Decision-Making under the Gambler's Fallacy:
Evidence from Asylum Judges, Loan Officers, and Baseball Umpires

Daniel Chen [1] Tobias Moskowitz [2] Kelly Shue [2]

[1]ETH Zurich

[2]University of Chicago, Booth School of Business

April 18, 2015

Fooled by Randomness

How people often imagine a sequence of coin flips:

THTHTTHTHHTTHTTHHTHTT

A real sequence of coin flips:

THTHTHHHHHTHHTTTTTHHT

Fooled by Randomness

Law of Small Numbers (Rabin 2002)

- Expect very small samples/short sequences to resemble the population
- Expect alternation even though streaks often occur by chance

Gambler's Fallacy (Tversky and Kahneman 1974)

- Seeing a 0 or 0s increases the odds of the next draw being a 1 and vice versa (e.g. a "fair" slot machine)

Decision-Making under the Gambler's Fallacy

Large literature explores these misperceptions of randomness

- Many studies focus on **predictions** in lab settings or **betting** behavior after agents observe previous outcomes
- Little field research on how misperceptions of randomness can affect agents making sequential **decisions** under uncertainty
  - We focus on field evidence with people making decisions in their primary occupation

Our hypothesis:

- Gambler's fallacy $\Longrightarrow$ Negatively autocorrelated decisions, avoidance of streaks

The Decision-Maker's Problem

Suppose an agent makes 0/1 decisions on randomly ordered cases

- If decisions are based on case merits, decision on the previous case should not predict decision on the next case (controlling for base rates)

If the decision-maker suffers from the gambler's fallacy

- After deciding 1 on the previous case, will approach the next case with a *prior belief* that it is likely to be a 0 (and vice versa)
- Also receives a noisy signal about the quality of the current case
- Decisions will be negatively autocorrelated if they depend on a mixture of prior beliefs and the noisy signal
- Similar patterns if agent is rational but judged by behavioral others

Decisions vs. **predictions/betting**: Greater confidence in the noisy signal $\Longrightarrow$ less negative autocorrelation in decisions

Three High-Stakes Real World Settings

1. Refugee court judge decisions to grant or deny asylum
   - Random assignment to judges and FIFO ordering of cases
   - High stakes decisions determining whether refugees are deported

2. Loan officer decisions to grant or deny loan applications
   - Field experiment with random ordering of loan files
     (Data from Cole, Kanz, and Klapper 2013)
   - Randomly assigned incentive schemes

3. Umpire calls of strike or ball for pitches in baseball games
   - Exact pitch location, speed, etc. to control for pitch quality
   - Know whether the decision was correct

Preview of Results

Negative autocorrelation in decisions and avoidance of streaks

- Up to 5% of decisions are reversed due to the gambler's fallacy
- Stronger bias for moderate decision-makers, similar or close-in-time cases
- Weaker bias for experienced or educated decision-makers, under strong incentives for accuracy

Less likely to be driven by potential alternative explanations

- Preference to be equally nice/fair to two opposing teams
- Sequential contrast effects
- Quotas and/or learning
- Not driven solely by concerns of external perceptions

Outline

1. Empirical Framework
2. Setting 1: Asylum Judges
3. Setting 2: Loan Officers
4. Setting 3: Baseball Umpires
5. Discussion and Alternative Explanations

Outline

Baseline Empirical Model

$$Y_{it} = \beta_0 + \beta_1 Y_{i,t-1} + Controls + \varepsilon_{it}$$

- If the ordering of cases is conditionally random, $\beta_1 < 0$ is evidence in favor of the gambler's fallacy affecting decisions

- Each decision-maker's tendency to be positive may be fixed or slowly changing over time, leading to upward bias for β_1 (bias against us)
  - Don't include individual FE, because that biases toward $\beta_1 < 0$
  - Control for average of previous n decisions, excluding current decision
  - Or, control for decision-maker's average Y in other sessions

Empirical Model Extensions

- We also test reactions to streaks using past two decisions

- Look at subsamples restricted to moderate decision makers – those with average grant rates closer to 0.5 (calculated excluding the current or recent decisions)
  - Mechanically, extreme decision makers are not negatively autocorrelated
  -

- Similar baseline results using logit or probit

Outline

Asylum Judges: Data

- Administrative universe, 1985-2013, 45 immigration courts, 357 judges
- High stakes: Denial of asylum usually results in deportation
  - "Applicant for asylum reasonably fears imprisonment, torture, or death if forced to return to her home country" (Stanford Law Review 2007)
- Cases filed within each court are randomly assigned to judges, and judges review the queue of cases following "first-in-first-out"
  - Control for time-variation in court-level case quality using recent approval rates of other judges in the same court (tends to be slow-moving positive autocorrelation)

Asylum Judges: Data

- Judges have a high degree of discretion
  - No formal or advised quotas (substantial heterogeneity in grant rates across judges in the same court)
  - Serve until retirement, fixed wage schedule w/o bonuses

- Average grant rate is 30%

- Control for recent approval rate of other judges in same court, judge grant rate over the past 5 decisions (excluding current decision), judge overall approval rate (excluding current decision), current case characteristics, and time of day FE

- Restrict sample to consecutive decisions (same day or across days)

Asylum Judges: Baseline Results

	Grant Asylum Dummy				
	(1)	(2)	(3)	(4)	(5)
Lag grant	-0.00544*	-0.0108***	-0.0155**	-0.0326***	
	(0.00308)	(0.00413)	(0.00631)	(0.00773)	
Lag grant - grant					-0.0549***
					(0.0148)
Lag deny - grant					-0.0367**
					(0.0171)
Lag grant - deny					-0.00804
					(0.0157)
Exclude extreme judges	No	Yes	Yes	Yes	Yes
Same day cases	No	No	Yes	Yes	Yes
Same defensive cases	No	No	No	Yes	Yes
N	150357	80733	36389	23990	10652
R^2	0.374	0.207	0.223	0.228	0.269

- Judges are up to 5 percentage points less likely to grant asylum if the previous case(s) were granted
- Up to 17% decline relative to the base rate of asylum grants

Asylum Judges: Heterogeneity

	Grant Asylum Dummy			
	(1)	(2)	(3)	(4)
Lag grant	-0.0196**	0.00180	-0.0484***	-0.0553***
	(0.00801)	(0.00900)	(0.0115)	(0.0115)
Same nationality	0.0336***			
	(0.0108)			
Lag grant x same nationality	-0.0421***			
	(0.0126)			
Moderate judge		0.0326***		
		(0.0116)		
Lag grant x moderate judge		-0.0700***		
		(0.0136)		
Experienced judge			0.0138	0.0253*
			(0.0106)	(0.0140)
Lag grant x experienced judge			0.0327**	0.0456***
			(0.0152)	(0.0156)
Judge FE	No	No	No	Yes
N	23990	23990	22965	22965
R^2	0.229	0.229	0.229	0.247

- Stronger bias when consecutive cases are same nationality and among moderate judges (grant rate, excl. current, is between 0.3 and 0.7)
- Weaker bias with experience

Asylum Judges: Ordering of Case Quality

	Quality Measure 1	Quality Measure 2	Lawyer Dummy	Lawyer Quality	Size of Family
	(1)	(2)	(3)	(4)	(5)
Lag grant	0.00273**	0.00307**	-0.0000772	-0.00117	-0.00927
	(0.00116)	(0.00134)	(0.00258)	(0.00293)	(0.0104)
N	23990	23980	23990	19737	23990
R^2	0.806	0.761	0.0858	0.451	0.159

- A previous grant decision does not predict that the next case will be lower in observed quality measures
- Case quality is weakly positively correlated – bias against our findings of negatively autocorrelated decisions

Outline

Loan Officers Field Experiment

- Real loan officers in India are paid to screen actual loan applications which either performed or defaulted/rejected in the past
  - Data from Cole, Kanz and Klapper (2013)
  - Paid for accuracy, but decisions do not affect actual loan origination
- In each session, the loan officer screens 6 randomly ordered loan files and decides whether to approve or reject each loan file
- Incentive schemes [app perf loan, app nonperf loan, reject loan]:
  1. Flat incentives $[20, 20, 0]$
  2. Stronger incentives $[20, 0, 10]$
  3. Strongest incentives $[50, -100, 0]$

Loan Officers: Baseline Results

	Approve Loan Dummy			
	(1)	(2)	(3)	(4)
Lag approve x flat incent	-0.0814**	-0.0712**	-0.225***	-0.228***
	(0.0322)	(0.0323)	(0.0646)	(0.0639)
Lag approve x stronger incent	-0.00674	-0.00215	-0.0525**	-0.0484**
	(0.0134)	(0.0134)	(0.0215)	(0.0214)
Lag approve x strongest incent	0.0102	0.0159	-0.0530	-0.0473
	(0.0298)	(0.0292)	(0.0468)	(0.0450)
p-value equality across incentives	0.0695	0.0963	0.0395	0.0278
Control for current loan quality	No	Yes	No	Yes
Sample	All	All	Moderates	Moderates
N	7640	7640	2615	2615
R^2	0.0257	0.0536	0.0247	0.0544

- Differences across incentive schemes are significant
- Under flat incentives, 8 pct points less likely to approve if the previous loan was approved (10% decline relative to the base rate of approval)
- Stronger effects among moderates

Loan Officers: Heterogeneity

	Approve Loan Dummy			
	(1)	(2)	(3)	(4)
Lag approve	-0.0247*	-0.127***	-0.376***	-0.0555**
	(0.0135)	(0.0329)	(0.136)	(0.0250)
Grad school	-0.0213			
	(0.0214)			
Lag approve x grad school	0.0448*			
	(0.0245)			
Log(time viewed)		-0.0968***		
		(0.0202)		
Lag approve x log(time viewed)		0.0858***		
		(0.0230)		
Log(age)			-0.0603*	
			(0.0329)	
Lag approve x log(age)			0.101***	
			(0.0375)	
Log(experience)				-0.0133
				(0.00985)
Lag approve x log(experience)				0.0226*
				(0.0116)
Sample	All	All	All	All
N	7640	7640	7640	7640
R^2	0.0256	0.0281	0.0260	0.0256

- Education, longer time spent reviewing the current loan file, age, and experience reduce negative autocorrelation

Loan Officers: Reaction to Streaks

	Approve Loan Dummy	
	(1)	(2)
Lag approve - approve	-0.0751***	-0.165***
	(0.0216)	(0.0329)
Lag approve - reject	-0.0691***	-0.0955***
	(0.0236)	(0.0347)
Lag reject - approve	-0.0322	-0.0832**
	(0.0225)	(0.0332)
Sample	All	Moderates
N	6112	2092
R^2	0.0290	0.0322

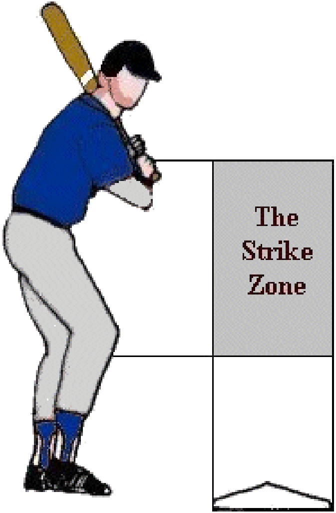
Loan Officers: Balanced Sessions

	Approve Loan Dummy		Performing Loan Dummy	
	(1)	(2)	(3)	(4)
Lag approve x flat incent	-0.0814**	-0.225***		
	(0.0322)	(0.0646)		
Lag approve x stronger incent	-0.00674	-0.0525**		
	(0.0134)	(0.0215)		
Lag approve x strongest incent	0.0102	-0.0530		
	(0.0298)	(0.0468)		
Lag perform x flat incent			-0.191***	-0.155***
			(0.0262)	(0.0529)
Lag perform x stronger incent			-0.131***	-0.142***
			(0.0123)	(0.0198)
Lag perform x strongest incent			-0.195***	-0.231***
			(0.0255)	(0.0407)
Sample	All	Moderates	All	Moderates
N	7640	2615	7640	2615
R^2	0.0257	0.0247	0.0235	0.0267

- Balanced sessions consisted of 4 performing loans and 2 bad loans
- Loan officers were NOT told of this
- If loan officers had "figured out" the balanced session design, we would expect more negative coefficients for the stronger incentive treatments

Outline

Baseball Umpires



Baseball Umpires: PITCHf/x Pitch Trajectory Data

Look at called pitches (batter does not swing): 30% are called strikes

- Detailed controls for location; speed, acceleration, curvature, spin in x, y, and z directions; **whether pitch is within strike zone**

Pitch characteristics are not randomly ordered

- Test whether, controlling for the true location and strike status, umpires make mistakes in the opposite direction of the previous call
- Controls completely determine the correct call – coefficients on lagged call reflect **mistakes**

Umpires may be biased in other ways, e.g. avoid game-determining calls

- Control for count (# balls and strikes so far), leverage (importance of current call for determining game), score, and home team

Baseball Umpires: Baseline

Strike	Full Sample		Consecutive Pitches	
	(1)	(2)	(3)	(4)
Lag strike	-0.00919***		-0.0146***	
	(0.000591)		(0.000972)	
Lag strike - strike		-0.0131***		-0.0212***
		(0.00104)		(0.00268)
Lag ball - strike		-0.00994***		-0.0189***
		(0.000718)		(0.00156)
Lag strike - ball		-0.00267***		-0.00689***
		(0.000646)		(0.00155)
Pitch location	Yes	Yes	Yes	Yes
Pitch trajectory	Yes	Yes	Yes	Yes
Game conditions	Yes	Yes	Yes	Yes
N	1536807	1331399	898741	428005
R^2	0.669	0.668	0.665	0.669

- Umpires are up to 2 percentage points less likely to call the current pitch a strike if the previous pitch(es) were called strikes
- 6.8% decline relative to the base rate of strike calls

Baseball Umpires: Endogenous Pitcher Response

	True Strike		Distance from Center			
	(1)	(2)	(3)	(4)	(5)	(6)
Lag strike	0.0168***		-0.275***		-0.00385	
	(0.00149)		(0.0236)		(0.00573)	
Lag strike - strike		0.0121***		-0.156**		-0.00403
		(0.00415)		(0.0701)		(0.0168)
Lag ball - strike		0.0200***		-0.361***		0.00651
		(0.00243)		(0.0367)		(0.00875)
Lag strike - ball		0.00308		-0.131***		0.00707
		(0.00241)		(0.0359)		(0.00854)
Pitch location	No	No	No	No	Yes	Yes
Pitch trajectory	Yes	Yes	Yes	Yes	Yes	Yes
Game conditions	Yes	Yes	Yes	Yes	Yes	Yes
N	898741	428005	898741	428005	898741	428005
R^2	0.0798	0.0924	0.171	0.188	0.952	0.952

Endogenous changes in pitch location are likely to be a bias against our findings: Following a strike, the next pitch is likely to be closer to the center of the strike zone, i.e. another strike

Baseball Umpires: Ambiguous vs. Obvious Calls

Strike	Current Pitch Ambiguous		Current Pitch Obvious	
	(1)	(2)	(3)	(4)
Lag strike	-0.0347*** (0.00378)		-0.00226*** (0.000415)	
Lag strike - strike		-0.0479*** (0.0113)		-0.00515*** (0.00101)
Lag ball - strike		-0.0324*** (0.00566)		-0.00442*** (0.000773)
Lag strike - ball		-0.000838 (0.00563)		-0.00283*** (0.000841)
Pitch location	Yes	Yes	Yes	Yes
Pitch trajectory	Yes	Yes	Yes	Yes
Game conditions	Yes	Yes	Yes	Yes
N	151501	73820	335318	153996
R^2	0.317	0.316	0.891	0.896

Negative autocorrelation is stronger if current pitch is ambiguous ($\pm 1.5''$ from edge of strike zone) rather than obvious ($3''$ around center or $6''$ from edge)

Outline

Preference for be Equally Nice/Fair to Two Teams

Asylum applicants and loan applicants are not on teams

In baseball, the umpire makes sequential calls on the same team at bat

- If the previous pitch was ambiguous (near the edge of the strike zone) and the umpire called it a strike, the umpire might tend to call the next pitch a ball to make it up to the team at bat
- Umpires might also wish to reverse previous mistakes

However, we find equally strong or stronger negative autocorrelation after previous calls that are obvious and correct

- In these cases, the umpire is less likely to feel guilt because the umpire could not have called the pitch any other way

Baseball Umpires: Fairness

Strike	Full Sample		Following Ambiguous/Obvious
	(1)	(2)	(3)
Lag strike x prev call correct	-0.0177***		
	(0.00101)		
Lag strike x prev call incorrect	-0.00663***		
	(0.00130)		
Lag strike x prev call obvious		-0.0180***	-0.0175***
		(0.00189)	(0.00216)
Lag strike x prev call ambiguous		-0.0120***	
		(0.00123)	
Lag strike x prev call not ambiguous/obvious		-0.0150***	
		(0.00103)	
Lag strike x prev call ambiguous and correct			-0.0140***
			(0.00175)
Lag strike x prev call ambiguous and incorrect			-0.00824***
			(0.00188)
Pitch location	Yes	Yes	Yes
Pitch trajectory	Yes	Yes	Yes
Game conditions	Yes	Yes	Yes
N	898741	895733	476819
R^2	0.665	0.665	0.666

Negative autocorrelation are slightly stronger after previous calls that are obvious or correct, suggesting that desire to undo marginal calls is not the sole driver of our results

Sequential Contrast Effects (SCE)

Criteria for quality while judging the current case may be higher if the previous case was particularly high quality (Bhargava and Fisman, 2012)

- After reading a really great book, my standard for judging the next book to be "good" on a 0/1 scale may be higher

$$Y_{it} = \beta_0 + \beta_1 Y_{i,t-1} + \beta_2 Quality_{i,t-1} + Controls + \varepsilon_{it}$$

If SCE causes negatively autocorrelated decisions, we expect $\beta_2 < 0$

- Controlling for discrete decision $Y_{i,t-1}$, decision-makers should be more likely to reject the current case if the previous case was of very high quality, as measured continuously using $Quality_{i,t-1}$

Loan Officers: Sequential Contrast Effects

	Approve Loan Dummy	
	(1)	(2)
Lag approve	-0.0223	-0.0736***
	(0.0148)	(0.0264)
Lag loan quality	0.00679	0.00692
	(0.00994)	(0.0201)
p-value lag loan quality rating < 0	0.247	0.365
Sample	All	Moderates
N	7495	2615
R^2	0.0252	0.0225

- Loan quality as reported on a 100 point scale by the loan officer (scaled down to 0-1)

Asylum Judges: Sequential Contrast Effects

	Grant Asylum Dummy	
	(1)	(2)
Lag grant	-0.0356***	-0.0352***
	(0.00788)	(0.00785)
Lag case quality	0.00691*	0.00520
	(0.00385)	(0.00360)
p-value lag case quality < 0	0.0367	0.0751
Quality Measure	1	2
N	23981	23973
R^2	0.228	0.228

- Case quality is predicted using a regression of asylum decisions on applicant characteristics

Quotas and/or Learning

Quotas or Learning may cause negatively autocorrelated decision-making

- Judges, loan officers (in field experiment), and umpires do not face explicit quotas or targets, but may self-impose these

We control for the fraction of the previous 2-10 decisions that were 1's

- Conditional on this fraction, the most recent decision still negatively predicts the next decision
- Unlikely to be explained by quotas/learning unless they operate unless agents can't remember beyond the most recent decision
- Agents are highly experienced, and quality bar is given in baseball

Concerns about External Perceptions

Decision-maker is rational but judged by others who suffer from the gambler's fallacy

- This is broadly consistent with our hypothesis
- Not likely to be a strong factor in the loan officers experiment where they are paid for accuracy
- Asylum judges typically serve until retirement, are paid fixed salary, and can discriminate by nationality of asylum applicant
- Negative autocorrelation in umpire calls does not vary dramatically by game attendance or leverage

Preference for Randomization

Agents may prefer to alternate being "mean" and "nice" over short time horizons

- Loan officers in the experiment are told that their decisions do not affect actual loan origination

- More generally, the gambler's fallacy may be the reason why agents feel more guilty after "1100" than "1010"

Conclusion

Gambler's fallacy $\Longrightarrow$ Negatively autocorrelated decisions, avoidance of streaks

- Stronger for moderate judges, similar or close-in-time cases
- Weaker for experienced or educated decision-makers, under strong incentives for accuracy
- Decisions vs. predictions/betting: Greater confidence in the signal of case quality reduces bias in decision-making (even if agents continue to suffer from the gambler's fallacy)

Pervasive phenomenon: Judicial courts, loan approval, referee calls

- May also apply to HR hiring, grading, admissions, medical diagnosis, auditing, investing, etc.

Model Setup (Based on Rabin 2002)

An agent makes 0/1 decisions for a randomly ordered series of cases

- True case quality is an i.i.d. sequence $\{y_t\}_{t=1}^{M}$ where $y_t = \{0,1\}$, $P(y_t = 1) = \alpha$
- Agent's prior about the current case: $P_t \equiv P\left(y_t = 1 \mid \{y_\tau\}_{\tau=1}^{t-1}\right)$
- Agent also observes a signal about current case quality $S_t \in \{0,1\}$ which is accurate with prob μ and uninformative with prob $1-\mu$
- By Bayes Rule, the agent's belief after observing $S_t = 1$ is

$$P\left(y_t = 1 \mid S_t, \{y_\tau\}_{\tau=1}^{t-1}\right) = \frac{[\mu S_t + (1-\mu)\alpha]\, P_t}{\alpha}$$

$$\textit{Decision}\ D_t = 1\left\{\frac{[\mu S_t + (1-\mu)\alpha]\, P_t}{\alpha} \geq X\right\}$$

The Rational Thinker

We compare the prior beliefs and decisions of a rational agent to those of a coarse thinker

- The rational agent understands that the y_t are i.i.d.
- Priors are independent of history

$$P_t^R = P\left(y_t = 1 \mid \{y_\tau\}_{\tau=1}^{t-1}\right) = P(y_t = 1) = \alpha$$

- By Bayes Rule, the agent's belief after observing S_t is

$$P\left(y_t = 1 \mid S_t = 1, \{y_\tau\}_{\tau=1}^{t-1}\right) = \mu S_t + (1-\mu)\alpha$$

The Coarse-Thinker

Degree of coarse-thinking is indexed by $N \in \mathbb{N},\ N \geq 6$
(lower N corresponds to more severe coarse thinking)

The coarse thinker believes that:

- For rounds 1, 4, 7, ... cases are drawn from an urn containing N cases, αN of which are 1's (and the remainder are 0's)
- For rounds 2, 5, 8, ... cases are drawn from an urn containing $N-1$ cases, $\alpha N - y_{t-1}$ of which are 1's
- For rounds 3, 6, 9, ... cases are drawn from an urn containing $N-2$ cases, $\alpha N - y_{t-1} - y_{t-2}$ of which are 1's

As $N \to \infty$, the coarse-thinker behaves likes the rational thinker

Fraction of Decisions Altered by Gambler's Fallacy

Simple regression

$$Y_{it} = \beta_0 + \beta_1 Y_{i,t-1} + \varepsilon_{it}$$

Base rate of affirmatives

$$\alpha \equiv P(Y = 1) = \frac{\beta_0}{1 - \beta_1}$$

Fraction of decisions altered

$$(\beta_0 - \alpha) \cdot P(Y_{i,t-1} = 0) + (\alpha - (\beta_0 + \beta_1)) \cdot P(Y_{i,t-1} = 1)$$

$$= 2\beta_1 \alpha (1 - \alpha)$$

Asylum Judges: First-in-First-Out

FIFO can be violated if asylum applicant claims work hardship, files additional applications, etc.

- Assume these violations of FIFO, which are driven by applicant behaviors, are not negatively correlated with the previous decision
- Asylum judges scheduling system usually picks the next available date

- We estimate the "quality" of each case by regressing grant decisions on case characteristics and using the predicted grant outcomes
    - Predicted case quality is positively autocorrelated
- Previous grant or deny decisions do not significantly predict whether the next case has a written decision, remote hearing, or non-decision

Reaction to Streaks

Is negative autocorrelation stronger following streaks of 1's or 0's?

$$Y_{it} = \beta_0 + \beta_1 I(1,1) + \beta_2 I(0,1) + \beta_3 I(1,0) + Controls + \varepsilon_{it}$$

- $I(Y_{i,t-2}, Y_{i,t-1})$ is an indicator representing the two previous decisions
- All β's measure behavior relative to the omitted group $I(0,0)$
- If negative autocorrelation increases with streaks we expect $\beta_1 < \beta_2 < 0$ and $\beta_1 < \beta_3 < 0$
- Under certain modeling assumptions, we also expect $\beta_2 < \beta_3$

Asylum Judges: Summary Statistics

	Mean	Median	S.D.
Number of judges	357		
Number of courts	45		
Years since appointment	8.41	8	6.06
Daily caseload of judge	1.89	2	0.84
Family size	1.21	1	0.64
Grant indicator	0.29		
Non-extreme indicator	0.54		
Moderate indicator	0.25		
Lawyer indicator	0.939		
Defensive indicator	0.437		
Morning indicator	0.47		
Lunchtime indicator	0.38		
Afternoon indicator	0.15		

Loan Officers: Summary Statistics

	Full Sample		Flat Incentives		Strong Incentives		Strongest Incentives	
Loan officer x loan observations	9168		1332		6336		1470	
Loan officers	188		76		181		89	
Sessions (6 loans per session)	1528		222		1056		245	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Fraction loans approved	0.73		0.81		0.72		0.68	
Fraction moderate	0.34		0.25		0.36		0.36	
Loan rating (0-1)	0.71	0.16	0.74	0.16	0.70	0.16	0.73	0.15
Fraction grad school education	0.29		0.30		0.29		0.26	
Time viewed (minutes)	3.48	2.77	2.84	2.11	3.70	2.96	3.09	2.23
Age (years)	37.70	11.95	37.37	11.93	38.60	12.17	34.13	10.21
Experience in banking (years)	9.54	9.54	9.67	9.41	9.85	9.76	8.09	8.50

Controls include:

- Mean loan officer approval rate within each incentive treatment (calculated excluding the current session)
- Incentive scheme type: Flat, stronger, or strongest

Baseball Umpires: Summary Statistics

Number of called pitches following a previous called pitch	1536807
Number of called pitches following a consecutive previous called pitch	898741
Number of games	12564
Number of umpires	127
Fraction of pitches called as strike	0.3079
Fraction of pitches called correctly	0.8664
Fraction of pitches categorized as ambiguous	0.1686
Fraction of pitches categorized as obvious	0.3731
Fraction of ambiguous pitches called correctly	0.6006
Fraction of obvious pitches called correctly	0.9924

Baseball Umpires: Heterogeneity

	(1)	(2)	(3)
Lag strike	-0.0146***	-0.0146***	-0.0143***
	(0.000972)	(0.000972)	(0.00108)
Leverage	0.000330		
	(0.000390)		
Lag strike x leverage	-0.00140**		
	(0.000625)		
Umpire accuracy		-0.00406***	
		(0.000451)	
Lag strike x umpire accuracy		0.00353***	
		(0.000621)	
High attendance			0.00441***
			(0.00115)
Low attendance			-0.00330***
			(0.00117)
Lag strike x high attendance			-0.00270*
			(0.00157)
Lag strike x low attendance			0.00123
			(0.00164)
Pitch location	Yes	Yes	Yes
Pitch trajectory	Yes	Yes	Yes
Game conditions	Yes	Yes	Yes
N	898741	898154	894779
R^2	0.665	0.665	0.665