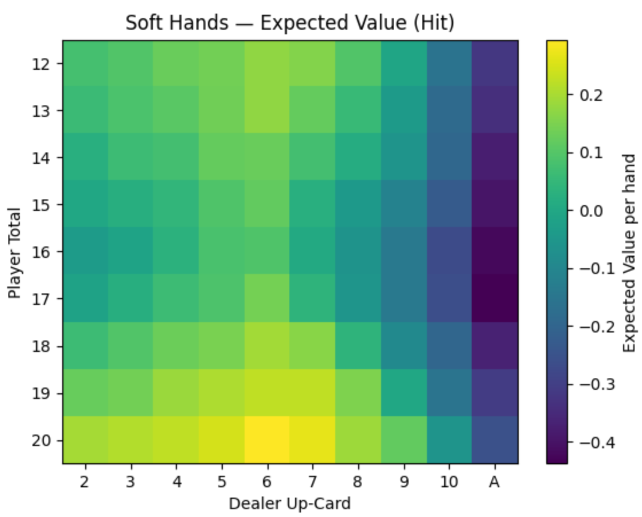

Figure 22: Soft Hands - Expected Profit (Hit) in Infinite Deck Blackjack, exhibiting largely the same heatmap patterns for the corresponding heatmap for single-deck Blackjack.

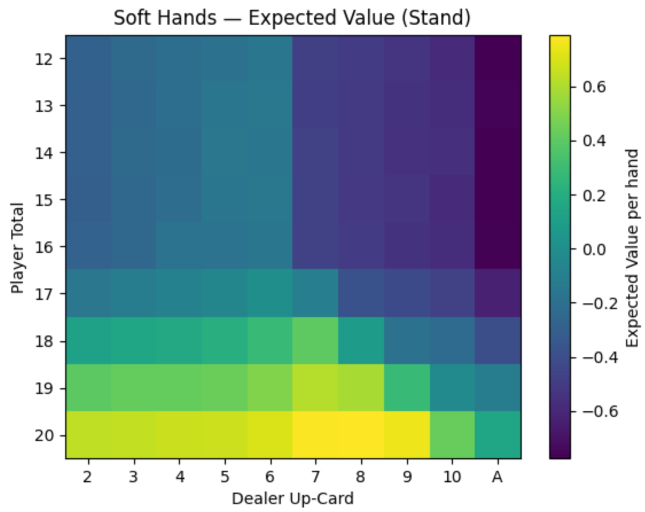

Figure 23: Soft Hands - Expected Profit (Stand) in Infinite Deck Blackjack, exhibiting largely the same heatmap patterns for the corresopnding heatmap for single-deck Blackjack.

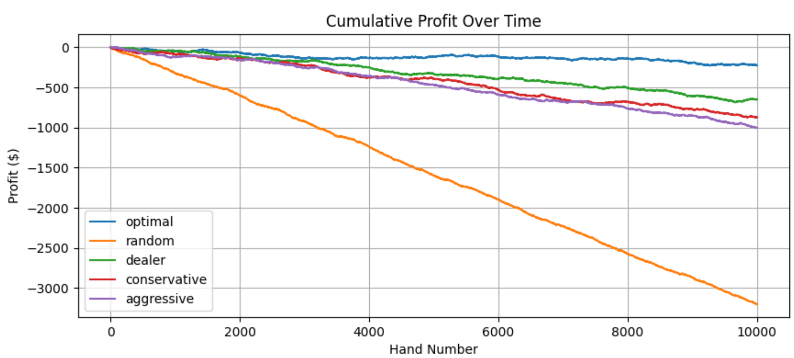

Figure 24: Cumulative Profit/Loss Over 10,000 Hands Across Different Blackjack Strategies on Infinite-Deck Blackjack. All strategies still lost money over time, with Optimal with -$232.5, Random with -$3200.5, Dealer with -$685.5, Conservative with -$871.0, and Aggressive with -$1007.0.

Strategy	EV	Win %	Loss %	Push %	Natural BJ %	Ruin @ $500
Optimal	-0.0224	43.5%	48.0%	8.56%	4.59%	No
Random	-0.3201	30.9%	65.1%	3.97%	4.33%	Yes
Dealer	-0.0648	40.7%	49.5%	9.81%	4.66%	Yes
Conservative	-0.0870	41.4%	52.4%	6.17%	4.66%	Yes
Aggressive	-0.1002	39.6%	51.8%	8.57%	4.42%	Yes

Table 2: Performance Comparison Across Different Blackjack Strategies for Infinite Deck Blackjack

We compare the results of our Infinite Deck Blackjack simulation with the Single Deck Blackjack simulation from before, both with optimal strategies determined from 10,000 trial Monte Carlo simulations and 10,000 Blackjack hands played. The optimal strategy table for Infinite Deck Blackjack (Figure 19) is mostly the same compared to Single Deck Blackjack (Figure 13), with the exception of standing when the player hard hand total is 12 and the dealer face-up card is 4 or 6. Otherwise, the boundaries between hitting and standing are exactly the same. The variation in the player hard hand total of 12 for dealer face-up cards of 4, 5, and 6 could suggest that this region is the most ambiguous area along the boundary to determine whether hitting or standing is more profitable. When examining the corresponding heatmaps in Figures 20 and 23, we can see that these cells yield roughly the same expected profit of ≈ -0.2 to -0.3; as such, either action approximately yields the same expected profit. Even more trials could be run to ensure maximal precision in determine whether hitting or standing is optimal for these states.
In both deck variants, the expected profits for hitting and standing for both hard and soft hands are mostly the same from a visual standpoint. The heatmaps all show the same gradient-like patterns. Turning to the performance of our optimal strategy against other suboptimal strategies in playing 10000 Blackjack hands with an infinite deck, we present quantitative results in Figure 24 and Table 3. The strategies' cumulative profit over time are in the same order as in single-deck Blackjack, where the best to worst was optimal, dealer, conservative, aggressive, and random. With infinite-deck Blackjack, the optimal strategy lost $\approx$ \$70 less money as compared to single-deck Blackjack, whereas the random strategy lost $\approx$ \$300 more. Dealer, conservative, and aggressive strategies all lost roughly the same amounts of money. Table 19 indicates that the optimal strategy experienced a slightly lower EV of -0.0224 as compared to -0.0258 in single-deck Blackjack, resulting in a 0.2% increase in win rate and push rate, and a 0.1% decrease in loss rate. The random strategy experienced the most drastic changes, with a $\approx 2\%$ decrease in win rate and $\approx 2\%$ increase in loss rate. From these results, we can see that playing with infinite decks slightly reduces the house edge from 2.58% to 2.24%, reflecting a minor yet measurable shift in player advantage due to the infinite availability of each card value when drawing with replacement. This results in more scenarios where it's possible for you to keep drawing lower cards without busting, as the probability of drawing them is independent per draw. Nevertheless, our results show an insignificant

change in optimal strategy between playing with a single deck and infinite decks, and that your expected profits (losses) roughly stay the same.

6.2 More Complex Actions

Our previous studies showed the results for playing simplified single-deck and infinite-deck Blackjack with only "hit" or "stand" as our possible actions in our action space. However, true Blackjack contains three more actions: double-down, split, surrender. We added these actions to our set of available actions to take, and outline how these were each implemented as part of discovering our optimal strategy.

Double-Down
To double-down, a player doubles their initial bet and draws exactly one more card. Afterwards, they must stand. To incorporate this into our existing algorithm, this was simply added as an option to our action space, as "double-down" can now be an optimal action for a particular state in any cell of our hard or soft optimal strategy tables. Specifically, this action was available only at the beginning of the hand when the player had two cards exactly. To double-down, the player's bet was doubled (from 1 to 2), an additional card was drawn, and the procedure to determine the winner between the player and dealer were the same as before.

Surrender
To surrender, the player forfeits half of their initial bet immediately after seeing their initial hand and the dealer's face-up card. Similar to how we compared other actions, "surrender" was added as an action to the action space, and its expected value was compared against alternative actions (hit, stand, double-down) to observe which action was the best when constructing our optimal strategy table.

Split
To split, the player's initial two cards must have the same value. This creates two separate hands, each with their own bet, and both hands are played out to determine the winners and payouts of both hands. Incorporating this action into our algorithm is not as simple as the others, since "split" is not an optimal action we can choose to put in our optimal hard and soft strategy tables. For example, when you receive two 5's, treating this as a hard hand total of 10 doesn't make sense, as these are being separated into two hands, and additional cards will be dealt to create two new full hands (two cards each). Thus, once we split, our new hands can transition to be played using our optimal hard / soft strategy tables, but we must define a separate state-action table for pair-splitting. Specifically, this state-action table is defined in Figure 25.
Each cell in Figure 25, instead of an action, should simply be either "Pair" (P) or "No Pair" (N). No matter whether you choose to split or not, each of your hands can now be played using the optimal strategy hard/soft tables found pre-

	2	3	4	5	6	7	8	9	10	A
A, A										
2,2										
3,3										
4,4										
5,5										
6,6										
7,7										
8,8										
9,9										
10,10										

Figure 25: Strategy (State-Action) Table for Split Pairs

viously. Specifically, if you decide to split your pair, additional cards will be drawn, and each of your hands can be played using the optimal hard/soft hand strategy. If you don't decide to split your pair, you can treat your current initial hand as a regular hard/soft hand, which can also be played using the optimal hard/soft hand strategy. As such, to determine the optimal strategy for split pairs, we first fill our optimal hard and soft state-action tables. Then, to fill out split-pairs strategy table, we consider each possible split-pair and dealer face-up card combination, and simulate 10000 trials of Monte Carlo simulation for that particular state to determine whether more split or unsplit pairs were won, for that particular pair of cards. The optimal action ("Pair" or "No Pair") is then chosen as the action that yielded the highest expected value.

With the three complex actions available to the user, we determined the optimal strategy for our hard hands table, soft hands table, and split-pairs table with the same procedure as before, by running 10000 trials per game state. Then, we simulated 10000 Single Deck Blackjack hands to compare various strategies under this new action space. We populated the split-pairs strategy table for each of the suboptimal strategies to mimic the nature of that strategy:

1. **Random Strategy:** At each state s = (*split pair, dealer up card*), the player randomly chooses to either split pair ("P") or not split ("N") with equal probability.

2. **Dealer Strategy:** Mimic the dealer. As the dealer cannot split, all cells are "N".

3. **Conservative Strategy:** A risk-adverse player would want to split if the resulting hands were surely stronger than the paired hand, as they could potentially lose twice the amount of money when splitting the pair. The

clearest case is to split A, A; the two resulting hands will become hands in the range $[12, 21]$, which is better than a single soft hand worth 12, as there is an increased chance of hitting until a hand total close to 21. Otherwise, this player does not split.

4. **Aggressive Strategy:** A risky player would choose to split often to make twice the amount of money. As such, this player splits for all card pairs besides $10, 10$, as standing is clearly optimal for a hard 20.

Hard Totals State-Action Table

	2	3	4	5	6	7	8	9	10	A
4	hit	stand	stand	hit	hit	hit	hit	hit	hit	hit
5	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
6	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
7	hit	hit	hit	hit	hit	hit	hit	hit	hit	hit
8	hit	hit	hit	hit	double	hit	hit	hit	hit	hit
9	double	double	double	double	double	hit	hit	hit	hit	hit
10	double	double	double	double	double	double	double	double	hit	hit
11	double	double	double	double	double	double	double	double	double	hit
12	hit	hit	hit	stand	stand	hit	hit	hit	hit	hit
13	stand	stand	stand	stand	stand	hit	hit	hit	hit	hit
14	stand	stand	stand	stand	stand	hit	hit	hit	hit	hit
15	stand	stand	stand	stand	stand	hit	hit	hit	hit	hit
16	stand	stand	stand	stand	stand	hit	hit	surrender	surrender	surrender
17	stand	stand	stand	stand	stand	stand	stand	stand	stand	stand
18	stand	stand	stand	stand	stand	stand	stand	stand	stand	stand
19	stand	stand	stand	stand	stand	stand	stand	stand	stand	stand
20	stand	stand	stand	stand	stand	stand	stand	stand	stand	stand

Soft Totals State-Action Table

	2	3	4	5	6	7	8	9	10	A
12	hit	hit	double	double	double	hit	hit	hit	hit	hit
13	hit	hit	double	double	double	hit	hit	hit	hit	hit
14	hit	hit	double	double	double	hit	hit	hit	hit	hit
15	hit	double	double	double	double	hit	hit	hit	hit	hit
16	hit	hit	double	double	double	hit	hit	hit	hit	hit
17	double	double	double	double	double	hit	hit	hit	hit	hit
18	double	double	double	double	double	stand	stand	hit	hit	stand
19	stand	stand	stand	stand	stand	stand	stand	stand	stand	stand
20	stand	stand	stand	stand	stand	stand	stand	stand	stand	stand

Split Strategy Table

Pairs	2	3	4	5	6	7	8	9	10	A
A,A	P	P	P	P	P	P	P	P	P	P
2,2	P	P	P	P	P	P	P	P	P	P
3,3	P	P	P	P	P	P	P	P	P	P
4,4	N	P	P	P	P	P	P	P	P	N
5,5	N	N	N	N	N	N	N	N	N	N
6,6	P	P	P	P	P	P	P	N	N	N
7,7	P	P	P	P	P	P	P	P	P	N
8,8	P	P	P	P	P	P	P	P	P	P
9,9	P	P	P	P	P	P	P	P	P	P
10,10	N	N	N	N	N	N	N	N	N	N

Figure 26: Optimal Strategy Table for Single Deck Composition with Complex Actions, derived from Monte Carlo simulation with 10000 trials per game state.

From the optimal strategy above, we observe noticeable changes from our previous optimal strategy tables that solely involved "hit" and "stand". For the hard totals table, we see an abundance of double-downing when your hand total is either 9, 10, or 11, and the dealer hand total is less than 7. This also applies for higher dealer face-up cards like 7, 8, 9, and 10, where we double down for player hand totals of 10 and 11. In these regions, doubling-down being chosen as the optimal action is clearly associated with a high probability of winning, as we're more confident that placing more risk will result in a benefit. As such, this suggests that in these states, the dealer could be more likely to bust while we end up with a hand total near 21. With soft hands, a majority of double-downs are played when the player total is anywhere from 12-18, and when the dealer's face-up card is either a 4, 5, or 6. In these scenarios, hitting on the next card will not bust our hand as it is soft, so we have a higher chance of drawing cards that lead us to near-21 hand totals.

As regards to surrendering, we notice three "surrender" cells in the hard totals table, where the player total is 16 and a the dealer face-up card is a high value, such as 9, 10, or A. Clearly, we're likely to be in trouble for these cells, as we'd rather lose half of our money than continue playing. This is because hitting could very likely result in busting our hand (if we draw anything greater than 5), and as the dealer has a high face-up card, they have a good chance of either already having a high hand total, or drawing a card to achieve a greater total than the player's hand total.

Examining the Split Strategy table, we notice a few patterns: we never split 5's or 10's. This makes sense in the latter case, where our hand total is already a 20, making it likely for us to stand and win. In the former case, a hand total of 10 yields greater returns than splitting them into two hands of 5 each based on our optimal hard/soft strategy tables. We also observe an increase in the number of no split-pairs for pairs 4, 5, and 6 when the dealer's face-up card is a 9, 10, or A. This suggests that against strong dealer up-cards, breaking a low pair into two weak hands actually lowers our expected profit. Since the dealer is unlikely to bust, it's better to keep the combined total rather than split into two inferior hands.

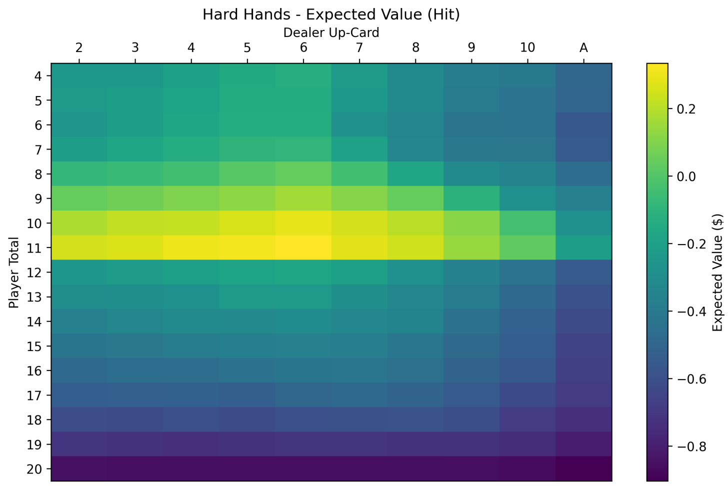

Figure 27: Hard Hands - Expected Profit (Hit) in Single Deck Blackjack with Complex Actions

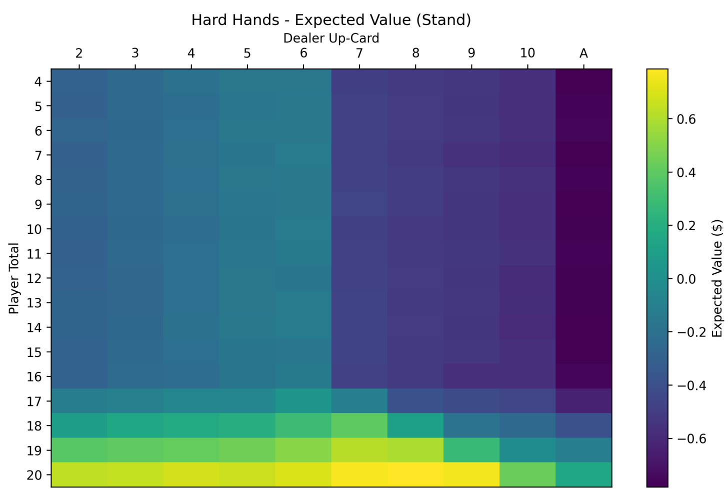

Figure 28: Hard Stands - Expected Profit (Stand) in Single Deck Blackjack with Complex Actions

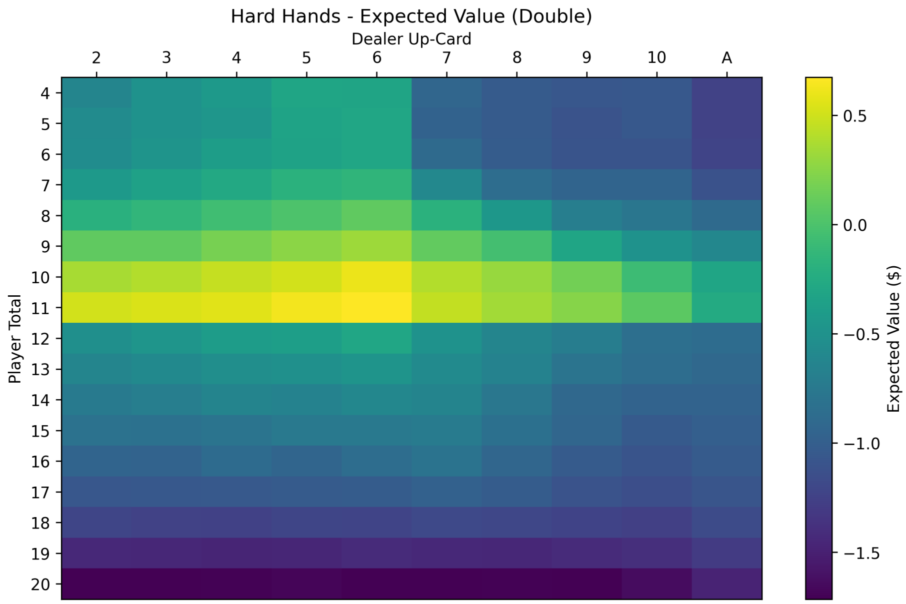

Figure 29: Hard Double Down - Expected Profit (Double Down) in Single Deck Blackjack with Complex Actions

Examining the hard hand heatmaps (Figures 27 - 29), we notice that the "hit" heatmap displays a familiar semi-circular peak around player totals of 10-12 against dealer face-up cards 4-6, similar to the simpler model (optimal strategy of only "hit" or "stand") as previously displayed. The "stand" heatmap likewise reproduces the steep vertical gradient favoring standing on player hand totals ≥ 17 and against low dealer face-up cards, which is also similar to our previous simpler model's heatmaps. Doubling-down is highly similar to the hit heatmap but with higher magnitude in the range where the player's hand total was 9-11 against dealer's face-up cards of 3-6, highlighting that doubling-down is simply leveraging those same high expected-profit "hit" scenarios to make extra profit.

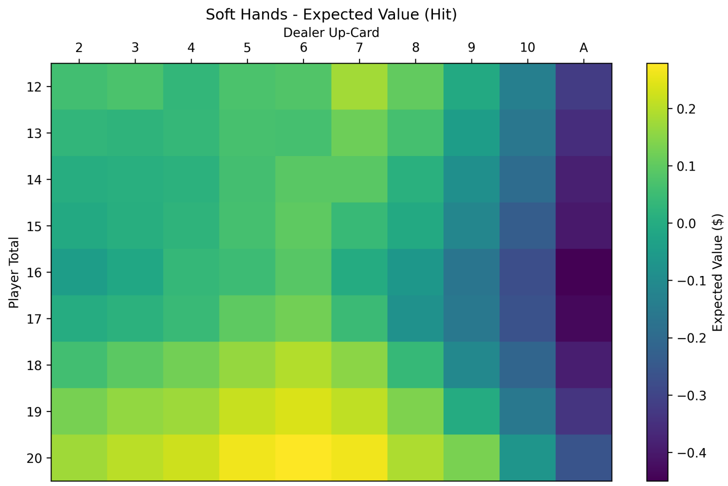

Figure 30: Soft Hit - Expected Profit (Hit) in Single Deck Blackjack with Complex Actions

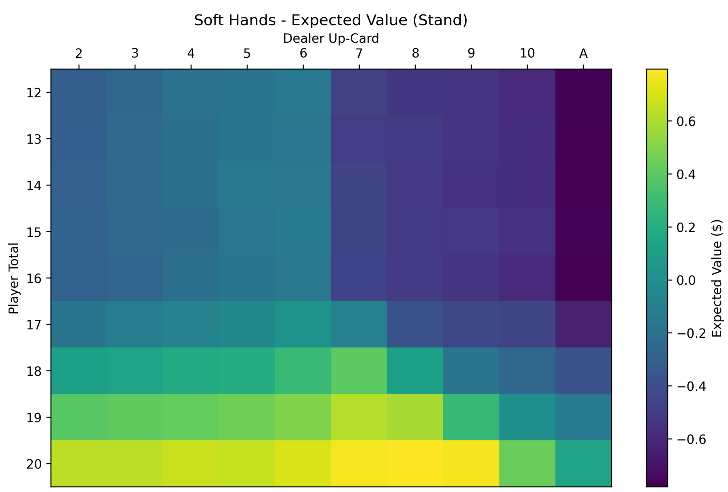

Figure 31: Soft Stands - Expected Profit (Stand) in Single Deck Blackjack with Complex Actions

Figure 32: Soft Double Down - Expected Profit (Double Down) in Single Deck Blackjack with Complex Actions

Examining the soft-hand heatmaps, we again observe similar smooth gradient patterns that we first observed in our simpler Blackjack setup with only "hit"s and "stands". The stand heatmap confirms that standing on soft 18–20 yields the highest expected profit, just as in the simpler stand map. Analyzing the soft double-down map, we see largely positive values on all rows of the left side of the map, indicating that a soft hand can't bust by drawing one more card if we double down. However, if the dealer has a high face-up card, this actually results in more losses for us, as constraining ourselves to only drawing one more card is at risk of resulting in a lower hand total than the dealer, especially when their face-up card is high.

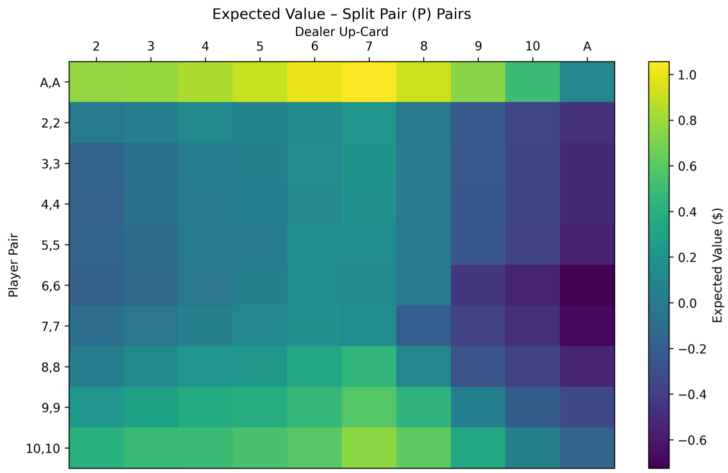

Figure 33: Split Pair (P) - Expected Profit in Single Deck Blackjack with Complex Actions

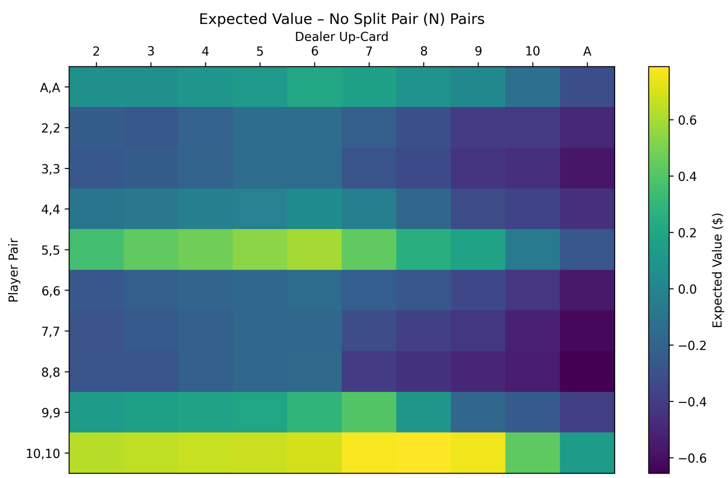

Figure 34: No Split Pair (N) - Expected Profit in Single Deck Blackjack with Complex Actions

We further analyze the heatmaps of our newest Split Pairs strategy state-action table. In the "Pair" heatmap, we clearly see that splitting A, A results in the highest expected profit regardless of the dealer's up-card. We further see more muted, yet still positive expected profits throughout the majority of the pairs on the left and middle portions of the heatmap, which eventually transition into losses on the right side. As such, this could explain how our optimal strategy table did not encourage splitting pairs for various pairs corresponding to high dealer up-cards. We see significantly clearer, less gradient-like results in the "No Split Pairs" heatmap, where not splitting 5s and 10s yielded the highest expected profit. We also see positive expected profits for splitting A, A, but from the "Split Pair" heatmap, splitting them clearly yields significantly ore profit. Across any split pair, we see that high dealer face-up cards always result in more losses for the player, as we see a horizontal gradient where the highest losses are experienced when the dealer has a 9, 10, or A.

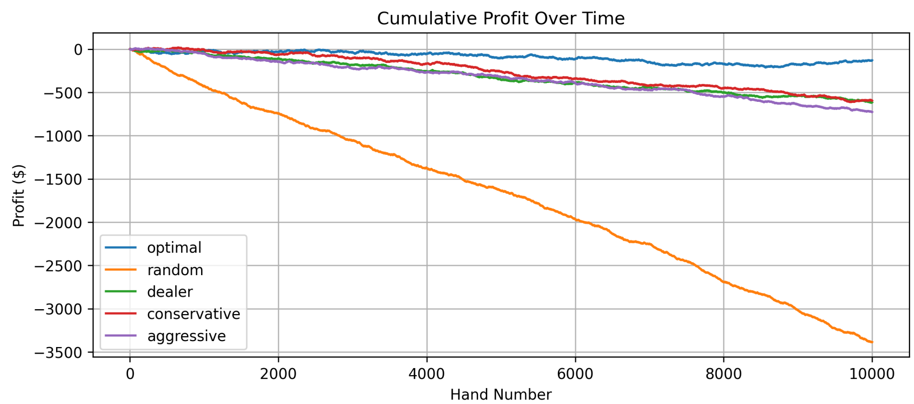

Figure 35: Cumulative Profit/Loss Over 10,000 Hands Across Different Blackjack Strategies for Single Deck Blackjack with Complex Actions. All strategies still lost money over time, with Optimal with −\$129.5, Random with −\$3387.5, Dealer with −\$615.5, Conservative with −\$595.5, and Aggressive with −\$725.5

Strategy	**EV**	**Win %**	**Loss %**	**Push %**	**Natural BJ %**	**Ruin @ \$500**
Optimal	-0.0130	42.5%	49.3%	8.21%	4.66%	No
Random	-0.3386	27.25%	68.86%	3.89%	4.66%	Yes
Dealer	-0.0616	41.0%	49.4%	9.58%	4.41%	Yes
Conservative	-0.0595	42.7%	51.1%	6.22%	4.78%	Yes
Aggressive	-0.0726	40.7%	50.9%	8.45%	4.55%	Yes

Table 3: Performance Comparison Across Different Blackjack Strategies for Single Deck Blackjack with Complex Actions.

Comparing our optimal strategy against other strategies in Single Deck Blackjack with complex actions, we clearly see an improvement in our optimal strategy's performance. Specifically, our expected profit loss is only −0.0130, translating to a house edge of 1.30%, a significant decrease from the previous house edge of 2.58% in single deck blackjack with only "hit" and "stand". We would expected our win percentage to increase, but it stays relatively the same; this is because as we double-down, the number of wins we have across our trials remains roughly the same, but we just win more per hand. Our loss, push, and natural blackjack rate remain roughly the same from the previous tables. Another notable observation is our random strategy becoming worse and worse; as we introduced more possible actions into the game, there were more opportunities for the random model to lose a hand, given that the cells for its state-action tables were populated randomly. We also note the small improved performance for the Conservative strategy over the Dealer strategy. Though this could be a result of pure chance, it could also be attributed to the fact that the conservative strategy allowed for splitting pairs on A, A, whereas the dealer was not allowed to do so. In these scenarios, the conservative player converted their one soft 12 into two strong hands each containing an Ace; hitting afterwards would then have a good chance of beating the dealer in the end game. Overall, by adding doubling, splitting, and surrendering to our action space, our optimal strategy was capable of nearly halving the previously reported house edge and significantly outperforming other suboptimal strategies.

7 Conclusion

In this project, we aimed to first find an optimal playing strategy for Blackjack using Monte Carlo simulations. By breaking down the casino game into state-action table and simulating thousands of hands, we were able to iteratively construct the strategy state-action table in a backwards-manner that maximizes our objective, which is expected profit. Testing with and analyzing our optimal strategy, we found that varying the number of Monte Carlo simulations in our optimal strategy significantly dictated the noise level and stability of our state-action tables and their corresponding expected profit heatmaps. After performing simulations for a simplified Blackjack setup with only "hit" and "stand" as possible actions, we were able to extend our simulation setup to include "double-down", "split", and "surrender", and find optimal hard, soft, and split strategy tables that accounted for all of these actions.

Evaluating our optimal strategy against other Blackjack strategies in simplified Blackjack, we found that the optimal strategy yielded the lowest loss in expected profit, highest win rate, lowest loss rate, and was the only strategy that did not go bankrupt. This strategy was capable of achieving a house edge of 2.58%. Comparing each strategy's state action tables, and even state-action tables across optimal strategy trained on different k trials, we found that even small deviations in decision making for particular states led to large differences in the amount of profit lost in the long run. When testing our optimal strategy on simplified Blackjack across single and infinite decks, our optimal strategy tables change only marginally, underscoring the robustness of our strategy under these deck compositions.

Evaluating our optimal strategy in complete Blackjack with all possible actions, our optimal strategy was capable of reducing the house edge by nearly 50% compared to the hit/stand-only baseline. Our optimal strategy achieved an expected loss of just 1.30 cents per \$1 bet, which equates to house edge of 1.30%. This aligns with the expected house edge when playing with optimal strategy in the real world, where house edge is expected to be $\approx 1\%$. This can be attributed to the introduction of doubling-down and splits, which leveraged particular game states to achieve high expected profits in returns. When pitted against suboptimal policies (random play, dealer play, conservative play, and aggressive play), our optimal policy yielded the highest expected profit (least loss) and experienced zero risk of ruin over $10,000$ simulated hands. By demonstrating improved results of our optimal strategy in complete Blackjack over simplified Blackjack and other suboptimal strategies, we're able to show that optimally assigning available actions at each game state is crucial for minimizing expected loss in the long term.

Through our computational simulations, experiments, and results, we have shown that even though Blackjack is a seemingly "random" game with casino edge, there is still a way to "win" by minimizing how much you lose. In regards to future work, we aim to extend our experiments to explore even more rule variations, such as dealer hitting soft 17, different payout structures, or even simple card-counting signals to improve and test the robustness of our optimal strategy.