A Description of Figure 3

Here we describe how the plots in Figure 3 are generated. There are six applications of the suspense and surprise models: tennis, soccer, blackjack, Clinton-Obama primary, optimal surprise, and optimal suspense. For each application, the second column depicts a scatterplot of surprise and suspense values for each observation, where $u(x) = \sqrt{x}$ and absolute surprise and suspense values are divided by the square root of the number of periods until beliefs first hit 0 or 1. Using the observation closest to the median surprise and suspense values, the first column traces actual and counterfactual belief paths during the periods within the median game. Below we provide an overview of the data and instructions for replicating these plots.

A.1 Tennis

We collect point-level data on all 2011 women’s Grand Slam tournament matches from www.jeffsackmann.com. There are a total of 43,471 points from 309 complete matches. Six incomplete matches are dropped from the analysis. Following the methodology at www.jeffsackmann.com, all players are assumed to be identical, and we assign a tournament-specific probability that a server wins the point. The probabilities are as follows: Australian Open, 0.61; French Open, 0.59; Wimbledon, 0.64; and US Open, 0.62. Each belief path begins the match at 0.5, and evolves according the probabilities of winning the current game and set, conditional on the tournament and the state of the game.

A.2 Soccer

We collect score data by minute for all 36,693 soccer matches on the ESPN website, http://soccernet.espn.com, from 99 different leagues between August 1, 2010 and July 31, 2012. We are interested in the probability of the home team winning (as opposed to losing or tying). The home team is designated on the website, although the distinction is arbitrary when the match is played on a neutral site. To restrict the data to matches that would end after regulation time without extra time or penalty kicks, we remove matches from our sample satisfying any of the following criteria:
• Non-group matches in leagues with the words “Cup”, “Copa”, “Coupe”, “Trophy”, or “Pokal”.

• Leagues that include an extra time match that is not designated as a leg or non-group stage.

• Matches with multiple legs.

• Matches abandoned or suspended before full-time.

• Matches where the goals are scored at minute 0', or the sum of listed goals does not match the team’s total.

• Leagues with a remaining sample of less than 10 matches.

The final sample contains 24,825 matches from 67 different leagues.

The data does not distinguish between goals scored at 45' and those scored in stoppage time at the end of the first half, or goals scored at 90' and those scored in stoppage time at the end of regulation. To account for this we randomly allocate goals listed at 45' as being scored at minutes 45 or 46 with equal probability, then treat the second half as beginning at minute 47. We allocate goals at 90' as being scored with equal probability at minutes 91, 92, 93, or 94. Each match thus lasts 94 minutes.

We assume that teams are identical within leagues, except for the status of home vs. away. In every minute from 1 to 94, the home team \( H \) and away team \( A \) are assumed to independently draw \( s = 0, 1, \) or 2 goals with probability \( p_s^T \) for \( T \in \{ A, H \} \), where the scoring probability remains constant over time. We estimate these six probabilities separately for each league. Each belief path begins the match at the league-specific simulated probability that the home team wins, and evolves according the probabilities of the home team’s winning the current match, conditional on the league and home minus away score differential. The tendrils in Figure 3 show the possible beliefs supposing that team \( A \) scores, team \( H \) scores, or neither team scores in the coming minute.

A.3 Blackjack

Blackjack data are based on 20,000 simulated “visits” to the casino. On every visit, the Player plays $10 hands of blackjack until his net winnings are above $100 or below $-100. The Player is allowed to split, double, and redouble his hand, but may not take insurance
or surrender.\textsuperscript{1} The Dealer stands on “soft 17” and is not allowed to peek.\textsuperscript{2} Simulated cards are drawn from a standard deck with replacement. The Player plays the optimal non-card counting strategy, taken from http://www.blackjackencyclopedia.com/blackjack-strategies.html under the following specifications: 8 decks, Dealer stands on soft 17, May Double After Split, No Surrender, No Peek.

The probability of winning the visit is based on the Player’s current money balance and the states of the Dealer and Player hands. Based on the value of the next drawn card, there are at most 10 different beliefs in the next period. The tendrils in Figure 3 show the maximum and minimum possible beliefs after the next card is drawn.

We repeat this exercise, without reporting the results on the graph, for the blackjack variation with no doubling down and no splits allowed. We find that eliminating these options reduces average suspense and surprise, after scaling down by the square root of the visit length.\textsuperscript{3} The averages over 20,000 visits each are shown below.

<table>
<thead>
<tr>
<th></th>
<th>Mean Suspense</th>
<th>Mean Surprise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Splits, Doubling</td>
<td>.521</td>
<td>.452</td>
</tr>
<tr>
<td>No Splits, No Doubling</td>
<td>.511</td>
<td>.444</td>
</tr>
</tbody>
</table>

\section*{A.4 Clinton-Obama Primary}

The data are results from the Iowa Electronic Markets\textsuperscript{4}, an online prediction market run by University of Iowa Henry B. Tippie College of Business for research purposes. The beliefs are based on the average daily price in this market from March 2, 2007 to August 28, 2008. Without explicitly modeling the belief process, we can only graph surprise. The scatter plot depicts a vertical line at the calculated surprise level, and the belief path does not include tendrils for counterfactual states.

\textsuperscript{1}Split: When a Player has exactly two cards of the same value, he may create two separate hands and play both hands. Each hand has the same wager as the original wager, so twice as much money is at risk. Double: After receiving two cards, the Player may double his bet and receive exactly one more card. Redouble: after splitting, a player is allowed to double either hand. Surrender: When the Player has two cards, he may forfeit his hand and only loses half the wager. The two variations are early and late surrender. In early surrender, the Player may fold before the Dealer checks her hand for blackjack (see peek below). In late surrender, the Player may only fold after the Dealer checks her hand for blackjack. Insurance: When the Dealer up card is an Ace, a Player may bet 2:1 if the Dealer has blackjack (exactly two cards). A Player may bet up to half his original wager. (cf http://www.blackjackinfo.com/blackjack-rules.php)

\textsuperscript{2}Dealer Peek: If the Dealer up card is an Ace or 10, the Dealer checks her hand and reveals Blackjack before a Player decides to split or double. This method is used in the US, but not in Europe. When Peek is allowed, the current version of the code allows the Dealer to immediately draw Blackjack. A Dealer may stand or hit on “soft 17”; that is, Ace + 6.

\textsuperscript{3}Looking at total suspense and surprise without scaling down, games which disallow doubling down tend to be more suspenseful and surprising for the reason that they tend to take longer. Doubling down allows the player to bet more money on a single hand.

\textsuperscript{4}http://iemweb.biz.uiowa.edu/pricehistory/pricehistory_SelectContract.cfm?market_ID=214
A.5 Surprise Optimum

We construct the optimal surprise policy by numerically solving the recursive maximization problem discussed in Section 5 and Appendix A to get $\mu_l$ and $\mu_h$ as functions of the current belief and the number of remaining periods. All beliefs are restricted to a grid of 501 uniformly spaced points on $[0,1]$. Starting at some belief $\mu_t$, in period $t+1$ we jump down to belief $\mu_l$ with probability $\frac{\mu_h - \mu_t}{\mu_h - \mu_l}$ and up to belief $\mu_h$ with the complementary probability. We simulate 1000 belief paths of this process, given $T = 50$ and starting belief $\mu_0 = \frac{1}{2}$.

A.6 Suspense Optimum

The optimal suspense policy is given analytically in Section 4, where the belief at period $t$ is either $H_t$ or $L_t$:

$$H_t = \frac{1}{2} + \sqrt{\left(\mu_0 - \frac{1}{2}\right)^2 + \frac{t}{T}\mu_0(1 - \mu_0)}$$

$$L_t = \frac{1}{2} - \sqrt{\left(\mu_0 - \frac{1}{2}\right)^2 + \frac{t}{T}\mu_0(1 - \mu_0)}$$

Starting at some belief $\mu_{t-1}$, in period $t$ we jump down to belief $L_{t+1}$ with probability $\frac{H_t - \mu_t}{H_{T+1} - L_{T+1}}$ and up to belief $H_{t+1}$ with the complementary probability. We simulate 1000 belief paths of this process, given $T = 50$ and starting belief $\mu_0 = \frac{1}{2}$.

B Costly Time and Random End Dates

Instead of having a fixed time $T$ at which the game ends, now suppose that the end time may be stochastic, and that the agent pays an opportunity cost for the time elapsed until the state is revealed. In particular, if the state is revealed after $T^*$ periods, the agent pays $c(T^*)$ in addition to receiving her suspense or surprise payoff, where $c(T^*)$ is increasing and weakly convex. Focusing in this section on suspense utility, the agent’s net payoff – realized suspense payoff minus opportunity costs – is

$$\sum_{t=0}^{T^*-1} u \left( E_t \sum_{\omega} (\hat{\mu}_{t+1}^{\omega} - \mu_t^{\omega})^2 \right) - c(T^*).$$

The principal’s problem is to construct an information process to maximize the net payoff of suspense minus opportunity costs. We are interested in comparing the solution to this problem with the one in the body of the paper, of finding an optimal information process subject to a fixed end date.

For martingales with a deterministic end date $T^* = T$, we show in Section 4 of the paper that the agent’s optimal suspense payoff is $Tu \left( \Psi(\mu_0) \right) / T$. In fact, an extension of that
argument shows that this same expression gives an upper bound for the agent’s suspense payoff over all martingales with expected end date $E[T^*] = T$.

**Lemma.** For any martingale $\tilde{\mu}$ with an expected end date of $E_{\tilde{\mu}}[T^*] = T$ in $[1, \infty)$, the suspense payoff is bounded above by $Tu\left(\frac{\Psi(\mu_0)}{T}\right)$.

**Proof.** As in Lemma 2 of the paper, given any martingale such that $E_{\tilde{\mu}}[T^*]$ is finite, we have $E\sum_{t=0}^{T^*-1} \sigma_t^2 = \Psi(\mu_0)$. Subject to this budget, doling out variance evenly over time would be optimal, if that were possible. That would correspond to a variance of $\sigma_t^2 = \Psi(\mu_0)/T$ in each period prior to $T^*$. So the resulting upper bound for the expected suspense payoff is $Tu\left(\frac{\Psi(\mu_0)}{T}\right)$.

Restricting attention to policies with an integer-valued expected end date, we see that there is an upper bound on the agent’s net payoff of

$$\max_{T \in \mathbb{N}} \left[ Tu\left(\frac{\Psi(\mu_0)}{T}\right) - c(T) \right].$$

This payoff is exactly achieved by a policy that is optimal for a fixed end date, of the form derived in Section 4 of the paper, choosing the end date $T$ to maximize the above expression. In other words, modulo integer constraints, the problem of finding a martingale that maximizes net payoffs can be separated into finding an optimal length of the information process; and then implementing the suspense-optimal information policy of that length. (There might be a small benefit of policies with noninteger-valued expected end dates, in which case we could improve the agent’s payoff with a policy that randomizes over two adjacent integer end dates.)

In Section 4 of the paper, we also found that given some fixed end date and a binary state space, the suspense-optimal martingale was unique. We now present an alternative martingale that has the same suspense payoff and lasts for the same expected number of periods, but has a stochastic end date. Under linear opportunity costs $c(T^*) = kT^*$, the agent would be indifferent between the optimal time-$T$ martingale and one of the class described below.

**Example.** Let there be two states, $\Omega = \{A, B\}$. Fix some $\mu_0$ and some probability $p \in (0, 1)$ and define a belief martingale in the following way. At period $t + 1$, move from $\mu_t$ to $\mu_{t+1} = 1$ with probability $p\mu_t$; to $\mu_{t+1} = 0$ with probability $p(1 - \mu_t)$; and remain at $\mu_{t+1} = \mu_t$ with probability $1 - p$.

This martingale represents belief paths under a sudden death overtime rule, or a so-called “golden goal”, where the first team to score wins. Team A has a probability $p\mu_0$ of scoring in each period and Team B has a mutually exclusive probability of $p(1 - \mu_0)$ of scoring in each period. With probability $1 - p$, no team scores, and we repeat this indefinitely.
This martingale reveals the state in an expected $E_r[T^*] = \frac{1}{T}$ periods, and beliefs have a constant variance of $\sigma^2_t = p(\mu_0)(1 - \mu_0) = p\Psi(\mu_0)$ in each period $t$ before the state is revealed. So it induces a suspense payoff of $\frac{1}{p} u(p\Psi(\mu_0))$, exactly the level of the upper bound from the lemma above.

For $p = \frac{1}{T}$, this martingale delivers the same suspense payoff as the suspense-optimal martingale for a fixed number of periods $T$, and under linear opportunity costs $c(T^*) = kT^*$ it has the same expected costs as well.

## C Suspense extensions

**Claim.** Under state-dependent significance, the optimal prior is no longer uniform for state-dependent significance. More significant states are given a prior closer to $\frac{1}{2}$. Suffi-

**Proof.** Without loss of generality, suppose all $\alpha^\omega > 0$.5

The payoff from the suspense-optimizing policy is now

$$Tu\left(\sum_\omega \alpha^\omega \bar{\mu}_0^\omega (1 - \mu_0^\omega)\right).$$

Maximizing over $\mu_0 \in \Delta(\Omega)$, the optimal prior will set $\alpha^\omega(1 - 2\mu^\omega)$ equal to a constant Lagrange multiplier $\lambda$ on all states with $\mu_0^\omega > 0$, and $\alpha^\omega(1 - 2\mu^\omega) \leq \lambda$ for any state with $\mu^\omega = 0$. Rewriting this, it holds that $\mu^\omega \geq \frac{\lambda}{2\alpha^\omega}$, with equality if $\mu_0^\omega$ is positive. That is, $\mu_0^\omega = \max\{0, \frac{1}{2} - \frac{\lambda}{2\alpha^\omega}\}$. So define

$$\bar{\mu}^\omega(\lambda) \equiv \max\{0, \frac{1}{2} - \frac{\lambda}{2\alpha^\omega}\}.$$  

This is 0 for $\lambda \geq \alpha^\omega$ and positive for lower $\lambda$.

Now find $\lambda^*$ such that $\sum_\omega \bar{\mu}^\omega(\lambda^*) = 1$. The value $\lambda^*$ is uniquely defined and nonnegative because $\sum_\omega \bar{\mu}^\omega(\lambda)$ is $|\Omega|/2 \geq 1$ at $\lambda = 0$, is strictly decreasing in $\lambda$ when positive, and goes to 0 at $\lambda = \max_\omega \alpha^\omega$. The optimal prior sets $\mu_0^\omega = \bar{\mu}^\omega(\lambda^*)$.

With two states, we get $\mu_0^\omega = \frac{1}{2}$ for both states (and $\lambda^* = 0$).

Consider now the case of three or more states.

As we hold other $\alpha^\omega$ fixed for $\omega \neq \omega'$ and increase $\alpha^{\omega'}$ to infinity, we get that $\mu_0^{\omega'}$ approaches 1/2 from below. To see this, let $\lambda' = \max_{\omega \neq \omega'} \alpha^\omega$, which does not vary with $\alpha^\omega$. It holds that $\sum_\omega \bar{\mu}^\omega(\lambda') = \bar{\mu}^\omega(\lambda') \leq \frac{1}{2}$, and so $\lambda^* \leq \lambda'$. Hence $\mu_0^{\omega'} \geq \bar{\mu}^{\omega'}(\lambda')$, and the right-hand side of the inequality goes to $\frac{1}{2}$ as $\alpha^\omega$ goes to infinity.

---

5If only a single state has positive $\alpha^\omega$, then it is optimal set the positive weight state at $\frac{1}{2}$ and allocate the rest of the probability arbitrarily across states. With two or more states of positive weight, delete any states with $\alpha^\omega = 0$ and continue the analysis with fewer states.
With four or more states, it is possible for a state $\omega'$ with a low weight to have priors of 0. In particular, find $\lambda > 0$ such that $\sum_{\omega \neq \omega'} \mu_{\omega}(\lambda) = 1$. If $\alpha^\omega < \lambda$, then $\lambda = \lambda^*$ and $\mu_{\omega}^* = 0$.

**Claim.** Under state-dependent significance with a choice of prior and significances, $\alpha^\omega$ and $\mu_0$ are optimal if and only if $\mu_{\omega}^0 = \frac{1}{2}$ for each state with $\alpha^\omega > 0$.

**Proof.** The principal’s problem is to maximize

$$\max_{\alpha, \mu_0} Tu \left( \frac{\sum_{\omega} \alpha^\omega \mu_{\omega}^0(1 - \mu_{\omega}^0)}{T} \right)$$

subject to $\mu_0 \in \Delta(\Omega)$, $\alpha^\omega \geq 0$ for all $\omega$, and $\sum_{\omega} \alpha^\omega$ fixed at some positive value. Equivalently, the principal maximizes $\sum_{\omega} \alpha^\omega \mu_{\omega}^0(1 - \mu_{\omega}^0)$.

For any $\mu_0$ and $\alpha$,

$$\sum_{\omega} \alpha^\omega \mu_{\omega}^0(1 - \mu_{\omega}^0) \leq \left( \sum_{\omega} \alpha^\omega \right) \cdot \max_{\omega} [\mu_{\omega}^0(1 - \mu_{\omega}^0)] \leq \left( \sum_{\omega} \alpha^\omega \right) \frac{1}{4},$$

and this holds with equality if and only if $\mu_{\omega}^0 = \frac{1}{2}$ for each state with $\alpha^\omega > 0$.

**Claim.** Under time-dependent significance, more important periods are more suspenseful. If $u(x) = \sqrt{x}$, then $\sigma_t$ is proportional to $\beta_t$.

**Proof.** To maximize

$$U_{\text{susp}} (\tilde{\mu}, \eta) = \sum_{t=0}^{T-1} \beta_t u \left( E_t \sum_{\omega} (\tilde{\mu}_{t+1}^\omega - \mu_{t}^\omega)^2 \right)$$

we move along circles of constant $\Psi(\mu)$ which do not necessarily have radius increasing linearly in $t$.

There is a total budget of variance of $\sum_t \sigma_t^2 = \sum_{\omega} \mu_{\omega}^0(1 - \mu_{\omega}^0)$, and the payoff in a period is $\beta_t u(\sigma_t^2)$. This is maximized by setting $\beta_t u'(\sigma_t^2)$ equal to a constant Lagrange multiplier $\lambda$ for all periods $t$ with $\sigma_t^2 > 0$, and $\sigma_t^2 = 0$ if $\beta_t u(0) < \lambda$. We see that $\sigma_t^2$ weakly increases in $\beta_t$, because $u'$ is decreasing in $\sigma_t^2$.

For the special case of $u(x) = x^\gamma$ for $0 < \gamma < 1$, we have $\gamma \beta_t (\sigma_t^2)^{\gamma-1} = \lambda$, so that $\sigma_t^2 = \left( \frac{\gamma \beta_t}{\lambda} \right)^{-\frac{1}{1-\gamma}}$; that is, $\sigma_t^2$ is proportional to $(\beta_t)^{\frac{1}{1-\gamma}}$. Plugging in $\gamma = \frac{1}{2}$, we have $\sigma_t^2$ proportional to $\beta_t^2$. \qed
D Optimizing surprise for alternative $u(x)$

When $\Omega = \{A, B\}$, the argument that the surprise-optimal policy can be achieved with binary signals goes through under arbitrary $u(\cdot)$. In each period, we continue to randomize over some low belief $\mu_l$ and some high belief $\mu_h$.

Moving beyond $u(x)$ proportional to $\sqrt{x}$, we can consider not just alternative functional forms of concave $u(\cdot)$, but convex ones as well. The convexity or concavity of $u(\cdot)$ tells us how to trade off the expected frequency of small belief changes versus large belief changes. The less concave or the more convex is the stage utility function, the stronger is the relative preference for large jumps. By considering utilities in the class $u(x) = x^\gamma$, we can reveal some novel behavior of surprise optimal martingales. When $u(\cdot)$ is very concave, full revelation might not be optimal. When $u(\cdot)$ is convex, the value function might not be maximized at a prior of $\frac{1}{2}$.

First consider the issue of full revelation by the end. Suppose that the belief is $\mu_{T-1} < \frac{1}{2}$ at the second-to-last period. Under full revelation, there is a large chance $(1 - \mu_{T-1})$ of a small jump to $\mu_l = 0$, and a small chance $(\mu_{T-1})$ of a large jump to $\mu_h = 1$. Now consider reducing $\mu_h$ below 1, so that we no longer have full revelation. This increases the probability of a large jump upwards at the cost of making this jump a little bit smaller. The relative cost of making a large jump a little smaller is high when $u(\cdot)$ is convex (expanding the jump size by a little adds a lot of utility because $u'(\cdot)$ is increasing) and low when $u(\cdot)$ is concave. In particular, full revelation is optimal for $u(x)$ more convex than $\sqrt{x}$, i.e., $u(x) = x^\gamma$ for $\gamma \geq \frac{1}{2}$. For $u(x)$ more concave than $\sqrt{x}$, non-full revelation – taking $\mu_h < 1$ – is optimal when the belief $\mu$ is close to a boundary.

**Proposition.** Suppose $u(x) = x^\gamma$ for some $\gamma > 0$. If $\gamma \geq \frac{1}{2}$, then the surprise maximizing policy will be fully revealing by the end. If $\gamma < \frac{1}{2}$, then there exists a nondegenerate belief $\hat{\mu}$ such that the policy does not fully reveal if $|\mu_{T-1} - \frac{1}{2}| > |\hat{\mu} - \frac{1}{2}|$.

**Proof.** Starting at period $T-1$ with belief $\mu_{T-1} = \mu$, the payoff from moving to an arbitrary $\mu_l$ and $\mu_h$ is

$$\frac{\mu_h - \hat{\mu}}{\mu_h - \mu_l} (\hat{\mu} - \mu_l)^{2\gamma} + \frac{\hat{\mu} - \mu_l}{\mu_h - \mu_l} (\mu_h - \mu)^{2\gamma}$$

1. Take $\gamma \geq \frac{1}{2}$. I will show that for any $\mu_l$, it is optimal to choose $\mu_h = 1$. The symmetric argument would show that $\mu_l$ is always chosen at 0.

Taking derivative of the above payoff expression with respect to $\mu_h$ gives

$$\frac{\mu_h - \mu}{(\mu_h - \mu_l)^2(\mu_h - \mu)} ((\mu_h - \mu)(\mu - \mu_l)^{2\gamma} + (\mu_h - \mu_l)^{2\gamma} (2\gamma(\mu_h - \mu_l) - (\mu_h - \mu)))$$

I seek to show that the derivative is positive, assuming $\mu_l < \mu < \mu_h$. It is sufficient to show that $2\gamma(\mu_h - \mu_l) - (\mu_h - \mu)$ is positive. This holds because $2\gamma(\mu_h - \mu_l) - (\mu_h - \mu) > (\mu_h - \mu_l) - (\mu_h - \mu) > 0$. 

8
2. Take \( \gamma < \frac{1}{2} \). I seek to show that for \( \mu \) close enough to 0, it is not optimal to take \( \mu_h = 1 \) if \( \mu_l = 0 \) in the coming period. (A symmetric argument would show that \( \mu_l = 0 \) is not optimal for \( \mu \) close to 1). The expected stage payoff at time \( T \) from choosing \( \mu_l = 0 \) and some general \( \mu_h > \mu \) is

\[
\frac{\mu_h - \mu}{\mu_h} \mu^{2\gamma} + \frac{\mu}{\mu_h} (\mu_h - \mu)^{2\gamma}
\]

It suffices to show that for small \( \mu \), the derivative of this expression with respect to \( \mu_h \) is less than 0 evaluated at \( \mu_h = 1 \). Calculating the derivative and simplifying, it suffices to show that

\[
\mu^{2\gamma} (1 - \mu)^{1-2\gamma} < 1 - \mu - 2\gamma
\]

for small \( \mu \). Indeed, as \( \mu \) goes to 0 the left-hand side approaches 0 and the right-hand side approaches \( 1 - 2\gamma > 0 \), so the inequality holds.

When full revelation is optimal, as it is for convex \( u(\cdot) \), the surprise value function \( W_1(\mu|u) \) at the second-to-last period (i.e., for \( T = 1 \)) is given by

\[
W_1(\mu|u) = (1 - \mu)u(\mu) + \mu u(1 - \mu).
\]

At the uniform belief, this evaluates to \( u(.5) \). Consider an alternative starting belief, say \( \mu = .1 \). This gives a surprise value of \( .1u(.9) + .9u(.1) \). As \( u(\cdot) \) becomes more and more convex, \( u(.9) \) comes to dominate both \( u(.1) \) and \( u(.5) \). So for convex enough \( u(\cdot) \), the surprise value \( W_1(\mu|u) \) is higher at \( \mu = .1 \) than at \( \mu = .5 \).

In the class of utilities \( u(x) = x^\gamma \), the second derivative of \( W_1(\mu|u) \) at \( \mu = .5 \) is positive (indicating that \( \mu = .5 \) does not maximize the payoff) for \( \gamma > 3.6 \)

\[\text{footnote}{The payoff is } (1 - \mu)\mu^{\gamma} + \mu(1 - \mu)^{\gamma}, \text{ which has second derivative at } \mu = \frac{1}{2} \text{ of } \gamma^{2-\gamma}(\gamma - 3).\]