THE REVIEW OF **ECONOMIC STUDIES**

Review of Economic Studies (2010) **77**, 882–913 0034-6527/10/00411011$02.00

doi: 10.1111/j.1467-937X.2009.00594.x

Moderation in Groups: Evidence from Betting on Ice Break-ups in Alaska

RENÉE ADAMS

University of Queensland and ECGI

and

DANIEL FERREIRA

London School of Economics, CEPR and ECGI

First version received August 2005; *final version accepted August* 2009 *(Eds.)*

We use a large sample of guessed ice break-up dates for the Tanana River in Alaska to study differences between outcomes of decisions made by individuals versus groups. We estimate the distribution of guesses conditional on whether they were made by individual bettors or betting pools. We document two major distinctions between the two sets of guesses: (1) the distribution of guesses made by groups of bettors appears to conform more to the distribution of historical break-up dates than the distribution of guesses made by individual bettors, and (2) the distribution for groups has less mass in its tails and displays lower variability than the distribution for individuals. We argue that these two pieces of evidence are consistent with the hypothesis that group decisions are more moderate, either because groups have to reach a compromise when their members disagree or because individuals with extreme opinions are less likely to be part of a group.

1. INTRODUCTION

Many economic decisions are made by groups. In firms, corporate boards and management committees make important decisions. In households, final decisions depend on the relative bargaining power of each family member. However, despite their obvious relevance for understanding economic choices, little is known about the differences between the decisions made by groups and the ones made by individuals.

Since individuals differ in the information they have, deliberation within groups presumably leads to information sharing among group members. If there are no conflicts of interest within a group, group decision making should be based upon all relevant information. This suggests that groups should make more predictable decisions than individuals, in the sense of relying more on "hard evidence" collected by information pooling. Some also argue that groups are more prone to reason-based choice, that is, group members favour choices that are easy to justify (see, for example, Barber, Heath and Odean, 2003). The importance of finding "good reasons" naturally leads groups to rely more on tangible evidence.

To be able to reach a consensus, groups need to balance individual opinions. This implies that groups should make less extreme decisions than the ones made by individuals. We call this a *compromise effect*. Groups may also choose to restrict membership of individuals with extreme opinions, or individuals with extreme opinions may choose not to join groups. We call this a *membership effect*. Taken together, these arguments imply that groups should make more moderate decisions than individuals.[1]

We develop a simple model in which these two effects interact and lead to moderation in groups that are endogenously formed. In order to avoid compromise, some individuals choose not to be group members.[2] Thus, the key insight from the model is that endogenous membership amplifies the moderating effect of groups.

Some experimental studies document cognitive biases specific to group decision making that may result in more extreme outcomes. For example, there is evidence that groups sometimes make riskier choices than individuals, a phenomenon that has been labelled "risky shifts" (Wallach and Kogan, 1965; Stoner, 1968). A related empirical regularity is known as "group polarization", which happens when group judgment deviates significantly from the average of the pre-discussion individual judgments, in the direction of higher "extremity" (Moscovici and Zavalloni, 1969; Kerr, 1992). Finally, there is the phenomenon of "groupthink" (Janis, 1982), which can be described as a "...dysfunctional mode of group decision-making characterized by a relentless striving for unanimity, resulting in a reduction in independent critical thinking" (Forbes and Milliken, 1999).

Such biases could counteract the moderating effects we describe above. Groupthink and group polarization could lead groups to discard some evidence aggregated through information pooling, which may make group decisions less predictable, while group polarization and risky shifts could lead groups to make more extreme decisions than the ones made by individuals. Research on small groups sometimes suggests that "small-group judgments tend to be more volatile and extreme" (Surowiecki, 2005, p. 176). Thus, the question "Are decisions made by groups more or less predictable and extreme than the ones made by individuals?" is, in principle, an open one.

The main contribution of this paper is to complement the experimental literature on group behaviour by presenting field evidence consistent with the idea that the outcomes of decisions made by individuals and groups differ systematically. We use data from an unusual betting game played in Alaska by a large number of individuals in 2002. The Nenana Ice Classic, 2002, named after the city of Nenana, started in 1917 when a group of surveyors for the Alaska Railroad made bets with each other on the date the ice on the Tanana River would break up. Nowadays, bets can be made on the Tanana River ice break-up date by buying tickets. The ticket holders whose guess is closest to the exact minute of break-up win the Jackpot. In 2002, the amount that could be won was $304,000.

Our primary data consist of the Nenana Ice Classic's *List of Guesses* for 2002. This is the official compilation of guesses that serves to determine the winners of the Ice Classic. It lists the names of each bettor in the order of predicted break-up date, hour and minute. Using bettors' names, we classify guesses into those made by individual bettors or by groups of bettors (*betting pools*). To examine whether the betting behaviour of groups is more or less extreme than that of individuals, we use estimates of the distribution of guesses conditional on this information.[3]

1. These ideas are well-known in the social psychology literature on groups. According to Levine and Moreland (1998), there are two mechanisms for the resolution of disagreement among group members: (i) *change of position*: members either compromise or adopt the position of others, and (ii) *redefinition of group boundaries*: non-conforming members leave the group.

2. Kocher, Strauß and Sutter (2006) provide experimental evidence that unwillingness to compromise and the desire to preserve autonomy are the main reasons why individuals choose not to join groups.

3. We discuss potential problems with our classification in detail in Sections 4 and 6.

There are many advantages of using this game to study differences between group and individual decision making. First, it is a real game that participants most likely understand and choose to play voluntarily. The rules are very simple, the game has been around for a long time, and all participants have access to the same information necessary for an educated guess. The tickets are accompanied with brochures that provide a tabulation of the historical break-up dates and times since 1917, and weather and ice condition information can be obtained from the Nenana Ice Classic website, as well as from online and print services. Thus, concerns about whether agents understand the problem or have the right incentives to behave in the expected manner are minimal.

Second, the decisions we examine are easy to measure and the dataset is very large. We are able to use 290,051 (virtually all) guesses for the 2002 break-up. Since we analyse essentially the entire population of guesses, we are able to obtain very precise estimates.

Of course, our data also have their limitations. Because we compare group and individual decision making in the context of a wagering game, our results may not generalize to situations in which participants do not behave in risk-seeking ways. It is also difficult to determine whether groups perform better or worse than individuals, because there is no obvious measure of good performance in this game. As Thaler and Ziemba (1988), Farrell *et al.* (2000) and Forrest, Simmons and Chesters (2002) point out, this issue is generally difficult to address in the context of wagering games. In this case, bettors face a trade-off between risk and return: betting on the most obvious dates would most likely imply a smaller prize, if right, because the prize is shared by all winners. Without knowledge of risk preferences, it is difficult to say that one guess is "better" than another one. This issue is complicated by the fact that the objectives of betting pools may be different from those of individuals, and that factors other than the expected return on the bet may play a role in bettors' objectives, such as, for example, fun.[4]

Because we rely on non-experimental data, we also cannot control for all aspects one may think are important. In particular, we cannot perfectly disentangle the direct effects of groups on outcomes from the effects of individuals' characteristics on group membership. Thus, we primarily provide evidence for the *joint hypothesis* that groups make more moderate decisions, either because they require a compromise among their members, or because individual characteristics may affect group membership, or both. However, we also provide some evidence that selection on individual characteristics alone is not sufficient to explain our findings.

Our main findings can be summarized as follows. First, the distribution of group guesses appears to conform more to the distribution of historical break-up dates than the distribution of individual guesses. In this sense, group guesses are more predictable. For example, individual bettors place more weight on earlier spikes in the historical break-up distribution than on later spikes, which is consistent with the idea that individual bettors may be biased towards earlier break-up dates. Second, we find that the group guess distribution places less weight on the tails and displays lower variability than the distribution for individuals.

By means of individual fixed-effects regressions, we also show that individuals who bet both alone and in groups are more likely to make their extreme guesses alone, suggesting that the moderating effect of groups works not only through the sorting of similar people

4. The results may also not generalize to situations in which utility functions and payoffs are invariant across individuals and groups or to situations in which individuals cannot choose to act alone as well as with a group. In addition, unlike in many experimental studies, per capita payoffs are not equal across individual bettors and betting pools. While our results seem to hold even if we restrict ourselves to groups and examine variation in group size, it is possible that having the choice to bet alone changes the way individuals negotiate about the pool's betting behaviour.

into the same groups. Although the fixed-effects regressions cannot hold the characteristics of the other group members fixed, it goes a long way towards illuminating the mechanism that leads to moderation in groups. Because we focus on a subsample of individuals who bet both alone and in groups, these results rule out the possibility that only people with moderate bet preferences choose to join groups. In this subsample, moderation in groups occurs either because of compromise or because group members cannot "agree" upon an extreme guess, and consequently choose to place their extreme guesses alone or leave groups that have guess preferences that are too different from their own.

To confirm that our findings are not special to the year 2002, we replicate our tests using 2008 data. Although we cannot use the same approach to classify bettors into individuals versus pools as in 2002, we were able to obtain a list of "registered pools" from the Ice Classic office, which was not available for 2002. A major advantage of using the 2008 data is that this subsample of pools is free from measurement error, which enables us to assess the impact of measurement errors on our results. Overall, the data from 2008 confirm the main results found in 2002: the guess distribution for pools displays a lower variance than the guess distribution for individuals.

Our evidence suggests that groups behave more conservatively than individuals. This statement, by itself, does not say anything about the relative *rationality* of group versus individual decision making. However, because group decisions appear to differ systematically from individual decisions and many economic decisions are made by endogeneously formed groups, our findings have economic implications, at least for situations similar to those we study.

The paper is structured as follows. In Section 2 we discuss the related literature. We present a simple model that illustrates the main ideas in Section 3. Section 4 describes the data. We discuss factors that are likely to affect decisions and present our results in Section 5. We address selection and other data-related issues in Section 6. We present the results for the 2008 data and their implications for potential problems with measurement errors in Section 7. We conclude in Section 8.

2. RELATED LITERATURE

The literature on wagering games, such as sports betting and lotteries, is extensive. Its primary focus has been to analyse the efficiency of betting markets and to estimate bettors' utility and demand functions (for surveys of this literature, see, for example, Sauer, 1998; Williams, 1999). Empirical papers in this literature have generally relied on aggregate data on bets. In contrast, we are able to identify the entire betting portfolios of individual bettors.

Few empirical papers in economics explicitly compare group to individual decision making. The main questions in the ones that do are the relative rationality and performance of groups versus individuals. Bone, Hey and Suckling (1999) find no evidence that groups conform more to expected-utility maximizing behaviour than individuals. Using field-data on mutual fund performance, Prather and Middleton (2002) also find no evidence that differences in fund performance can be attributed to differences in group versus individual decision making. Work by Blinder and Morgan (2005) documents contrasting results. Using experiments that are specifically designed to simulate how decisions about monetary policy are conducted in central banks, they provide experimental evidence that groups make both faster and better decisions than individuals. In experimental beauty-contest games, Kocher and Sutter (2005) find that groups do not appear to be more rational than individuals, but they tend to learn faster. In a similar setting, Sutter (2005) finds that larger groups perform better. None of these papers addresses our hypothesis that groups are more moderate.[5]

Barber *et al.* (2003) provide field evidence that stock clubs, relative to individual investors, are more likely to engage in reason-based choice when picking stocks. Their findings are consistent with the hypothesis that group decisions are more predictable on the basis of hard evidence (i.e. "good reasons"). Adams, Almeida and Ferreira (2005) find that stock returns are significantly more variable for firms run by powerful CEOs. They stress that in firms in which the CEO has less discretion, decisions are more clearly the product of consensus among the top executives. Bär, Kempf and Ruenzi (2005) provide evidence that team management reduces portfolio risk in mutual funds. Rockenbach, Sadrieh and Mathauschek (2007) present experimental evidence that for a given expected return, groups, relative to individuals, are more likely to invest in portfolios with lower standard deviations. These findings are broadly consistent with the ones in this paper.

The social psychology literature on group decision making has also analysed the specific effect of group processes on different dimensions of group decisions, such as their extremity and riskiness. As discussed by Moscovici and Zavalloni (1969), a natural hypothesis is that the "group consensus" (the final choice made by a group) represents "an averaging, a compromise among individual positions". This idea is supported by a number of experimental research findings, such as those of Kogan and Wallach (1966), who find that group judgment represents the average of the prior individual judgments, even when consensus is achieved via group discussion of each prior judgment.

While this evidence is consistent with a moderating effect of groups, as we described in the introduction, there is also experimental evidence for cognitive biases, such as "risky shifts", "group polarization", and "groupthink", which may make group decisions less predictable and more extreme than individual ones. Risky shifts in groups occur when groups make riskier choices than those that would represent a compromise between the choices of the individuals comprising the group (Wallach and Kogan, 1965). Groups may also display "cautious shifts" (Stoner, 1968). Eliaz, Ray and Razin (2006) develop a model that links the phenomenon of choice shifts to well-known failures of expected-utility theory.

Group polarization arises when "members of a deliberating group move toward a more extreme point in whatever direction is indicated by the members' predeliberation tendency" (Sunstein, 2002). See also Cason and Mui (1997), who are among the first to incorporate the idea of group polarization in economics. For a recent model of group polarization, see Glaeser and Sunstein (2007).

Finally, the phenomenon of "groupthink" (Janis, 1982) may lead groups to ignore "hard evidence" at their disposal to irrationally pursue unanimity, which may lead to an increase in the riskiness and unpredictability of group decisions. Bénabou (2008) provides a formal model of this phenomenon.

Ultimately, whether group decisions are more or less risky than individual decisions is an empirical issue. It is possible that group polarization, risky shifts or groupthink may attenuate or even reverse the *compromise effect* of group decision-making. What is less clear from the experimental literature is how important these biases are when group formation is endogenous. Because the issue of group membership has received little attention in the previous literature,

5. From a theoretical perspective, the most closely related argument is in the work of Sah and Stiglitz (1986, 1991). In their models, because group members may disagree, group decision making entails a *diversification-of-opinions* effect. The final group decision will be a compromise, which reflects the different opinions of the group members.

in the next section we develop a simple model to clarify the relation between the *compromise* and the *membership effect*.

3. A SIMPLE MODEL OF MODERATION IN GROUPS

In this section, we provide a simple model to illustrate the two effects that may lead to moderation in naturally occurring groups: the *compromise effect* and the *membership effect*. Although these effects are quite intuitive, the model helps us clarify the relationship between them. In particular, we show that they are closely related: the membership effect decreases the variability of group decisions only when there is compromise in decision making.[6]

We model a situation in which group members must agree upon a single decision when all members have access to the same material information. We assume that there are no strategic communication issues.[7]

In order to make the analogy with our empirical application clear, we develop a model of a betting game in which the main decision is to make a guess x. A *bettor* (group or individual) must guess the realization of a single draw from the distribution of D. In a standard "guess-the-number" game, bettors win if they guess the exact number that is chosen. We call a possible draw of D a *break-up date*, and denote it by d. Break-up dates are integers drawn from a finite support $\mathcal{D} \equiv \left\{-\overline{D}, ..., 0, ..., \overline{D}\right\}$. For simplicity, we assume that the distribution of dates is symmetric with zero mean, and denote the probability of a break-up date d by $p_d > 0$, for all $d \in \mathcal{D}$.[8]

There is a large number of individuals, whose types $t \in \mathcal{D}$ represent the date they would like to guess, all else constant. That is, we assume that each individual has a preference for a specific guess. This is particularly realistic in our empirical application.[9] In general, group members could disagree on decisions due to differential information, heterogeneous priors, overconfidence, and the like, but assuming a preference for decisions is the simplest way of generating disagreement among group members.

For simplicity, we assume that the distribution of t in the population of individuals who may become group members is the same as the distribution of break-up dates:

Assumption 1. $\Pr(t) = p_t$, *all* $t \in \mathcal{D}$.

The main role of this assumption is to simplify notation, but incidentally it also implies that bettors do not have biased preferences, either for extreme or conservative guesses. Thus, in our model moderation can only arise due to the decision-making process and membership decisions.[10]

We now define the utility of an individual of type t who guesses as a member of a group of size n (an individual who bets alone is equivalent to a group of size $n = 1$). We assume that each individual only has enough funds to make a single guess, which costs c. If guesses

6. Although we model the membership effect as a consequence of individuals' decisions to be part of a group, it can also be modelled as the consequence of groups' decisions to restrict membership.

7. Alternatively, see Ottaviani and Sorensen (2006) for a model of strategic information transmission that also has implications for the variability of decisions.

8. Symmetry simplifies notation and facilitates the computation of equilibria.

9. For example, the sole winner of the 2008 Ice Classic was an Anchorage woman who placed guesses on the exact time that she was born.

10. Relaxing this assumption has no important implications for the results; it merely introduces a group "bias" towards some dates by assumption. Our results should be understood as being "net" of such biases.

are made by a group of size n, the group as a whole enters n guesses of a single date x.[11] We denote by $x_n(t_1, ..., t_n) : \mathcal{D}^n \to \mathcal{D}$ the rule that determines the guess of a group of size $n \geq 1$ with member types given by $(t_1, ..., t_n)$.

An individual t whose group makes a guess according to $x_n(t_1, ..., t_n)$ enjoys utility:

$$U_t = V[x_n(t_1, ..., t_n), t] - c + f - |t - E_t[x_n(t_1, ..., t_n)]|, \tag{1}$$

where $V[x_n(t_1, ..., t_n), t]$ is the expected payoff to a group guess made according to the rule $x_n(t_1, ..., t_n)$ conditional on knowing one's own preferred guess t, c is the cost of one guess, f is a positive number, and $E_t[x_n(t_1, ..., t_n)]$ is the expected group guess conditional on t. This utility function is valid both for individual bettors ($n = 1$) and for members of proper groups ($n > 1$).

This utility function can be broken down into two parts: the first one, $V[x_n(t_1, ..., t_n), t] - c$, measures the net pecuniary payoffs from betting, while the second one, $f - |t - E_t[x_n(t_1, ..., t_n)]|$, measures net non-pecuniary or psychological payoffs. The term f can be interpreted as a direct benefit from betting, or the value of "fun". Because such types of lotteries have negative expected monetary returns, it is necessary to assume either that bettors derive utility from betting directly or that they are risk-lovers. Therefore, f can also be interpreted as the (negative of the) risk premium if bettors are risk-lovers, in an alternative certainty-equivalent representation.

The term $|t - E_t[x_n(t_1, ..., t_n)]|$ represents the costs from not guessing one's preferred date t. Because the group guess $x_n(t_1, ..., t_n)$ when $n > 1$ will not necessarily be equal to the individual's preferred guess t, group bettors are usually worse off than individuals due to the need for compromise. We assume that compromise is more costly the farther the group guess is from the individual's preferred guess. Thus, we need to explain why individuals choose to join or remain as members of groups.

There are many ways of modelling group formation. Here we choose a simple framework that is sufficient to illustrate the main effects we wish to emphasize. We assume that each individual is randomly assigned either to a group of size $N > 1$ or to no group at all.[12] If assigned to a group, an individual must choose whether to stay with that group or to leave and make an individual guess, or perhaps not bet at all. We denote the (endogenous) group size after membership decisions are made by $n \leq N$. Each individual chooses whether to stay after learning his own type t but before the group decides the group guess $x_n(t_1, ..., t_n)$ (that is, we assume that individuals cannot leave their groups after group guesses become known). If an individual chooses to leave a group, he pays a "defection" cost of y. An individual who leaves a group pays this cost regardless of whether he bets alone or chooses not to bet. We normalize the utility (net of defection costs) of not making a guess to zero. Similarly, we could also have assumed that the "fun" value of betting in groups is larger for groups than individuals, with identical results.

If J is the jackpot and w_d is the number of winning tickets on date d, the payoff to one guess on date d is given by J/w_d. Let T be the total number of guesses. The jackpot is usually a fraction z of the revenue from bets cT, that is, $J = zcT$. In the Ice Classic, z is roughly 0.5.

11. This assumption is made for simplicity. Allowing groups to guess multiple dates creates no difficulties for the analysis, because the utility function (defined below) depends on the expected group guess. However, the analysis would change if groups could make fewer than n guesses. This would create an additional benefit from group membership, which is currently not modelled.

12. Strictly speaking, it is not necessary that groups are randomly formed, but only that group membership is not correlated with t.

We can now characterize the equilibrium distribution of guesses. A crucial assumption we make here is that there is a (potentially very) large number of unmatched individuals (i.e. individuals that are required to bet alone or not at all) relative to the number of groups. We focus only on equilibria with "large" numbers of guesses, so that the number of groups is relatively small compared to the total number of guesses.[13] Under these assumptions, the equilibrium distribution of guesses and the expected pecuniary payoffs to guessing a given day are determined by the behaviour of individual bettors alone. Thus it is possible to describe the equilibrium distribution of guesses without reference to $x_n(t_1, ..., t_n)$, unless $n = 1$, because the "marginal" players are individual bettors.

We consider a rational expectations equilibrium in which: (i) all players know the number of guesses of each date d in equilibrium, w_d^*, and the equilibrium jackpot, J^*, (ii) all players consider themselves too small to affect the equilibrium jackpot and the number of guesses of each date, (iii) all players form $E_t[x_n(t_1, ..., t_n)]$ in a Bayesian rational manner, and (iv) all individual bettors choose their guesses (i.e. choose their $x_1(t)$) in order to maximize their expected utilities as defined in (1).

The utility of an individual player of type t guessing date d in equilibrium is

$$U_t^* = p_d \frac{J^*}{w_d^*} - c + f - |t - d|. \tag{2}$$

It is straightforward to see that, in a rational expectations equilibrium, it must be that $x_1(t) = t$ for all $t \in \mathcal{D}$, all guesses from unmatched players should yield the same expected payoff, and (due to free entry of bettors) expected payoffs should be zero, implying that:[14]

$$p_d \frac{zcT^*}{w_d^*} = c - f, \tag{3}$$

which implies that the fraction of guesses of date d must be

$$\pi_d^* \equiv \frac{w_d^*}{T^*} = kp_d, \tag{4}$$

where $k \equiv zc/c - f$. Thus, the equilibrium distribution of guesses should be proportional to the true distribution of break-up dates. Because the condition above must hold for all d, it is clear that the only possible equilibrium requires $k = 1$. We assume that the game managers choose z and c so that this condition is satisfied.[15]

Result 1: *In a rational expectations equilibrium with a (sufficiently) large number of individual bettors: (1a) the proportion of guesses of a given date d should be equal to the true probability of break up on that date:* $\pi_d^* = p_d$*, for all* $d \in \mathcal{D}$*; and (1b) all guesses should yield the same net pecuniary payoff:* $p_d \frac{J^*}{w_d^*} - c = -f$*, all* $d \in \mathcal{D}$*.*

Now we turn to the question of individual versus group betting.

13. This assumption is realistic in our empirical application.

14. For simplicity, we ignore possible discontinuities here that may arise because of the discreteness of the setup. Formally, the problem can be made continuous by allowing for mixed strategies for the participation decision.

15. This does not eliminate multiple equilibria: the total number of guesses T is indeterminate. Thus, choosing $k = 1$ only guarantees that a rational expectations equilibrium is possible. Because any feasible profit for the game managers can be achieved with $k = 1$, there is no reason to assume that they will choose a different k.

Case 1: Exogenous membership.

We start by analysing the simplest case in which groups are formed exogenously and defection is not possible (i.e. the defection cost is $y = \infty$). When a group of n individuals has to decide on a single guess x, there must be a rule to determine the group's choice. Let t_i denote the preferred break-up date of member i of a given group. We assume a simple rule to model compromise:

Assumption 2. *A group's guess is the average of each individual's preferred guesses:*

$$x_n\,(t_1, ..., t_n) = \frac{1}{n}\sum_{i=1}^{n} t_i. \tag{5}$$

Clearly, given Assumptions 1 and 2, the unconditional expected group guess is $E[x_n(t_1, ..., t_n)] = 0$. If we denote the unconditional guess variance for an individual by σ^2, we have that the unconditional variance of a group guess is σ^2/n.[16] Although this is a straightforward consequence of the assumption that groups compromise, we state it as a result due to its importance for the empirical part:

Result 2 – The compromise effect: *If group size is exogenously determined, and if a group's guess is the average of its members' preferred guesses, then the variance of group guesses is lower than the variance of each of the preferred guesses of its members.*

Case 2: Endogenous membership.

The exogenous membership case neatly illustrates the compromise effect. When group membership is endogenous, however, the membership decisions of individuals can also affect the statistical distributions of guesses. Here we develop a very simple model of endogenous membership decisions and their effects on group moderation.

An individual of type t matched to a group of a still unknown size n needs to estimate the group bet $x_n\,(t_1, ..., t_n)$ before deciding whether to stay or leave the group. From the individual t's perspective, $x_n\,(t_1, ..., t_n)$ is a random variable because both n and $t_i \neq t$ are unknown. We assume that an individual cannot learn the types of the other group members before decisions are made. We economize on notation by denoting $x_n\,(t_1, ..., t_n)$ by x_n.

The equilibrium utility of a type t individual when deciding whether to stay with the group is

$$U_t^*\,(\text{stay}) = E_t\left[p_{x_n}\frac{J^*}{w_{x_n}^*}\right] - c + f - |t - E_t\,[x_n]|\,. \tag{6}$$

If he leaves, his expected utility is

$$U_t^*\,(\text{leave}) = \max\left\{p_t\frac{J^*}{w_t^*} - c + f, 0\right\} - y. \tag{7}$$

16. More generally, the compromise effect can be modelled as a weighted average of individual guesses, in which some members could be more influential than others. The qualitative Result 2 (below) would still go through.

The individual will compare these two utilities when deciding whether to stay or leave. Result 1 simplifies the problem: because the net expected pecuniary payoffs are always equal to $-f$ regardless of the date one guesses, the individual will choose to stay only if

$$|t - E_t[x_n]| \leq y. \tag{8}$$

We can now state the conditions for an equilibrium. Let $s(t) : \mathcal{D} \to [0, 1]$ be a function that assigns a probability that a player will stay with his group for each possible type in $\mathcal{D}$ and let $F(x)$ be the distribution of x_n. The functions $s^*(t)$ and $F^*(x)$ represent strategies and beliefs in equilibrium if and only if $F^*(x)$ is equal to the true distribution of x_n implied by Assumptions 1 and 2 and the equilibrium strategy profile $s^*(t)$, and

$$s^*(t) = \begin{cases} 1 \text{ if } \left|t - E_t^*[x_n]\right| < y \\ 0 \text{ if } \left|t - E_t^*[x_n]\right| > y \\ \theta \in [0, 1] \text{ if } \left|t - E_t^*[x_n]\right| = y \end{cases}, \tag{9}$$

where

$$E_t^*[x_n] = \int_{-\overline{D}}^{\overline{D}} x_n dF^*(x_n \mid t), \text{ all } t \in \mathcal{D}. \tag{10}$$

We restrict attention to symmetric equilibria, that is, to equilibria where type t's equilibrium strategy is the same as type $-t$'s equilibrium strategy. Once restricted to symmetric equilibria, it is possible to show that a unique equilibrium exists, and has a "threshold property":[17]

Result 3 – The membership effect: *There exists a unique symmetric equilibrium. The symmetric equilibrium is characterized by a unique integer t^* such that individuals with $t \in (-t^*, t^*)$ choose to stay as group members, while $t \notin [-t^*, t^*]$ leave the group.*

This result shows that endogenous membership generates a truncation of the distribution of guesses, such that individuals with extreme guess preferences (either too high or too low) choose not to participate. As a consequence, groups will never make extreme guesses. Intuitively, members leave the group in order to avoid the negative effects of compromise. Individuals with more extreme guess preferences are more likely to suffer from compromise and thus will choose to leave the group and bet alone (or not at all).[18]

Proof. Notice that

$$E_t^*[x_n] = E_t^*\left[\frac{t + (n-1)e}{n}\right], \tag{11}$$

where e is the average type of all other individuals who stay and n is the size of the group after membership decisions are made. Symmetry implies that $E^*[e] = 0$ for all groups and

17. Asymmetric equilibria may also exist and will all have a similar threshold property, although an asymmetric one. Without any ad hoc reasons for choosing an equilibrium, the unique symmetric equilibrium we describe below is appealing because it is the "average" or "median" equilibrium, if, for example, we consider each equilibrium to be equally likely.

18. Note that the exogenous membership case is a special case of the threshold equilibrium in which $t^* > \overline{D}$. This is the unique equilibrium for sufficiently high y.

also that e is mean-independent of group size n. Thus, condition (10) simplifies to

$$E_t^*[x_n] = tE^*\left[\frac{1}{n}\right] + E^*\left[\frac{(n-1)}{n}\right]E^*[e]$$
$$= tE^*\left[\frac{1}{n}\right] \equiv t\alpha.$$

We drop the conditioning on t for the expectation because n and e are independent of t. Let q_s denote the probability that a randomly chosen original group member (other than the member making this calculation) chooses to stay in an equilibrium. Define

$$\alpha(q_s) \equiv E^*\left[\frac{1}{n}\right] = \lim_{q \to q_s} \sum_{j=1}^{N} \frac{N-1!}{(N-j)!\,(j-1)!} \frac{q^{j-1}(1-q)^{N-j}}{j}. \tag{12}$$

It can be shown that $\alpha(q_s)$ is continuous and strictly decreasing in q_s (since increases in q_s make larger groups more likely, so that $E\left[1/n\right]$ is smaller), with $\alpha(0) = 1$ and $\alpha(1) = 1/N$.[19]

Using condition (8), an individual of type t will choose to stay with probability 1 in equilibrium if

$$\left|\left[1-\alpha(q_s)\right]t\right| < y. \tag{13}$$

If a symmetric equilibrium exists, this condition implies that there must be a threshold t^* such that all $|t| < t^*$ choose to stay while all $|t| > t^*$ choose to leave. A symmetric equilibrium is then fully characterized by a triplet $\left(t^*, q_s^*, \theta^*\right)$ such that the conditions below are satisfied:

$$q_s^* = p_0 + 2p_1 + \ldots + 2\theta^* p_{t^*}, \tag{14}$$

$$t^* = \left\lfloor \frac{y}{1-\alpha\left(q_s^*\right)} \right\rfloor, \tag{15}$$

$$\theta^* = \begin{cases} \theta \in [0,1] \text{ if } t^* = \frac{y}{1-\alpha\left(q_s^*\right)} \\ 1 \text{ if } t^* < \frac{y}{1-\alpha\left(q_s^*\right)} \end{cases}, \tag{16}$$

where the lower brackets denote the integer part function, that is, $\lfloor x \rfloor$ is the largest integer less than or equal to x.

Define the correspondence $f : [0,1] \to [0,1]$ as

$$f(q_s) = p_0 + 2p_1 + \ldots + 2\lambda p_{\left\lfloor \frac{y}{1-\alpha(q_s)} \right\rfloor},$$

where

$$\lambda = \begin{cases} \tau \in [0,1] \text{ if } \left\lfloor \frac{y}{1-\alpha(q_s)} \right\rfloor = \frac{y}{1-\alpha(q_s)} \\ 1 \text{ if } \left\lfloor \frac{y}{1-\alpha(q_s)} \right\rfloor < \frac{y}{1-\alpha(q_s)} \end{cases}.$$

19. The minimum group size n is 1 (and not zero), because we are considering the perspective of a member of type t assuming that he will stay with the group.

For each $q_s \in [0, 1]$, the set $f(q_s)$ is non-empty and convex, and the graph of f is closed. Thus, Kakutani's theorem applies, so there exists at least one fixed point $q_s' \in f\left(q_s'\right)$. For each q_s', condition (15) implies a unique $t^*\left(q_s'\right)$, and consequently condition (14) implies a unique $\theta^*\left(q_s'\right)$.

To confirm that the equilibrium is indeed unique, notice that f is non-increasing in q_s (in the sense that if $x > y$, $\max_\lambda f(x) \leq \min_\lambda f(y)$), which implies that the graph of f crosses the 45 degree line only once, implying a unique q_s^*. ‖

4. DATA

Bets can be placed on the Tanana River ice break-up date either by depositing tickets in red cans, which are located in more than 200 locations throughout Alaska, or by mailing in a guess. The tickets cost \$2.00 (in 2002) and are sold from 1 February to midnight on 5 April.

To determine the winner, a tripod is erected on the river ice. The tripod is connected to a clock so that a tripwire will signal the exact minute of break-up. To prevent bettors from tampering with the tripod, the tripod is guarded by watchmen 24 hours a day. The ticket holders whose guess is closest to the exact minute of break-up win the jackpot.

Our data consist of the Nenana Ice Classic's *List of Guesses* for 2002. This is the official compilation of guesses which serves to determine the winners of the Ice Classic. It lists, in order of date, hour and minute, the names of each bettor who guessed the ice would break at that time. Thus, we can track bettors' guesses using their names. In addition, bettors' names can be used to identify betting pools, which makes it possible for us to compare the betting strategies of individuals to those of groups. For example, we can classify bettors with names such as "7 Lucky Ladies", "Fat Freddie's 2 p.m. Coffee Club", "Chilly Dogs & Co." and "Gene Pool 2002" as pools.

Of course, we cannot tell if a given bettor is a pool if the pool's guesses are registered under the name of an individual. In 1997, for example, a winning pool containing two members was listed under an individual's name. We also cannot tell if a bettor is an individual if the bettor uses an imaginative pseudonym that appears to be a pool name. However, we believe that bettors will not "disguise" their identities very frequently for several reasons. First, pools have an incentive not to bet using an individual pool member's name to avoid potential conflicts over the distribution of the jackpot, especially since pool members may bet on their own as well as with the pool. Using a name distinct from an individual's name also makes it easier for a pool to make multiple guesses, because the pool need only enter its name on each ticket and register the pool by sending in a list of pool members to the Nenana Office. If multiple guesses are listed under an individual's name who acts as a representative of the pool and that individual also bets alone, the names of all other pool members must be listed on the back of *each* ticket belonging to the pool. Finally, it is clear from the names of some of the pools that part of the fun of being in a pool is to come up with a betting name. Because the Ice Classic tickets require bettors to fill in a name and mailing address (see the example at http://www.nenanaakiceclassic.com/TicketsBrochures.html), we also think that individuals betting alone are more likely to use their real names and addresses than to use pseudonyms. Thus, although we may sometimes misclassify pools as individuals if they bet under an individual's name and vice versa, we believe that such misclassification may add noise to our estimates but not systematically bias them. Nevertheless, we try to address potential measurement error in our pool classification in Sections 6 and 7.

We were only able to obtain a hard copy of the *List of Guesses*, which contained roughly 295,667 guesses.[20] We scanned the data using the scanning software OmniPage. Since the quality of the type in the hardcopy varied, and the scanning process introduced certain systematic errors (e.g. sometimes replacing the letter "M" with "14"), we subjected the data to an extensive cleaning process. It was straightforward to clean the guesses since they are listed in order by date and time in the *List of Guesses.* We end with a sample of 294,176 usable guess dates and 284,724 usable guess times.[21]

To clean the names, we developed a cleaning program which we describe in more detail in the Appendix. After applying the program, we also checked the bettors' names by hand to ensure that they were spelled uniformly. Since each guess is entered on a separate ticket, it is possible that name changes could occur which would make it difficult to track bettors' betting profiles. Since guesses made by betting pools may be entered by different people in the pool, we believed it to be more likely that spelling changes would occur for betting pools than for individuals.[22] Thus, we focussed primarily on making names with four or more words uniform. To make the bettors' names consistent, we changed names that occurred in multiple variations, for example, "Dave & Linda Huffaker" and "Linda & Dave Huffaker", to one of the forms in which the name occurred.

Because of the size of the dataset, it was impossible to examine each name to determine whether it corresponded to a pool or not. We proceeded as follows to identify individuals and pools. First, we classified names that contained obvious pool identifiers as pools. Such identifiers consisted of words such as "pool", "group", "company", "team", "family", and the word "and", as well as symbols such as "&" and "/", which were often used to separate individual names in a betting pool. Remaining observations were classified as individuals if the first word in the name appeared in the list of first names in the 2000 census. The observations not classified were then checked by hand to see if they belonged to pools with unusual names, e.g. "Alaska's Point of View". Checking by hand also enabled us to classify as individuals bettors who listed their last names first or who abbreviated their first names. With this procedure, we identify a total of 3093 pools. The most common pool identifiers are the character "&" (1030 pools), the words "co" (591) and "pool" (395), and the character "/" (243 cases). After cleaning, we are able to classify 290,051 guess dates belonging to 34,656 different bettors and 281,018 guess times according to whether they were entered by individuals or pools.[23]

20. The *List of Guesses* has 1479 pages with approximately 200 bets per page, except for page 1479 which contains only 67 bets.

21. It is easier to clean guess dates than guess times for several reasons. First, there are fewer dates in a year than minutes in a day, which makes it easier to infer correct dates (for example, by referring to the original text). Second, scanning errors often occur at the end of strings; for example, the last character is often dropped. Dates are provided in the format dd/mm/yyyy, where yyyy is identical for all bets, so scanning problems in the last four characters do not pose a problem. However, scanning errors at the end of times make it impossible to infer the correct time. Finally, times are provided in the format hh:mm A or hh:mm P, where A refers to a.m. and P to p.m. In general, it is impossible to rectify scanning error in the letter A or P, which means that the correct time cannot be inferred for those observations. We attribute the fact that we lose fewer guess times in our 2008 data to improvements in scanning software since 2002.

22. The Nenana Ice Classic also allows bettors to rubber stamp their names and addresses on tickets. Individuals are probably more likely to have such rubber stamps which makes name changes across bets less likely for individuals.

23. We have information on 553 non-missing guess dates and 498 non-missing guess times for dates between 1 February and 5 April 2002. Of these, 48 date and 45 time guesses were made by pools. These guesses may potentially be considered irrational, since the guessed break-up date occurs during the time in which tickets are sold, which means that the submission date for the bet potentially succeeds the guessed date. However, we do not exclude these from our analysis because we have no way of verifying whether they are irrational or not. Furthermore, we view this as further evidence that individuals place more weight on tails than pools do.

Because our distinction between individuals and pools is not error free, we analyse similar data for 2008 in Section 8. The 2008 data has special features which we can use to validate our bettor classification method, as well as to perform robustness checks. As we will discuss in Section 8, we do not find evidence of substantial mismeasurement induced by our classification procedure.

As is to be expected, pools make more guesses than individuals.[24] The mean (standard deviation) number of guesses entered by 3093 pools was 24.03 (77.70), with a maximum of 2158. The mean (standard deviation) number of guesses entered by 31,563 individuals was 6.84 (15.15), with a maximum of 1161.

5. ANALYSIS AND RESULTS

5.1. *Factors influencing guesses*

There are several key factors that we expect should play a role in a bettor's guesses. Since the ice break-up date is determined by environmental factors, information about weather and ice conditions should be important, although the extent to which the break-up day can be predicted is, of course, limited. Many bettors do appear to actively search out this information in order to make their guesses. For example, many of them call the Nenana Ice Classic office, asking about weather and snow conditions, according to the manager of the Ice Classic, Cherrie Forness (*Arctic Science Journeys*, 1997). Newspaper articles describe how bettors gather environmental information for their guesses (e.g. *The Seattle Times*, 1986; Richards, 1995; *Arctic Science Journeys*, 1997). The Nenana Ice Classic also posts a record of historical (since 1989) and current ice measurements together with the historical break-up dates on its website (www.nenanaakiceclassic.com); the historical break-up dates together with their frequencies are printed on brochures accompanying Ice Classic tickets. In addition, Alaskan television stations include updates about the thickness of the ice on the Tanana River in their weather coverage as the deadline for betting approaches (Finkel, 1998). Since bettors can only rely on information on current environmental conditions up until the day that betting closes, we expect bettors to take both current and historical information into account.

5.2. *The distribution of guesses in 2002*

Because we will compare the distributions of pool and individual guesses to examine whether pools are more moderate in their betting behaviour, it is useful to discuss some features of the overall distribution of guesses. Although it is difficult to say that one guess is "better" than another without knowledge of individual risk preferences, we can characterize the expected return and the shape of the guess distribution by making some simplifying assumptions.

In the Ice Classic, the optimal betting strategy for a single bettor depends on the strategies of all other bettors because the jackpot is evenly divided if there are multiple winners. According to the model in Section 3, if we assume a rational expectations equilibrium with free entry of bettors, all guesses should yield exactly the same net expected payoff in equilibrium. This is a straightforward consequence of our model (see Result 1b). Thus, the net expected payoff

24. The fact that bettors we classify as pools make more guesses than bettors we classify as individuals suggests that our classification reflects bettors' identities reasonably well and should lessen concerns about measurement error due to misclassification.

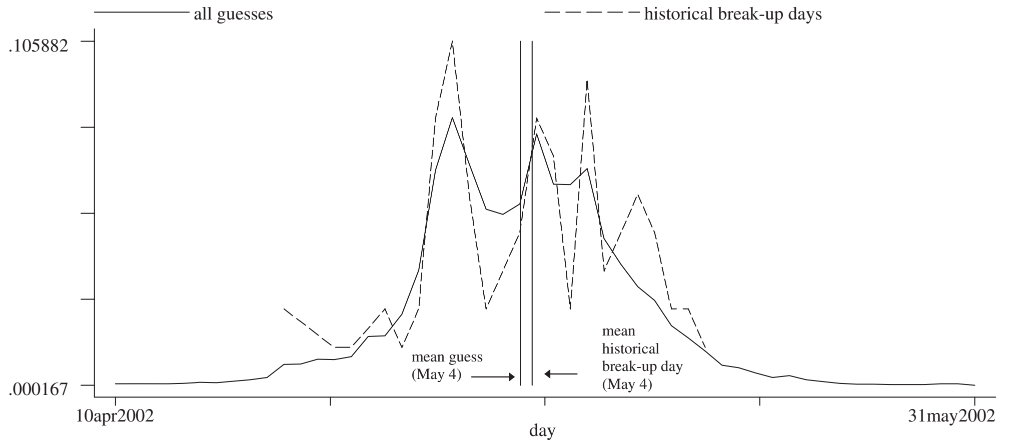

FIGURE 1

Comparison of frequency distribution of guessed dates in 2002 to historical break-up days (entire sample: 294,170 observations; historical dates from 1917 to 2001; only guesses from 10 April to 31 May displayed)

of any guess should be equal to −\$0.97, which is the total jackpot (\$304,000) divided by the total number of guesses (295,667) minus the cost of a ticket (\$2). As in many other wagering games, bettors are willing to participate in the Ice Classic even though the expected payoff is negative, either because they exhibit risk-seeking behaviour or they derive satisfaction from playing (i.e. "fun"). If bettors are risk-neutral, as in our model, our estimates suggest that the value of "fun" (f) per guess is −\$0.97.

These numbers are only meaningful if the actual betting behaviour is similar to the one implied by equilibrium play. In our model, this requires bettors to know the distribution of break-up dates. In reality, bettors do not know the true distribution of break-up dates. However, they do know the historical distribution of break-up dates, which we plot in Figure 1 together with the distribution of 2002 guesses.[25] If we are willing to assume that, as a first approximation, bettors view the historical distribution as the true one, Result 1a implies that the actual distribution of guesses should mimic the historical distribution of break-ups exactly ($\pi_d^* = p_d$, for all $d \in \mathcal{D}$). Because many assumptions of this simple model will not hold, we do not expect the two distributions to be exactly the same in our data. Still, the similarities are striking, as we can see from Figure 1, which suggests that this simple model provides a good approximation for the behaviour of bettors in this game. Figure 1 strongly suggests that bettors take the historical data into account when making their guesses, which gives rise to the unusual shape of the guess frequency distribution. In particular, the sample frequency appears to mimic all but the last spike in the historical data on 11 May.

To get a feeling for how well our model works in explaining the guess distribution in 2002, consider the historical mean break-up date (4 May). The historical probability of break-up on that day is roughly 7%, which is also the proportion of 2002 guesses for that day (see Figure 1). Thus, for the strategy of guessing the mean day, the model's prediction is fulfilled exactly. Looking at all dates, the theoretical prediction is off by at most 3.7 percentage points, and in most cases significantly less.

25. To make the graph easier to read, we report only the distribution of guesses between 10 April 2002 and 31 May 2002, a period which contains 99.53% (292,807) of the guesses.

TABLE 1
Summary statistics

Variable	Observations	Mean	Std. dev.	Min	Max
Historical break-up days	85	124.74 (4 May)	5.95	110 (20 Apr)	140 (20 May)
		Guess days			
Entire sample	294,170	124.04 (4 May)	6.64	33 (2 Feb)	363 (29 Dec)
Individual guesses	215,713	124.03 (4 May)	6.84	33 (2 Feb)	363 (29 Dec)
Pool guesses	74,332	124.25 (4 May)	5.94	33 (2 Feb)	320 (16 Nov)
Historical break-up time	85	858.34	281.38	4	1404
		Guess times			
Entire sample	284,718	839.87	260.67	0	1439
Individual guesses	208,765	838.44	263.89	0	1439
Pool guesses	72,247	844.21	251.22	0	1439

Notes: *Guesses* consist of guesses of the day, hour and minute that the ice on the Tenana River was going to break in 2002. We exclude six guesses of days prior to the date guesses could be made by buying tickets (1 February). *Historical break-up days* are the days the ice broke from 1917 to 2001. *Historical break-up time* is the minute within a day the ice broke from 1917 to 2001. *Individual guesses* are guesses made by individual bettors. *Pool guesses* are guesses made by betting pools. All dates are expressed in the number of days since 31 December. Standard deviations of days are measured in number of days. Time is measured in terms of minutes in a day, from 0 to 1439. Standard deviations of time are measured in minutes.

The evidence displayed in Figure 1 is quite striking. Despite its simplicity, the main prediction (Result 1a) of the equilibrium model in Section 3 seems to be supported by the evidence. In spite of the very low stakes involved and the rather lighthearted nature of the game, bettors appear to take the historical evidence seriously. Of course, it is impossible to determine exactly which factors affect bettors' choices, but Figure 1 suggests that history is an important one. In particular, the dates of 30 April, 5 May, and 8 May are clear attractors and guesses have a tendency to cluster around them.

5.3. *Pools versus individuals*

In Table 1, we present summary statistics for the historical ice break-up days from 1917 to 2001, the historical break-up times, the entire sample of guesses, and the guesses made by individuals and pools. All dates are expressed in the number of days since 31 December 2001, and all times are expressed in minutes of the day.[26]

The historical mean break-up day is 4 May, which is also the mean guess for the entire sample. Although prior literature has documented that groups may perform differently than individuals, *on average* pools did not behave much differently than individuals in our data. The mean guessed date for individuals (124.03) is earlier than for pools (124.25), but both fall within the same day. Thus, although the difference between the mean guess for individuals and for pools is statistically significant (the t-statistic is 7.8), it is arguably not economically significant. The historical mean break-up time is at 2:18 p.m. (858.34 minutes), while the mean guessed time in 2002 was 2:00 p.m. In 2002, the break-up time was at 9:27 p.m. Similarly to the pattern for the dates, the individuals' mean guessed break-up time is earlier (1:58 p.m.)

26. In total, six guesses were for dates prior to the opening of the ticket selling season, one of which was made by a pool. We exclude these guesses from our analysis, since it is irrational to guess a date that has already passed. Results do not change if these guesses are included.

than for pools (2:04 p.m.). These differences are again statistically significant, but whether a 6-minute difference is substantial is less clear.[27]

The break-up date in 2002 was 7 May. Individuals and pools display similar rates of ex post success: roughly 6% of each guessed 7 May. The ex post performance of both groups and individuals is quite good: the proportion that guessed the right day is significantly higher than the historical probability of break-up on that day (which is 2.35%). Interestingly, the largest difference between guess and historical probabilities for all days in our sample occurs on 7 May.

Result 2 from the model implies that the variance of group guesses should be lower than that of individual guesses. We find indeed that the standard deviation of pool guess dates is smaller than the standard deviation of individual guess dates by approximately one day (see Table 1). This result is not driven by differences in sample sizes, since the sample of individual guesses is much larger than the sample of pool guesses. As a consequence, the differences in standard deviations are also statistically significant. The p-value in a variance ratio test of the hypothesis that the standard deviation of individual guess dates is greater than the standard deviation of pool guess dates is less than 0.0001. Thus, group guesses appear to be more moderate than individual guesses.

To characterize the betting strategies in more detail, we analyse the frequency distribution of guesses. We focus primarily on guess dates, since the distributions for guess times, as measured in minutes, are less smooth which complicates the analysis. Figure 2 compares the frequency distribution of pool guesses to individual guesses. While pools place very similar weights on the three main historical spikes, individuals appear to place too much weight on the first spike (30 April) and too little on the third (8 May). Thus, the distribution of guesses made by groups appears to conform more to the distribution of historical break-up dates than the distribution of guesses made by individual bettors. This suggests that individual bettors rely relatively less on historical data than pools.

This pattern is not driven by individuals placing more weight on recent historical information. A similar pattern emerges if we use only the most recent 42 years (50% of the sample) of break-up history (results not reported).[28] It also appears that individuals placed less weight on current environmental information. According to the ice measurements available on the Nenana Ice Classic website, the average ice thickness of the Tanana River in 2002 was higher in the two full months during which tickets could be purchased, February (47″) and March (57.26″), than in 2001 (36.83″ and 39.72″, respectively). Since the break-up day in 2001 occurred on 8 May, four days after the historical mean break-up day, this was at least suggestive that the break-up day in 2002 might also be after 4 May, rather than before (as was indeed the case, since the ice broke on 7 May).

Result 3 of the model suggests that groups should be less likely to guess very early or very late dates than individuals. Consistent with this prediction, we find that the cumulative frequency of individual guesses of dates prior to the earliest recorded break-up day of 20 April (0.0156) is larger than the respective number for pools (0.0088). The cumulative frequency of

27. In most of our analysis, we do not emphasize the concept of statistical significance. There are two reasons for this. First, our sample comes very close to covering the entire relevant population. Thus, it is unclear what the gains from applying sample theory are in such cases. Second, with more than 290,000 observations, almost any difference we encounter will be highly statistically significant, even though it may not be economically significant.

28. In an article that appeared in *Science*, Sagarin and Micheli (2001) argue that the trend in historical ice break-up dates recorded by the Nenana Ice Classic indicates the presence of global warming. However, the article raised some controversy (Daly, 2001; O'Ronain, 2002), because, for example, it did not include the 2001 break-up day, which occurred four days after the mean historical break-up day of 4 May, in its analysis. Thus, the distribution of historical break-ups days until 2002 does not show clear evidence that betting earlier was a better strategy.

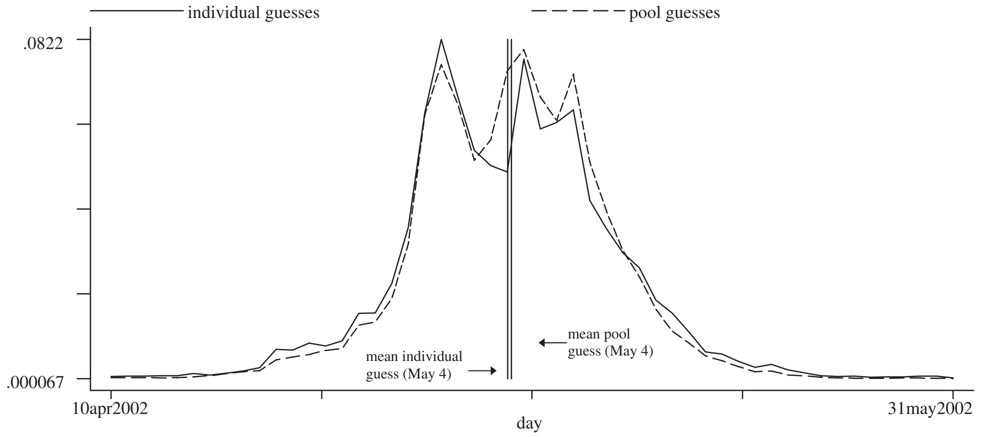

FIGURE 2

Comparison of frequency distribution of days of individual guesses to days of pool guesses in 2002 (sample of individual guesses: 215,713 observations; sample of pool guesses: 74,332 observations; only guesses from 10 April to 31 May displayed)

individual guesses of dates after the last recorded break-up day of 20 May (0.0015) is also larger than for pools (0.0008).

The fact that individuals are more likely to guess dates both earlier than 20 April and later than 20 May is consistent with our hypothesis that groups tend to make less extreme decisions than individuals. Figure 2 provides further evidence in favour of the idea that groups are more moderate. As is clear from the figure, individuals place more weight on the tails of the distribution than groups do. The two distributions cross only twice. To the left and right of the crossing points, the individual distribution always lies above the pool distribution, whereas the pool distribution lies above the individual distribution in between the crossing points. The cumulative frequency of individual (pool) guesses to the left of the earliest crossing point (on 3 May) is 0.417 (0.377); the cumulative frequency of individual (pool) guesses to the right of the latest crossing point (11 May) is 0.102 (0.082). Furthermore, the cutoffs for 5%, 10% and 20% of the individual guess distribution are in each case a day earlier than the corresponding cutoffs for the pool guesses. The cutoffs for 95% and 90% of the individual guess distribution are both a day later than for the pool guess distribution (the 80% cutoff is the same for both distributions).

Our strategy of estimating the probability density functions of guesses with histograms has potential pitfalls. In particular, the excessively "jagged" appearance of the functions displayed in Figures 1 and 2 suggests that one could improve the precision of the estimates by using a non-parametric smoothing procedure. In Figure 3, we display kernel density estimates of the individual and pool guess distributions using the Epanechnikov kernel function with a bandwidth of two days. This choice of bandwidth is just sufficient to make the distributions look "single-peaked". The key feature of Figure 3 is that it highlights, even more strongly, the fact that individuals tend to place more weight on the tails of the distribution. Our results, therefore, do not appear to be driven by undersmoothing.[29] If anything, the differences between

29. Because we do not use information about the guessed minutes in Figures 1 and 2, our graphs of frequency distributions also represent smoothed versions of the true frequency distribution of guesses. The same is true of the historical break-up frequency.

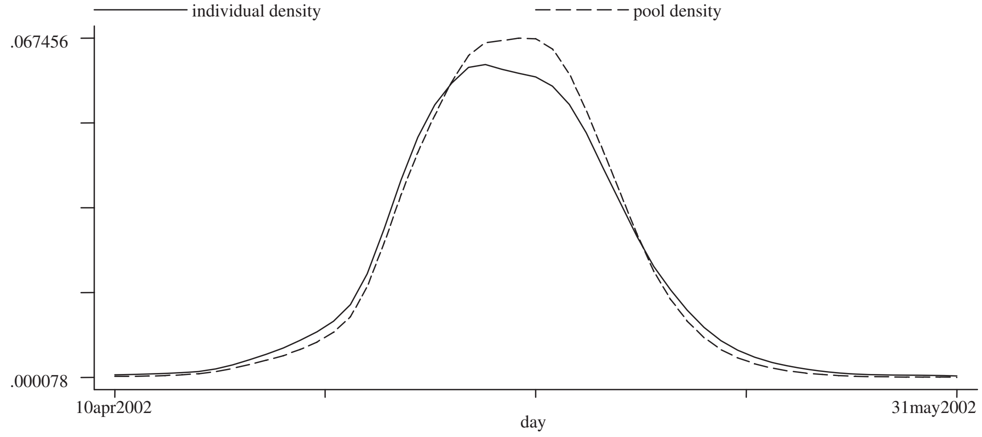

FIGURE 3

Kernel density estimates of days of individual and pool guess distributions in 2002 (sample of individual guesses: 215,713 observations; sample of pool guesses: 74,332 observations; only guesses from 10 April to 31 May displayed. Densities estimated using the Epanechnikov kernel function with a bandwidth of 2 days)

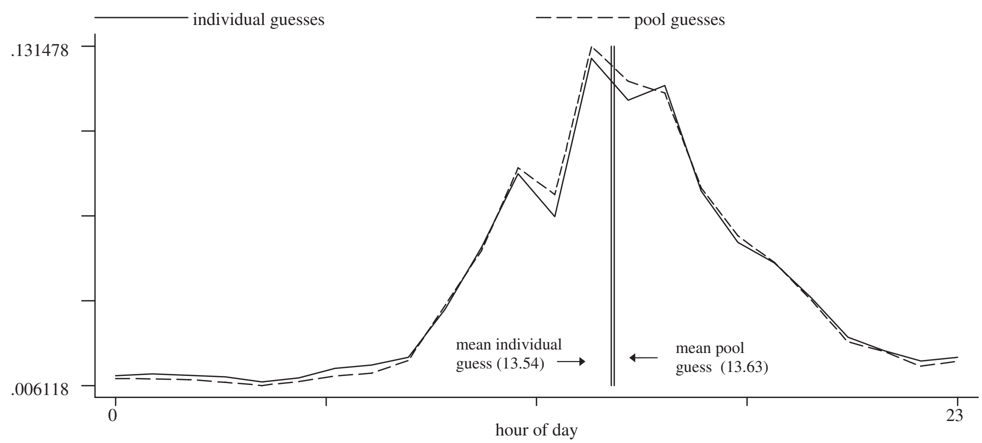

FIGURE 4

Comparison of frequency distribution of hours of individual guesses to hours of pool guesses in 2002 (sample of data on hours of individual guesses: 208,765 observations; sample of data on hours of pool guesses: 72,247 observations)

the weights which groups and individuals assign to the tails appear to be much more pronounced when we use kernel estimation techniques.

From Table 1, we see that the standard deviation of individual guess times is 12 minutes larger than the standard deviation of pool guess times (this difference is statistically significant), which is consistent with the pattern we find for the guess days. In Figure 4, we plot the histograms for individuals and pools based on the hours of the guessed times. A similar pattern as in Figure 2 emerges: individuals are more likely to guess both earlier and later hours of the day than pools. The differences between the two histograms are much less pronounced, however, with the distributions crossing each other six times, and overall behaving very similarly. We conclude that the evidence for the hours and minutes of the day also suggests that pools make less extreme guesses, although the differences between the guess time distributions are less pronounced than for the guess dates.

6. SELECTION, MEASUREMENT ERRORS AND ADDITIONAL EVIDENCE

So far, we can only interpret our evidence in terms of the *joint hypothesis* that groups make more moderate decisions either because they require compromise among their members or because individual characteristics that are correlated with betting behaviour affect group membership (or both). But, in order to obtain a clearer picture of the differences between group and individual behaviours, it would be nice to know whether our results are completely driven by selection, for reasons other than the ones we have discussed in the theory section.

Our hypothesis is that we observe more moderate guesses in groups either because group members compromise or because individuals choose not to be group members to avoid compromising (or equivalently that groups terminate membership of individuals not willing to compromise). Our model clarifies the role of actual or potential compromise for membership decisions. However, it could also be that personal characteristics (and in particular characteristics other than betting preferences) bring similar people together, so that groups may appear more moderate despite the absence of actual or potential compromise. In order to address this possibility, in this section we attempt to control for potential selection on personal characteristics.

There are two main types of selection mechanisms that are relevant for our analysis. First, individuals choose whether or not to participate in the game by making a guess. Second, conditional on participation, individuals choose whether to make guesses individually, in groups, or both.

The first selection problem is virtually impossible to address. This is not necessarily a problem, because it is still an interesting question to compare the betting behaviour of groups versus individuals conditional on the decision to participate. Of course, if the characteristics that affect participation also affect the behavioural differences between groups and individuals, then it would be difficult to generalize our results to different settings.[30]

It is also difficult to examine the second selection issue, because it is not clear which types of individuals self-select into groups. Barber *et al.* (2003) face similar problems in their study of stock clubs, but argue that their experimental evidence suggests that selection is not likely to be a major issue. Hamilton, Nickerson and Owan (2003) provide evidence that more productive workers self-select into teams. However, Kocher, Strauß and Sutter (2006) show the opposite in the context of a beauty-contest game. Individuals self-selected into groups hoping to improve performance, while individuals who opted to stay alone appeared to value autonomy in decision making. Nevertheless, we use three complementary methods to examine whether selection alone can explain our results.

1. *Individual fixed effects*: we use individual fixed effects in a subsample of individuals who appeared to make guesses both as individuals and as part of a pool. This analysis suggests

30. Because the Ice Classic has a longstanding tradition in Nenana and Alaska, selection due to the decision to participate may not be as important a concern as in other games. In our data, we identify 31,563 different individual bettors. The 2000 US census estimated the populations of Alaska and Nenana to be 626,932 and 402, respectively. Even though people outside Alaska can bet, it is more difficult because they must contact the Ice Classic office directly. If we assume that most bettors are from Alaska, this suggests that roughly 5% of the population placed a bet. This is a substantial fraction of the population. Because accounts of the Ice Classic picture it as a community affair that provides part-time employment for roughly 100 residents, it is likely that most Nenana residents bet.

that individual characteristics other than betting preferences do not fully explain the difference between group and individual guesses.

2. *Family pools*: we identify groups that appeared to be formed on the basis of family ties. Such groups are less likely to be formed as a consequence of the similarity of opinions of their members, although family pools are also possibly more homogeneous since family members may have similar opinions.

3. *Pool size*: we identify a subsample of pools for which we can infer group size. We then analyse the effect of group size on the variability of guesses.

None of these approaches rules out the importance of selection, but together they provide evidence that selection alone is not sufficient to explain our results. Because family pools and pools for which we can infer the number of members are less likely to be misclassified as pools, this analysis also acts as a robustness check that measurement error in our pool classification is not driving the results.

6.1. *Individual fixed effects*

It is possible that individuals with more moderate predictions will be more likely to join a betting pool. One reason is the membership effect we discussed in Section 3: people with more extreme predictions may choose not to join groups if they expect groups to compromise (or similarly groups may refuse membership to individuals with extreme predictions). Another possibility is that omitted personal characteristics may be correlated both with moderation in betting preferences and with group membership decisions. If the latter is true, then groups may appear more moderate even if there is no (actual or expected) compromise effect.

To address this issue, we construct a dataset of guesses placed by individuals who also bet in groups. We then use individual fixed effects to identify within-individual variation in betting strategy that can be attributed to betting in a group. For example, if we define b_{ijk} to be a measure of betting strategy for guessed date k of individual i who also guesses as part of pool j (where pools include pools with multiple members and the extreme "pool" of betting alone), POOL_j to be a dummy variable which is equal to 1 if POOL_j has more than one member and 0 otherwise, γ_i to be individual fixed effects and ε_{ijk} to be error terms, then *for the same individual* we can identify the effect of betting in a group on betting strategy by estimating the following regression:

$$b_{ijk} = \alpha + \beta \text{POOL}_j + \gamma_i + \varepsilon_{ijk}. \tag{17}$$

While ex ante it seems difficult to identify individuals who also bet in groups in our dataset, we often observe what appear to be married couples betting together as well as individually. For example, we observe guesses placed by "Ray & Ruth Birky", as well as guesses placed by "Ray Birky" and "Ruth Birky". We also observe individuals who appear to be betting with a pool, for example, "Andrea F. Staniforth & Co" and "Helen Lazeration Pool" as well as on their own, for example, "Andrea F. Staniforth" and "Helen Lazeration". Under the assumption that no other individual with exactly the same name is betting, for example, there are not two different bettors called "Andrea F. Staniforth", and that guesses associated with such individuals' names are truly individual guesses, we can construct a dataset of guesses placed by individuals who also bet in pools. While these assumptions are strong, we believe the second assumption would simply bias us against finding a result. If guesses associated with an individual's name who also bets in a pool are also pool guesses, then β should not be significantly different from 0. It is less clear how the first assumption would affect our results.

If the guesses we associate with, for example, "Andrea F. Staniforth", are individual guesses but are not those belonging to the Andrea F. Staniforth of "Andrea F. Staniforth & Co", then we are still comparing guesses of individuals to guesses of pools but in this case the individual fixed effect only captures the effect of having the same name and is essentially meaningless. We believe it unlikely that many bettors with *exactly* the same names will be present in our data,[31] but our results should still be interpreted with care given that we cannot cross-check the identity of the individuals.

To construct our sample for this section, we examined all pools that contain the names of individuals and searched for other occurrences of the individual's name. We were able to identify 351 individuals whose name appeared alone as well as in the name of a pool. Our sample of guesses consists of all guesses made by individuals whose name also occurred in the name of a pool as well as those of the associated pools.

We examine three measures of an individual's betting strategy: the guess, the absolute value of the distance of the guess to the sample mean guess and the square of this absolute value. We focus primarily on the latter two measures of guess dispersion.

In Table 2, we report the results of estimating the above equation for both guess dates and guess times (minutes of the day).[32] From columns I and IV, we see that groups guessed earlier dates but later times. However, these differences are not large: mean pool and individual guesses are on the same day, and pool guesses are on average 17 minutes later than individual guesses. More important are the results for the variability measures. Columns II, III, V and VI all show that both the absolute value and the square of the deviation from the mean are lower for pool guesses, both for the guess day and for the guess time. Consider, for example, column III. The coefficient on the constant can be interpreted as the variance of individual guesses, so the implied standard deviation of individual guesses is 6.72 days, a number that is slightly lower than in Table 1 (6.86 days). The implied standard deviation for pools is 6.29 days, which is higher than in Table 1 (5.95 days). If we consider the guess time instead, the implied standard deviation for individuals is 278.86 minutes, which is higher than in Table 1 (263.89 minutes), while the implied standard deviation for pools is 238.87, which is lower than in Table 1 (251.23).

The results show that group guesses display lower standard deviations than individual guesses, for both guess days and guess times, in a subsample of guesses by individuals who made guesses both individually and as part of a group. Thus, individual characteristics may explain some of the difference between the standard deviations of the two groups, but not all of it. These results appear even more striking if one recalls that all regressions include 351 individual dummy variables.

Why would an individual who chooses to make multiple guesses make his more extreme guesses alone? Either groups choose a betting rule that leads to compromise or individuals choose not to make their extreme guesses in groups.[33] Similarly, the group members betting together with such an individual may refuse to be part of a betting pool that places an extreme

31. As robustness checks, we created measures of "common" names, like "Johnson", and "uncommon" names, like "Arrants". To help in our classification, we used the number of unique individuals with the same last name as a guide. For example, there were roughly 233 individual bettors with the last name "Johnson", but only one bettor with the last name "Arrants". Thus, "Johnson" is common, while "Arrants" is uncommon. We redid our analysis after excluding the common names, since for the uncommon names it was less likely that another individual with exactly the same name was betting. Naturally, this procedure reduces our sample substantially, so our results are statistically weaker. Nevertheless, the signs of our coefficient estimates are the same as when we use our full subsample.

32. Reported constants are averages over all individual effects.

33. Of course, this does not rule out the possibility that individuals expect groups to compromise even when they do not.

TABLE 2
Individual fixed effect results for individuals betting alone and in groups

	I	II	III	IV	V	VI
	Guess date (in days)			Guess time (in minutes)		
	Guess	Abs. distance to mean guess	Square distance to mean guess	Guess	Abs. distance to mean guess	Square of distance to mean guess
Pool dummy	−0.712***	−0.242*	−5.657**	17.582*	−29.381***	−20,704.424***
	[3.50]	[1.84]	[2.09]	[1.89]	[4.86]	[4.96]
Constant	124.200***	5.110***	45.168***	836.586***	210.030***	77,761.704***
	[1003.41]	[62.61]	[26.15]	[139.80]	[54.21]	[28.49]
Observations		5431			5268	

Notes: The sample consists of observations on individual and group guesses for 351 individuals whom we classified as betting on their own and in a group. We exclude six guesses of days prior to the date guesses could be made by buying tickets (1 February). An individual was considered to bet on her own and in a group if her name appeared both on its own and as part of a group name in the list of bettors, for example, "Lois Swanberg" and "Lois Swanberg & Bob Hager". Columns I–III show individual fixed effect regressions of the guessed day, the absolute value of the distance to the sample mean guessed day and the square of that distance on a group dummy. This dummy is defined to be equal to one if the bettor's name suggests it is a group with more than one member, for example, "Lois Swanberg & Bob Hager" and is 0 otherwise. The guess day is measured in days of the year since 31 December. Columns IV–VI replicate the regressions in columns I–III using the guessed time. The guess time is measured in minutes of the day. All standard errors are corrected for heteroscedasticity. The absolute values of the robust t-statistics are in brackets. *, **, *** indicate significance at the 10%, 5% and 1% level, respectively.

bet. All these possibilities are plausibly related to betting preferences and the expectation of compromising in groups. However, what the fixed effects results rule out is the possibility that moderation in groups is fully driven by personal characteristics (other than betting preferences) that determine the match between group members.[34]

6.2. *Family pools*

To address the possibility that our results may be driven by the fact that individuals self-select into groups based on similarities in betting preferences, we try to identify pools that are formed for reasons unrelated to betting preferences. Family pools are a natural example of pools formed for "exogenous" reasons, because families are likely to bet together even if they disagree about dates. While we cannot control for initial selection (whether families choose to participate at all) or family member inaction (some family members may not participate in the decision making process), we can be reasonably confident that family composition is not driven by betting preferences. However, it is still possible that family pools are more homogeneous in their opinions than non-family pools.

34. The fixed-effect analysis does not hold the characteristics of other group members constant. However, we can control for partner effects to a certain extent by restricting our sample to individuals who bet alone and with members of the opposite sex. For example, we can restrict our sample to women who bet alone and in groups with only one other person who is male, for example, Barbara Bluekens versus Barbara & Thomas Bluekens. In this restricted sample, we find a coefficient of −0.449 (significant at the 10% level) on the pool dummy in an individual fixed effect regression with the absolute value of the distance to the mean guess as the dependent variable. This is at least suggestive that our results are not fully driven by characteristics of partners that may determine the match between group members.

TABLE 3
Comparison of individual guesses to guesses of family pools

	I	II	III	IV	V	VI
	Guess date (in days)			Guess time (in minutes)		
	Guess	Abs. distance to mean guess	Square distance to mean guess	Guess	Abs. distance to mean guess	Square of distance to mean guess
Family pool	−0.322***	−0.159***	−6.485***	8.364***	−9.762***	−9154.122***
	[5.13]	[3.90]	[5.61]	[3.38]	[6.04]	[8.66]
Constant	124.029***	4.988***	46.787***	838.444***	196.225***	69,636.798***
	[8421.59]	[495.03]	[64.81]	[1451.72]	[508.12]	[264.35]
Observations		226,511			219,202	

Notes: The sample consists of observations on guesses made by individuals and 965 family pools. We exclude six guesses of days prior to the date guesses could be made by buying tickets (1 February). A pool is classified as a family pool if it is clear that the members of the pool have the same last name, e.g. the bettor's name is "Ray & Ruth Birky", or are members of the same family, for example, "John Pickett & Sons". Columns I–III show regressions of the guessed day, the absolute value of the distance to the sample mean guess day and the square of that distance on a "Family pool" dummy. This dummy is defined to be equal to one if the bettor is classified as family pool and is 0 if the bettor is classified as an individual. The guess day is measured in days of the year since 31 December. Columns IV–VI show the same regressions using the guess time. The guess time is measured in minutes of the day. All standard errors are corrected for heteroscedasticity. The absolute values of robust t-statistics are in brackets.
*, **, *** indicate significance at the 10%, 5% and 1% level, respectively.

Our subsample for this section is the set of observations on guesses made by individuals and family pools. We classify a pool as a family pool if it is clear that the members of the pool have the same last name, e.g. the bettor's name is "Ray & Ruth Birky", or are members of the same family, for example, "John Pickett & Sons". Of the 3093 pools in our sample, 965 (31.12%) are family pools.

Table 3 shows the results of similar regressions as in Table 2 (without fixed effects). The results are similar to those for the full sample. For the guess days, the implied standard deviation for individuals is 6.86 days, which is identical to the number reported in Table 1 (as it should, because the sample of individuals is not restricted), while the implied standard deviation for pools is 6.35 days (higher than the number in Table 1: 5.95). For the guess times, the implied standard deviation for individuals is 263.89 minutes, while the implied standard deviation for pools is 245.93 days (lower than the number in Table 1: 251.23).

We conclude that family pools exhibit less extreme betting behaviour than individual bettors. These results also help to address concerns about misclassification, as family pools (or pools that use family names or family identifiers) are straightforward to classify as pools.

6.3. *Variation in pool size*

Our model implies that guess variance should decrease with pool size if membership is exogenously determined (see Result 2).[35] Thus, we could reject the exogenous membership hypothesis if we find evidence that the variance is invariant or increasing in pool size. We investigate this issue in a subsample of observations on guesses placed by pools whose names

35. By exogenous membership we mean that betting preferences are not the reason groups are formed, not that groups are formed randomly.

TABLE 4
Variation in guesses according to pool size

	I	II	III	IV	V	VI
	Guess date (in days)			Guess time (in minutes)		
	Guess	Abs. distance to mean guess	Square distance to mean guess	Guess	Abs. distance to mean guess	Square of distance to mean guess
Pool size	0.002	−0.052***	−0.796***	1.014***	−1.054***	−562.344***
	[0.95]	[36.79]	[8.68]	[8.64]	[13.07]	[11.97]
Constant	123.868***	4.540***	40.849***	825.852***	219.411***	84,446.079***
	[2268.21]	[119.87]	[13.15]	[333.56]	[134.40]	[74.37]
Observations		18,964			18,434	

Notes: The sample consists of observations on guesses made by 1244 pools whose names could be used to proxy for their sizes. We exclude six guesses of days prior to the date guesses could be placed by buying tickets (1 February). A pool's name was considered to reflect its size if its name consisted of a list of individual names, for example, "Brett Miller/Bruce Atkinson" (a size of 2) or its name suggested the number of members, for example, "7 Lucky Ladies" (a size of 7). To control for numbers that were potentially introduced into a pool's name by scanning error, we restrict ourselves to pools whose name occurs more than once as a bettor. Columns I–III show regressions of the guessed day, the absolute value of the distance to the sample mean guess day and the square of that distance on pool size. The guess day is measured in days of the year since 31 December. Columns IV–VI show the same regressions using the guessed time. The guess time is measured in minutes of the day. All standard errors are corrected for heteroscedasticity. The absolute values of the robust t-statistics are in brackets.
*, **, *** indicate significance at the 10%, 5% and 1% level, respectively.

can be used to proxy for their sizes. We assume a pool's name reflects its size if its name consists of a list of individual names, for example, "Brett Miller/Bruce Atkinson" (a size of 2) or its name suggests the number of members, for example, "7 Lucky Ladies" (a size of 7). We were able to assign a pool size to 40.22% of pools (1244) because a large number of pools consists of couples (1149). Average pool size is 2.53 with a maximum of 50. To control for numbers that were potentially introduced into a pool's name by scanning error, we restrict ourselves to pools whose name occurs more than once in the list of bettors.

In Table 4, we report the results of the same types of regressions as in Table 3. For both dates and times, guess variability decreases with pool size. Our results imply that a pool of size 2 would have a standard deviation of guess dates of 6.27 days, while a pool of size 7 would display a standard deviation of 5.93 days. The analogous effect for guess times is smaller, with a change in the standard deviation of roughly 6 minutes.

The evidence from this table suggests that group size is negatively related to the variability of guesses. Thus, larger groups may be characterized by more moderation in decisions, or larger groups attract individuals with more moderate opinions, or both. Since we find similar results when we compare pools of different sizes as when we compare pools and individuals, misclassification of bettors into individuals and pools does not appear to be the primary explanation of our results.

7. THE 2008 DATA: ADDITIONAL EVIDENCE AND ROBUSTNESS

Although we discussed potential measurement error problems and addressed them in different ways in the previous section, in this section we reexamine the measurement error problem in order to get a sense of its magnitude. As we discussed earlier, there are several reasons to believe that our results are not generated by measurement error. Instead, it is more likely that measurement error introduces noise that will cause attenuation bias, making it harder for us to find that pools are more moderate. The question is, how large will this bias be?

To examine this issue, we tried to obtain a list of pools that registered their members with the Nenana Ice Classic office in 2002.[36] For these pools, there can be no doubt that they are pools, so there is no measurement error associated with the assignment of their guesses. Unfortunately, we were unable to obtain this information for 2002. However, we were able to obtain this information for 2008 along with the Nenana Ice Classic's *List of Guesses* for 2008. We define an *official pool* to be a pool that registered its name with the Nenana Ice Classic office. We followed a similar procedure for scanning the data as for the 2002 *List of Guesses* and used a similar pool classification method to assign guesses to pools (other than the official ones) and individuals as in 2002. However, we used only a simplified cleaning procedure, because the full cleaning procedure we employed for the 2002 data was too time intensive. This limits the usefulness of the 2008 data (e.g. we do not have enough detailed data to replicate the analysis from Section 6 fully). However, the availability of registered pool data is a major advantage.

The 2008 *List of Guesses* contains 239,579 guesses. After scanning and cleaning, we are able to classify 238,684 guess dates placed by 41,977 different bettors and 238,307 guess times according to whether they were placed by individuals or pools.[37] We identify a total of 4998 pools, 150 of which are official pools. These 150 pools account for 27.56% of all pool date guesses (49,336) and 27.54% of pool time guesses (49,262). This suggests that pools are more likely to register their members with the Nenana Ice Classic office if they place many guesses. Indeed, official pools placed on average 83.7 more guesses than other pools, a difference which is significant at lower than the 1% level.

As a first step towards assessing the measurement error problem, we examine the list of official pools to see whether our method of assigning bettors to the pool category in 2002 is reasonable. Two pool names contain the identifier "&", eight names contain one of the words "bunch", "clan", "committee", "company", "pack", "partners", "kids" and "sisters", two names contain the word "gang", two names contain the word "crew", two names contain the word "group", two names contain the word "staff", two names contain the word "team", five names contain the word "family", 15 names contain a number, for example, "6 R US" and "The Three Amigos", and 81 names contain the word "pool". The remaining names are names such as "CHILLY DOGS" and "Northwest Airlines" that do not contain a clear pool identifier, but that also do not contain the name of an individual in it. Thus, these are names that we would not have assigned to the individual category based on our merge with census data. Instead, we would have classified these as pools by hand. We conclude that we would have correctly identified each official pool as a pool based on our 2002 classification method. We also note that no official pool is registered under the name of an individual without an accompanying pool identifier. This is at least suggestive that bettors who bet under individual names are not likely to be pools.

Next we use the 2008 data to compare the betting behaviour of pools and individuals.[38] This serves as a robustness check that our findings are not specific to a given year. We replicate our previous figures using the entire sample of guesses. The figures are similar to those in 2002.

36. The Nenana Ice Classic encourages pools to register their members to avoid disputes.

37. In 2008, ten guesses were for dates prior to the opening of the ticket selling season, none of which were placed by pools. We exclude these bets from our analysis, since it is irrational to guess a date that has already passed.

38. We have information on 511 non-missing guess dates and 510 non-missing guess times placed for dates between 1 February and 5 April, which is the period during which guesses could be made by buying tickets. Of these, 66 date and time guesses were made by pools. These numbers are comparable to the numbers for 2002.

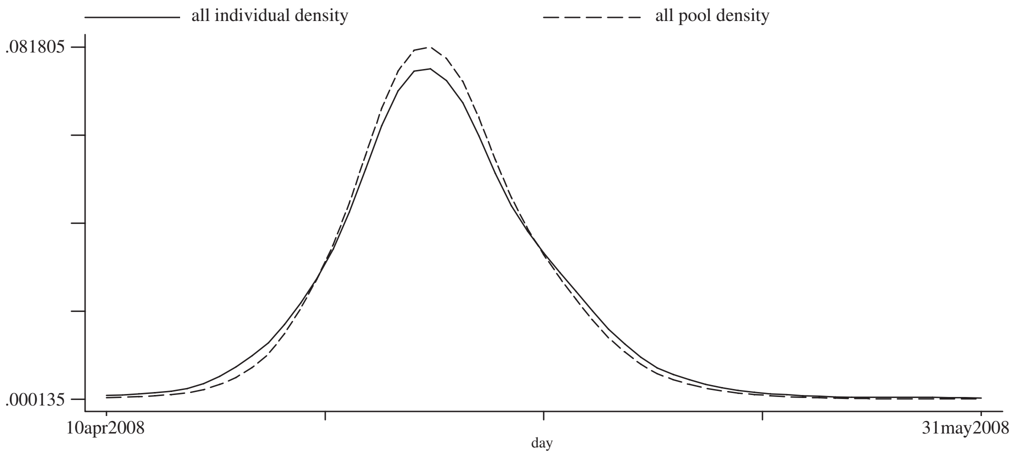

Figure 5

Kernel density estimates of days of individual and pool guess distributions in 2008 (sample of individual guesses: 189,348 observations; sample of pool guesses: 49,336 observations; only guesses from 10 April to 31 May displayed. Densities estimated using the Epanechnikov kernel function with a bandwidth of 2 days)

For the sake of brevity, we provide only the replication of Figure 3 in Figure 5. From Figure 5, it is clear that in 2008 individuals also tend to place more weight on the tails of the distribution than pools.

Finally, as an alternative method of addressing potential measurement error problems, we compare the behaviour of *official* pools to the behaviour of individuals. Because we are certain that official pools are indeed pools, this comparison will be subject to measurement error problems only if pools bet under the names of individuals. The official pool classification thus represents the cleanest measure of pools available and we can use it both to verify that our previous results are not an artifact of measurement error and to assess the magnitude of the measurement error problem.

Because the official pools represent a subsample of all pools, we first compare official pools to a random subsample of individuals with roughly the same amount of guesses as the official pools. We construct the sample of individuals by sorting them alphabetically and truncating the list of names when the cumulative number of individual guesses is at least as large as the number of official pool guesses (13,598 guess dates and 13,568 guess times). This gives us a sample of 2612 individuals with 13,617 guess dates and 2607 individuals with 13,571 guess times. To ensure that our choice of random sample is not biasing the results, we also compare official pools to the sample of all individuals.

We report the results in Table 5. We first regress the guessed date, the absolute value of the distance of the guess date to the sample mean guess, and the square of this absolute value on a pool dummy variable. "Pool dummy" is defined to be one if a bettor is classified as a pool using our 2002 methodology. We show the results for three different sample types: the full sample, official pools plus random individuals, and official pools plus all individuals. The top panel of Table 5 shows the results. The bottom panel of Table 5 shows the same regressions using the guess time.

Overall, we find that the 2008 data deliver results that are qualitatively identical to the 2002 results (i.e. groups are more moderate). The pool dummy enters with a negative sign in

TABLE 5

Variation in guesses in 2008

	Full sample			Random individuals + official pools			All individuals + official pools		
	I	II	III	IV	V	VI	VII	VIII	IX
	Guess	Abs. distance to mean guess	Square distance to mean guess	Guess	Abs. distance to mean guess	Square distance to mean guess	Guess	Abs. distance to mean guess	Square distance to mean guess
	Guess date (in days)								
Pool dummy	−0.045	−0.538***	−11.493***	−0.672***	−0.932***	−25.752***	−0.675***	−0.775***	−16.590***
	[1.50]	[25.07]	[8.05]	[8.77]	[16.31]	[4.70]	[14.35]	[23.33]	[6.72]
Constant	121.294***	4.525***	43.966***	121.292***	4.683***	53.127***	121.294***	4.525***	43.966***
	[7959.99]	[406.34]	[58.76]	[1941.77]	[97.84]	[10.72]	[7959.98]	[406.34]	[58.76]
Observations	238,684			27,215			202,946		
	Guess time (in minutes)								
Pool dummy	3.041**	−4.436***	−4,849.930***	9.282***	−6.515***	−6,249.972***	3.413*	−8.194***	−8,033.334***
	[2.50]	[5.44]	[8.79]	[3.21]	[3.37]	[4.83]	[1.65]	[5.98]	[8.99]
Constant	867.780***	181.822***	61,628.443***	861.911***	180.142***	59,845.072***	867.780***	181.822***	61,628.443***
	[1519.85]	[467.71]	[229.55]	[410.56]	[126.78]	[61.54]	[1519.85]	[467.71]	[229.55]
Observations	238,307			27,138			202,613		

Notes: The full sample consists of observations on guesses in 2008. We exclude 10 guesses of days prior to the date guesses could be made by buying tickets (1 February).The top panel shows regressions of the guess day, the absolute value of the distance to the sample mean guess day and the square of that distance on the pool dummy in three different samples: the full sample, official pools plus random individuals, and official pools plus all individuals. The bottom panel shows the same regressions using the guess time. "Pool dummy" is defined to be one if a bettor's name contains a pool identifier, such as the word "pool", "group", "family", "and", or a symbol such as "&" or "/", or is an official pool. An official pool is one of 150 pools that officially registered its members with the Nenana Ice Classic office. The "official pools plus random individuals" sample includes only official pools and a random subsample of 2612 (in the top panel) or 2607 (in the bottom panel) individuals with roughly the same amount of date guesses (13,617 in the top panel) or minute guesses (13,571 in the bottom panel) as official pools. The "official pools plus all individuals" sample includes only official pools and the subsample of all individuals. The guess day is measured in days of the year since 31 December. The guess time is measured in minutes of the day. All standard errors are corrected for heteroscedasticity. The absolute values of the robust t-statistics are in brackets.

*, **, *** indicate significance at the 10%, 5% and 1% level, respectively.

all regressions using absolute and square distances as dependent variables.[39] The difference in variances between groups and individuals is similar to the one found in the 2002 data. For example, from Column III we find that the standard deviation for individuals is 0.93 day larger than the one for pools (in 2002, this difference was 0.9 day; see Table 1). Once we use the official pool classification compared to the full sample of individuals (Column IX), this difference increases to 1.4 days. Thus, our results suggest that potential measurement errors may have an effect of 0.5 day on the difference in standard deviations.[40]

Our conclusions from this section are as follows. First, as in 2002, the 2008 data unambiguously show that the distribution of pool guesses displays lower variability than the distribution of individual guesses, both for guess dates and for guess times. Thus, our results are robust to the choice of year. Second, our 2002 classification procedure correctly classifies all official pools in 2008 as pools. Third, at least in the 2008 data, consistent with an attenuation bias, potential measurement error in the classification of pools seems to reduce the differences in variances.

8. FINAL REMARKS

The main contribution of this paper is to compare individual and group outcomes in a setting in which group formation is endogenous. We use a unique dataset on guess portfolios to compare the decisions made by individual bettors to those by betting pools. We find that the decisions made by pools are more moderate than the decisions made by individuals. On average, individuals appear to rely less on historical information than groups. They also place more weight on outlying days and times than groups do. Both these pieces of evidence are consistent with the idea that group decisions are more moderate, either due to the fact that groups have to reach a compromise when their members disagree, or because individuals with extreme opinions are less likely to be part of a group, or both.

Our results may be special to situations in which participants behave in a risk or fun seeking way, and situations in which individuals can choose to act alone or with a group. But they are an important complement to the existing (usually experimental) literature on group decision making. Motivated by some of the findings in this literature, economists have recently displayed renewed interest in group decision-making biases such as risky shifts, group polarization and groupthink (see e.g. Eliaz, Ray and Razin, 2006; Glaeser and Sunstein, 2007; Bénabou, 2008). Consistent with the modern psychology literature, our results highlight that cognitive biases such as group polarization, risky shifts, and groupthink may not always be present, or may not always dominate other characteristics of group decision making that can lead to moderation in groups.

APPENDIX: DATA CLEANING FOR 2002 DATA

We labelled name observations as possibly problematic if they were unique observations. For example, there were a series of guesses by BOB JOHNSON and a single guess by 8OB JOHNSON. This single guess was identified as problematic. These problem observations were checked manually one-by-one. Observations with wrong entries were fixed when the correct name was clear (as in 8OB JOHNSON).

39. In 2008, pools guessed on average earlier dates but later times of the day. As in 2002, these differences are economically small (again, on average pools and individuals guess the same day).

40. The measurement error in 2008 (except when using the official pools) is likely to be larger than in 2002, because we did not apply the same cleaning procedure nor cleaned as many observations by hand as we did in 2002 (due to the amount of time it would have taken).

For the final phase of the cleaning process we used a program to clean the names as follows. Each name was compared to every other name by comparing each character in the same position in each word; the number of character matches was saved. For instance, a comparison of BOB JOHNSON to 8OB JOHNSON results in ten matches, as all but the first characters match. These two words are then given a matching percentage of 91 (total matches = 10)/(length of word = 11). All words are compared and the highest matching percentage is recorded so that each word is associated with the word which it best matches. However, scanning errors sometimes change the length of the names, which creates a shifting problem for these comparisons. A common scanning error occurred when the character "M" scanned as "14". In a comparison of a name like MARY to 14ARY, the scanning error would not be picked up, because the program would make the following comparisons. First, "M" would be compared to "1", which is not a match. Then, "A" would be compared to "4", "R" to "A", and "Y" to "R", yielding zero matches. To help remedy this problem, we amended the program to also check for matches starting from the back of the words. Whichever matching percentage was higher between the forward check and the backward check was used as the matching percentage for those names. In order to increase the program's speed during this checking process, we only ran the backward check if the words were of different lengths. We also skipped checking words that differed in length by three or more letters, as it was highly unlikely that scanning errors could have accounted for such a large difference.

The next step was to determine whether we wanted to replace a name by its best match. If the matching name occurred more frequently than the original name in the list of guesses, it was less likely the product of scanning error than the original name. We therefore replaced the original name with its match only if the match percentage was greater than 85% and the matching name occurred more frequently than the original name.

One final issue was the possibility of an original name being replaced by a match name, which itself was replaced by its match name, as in the following example:

Observation	Name	Frequency in list of bets
1	Bob Johnso	2
2	Bob Johnson	10
3	8ob Johnso	1

In this case, the program will match Observation 1 with Observation 2, Observation 2 with Observation 1, and Observation 3 will match Observation 1. We will end up with the following:

Observation	Name	Frequency	Match name	Match percentage
1	Bob Johnso	2	Bob Johnson	1.0 (10 matches/10 length)
2	Bob Johnson	10	Bob Johnso	0.91 (10 matches/11 length)
3	8ob Johnso	1	Bob Johnso	0.9 (9 matches/10 length)

The program would then replace the name in Observation 1 with the name in Observation 2, as they have a match percentage above 85 and Observation 2 occurs more frequently than Observation 1. Observation 2 will not be replaced as there is no name that occurs more frequently. The problem arises when Observation 3 is replaced with Observation 1, which is itself an incorrect name. We want Observation 3 to be replaced by Observation 1 only if Observation 1 is correct. Otherwise, we want Observation 3 to be replaced by the same replacement as Observation 1. The final portion of the code implements a chaining procedure, which solves this by replacing an original name with the final match of its match name.

After running the program, we determined that removing all spaces from the names would increase the efficiency and accuracy of the program. For example, two names that are clearly the same, such as BOB JO HNSO N and BOBJOHNSON would not match because their percentages were below the threshold due to errant spaces input by the scanner. We therefore reran the program after removing spaces in the names to obtain the final output of the cleaning program.

Acknowledgements. We benefited from comments on previous versions by Heitor Almeida, Jordi Blanes-i-Vidal, Tore Ellingsen, Thomas Kittsteiner, Don Morgan, Marco Ottaviani, Maitreesh Ghatak and Bruno Biais (the editors), and especially two anonymous referees. We thank Cherrie Forness, Manager of the Nenana Ice Classic, for providing us with copies of the List of Guesses and Amanda Wubbold for her assistance with the 2008 data. For excellent research assistance, we thank Elise Aatenir and especially Kyle Lewis.

REFERENCES

ADAMS, R., ALMEIDA, H. and FERREIRA, D. (2005), "Powerful CEOs and their Impact on Corporate Performance", *Review of Financial Studies,* **18**, 1403–1432.

Arctic Science Journeys (1997), "Nenana Ice Classic", Radioscript, http://www.uaf.edu/seagrant/NewsMedia/97ASJ/04.29.97_IceClassic.html.

BÄR, M., KEMPF, A. and RUENZI, S. (2005), "Team Management and Mutual Funds" (CFR Working Paper No. 05-10, Cologne).

BARBER, B. M., HEATH, C. and ODEAN, T. (2003), "Good Reasons Sell: Reason-Based Choice Among Group and Individual Investors in the Stock market", *Management Science,* **49**, 1636–1652.

BÉNABOU, R. (2008), "Groupthink: Collective Delusions in Organizations and Markets" (Working Paper, Princeton University).

BLINDER, A. S. and MORGAN J. (2005), "Are Two Heads Better than One? Monetary Policy by Committee", *Journal of Money, Credit and Banking,* **37**, 789–812.

BONE, J., HEY, J. and SUCKLING, J. (1999), "Are Groups More (or Less) Consistent than Individuals?" *Journal of Risk and Uncertainty,* **18**, 63–81.

CASON, T. N. and MUI, V. L. (1997), "A Laboratory Study of Group Polarization in the Team Dictator Game", *Economic Journal,* **107**, 1465–1483.

DALY, J. (2001), "The Nenana Ice Classic: Betting on Warming", Web Document, http://www.john-daly.com/nenana.htm.

ELIAZ, K., RAY, D. and RAZIN, R. (2006), "A Decision-Theoretic Basis for Choice Shifts in Groups", *American Economic Review,* **96**, 1321–1332.

FARRELL, L., HARTLEY, R., LANOT, G. and WALKER, I. (2000), "The Demand for Lotto: The Role of Conscious Selection", *Journal of Business and Economic Statistics,* **18**, 228–241.

FINKEL, M. (1998), "This Break-up Is the Talk of the Town: When the River Thaws in Nenana, Alaska, a Pile of Money Changes Hands", *Sports Illustrated Magazine,* **88**, R12.

FORBES, D. and MILLIKEN, F. (1999), "Cognition and Corporate Governance: Understanding Boards of Directors as Strategic Decision-Making Groups", *Academy of Management Review,* **24**, 489–505.

FORREST, D., SIMMONS, R. and CHESTERS, N. (2002), "Buying a Dream: Alternative Models of Demand for Lotto", *Economic Inquiry,* **40**, 485–496.

GLAESER, E. L. and SUNSTEIN, C. R. (2007), "Extremism and Social Learning" *Journal of Legal Studies,* (forthcoming).

HAMILTON, B. H., NICKERSON, J. A. and OWAN, H. (2003), "Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation", *Journal of Political Economy,* **111**, 465–497.

JANIS, I. (1982), *Groupthink: Psychological Studies of Policy Decisions and Fiascoes* (Boston: Houghton Mifflin).

KERR, N. (1992), "Group Decision Making at a Multialternative Task: Extremity, Interfaction Distance, Pluralities, and Issue Importance", *Organizational Behavior and Human Decision Processes,* **52**, 64–95.

KOCHER, M., STRAUß, S. and SUTTER, M. (2006), "Individual or Team Decision-Making–Causes and Consequences of Self-Selection", *Games and Economic Behavior,* **56**, 259–270.

KOCHER, M. G. and SUTTER, M. (2005), "The Decision Maker Matters: Individual versus Group Behavior in Experimental Beauty-Contest Games", *Economic Journal,* **115**, 200–223.

KOGAN, N. and WALLACH, M. (1966), "Modification of Judgmental Style Through Group Interaction", *Journal of Personality and Social Psychology,* **4**, 165–174.

LEVINE, J. M. and MORELAND, R. L. (1998), "Small Groups", in Gilbert, D. T., Fiske, S. T. and Lindzey, G. (eds.) *The Handbook of Social Psychology*, 4th edn, vol 2 (New York: McGraw-Hill) 415–469.

MOSCOVICI, S. and ZAVALLONI, M. (1969), "The Group as a Polarizer of Attitudes" *Journal of Personality and Social Psychology,* **12**, 125–135.

Nenana Ice Classic (2002), *List of Guesses* (Alaska: Nenana).

OTTAVIANI, M. and SORENSEN, P. N. (2006), "The Strategy of Professional Forecasting", *Journal of Financial Economics,* **81**, 441–466.

O'RONAIN, M. (2002), "The Nenana Ice Classic 2002 or How to Lose Money Gambling on Global Warming", Web Document, http://www.numberwatch.co.uk/nenana_ice_classic_2002.htm.

PRATHER, L. J. and MIDDLETON, K. L. (2002), "Are N+1 Heads Better than One? The Case of Mutual Fund Managers", *Journal of Economic Behavior and Organization,* **47**, 103–120.

RICHARDS, B. (1995), "Forget the Calendar, Alaskans Bet They Can Spot Spring: This Is the Season for Wagers On When the Tanana Ice–and Winter–Will Break", *The Wall Street Journal,* 20 April, A1.

ROCKENBACH, B., SADRIEH, A. and MATHAUSCHEK, B. (2007), "Teams Take the Better Risks", *Journal of Economic Behavior and Organization,* **63**, 412–422.

SAGARIN, R. and MICHELI, F. (2001), "Climate Change in Nontraditional Data Sets", *Science,* **294**, 811.

SAH, R. K. and STIGLITZ, J. (1986), "The Architecture of Economic Systems: Hierarchies and Polyarchies", *American Economic Review,* **76**, 716–727.

SAH, R. K. and STIGLITZ, J. (1991), "The Quality of Managers in Centralized versus Decentralized Organizations", *Quarterly Journal of Economics,* **106**, 289–295.

SAUER, R. (1998), "The Economics of Wagering Markets", *Journal of Economic Literature,* **36**, 2021–2064.

STONER, J. F. A. (1968), "Risky and Cautious Shifts in Group Decisions" *Journal of Experimental Social Psychology,* **4**, 442–459.

SUNSTEIN, C. R. (2002), "The Law of Group Polarization", *Journal of Political Philosophy,* **10**, 175–195.

SUROWIECKI, J. (2005), *The Wisdom of Crowds* (New York: Anchor Books).

SUTTER, M. (2005), "Are Four Heads Better Than Two? An Experimental Beauty-Contest Game with Teams of Different Size", *Economics Letters,* **88**, 41–46.

THALER, R. and ZIEMBA, W. (1988), "Anomalies: Parimutuel Betting Markets: Racetracks and Lotteries", *Journal of Economic Perspectives,* **2**, 161–174.

THE SEATTLE TIMES (1986), "All Bets are on Ice in Quest for Alaska Cash", *The Seattle Times,* 4 May, E1.

WALLACH, M. and KOGAN, N. (1965), "The Roles of Information, Discussion, and Consensus in Group Risk Taking", *Journal of Experimental Social Psychology,* **1**, 1–19.

WILLIAMS, L. (1999), "Information Efficiency in Betting Markets: A Survey", *Bulletin of Economic Research,* **51**, 1–30.