The Structure of Intrinsic Prosociality: Field Evidence from Online Chess

Ravideep Sethi*

October 30, 2026

Abstract

I study costly, anonymous, unsolicited, unrewarded concessions to a rival. On a chess platform, a player whose opponent disconnects can claim the win or grant a draw at a rating price set before play; 1.4 percent of 41.5 million decisions are grants. Granting rises with the opponent's position and falls with game length, the two dictator-game entitlement manipulations, generated by play. Merited concessions, to an opponent still in the game, and unmerited ones, to an opponent clearly lost, are distinct traits: each predicts mainly itself, and they relate to paying for the platform with opposite signs, which one-dimensional models cannot produce. Receiving a draw, instrumented by the assigned opponent's leniency, raises later granting toward third players by a fifth of the base rate; a lenient opponent met in a game without a disconnection, or a regular draw, does not.

Keywords: generosity, prosocial behavior, competition, desert, fairness, social preferences, reciprocity
JEL: D63, D64, D91, L86

*Division of Quantitative Analysis of Markets and Organizations, David Eccles School of Business, University of Utah, Salt Lake City, UT 84112, USA. Email: ravi@eccles.utah.edu. Code and replication materials: https://github.com/ravideepsethi/intrinsic-prosociality. I thank Alexander Coutts, Decio Coviello, Anil Kshatriya, Bradley Ruffle, Erik Snowberg, and participants at the Advances with Field Experiments (AFE) conference for helpful comments and suggestions. All errors are my own.

1 Introduction

On Lichess, a popular online chess platform, a player whose opponent disconnects from a rated game waits a few seconds and then sees two buttons: *Claim victory* and *Call draw*.[1] The first takes the win. The second ends the game as a draw and gives the absent opponent a better result at the chooser's expense, forgoing a win premium that averages 11.8 rating points, the currency the two were competing for a moment earlier.[2] Every ordinary support for giving is muted, removed, or reversed: the opponent is a stranger assigned by a matchmaking algorithm, is not there to ask, and will almost never be met again; the platform neither rewards the draw nor announces it; and pressing neither button lets the absent player's clock run out, which ordinarily produces a win.[3]

Players chose the draw 570,000 times in 41.5 million such decisions, 1.37 percent, and 276,000 of the 3.2 million accounts that faced the choice, 8.6 percent, did so at least once.[4] Together those accounts gave up 6.9 million rating points, the modeled difference between the win each could have claimed and the draw it granted (Section 2.5). The behavior is uncommon, and it is structured. The grant rate is 2.76 percent when the absent opponent's position was still competitive and 0.47 percent when it was clearly lost, and within competitive positions it rises with the opponent's advantage and falls with the length of the game. Players give up a win most readily where the opponent's claim on the result was strongest.

Much of what is known about prosocial behavior concerns its external supports. People avoid situations [Dana et al., 2006] and information [Dana et al., 2007] that would compel giving. Observability increases it [Ariely et al., 2009], as concern for social image predicts [Andreoni and Bernheim, 2009]. In-person giving partly reflects the difficulty of avoiding the ask [DellaVigna et al., 2012, Andreoni et al., 2017]. People also give more when they have received something to repay [Falk, 2007, Chandar et al., 2019], when a norm says to give [Krupka and Weber, 2013], and when giving is what happens unless they act to stop it [Haggag and Paci, 2014]. Each of these designs removes one support

[1] Above the two buttons, the chooser sees the sentence "*Your opponent left the game. You can claim victory, call the game a draw, or wait.*"

[2] The median win premium is 5.7 rating points, about 0.29 of the chooser's own rating standard deviation over the preceding 100 games. Ratings are valued by players [Anderson and Green, 2018, Avoyan et al., 2025], making this a modest but measurable cost.

[3] The exception is a chooser without enough material to deliver mate, a bare king or a king with one minor piece; a flag fall against such a chooser is recorded as a draw rather than a win for the chooser, and Section 2.1 uses the same rule to classify draws.

[4] The decisions are drawn from the 2.2 billion standard rated games Lichess hosted from November 2023 through October 2025: those from the automated matchmaking pool that end with the chooser actively choosing between calling a draw and claiming the win, excluding immediate rematches and games in which the absent player's clock showed under five seconds (Section 2.3).

and measures what changes. The disconnect screen mutes, removes, or reverses them together. I call the costly prosocial behavior that remains *intrinsic prosociality*. The term names a condition, not a motive: fairness, altruism, sportsmanship, and self-image all remain candidates.[5] The act is prosocial in what it does: it transfers value, at a cost to the giver, to a rival who may have been struck by bad luck.

Each support is muted, removed, or reversed in a different way. The audience and the ask are muted: the opponent is gone at the moment of choice and cannot request anything, and the platform does not announce the result; usernames are visible and the returning opponent sees the result, so social image is weakened rather than gone. But granting is not concentrated among the accounts most likely to be watched: those rated 2400 or above, titled, or known to stream supply 0.6 percent of opportunities and 0.4 percent of grants, and removing every opportunity that involves an account with an audience proxy leaves the pattern nearly unchanged (Appendix B.13).

The prospect of meeting again is removed, since the opponent is assigned by the matchmaker and direct repayment is almost never possible. Obligation is removed: the opponent has provided no service and is owed nothing, and a moment earlier the two were trying to beat each other. The reward is reversed, because granting is not compensated and costs rating points in the currency players compete for. And the norm, the rule, and the default all favor claiming: competitive framing makes self-interest more acceptable to chess players [Bühren et al., 2012], and the interface attaches no suggestion to either button that would say otherwise; online chess regulations place responsibility for a connection on the player who lost it; and waiting ordinarily produces a win.[6]

Both gradients, with the opponent's advantage and with the length of the game, hold within the same chooser. Within competitive positions, after accounting for the chooser and the stage of the game, granting rises from 1.9 percent in an equal position to 3.7 percent when the opponent held a substantial winning advantage.[7] Granting also falls as the game lengthens: within competitive positions, the rate is 4.6 percent when the opponent leaves after one move each, 2.5 percent after ten plies (five moves by each side), and keeps falling through the eightieth ply; the same player, at the same position, grants 2.3 percentage points less often after forty plies than in the first four when the opponent was still competitive, and 1.6 less often when the opponent had lost. The opponent's rat-

[5] The term follows the intrinsic–extrinsic distinction of Bénabou and Tirole [2003]. The usage differs from that of Bénabou and Tirole [2006], who treat self-image as part of the concern for reputation rather than as an intrinsic motive, and from that of Dwenger et al. [2016], for whom intrinsic motivation is compliance without enforcement while the ask, the recipient, and the norm remain in place.

[6] FIDE Online Chess Regulations, Arts. 11.1 and 11.4.2, `https://handbook.fide.com/chapter/OnlineChessRegulations`.

[7] The latter position has an engine evaluation of +500 centipawns from the disconnected player's perspective. Section 2.4 defines the evaluation and the competitive region.

ing matters where the board does: toward an opponent 100 points stronger rather than 100 points weaker, the same player grants 1.0 point more often in competitive positions and 0.04 more in lost ones.

The two gradients correspond to the two classic manipulations of entitlement in dictator games. Giving rises when the recipient has earned a claim [Ruffle, 1998, Oxoby and Spraggon, 2008] and falls when the giver has exerted effort for the stakes [Hoffman and Spitzer, 1985, Hoffman et al., 1994, Cherry et al., 2002, Carlsson et al., 2013]; here both arise from the same board, earned in the course of the game itself rather than in a task set for the purpose.

The disposition persists over time and across decision types. An account that granted at least once in the first year of the panel is almost seventeen times as likely to grant in the second as one that did not.[8] It also extends to a decision that is monetary and public rather than rating-denominated and unannounced: accounts that grant in competitive positions are about twice as likely as matched never-granting accounts to hold the platform's paid supporter badge.[9] The link runs through the circumstances of the grant. A merited concession, a draw granted to an opponent who could still have come back, is associated with a higher badge rate; an unmerited concession, a draw granted to an opponent whose position was clearly lost, with a lower one. One disposition to give cannot produce that sign reversal; two traits, the value placed on a deserved result and the weight on rating points, do (Proposition 2).

The disposition is also passed on. Receiving a draw raises the recipient's own later granting toward a third player by about a fifth of the base rate, an effect identified from the leniency the matchmaker-assigned opponent showed toward other players. Two placebo tests say what is passed on. Meeting a lenient opponent in a game that ends without a disconnection does not change the recipient's later granting, so the channel is the concession rather than the opponent's company. A regular draw, without a disconnect screen, does not measurably change it either, which points to the act rather than the result.

The paper makes four contributions. First, it measures generosity in a hard case. Field designs keep at least one support for giving in place [Karlan, 2005, DellaVigna et al., 2012, Cohn et al., 2019, Chandar et al., 2019], competition erodes moral behavior on average [Huber et al., 2023], and concessions between rivals have been documented mainly where they are repaid [Duggan and Levitt, 2002]. Here the recipient is a rival, what is given is a

[8] Among accounts observed with a competitive opportunity in both years, those with a first-year grant grant at 15.8 percent of their second-year opportunities and those without at 0.9 percent, a ratio that partly reflects how many chances each account had in the first year and that describes the cohort returning in both periods.

[9] 0.64 against 0.33 percent (Section 5.2).

share of the contested result, and repayment is almost never possible; even if every grant in a clearly lost position were noise, competitive positions produce some 350,000 grants beyond that rate.

Second, it is a field dictator game in which entitlement is earned in play rather than in a task set for the purpose. The engine evaluation measures the recipient's claim and the ply count the giver's investment, the two classic manipulations [Hoffman and Spitzer, 1985, Ruffle, 1998, Cherry et al., 2002, Oxoby and Spraggon, 2008], while the rating rules fix the price before play, so the claim does not move with the stakes as it does when an earned claim enlarges the pie.

Third, it shows that the same button, pressed at the same price, records two distinct kinds of concession. Merited and unmerited concessions each predict themselves far better than each other and relate to paying for the platform with opposite signs. Two-dimensional models of type predict this [Bénabou and Tirole, 2006, Cappelen et al., 2007], and the board supplies the restriction that separates the dimensions: in a lost position, the value of a deserved result cannot act. Because these patterns, documented elsewhere in separate designs, appear here in one decision, population, and currency, they restrict models jointly: models that value only the rating outcomes, whether through altruism, warm glow, or inequity aversion, cannot produce the board patterns, and a single trait that raises both kinds of granting and support for the platform cannot produce the badge ordering (Section 7; Appendix E.5).

Fourth, it estimates, to my knowledge for the first time between rivals, the effect of receiving a concession on the recipient's later concessions to third players, from a leniency design in which the matchmaker, not the need for a decision, assigns the judge. Most assigned pairs therefore never face a decision, and in those pairs a lenient opponent has no effect, a test of exclusion that court designs cannot run.

Two measurement tools make these comparisons possible. A chess engine measures the claim of the player who did not decide, where economics has used engines to grade a player's own moves [Anderson et al., 2016, Strittmatter et al., 2020, Künn et al., 2023, Salant and Spenkuch, 2026]; and a replay of 7.76 billion rating records prices the button not pressed (Section 2.5).

Two limits apply. The design does not identify the motive behind a concession; fairness, altruism, sportsmanship, and self-image all remain candidates. And although the concession and the supporter badge differ in observability and currency, both are decisions on the same platform, so their association is stable heterogeneity within Lichess and does not establish that the disposition extends beyond it.

1.1 Related Literature

The paper builds on the literature on social preferences [Fehr and Charness, 2025], in particular on models in which prosocial acts signal type to others or to oneself [Bénabou and Tirole, 2006, Bursztyn and Jensen, 2017] and on evidence that people sort out of situations that would compel giving [Lazear et al., 2012]. Field designs that measure giving outside the laboratory keep at least one support in place: lab-in-the-field games keep experimenter observation [Karlan, 2005]; natural field experiments on charitable giving keep solicitation, a named recipient, and a charitable norm [Benz and Meier, 2008, DellaVigna et al., 2012, Andreoni et al., 2017]; wallet-drop experiments limit observability but keep a named owner and a civic norm [Cohn et al., 2019]; and tipping on a ride-sharing platform removes the audience at the moment of choice but keeps the ask, the norm, and a recipient who has just provided a service [Chandar et al., 2019]. Misdirected-letter designs come closest, removing the ask, the audience, and the reward, but they keep a named owner and a norm against keeping another's money [Franzen and Pointner, 2013, Stoop, 2014]. In the laboratory, giving falls when taking is an option [List, 2007, Bardsley, 2008]; here taking, the claim, is the default. What is given up is the contested outcome itself rather than a payment added to a completed exchange, and the design has no variation in the supports, so it measures what remains when they are weakened together rather than the contribution of any one of them.

The position gradient connects the field evidence to work on fairness, merit, and deservingness, where desert is elicited through allocation choices, by stakeholders or impartial spectators, after the experimenter induces differences in luck, effort, or choice [Konow, 2000, Cappelen et al., 2007, 2013, Almås et al., 2020], or through information about a real recipient's work attitudes [Fong, 2007]. Here the disconnection is the luck shock and the board the record of performance, measured from the realized position in a consequential decision whose price, unlike the pie in earned-entitlement experiments [Ruffle, 1998, Oxoby and Spraggon, 2008], does not vary with it. The paper also complements research on whether market interaction erodes moral values, as laboratory market games suggest [Falk and Szech, 2013, Bartling et al., 2023], or sustains them, as the correlation between market integration and fair offers across small-scale societies suggests [Henrich et al., 2001, 2010]; on how competitive incentives shape sabotage, helping, and cooperation among rivals [Carpenter et al., 2010, Harbring and Irlenbusch, 2011, Drago and Garvey, 1998, Bandiera et al., 2005, Buser and Dreber, 2016]; and observational studies of prosocial acts and collusive accommodation in sports [Kavussanu et al., 2006, Duggan and Levitt, 2002, Moul and Nye, 2009]; competition erodes moral behavior on average [Huber et al., 2023], winners feel entitled to more [Schurr and Ritov, 2016], and

competitive framing makes self-interest more acceptable to chess players [Bühren et al., 2012], which is why a concession here is the exception.

Competition in the field has supplied behavioral economics with some of its cleanest evidence, such as loss aversion on the putting green [Pope and Schweitzer, 2011]. Chess has served economics as a setting for observing decisions at scale [Salant and Spenkuch, 2026, Backus et al., 2023, Anderson and Green, 2018, Avoyan et al., 2025], and chess players as expert subjects in experiments on backward induction [Palacios-Huerta and Volij, 2009, Levitt et al., 2011]. The large-sample papers study moves or whether players keep playing, and engine evaluations have graded a player's own moves [Anderson et al., 2016, Strittmatter et al., 2020, Künn et al., 2023] or the decision to resign [Purohit and Srivastava, 2026]; here the engine measures the claim of the player who did not decide, and chess supplies only the measured state of the world behind a binary choice: what the opponent had earned, how far the contest had gone, and what the concession cost.

The persistence and supporter results speak to heterogeneity in prosocial behavior that is stable over time [Bruhin et al., 2019] and predicts behavior in other domains [Peysakhovich et al., 2014]. Year-to-year persistence alone could also arise from habit formation [Bruhin et al., 2026]; the supporter result, which links two distinct decisions, points toward stable types, and it bears on the question Levitt and List [2007] raise about whether prosociality measured in one setting carries over to another. Identity and self-signaling models predict that costly prosocial acts reveal type and shape later behavior [Bénabou and Tirole, 2006, 2011, Bodner and Prelec, 2003, Grossman and van der Weele, 2017], and Gneezy et al. [2012] show that costly, but not costless, initial prosocial acts predict subsequent prosocial behavior in a different domain.

The two concessions studied here have the same cost. In a model with one prosocial parameter the unmerited one is at least as diagnostic of type as the merited one, since it gives the same amount to someone with no claim, and under warm glow [Andreoni, 1990] the two are the same act. Section 5 shows they are not, as two-dimensional models predict, with one weight on fairness and another on one's own payoff [Bénabou and Tirole, 2006, Cappelen et al., 2007]; the board separates the two, since in a lost position the first cannot act. As among the self-described altruists of Fong [2007], giving that tracks the recipient's desert marks the more prosocial account.

The receipt result contributes field evidence on upstream reciprocity, in which help received is passed on to a third party rather than returned to the giver [Nowak and Sigmund, 2005, Nowak and Roch, 2007]. Laboratory evidence shows that generosity received is passed on and travels along chains of interaction [Dufwenberg et al., 2001, Stanca, 2009, Fowler and Christakis, 2010]. Field evidence in which receipt is assigned points the same way: a courtesy is passed on in traffic [Mujcic and Leibbrandt, 2018], a

randomly assigned upvote raises answers to others' questions [Wachs et al., 2026], gift subscriptions the platform allocates among viewers raise gifts to others [Chaudhry et al., 2025, Kim et al., 2026], and randomized shares of a gift raise giving to the same group [Yuan et al., 2024]. Whether the recipient responds to the act or to the outcome is the question of the intentions literature, which tests it by comparing outcomes someone chose with the same outcomes arising by chance [Falk and Fischbacher, 2006, Falk et al., 2008, Charness and Levine, 2007]. Here the benefactor and the later beneficiary are both rivals, the act is a deliberate concession to a particular rival, costly in the competition's own currency, rather than feedback or a gift the platform allocates, the design identifies the effect from the opponent the matchmaker assigned, in the manner of examiner designs [Dobbie et al., 2018, Goldsmith-Pinkham et al., 2026], and a comparison with regular draws, reached without a disconnect screen, is a field version of the intentions test.

Section 2 describes the setting and measurements and Section 3 the framework. Section 4 studies when players grant, Section 5 persistence and platform support, and Section 6 behavior after receipt. Section 7 states what the evidence settles and what it leaves open, and Section 8 concludes.

2 Setting, Data, and Measurement

2.1 The Disconnect Decision

Lichess is a free online chess platform. It allows users to play rated or unrated games in a variety of time formats as well as non-standard variations of chess. I restrict my analysis to standard rated chess games in which players join using the pool matchmaker.

Given that players can choose to play unrated games, they enter rated games in order to compete, and their ratings change with the result. A win increases the player's rating, while a loss reduces it. A draw against a higher-rated (resp. lower-rated) opponent increases (resp. decreases) a player's rating.

When one player disconnects, the game clock continues to run. An invisible countdown runs in the background. If it expires before the disconnected player returns, the connected player is offered two buttons: *Claim victory* and *Call draw* (Figure 1) presented alongside the instruction "*Your opponent left the game. You can claim victory, call the game a draw, or wait.*" The first claims a win; the second ends the game as a draw. Pressing neither button leaves the game running, so the absent player can still reconnect.[10]

[10]Before a change in July 2025 the base wait was 40 seconds in bullet and blitz, 80 in rapid, and 120 in classical games; since then it ranges from 30 to 300 seconds by speed. It is divided by three when the disconnected player is down at least four points of material, and other conditions can shorten it (Lichess pull request 17866, https://github.com/lichess-org/lila/pull/17866). The records do not observe the

Figure 1: **The disconnect decision.** The screen shown to the connected player after an opponent's disconnection countdown has elapsed. Two actions are available: *Claim victory* and *Call draw*. Leaving both unpressed lets the disconnected player's clock continue to run. Account names are replaced with role labels.



An active resolution carries the status `timeout`; an ordinary time loss carries `outoftime`, and only the first reveals the choice between a win and a draw. The connected player is the *chooser*, and the absent opponent is the potential recipient. A recorded chooser win identifies a claim. A draw is classified as a grant when the chooser had enough material to deliver mate. This rule, which excludes draws in which the chooser has only a bare king or a king with one minor piece, approximates the platform's own adjudication of a flag fall; draws that fail it, 20,902 or 0.05 percent of opportunities, remain in the sample as non-grants, and dropping them moves no reported rate (Appendix A).

2.2 Incentives and Observability

Granting gives up an advantage in the competition's own currency. Ratings are displayed on profiles and beside names during play, govern whom the matchmaker pairs, and matter to players, who respond to rating milestones [Anderson and Green, 2018] and keep playing or stop according to the previous result [Avoyan et al., 2025]. The benefit of a granted draw goes to the opponent the chooser was trying to beat, and only the chooser can provide it. The decision is a dictator game: the chooser alone divides a fixed stake

countdown itself.

between self and a passive recipient, with no response possible. What is divided is not an endowment but the result of the contest the two were just playing, and the division is binary, a win or a draw, with the transfer's size fixed by the rating rules before play.

The institutional incentives favor claiming. If the chooser waits and the opponent never returns, the clock ordinarily produces a win. In sanctioned online chess, maintaining a connection is the player's responsibility, and failure to reconnect before the clock expires ordinarily results in a loss.[11] Lichess offers no additional rating reward for making a concession. The two buttons also differ in what they do to the absent player. A claim records a departure against the absent player's account, which counts toward the platform's temporary play bans and, when that player was down four or more points of material, lowers a conduct counter that the matchmaker uses to pair players with similar records; a granted draw records nothing.[12] A concession therefore also spares the opponent a mark. Requests can themselves induce giving [DellaVigna et al., 2012, Andreoni et al., 2017]; here the interface supplies the choice without requiring one.

Social-image incentives [Andreoni and Bernheim, 2009, Ariely et al., 2009] are limited here but not eliminated: the opponent can see the result on returning, games can have spectators, and players can exchange chat messages, none of which the data record. Proxies for an audience do not organize the behavior. Opportunities involving a player rated 2400 or above, holding a recognized title, or listed in the platform's streamer directory are 0.57 percent of opportunities and 0.43 percent of grants. Removing every opportunity in which either player carries any audience proxy, including a profile channel link or recorded time in games featured for spectators, drops 18 percent of opportunities and 20 percent of grants. It leaves the within-chooser competitive-minus-clearly-worse gap at 2.09 percentage points against 2.17, and the removed accounts show the same gap. The groups do not share a sign in levels: choosers rated 2400 or above or titled grant less often overall than the rest, titled choosers alone grant more often, and known streamers are too few to compare (Appendix B.13). *Intrinsic prosociality*, as used here, describes costly giving where the direct competitive incentive favors keeping the advantage; it does not distinguish fairness, altruism, sportsmanship, or self-image as the motive [Bénabou and Tirole, 2006].

The measured choice is conditional on an active resolution. Waiting could itself be generous if a player deliberately leaves the opponent time to return, and the interface also

[11] FIDE Online Chess Regulations, Arts. 11.1 and 11.4.2, `https://handbook.fide.com/chapter/OnlineChessRegulations`. Competition-specific regulations and material rules can qualify the outcome.

[12] Lichess source code, `modules/round/src/main/RoundAsyncActor.scala` and `Finisher.scala`, `modules/playban/src/main/PlaybanApi.scala` and `RageSit.scala`, and `modules/pool/src/main/MatchMaking.scala`, at revisions from before and after the panel. Sections 6.2 and 7 take up what this means for a recipient.

lets the connected player add time to the absent opponent's clock ("Give 15 seconds" in Figure A.1), an accommodation the game record preserves only if the opponent returns and moves. A clock expiration does not reveal whether the claim buttons appeared or why the player waited. The data therefore measure granted draws among observed active decisions, rather than every way in which a connected player might accommodate a disconnected rival.

2.3 The Analysis Sample

The analysis covers live rated standard games from November 2023 through October 2025. An opportunity must end in an active resolution and show at least five seconds on the disconnected player's last recorded clock. The five-second threshold is my restriction, not the platform's; the data do not establish when the disconnection occurred or how much time remained when the chooser acted. It excludes almost no decisions: in a random sample of 150,000 pool games per month that ended with a clock-related status and a reading under five seconds, 3.6 million games in all, 65 were active resolutions, which implies about 7,900 excluded opportunities over the panel, 0.02 percent of the analysis sample, with an exact 95 percent upper bound of 0.09 percent (Appendix A). What it does is remove about three quarters of the games with a clock-related status before their metadata are retrieved.

I retain games created by the automated matchmaking pool and remove observed immediate rematches. The rematch restriction excludes a game when both players' immediately preceding observed rated games were against each other; cases with an unavailable predecessor or ambiguous ordering are also excluded. This focuses the sample on rivals assigned by the matchmaker. Earlier encounters can still occur: 7.07 percent of the retained opportunities involve accounts with a previous recorded rated meeting. Appendix A gives the complete sample flow.[13]

The resulting sample contains 41,534,289 opportunities faced by 3,224,171 chooser accounts. Of these opportunities, 570,090 end in a granted draw, a rate of 1.373 percent. Across accounts, 276,481 grant at least once, or 8.58 percent (Table 1). Account identifiers link decisions over time; they do not establish that each account corresponds to a different person.

[13]Over the 24 months from November 2023 through October 2025, Lichess recorded 2.23 billion standard rated games, of which 47.59 million were active resolutions of an opponent's disconnection (`timeout`) with the absent player's clock at five seconds or more. Keeping pool games and removing immediate rematches, games with missing or ambiguous preceding-game information, and the 42 games of my own two accounts leaves 41,534,289 (Table A.1).

Table 1: **Analysis sample and core outcomes**

Panel A. Opportunities and granted draws			
	Opportunities	Granted draws	Rate (%)
All positions	41,534,289	570,090	1.373
Competitive positions	15,207,414	419,462	2.758
Clearly-worse positions	22,664,787	106,556	0.470
Competitive minus clearly worse (pp)			2.288
Competitive / clearly worse			5.867
Panel B. Accounts and the forgone win premium			
Distinct chooser accounts			3,224,171
Accounts with at least one granted draw, any position			276,481
Accounts with at least one competitive-position granted draw			212,859
Mean forgone win premium (rating points)			11.846
Median forgone win premium (rating points)			5.709

Notes: Competitive positions have an engine evaluation of at least -100 centipawns from the disconnected player's side and clearly-worse positions at most -300; the region between enters neither the difference nor the ratio. The forgone win premium rows summarize $\kappa = K/2$ (Section 2.5) over the 419,462 competitive-position granted draws. Appendix A gives construction and composition.

2.4 Board Position and Game Stage

I evaluate the final recorded position with Stockfish 18 at a search budget of 100,000 nodes. The evaluation is measured from the disconnected player's side: positive values favor the recipient and negative values favor the chooser. A centipawn is one hundredth of the engine's pawn unit. The evaluation summarizes the position's strength; it is not a direct measure of the chooser's beliefs or a win probability.

For descriptive comparisons, I divide positions into five bins: clearly better for the disconnected player, at least $+300$ centipawns; better, $+101$ to $+299$; roughly equal, -100 to $+100$; modestly worse, -299 to -101; and clearly worse, at most -300. I call the first three bins the *competitive region*. It includes positions in which the disconnected player remained approximately equal or was ahead. A draw granted in the competitive region is a *merited concession*: the opponent could still have come back. A draw granted in a clearly worse position is an *unmerited concession*. The modestly worse bin lies between the two and enters neither comparison. The terms describe the opponent's claim on the result, not the chooser's motive.

Game stage is measured by the number of recorded plies, where a ply is one player's move. Ply two therefore means one move by each player.

2.5 Rating Consequences

Lichess updates ratings with Glicko-2 rather than the more familiar Elo system. The two share the expected-score formula, but Glicko-2 scales each update by a rating deviation and a volatility that the game records do not report. These states are coupled across players and time: each player's deviation depends on previous opponents' deviations, which depend in turn on their own histories, so recovering them for any one game requires replaying every rated game before it. I reconstruct them by replaying the public archive of rated games from January 2013 onward, about 7.76 billion records, in chronological order within each rating pool, through a Glicko-2 implementation based on the platform's published code, anchoring each player's rating to its recorded pre-game value at every game. Exact recovery is not possible, since histories before the archive and games missing from it are unobserved, but the reconstruction closely reproduces realized rating changes: the correlation between predicted and observed changes is 0.993 for granted draws and 0.996 for claimed wins, with mean absolute errors of 0.34 and 0.46 rating points. I use the recovered states to calculate the counterfactual rating consequences of granting a draw rather than claiming a win. Appendix C describes the replay, its initialization, and its validation.

Let R_i and R_j denote the pre-game ratings of the chooser and the disconnected opponent, respectively. Let $K > 0$ be the rating-update scale, determined primarily by the chooser's rating uncertainty, and let g_j be the Glicko-2 adjustment for the opponent's rating uncertainty.[14] The reconstructed rating changes from granting a draw and claiming a win are

$$r_G = K\left(\frac{1}{2} - \frac{1}{1+10^{g_j(R_j-R_i)/400}}\right),$$
$$r_C = K\frac{10^{g_j(R_j-R_i)/400}}{1+10^{g_j(R_j-R_i)/400}},$$
$$\kappa = r_C - r_G = \frac{K}{2}.$$

The *draw payoff*, r_G, measures the change from the chooser's pre-game rating. It is positive when the opponent is higher-rated ($R_j > R_i$), negative when the opponent is lower-rated ($R_j < R_i$), and zero when their ratings are equal. The *forgone win premium*, κ, measures the additional rating points the chooser would receive by claiming instead.[15] It is always positive and, conditional on K, does not depend on the rating difference. A player can

[14] Specifically, $g_j = [1 + 3(\mathrm{RD}_j/173.7178)^2/\pi^2]^{-1/2}$, where RD_j is the opponent's reconstructed pre-game rating deviation. Setting $g_j = 1$ gives the familiar Elo expected-score formula.

[15] The two hypothetical changes use the same update scale by construction. The platform's game update applies the rating deviation already adjusted for elapsed time, and the volatility it computes after the game affects only later deviation growth, so the scale does not depend on which result is realized.

therefore gain rating points from a draw while still sacrificing points relative to a win.

The premium averages 11.85 rating points among competitive-position grants, with a median of 5.71. On two further scales, the median difference between the modeled win and draw outcomes is 0.45 percentile points of the preceding month's observed same-speed rating distribution, and among grants with sufficient rating history the median premium is 0.29 of the chooser's own rating standard deviation over the preceding 100 games; these are separate medians (Appendix C.4). The ratings, rating deviations, and update scale are determined before play, so the reconstructed consequences do not change with the subsequent board position. They do vary across games with the chooser's rating state. The update scale is set chiefly by the chooser's rating deviation: $\log K$ correlates 0.897 with the chooser's pre-game rating deviation and 0.017 with the opponent's (Appendix C). Across accounts, κ therefore varies with each account's typical rating state, which chooser fixed effects absorb. Within an account, κ moves with that deviation, most in an account's first games and after inactivity, when who is playing also changes (Appendix B.8); the rating-deviation controls of equation (4) absorb those movements. The paper therefore conditions on κ rather than interpreting its coefficient as a response to the points given up. The position profile is the same with and without these controls (Appendix B.1).

3 Framework

When the opponent disconnects, the chooser must end the game by claiming the win or granting a draw. This section models that choice. A chooser who construes the screen as a choice asks two questions: whether the win is deserved, and how much the chooser has put into the game. The answers are weighed against the rating points a draw gives up. The model is reduced-form; its purpose is to say which comparisons in Sections 4 to 6 bear on which part of the decision, and what the concession and the supporter badge reveal about an account.

3.1 The Decision

Chooser i faces opportunity o. The state of the game at the disconnection is the vector $(v_o, \Delta r_o, k_o, t_o, \tau_o)$: the opponent's engine evaluation v, the displayed rating gap Δr (opponent minus chooser), the two clocks k, the ply count t, and the time control τ. Claiming changes the chooser's rating by $r_{C,o}$ and granting by $r_{G,o}$; the difference $\kappa_o = r_{C,o} - r_{G,o} = K_o/2 > 0$ is the number of rating points a draw gives up (Section 2.5). The opponent's rating changes are fixed by the same pre-game states.

3.2 Two Questions

Construal. Not every chooser sees a choice. Let $c_i \in \{0, 1\}$ indicate whether account i construes the screen as offering one. An account with $c_i = 0$ claims without weighing anything; an account with $c_i = 1$ asks the two questions below. Most accounts have $c_i = 0$, which is the model's account of why most accounts never grant. A construer can still let a given opportunity pass without deliberation, in a hurry or distracted; that is part of the opportunity-specific shock introduced below.

Do I deserve this win? A chooser who construes the choice first asks whether the win is deserved, and answers by asking whether the opponent could have come back. Let

$$\pi_o \;=\; \pi(v_o, \Delta r_o, k_o, t_o, \tau_o) \tag{1}$$

be the probability that the opponent would have avoided defeat had play continued. It rises with the opponent's evaluation, with the opponent's rating advantage, and with the opponent's remaining time, and falls with the chooser's remaining time; how decisive a given evaluation is depends on the ply count and the time control. The function π describes how games on the platform continue. The chooser is not assumed to compute it, only to judge the board, the ratings and the clocks in a way that is monotone in it.

How undeserved the win is, from the chooser's point of view, is $m(\pi_o)$, with $m(\pi) = 0$ for $\pi \leq \underline{\pi}$ and m strictly increasing above that threshold. Below the threshold the opponent had no realistic chance, the win is fully deserved, and the question has a definite answer: claim. Above it the answer is a matter of degree.

How much have I put into this game? The second question concerns the chooser's investment. Let $s_o \geq 0$ measure it. The model does not fix how a chooser counts investment; three measures are natural, and they make different predictions across time controls (Section 3.4): the number of moves played, the time the chooser has spent, and the share of a typical game at this time control and rating that has been completed. Investment is sunk whichever option the chooser takes, so it bears on neither outcome. It enters because a win makes the effort count and a draw does not, the same sense of entitlement that makes people give away less of what they have earned, first shown by Hoffman and Spitzer [1985], whose subjects kept more when they had won the right to divide by winning a game than when a coin had assigned it [see also Hoffman et al., 1994, Cherry et al., 2002].

3.3 The Choice

The two options give the chooser

$$\begin{aligned} U_{io}(\text{claim}) &= \lambda_i\, r_{C,o} \;+\; \gamma\, s_o \;+\; a \;+\; \eta^{C}_{io}, \\ U_{io}(\text{grant}) &= \lambda_i\, r_{G,o} \;+\; b_i\, m(\pi_o) \;+\; \eta^{G}_{io}. \end{aligned} \tag{2}$$

The weight $\lambda_i \geq 0$ is the value account i places on a rating point. The weight $b_i \geq 0$ is the value it places on a result the game deserved, whatever its source: fairness, sportsmanship, altruism toward a rival struck by bad luck, or the self-image of someone who does not take undeserved wins. A draw delivers that value in proportion to how undeserved the win would have been. The term γs_o, with $\gamma > 0$, is the entitlement from the chooser's investment, and the constant $a > 0$ is the pull of the default and the formal norm, both of which favor claiming (Section 2.2). The shocks are opportunity-specific, independent of the account's traits, and their difference $\varepsilon_{io} = \eta^{C}_{io} - \eta^{G}_{io}$ is continuously distributed with full support and a log-concave density, with distribution function F.

A chooser who construes the choice therefore grants if

$$b_i\, m(\pi_o) \;-\; \gamma\, s_o \;-\; \lambda_i\, \kappa_o \;-\; a \;\geq\; \varepsilon_{io}, \tag{3}$$

and the probability of a grant is

$$P_{io} \;=\; c_i\, F\big(b_i\, m(\pi_o) - \gamma\, s_o - \lambda_i\, \kappa_o - a\big).$$

The grant happens when the win is undeserved enough, and the investment small enough, that the value of a deserved result exceeds the rating points a draw gives up together with the pull of the default. Equation (3) states that sentence.

Lemma 1 (Only the points given up matter)**.** *The choice depends on the rating consequences only through κ_o: holding κ_o fixed, the level of the draw payoff $r_{G,o}$ does not affect it.*

The rating terms in (2) differ by $\lambda_i(r_{C,o} - r_{G,o}) = \lambda_i \kappa_o$, and nothing else in either utility depends on the rating payoffs. Whether a draw gains or loses points therefore plays no role in the choice; the opponent's rating enters only through π, that is, through the opponent's chance of a comeback.[16]

[16] Loss aversion over rating points, or any value of them concave around the pre-game rating, would break this: a draw that costs points would weigh more than one that gains them, and the cost of a concession would grow with the size of the loss. Section 4.5 tests this through the shape of the gap's effect.

3.4 What the Model Predicts

Proposition 1 (Comparative statics)**.** *An account with $c_i = 0$ never grants. For an account with $c_i = 1$:*

(i) P_{io} is nondecreasing in the opponent's evaluation, in the opponent's rating advantage and in the opponent's remaining time, strictly so where $b_i > 0$ and $\pi_o > \underline{\pi}$, and it does not depend on any of them where $\pi_o < \underline{\pi}$.

(ii) P_{io} is decreasing in the investment s_o, in the points given up κ_o, and in λ_i, and increasing in b_i where $\pi_o > \underline{\pi}$.

Part (i) is the first question: the board, the rating gap and the opponent's clock matter only where the opponent could have come back, and not at all where the win is fully deserved. Part (ii) is the second question and the points given up. The empirical work tests part (i) and the investment part of (ii); it conditions on κ rather than testing it, because κ moves with the account's rating state (Section 2.5).

Three further implications follow from equation (3). Where the opponent had clearly lost, π lies below $\underline{\pi}$ and neither the rating gap nor the clocks affect the choice, whereas in competitive positions both do; in the data the contrast between an opponent 100 points stronger and one 100 points weaker is 1.02 points in competitive positions and 0.04 in lost ones, and a log second more on the opponent's clock is worth 1.37 points in the one and -0.14 in the other (Table 2; Section 4.5). The ply count enters twice, through investment, which rises with every move, and through π: late in a game a given advantage is harder to overturn, but in long, level games a draw becomes the likely result, so the opponent's chance of avoiding defeat rises and granting with it (Table 2, Panel B). And the three measures of investment make different predictions across time controls at a fixed π: if investment is counted in moves, granting declines at the same rate per move in every time control; if in time spent, it declines faster per move where moves take longer; if in the share of a typical game completed, it declines faster per move where a typical game is shorter (Appendix B.5).

3.5 What a Grant Reveals

The two kinds of position reveal different traits. Where the opponent had clearly lost, $m = 0$ and a grant occurs with probability $c_i F(-\gamma s_o - \lambda_i \kappa_o - a)$: it requires that the account construe the choice and that the points given up weigh little, and it does not depend on b_i. Where the opponent was still in the game, a grant occurs with probability $c_i F(b_i m(\pi_o) - \gamma s_o - \lambda_i \kappa_o - a)$, larger the larger b_i and the smaller λ_i. Both require $c_i = 1$, so it is the contrast between the two, not the level of either, that separates b_i from λ_i.

A second decision loads on the same traits. Let S_i indicate the supporter badge, with $S_i = 1$ if $p_0 + \tau_b b_i + \tau_\lambda \lambda_i + \nu_i \geq 0$, where ν_i is continuously distributed with full support: τ_b links valuing a deserved result to supporting the platform, and τ_λ links taking the platform's competition seriously to supporting it. The badge index does not depend on c_i. Write G_i^C and G_i^L for a grant at a competitive and at a clearly-lost opportunity, within a κ cell and at a fixed investment, and assume that c_i, b_i, λ_i, ε_{io} and ν_i are mutually independent. The second ranking in part (iii) uses one further condition, *lost-only monotonicity*: the probability that a construing account grants at the lost opportunity and claims at the competitive one does not rise with the weight on rating points. It holds whenever no construing account grants a competitive opportunity with probability above one half (Appendix E.4).

Proposition 2 (What a grant reveals)**.**

(i) If $\tau_\lambda > 0$ and λ_i is nondegenerate, then $\mathrm{Cov}(S_i, G_i^L) < 0$ for any $\tau_b \geq 0$.

(ii) If $\tau_\lambda = 0$, $\tau_b > 0$, and b_i is nondegenerate, then $\mathrm{Cov}(S_i, G_i^C) > 0$; if $\tau_b = 0$, $\tau_\lambda > 0$, and λ_i is nondegenerate, then $\mathrm{Cov}(S_i, G_i^C) < 0$. With both loadings positive, the sign depends on their relative size.

(iii) If $\tau_\lambda > 0$, $\tau_b \geq 0$, and λ_i is nondegenerate, then among accounts that grant in competitive positions, those that never grant in lost positions hold the badge more often than those that also grant in lost positions. If in addition lost-only monotonicity holds, then among accounts that never grant in competitive positions, those that never grant hold the badge more often than those that grant only in lost positions; if every account construes the choice, this second ranking holds without the condition.

A negative association in lost positions is the signature of a badge that loads on how seriously an account takes its rating; the sign in competitive positions is a race between the two loadings; and part (iii) orders the four granting histories of Table 4, Panel C, at fixed opportunity counts and conditions, leaving the comparison of accounts that grant in both kinds of position with accounts that grant in neither undetermined. A single trait that raised granting everywhere, c_i included, would give the same sign in both kinds of position. The proposition does not say what b_i is, and λ_i near zero covers any account for which the rating carries no weight.

3.6 Receipt and Persistence

Receiving a granted draw raises, for a time, the value the recipient places on a deserved result: $b_{io} = b_i + \rho_{io}$, where $\rho_{io} \geq 0$ rises after the account receives a grant and decays

with the time since, as generosity received does in experiments on upstream reciprocity (Section 1.1). Three implications follow. The increase follows the other player's act, not the result, so a regular draw, without a disconnect screen, should not raise later granting (Section 6.3). It does not depend on whether the recipient deserved the draw it received (Section 6.3). It raises the weight on a deserved result and not the weight on rating points, so it should raise merited granting and, where desert does not act, leave unmerited granting alone (Appendix D.5).

And because it raises a weight and does not switch construal on, it should raise later granting by a similar proportion among recipients who had faced the choice before and among those who had not, whatever the share of construers in each, and fade with time; a receipt that switched c_i from zero to one, by showing the recipient that the option exists, would instead concentrate the effect among recipients who had never faced the choice. The matched comparison shows the fade (Section 6.1), and Section 6.2 reports the comparison by prior experience of the choice.

The traits c_i, b_i and λ_i are fixed over the panel, so an account's grant rate in each kind of position persists. Because c_i switches both kinds on together, the two persist separately only through b_i and λ_i: a year-one rate in one kind of position should predict the year-two rate in the same kind far better than in the other (Section 5.1).

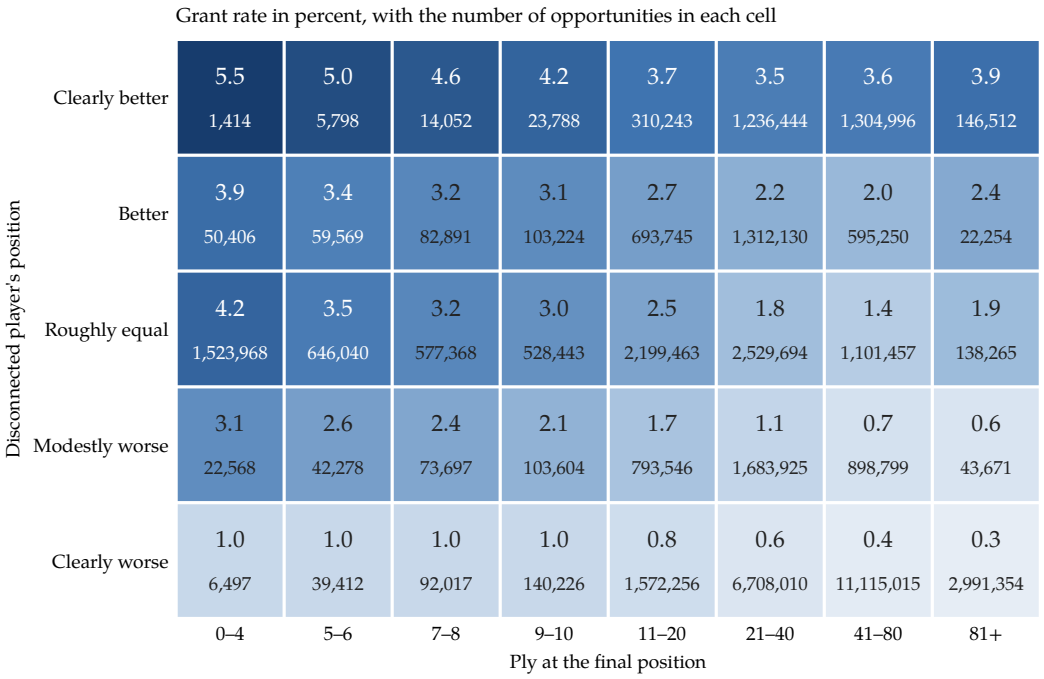
4 When Players Grant a Draw

4.1 The Raw Pattern

Section 3 reduces the decision to two questions. Is the win deserved? How much has the chooser put into the game? The board answers both. The opponent's final position and the rating gap speak to the first and the stage of the game to the second. Proposition 1 says how the grant probability should move in each.

Granting is 2.76 percent in competitive positions and 0.47 percent in clearly worse positions, a ratio of almost six. Across the five bins it ranges from 3.58 percent, when the disconnected opponent's position was clearly better, to 0.47 percent when it was clearly worse (Table A.3). Figure 2 crosses the five bins with the stage of the game. Granting rises from clearly worse to clearly better within every stage band, and falls with stage within every position. The opportunity to claim is available in every cell. The concession varies with what happened on the board. A position-independent tendency to press the wrong button would not produce the pattern. Nor would it differ by format: the same chooser grants less often in arena tournament games than in pool games, and the shortfall falls

Figure 2: **Granting by board position and game stage.** Each cell shows the raw grant rate, in percent, among opportunities in which the disconnected player's final position fell in the row's bin and the game ended in the column's ply band, with the number of opportunities beneath it. Rows follow the five bins of Section 2.4; the first three form the competitive region.



Notes: All 41,534,289 opportunities; no confidence intervals. Table 2 conditions on the chooser, the format, and the month.

on competitive positions (Appendix B.9).[17]

4.2 One Regression

The raw cells mix choosers, formats, and months. I estimate both movements from the same chooser, in one regression,

$$\begin{aligned} 100\, G_{ig} = \alpha_i + \sum_k \beta_k \, \mathbf{1}[v_g \in k] + \sum_b \left(\delta_b^C C_g + \delta_b^L (1 - C_g)\right) \mathbf{1}[t_g \in b] \\ + C_g \, h_C(d_g) + (1 - C_g)\, h_L(d_g) + x_g' \theta + \psi_{c(g)} + \omega_{m(g)} + u_{ig}, \end{aligned} \tag{4}$$

[17] The shorter countdown when the disconnected player is materially behind can affect both the wait before a decision and which departures last long enough to enter the opportunity sample. Appendix B.7 tests for a step in the grant rate at the material threshold that shortens it and finds none.

where G_{ig} indicates a grant by chooser i in game g and α_i is a chooser fixed effect. The indicator $\mathbf{1}[v_g \in k]$ marks the opponent's final evaluation as falling in position bin k, one of the five of Section 2.4; clearly worse is the reference. The indicator $\mathbf{1}[t_g \in b]$ marks the game as ending in stage band b, one of the eight of Figure 2; plies 0–4 are the reference. The rating gap d_g is the opponent's displayed pre-game rating minus the chooser's, $R_j - R_i$, and h is piecewise linear in it, with knots at -300, -100, 0, $+100$, and $+300$ and capped at ± 600. The indicator $C_g = \mathbf{1}[v_g \geq -100]$ marks the competitive region.

Proposition 1 says the stage and the gap should work differently where the opponent could still have come back, so the stage profile and the gap function each take one form in competitive positions, δ^C and h_C, and another in the rest, δ^L and h_L. The vector x_g holds the log of each player's last recorded clock, the log of the chooser's forgone premium κ and its square, and the log of each player's rating deviation and its square (Section 2.5). The ratings themselves are not separate controls: the chooser's is largely absorbed by α_i, and the opponent's enters as the gap d_g. The terms ψ and ω are speed-category and month effects.

The sample is all 41.5 million opportunities. Standard errors are clustered by chooser. The β_k estimates answer the first question: how much more often the same chooser grants, at a given stage and rating gap, as the opponent's claim on the result grows. The δ_b estimates answer the second: how much less often the same chooser grants, at a given position, the longer the game has run. The estimated functions h_C and h_L give the gap's part in the first question in each region. Table 2 reports all three from the one estimation. Table B.1 compares it with the version without x_g and with a single gap function; neither changes the picture.

4.3 Evaluation

The regression confirms the position gradient within chooser. Hold the chooser, the stage, the rating gap, the format, and the month fixed, and average over the sample's stages and gaps. The adjusted rate then runs from 0.63 percent in clearly worse positions to 3.63 in clearly better ones, a difference of 3.00 percentage points (standard error 0.014), and each step up the ordering is significant: 0.55 points from clearly to modestly worse, 0.86 to roughly equal, 0.30 to better, and 1.29 to clearly better (Table 2, Panel A). The binary version of the contrast, competitive against clearly worse within accounts observed in both regions, is 2.17 percentage points (Appendix B.2).

Within the competitive region the relationship is continuous. Among the 11.6 million competitive opportunities after ply ten, a natural cubic spline in the evaluation replaces

Table 2: **Position, stage, and rating gap within chooser**

Panel A. Does the chooser deserve the win? Position

	Adjusted rate (%)	Step from bin below
Clearly worse (≤ -300)	0.63 (0.005)	—
Modestly worse (-299 to -101)	1.19 (0.008)	0.55 (0.006)
Roughly equal (-100 to $+100$)	2.05 (0.009)	0.86 (0.009)
Better ($+101$ to $+299$)	2.34 (0.012)	0.30 (0.009)
Clearly better ($\geq +300$)	3.63 (0.015)	1.29 (0.013)
Clearly better minus clearly worse		3.00 (0.014)

Panel B. How much has the chooser put into the game? Stage, relative to the region's plies 0–4

	Competitive	Not competitive
Plies 5–6	-0.766 (0.024)	-0.659 (0.095)
Plies 7–8	-1.090 (0.024)	-0.828 (0.089)
Plies 9–10	-1.314 (0.024)	-0.976 (0.088)
Plies 11–20	-1.753 (0.018)	-1.233 (0.085)
Plies 21–40	-2.253 (0.018)	-1.497 (0.085)
Plies 41–80	-2.331 (0.019)	-1.562 (0.085)
Plies 81 or more	-1.961 (0.033)	-1.633 (0.085)
Plies 81 or more minus plies 41–80	0.370 (0.029)	-0.070 (0.005)
Decline to 41–80, share of the 0–4 rate	0.55 (0.003)	0.60 (0.016)

Panel C. Rating gap, opponent minus chooser: differences in h

	Competitive	Not competitive	Difference
$+100$ versus -100	1.018 (0.018)	0.037 (0.007)	0.980 (0.019)
Equal versus -100	0.629 (0.016)	0.001 (0.007)	0.627 (0.017)
$+100$ versus equal	0.389 (0.017)	0.036 (0.007)	0.353 (0.019)
$+300$ versus equal	0.183 (0.104)	0.345 (0.046)	-0.162 (0.112)
-300 versus equal	-0.534 (0.085)	0.312 (0.035)	-0.845 (0.092)

Notes: Equation (4) on all 41,534,289 opportunities from 3,224,171 chooser accounts; percentage points unless stated, standard errors clustered by chooser in parentheses. Panel A: adjusted rates standardize every opportunity to the bin, resetting its region, stage, and gap interactions; steps are differences of adjacent adjusted rates with delta-method standard errors from the joint covariance. Panel B: each region's profile relative to its own 0–4 band; the decline row divides the 41–80 coefficient by the region's raw 0–4 rate (4.20 and 2.61 percent), and the two late changes differ by 0.440 (0.029). Panel C: differences in h_C and h_L at the stated gaps in rating points. Table B.1 compares specifications.

the bins of equation (4), with the same fixed effects (Appendix B.3). The adjusted rate is 1.92 percent at equality, 2.69 percent at +300 centipawns, 3.71 percent at +500, and 4.75 percent when the opponent had a forced mate. The rise from +300 to +500 alone is 1.02 percentage points, with a standard error of 0.018 (Table B.4). Concessions become more common even as an already winning opponent's advantage grows.

Material, which a player can read without an engine, carries a gradient of its own within the competitive region, and material and the evaluation each retain their relationship when they enter together (Appendix B.3). Signs that the departure was deliberate matter: an opponent whose last move lost at least 200 centipawns receives a draw 0.55 points less often, but holding such signs fixed leaves the rise from +300 to +500 in the 20,000-account sample at 1.37 points against 1.41 without them (Appendix B.6). The position gradient is the same among accounts more than a year old (Appendix B.8), 0.88 points flatter in arena tournament games (Appendix B.9), and lower, not higher, toward opponents met before, and intact where another meeting is unlikely, with direct repayment essentially absent (Appendix B.10).

4.4 Game Stage

The stage gradient is as steep. Relative to a game that ended within the first four plies, the same chooser in a competitive position grants 0.77 percentage points less often at plies 5–6, 1.75 less often at plies 11–20, and 2.33 less often at plies 41–80 (Table 2, Panel B). In a lost position the declines are 0.66, 1.23, and 1.56. The decline runs through both regions at a similar proportional rate: by plies 41–80 it is 55 percent of the 0–4 rate in competitive positions and 60 percent in losing ones, and in both about three quarters of it is complete by ply twenty. That is what the framework expects if the stage enters through investment in both.

The two profiles part at the end. In competitive positions the rate turns up in games of 81 or more plies, by 0.37 percentage points (0.03) relative to plies 41–80. In losing positions it continues down, by 0.07 (0.005). That is the pattern Section 3.4 describes. In a long, level game a draw becomes the likely result, so the opponent's chance of avoiding defeat rises and granting with it. In a lost position the stage enters through investment alone.

The raw rates are of the same order. Within competitive positions, 4.57 percent of opportunities are granted after one move each, 3.48 percent between plies three and ten, and 2.47 percent after ply ten (Appendix B.4). A win after one move each is a windfall. Later in the game the chooser has competed for it. The early rate can also reflect a wish to undo a game that has barely begun, since Lichess allows an abort only at the start. But

the decline continues beyond the opening, and the position difference remains in games ending after ply ten, 2.47 against 0.46 percent. How the chooser counts the investment is taken up in Appendix B.5, which estimates the profile separately by time format: the decline per ply band differs across formats, which rules out a count in moves alone.

4.5 Rating Gap

The gap enters the same regression, separately by region. In competitive positions the chooser grants 1.02 percentage points (0.02) more often toward an opponent 100 rating points stronger than toward one 100 points weaker. In lost positions the same contrast is 0.04 (0.01). The difference is 0.98 (0.02) (Table 2, Panel C). That is the pattern Proposition 1 predicts. The framework reads the gap as part of the opponent's chance of a comeback, so it should matter where the opponent could still have held and not where the win was already deserved.

Within competitive positions the grant rate rises with the opponent's rating up to about 100 points either side of equality and flattens beyond: 0.63 points from 100 points weaker to equal, 0.39 from equal to 100 points stronger, and little further out. An opponent 300 points stronger receives 0.18 (0.10) more than an equal one, and an opponent 300 points weaker 0.53 (0.09) less, about the same as one 100 points weaker. In lost positions the function is flat within 100 points of equality and rises by about 0.3 points at both extremes. Appendix C.2 gives the finer profile in roughly equal positions after ply ten, where the same chooser grants 0.47 points less often toward an opponent about 100 points weaker and 0.25 points more often toward one about 100 points stronger.

The gap also moves the chooser's own payoff: a draw against a stronger opponent raises the chooser's rating, and against a weaker one lowers it. If players weighed those points around the pre-game rating, with losses weighing more than gains, the gap would also enter through the points given up. Granting would fall faster below equality than it rises above it, and the fall would continue as the loss grows, since the extra cost of a draw rises with the opponent's expected score. The first pattern appears: the slope is 0.63 just below equality against 0.39 just above. On its own it cannot separate a kink at equality from a smooth response whose steepest part lies just below it, and the opponent's expected score, through which the gap moves the opponent's chances, is itself steepest near equality.

The second pattern separates them, and it is absent. From -100 to -300 the opponent's expected score rises from 0.64 to 0.85, so a draw costs two and a half times as many points; if the asymmetry at equality reflected the points given up, granting would fall by about a third of a point more over that range. Instead both ends of the profile turn back

toward zero: granting is 0.10 points higher at -300 than at -100 (standard error 0.09) and 0.21 points lower at $+300$ than at $+100$ (0.11). In lost positions, where a draw costs the same points, granting rises rather than falls toward much weaker opponents. Nor do players treat the sign of the payoff as a reference point, in the manner of par: a chooser for whom the draw would gain a point grants no more often than one for whom it would cost a point, -0.03 points (0.07; Appendix C.2).

The third input to the opponent's chances, the opponent's remaining time, enters equation (4) through x_g. Letting the two clock terms differ by region gives the same contrast as the gap (Appendix B.1). In competitive positions a log second more on the opponent's clock raises granting by 1.37 points (0.02), and a log second more on the chooser's own clock, which also measures time not yet invested, lowers it by 1.21 (0.01). In lost positions the same coefficients are -0.14 (0.005) and 0.15 (0.005), about a tenth as large and of the opposite sign. Proposition 1(i) also says that below the threshold nothing that moves the opponent's chances should matter. The gap and the clocks behave that way. The engine evaluation itself does not: granting rises by 0.55 points from clearly to modestly worse positions, and the full-range profile keeps rising from -600 to -100 (Appendix B.3). The chooser responds to the position throughout the evaluation range, from -600 to $+600$.[18]

5 One Disposition or Two?

The same button records two kinds of concession. A merited concession is a draw granted to an opponent who could still have come back, a position in the competitive region of Section 2.4. An unmerited concession is a draw granted to an opponent whose position was clearly lost, an evaluation of -300 or worse. Modestly worse positions lie between and enter neither comparison. Section 4 showed that merited concessions rise with the opponent's advantage, rating, and remaining time. Below the competitive region the same inputs do little: the rating gap near equality and the clocks move granting by about a tenth as much, and the gap's small movement at the extremes rises toward stronger and weaker opponents alike. This section asks whether the two come from one disposition to give or from two traits. It compares them on what they predict: whether each kind of grant persists within an account from one year to the next, and whether each is associated with paying to support the platform.

[18] One potential concern is the platform's shortening of the countdown when the disconnected player is down four or more points of material (Section 2.1), which could change which disconnections in lost positions reach the chooser. Appendix B.7 tests for a step in the grant rate at that threshold and finds none; the countdown does not drive the gradient.

5.1 Persistence by Kind of Grant

An account's grant rate in one year predicts its rate in the next. I divide the panel into November 2023–October 2024 and November 2024–October 2025 and compute each account's competitive-position grant rate in each period. Among the 666,000 accounts with at least one such opportunity in each year and at least four in total, the correlation is 0.545 (standard error 0.003), or 0.618 (0.003) weighting accounts by their opportunities (Table 3, Panel A). Accounts with a first-year grant go on to grant at 15.8 percent of their second-year opportunities; accounts without, at 0.9 percent.[19] Accounts that never grant remain in these comparisons.

It persists by kind. Among the 266,000 accounts with at least four opportunities in each board state in each year, the competitive-position rate correlates 0.727 (0.003) across years and the clearly-worse rate 0.494 (0.018), while each kind correlates only 0.24 to 0.27 (0.007) with the other kind a year later (Panel B). If each year measures a fixed propensity with error, the two propensities correlate 0.40 to 0.45 (0.008); the figure is the same at thresholds of one and ten opportunities and in games ending after ply ten. Each kind of grant predicts itself far better than it predicts the other. That is what the framework's two weights produce, though it does not by itself rule out one trait: a single disposition with a much more convex loading on lost grants can also leave the two rates weakly correlated (Appendix E.5). The comparison that does rule it out is the badge ordering of Section 5.2.

Transitions. Panel C shows the same in counts. Accounts that granted only in competitive positions in the first year are competitive-only granters again in the second at 40.6 percent, against 4.7 percent of first-year non-granters, and clearly-worse-only granters at 2.3 percent against 2.1. Accounts that granted only in clearly worse positions repeat at 7.3 percent against 2.1 and become competitive-only granters at 8.7 percent against 4.7. A merited concession predicts another a year later at nearly nine times the base rate, and almost nothing else. An unmerited concession predicts another at three and a half times the base rate and a merited one at about twice; the doubling from 4.7 to 8.7 percent is what construal implies, since a grant in a lost position marks an account that sees the choice. Between 50 and 52 percent of each first-year group returns with enough opportunities in both states, so the cohort is not selected differently by group.

Sampling variation. A player with few opportunities may grant in one period and not the other at a constant propensity, which bounds how high any year-to-year correlation can be. Dividing each account's opportunities into two groups at random and correlating

[19] These pooled rates weight accounts by second-year opportunities, the grouping partly reflects how many chances an account had in the first year, and the cohort is the one that returns for a competitive choice in both periods; accounts with a first-year grant return more often, 65.9 against 58.3 percent. Appendix F.1 reports the cohort construction and counts.

Table 3: **Persistence of granting across years**

Panel A. Competitive-position grant rate, 666,131 accounts	Estimate (SE)	Accounts
Correlation of yearly grant rates, equal account weights	0.545 (0.003)	666,131
Correlation of yearly grant rates, total-opportunity weights	0.618 (0.003)	666,131
Second-year grant rate, no first-year competitive-position grant (%)	0.938 (0.007)	603,597
Second-year grant rate, at least one first-year competitive-position grant (%)	15.824 (0.123)	62,534
Rate ratio: prior grant / no prior grant	16.865 (0.176)	—

Panel B. Persistence by board state, 266,399 accounts	Estimate (SE)
Correlation across years, competitive-position rate	0.727 (0.003)
Correlation across years, clearly-worse rate	0.494 (0.018)
Correlation, first-year competitive rate with second-year clearly-worse rate	0.239 (0.007)
Correlation, first-year clearly-worse rate with second-year competitive rate	0.271 (0.007)
Implied correlation of the two propensities, from the first-year competitive rate	0.399 (0.008)
Implied correlation of the two propensities, from the first-year clearly-worse rate	0.452 (0.008)

Panel C. Second-year granting group by first-year group, same 266,399 accounts (% of row)

First-year group	Accounts	Neither	Competitive only	Clearly worse only	Both
Neither	231,905	92.7 (0.05)	4.7 (0.04)	2.1 (0.03)	0.4 (0.01)
Competitive only	24,896	52.6 (0.32)	40.6 (0.31)	2.3 (0.09)	4.6 (0.13)
Clearly worse only	5,848	81.8 (0.53)	8.7 (0.38)	7.3 (0.35)	2.2 (0.20)
Both	3,750	33.5 (0.77)	37.9 (0.78)	4.6 (0.35)	23.9 (0.68)

Notes: Periods are November 2023 through October 2024 and November 2024 through October 2025. Panel A: accounts with at least one competitive-position opportunity in each period and four in total; Pearson correlations across accounts, unweighted and weighted by competitive opportunities. Panels B and C: accounts with at least four competitive and four clearly-worse opportunities in each period. The implied rows divide each cross-state correlation by $\sqrt{0.727 \times 0.494}$. Standard errors, in parentheses, come from 1,999 bootstrap resamples of accounts, each retaining its complete two-year history and the original opportunity weights where applicable; Panels B and C share each resample. Counts and variants are in Appendices F.1 and F.2.

the resulting rates gives 0.600; the year-to-year correlation of 0.545 is about nine tenths of it. The propensity is close to as stable as the data can show.

Game conditions. An account that repeatedly faces positions where grants are common would persist without a stable propensity. Adjusting each opportunity for board state, game stage, and speed before computing the annual rates leaves the correlation among established returning accounts at 0.527, the same as without adjustment. The conditions an account meets do not account for it.

Horizon. Two years is the panel's length, and it bounds how much persistence can be seen. It shortens the window for both kinds of grant alike and does not by itself produce the gap between own-kind and cross-kind prediction. These are year-to-year associations within returning accounts rather than identified types. They are the pattern Section 3.6 predicts if what separates the two kinds of grant is stable within an account: whether an account construes the screen as a choice switches both on, and the value of a deserved result and the weight on rating points move them apart.

5.2 Support for the Platform

Lichess is free, and Patron status is a public badge associated with monetary support for the platform, a different outcome from giving up rating points to an opponent. I link granting during the panel to badge possession in an August 2026 profile snapshot. Because a badge can be purchased or gifted and the snapshot does not record when support began, the measured outcome is possession of the badge at that date.

The matched cohort was fixed before querying the profiles. Each account with a grant was assigned up to three accounts with no grants anywhere in the panel, matching on competitive-opportunity counts and modal game speed. The main comparison retains 213,000 accounts that granted in a competitive position and 638,000 controls. Patron rates are 0.64 and 0.33 percent, respectively. With matching-cell fixed effects and account controls, the association is 0.28 percentage points (HC1 standard error 0.019; Table 4, Panel A). Appendix E describes the matching, profile coverage, and controls, including those measured at the later query.

The more informative comparison distinguishes the circumstances of granting. A broader acquisition cohort includes accounts that granted in any state and their selected controls. Among those with available August profiles, 599,000 have at least four opportunities in each of the competitive and clearly-worse regions. I enter indicators for ever granting in each region in the same model for these accounts. Competitive-position granting has a positive association with Patron status, $+0.39$ percentage points; clearly-worse-position granting has a negative association, -0.19 points (Panel B). The raw rates

tell the same story. Badge possession is 0.90 percent among accounts that granted only in competitive positions, 0.26 percent among those that granted only in clearly worse positions, and 0.43 percent among accounts that granted in neither region (Panel C). Accounts that granted in both regions hold the badge at 0.49 percent, close to the neither-region rate, so a competitive-position grant is associated with a higher badge rate mainly among accounts that never granted in a lost position, as Section 3.5 anticipates.

Table 4: **Granting and Patron status**

Panel A. Matched comparison

	Granting accounts	Controls	Difference (pp)
Accounts	212,828	638,298	—
Patrons	1,369	2,114	—
Patron rate (%)	0.643	0.331	0.312
Adjusted association, cell fixed effects and full controls	—	—	0.277
HC1 standard error	—	—	0.019
Standard error clustered by matching cell (36 cells)	—	—	0.035

Panel B. Granting by board state, one model (percentage points)

	Estimate	HC1 SE	Cell-clustered SE
Ever granted in a competitive position	0.388	0.028	0.032
Ever granted in a clearly-worse position	−0.195	0.031	0.033
Competitive minus clearly worse	0.583	0.046	0.059

Panel C. Raw Patron rates by granting history, same accounts

	Patrons / accounts	Rate (%)
Granted only in competitive positions	908 / 100,671	0.902
Granted in both states	89 / 18,098	0.492
Granted in neither state	1,957 / 454,045	0.431
Granted only in clearly-worse positions	68 / 26,231	0.259

Notes: Patron status is read from public profiles queried in August 2026; it can be bought, gifted, or lapse, and the profile does not show when it began. Controls, which granted in no position, were drawn within cells of competitive-opportunity count and modal speed before any profile was observed, up to three per granting account. Account-level linear probability models with cell fixed effects and the controls of Appendix E; HC1 standard errors. Panel B: the 599,045 accounts with at least four opportunities in each state. Missingness is reported in Appendix E.

Proposition 2 makes the sign pattern informative. A trait that raised granting in every position, whether the account construes the screen as a choice among them, would give the badge the same sign in both kinds of position; the estimates have opposite signs. Under the proposition, the negative association in lost positions marks a badge that loads on the weight an account places on its rating, and the positive association in competitive positions then requires that the badge also load, and more strongly, on the value the account places on a deserved result. Part (iii) gives two rankings among the four

granting histories of Panel C, the second under lost-only monotonicity, and the raw rates follow both: accounts that granted only in competitive positions hold the badge more often than those that granted in both, and accounts that granted in neither more often than those that granted only in lost positions. The raw groups pool accounts with different opportunity counts and game conditions, which the proposition holds fixed, so the agreement is consistent with the proposition rather than implied by it. The regression identifies neither loading separately.

Account tenure provides a concrete competing explanation. Established users may both support the platform and grant selectively, while newer users may grant in lost positions and subsequently leave. I repeat the comparison using only opportunities faced after an account was at least a year old, with a smaller control set than Table 4, Panel B. The associations remain positive for competitive-position granting, $+0.60$ percentage points (standard error 0.04), and negative for clearly-worse-position granting, -0.27 points (0.05). These results connect the circumstances of giving to a separate marker of support, without requiring that granting itself causes the later badge outcome.

The badge also marks the recipient. Opponents who later hold the platform's Patron badge receive draws 0.32 percentage points (standard error 0.11) more often in competitive positions and no more often in lost ones (Appendix E); badge possession is observed in the snapshot rather than at the game, so the comparison does not identify a response to the badge as displayed.

6 Behavior After Receiving a Draw

An account can appear on both sides of the decision, so whether generosity received is passed on to a different opponent, as generalized reciprocity would suggest [Fowler and Christakis, 2010, Gray et al., 2014], can be examined directly. It is. A matched comparison and an instrumental-variables design that uses the assigned opponent's leniency toward other players agree.

6.1 The Matched Comparison

I begin with each account's earliest competitive-position, non-tournament disconnection in the panel, retained if the connected opponent granted or claimed and the pair had no earlier recorded rated game, and compare draw recipients with claim recipients facing similar exposure-game circumstances: board position, the sign of the chooser's draw payoff, rating gap, and game speed. The primary outcome is whether the recipient grants

at the first subsequent competitive choice within 90 days, with a complete window without such a choice coded as zero.

Draw recipients grant at their next competitive choice 0.40 percentage points more often than claim recipients (standard error 0.06), a 25 percent increase over the 1.56 percent grant rate of the reweighted claim recipients. The comparison holds the circumstances of the exposure game fixed: claim recipients are reweighted so that their distribution across cells of position, draw-payoff sign, rating gap, and speed matches that of draw recipients. The regression accounts for exposure-cell and month effects and the recorded conditions of the exposure game; standard errors are clustered by the opponent who made the original decision (Table 5).

Table 5: **Granting after receipt**

Outcome (percentage points)	Estimate	SE	Exposures
Grant at first later competitive choice, no later choice coded zero	0.399	0.058	2,110,399
same, censored at next competitive disconnection	0.308	0.047	2,110,399
Reach a competitive choice within 90 days	−0.197	0.194	2,110,399
Grant at first later competitive choice, conditional on reaching one	0.882	0.122	993,762

Notes: Draw recipients minus claim recipients at each account's first competitive-position, non-tournament exposure against a new opponent, within cells of position, draw-payoff sign, rating gap, and speed, with claim recipients reweighted to the cell's grant-to-claim ratio, cell and month fixed effects, and thirteen exposure-game controls. Weighted control means are 1.564 percent in the first row, 0.996 in the second, and 3.331 in the fourth. Standard errors clustered by exposure opponent. Appendix D gives the construction.

Returning to a choice and granting at it are distinct. Receipt has an estimated association of −0.20 percentage points with reaching such a choice within 90 days, with a standard error of 0.19. Among recipients who do reach a later competitive choice, draw recipients grant 0.88 percentage points (0.12) more often than reweighted claim recipients, whose rate is 3.33 percent; the increase is again about 25 percent of the baseline rate.

A common connectivity incident, such as an outage, could make opponents accommodating and recipients generous at once. Comparing draw and claim recipients whose exposure games began in the same clock hour leaves the association at 0.39 percentage points (standard error 0.06), and 0.39 using the hour of the last move. More localized shocks, and selection on the opponent's decision at the exposure itself, are the reasons to seek variation in receipt that does not depend on that decision.

The association is largest at the first competitive choice within six hours, 1.49 percentage points, and smaller at the first choice between 30 and 90 days, 0.50 points; these interval estimates use the accounts with a choice in each interval and describe a time profile, not a decay rate. Counting a later grant only if it comes before the recipient's next

competitive disconnection gives 0.31 points (0.05) against a control rate of 1.00 percent, and adding controls measured at the later choice, which requires that the recipient reach one, gives 0.90 (0.12) on the conditional outcome (Appendix D).

6.2 An Instrument from the Opponent's History

Some opponents are persistently more likely to grant than others. If the matchmaker pairs a recipient with a more lenient opponent, that recipient is more likely to receive a draw for a reason that has nothing to do with the recipient. I use this variation to instrument receipt, measuring leniency from the opponent's other decisions during the panel. The design follows the leniency designs that use an assigned judge or examiner as an instrument for a decision [Dobbie et al., 2018, Maestas et al., 2013, Chyn et al., 2025, Goldsmith-Pinkham et al., 2026]; here the opponent is the judge and the granted draw is the decision. A teammate's toxic speech in other matches has similarly instrumented exposure to toxic speech in an online game [Morrier et al., 2024]; here the counterpart's decision is a prosocial act.

Assignment works as follows. A player chooses one of eleven time controls and enters its pool. Every 12 to 60 seconds, or sooner once 20 to 40 players are waiting, the server pairs the players present by a minimum-weight matching on the absolute rating difference, subject to a cap on the allowed difference, a credit for waves already missed, and preferences for pairs whose optional user-set rating ranges are compatible or who have similar records under an internal conduct measure that counts abandoned games; blocked and flagged accounts are kept apart.[20] Neither player sees or chooses the other.

Cells of exposure month, exact time control, and rating bin therefore capture the main observed assignment dimensions, and exposure-date and hour-of-week effects absorb who is in the pool when; the user-set ranges, block lists, and the conduct measure are not reproduced by the cells. The conduct preference matters: players with similar records of abandoning games are paired together, which is why the recipients of frequently disconnecting opponents tend to be frequent disconnectors themselves and why the opponent's recent disconnection rates belong among the controls (Appendix D.3). Assignment is not entry. A game becomes an exposure only after the recipient's disconnection, a competitive final position, and an active resolution by the opponent, all of which follow assignment; that is why the design also tests entry.

The instrument is the opponent's grant score, adjusted for the board states of its de-

[20] Lichess source code, `modules/pool/src/main/MatchMaking.scala` at commits `99b72ad` (July 2023) and `bddf2f7` (July 2025), with the wave timing in `PoolActor.scala` and `PoolList.scala`, https://github.com/lichess-org/lila. The two snapshots differ only in the provisional-rating preference; repository snapshots document the rule, not its deployment dates.

cisions. It is the share of the opponent's other grant-or-claim decisions during the panel that ended in a draw, minus the share the positions in which they were made would predict. The score uses every such decision, before and after the exposure, at least five of them. It excludes every decision between the two accounts in either direction of time, so the recipient cannot affect their own instrument; this is the leave-out construction of examiner designs [Dahl et al., 2014, Dobbie et al., 2018, Bhuller et al., 2020]. The score is constructed from the records; the platform does not display it. The identifying comparison is between recipients who draw opponents with different scores, within cells of month, exact time control, and rating, conditional on the recorded pre-exposure histories and other controls.

Three regressions carry the design. For exposure e, let D_e indicate that the recipient received a draw, Y_e that the recipient granted at its first later competitive choice, and Z_e the opponent's score in units of its residual standard deviation. The first stage, the reduced form, and the second stage are

$$D_e = \phi\, Z_e + W_e'\zeta_D + \mu_{c(e)} + \xi_{t(e)} + v_e, \tag{5}$$

$$Y_e = \phi\chi\, Z_e + W_e'\zeta_Y + \mu_{c(e)} + \xi_{t(e)} + w_e, \tag{6}$$

$$Y_e = \chi\, D_e + W_e'\zeta + \mu_{c(e)} + \xi_{t(e)} + \varepsilon_e, \tag{7}$$

where μ absorbs the cells of exposure month, exact time control, and 100-point rating bin within which opponents are assigned, ξ holds exposure-date and hour-of-week effects, and W_e holds the recipient's recorded pre-exposure history and volume, the size of the opponent's history behind the score, and the opponent-environment measures (Appendix D). The second stage instruments D_e with Z_e. With one instrument and the same controls throughout, the reduced-form coefficient is the product of the other two, which is why equation (6) is written with $\phi\chi$. The quantity of interest is χ, the change in the probability of a later grant per draw received. Standard errors allow dependence among exposures that share an account in either role or an exposure date, the level at which the matchmaker's waves assign opponents. Table 6 reports 100ϕ, $100\phi\chi$, and 100χ.

A causal reading rests on four conditions, each with observable implications that are examined here, together with a diagnostic for selection into the sample. Assignment: within cells, the opponent's score must be unrelated to the recipient's other determinants of later granting; the recipient's own recorded behavior before the exposure should then be unrelated to the score (Table 6, Panel B, row 1). Selection: because a game enters the sample only after events that follow assignment, which the opponent's traits could affect, the score should not predict entry (row 4); this is a separate diagnostic, not an implication of random assignment. Relevance: the score must predict receipt (Panel A).

Exclusion: a lenient opponent must affect later granting only through the draw. Because the matchmaker assigns opponents before anyone knows whether a decision will arise, and most games end without one, the design contains assigned pairs in which the judge decides nothing; in court and examiner designs a judge is assigned because a decision is needed. To my knowledge, no leniency design has used such a sample to test exclusion. The first placebo test uses games that ended without a disconnection, where the opponent made no decision. There the score should not predict later granting (row 5). The second placebo test, a regular draw without a disconnect screen, asks a different question, whether recipients respond to the act or to the result (Panel D; Section 6.3).

Monotonicity: a higher score should not make any opponent less likely to grant at a given opportunity; the first stage should then be positive in every subsample and rise across the score distribution (Figure 3; Appendix D.2.1). Under these conditions the estimate is the average effect of receipt for the recipients whose receipt the opponent's leniency decides [Imbens and Angrist, 1994].

The sample contains 9.74 million competitive-position exposures from 1.88 million recipients, with the author's accounts excluded in every role. Exposures involve first-observed pairs whose opponent has at least five other decisions in the panel, the support the score requires. Every eligible exposure is retained: a recipient contributes each of its exposures, 1.15 million recipients contribute more than one, and a recipient's follow-up windows may overlap and share a later choice. The outcome is a grant at the recipient's first later competitive choice against a different opponent within 90 days, with zero assigned to a complete window without such a choice; its mean is 1.72 percent. Keeping those windows avoids selecting the sample on reaching a post-exposure decision, the same choice leniency designs make when an outcome is unrecorded for some of the assigned [Aizer and Doyle, 2015, Dobbie et al., 2018]. The estimates are exposure weighted, and the covariance allows dependence among all exposures of an account; restricting each recipient to exposures at least 90 days apart gives similar estimates (Appendix D).

Controls include the recipient's rating and rating gap; the size and composition of the opponent's decision history, measured by its number of decisions, its number of distinct counterparts, and the grant rate its positions would predict; the recipient's recorded behavior and playing volume during the 90 days before the exposure game's start; and the opponent's disconnection environment, measured by its recent game volume and the shares of its recent games that ended with it or its counterpart disconnected, with an indicator where that history is unavailable (Appendix D.3). The model also includes assignment-cell, exposure-date, and hour-of-week effects. I do not control for the exposure game's final position, clocks, or length because these are determined after opponent assignment. Appendix D gives the construction and the covariance estimator.

Table 6: **Instrumenting receipt with the opponent's history**

	Estimate	SE	Units
Panel A. Primary estimates: 9,736,182 exposures, 1,877,395 recipients			
Least squares: later grant on receipt	0.270	0.028	pp per receipt
First stage: receipt of a draw (100ϕ)	7.253	0.024	pp per score unit
Reduced form: grant at first later competitive choice ($100\phi\chi$)	0.024	0.004	pp per score unit
Two-stage least squares (100χ)	0.325	0.056	pp per receipt
95% interval $[0.214, 0.435]$			
Panel B. Assignment, entry, and exclusion			
Recipient's prior behavior: joint test on eight 90-day traits, history controls omitted, p-value	< 0.001	—	$\chi^2(8) = 39.7$
Recipient's prior grant rate, single-trait coefficient	0.0000	0.0003	trait SD per score SD
Recipient's prior receipts, single-trait coefficient	0.0009	0.0003	trait SD per score SD
Entry into the exposure sample, from games of all endings	-0.013	0.001	pp per score unit
Placebo 1: later granting after a game that ended without a disconnection	-0.001	0.003	pp per score unit
Reduced form minus the ordinary-ending association	0.025	0.005	pp per score unit
Panel C. Alternatives, two-stage least squares (pp per receipt)			
Without the recipient-history controls	0.335	0.060	95% $[0.217, 0.452]$
Each recipient given total weight one	0.293	0.082	95% $[0.131, 0.454]$
Spaced: one exposure per recipient per 90 days, 3,103,280 exposures	0.363	0.096	95% $[0.174, 0.552]$
same rows, with the 90-day histories of Panel A	0.371	0.096	95% $[0.183, 0.559]$
same rows, without the opponent-environment controls	0.372	0.096	95% $[0.183, 0.561]$
same rows, score from decisions before the exposure game began, on its 2,629,196 supported exposures	0.236	0.119	95% $[0.002, 0.470]$
same rows, exposures with a final material difference below four points	0.317	0.099	95% $[0.122, 0.512]$
Panel D. What recipients respond to: 1,290,237 ordinary games			
Placebo 2: later granting after an ordinary draw, without a disconnect screen, rather than an ordinary loss	0.043	0.041	pp, draw minus loss

Notes: The score is the opponent's state-adjusted grant rate over its other decisions in the panel, excluding every decision between the two accounts; first stages and reduced forms are per 0.0498 of the score, the paper's common unit, which is 0.93 of the score's residual standard deviation in the primary sample; balance coefficients are per residual standard deviation. Panel A uses the cells, time effects, and controls of the text; first-stage $F = 91{,}156$. Standard errors allow dependence among exposures sharing an account or a date. Panel B: row 1 is the joint test on the eight 90-day traits with the recipient-history controls omitted (Table F.5); rows 2 and 3 are two of those traits; rows 4 and 5 are the score's association with entry in a frame of 31,836,314 games and with later granting after 21,025,393 ordinary endings, both built under the sample rule of Panel A (Tables D.4 and D.5); row 6 is the reduced form minus row 5 ($p < 0.001$). Panel C: Appendix D describes each row; the last three rows are estimated on the spaced rows; intervals are 95 percent Wald intervals. Panel D: the overlap-weighted comparison of ordinary draws with ordinary losses of Section 6.3, with a 95 percent interval of $[-0.037, 0.122]$ (Table D.8).

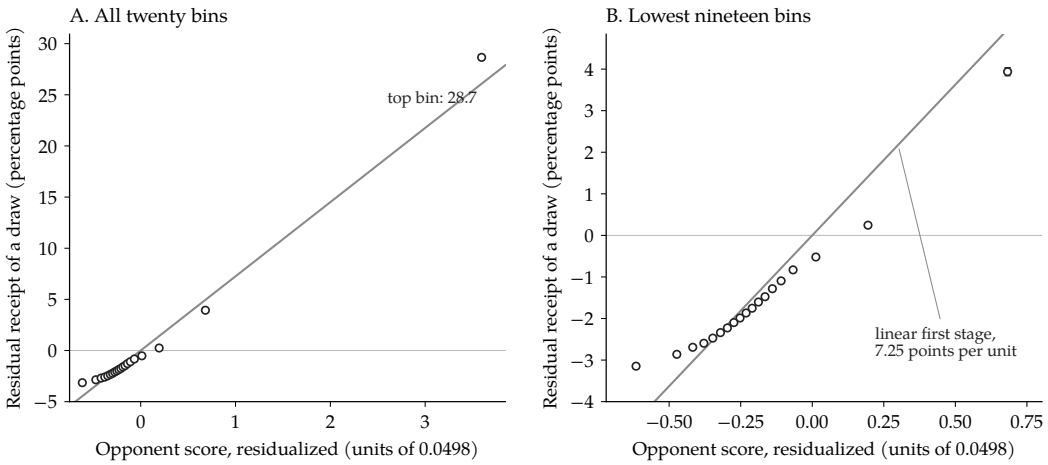
An increase of one unit in the opponent's score raises receipt by 7.25 percentage points, with a first-stage F-statistic of 91,156 (Table 6); that is 100ϕ in equation (5). The unit is 0.0498, five percentage points of the opponent's adjusted grant rate and 0.93 of the score's residual standard deviation in the primary sample; the same unit is used in every frame of the paper. The corresponding reduced-form increase in later granting, $100\phi\chi$, is 0.024 percentage points. Two-stage least squares gives $100\chi = 0.325$ percentage points per receipt (standard error 0.056), with a 95 percent interval of $[0.214, 0.435]$. Least squares on the same rows gives 0.270 (0.028), a sixth below the instrumented estimate, a difference of -0.055 (0.048) under the joint covariance. The estimate is about a fifth of the sample's later-granting rate and below the matched comparison of Section 6.1. Without the recipient-history controls the estimate is 0.335 (0.060), and with each recipient given a total weight of one, spread equally over its exposures, 0.293 (0.082) (Table 6, Panel C).

The primary sample keeps every eligible exposure and has no recipient fixed effects. A spaced design that admits one exposure per recipient per 90 days gives 0.363 (0.096) on its 3.1 million exposures, and across the other ways of building the sample and the score, which change the spacing, add recipient fixed effects, use first exposures only, vary the history window and the minimum number of opponent decisions, or build the score from competitive-position decisions, from earlier or later decisions, or from both at once, the estimate lies between 0.24 and 0.51 (Table 6, Panel C; Appendices D and F.4). The within-recipient estimate on repeated recipients, 0.264 (0.140), is the only one whose interval includes zero. With the pooled and competitive-position scores as two instruments, the overidentification test does not reject (Hansen $J = 0.42$, $p = 0.52$; Appendix F.7), the comparison that Mogstad et al. [2021] recommend before pooling instruments under monotonicity. The instrument has no association with reaching a later choice at all, -0.012 points per score unit (standard error 0.013), so the zero-coded outcome does not move through entry into a later choice.

Receipt rises across twenty equal-count bins of the residualized score (Figure 3): mean residual receipt increases at every step, from -3.1 percentage points in the lowest bin to $+28.7$ in the highest, and most of the slope sits in the top two bins, where opponents with past grants concentrate. Between the lowest and highest bins receipt differs by 32 points, so roughly three exposures in ten sit at a margin the opponent's leniency decides, the counterpart of the complier share that leniency designs compute from their strictest and most lenient judges [Dahl et al., 2014].

On the spaced rows, the first stage is positive in all 30 subgroups of game speed, recipient rating, prior competitive choices, and exposure month, from 6.3 to 8.3 points per score unit with every interval above zero (Appendix D.2.1). A score built only from the opponent's decisions in lost positions, the counterpart of the reverse-sample instru-

Figure 3: **Receipt rises with the opponent's score at every step**



Notes: Twenty equal-count bins of the residualized score, 486,809 or 486,810 exposures each, in the 9,736,182 primary exposures; bars are 95 percent intervals. The line is the linear first stage, 7.25 points per score unit. Panel B omits the top bin.

ment of Bhuller et al. [2020], also predicts receipt in competitive exposures, by 4.74 points per its own standard deviation (0.04; Appendix F.7). Compliers, the weighted margin whose receipt the instrument moves, resemble the sample, with somewhat lower ratings [Table D.6; Abadie, 2003]. On the primary rows, the compliers' later-granting rate is 2.19 percent with a draw received and 1.86 percent without (Appendix D.2.1).

Opponents of differing scores are matched with recipients who differ slightly, and in a known way, in what they have done before. The test is the one leniency designs use: each recipient trait recorded before the exposure is regressed on the residualized score under the model's own adjustment, with the eight 90-day traits tested with those controls omitted and the environment controls retained.[21] The joint test rejects ($\chi^2(8) = 39.7$, $p < 0.001$), and three traits carry it: the recipient's logged disconnections and claims in the prior 90 days each move about -0.0008 trait standard deviations per score standard deviation (standard errors 0.0003), and its logged receipts of a draw $+0.0009$ (0.0003). The traits that predict later granting most strongly, prior granting and its rate (Table 3), are flat, 0.0001 and 0.0000 (0.0003), and the remaining three are within 0.0004 of zero (Appendix F.5).

The source is the matchmaker's conduct preference: players with similar records of abandoning games are paired together, so recipients of higher-score opponents play in

[21] Leniency designs report the same test with the judge's or examiner's leniency as the instrument [Dobbie et al., 2018, Bhuller et al., 2020]; Goldsmith-Pinkham et al. [2026] recommend the trait-as-outcome form used here and reading any imbalance by the bias it implies.

slightly lower-disconnection conditions. The environment controls absorb part of the association and the recipient's own history, among the controls, the rest: adding the eight traits moves the two-stage estimate from 0.335 (0.060) to 0.325 (0.056), three percent, and leaves the first stage at 7.25.

Entry is the other margin. In a separately sampled frame of 31.8 million games of all endings, built under the same sample rule, a score higher by one unit predicts entry into the exposure sample 0.013 points less often (0.001) on an entry rate of 0.46 percent, 2.9 percent of entrants (Table D.4); lenient opponents may wait or add time, so fewer of their counterparts' disconnections end with a button press. For that margin to produce the reduced form, the recipients at it would have to grant about 0.8 points less than the rest, nearly half the base rate, and a bound that trims the affected share from one tail of the outcome [Lee, 2009] spans more than the base rate, since 2.9 percent exceeds the outcome rate of 1.72. In the entry frame the eight-trait test rejects on the same disconnection and claim counts ($\chi^2(8) = 33.8$), while the grant rate moves 0.0001 (0.0002) and whether the recipient granted at all -0.0001 (0.0002) (Appendix D).

The first placebo test checks an implication of exclusion directly: I repeat the reduced form on games that ended in checkmate, resignation, or stalemate, where the opponent made no disconnection decision. If lenient opponents affected later granting through their conduct of the game rather than through the draw, the score would predict later granting here too. It does not. The association is -0.001 percentage points per score unit (standard error 0.003), with a 95 percent interval of $[-0.007, 0.004]$ on a sampled frame of 21.0 million games built under the exposure sample's rule. The reduced form in the exposure sample, 0.024, lies outside that interval, and the upper end of the 90 percent interval, 0.003, is less than a sixth of it. The difference between the reduced form and the placebo is 0.025 (standard error 0.005, $p < 0.001$), with a covariance that lets the two samples share accounts and dates.

The placebo covers conduct in games without a disconnection, including whatever a lenient opponent says in the chat of an ordinary game. Waiting is not a channel to the recipient, who is away while the opponent waits and finds little trace of it afterward; it changes which disconnections end in a decision, which is the entry question above. What it cannot cover is conduct the recipient sees on returning, which the data do not record (Section 7).

The estimate also bundles a response to generosity received with learning that the option exists, and only a recipient who had never faced the choice could learn that from the exposure. Splitting the spaced rows by whether the recipient had faced a chooser-side competitive decision earlier in the panel, the two-stage estimate is 0.45 percentage points (standard error 0.13) on the 2.13 million exposures of recipients who had, a quarter of

their later-granting rate of 1.77 percent, and 0.22 (0.15) on the 975,000 exposures of recipients who had not, a sixth of their rate of 1.27 percent; the difference is 0.23 (0.20; $p = 0.25$) under a covariance shared across the two groups. Splitting instead on a choice within the 90 days before the exposure gives 0.47 (0.15) and 0.27 (0.12). Learning would concentrate the effect among recipients who had never faced the choice; the estimates lean the other way, imprecisely, and the similar proportional effects that Section 3.6 predicts are within their intervals.

Overall, the matchmaker pairs some recipients with more lenient opponents. Those recipients are more likely to receive a draw. For that reason alone, they are more likely to grant one to a third player in the next 90 days. The effect is about a third of a percentage point per draw received, a fifth of the base rate. It is of similar size across the ways of building the sample and the score in Table 6 and Appendix D. The recipients of lenient opponents differ little from other recipients, and not in the traits that best predict later granting. Their counterparts enter the sample slightly less often, but the margin is too small to produce the effect. A lenient opponent does not predict later granting after a game that ended without a disconnection, so the channel is the draw and not the opponent's company. The effect is no larger among recipients facing the choice for the first time, which leans against learning that the option exists.

6.3 What Recipients Respond To

A granted draw differs from a claimed loss in the opponent's act and in the result; two comparisons, mechanism checks rather than identified effects, ask which recipients respond to. The first, the second placebo test, holds the act out, as the intentions experiments do by letting the same outcome arise by chance [Falk et al., 2008, Charness and Levine, 2007]: if the result is what matters, a draw instead of a loss should be followed by more granting even when it comes without a disconnect screen. Among players whose ordinary games ended in a draw rather than a loss, with no disconnection and no decision by the opponent, later granting is 0.04 points higher (standard error 0.04; 95 percent interval -0.04 to 0.12) across 1.29 million games reweighted to the same ratings, time controls, and recent histories (Appendix D.4). The comparison is between different players, and ordinary draws arise in different circumstances from granted ones, so it does not identify the response to the act; what it shows is that a draw rather than a loss, on its own, is not followed by detectably more granting, with a difference above 0.12 points excluded against receipt estimates of 0.3 to 0.4.

The second comparison asks whether the response depends on how much the recipient had earned. Instrumenting receipt separately, on the spaced rows, for recipients

who were slightly behind, roughly equal, ahead, and well ahead when the connection dropped, the receipt effect is 0.44, 0.23, 0.48, and 0.43 percentage points (standard errors 0.16 to 0.23), and the difference between the well-ahead band and the pooled near-equal bands is 0.10 (standard error 0.25) (Table D.8). The interval allows differences of up to half a point in either direction, so the result is no detectable dependence on the recipient's position rather than its absence. Competing can reduce later cooperation, most of all for those who lose [Buser and Dreber, 2016]. That channel runs through the result of the game. Here the result alone does nothing: an ordinary draw in place of a loss is not followed by more granting. And the receipt effect does not depend on whether the recipient was losing when the connection dropped. The channel cannot account for the receipt effect.

For later unmerited granting the estimate is 0.06 percentage points (standard error 0.03), 11 percent of its 0.50 percent base against 19 percent for merited granting. The framework predicts no effect there (Section 3.6), and the data do not distinguish that prediction from an equal proportional rise ($p = 0.27$; Appendix D.5). Receipt does not extend to paying for the platform. It does not detectably change the recipient's Patron badge a median of 1.8 years later: -0.02 percentage points per receipt (95 percent interval -0.06 to 0.02) on a base of 0.16 percent (Appendix D.6).

7 Discussion

7.1 What the Patterns Rule Out

The rating consequences of the two buttons are fixed before play, so any model in which the chooser weighs only rating outcomes, selfish or not, predicts the same choice in every position. A model with one prosocial taste, whether for the rival's rating, for the act of giving [Andreoni, 1990], for equal outcomes [Fehr and Schmidt, 1999], or for seeing oneself as a giver [Bodner and Prelec, 2003, Grossman and van der Weele, 2017], fares no better: in none of them can the board matter, and a single trait that raises both kinds of granting and badge holding implies that accounts granting in both kinds of position hold the badge at least as often as accounts granting only in competitive positions, whatever the shape of the loadings (Appendix E.5); they hold it at 0.49 against 0.90 percent. The data reject each implication.

Two ingredients are needed: a motive that depends on the opponent's chances, which produces the position gradient and confines the gap and the clocks to competitive positions, and a second trait, the weight on rating points, which separates the two kinds of grant and gives the badge its opposite signs. The framework of Section 3 is one such

pair, a member of the two-dimensional family of Bénabou and Tirole [2006] and Cappelen et al. [2007]; guilt aversion [Battigalli and Dufwenberg, 2007], a fairness self-image, or a sportsmanship norm scoped to opponents still in the game would each supply the first ingredient and are readings of b_i rather than rivals. The framework misses one pattern, the rise in granting below the competitive threshold.

Expectations-based loss aversion over the result [Kőszegi and Rabin, 2006], with a constant taste for giving, produces the position, gap, and clock patterns, including that rise, and with heterogeneous tastes for giving and weights on rating points it can match the badge signs as well. With the same investment term, which neither motive supplies, the two readings share every prediction tested here. They differ in one: under loss aversion the position gradient steepens with the points at stake, and under desert it does not (Appendix E.5). Receipt lies outside all of these models: a granted draw adds three to four grants per thousand receipts at the recipient's next competitive choice, and neither placebo moves later granting.

7.2 What the Design Does Not Identify

The denominator. The sample conditions on an active resolution. A chooser who waits for the opponent to return, or who adds time to the absent opponent's clock, never enters it. The grant rate is therefore the share of impatient resolutions that were concessions, not the share of disconnections that were accommodated, and every gradient in the paper is estimated on that margin.

The motive and the reading. Because a claim also marks the absent player's account (Section 2.2), a claim in a lost position can punish a deliberate departure as well as take a deserved win. That can add to the contrast between the two kinds of position, but not to the gradient within competitive positions, which survives the proxies for a deliberate departure (Appendix B.6). The design does not name the motive. Fairness, altruism, sportsmanship, and self-image all remain candidates, and the framework's weight on a deserved result stands for any of them. Nor does it single out the forward reading of the board, as the opponent's chance of avoiding defeat, from readings that share its predictions, such as a judgment of how well the opponent played or an expectations-based loss from forgoing the win [Kőszegi and Rabin, 2006]. What the evidence establishes is the predictions they share, including the two that follow from reading the board through the opponent's chances rather than its surface: the gap, the clocks and the late-game turn matter only where the opponent could still have come back. Nor does the design show that an unmerited concession is not prosocial; it shows that the act carries none of the markers that a merited concession carries, and the framework attributes it to apathy

about the rating points at stake.

What the receipt estimate is. The instrument identifies the effect of being paired with a lenient opponent in a disconnection game, per draw that pairing produced, for the recipients whose receipt the opponent's leniency decided. The placebo rules out a direct effect of lenient opponents' conduct in games without a disconnection. Outside its reach is conduct specific to a disconnection that the recipient sees on returning, chiefly a message left in the game's chat. A message sent with a granted draw is part of the treatment the estimate measures; one sent with a claim would be a direct effect of the opponent's leniency that the placebo cannot detect. The estimate may include learning that the option exists; the split by prior experience of the choice leans against that reading, imprecisely (Section 6.2). A claim also marks the absent player's account (Section 2.2), so a draw spares the recipient a step toward a play ban and, when the recipient was down four or more points of material, a change in whom the matchmaker pairs it with next. Reaching a later competitive choice does not move with the instrument, and on the exposures with a material difference below four points, where a claim leaves the pairing rule untouched, the estimate is 0.317 against 0.363 (Table 6, Panel C), so neither consequence appears to carry the effect; but a concession is more than the rating points it gives up, and the estimate is the effect of the concession as the platform delivers it.

Scope. Both decisions are observed on one platform. The supporter association is stable heterogeneity within Lichess, and the paper does not claim that the disposition it measures carries to other settings. Social image is attenuated, not absent: usernames are visible, the returning opponent sees the result, and self-image is intact.

8 Conclusion

Generosity is usually studied where it is easiest to find: solicited, visible, rewarded, or directed at someone in need. This paper studies it where those supports are muted, removed, or reversed together. The act is anonymous, unsolicited, and unrewarded. It costs the currency the players compete for, it goes against the default, and it goes to a rival who a moment earlier was trying to win. I know of no other setting in which prosocial behavior has been measured at scale under all of these conditions together. In 41.5 million opportunities to claim a win after an opponent disconnects, players grant a draw in 1.37 percent of cases. That is what remains of generosity when little holds it up.

What remains is structured. The same chooser grants more often the stronger the opponent's position, at every step from clearly lost to clearly winning, and less often the longer the game has run. The opponent's rating and remaining time matter where the opponent could still have come back and far less where the position was clearly lost.

These are the two manipulations of entitlement that dictator-game experiments induce with a task set for the purpose, the recipient's claim and the giver's effort; here both are earned in the contest itself. A chooser who asks whether the win is deserved and how much has been put into the game accounts for the pattern. Two of the framework's sharper predictions hold: the rating gap and the clocks matter only where the opponent could still have come back, and the late turn appears in competitive positions only. One does not: granting keeps rising with the opponent's position below the threshold, where the framework says it should be flat.

Granting in each kind of position persists from one year to the next, and each kind predicts itself far better than the other. Merited concessions are associated with double the rate of paying to support the platform; unmerited concessions with a rate 40 percent lower. One disposition to give cannot produce that reversal; two traits can. The value placed on a deserved result and the weight on the platform's own currency are one such pair; a taste for giving and a weight on rating points under an expectations-based reference are another. The result suggests a test for platforms that depend on voluntary support and keep a record of what users do, from Wikipedia edits to open-source commits: whether the record predicts support by the quality of what users do rather than by how much.

The disposition is also passed on. A concession received raises the recipient's later granting toward a third player by about three grants per thousand receipts, a fifth of the base rate. The estimate comes from the leniency of the opponent the matchmaker assigned, measured in that opponent's other games, with a matched comparison in agreement. Two placebo tests say what is passed on, the first of a kind that court-based leniency designs cannot run. Lenient opponents' conduct in games without a disconnection has no measurable effect on later granting, so the channel is the concession and not the company; and a regular draw, without a disconnect screen, is not followed by more granting, which points to the act rather than the result. Generosity received is passed on, between rivals, in the competition's own currency, and from an act that cost its giver a few rating points.

References

Alberto Abadie. Semiparametric instrumental variable estimation of treatment response models. *Journal of Econometrics*, 113(2):231–263, 2003.

Anna Aizer and Joseph J. Doyle. Juvenile incarceration, human capital, and future crime: Evidence from randomly assigned judges. *Quarterly Journal of Economics*, 130(2):759–803, 2015. doi: 10.1093/qje/qjv003.

Ingvild Almås, Alexander W. Cappelen, and Bertil Tungodden. Cutthroat capitalism versus cuddly socialism: Are Americans more meritocratic and efficiency-seeking than Scandinavians? *Journal of Political Economy*, 128(5):1753–1788, 2020. doi: 10.1086/705551.

Ashton Anderson and Etan A. Green. Personal bests as reference points. *Proceedings of the National Academy of Sciences*, 115(8):1772–1776, 2018. doi: 10.1073/pnas.1706530115.

Ashton Anderson, Jon Kleinberg, and Sendhil Mullainathan. Assessing human error against a benchmark of perfection. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016. doi: 10.1145/2939672.2939803.

James Andreoni. Impure altruism and donations to public goods: A theory of warm-glow giving. *Economic Journal*, 100(401):464–477, 1990. doi: 10.2307/2234133.

James Andreoni and B. Douglas Bernheim. Social image and the 50-50 norm: A theoretical and experimental analysis of audience effects. *Econometrica*, 77(5):1607–1636, 2009. doi: 10.3982/ECTA7384.

James Andreoni, Justin M. Rao, and Hannah Trachtman. Avoiding the ask: A field experiment on altruism, empathy, and charitable giving. *Journal of Political Economy*, 125 (3):625–653, 2017. doi: 10.1086/691703.

Dan Ariely, Anat Bracha, and Stephan Meier. Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *American Economic Review*, 99 (1):544–555, 2009. doi: 10.1257/aer.99.1.544.

Ala Avoyan, Robizon Khubulashvili, and Giorgi Mekerishvili. Behavioral market design for online gaming platforms. *Management Science*, 71(6):4552–4565, 2025. doi: 10.1287/mnsc.2021.03628.

Peter Backus, Maria Cubel, Matej Guid, Santiago Sánchez-Pagés, and Enrique López Mañas. Gender, competition, and performance: Evidence from chess players. *Quantitative Economics*, 14(1):349–380, 2023. doi: 10.3982/QE1404.

Mark Bagnoli and Ted Bergstrom. Log-concave probability and its applications. *Economic Theory*, 26(2):445–469, 2005.

Oriana Bandiera, Iwan Barankay, and Imran Rasul. Social preferences and the response to incentives: Evidence from personnel data. *Quarterly Journal of Economics*, 120(3): 917–962, 2005. doi: 10.1093/qje/120.3.917.

Nicholas Bardsley. Dictator game giving: Altruism or artefact? *Experimental Economics*, 11(2):122–133, 2008.

Björn Bartling, Ernst Fehr, and Yagiz Özdemir. Does market interaction erode moral values? *Review of Economics and Statistics*, 105(1):226–235, 2023. doi: 10.1162/rest_a_01021.

Pierpaolo Battigalli and Martin Dufwenberg. Guilt in games. *American Economic Review*, 97(2):170–176, 2007.

Roland Bénabou and Jean Tirole. Intrinsic and extrinsic motivation. *Review of Economic Studies*, 70(3):489–520, 2003. doi: 10.1111/1467-937x.00253.

Roland Bénabou and Jean Tirole. Incentives and prosocial behavior. *American Economic Review*, 96(5):1652–1678, 2006. doi: 10.1257/aer.96.5.1652.

Roland Bénabou and Jean Tirole. Identity, morals, and taboos: Beliefs as assets. *Quarterly Journal of Economics*, 126(2):805–855, 2011. doi: 10.1093/qje/qjr002.

Matthias Benz and Stephan Meier. Do people behave in experiments as in the field? Evidence from donations. *Experimental Economics*, 11(3):268–281, 2008. doi: 10.1007/s10683-007-9192-y.

Manudeep Bhuller, Gordon B. Dahl, Katrine V. Løken, and Magne Mogstad. Incarceration, recidivism, and employment. *Journal of Political Economy*, 128(4):1269–1324, 2020. doi: 10.1086/705330.

Ronit Bodner and Drazen Prelec. Self-signaling and diagnostic utility in everyday decision making. In Isabelle Brocas and Juan D. Carrillo, editors, *The Psychology of Economic Decisions, Volume 1: Rationality and Well-Being*, pages 105–124. Oxford University Press, 2003. doi: 10.1093/oso/9780199251063.003.0006.

Adrian Bruhin, Ernst Fehr, and Daniel Schunk. The many faces of human sociality: Uncovering the distribution and stability of social preferences. *Journal of the European Economic Association*, 17(4):1025–1069, 2019. doi: 10.1093/jeea/jvy018.

Adrian Bruhin, Simon Haenni, Lingqing Jiang, Adrian Roethlisberger, Regula Buchli, Beat M. Frey, and Lorenz Goette. Understanding mechanisms of persistence in prosocial behavior: Evidence from a large-scale field experiment. *Journal of Political Economy Microeconomics*, 4(1):1–37, 2026. doi: 10.1086/735782.

Christoph Bühren, Björn Frank, Stefan Krabel, and Alexander Werner. Decision-making in competitive framings—strategic behavior of chess players in mini-ultimatum game chess puzzles. *Economics Letters*, 115(3):356–358, 2012. doi: 10.1016/j.econlet.2011.12.080.

Leonardo Bursztyn and Robert Jensen. Social image and economic behavior in the field: Identifying, understanding, and shaping social pressure. *Annual Review of Economics*, 9(1):131–153, 2017. doi: 10.1146/annurev-economics-063016-103625.

Thomas Buser and Anna Dreber. The flipside of comparative payment schemes. *Management Science*, 62(9):2626–2638, 2016. doi: 10.1287/mnsc.2015.2257.

A. Colin Cameron, Jonah B. Gelbach, and Douglas L. Miller. Robust inference with multiway clustering. *Journal of Business & Economic Statistics*, 29(2):238–249, 2011.

Alexander W. Cappelen, Astri Drange Hole, Erik Ø. Sørensen, and Bertil Tungodden. The pluralism of fairness ideals: An experimental approach. *American Economic Review*, 97 (3):818–827, 2007. doi: 10.1257/aer.97.3.818.

Alexander W. Cappelen, James Konow, Erik Ø. Sørensen, and Bertil Tungodden. Just luck: An experimental study of risk-taking and fairness. *American Economic Review*, 103(4):1398–1413, 2013. doi: 10.1257/aer.103.4.1398.

Fredrik Carlsson, Haoran He, and Peter Martinsson. Easy come, easy go: The role of windfall money in lab and field experiments. *Experimental Economics*, 16(2):190–207, 2013.

Jeffrey Carpenter, Peter Hans Matthews, and John Schirm. Tournaments and office politics: Evidence from a real effort experiment. *American Economic Review*, 100(1):504–517, 2010. doi: 10.1257/aer.100.1.504.

Bharat Chandar, Uri Gneezy, John A. List, and Ian Muir. The drivers of social preferences: Evidence from a nationwide tipping field experiment. NBER Working Paper 26380, National Bureau of Economic Research, 2019. URL https://www.nber.org/papers/w26380.

Gary Charness and David I. Levine. Intention and stochastic outcomes: An experimental study. *Economic Journal*, 117(522):1051–1072, 2007.

Alexander Chaudhry, Yang Wang, and Erya Ouyang. Community gifting and recipients' engagement with peers and influencers: Evidence from natural experiments on Twitch. *Journal of Interactive Marketing*, 60(1):83–99, 2025.

Todd L. Cherry, Peter Frykblom, and Jason F. Shogren. Hardnose the dictator. *American Economic Review*, 92(4):1218–1221, 2002.

Eric Chyn, Brigham Frandsen, and Emily C. Leslie. Examiner and judge designs in economics: A practitioner's guide. *Journal of Economic Literature*, 63(2):401–439, 2025.

Alain Cohn, Michel André Maréchal, David Tannenbaum, and Christian Lukas Zünd. Civic honesty around the globe. *Science*, 365(6448):70–73, 2019. doi: 10.1126/science.aau8712.

Gordon B. Dahl, Andreas Ravndal Kostøl, and Magne Mogstad. Family welfare cultures. *Quarterly Journal of Economics*, 129(4):1711–1752, 2014. doi: 10.1093/qje/qju019.

Jason Dana, Daylian M. Cain, and Robyn M. Dawes. What you don't know won't hurt me: Costly (but quiet) exit in dictator games. *Organizational Behavior and Human Decision Processes*, 100(2):193–201, 2006. doi: 10.1016/j.obhdp.2005.10.001.

Jason Dana, Roberto A. Weber, and Jason Xi Kuang. Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness. *Economic Theory*, 33(1): 67–80, 2007. doi: 10.1007/s00199-006-0153-z.

Stefano DellaVigna, John A. List, and Ulrike Malmendier. Testing for altruism and social pressure in charitable giving. *Quarterly Journal of Economics*, 127(1):1–56, 2012. doi: 10.1093/qje/qjr050.

Will Dobbie, Jacob Goldin, and Crystal S. Yang. The effects of pretrial detention on conviction, future crime, and employment: Evidence from randomly assigned judges. *American Economic Review*, 108(2):201–240, 2018. doi: 10.1257/aer.20161503.

Robert Drago and Gerald T. Garvey. Incentives for helping on the job: Theory and evidence. *Journal of Labor Economics*, 16(1):1–25, 1998. doi: 10.1086/209880.

Martin Dufwenberg, Uri Gneezy, Werner Güth, and Eric van Damme. Direct versus indirect reciprocity: An experiment. *Homo Oeconomicus*, 18(1/2):19–30, 2001.

Mark Duggan and Steven D. Levitt. Winning isn't everything: Corruption in sumo wrestling. *American Economic Review*, 92(5):1594–1605, 2002. doi: 10.1257/000282802762024665.

Nadja Dwenger, Henrik Kleven, Imran Rasul, and Johannes Rincke. Extrinsic and intrinsic motivations for tax compliance: Evidence from a field experiment in germany. *American Economic Journal: Economic Policy*, 8(3):203–232, 2016.

Christine L. Exley. Excusing selfishness in charitable giving: The role of risk. *Review of Economic Studies*, 83(2):587–628, 2016. doi: 10.1093/restud/rdv051.

Armin Falk. Gift exchange in the field. *Econometrica*, 75(5):1501–1511, 2007.

Armin Falk and Urs Fischbacher. A theory of reciprocity. *Games and Economic Behavior*, 54(2):293–315, 2006.

Armin Falk and Nora Szech. Morals and markets. *Science*, 340(6133):707–711, 2013. doi: 10.1126/science.1231566.

Armin Falk, Ernst Fehr, and Urs Fischbacher. Testing theories of fairness—intentions matter. *Games and Economic Behavior*, 62(1):287–303, 2008.

Ernst Fehr and Gary Charness. Social preferences: Fundamental characteristics and economic consequences. *Journal of Economic Literature*, 63(2):440–514, 2025. doi: 10.1257/jel.20241391.

Ernst Fehr and Klaus M. Schmidt. A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114(3):817–868, 1999.

Christina M. Fong. Evidence from an experiment on charity to welfare recipients: Reciprocity, altruism and the empathic responsiveness hypothesis. *Economic Journal*, 117 (522):1008–1024, 2007. doi: 10.1111/j.1468-0297.2007.02076.x.

James H. Fowler and Nicholas A. Christakis. Cooperative behavior cascades in human social networks. *Proceedings of the National Academy of Sciences*, 107(12):5334–5338, 2010. doi: 10.1073/pnas.0913149107.

Brigham Frandsen, Lars Lefgren, and Emily Leslie. Judging judge fixed effects. *American Economic Review*, 113(1):253–277, 2023. doi: 10.1257/aer.20201860.

Axel Franzen and Sonja Pointner. The external validity of giving in the dictator game: A field experiment using the misdirected letter technique. *Experimental Economics*, 16(2): 155–169, 2013.

Mark E. Glickman. Dynamic paired comparison models with stochastic variances. *Journal of Applied Statistics*, 28(6):673–689, 2001.

Ayelet Gneezy, Alex Imas, Amber Brown, Leif D. Nelson, and Michael I. Norton. Paying to be nice: Consistency and costly prosocial behavior. *Management Science*, 58(1):179–187, 2012. doi: 10.1287/mnsc.1110.1437.

Paul Goldsmith-Pinkham, Peter Hull, and Michal Kolesár. Leniency designs: An operator's manual. *Journal of Economic Perspectives*, 40(3):213–240, 2026. doi: 10.1257/jep.20251480.

Kurt Gray, Adrian F. Ward, and Michael I. Norton. Paying it forward: Generalized reciprocity and the limits of generosity. *Journal of Experimental Psychology: General*, 143(1): 247–254, 2014. doi: 10.1037/a0031047.

Zachary Grossman and Joël J. van der Weele. Self-image and willful ignorance in social decisions. *Journal of the European Economic Association*, 15(1):173–217, 2017. doi: 10.1093/jeea/jvw001.

Kareem Haggag and Giovanni Paci. Default tips. *American Economic Journal: Applied Economics*, 6(3):1–19, 2014.

Christine Harbring and Bernd Irlenbusch. Sabotage in tournaments: Evidence from a laboratory experiment. *Management Science*, 57(4):611–627, 2011. doi: 10.1287/mnsc.1100.1296.

Joseph Henrich, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, and Richard McElreath. In search of homo economicus: Behavioral experiments in 15 small-scale societies. *American Economic Review*, 91(2):73–78, 2001.

Joseph Henrich, Jean Ensminger, Richard McElreath, Abigail Barr, Clark Barrett, Alexander Bolyanatz, Juan Camilo Cardenas, Michael Gurven, Edwins Gwako, Natalie Henrich, Carolyn Lesorogol, Frank Marlowe, David Tracer, and John Ziker. Markets, religion, community size, and the evolution of fairness and punishment. *Science*, 327 (5972):1480–1484, 2010.

Elizabeth Hoffman and Matthew L. Spitzer. Entitlements, rights, and fairness: An experimental examination of subjects' concepts of distributive justice. *Journal of Legal Studies*, 14(2):259–297, 1985.

Elizabeth Hoffman, Kevin McCabe, Keith Shachat, and Vernon Smith. Preferences, property rights, and anonymity in bargaining games. *Games and Economic Behavior*, 7(3): 346–380, 1994.

Christoph Huber, Anna Dreber, Jürgen Huber, Magnus Johannesson, Michael Kirchler, Utz Weitzel, et al. Competition and moral behavior: A meta-analysis of forty-five crowd-sourced experimental designs. *Proceedings of the National Academy of Sciences*, 120(23):e2215572120, 2023.

Guido W. Imbens and Joshua D. Angrist. Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475, 1994.

Dean S. Karlan. Using experimental economics to measure social capital and predict financial decisions. *American Economic Review*, 95(5):1688–1699, 2005. doi: 10.1257/000282805775014407.

Maria Kavussanu, Alistair R. Seal, and Daniel R. Phillips. Observed prosocial and antisocial behaviors in male soccer teams: Age differences across adolescence and the role of motivational variables. *Journal of Applied Sport Psychology*, 18(4):326–344, 2006. doi: 10.1080/10413200600944108.

Ji Eun Kim, Seura Ha, Sangmi Kim, and Libby Hemphill. The spread of virtual gifting in live streaming: The case of Twitch. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 20, pages 1243–1255, 2026.

James Konow. Fair shares: Accountability and cognitive dissonance in allocation decisions. *American Economic Review*, 90(4):1072–1091, 2000. doi: 10.1257/aer.90.4.1072.

Botond Kőszegi and Matthew Rabin. A model of reference-dependent preferences. *Quarterly Journal of Economics*, 121(4):1133–1165, 2006.

Erin L. Krupka and Roberto A. Weber. Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association*, 11(3):495–524, 2013. doi: 10.1111/jeea.12006.

Steffen Künn, Juan Palacios, and Nico Pestel. Indoor air quality and strategic decision making. *Management Science*, 69(9):5354–5377, 2023.

Edward P. Lazear, Ulrike Malmendier, and Roberto A. Weber. Sorting in experiments with application to social preferences. *American Economic Journal: Applied Economics*, 4 (1):136–163, 2012. doi: 10.1257/app.4.1.136.

David S. Lee. Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies*, 76(3):1071–1102, 2009. doi: 10.1111/j.1467-937X.2009.00536.x.

Steven D. Levitt and John A. List. What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives*, 21(2):153–174, 2007. doi: 10.1257/jep.21.2.153.

Steven D. Levitt, John A. List, and Sally E. Sadoff. Checkmate: Exploring backward induction among chess players. *American Economic Review*, 101(2):975–990, 2011. doi: 10.1257/aer.101.2.975.

Fan Li, Kari Lock Morgan, and Alan M. Zaslavsky. Balancing covariates via propensity score weighting. *Journal of the American Statistical Association*, 113(521):390–400, 2018.

John A. List. On the interpretation of giving in dictator games. *Journal of Political Economy*, 115(3):482–493, 2007.

Nicole Maestas, Kathleen J. Mullen, and Alexander Strand. Does disability insurance receipt discourage work? using examiner assignment to estimate causal effects of SSDI receipt. *American Economic Review*, 103(5):1797–1829, 2013. doi: 10.1257/aer.103.5.1797.

Magne Mogstad, Alexander Torgovitsky, and Christopher R. Walters. The causal interpretation of two-stage least squares with multiple instrumental variables. *American Economic Review*, 111(11):3663–3698, 2021.

Jacob Morrier, Amine Mahmassani, and R. Michael Alvarez. Uncovering the viral nature of toxicity in competitive online video games. *arXiv preprint arXiv:2410.00978*, 2024.

Charles C. Moul and John V. C. Nye. Did the Soviets collude? A statistical analysis of championship chess 1940–1978. *Journal of Economic Behavior & Organization*, 70(1–2): 10–21, 2009. doi: 10.1016/j.jebo.2009.01.009.

Redzo Mujcic and Andreas Leibbrandt. Indirect reciprocity and prosocial behaviour: Evidence from a natural field experiment. *The Economic Journal*, 128(611):1683–1699, 2018. doi: 10.1111/ecoj.12474.

Martin A. Nowak and Sébastien Roch. Upstream reciprocity and the evolution of gratitude. *Proceedings of the Royal Society B: Biological Sciences*, 274(1610):605–610, 2007. doi: 10.1098/rspb.2006.0125.

Martin A. Nowak and Karl Sigmund. Evolution of indirect reciprocity. *Nature*, 437(7063): 1291–1298, 2005. doi: 10.1038/nature04131.

Robert J. Oxoby and John Spraggon. Mine and yours: Property rights in dictator games. *Journal of Economic Behavior & Organization*, 65(3–4):703–713, 2008.

Ignacio Palacios-Huerta and Oscar Volij. Field centipedes. *American Economic Review*, 99 (4):1619–1635, 2009.

Alexander Peysakhovich, Martin A. Nowak, and David G. Rand. Humans display a 'cooperative phenotype' that is domain general and temporally stable. *Nature Communications*, 5:4939, 2014.

Devin G. Pope and Maurice E. Schweitzer. Is Tiger Woods loss averse? Persistent bias in the face of experience, competition, and high stakes. *American Economic Review*, 101(1): 129–157, 2011.

Hariharan Purohit and Nisheeth Srivastava. A metacognitive appraisal of quitting in chess. *Cognitive Science*, 50:e70193, 2026.

Bradley J. Ruffle. More is better, but fair is fair: Tipping in dictator and ultimatum games. *Games and Economic Behavior*, 23(2):247–265, 1998.

Yuval Salant and Jörg L. Spenkuch. Complexity and satisficing: Theory with evidence from chess. *Review of Economic Studies*, 93(2):1296–1322, 2026. doi: 10.1093/restud/rdaf041.

Amos Schurr and Ilana Ritov. Winning a competition predicts dishonest behavior. *Proceedings of the National Academy of Sciences*, 113(7):1754–1759, 2016.

Luca Stanca. Measuring indirect reciprocity: Whose back do we scratch? *Journal of Economic Psychology*, 30(2):190–202, 2009. doi: 10.1016/j.joep.2008.07.010.

Jan Stoop. From the lab to the field: Envelopes, dictators and manners. *Experimental Economics*, 17(2):304–313, 2014.

Anthony Strittmatter, Uwe Sunde, and Dainis Zegners. Life cycle patterns of cognitive performance over the long run. *Proceedings of the National Academy of Sciences*, 117(44), 2020. doi: 10.1073/pnas.2006653117.

Johannes Wachs, Leonore Röseler, Tobias Gesche, Elliott Ash, and Anikó Hannák. Good question! The effect of positive feedback on contributions to online public goods. *arXiv preprint arXiv:2604.10360*, 2026.

Yuan Yuan, Tracy Xiao Liu, Chenhao Tan, Qian Chen, Alex Pentland, and Jie Tang. Gift contagion in online groups: Evidence from virtual red packets. *Management Science*, 70 (7):4465–4479, 2024.

A Sample Construction and Composition

This appendix documents the construction of the analysis sample described in Section 2: the recorded statuses and rules that define an opportunity and the grant indicator, the source and observed-adjacency restrictions, the composition of the resulting sample, and the engine-evaluation and game-stage conventions used throughout.

A.1 Opportunities, choosers, and the grant indicator

The platform records how each game ended, and two recorded statuses matter here. A game ended by the connected player's action after an opponent's disconnection carries the status `timeout`; a game ended because a player's clock ran out carries the status `outoftime`. Only the first reveals a choice between claiming the win and calling a draw, and the analysis sample is built from games with that status. The analysis sample covers bullet, blitz, rapid, and classical games. Ultrabullet is absent because the sample is restricted to automated pool pairings, whose available time controls do not include ultrabullet. The sample spans the 24 complete UTC calendar months from November 2023 through October 2025, a period in which Lichess hosted 2,231,251,938 standard rated games.[22]

The extraction applies a clock restriction, retaining a game with either status when the last recorded clock reading of the side to move in the final position is at least five seconds or is missing. An `outoftime` status records any flag fall, including ordinary time losses with both players connected, so it identifies neither a disconnection nor a game in which the claim screen was visible; the records therefore do not identify waiting. Of the 47,587,020 active resolutions, 262 lack a finite final clock reading for the disconnected player; all 262 are non-pool games and leave with the source restriction, so every opportunity in the analysis sample has an observed reading of at least five seconds. The threshold is mine rather than the platform's, and it does not guarantee a margin between the chooser's action and any automatic resolution.

The threshold excludes almost no active resolutions. From the pool games that ended with a clock-related status and a last recorded clock under five seconds, excluding immediate rematches and the author's accounts, I drew a random sample of 150,000 games in each of the 24 months, 3,600,000 in all. Of these, 65 carried the `timeout` status and the other 3,599,935 were ordinary flag falls. Expanding by each month's sampling fraction gives an estimated 7,893 active resolutions excluded by the restriction over the panel,

[22] Sum of the official monthly standard-rated game counts published at https://database.lichess.org/.

0.019 percent of the 41,534,289 opportunities, with an exact 95 percent interval, simultaneous across the 24 months, of 629 to 38,390, at most 0.092 percent. The restriction's purpose is practical: the games it removes are about three quarters of those with a clock-related status, none of which would then need their metadata retrieved.

Within a retained game, the chooser is the connected player, identified operationally as the final recorded mover, and the disconnected player is the side to move in the final recorded position. The grant indicator equals one when the opportunity ends in a draw and the chooser passes the material rule, and zero otherwise. The rule accepts any pawn, rook, or queen; at least two bishops; a bishop and a knight; or at least two knights, and it excludes a bare king and a king with only one bishop or knight. A draw recorded with excluded material is therefore not dropped: it remains an opportunity with an indicator of zero. The rule is my own and does not reproduce the platform's position-specific adjudication of insufficient material. A recorded win for the chooser identifies a claim.

The author's two accounts are excluded from every analysis: from the analysis sample, where they held 42 opportunities, nine of them in competitive positions, and one competitive-position grant; from every role in the instrumental-variables analysis of Section 6.2, including its exposures, the opponents' decision histories, the state reference, and later choices; and from the behavioral histories and cohorts built from the sample. Outside the instrumental-variables design, where they were never included, adding them back changes no reported estimate by more than 0.03 standard errors.

A.2 Draws that fail the material rule

The full-panel inventory identifies 20,902 active-resolution draws in which the chooser fails the material rule, all in the competitive region; the flag is the conjunction of a recorded draw and failure of the chooser material rule, and the inventory applies no competitive-position restriction of its own. Dropping these rows leaves 41,513,429 opportunities and 15,186,521 competitive opportunities with 419,462 grants, a competitive rate of 2.762 percent against 0.470 percent in clearly-worse positions: a raw difference of 2.292 percentage points and a ratio of 5.875, against 2.288 and 5.867 in Table 1. This sensitivity excludes the flagged draws; the all-source precursor also contains 75 chooser-loss outcomes, too few to move any rate reported here.

A.3 Source and observed-adjacency restrictions

The platform also records how each game was created, and the analysis sample retains only games created through the automated matchmaking pool. The other creation sources observed among active resolutions are arena tournaments, the lobby, direct chal-

lenges, and games created through the API; these are excluded. The pool label records that the matchmaker created the game and does not establish that both players entered the pool through the same control.

Among pool games, the observed-adjacency restriction looks at each player's immediately preceding observed rated game. That chronology crosses the recorded speed categories and is ordered by recorded game-start timestamp, with archive order and game identifier as tie-breakers. A focal game is excluded when both players' immediately preceding observed rated games are the same game against one another, and a game is retained only with a determinate finding that this is not the case. Two further groups are therefore excluded rather than retained: games for which tied game-start timestamps prevent a determinate ordering around the focal game or its predecessor, and games for which an observed predecessor is unavailable for at least one player.

The rule concerns only the immediately preceding game. It does not exclude a game because a later game was played against the same opponent, it imposes no elapsed-time cutoff, and it does not eliminate all repeat encounters between accounts. It is also distinct from the first-observed-meeting classification of Appendix B.11: a game can pass the adjacency restriction even if the two accounts had met earlier in the recorded history.

Table A.1 reports the resulting flow. Game or opportunity counts are reported at each stage; the grant count and rate are reported for the analysis sample only.

A.4 Composition

Table A.2 reports the composition of the analysis sample by twelve-month period, game speed, and the chooser's displayed rating. The grant rate is 1.37 percent in the first twelve-month period and 1.38 percent in the second, and the competitive-position rate is 2.76 percent in both. Across speeds the rate rises from 1.19 percent in bullet games to 2.06 percent in classical games, and across chooser ratings it falls from 1.55 percent below 1600 to 0.87 percent at 2400 and above. These rows describe the composition of the sample and are not estimates of effects.

The 41,534,289 opportunities are faced by 3,224,171 distinct chooser accounts, of which 2,442,574 face at least one competitive-position opportunity. Across the sample, 276,481 accounts grant at least one granted draw and 212,859 do so in a competitive position.

A.5 Engine evaluation and game stage

Every final recorded position is evaluated with Stockfish 18 at a fixed search budget of 100,000 nodes, from the disconnected player's perspective, so that positive values favor

Table A.1: Sample construction

Stage	Games or opportunities
Standard rated games, November 2023 through October 2025	2,231,251,938
Games with `timeout` or `outoftime` status passing the clock restriction	170,648,691
Clock expirations (`outoftime`, any flag fall), excluded	123,061,671
Active resolutions (`timeout`), all creation sources	47,587,020
Excluded: non-pool sources (arena, lobby, direct challenge, API)	3,448,799
Excluded: pool games with observed adjacency	2,247,688
Excluded: ambiguous ordering	582
Excluded: missing previous-game status	355,620
Excluded: the author's two accounts	42
Analysis sample	41,534,289
Granted draws	570,090
Grant rate (%)	1.373

Notes: Games enter with either status when the last recorded clock of the side to move is at least five seconds or missing; every opportunity in the analysis sample has a reading of at least five seconds. An `outoftime` status records any flag fall and reveals neither a disconnection nor a claim screen. The five exclusion rows partition the active resolutions outside the analysis sample: "ambiguous ordering" means tied game-start timestamps, and "missing previous-game status" an unavailable predecessor for at least one player.

Table A.2: Composition of the analysis sample

	Opportunities	Granted draws	Rate (%)
Panel A. Twelve-month period			
November 2023–October 2024, all positions	20,491,614	279,737	1.365
competitive positions	7,494,367	206,695	2.758
November 2024–October 2025, all positions	21,042,717	290,354	1.380
competitive positions	7,713,047	212,767	2.759
Panel B. Game speed			
Bullet	5,295,012	62,916	1.188
Blitz	23,435,814	308,671	1.317
Rapid	12,361,435	189,384	1.532
Classical	442,028	9,119	2.063
Panel C. Chooser's displayed rating			
Below 1600	26,791,515	414,362	1.547
1600–1999	11,779,288	128,438	1.090
2000–2399	2,787,338	25,766	0.924
2400 and above	176,148	1,524	0.865

Notes: Granted draws per hundred opportunities. Each panel partitions the 41,534,289 opportunities; the competitive rows in Panel A partition the 15,207,414 competitive-position opportunities.

the disconnected player; the same requested budget is used for every position. The processing retains the uncapped evaluation and a version capped at ± 600 centipawns, and positions with a forced mate receive a signed sentinel value of $\pm 10{,}000$ before capping.

The five descriptive bins used in Section 4 are defined on integer centipawn values with inclusive boundaries: clearly better for the disconnected player, at least $+300$; better, $+101$ to $+299$; roughly equal, -100 to $+100$; modestly worse, -299 to -101; and clearly worse, at most -300. The competitive region comprises evaluations of at least -100 and the clearly-worse region evaluations of at most -300. Particular analyses use their own intervals, which differ from these bins at some boundaries: the regressions of Appendix C.2 use the evaluation bands stated with their results, and the exposure cells of Section 6.1 use the bands stated in Appendix D.

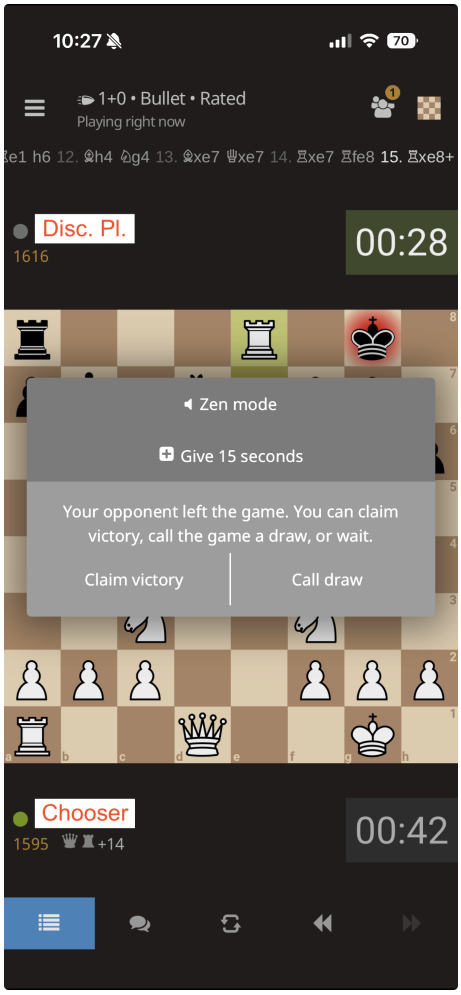
Game stage is the number of plies recorded before the final position; a ply is a half-move, so ply 2 means one move by each player. The main text groups competitive-position opportunities into games ending at ply 2, between plies 3 and 10, and after ply 10, and Figure 2 crosses the five bins with the bands 0–4, 5–6, 7–8, 9–10, 11–20, 21–40, 41–80, and 81 or more plies; Table A.3 gives the opportunities in each of its forty cells. Individual regressions use the stage controls stated with their results.

Table A.3: Opportunities and granted draws by board state and ply band

Ply band	Clearly better	Better	Roughly equal	Modestly worse	Clearly worse
0–4	1,414	50,406	1,523,968	22,568	6,497
5–6	5,798	59,569	646,040	42,278	39,412
7–8	14,052	82,891	577,368	73,697	92,017
9–10	23,788	103,224	528,443	103,604	140,226
11–20	310,243	693,745	2,199,463	793,546	1,572,256
21–40	1,236,444	1,312,130	2,529,694	1,683,925	6,708,010
41–80	1,304,996	595,250	1,101,457	898,799	11,115,015
81 or more	146,512	22,254	138,265	43,671	2,991,354
All plies	3,043,247	2,919,469	9,244,698	3,662,088	22,664,787
Granted draws, all plies	108,837	69,801	240,824	44,072	106,556
Rate, all plies (%)	3.576	2.391	2.605	1.203	0.470

Notes: Denominators of the rates in Figure 2; the forty cells partition the 41,534,289 opportunities. The smallest cell, clearly better by ply 4, holds 1,414 opportunities and 78 grants; every other holds at least 5,798. The last two rows give each bin's granted draws and grant rate over all plies.

Figure A.1: The disconnection decision as presented on the mobile interface, in a screenshot from one of the author's own games. Account names are replaced with role labels.



B Board Position: Specifications and Additional Comparisons

This appendix reports the specifications behind Section 4: the comparison of specifications for Table 2 and the account-weighted versions of its central contrast; the continuous position profile, its sensitivity models, and the material comparison; the game-stage counts and the stage profile by time format; the proxies for a deliberate departure; the countdown threshold; the tenure and arena comparisons; and the repeat-opponent, re-encounter, and repayment comparisons.

B.1 Specifications behind Table 2

Table B.1 re-estimates equation (4) without the control vector x_g and with a single gap function h in place of h_C and h_L, and sets the focal contrasts beside those of Table 2. Dropping the controls moves the position and gap contrasts by at most 0.04 points and the stage contrasts by at most 0.06; the late changes move most, by about 0.06 in both regions, and the clocks carry that. Pooling the gap replaces the two region profiles with their opportunity-weighted average, 0.39 points for $+100$ against -100, and hides the region difference; the stage profiles are unchanged. The table's position rows are contrasts with clearly worse at the references of the two region interactions, plies 0–4 and equal ratings, where the region difference is largest and the lost-region reference cell is thin. Evaluated at plies 21–40 and equal ratings, the same contrasts under Table 2's specification are 0.55, 1.36, 1.66, and 2.95 points (standard errors 0.006 to 0.016), and at plies 41–80 they are 0.55, 1.35, 1.64, and 2.94; the standardized steps of Table 2, Panel A, cumulate to 3.00.

Table B.2 lets the two clock terms of x_g differ by region, everything else as in Table 2. Pooled, the opponent's clock carries 0.165 points per log second (0.005) and the chooser's -0.207 (0.006). By region, both are an order of magnitude larger in competitive positions, 1.373 (0.015) and -1.212 (0.014), and small with the opposite sign in lost positions, -0.140 (0.005) and 0.153 (0.005).

B.2 Weighting accounts equally

The competitive-minus-clearly-worse comparison of Section 4.2 weights opportunities, so accounts with more opportunities contribute more to it. Table B.3 reports the comparison under three weighting rules in the binary sample of 1,982,343 accounts observed in both competitive and clearly-worse positions, of which 1,757,863 grant in neither. For account i, write F_i for its competitive opportunities, L_i for its clearly-worse opportunities, $N_i = F_i + L_i$, and d_i for its competitive-minus-clearly-worse difference in grant rates;

Table B.1: Specifications behind Table 2

	Table 2	Without x_g	Single h
Position, contrast with clearly worse at plies 0–4 and equal ratings			
Modestly worse	0.552 (0.006)	0.559 (0.006)	0.547 (0.006)
Roughly equal	2.117 (0.087)	2.132 (0.087)	2.044 (0.086)
Better	2.412 (0.087)	2.436 (0.087)	2.344 (0.087)
Clearly better	3.704 (0.088)	3.743 (0.088)	3.640 (0.087)
Stage			
Plies 41–80 minus 0–4, competitive	−2.331 (0.019)	−2.278 (0.019)	−2.324 (0.019)
Plies 41–80 minus 0–4, not competitive	−1.562 (0.085)	−1.505 (0.085)	−1.566 (0.085)
Plies 81 or more minus 41–80, competitive	0.370 (0.029)	0.433 (0.029)	0.370 (0.029)
Plies 81 or more minus 41–80, not competitive	−0.070 (0.005)	−0.015 (0.005)	−0.070 (0.005)
Rating gap, +100 *versus* −100			
Competitive	1.018 (0.018)	1.025 (0.018)	0.392 (0.008)
Not competitive	0.037 (0.007)	0.050 (0.007)	0.392 (0.008)

Notes: All 41,534,289 opportunities; percentage points, standard errors clustered by chooser. The first column repeats Table 2. The second drops x_g. The third keeps the region-specific stage profiles and a single piecewise-linear gap function, so its two gap rows coincide.

Table B.2: The clocks by region

Percentage points per log second	Competitive	Not competitive	Difference
Opponent's clock	1.373 (0.015)	−0.140 (0.005)	1.513 (0.017)
Chooser's clock	−1.212 (0.014)	0.153 (0.005)	−1.365 (0.015)

Notes: Equation (4) with the two log-clock terms of x_g interacted with the region indicator; all 41,534,289 opportunities; standard errors clustered by chooser.

these counts exclude the modestly-worse region.

Each estimator is a chooser-fixed-effects regression on this sample and is a normalized weighted average of the d_i, so the three differ only in the weight each account's difference receives. Weighting every opportunity equally places weight $F_i L_i / N_i$ on d_i and gives a gap of 2.174 percentage points; this estimate is a within-account regression and is distinct from the raw pooled rate difference reported for the full main sample. Weighting each opportunity by $1/N_i$ places weight $F_i L_i / N_i^2$ on d_i and gives 2.831, which is not an equal-account average. Weighting each opportunity by $N_i/(F_i L_i)$ places equal weight on every account's d_i, making the estimator the ordinary mean of the account differences, and gives 2.820. The equal-account gap exceeds the opportunity-weighted gap by 0.646 points, with uncertainty from the two estimators' joint chooser-level covariance.

Table B.3: The competitive-minus-clearly-worse gap with equal account weights

	Gap (pp)	SE	95% interval
Opportunities weighted equally (weight $F_i L_i / N_i$ on d_i)	2.174	0.009	[2.156, 2.193]
Inverse-count weights (weight $F_i L_i / N_i^2$ on d_i)	2.831	0.010	[2.811, 2.851]
Equal weight per account (mean of d_i)	2.820	0.010	[2.800, 2.840]
Equal-account minus opportunity-weighted	0.646	0.009	[0.628, 0.664]

Notes: The binary sample contains the competitive- and clearly-worse-position opportunities of the 1,982,343 accounts observed in both states, of which 1,757,863 grant in neither. F_i, L_i, $N_i = F_i + L_i$, and d_i are an account's competitive opportunities, clearly-worse opportunities, their sum, and its competitive-minus-clearly-worse grant-rate difference. The table reports chooser-fixed-effects estimators that weight each opportunity by 1, $1/N_i$, and $N_i/(F_i L_i)$, respectively, with the implied weight on d_i shown in the row label; standard errors are clustered by chooser account, and the last row's uncertainty uses the joint chooser-level covariance.

B.3 The continuous profile within competitive positions

This subsection reports the estimates behind Section 4.3 and the sensitivity models around it. All models use the analysis sample of Section 2.3. The outcome is the grant indicator in percentage points, every model absorbs chooser fixed effects and fixed effects for stage band, speed category, and month, and standard errors are clustered by chooser account with the $G/(G-1)$ correction; accounts with one opportunity and accounts that never grant remain in the sample, the former contributing no within-account variation. The primary sample is the 11,590,453 competitive opportunities (evaluation of at least -100 centipawns from the rival's side) that arose after ply ten, from 2,183,097 accounts, of which 2.469 percent were granted. Adjusted rates are opportunity-weighted

predictions standardized to each model sample's own composition; their intervals let the composition vary when chooser clusters are resampled.

The primary specification is

$$100G_{ig} = \alpha_i + f(v_g) + \phi_{b(g)} + \psi_{c(g)} + \omega_{m(g)} + u_{ig}, \tag{8}$$

where G_{ig} indicates a grant by chooser i in game g, v_g is the rival's engine evaluation, and $b(g)$, $c(g)$, and $m(g)$ index stage band, speed category, and month. The function f is a natural cubic spline in the raw centipawn score with boundaries at -100 and $+600$ and interior knots at 0, $+100$, and $+300$; finite scores of $+600$ and above and forced mates in the rival's favor receive separate indicators rather than being fitted as large centipawn values. The curve is unconstrained: positive contrasts do not by themselves establish monotonicity everywhere.

Table B.4 reports the contrasts across specifications and the eight-bin profile. The contrasts are stable to the clock controls and the rating-environment controls of equation (4), the two log clocks and the piecewise-linear gap with the log premium and log rating deviations and their squares, to an alternative knot placement at -50, $+100$, $+250$, and $+450$, and to dropping games beyond ply eighty. Pooling all game stages lowers the first contrast to $+0.20$ points; an interaction model shows that the step from equality to $+100$ is absent within the first ten plies ($+0.03$ points, standard error 0.026) while the two larger steps remain ($+0.54$ and $+0.91$), and a joint test rejects equality of the contrasts across the two stage groups ($\chi^2(3) = 67.5$). The same split appears in the two-bin comparison of roughly equal with better positions, with chooser, month, and speed fixed effects and the premium and clock controls: the difference is $+0.219$ points (standard error 0.034) in games ending by ply 10 and -0.372 (0.010) in games ending after it.

The profile has the same shape within each chooser rating tier below 2400: the three contrasts are $+0.25$, $+0.47$, and $+0.98$ below 1600, $+0.21$, $+0.60$, and $+1.18$ from 1600 to 1999, and $+0.27$, $+0.80$, and $+1.01$ from 2000 to 2399, with standard errors between 0.014 and 0.091; the tier at 2400 and above holds too few opportunities for precise estimates. Extending the spline over the full range of positions, from -600 to $+600$ with additional knots at -300 and -100, gives adjusted rates of 0.51 percent at -600, 1.00 at -200, 1.43 at -100, 1.94 at equality, and 4.02 just below $+600$; the two regions join smoothly.

Material provides a measure of advantage that a player can recognize without an engine. I count the disconnected opponent's material lead in pawn equivalents, valuing a knight or bishop at three, a rook at five, and a queen at nine. Within competitive positions after ply ten, the adjusted grant rate is 2.22 percent at equal material, 3.00 percent when the opponent is two pawns ahead, and 4.61 percent when ahead by five or more. When

material and engine evaluation enter together, both retain a relationship with granting. The highest material category receives 0.61 percentage points more than equal material at a given evaluation. The evaluation difference from +300 to +500 remains 0.87 points at a given material category.

Table B.4: The continuous profile within competitive positions: contrasts and the eight-bin profile

Panel A: Contrasts in the adjusted grant rate (percentage points)

Specification	$f(100) - f(0)$	$f(300) - f(100)$	$f(500) - f(300)$	Opportunities
Primary: competitive, after ply 10	0.241 (0.010)	0.529 (0.014)	1.021 (0.018)	11,590,453
plus clock controls	0.227 (0.010)	0.509 (0.014)	0.998 (0.018)	11,590,453
plus rating controls	0.222 (0.010)	0.507 (0.014)	0.995 (0.018)	11,590,453
alternative knots	0.268 (0.011)	0.495 (0.015)	1.046 (0.019)	11,590,453
plies 11 to 80 only	0.244 (0.011)	0.526 (0.014)	1.019 (0.018)	11,283,422
holding material fixed	0.216 (0.011)	0.498 (0.015)	0.866 (0.019)	11,590,453
Competitive, all stages	0.202 (0.010)	0.548 (0.014)	1.026 (0.018)	15,207,414
All positions, after ply ten	0.141 (0.010)	0.430 (0.014)	0.976 (0.019)	37,397,029

Panel B: Eight-bin profile, competitive positions after ply ten

Evaluation bin (centipawns)	Adjusted rate (%)	Step from the bin below (pp)
$[-100, -50)$	1.548 (0.013)	—
$[-50, 0)$	1.772 (0.012)	0.224 (0.015)
$[0, 50)$	2.033 (0.012)	0.261 (0.014)
$[50, 100]$	2.105 (0.014)	0.072 (0.015)
$(100, 200]$	2.264 (0.013)	0.159 (0.015)
$(200, 300)$	2.509 (0.016)	0.245 (0.016)
$[300, 600)$	3.334 (0.015)	0.825 (0.017)
≥ 600 or forced mate	4.813 (0.026)	1.479 (0.024)

Notes: Standard errors clustered by chooser; every model has chooser, stage-band, speed, and month fixed effects. The primary sample has 2,183,097 chooser accounts. Panel A: natural cubic spline in the evaluation with boundaries at −100 and +600 and knots at 0, +100, and +300 (alternative: −50, +100, +250, +450); scores of +600 and above and forced mates have separate indicators; the all-positions model adds knots at −300 and −100 and lower-tail indicators. Clock and rating-environment controls are those of equation (4). Panel B: bins as indicators, the last pooling +600 and above with mates; adjusted rates standardize to the sample. All contrasts and steps are significant.

B.4 Game stage

Table B.5 reports grant rates by the stage at which the game ended, in plies, where a ply is a half-move. Panel A covers competitive positions. Its ply-2, plies 3–10, and after-ply-10 rows are mutually exclusive and exhaust the region, while the by-ply-4 and by-ply-10 rows are cumulative: of the 15,207,414 competitive-position opportunities, 1,575,788 (10.36 percent) end by ply 4 and 3,616,961 (23.78 percent) end by ply 10.

Panel B partitions the roughly equal bin into the eight ply bands of Figure 2. The rate falls from 4.210 percent in the 0–4 band to 1.374 percent in the 41–80 band, and then

turns up to 1.882 percent in the 81-or-more band, which contains 138,265 opportunities. The profile across the eight bands is therefore not monotonic; the figure shows the same bands for the other four position bins, and Appendix A gives the opportunity count behind each of its forty cells.

Table B.5: Granted draws by game stage

	Opportunities	Granted draws	Rate (%)
Panel A. Competitive positions by stage			
Ply 2	677,538	30,952	4.568
Plies 3–10	2,939,426	102,306	3.480
By ply 10 (cumulative)	3,616,961	133,257	3.684
After ply 10	11,590,453	286,205	2.469
By ply 4 (cumulative)	1,575,788	66,224	4.203
Panel B. Roughly equal positions by ply band			
Plies 0–4	1,523,968	64,156	4.210
Plies 5–6	646,040	22,793	3.528
Plies 7–8	577,368	18,627	3.226
Plies 9–10	528,443	15,735	2.978
Plies 11–20	2,199,463	55,793	2.537
Plies 21–40	2,529,694	45,989	1.818
Plies 41–80	1,101,457	15,129	1.374
Plies 81 or more	138,265	2,602	1.882

Notes: A ply is a half-move. In Panel A, the ply 2 and plies 3–10 rows partition the by-ply-10 row, and the by-ply-10 and after-ply-10 rows partition the 15,207,414 competitive-position opportunities; the by-ply-4 row is cumulative. The ply 2 and plies 3–10 rows predate the exclusion of the author's accounts, which removes three opportunities and one grant from their total; the other rows are computed after it. Panel B partitions the 9,244,698 roughly-equal opportunities.

B.5 The stage profile by time format

Table B.6 estimates equation (4) with a single stage profile and a single gap function in place of the region-specific ones, the seven stage bands interacted with the speed category, and the position bins, the gap, and the controls retained; it reports each format's profile relative to its own 0–4 band. The pool-paired analysis sample contains no ultrabullet games, so four formats are estimated. The bottom rows give the decline from plies 0–4 to plies 41–80 in each format and its difference from the next faster format, the comparison read below against the three measures of investment in Section 3.4, together with the raw 0–4 grant rate, the opportunity count, and the share of each format's opportunities in games of 81 or more plies.

Section 3.4 says where to look for how a chooser counts the investment: the three

measures differ in how the decline per band should vary across time formats. The decline from plies 0–4 to plies 41–80 is 2.00 points in bullet, 2.37 in blitz, 2.09 in rapid, and 1.85 in classical games. Blitz is 0.36 points steeper than bullet (standard error 0.05). Rapid is 0.27 points shallower than blitz (0.04), and classical 0.25 shallower again (0.15), the last difference imprecise.

A count in moves alone predicts the same decline in every format; it is rejected. A count in time spent alone predicts a steeper decline as the format slows; it is rejected too. A count in the share of a typical game completed predicts a steeper decline where typical games are shorter, a comparison that needs the length of a typical game in each format, which the analysis sample does not supply. Within a format, pooling all positions, the rate does not turn up in the longest games. The late turn of Table 2, Panel B, belongs to competitive positions, which are fewer than one in ten of the opportunities in games past eighty plies (Table A.3).

Table B.6: The stage profile by time format

	Bullet	Blitz	Rapid	Classical
Panel A. Stage coefficients, relative to the format's plies 0–4 (percentage points)				
Plies 5–6	−0.918 (0.052)	−0.819 (0.032)	−0.730 (0.044)	−0.683 (0.194)
Plies 7–8	−1.218 (0.049)	−1.194 (0.030)	−1.052 (0.042)	−0.574 (0.201)
Plies 9–10	−1.347 (0.048)	−1.451 (0.029)	−1.279 (0.041)	−1.076 (0.184)
Plies 11–20	−1.668 (0.037)	−1.905 (0.023)	−1.666 (0.032)	−1.207 (0.148)
Plies 21–40	−1.949 (0.038)	−2.326 (0.024)	−2.016 (0.033)	−1.631 (0.146)
Plies 41–80	−2.005 (0.040)	−2.366 (0.024)	−2.094 (0.033)	−1.847 (0.148)
Plies 81 or more	−2.240 (0.047)	−2.382 (0.026)	−2.098 (0.035)	−1.890 (0.154)
Panel B. Contrasts (percentage points)				
Plies 81 or more minus plies 41–80	−0.235 (0.023)	−0.016 (0.007)	−0.003 (0.009)	−0.044 (0.058)
Decline, plies 0–4 to 41–80	2.005 (0.040)	2.366 (0.024)	2.094 (0.033)	1.847 (0.148)
minus the next faster format's	—	0.362 (0.046)	−0.272 (0.041)	−0.248 (0.151)
Panel C. Composition				
Raw grant rate, plies 0–4 (%)	3.84	4.24	4.29	4.25
Share at 81 or more plies (%)	2.9	8.1	10.2	8.9
Opportunities	5,295,012	23,435,814	12,361,435	442,028

Notes: Equation (4) with a single stage profile and a single gap function, the stage bands interacted with speed, on all 41,534,289 opportunities; standard errors clustered by chooser, between-format differences from the model's covariance. Panel B's decline is the negative of the 41–80 coefficient; the last row subtracts the next faster format's decline.

B.6 Proxies for a deliberate departure

The position gradient could partly reflect how choosers interpret a disconnection. A player who disappears after losing a queen may look as though they have quit. A chooser may then claim because the departure seems deliberate. A departure that looks deliberate

also supplies a justification for keeping the win, in the way that excuses reduce giving elsewhere [Dana et al., 2007, Exley, 2016]. Comparing competitive and lost positions alone cannot distinguish this response from a response to performance.

Could this account explain the rise within competitive positions? Only if the chance of an intentional departure, or the chooser's belief in it, falls as an already-winning opponent pulls further ahead. Frustration at a losing position cannot explain the 1.02-point rise between $+300$ and $+500$ centipawns. The path to the final position could still matter. An opponent can blunder away part of an advantage and remain ahead, and the preceding moves allow that to be examined.

The analysis, summarized in Section 4.3, uses a uniform deterministic sample of 20,000 chooser accounts from the analysis sample, with their complete opportunity histories: 260,863 opportunities and 3,755 granted draws. For each opportunity the game record was replayed and the positions before the final one were evaluated with the engine and search budget of Appendix A, using the last exact score of each search; board-state membership uses the final-position evaluation of the full sample. Re-evaluating the final positions of this sample the same way changes the evaluation by 23 centipawns on average and would move 1,775 of 234,811 positions, 0.76 percent, into or out of the competitive region.

The cost of the rival's own last move is the evaluation before that move minus the evaluation after it, from the rival's side, clipped at $\pm$1,000 centipawns and entered per 100 centipawns; a blunder indicator marks a cost of at least 100, 200, or 300 centipawns. Moves into or out of a forced mate enter as separate categories rather than as centipawn values. The two- and four-ply swings are the change in the evaluation from the rival's side over the last two and four plies, and they share the final evaluation mechanically, so the last move is the more direct reading. The clock proxies are $\log(1 + \text{seconds})$ on the rival's last recorded clock, the ratio of the rival's to the chooser's seconds, and an indicator that the rival had less than a tenth of the base time; they are last recorded readings, not departure times. The material swing is the change in the rival's material lead, in pawn equivalents, over the last four plies.

Every model absorbs chooser fixed effects and fixed effects for stage band, speed category, and month, and standard errors are clustered by chooser account. Competitive-region models add eight categories of the final evaluation, and clearly-worse-region models an unrestricted spline in it. The one-proxy models enter one proxy at a time; the profile and region models with all proxies enter the seven continuous and indicator proxies together. The asymmetry model pools the two regions after ply ten, enters all seven proxies together, interacts each with the competitive region, and lets the evaluation function and the month, speed, and stage effects differ by region; its interaction coefficients estimate

the difference in slopes directly and are not differences of the one-proxy estimates. In this sample the baseline steps are $+0.40$, $+0.34$, and $+1.41$ points, against $+0.24$, $+0.53$, and $+1.02$ in the full sample, differences within sampling error; Panel C therefore compares specifications on identical rows.

Recent blunders matter. An opponent whose last move lost at least 200 centipawns receives a draw 0.55 percentage points less often than an opponent without such a blunder, with a standard error of 0.19 (Table B.7, Panel B). The corresponding difference for a loss of at least 300 centipawns is 0.83 points. A mistake may make quitting seem more plausible, or it may weaken the chooser's judgment of what the opponent deserves. On either reading the history leading to a position matters alongside the position. In clearly lost positions, however, the last move's evaluation cost has little association with granting: -0.002 percentage points per 100 centipawns (standard error 0.010), compared with -0.181 (0.046) in competitive positions. This fits a judgment about the opponent's performance. It also fits a blame account in which a further blunder in an already-lost position says nothing about intent. The comparison does not separate them.

These measures do little to explain the continuous gradient. In the sampled opportunities, the adjusted increase from $+300$ to $+500$ centipawns is 1.41 percentage points before the departure proxies enter and 1.37 points afterward. It remains 1.24 points when the sample is restricted to opponents whose last move cost less than 100 centipawns and who had more than a tenth of the base time remaining. The lower contrasts are similarly stable.

B.7 The countdown threshold

At the material threshold where the platform shortens the countdown, the grant rate does not step. The platform divides the wait before the buttons appear by three when the disconnected player is down at least four points of material (Section 2.1). If the shorter wait changed which disconnections in lost positions reach the chooser in a way that moved the grant rate, the rate would step between a deficit of three and a deficit of four. Figure B.1 traces the adjusted grant rate across deficits of zero to eight within lost positions, by speed, before and after the platform revised the countdown in July 2025; Table B.8 reports the contrasts.

Design. The sample is the analysis sample's opportunities with a finite, non-mate evaluation below -100 centipawns from the disconnected player's side and a material deficit of zero to eight at the final position: 13,234,511 opportunities before July 19, 2025, and 1,669,732 from August 11, 2025, with the intervening weeks excluded. The model is linear in saturated deficit indicators, with chooser fixed effects, fixed effects for the

Table B.7: Proxies for a deliberate departure, after ply ten

Panel A. Coverage

	Opportunities	Evaluated	Grant rate (%)
Competitive positions	72,575	72,573	2.563
Clearly-worse positions	141,029	140,815	0.535

Panel B. Proxy coefficients (percentage points)

	Competitive	Clearly worse	Difference, joint
Cost of the rival's last move, per 100 centipawns	−0.181 (0.046)	−0.002 (0.010)	−0.332 (0.058)
Last move cost at least 100 centipawns	−0.474 (0.133)	−0.034 (0.044)	—
Last move cost at least 200 centipawns	−0.547 (0.189)	−0.035 (0.046)	—
Last move cost at least 300 centipawns	−0.826 (0.245)	−0.036 (0.051)	—
Evaluation change over two plies, per 100 centipawns	0.093 (0.038)	−0.006 (0.010)	0.198 (0.061)
Evaluation change over four plies, per 100 centipawns	0.044 (0.033)	−0.010 (0.009)	−0.011 (0.049)
$\log(1 + \text{rival's remaining seconds})$	0.363 (0.198)	−0.021 (0.041)	0.699 (0.216)
Rival's seconds relative to the chooser's	0.001 (0.002)	0.036 (0.019)	−0.001 (0.015)
Rival below a tenth of the base time	−0.062 (1.423)	−0.193 (0.154)	1.162 (1.284)
Rival's material gain over four plies, per pawn	0.105 (0.030)	−0.002 (0.005)	0.082 (0.032)

Panel C. Steps in the competitive profile, same rows (percentage points)

	$f(100) - f(0)$	$f(300) - f(100)$	$f(500) - f(300)$
Baseline	0.404 (0.141)	0.341 (0.189)	1.409 (0.259)
All seven proxies held fixed	0.420 (0.140)	0.425 (0.193)	1.367 (0.260)
plus clock and rating controls	0.400 (0.140)	0.414 (0.193)	1.343 (0.259)
Last move cost below 100 and ample clock	0.419 (0.158)	0.449 (0.214)	1.236 (0.276)

Panel D. Competitive minus clearly worse, same rows (percentage points)

	Baseline	All proxies	Opportunities
Pooled after ply ten	1.901 (0.117)	1.750 (0.120)	213,388

Notes: Standard errors clustered by chooser. Sample: 20,000 chooser accounts drawn uniformly, opportunities after ply ten; evaluated opportunities have a score or mate for every requested position. Panel B: each row in the first two columns is a separate one-proxy model on 72,575 competitive and 141,029 clearly-worse opportunities (blunder rows: 71,389 and 113,620 with finite endpoints); the third column is the proxy's competitive-region interaction in a model with all seven proxies on 213,388 opportunities. Panel C: the spline of Appendix B.3 on 72,573 evaluated competitive opportunities; the last row keeps the 58,406 with a last-move cost below 100 centipawns and more than a tenth of the base time left. Panel D: the region contrast with and without the proxies.

10-centipawn evaluation bin by stage band by period, month fixed effects, the logs of both players' last recorded clocks, the rating gap, and the chooser's rating, all by speed, with the numeric slopes also by period. Adjusted profiles standardize each deficit to the retained sample's other characteristics within speed and period. The threshold contrasts are the step from three to four; the difference between the rate at four and a straight line through zero to three; the symmetric contrast $0.5\mu_2 - 1.5\mu_3 + 1.5\mu_4 - 0.5\mu_5$; and the same local contrasts centered on deficits that carry no rule, as placebos. Standard errors are clustered by chooser.

The profile. In blitz before the revision the adjusted rate is 0.68 percent at a deficit of zero, 0.55 at one, and between 0.51 and 0.53 at every deficit from two to eight. The step from three to four is -0.019 percentage points (standard error 0.009); the steps from four to five, five to six, six to seven, and seven to eight are $+0.005$, $+0.006$, -0.010, and -0.004, and the placebo contrasts at six, seven, and eight are 0.001, -0.016, and $+0.007$ (standard errors 0.016 to 0.019). The straight-line contrast is $+0.06$ (0.01) pooled across speeds, because the profile falls from zero to one and flattens by two, so a line through zero to three undershoots four; the symmetric contrast is -0.03 (0.01) pooled, and in blitz the quadratic version is -0.12 (0.02). Allowing the untreated profile a discrete second difference of up to 0.025 points per deficit squared widens the pooled interval for the straight-line contrast to $[-0.02, 0.14]$. Against a within-lost evaluation gradient of 0.55 points from clearly to modestly worse positions (Table 2), a step of a few hundredths of a point at the rule's own threshold cannot account for it.

The revision. The countdown's base waits changed in July 2025. The pooled symmetric contrast is -0.03 (0.01) before and -0.05 (0.03) after, a change of -0.02 (0.04); the straight-line contrast changes by -0.01 (0.03). Restricting to opportunities with at least 40, 60, 120, or 300 seconds on the last recorded clock by speed, or twice those, and moving the calendar windows to August 1 or September 1, leaves every pooled contrast within 0.02 points of the primary. Monthly contrasts before July 2025 are jointly equal in bullet, blitz, and classical (p between 0.22 and 0.88) and not in rapid ($p = 0.005$).

What the profile does show. Positions lost on the engine but level in material are granted about 0.15 points more often than positions a pawn or more down, in every speed before the revision and in all but the thinly populated classical format after it. The chooser reads the visible board.

These are adjusted descriptive contrasts. They do not separate the countdown's effect on who reaches the chooser from any effect on the chooser's choice or on the opponent's reconnection, and material is counted at the final position rather than at the moment of disconnection.

Figure B.1: **Grant rate by material deficit within lost positions.** Adjusted grant rate in percent by the disconnected player's material deficit at the final position, by speed, before (left) and after (right) the platform's July 2025 revision of the countdown. The dashed line marks the deficit of four at which the countdown is divided by three. Bars are 95 percent intervals with standard errors clustered by chooser; the label above each point is the number of retained opportunities.

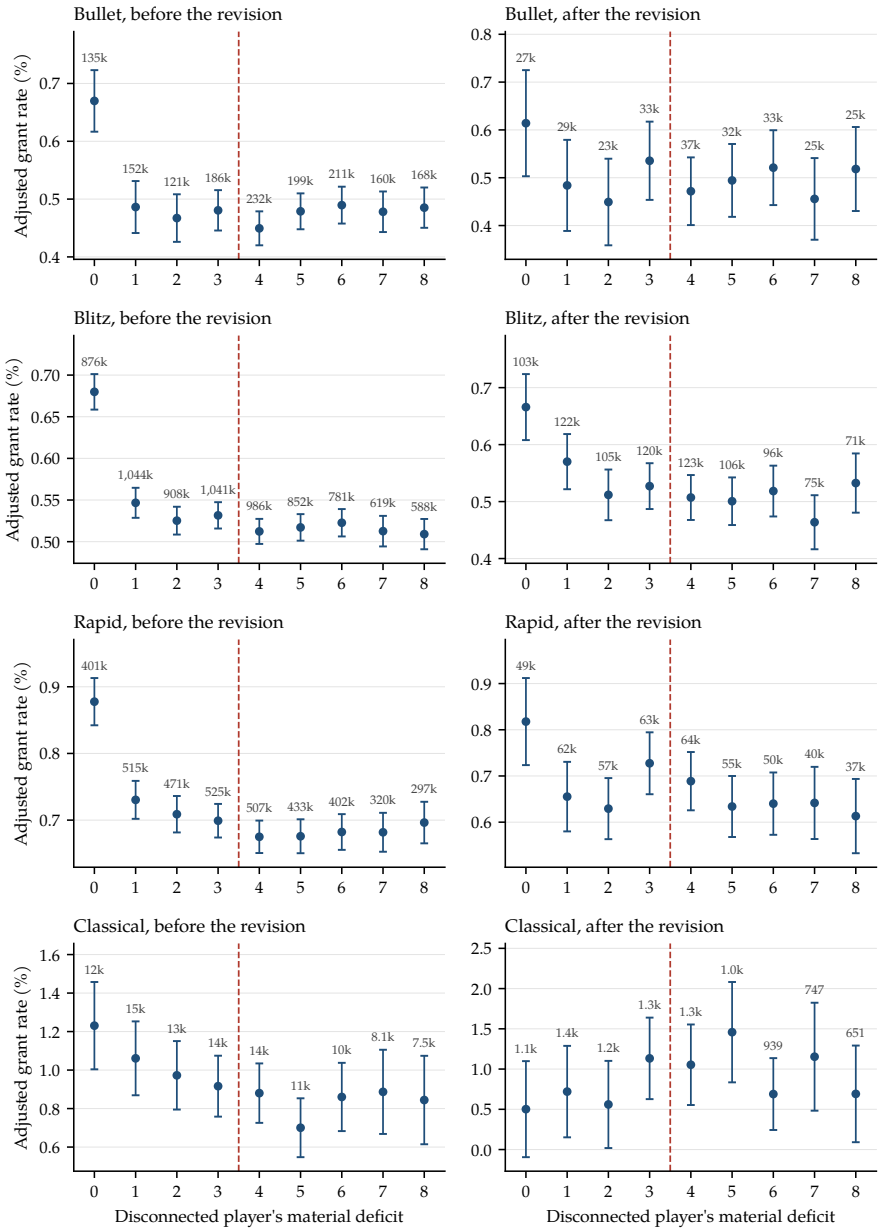


Table B.8: **Contrasts at the countdown threshold**

	Before	After	Change
Panel A. Pooled across speeds, fixed speed weights (pp)			
Rate at deficit 4 minus a line through deficits 0–3	0.060 (0.009)	0.051 (0.025)	−0.010 (0.027)
Symmetric contrast, deficits 2–5	−0.026 (0.011)	−0.051 (0.034)	−0.025 (0.035)
Panel B. Blitz, before the revision, 7,694,827 opportunities (pp)			
Step from deficit 3 to 4	−0.019 (0.009)		
Rate at 4 minus a line through 2 and 3	−0.026 (0.016)		
Rate at 4 minus a quadratic through 0–3	−0.116 (0.021)		
Step from deficit 0 to 1	−0.133 (0.013)		
Step from deficit 1 to 2	−0.021 (0.011)		
Step from deficit 2 to 3	0.006 (0.010)		
Step from deficit 4 to 5	0.005 (0.009)		
Step from deficit 5 to 6	0.006 (0.010)		
Step from deficit 6 to 7	−0.010 (0.010)		
Step from deficit 7 to 8	−0.004 (0.011)		
Placebo local contrast at deficit 6	0.001 (0.016)		
Placebo local contrast at deficit 7	−0.016 (0.017)		
Placebo local contrast at deficit 8	0.007 (0.019)		

Notes: Linear probability models in percentage points on lost-position opportunities (finite non-mate evaluation below −100 centipawns) with material deficits 0–8, before July 19, 2025 and from August 11, 2025; chooser, evaluation-bin-by-stage-by-period, and month fixed effects, clock, gap, and rating controls, all by speed. Pooled rows weight speeds by their share of opportunities (blitz 0.58, rapid 0.29, bullet 0.12, classical 0.01). Standard errors clustered by chooser in parentheses. A placebo local contrast at deficit d is the rate at d minus a line through $d-2$ and $d-1$.

B.8 Account tenure

Account creation dates from the profile query give each chooser's age at the time of each opportunity. The tenure comparisons use a labeled sample of 100,000 chooser accounts rather than the full analysis sample.

Within that sample, the competitive-minus-clearly-worse gradient, adjusted with chooser fixed effects and controls for the stage and rating environment of the game, lies between +1.93 and +2.01 percentage points across tenure bands, and is +2.013 (95 percent interval 1.891 to 2.136) among accounts more than a year old at the opportunity. Raw grant rates fall with tenure in both board states, so the raw gap is not constant across bands while the adjusted gap is; the novice surplus is not a uniform additive component, and the defensible description is that established accounts grant less often overall while retaining the same selectivity by board state. Among year-plus accounts, the grant rate in competitive positions is 1.63 points lower in games ending between plies 11 and 40 than in games ending by ply ten. Census bounds computed on the full analysis sample place at most 9.6 percent of grants in an account's first 30 days and at most 35.6 percent in its first year.

B.9 Arena tournament games

Arena tournament games are excluded from the analysis sample by the source restriction of Section 2.3. A companion sample compares them with pool games for the choosers observed in both formats: 6,078,928 opportunities from 301,489 accounts, at all game stages, with both formats passing the observed-adjacency restriction. The raw grant rate is 0.84 percent in arena games and 1.15 percent in pool games. With chooser and time-control-by-month fixed effects and the recorded rating, board-stage, clock, rating-deviation, and win-premium controls, the arena grant rate is 1.06 percentage points lower than the pool rate in competitive positions (standard error 0.03) and 0.18 points lower in clearly-worse positions (0.02), so the competitive-minus-clearly-worse gradient is 0.88 points flatter in arena games (standard error 0.03; 95 percent interval 0.82 to 0.94). Standard errors are clustered by chooser.

Granting therefore depends on the context of the game, and the shortfall in arena games falls mostly on rivals who still had a claim. The comparison holds the chooser fixed, not the game: it does not identify which feature of arena play (pace, berserk, tournament standing, or norms) produces the difference, or an effect of competitive intensity alone.

B.10 Repeat opponents and repayment

The position gradient does not reflect goodwill cultivated toward opponents who may be met again. The raw grant rate is lower at repeat meetings, 1.00 against 1.40 percent at first recorded meetings. Within competitive positions, the adjusted repeat-minus-first difference in a one-in-five sample of account pairs is -0.017 percentage points (standard error 0.030). The competitive-minus-clearly-worse contrast is 0.46 points flatter at repeat meetings than at first meetings, 1.36 against 1.82. Selectivity is lower, not higher, toward opponents met before. It also survives where another meeting is unlikely. Among the 252,423 opportunities with a predicted 90-day re-encounter probability below 0.9 percent, where the observed re-encounter rate is 0.68 percent, the adjusted competitive-minus-clearly-worse difference is 1.94 points (0.17). Direct repayment is rare. After 85,478 initial grants followed for six months, the original recipient has an opportunity to reciprocate in 39 cases, none taken. The rating points available through them are 0.08 percent of the points given up. The history covers rated games on Lichess only (Appendix B.12).

B.11 First and repeat observed meetings

A first observed meeting is the first recorded rated game between the unordered pair of accounts in the available game history, which begins in 2013 and covers ultrabullet, bullet, blitz, rapid, and classical games, with games ordered by timestamp, archive order, and game identifier. A repeat meeting has at least one earlier recorded game between the two accounts. The history predates the analysis window, but it covers only rated games on this platform, so a first observed meeting establishes neither first contact between the players nor that they were strangers.

Table B.9 describes opportunities in the full analysis sample: repeat meetings are 2,934,481 of the 41,534,289 opportunities, or 7.07 percent, and the grant rate is lower at repeat meetings overall and in both comparison regions, 1.000 against 1.401 percent overall, 2.025 against 2.817 percent in competitive positions, and 0.297 against 0.483 percent in clearly-worse positions.

Table B.9: First and repeat observed meetings

	Opportunities	Granted draws	Rate (%)
First observed meeting, all positions	38,599,808	540,735	1.401
Repeat meeting, all positions	2,934,481	29,355	1.000
First observed meeting, competitive positions	14,078,262	396,595	2.817
Repeat meeting, competitive positions	1,129,152	22,867	2.025
First observed meeting, clearly-worse positions	21,119,015	101,967	0.483
Repeat meeting, clearly-worse positions	1,545,772	4,589	0.297

Notes: The table partitions the analysis sample by whether the two accounts had an earlier recorded rated game in the history beginning in 2013; the competitive and clearly-worse rows exclude the modestly-worse region.

B.12 Predicted re-encounters and observed repayment capacity

The comparisons in Appendix B.10 use a fixed one-in-five sample of unordered account pairs, 8,300,865 focal opportunities and 8,279,179 pairs, selected by a rule that uses neither later meetings nor granting. Table B.10 reports the estimates. The repeat-minus-first comparison within competitive positions uses 3,038,973 opportunities from 1,258,805 chooser accounts. A joint model with first-meeting and repeat-meeting terms gives a competitive-minus-clearly-worse contrast of 1.822 points (standard error 0.014) at first meetings and a repeat-minus-first change of -0.464 (0.033); it is a different specification from the competitive-only repeat premium, and the two are not combined.

The predicted re-encounter analysis uses information available before the focal game to predict any recorded rated real-time meeting of the pair within 90 days. Training uses

games before May 2024 in pair folds 0 and 1, calibration August to October 2024 in fold 2, and evaluation November 2024 onward in folds 3 and 4, with complete 90-day follow-up and no pair in more than one fold; the grant outcome is not a predictor. The evaluation sample has 1,281,253 opportunities. The low-risk group is the 252,423 with a predicted probability below 0.887 percent; their mean predicted rate is 0.53 percent and their observed rate 0.68 percent. The prediction is an out-of-sample forecast of observed meetings under existing behavior, not a measure of beliefs. The low-risk contrast excludes the modestly-worse region, which is why its denominator is below 252,423; the raw low-risk rates are 3.565 percent in competitive positions (2,956 of 82,911) and 0.747 percent in clearly-worse positions (1,107 of 148,100).

Among 85,478 initial grants with complete six-month follow-up, a later rated meeting of the pair is observed after 3.321 percent, and an opportunity for the original recipient to reciprocate, the original chooser disconnecting against the original recipient, after 0.0456 percent: 39 opportunities, none ending in a reciprocal grant. If every observed repayment opportunity had been used, the modeled rating benefit would sum to 770 points against 1,019,578 points given up at the original grants, a ratio of 0.0755 percent. Later meetings come from the rated-game chronology and repayment opportunities from the all-source active-decision chronology, including rematches and challenges, anchored at the recorded last-move time; the 30- and 90-day horizons give the same picture, and a one-hour timing sensitivity on 7,298,937 complete-window focal opportunities, not specific to grants, gives later-meeting rates of 3.18 against 2.97 percent. This is observed repayment capacity over six months: not an ex ante expected return, not a lifetime bound, and not evidence against every reputational motive, since later meetings can depend on the initial action, and recipient approval, spectators, chat, and off-platform relationships are not measured.

Table B.10: Repeat meetings and predicted re-encounter risk, one-in-five pair sample

Comparison (percentage points)	Estimate	SE	95% interval	Opportunities
Repeat minus first recorded meeting, competitive positions	−0.017	0.030	[−0.076, 0.042]	3,038,973
Competitive minus clearly worse, predicted low-risk sample	1.942	0.170	[1.610, 2.275]	231,011
Same contrast, low-risk first recorded meetings only	1.946	0.170	[1.613, 2.280]	230,739
Joint model: competitive minus clearly worse at first meetings	1.822	0.014		7,569,155
Joint model: repeat-minus-first change in that contrast	−0.464	0.033		7,569,155

Notes: Linear probability models absorbing chooser, month, speed, hour-of-week, and ply-band fixed effects (ply bands at most 4, 5 to 10, 11 to 20, 21 to 40, 41 to 60, 61 to 100, and above 100), with both ratings, both rating deviations, the forgone premium, the draw payoff, and logged clock readings with missingness indicators as controls; the competitive-only repeat comparison also absorbs evaluation bands, which the joint model does not. Standard errors are clustered by chooser account and condition on the fitted prediction. The joint model uses 7,569,155 opportunities from 1,895,845 chooser accounts. The shared-account-or-date covariance gives standard errors of 0.030 and 0.163 for the first two rows. The low-risk sample is the 252,423 evaluation opportunities with a predicted 90-day re-encounter probability below 0.887 percent, excluding the modestly-worse region in the contrasts. The joint-model rows come from a different specification from the first row.

B.13 Audience proxies

Section 2.2 notes that games can have spectators. The data do not record who watched a game, so this appendix asks a narrower question: whether granting is concentrated among the accounts most likely to have an audience. Five proxies are available: a rating of 2400 or above at the focal game; a recognized chess title (GM, IM, FM, CM, NM, and the women's titles) in the account's profile; membership in the platform's streamer directory; a Twitch, YouTube, or Kick link in the profile text; and positive recorded playing time in games featured on the platform's spectator channel. Profiles and the directory were collected in August and September 2026, after the games. A title or link is therefore measured at that time and not at the focal decision, and featured-game time says that the account was broadcast at some point, not that the focal game was watched. The analysis uses the sample as it stood before the author's accounts were excluded, 41,534,331 opportunities and 570,091 grants, 42 opportunities and one grant more than the main sample.

Table B.11, Panel A, removes opportunities in which either player carries a proxy, cumulatively, and reports the remaining sample's grant rates by position and its within-chooser competitive-minus-clearly-worse gap, the estimator of Appendix B.2 with chooser fixed effects and no other controls. Opportunities involving a player rated 2400 or above, titled, or listed as a streamer are 0.57 percent of opportunities and 0.43 percent

of grants, and removing them leaves the gap at 2.173 points against 2.174. Adding the channel-link and featured-game proxies removes 17.6 percent of opportunities and 20.1 percent of grants and leaves the gap at 2.090 (0.010). Within the subsample in which both players' profiles and the featured-game measure are observed, the corresponding figures are 2.859 and 2.879. Profiles were collected with priority to accounts with granting history, so that subsample's levels are higher than the full sample's; the comparison within it is between the subsample and its own screened remainder. The excluded groups show within-chooser gaps of 2.32 to 2.59 points, so the pattern is present among the accounts with a proxy and not carried by them.

Panel B reports grant rates by chooser group, with each chooser assigned to one group: known streamer first, then rated 2400 or above or titled, then the remainder. Choosers rated 2400 or above or titled grant less often overall than the remainder, 0.91 against 1.37 percent, a difference of -0.46 points (standard error 0.03, clustered by chooser). The difference is concentrated in clearly-worse positions, -0.23 (0.02), and is not significant in competitive ones, -0.16 (0.11); the group's within-chooser gap is 2.24 (0.12). Known streamers are 178 choosers with 29 grants, and their rate of 1.16 percent is -0.21 (0.34) from the remainder.

The individual indicators do not share a sign (Panel C): choosers with a recognized title grant at 1.81 percent, those with a profile channel link at 2.27 percent, those with featured-game time at 1.78 percent, and those rated 2400 or above at 0.87 percent, in unadjusted and overlapping counts. The conclusion is the one stated in Section 2.2: granting is not concentrated among the players these proxies identify. The proxies do not measure whether the focal decision was observed, and none of these comparisons controls for the draw payoff or estimates the effect of being watched.

Table B.11: Audience proxies

	Opportunities	Grants	Rate (%)	Competitive rate (%)	Clearly worse rate (%)	Within-chooser gap (pp)
Panel A. Cumulative exclusion of opportunities in which either player carries a proxy						
Full sample	41,534,331	570,091	1.373	2.758	0.470	2.174 (0.010)
Exclude rating $\geq$ 2400, title, or streamer directory	41,297,479	567,613	1.374	2.758	0.472	2.173 (0.010)
excluded group	236,852	2,478	1.046	2.815	0.258	2.321 (0.111)
Also exclude profile channel link	41,235,640	566,395	1.374	2.756	0.471	2.172 (0.010)
excluded group, cumulative	298,691	3,696	1.237	3.077	0.322	2.594 (0.115)
Also exclude positive featured-game time	34,208,788	455,609	1.332	2.685	0.466	2.090 (0.010)
excluded group, cumulative	7,325,543	114,482	1.563	3.076	0.490	2.551 (0.025)
Both profiles and featured-game time observed	16,007,874	319,436	1.995	3.764	0.639	2.879 (0.015)
after all three exclusions	11,758,035	238,732	2.030	3.824	0.666	2.859 (0.016)
Panel B. Chooser groups, one group per chooser						
Remainder	41,345,501	568,366	1.375	2.759	0.471	2.174 (0.010)
Rated $\geq$ 2400 or titled, not a known streamer	186,338	1,696	0.910	2.595	0.240	2.236 (0.123)
Known streamer	2,492	29	1.164	2.654	0.077	2.894 (1.022)
Panel C. Individual indicators carried by the chooser, overlapping						
Rated $\geq$ 2400 at the game	176,148	1,524	0.865			
Recognized title	19,088	345	1.807			
Streamer directory	2,492	29	1.164			
Profile channel link	42,523	967	2.274			
Positive featured-game time	4,265,814	75,726	1.775			

Notes: Sample before the author-account exclusion. The gap is the competitive-minus-clearly-worse difference with chooser fixed effects and no other controls; standard errors in parentheses, clustered by chooser. Panel A excludes an opportunity when either player carries the proxy; the last two rows restrict to opportunities with both players' profiles and the featured-game measure observed. Panel B assigns each chooser to one group. Panel C counts are unadjusted and overlap across rows. Known streamers are 178 choosers.

C Relative Strength and Rating Consequences

This appendix documents the rating reconstruction of Section 2.5, the rating-gap profile by region summarized in Section 4.5, the adjacent-rating comparisons, and the forgone premium on percentile and personal scales.

C.1 Rating reconstruction

Both players' displayed pre-game ratings are recorded in the game headers and are used as recorded. The potential rating changes from a draw and a claimed win are reconstructed from pre-game information because the two outcomes cannot both be observed for the same game. For each player I replay earlier rated games separately within each rating pool, recovering the Glicko-2 [Glickman, 2001] rating deviation, the volatility, and the timing of previous updates, and allowing rating uncertainty to increase with elapsed time between games; at each game the displayed rating is anchored again to its recorded pre-game value. The replay runs from January 2013, the first month of the public archive, and keeps state separately for each speed pool. When an account first appears in a pool, its rating deviation is set to 500 and its volatility to 0.09, the platform's values for a new account, with no elapsed-time adjustment before that first game; both then evolve through the account's recorded games in the pool. The implementation takes its constants from the platform's code: a scale of 173.7178, a rating deviation bounded between 45 and 500, a volatility cap of 0.10, a system constant τ of 0.75, and 0.21436 rating periods per day for the elapsed-time adjustment.

Games whose recorded rating changes are incomplete are handled by two rules. When both players' recorded changes are missing, the replay does not update either player's rating deviation or volatility, and neither player's last-update timestamp advances; hypothetical draw and win changes are nevertheless computed from the reconstructed pre-game state when the game is an eligible opportunity. When only one player's recorded change is missing, the observed change on the other side is treated as evidence that a rated update occurred, and both players' hidden states and update timestamps advance. Missing changes are not imputed as zero. The subsequent elapsed-time adjustment runs from the last update retained under these rules. The reconstruction applies neither additional same-game inflation of rating uncertainty nor the platform's post-update rating regulator.

The reconstructed pre-game states supply, for each opportunity, the chooser's expected score E and update scale K, and hence the hypothetical draw and win changes $r_G = K(\frac{1}{2} - E)$ and $r_C = K(1 - E)$ and the forgone win premium $\kappa = K/2$ of Sec-

tion 2.5. Those quantities satisfy the identity by construction. The reconstruction is validated against the recorded rating change of the action actually taken, which the game headers record for both players. Across the 564,118 granted draws with a recorded update, the mean absolute error between the reconstructed draw change and the recorded one is 0.344 rating points, 99.12 percent of reconstructed changes fall within one point of the recorded change, and the 99th percentile of the absolute error is 0.968 points. Across the 40,707,319 claimed wins, the corresponding figures are 0.459 points, 96.28 percent, and 3.252 points. The larger claim-error tail is concentrated in cells with high rating deviation, which carry a signed error of about -2 points; the replay omits the platform's post-update regulator, which matters most for such accounts. Validation of the realized action does not independently validate its unchosen counterfactual, which is never observed; the counterfactual inherits the replay's accuracy on the realized side only under the maintained Glicko-2 update. Post-change validation after the platform's November 2025 change to the expected-score formula matches the pre-change error cell by cell within rating-deviation bins (0.334 against 0.338 points at the floor; 4.83 against 4.90 at deviations of 200 and above).

The update scale is closely tied to the chooser's own rating uncertainty. Over the 41,534,289 opportunities of the analysis sample, the correlation between $\log K$ and the chooser's pre-game rating deviation is 0.897, and the correlation with the opponent's rating deviation is 0.017. Distance from zero on the draw-payoff axis therefore reflects the chooser's account state as well as the displayed rating gap.

C.2 The rating gap by region

The opponent's pre-game rating also predicts granting. In roughly equal positions after ply ten, the same chooser grants less often to weaker-rated opponents and more often to stronger-rated ones. Relative to opponents of approximately equal rating, the adjusted grant rate is 0.47 percentage points lower toward opponents about 100 points weaker and 0.25 points higher toward opponents about 100 points stronger (Table C.1, Panel A). The rating gap combines information about the opponent with information about the chooser's own payoff. A stronger opponent may be viewed as more likely to have won had play continued. Drawing against that opponent also raises the chooser's rating, whereas drawing against a weaker opponent lowers it, with the forgone premium $K/2$ unchanged. Equality as a reference point, in the manner of par in Pope and Schweitzer [2011], finds no support: displayed gaps of -1, 0, and $+1$ show no level difference (Table C.1, Panel B; Appendix C.3).

Table C.1: Relative strength and the rating consequences of granting a draw

A. Grant rate by displayed rating gap, 50-point bins, relative to the bin from −25 to below +25 (pp)			
Gap bin center	Coefficient (SE)	Gap bin center	Coefficient (SE)
−550	−1.053 (0.510)	+50	0.153 (0.016)
−500	−0.534 (0.586)	+100	0.245 (0.027)
−450	−1.229 (0.561)	+150	0.431 (0.056)
−400	−0.697 (0.595)	+200	0.437 (0.173)
−350	−0.881 (0.363)	+250	0.747 (0.322)
−300	−0.013 (0.363)	+300	0.428 (0.340)
−250	−0.647 (0.222)	+350	1.320 (0.556)
−200	−0.395 (0.139)	+400	1.433 (0.749)
−150	−0.580 (0.044)	+450	0.312 (0.753)
−100	−0.469 (0.023)	+500	0.083 (0.869)
−50	−0.295 (0.015)	+550	1.239 (1.032)
0	0 (reference)		
B. Near equal ratings			
Level, equal ratings minus a displayed gap of −1 (pp)			0.012 (0.066)
Level, gap of +1 minus gap of −1 (pp)			−0.031 (0.069)

Notes: The displayed rating gap is the disconnected opponent's rating minus the chooser's. Panel A: 5,940,861 opportunities in roughly equal positions after ply 10; chooser fixed effects with controls for the update scale, month, speed, stage, evaluation, and the premium; standard errors clustered by chooser. Panel B: 591,108 competitive-position opportunities at gaps of exactly −1, 0, and +1. p-values and variants are in Appendix C.3.

C.3 Adjacent displayed-rating levels

Table C.2 reports the level comparisons at adjacent displayed ratings described in Appendix C.2, together with two variants. The sample contains competitive-position opportunities, with evaluation at least −100, at all game stages in which the displayed rating gap is exactly −1, 0, or +1; the two reported contrasts are equal ratings minus a gap of −1, and a gap of +1 minus a gap of −1.

The linear probability model includes chooser and month fixed effects; standardized board evaluation and its square; standardized $\log(1 + \text{seconds})$ on both clocks, with missing values median-filled before standardization and indicators for missingness; speed indicators; and indicators for the game-stage bands 0–4, 5–6, 7–8, 9–10, 11–20, 21–40, 41–80, and 81 or more plies. Standard errors are clustered by chooser account. The first variant restricts the sample to games ending after ply 10. The second retains pairs in which both pre-game rating deviations are at most 110 and the two displayed ratings are not both 1500; it does not require each rating individually to differ from 1500. A low rating deviation is a recorded rating-state restriction rather than a statement about how long an account has existed, and a displayed rating of 1500 does not establish that an account is new.

None of the comparisons detects a level difference. That is not an equivalence test, does not establish exact continuity, and does not reject reference dependence, since a reference-dependent valuation can be continuous at its reference point while changing slope.

Table C.2: Grant rate at adjacent displayed rating gaps

Comparison	Estimate (pp)	SE	p-value	Opportunities	Accounts
Equal ratings minus gap -1	0.012	0.066	0.859	591,108	106,964
Gap $+1$ minus gap -1	-0.031	0.069	0.654	591,108	106,964
Equal minus gap -1, games after ply 10	0.065	0.073	0.379	456,191	76,268
Equal minus gap -1, both rating deviations at most 110 and ratings not both 1500	-0.005	0.067	0.946	566,712	103,227

Notes: Competitive-position opportunities (evaluation at least -100 centipawns) with displayed gaps of exactly -1, 0, or $+1$. Linear probability models with chooser and month fixed effects and the controls described in the text; standard errors are clustered by chooser account. The first two rows come from one model on the full sample; the third and fourth are the same specification on the indicated subsamples.

C.4 The forgone premium in percentile points and personal rating variation

Table C.3 puts the forgone win premium on two further scales. The percentile scale compares the modeled win and draw outcomes against an empirical rating distribution formed, for each speed and preceding calendar month, from one observation per account, its latest supported pre-game rating in that month; the percentile difference is $100\,[F(\text{win rating}) - F(\text{draw rating})]$, computed for each opportunity before taking medians. The reference is the observed active-account history, not the platform's live leaderboard or its full population. The personal scale divides the premium by the sample standard deviation of the chooser's preceding 100 observed same-speed pre-game ratings, using strictly earlier game starts and denominator 99; zero or unsupported standard deviations are missing, and histories begin with the available archive rather than at account creation.

All entries are medians of opportunity-level quantities and are separate summaries, not a description of one representative observation. They are scales for the modeled forgone premium; they do not identify utility, willingness to pay, or price sensitivity. On the 13,017,592 opportunities that support every denominator, the median ratio is 0.294 with the preceding 100 ratings' standard deviation, 0.410 with the preceding 50, 0.207 with the 100-rating interquartile range, and 0.415 with a detrended 100-rating standard deviation; the equal-account and common-support sensitivities give the same picture.

Table C.3: The forgone win premium in percentile points and personal rating variation

Sample	Opportunities or grants	Premium (rating points)	Percentile points	Premium / prior rating SD	Supported SD histories
Competitive opportunities	15,207,414	5.683	0.429	0.295	13,032,714
Competitive grants	419,462	5.709	0.446	0.293	328,347
Competitive opportunities, chooser RD $\leq$ 110	14,692,899	5.682	0.425	0.294	13,020,215
Competitive grants, chooser RD $\leq$ 110	394,652	5.696	0.434	0.293	327,848

Notes: Medians of opportunity-level quantities. The premium is the chooser's modeled win-minus-draw rating change; percentile points are the difference in the empirical distribution function of the preceding month's same-speed active-account ratings between the modeled win and draw ratings; the last ratio uses the standard deviation of the chooser's preceding 100 same-speed pre-game ratings and is computed on the opportunities with a supported positive standard deviation, counted in the final column. RD is the chooser's rating deviation.

D Behavior After Receipt: Designs and Diagnostics

This appendix gives the designs and the diagnostics behind Section 6: the receipt design and its companion estimates, the instrumental-variables analysis with its alternative constructions and its entry, placebo, balance, and monotonicity diagnostics, the opponent's disconnection environment, and the comparisons that separate the act from the result. The persistence cohort of Section 5.1 and the persistence comparisons by board state, the interval profile, the exposure-specification variants, the trait-by-trait balance coefficients, the subgroup first stages, and the competitive-only score are in Appendix F.

D.1 Receiving a draw: design and companion estimates

The receipt comparison of Section 6.1 follows recipient accounts forward from a single index exposure, and the order of construction matters. For each account I first select its earliest competitive-position, non-tournament disconnection in the panel, ordered by recorded game-start timestamp with archive order and game identifier as tie-breakers. That exposure is then retained only if the connected opponent granted a draw or claimed a win and the unordered pair had no earlier recorded rated game; draws that fail the material rule and outcomes in which the exposure opponent lost are excluded. An index that fails these restrictions is not replaced by a later one, so there is one exposure per recipient account at most.

Three nested samples follow. The supported index cohort contains 2,294,011 recipient accounts; the index game is the game from which a recipient's 90-day follow-up is measured. Requiring its recorded last-move timestamp plus 90 days to fall within the panel leaves 2,110,399 exposures with complete follow-up; exposures with incomplete windows are excluded rather than coded zero. Requiring a subsequent competitive-position

choice within that window leaves the 993,762 observations of the conditional comparisons. The subsequent choice is the recipient's first later competitive-position choice as chooser in the panel, skipping other board states, with recorded last-move time strictly after the exposure anchor and at most 90 days later; it does not impose the instrumental-variables design's different-opponent requirement.

Draw and claim recipients are compared within exposure cells. The cells cross eight bands of the disconnected-player evaluation in the exposure game, $[-100, -51]$, $[-50, -1]$, $[0, 50]$, $[51, 100]$, $[101, 200]$, $[201, 400]$, $[401, 800]$, and 801 or more centipawns; the sign of the exposure opponent's draw payoff, negative, zero, or positive; eight bands of the rating gap, exposure opponent minus recipient, below -400, $[-400, -201]$, $[-200, -101]$, $[-100, -1]$, $[0, 99]$, $[100, 199]$, $[200, 399]$, and 400 or more; and game speed. A cell is retained if it holds at least five grants and twenty claims. Grant recipients receive weight one and controls the cell's grant count divided by its claim count. This support and these weights are formed in the index cohort, before the follow-up and outcome restrictions are applied, and are reused unchanged across the comparisons that follow.

The primary regression absorbs exposure-cell and exposure-month fixed effects and includes linear and squared terms for thirteen exposure-game variables: the evaluation; the exposure opponent's draw payoff and forgone win premium; both ratings, both rating deviations, and both remaining clocks; the ply count; the chooser-side material advantage; and the time-control base and increment. It contains no recipient fixed effects and is not a before-and-after comparison. Standard errors cluster by the exposure-opponent account that made the decision.

Table D.1 collects the primary estimate and its companions, which differ in outcome and denominator. The later-choice-adjusted specification adds controls measured at the later choice and therefore after the exposure: its evaluation, draw payoff, premium, both clocks, both ratings, both rating deviations, elapsed time, and later-speed and month categories. The exposure-hour specifications add fixed effects for the UTC hour in which the exposure game started or, alternatively, the hour of its last move; both nest the exposure-month effects, so that draw and claim recipients are compared only with others disconnected in the same hour. They remove any shock common to an hour of exposures, such as a connectivity incident. A further specification also adds fixed effects for the hour of the later choice's last move, which removes a common shock that persists until that choice.

Because draw recipients return to rated play sooner (Panel C), the timing of the later choice follows the exposure, and that specification is a sensitivity check rather than an adjustment. With exposure-hour effects, the probability of reaching a later competitive choice changes by -0.22 (0.20) or -0.21 (0.20) points under the start-hour and last-move-

hour definitions, and the zero-coded outcome is 0.39 (0.06) under either definition.

The reach, any-grant, and zero-coded outcomes all use the 2,110,399 exposures with complete follow-up; among them, the any-grant outcome is an indicator for at least one competitive-position grant anywhere in the 90-day window and is distinct from the zero-coded first-choice outcome. The censored comparison retains accounts of the conditional sample with no further competitive-position disconnection within 90 days of the index, or whose first later competitive choice occurs no later than that disconnection, with timestamp ties retained; the censoring disconnection need satisfy neither a grant-or-claim nor a first-observed-pair restriction. The zero-coded censored row keeps all 2,110,399 exposures and counts a later grant only if it comes strictly before the recipient's next competitive disconnection. The engagement models use the 2,294,011-account support cohort, because their separate rated-game chronology supplies complete 30-day follow-up without the panel's 90-day requirement, and their outcomes concern any rated standard game rather than competitive choices.

D.2 The instrumental-variables analysis

Opponents differ in how often they grant, and the instrumental-variables analysis of Section 6.2 uses that variation, measured from each opponent's decisions at other players' disconnections rather than from the exposure at hand. The instrument is the exposure opponent's state-adjusted grant score over all of its other grant-or-claim decisions in the 24-month panel, before and after the exposure, with every decision between the two accounts excluded in either direction of time and at least five decisions remaining; that support requirement is also the sample's eligibility rule. The state-reference construction excludes both participants' cross-fitting folds. The score inputs are treated as fixed measured quantities, so their estimation uncertainty is not incorporated. The score for opponent j at an exposure with recipient i is

$$Z_{ij} = \frac{1}{n_{ij}} \sum_{h \in H_{ij}} (g_h - \hat{p}_h),$$

where H_{ij} is the set of j's recorded grant-or-claim decisions as chooser, in any board state and from any game source, in the panel's 24 months, excluding every decision between j and i in either role and either direction of time, and $n_{ij} = |H_{ij}| \geq 5$; g_h is one for a granted draw and zero for a claim; and $\hat{p}_h$ is the expected grant probability of decision h's cell. Decisions in which j was the disconnected player cannot enter j's history, ordinary endings are not counted as refusals, and the author's accounts are excluded from H_{ij} and from the reference. Cells cross five bands of the disconnected player's evaluation (at most

Table D.1: Receiving a draw: primary estimate and companion outcomes

Outcome	Estimate	SE	p-value	Exposures	Clusters
Panel A. Conditional on a later competitive choice (percentage points)					
Grant at first later competitive choice	0.882	0.122	< 0.001	993,762	644,491
Same, with later-choice controls	0.898	0.122	< 0.001	993,762	644,491
Same, exposure start-hour effects	0.854	0.127	< 0.001	993,762	644,491
Same, exposure last-move-hour effects	0.820	0.125	< 0.001	993,762	644,491
Same, adding later-choice-hour effects	0.839	0.117	< 0.001	993,762	644,491
Censored at next competitive disconnection	1.095	0.159	< 0.001	624,521	461,377
Panel B. Complete 90-day follow-up (percentage points)					
Reach a competitive choice within 90 days	-0.197	0.194	0.309	2,110,399	1,052,419
Any competitive-position grant within 90 days	0.468	0.072	< 0.001	2,110,399	1,052,419
Grant at first later choice, no choice coded zero	0.399	0.058	< 0.001	2,110,399	1,052,419
Same, exposure start-hour effects	0.392	0.059	< 0.001	2,110,399	1,052,419
Same, exposure last-move-hour effects	0.386	0.059	< 0.001	2,110,399	1,052,419
Same, censored at next competitive disconnection	0.308	0.047	< 0.001	2,110,399	1,052,419
Panel C. Engagement, support cohort with 30-day rated-game chronology					
$\log(1 + \text{rated games within 30 days})$	0.0075	0.0059	0.198	2,294,011	1,127,372
Any rated game within 10 minutes (pp)	0.915	0.168	< 0.001	2,294,011	1,127,372
Any rated game within 30 minutes (pp)	0.950	0.184	< 0.001	2,294,011	1,127,372
Any rated game within 60 minutes (pp)	0.950	0.191	< 0.001	2,294,011	1,127,372

Notes: All rows use the receipt design's cells, weights, fixed effects, and thirteen exposure-game controls; standard errors clustered by exposure opponent. Panel A: weighted means in the first row are 4.386 (treated) and 3.331 percent (control); rows three to five add fixed effects for the UTC hour of the exposure's start, of its last move, and of both games' last moves, with two-way clustering by account and hour; the sixth row retains 624,521 of 993,762 observations. Panel B: the zero-coded row has means of 2.025 and 1.564 percent and the censored row 1.330 and 0.996; the hour rows cluster two ways, as in Panel A. Panel C: outcomes from a separate rated-game chronology with complete 30-day follow-up; the first row is in log points.

-300; -300 to -100; -100 to $+100$; $+100$ to $+300$; at least $+300$), four ply bands (at most 10; 11 to 30; 31 to 60; above 60), five speeds (ultrabullet, bullet, blitz, rapid, classical), and five rating-gap bands with cut points at -200, -100, $+100$, and $+200$, 500 cells in all.

Each $\hat{p}$ is a cell grant rate from the reference model of the decision's own month, smoothed toward the global rate, $\hat{p}_{cm} = (S_{cm} + 50\,\hat{p}_{0m})/(N_{cm} + 50)$ with $\hat{p}_{0m} = (S_m + \frac{1}{2})/(N_m + 1)$, where S and N count grants and decisions within the eligible reference folds in a window ending seven days before month m begins. The reference uses four account folds and excludes observations with either participant in either focal account's fold, so two reference folds remain when the focal folds differ and three when they coincide; a wholly empty reference is unsupported. Where a decision's month has no usable reference, the nearest usable month's is used, the latest earlier one where it exists and otherwise the earliest. There is no further shrinkage of the score itself.

On the spaced rows the median exposure opponent has 41 other decisions and the mean 61, and half of them fall after the exposure. The opponent's expected grant rate over H_{ij}, $\log(1 + n_{ij})$, and the logged number of distinct counterparts in H_{ij} enter the model as controls. First stages, reduced forms, and entry and placebo coefficients are reported per 0.0498 of the score, a fixed unit used in every frame and close to one residual standard deviation under the primary controls: the residual standard deviation is 0.0535 in the primary sample and 0.0545 on the spaced rows. The two-stage ratio does not depend on the unit.

Exposures are competitive-position pool games with grant-or-claim outcomes between first observed unordered pairs. An exposure is supported if its opponent has at least five eligible decisions involving other recipients anywhere in the panel, after the pair exclusion, and a finite score on them; no part of the rule looks at when those decisions fall relative to the exposure. Every supported exposure with complete recorded history over the 90 days before the exposure game's recorded start and complete 90-day outcome follow-up is retained; there is no spacing within recipient and no cap on exposures per recipient. The primary sample contains 9,736,182 exposures from 1,877,395 recipient accounts. Of the recipients, 727,803 contribute one exposure and 1,149,592 contribute more than one, together 9,008,379 of the rows; the mean is 5.2 exposures per recipient, the median 2, the 90th percentile 12, the 99th 46, and the most prolific 444, so that the 11.8 percent of recipients with more than ten exposures supply 53.7 percent of the rows and the 0.8 percent with more than fifty supply 11.4 percent; 4,746,043 rows share their later choice with another exposure of the same recipient, in 1,372,428 groups of which the largest has 82 members, and 3,374,050 rows are followed by another retained exposure before their outcome game starts.

The estimates are exposure weighted, and the covariance below allows dependence

among all exposures of an account; more rows are not more independent observations. Giving each recipient a total weight of one, spread equally over its exposures, lowers the first stage to 6.96 (0.03) and the reduced form to 0.020 (0.006) and gives a two-stage estimate of 0.293 (0.082), against 0.325 (0.056) with exposure weights; the difference is -0.032 (0.063) under the joint covariance. Least squares on the primary rows with the same controls gives 0.270 (0.028), -0.055 (0.048) below the instrumented estimate. The model contains no recipient fixed effects; recipient fixed effects appear only in the variants reported in Appendix F.4.

The spaced design used for comparison admits a supported exposure only if its game starts more than 90 days after the recipient's previous admitted exposure, applies that rule before the history and follow-up cuts, and measures the recipient's history over the window from 97 to 7 days before the game; its 3,103,280 exposures from 1,835,628 recipients are a subset of the primary rows.

The outcome is a grant at the recipient's first later competitive-position choice against a different opponent, in a game that starts strictly after the exposure game's recorded last-move timestamp and with a decision recorded within 90 days of that timestamp. Complete windows without a qualifying choice receive zero, and there is no next-exposure censoring. This outcome is defined over all retained complete-follow-up exposures, not only those followed by a qualifying choice. It is the outcome of the matched comparison's first row in Section 6.1; the exposure construction and controls differ.

The model absorbs cells crossing exposure month, exact time control, and a 100-point bin of the pair's mean displayed rating, together with exposure-date and hour-of-week effects. Its additional controls are the recipient's rating; the rating gap, its square, and its absolute value; the opponent's expected grant rate over its donor decisions, $\log(1 + \text{donor decisions})$, and $\log(1 + \text{distinct donor counterparts})$; for the recipient's own history in the 90 days before the exposure game's start, the earlier endpoint included and the start excluded, an indicator for any prior competitive opportunity, $\log(1 + \text{prior competitive opportunities})$, an ever-granted indicator, and the prior competitive grant rate, with a zero rate accompanied by the availability indicator when no prior competitive opportunity exists; $\log(1 + \text{prior disconnections})$, $\log(1 + \text{prior receipts})$, and $\log(1 + \text{prior claims})$; and $\log(1 + \text{game starts in the same 90 days})$, the recipient's playing volume; and the three opponent-environment measures of Appendix D.3 with their availability indicator. The final evaluation, clocks, and ply of the exposure game are not controls in this model.

Inference allows dependence between exposures that share an account in either role or that share an exposure date. The covariance combines account contributions, subtracts repeated unordered-pair contributions, and applies account-date and pair-date

inclusion–exclusion for common dates, in the manner of multiway clustering [Cameron et al., 2011]; it is therefore not ordinary separate clustering by recipient account, opponent account, and date. Reported intervals are Wald intervals from this asymptotic covariance; the procedure maintains the specified IV model and dependence assumptions and treats the historical scores as fixed. The first-stage F-statistic is 91,156, and on the spaced rows the partial first-stage R^2 of the instrument, net of the cells, time effects, and controls, is 0.233; with a first-stage F of this size weak-instrument corrections are immaterial. Neither design removes intervening receipts or carry-over between a recipient's exposures.

Table D.2 reports the estimate under the spaced and the primary sample designs, each a separate model, with the differences estimated under a covariance shared across the two samples compared. Measuring the recipient's history over the 90 days before the game, on the spaced rows themselves, moves the two-stage estimate from 0.363 to 0.371, a change of 0.009 (0.011). Admitting every eligible exposure, at the spaced design's calendar range, moves it to 0.334 (0.057), a change of -0.038 (0.077), while the first stage rises by 0.129 (0.019) points as the added rows enter. Extending the calendar to the games that the 97-day history requirement had excluded adds 122,055 exposures and moves the estimate by -0.009 (0.005).

Table D.2: The spaced design and the primary sample

Sample	Exposures	First stage	Reduced form	2SLS (SE)	95% interval
Spaced, histories 97 to 7 days before the game	3,103,280	7.125 (0.030)	0.0258 (0.0069)	0.363 (0.096)	[0.174, 0.552]
Same rows, histories over the 90 days before the game	3,103,280	7.125 (0.030)	0.0265 (0.0068)	0.371 (0.096)	[0.183, 0.559]
Spaced, 90-day histories, full calendar	3,140,825	7.120 (0.030)	0.0255 (0.0068)	0.358 (0.095)	[0.171, 0.545]
All eligible exposures, spaced design's calendar	9,614,127	7.253 (0.024)	0.0242 (0.0041)	0.334 (0.057)	[0.222, 0.445]
All eligible exposures, full calendar (primary)	9,736,182	7.253 (0.024)	0.0236 (0.0041)	0.325 (0.056)	[0.214, 0.435]
same, without the recipient-history controls	9,736,182	7.254 (0.024)	0.0243 (0.0044)	0.335 (0.060)	[0.217, 0.452]
Differences under the shared covariance (two-stage least squares, pp per receipt)					
Histories, same rows (row 2 minus row 1)				0.009 (0.011)	$p = 0.44$
Spacing, same calendar (row 4 minus row 1)				-0.038 (0.077)	$p = 0.62$
Calendar, unspaced (row 5 minus row 4)				-0.009 (0.005)	$p = 0.05$
Spacing, full calendar (row 5 minus row 3)				-0.033 (0.076)	$p = 0.66$

Notes: All-time score; cells, time effects, and controls of Table 6 with the history definition of each row; first stages and reduced forms per 0.0498 of the score, the common unit; shared-account-or-date covariance, Wald intervals. The spaced design's calendar excludes games whose 97-day history window is incomplete. Recipients: 1,835,628 in rows 1 and 2, 1,850,392 in row 3, 1,862,144 in row 4, and 1,877,395 in rows 5 and 6.

On the spaced rows, adding fixed effects for each UTC hour in which an exposure

game started, which absorb the exposure-date and hour-of-week effects, leaves the estimates unchanged on the same rows: the first stage is 7.12 (0.03), the reduced form for later granting 0.026 (0.007), the two-stage estimate 0.368 (0.097) with interval $[0.178, 0.557]$, and least squares 0.300 (0.047). Using the hour of the exposure game's last move instead gives a two-stage estimate of 0.366 (0.097), with interval $[0.176, 0.555]$.

Seven further exposure specifications of the spaced design, which vary the spacing between a recipient's exposures, restrict to repeated recipients with and without recipient fixed effects, use the first-exposure rows, or omit the recipient history and volume controls, are reported in Appendix F.4, beside the primary sample. All retain the 90-day outcome horizon and the different-opponent requirement. The two-stage point estimates are positive throughout, from 0.26 to 0.51. The repeated-recipient estimate with recipient fixed effects, 0.264 (0.140), has an interval that includes zero; the first-exposure estimates, 0.506 (0.149) with the common history controls and 0.506 (0.153) without them, are the least precise; and omitting the recipient history and volume controls from the spaced rows moves the estimate from 0.363 to 0.371 (0.100).

Table D.3 reports, on the spaced rows, the alternatives that change the score or the exposure restrictions rather than the sample design. Dropping the opponent-environment controls gives 0.372 (0.096). Building the score only from the opponent's decisions made strictly before the exposure game's start, with the same exclusions and its own three history controls, requires five such decisions and so supports 2,629,196 of the exposures; on them it gives 0.236 (0.119), against 0.247 (0.106) for the all-time score on the same rows, with a first stage of 6.11 against 7.59 percentage points per score unit; the longer history attenuates the score less. Restricting to the 2,832,568 exposures whose final material difference is below four points, where the countdown is not shortened for a material deficit, gives 0.317 (0.099).

Panel C uses the 2,245,874 exposures whose opponent has at least five decisions both before and after the exposure. On those rows, scores built only from earlier decisions and only from later ones give 0.247 (0.134) and 0.253 (0.127), against 0.270 (0.117) for the all-time score; their paired difference, 0.007 (standard error 0.112, $p = 0.95$) from the two equations estimated together under the shared covariance, is within sampling error of zero, so a response of the opponent to the exposure game, which would make the later decisions the more informative, leaves no trace. These rows are selected on the opponent having later decisions and are a diagnostic, not an alternative headline.

Panel D varies the minimum number of decisions behind the score. It rebuilds the spaced sample from the full universe of first-pair exposures under requirements of 3, 5, 10, and 20 of the opponent's decisions over the panel, recomputing the score, its history controls, the spacing, and the history and follow-up restrictions each time; the five-

decision row is the spaced sample. The estimates lie between 0.26 and 0.36, and the first stage rises with the requirement, from 6.72 to 8.09, as the score is measured from more decisions.

Table D.3: Alternative constructions of the score and the exposure restrictions, spaced rows

Specification	First stage	2SLS	SE	95% interval
A. All 3,103,280 spaced exposures				
All-time score, environment controls	7.125	0.363	0.096	$[0.174, 0.552]$
All-time score, without the environment controls	7.128	0.372	0.096	$[0.183, 0.561]$
B. 2,629,196 exposures with at least five opponent decisions before the exposure, environment controls				
All-time score	7.586	0.247	0.106	$[0.038, 0.456]$
Score from decisions before the exposure	6.107	0.236	0.119	$[0.002, 0.470]$
All-time minus before, paired	—	0.011	0.052	$p = 0.84$
C. 2,245,874 exposures with at least five opponent decisions before and after, environment controls				
All-time score	8.050	0.270	0.117	$[0.042, 0.499]$
Decisions before the exposure only	6.233	0.247	0.134	$[-0.016, 0.509]$
Decisions after the exposure only	6.704	0.253	0.127	$[0.005, 0.502]$
After minus before, paired	—	0.007	0.112	$p = 0.95$
D. Minimum opponent decisions, sample and score rebuilt under each requirement				
At least 3: 3,162,520 exposures	6.722	0.302	0.095	$[0.115, 0.489]$
At least 5 (spaced sample): 3,103,280 exposures	7.125	0.363	0.096	$[0.174, 0.552]$
At least 10: 2,957,506 exposures	7.636	0.316	0.099	$[0.122, 0.509]$
At least 20: 2,670,289 exposures	8.094	0.263	0.104	$[0.060, 0.466]$

Notes: First stages are percentage points of receipt per 0.0498 of the score, the paper's common unit; two-stage estimates are per receipt. All rows use the spaced rows with the spaced design's histories, the cells, time effects, and controls of Table 6, Panel A, and the shared-account-or-date covariance, with Wald intervals. Each score carries its own history controls. The before score uses the opponent's decisions strictly before the exposure's start and the after score those strictly after its last move; the paired rows are differences under the joint covariance. The material-difference restriction of Table 6, Panel C, retains 2,832,568 exposures.

The entry diagnostics ask a different question on a different population: whether the instrument predicts appearing in the exposure sample at all. They use an independently sampled frame of 31,836,314 games with all ending types, with the original recipient orientation retained. This frame and the ordinary-ending frame below are built under the primary sample's rule: the opponent's score is supported by at least five eligible decisions anywhere in the panel, with no restriction on the opponent's history before the game; no receipt spacing applies; and the recipient's history and volume are measured over the 90 days before the game. Eligibility requires observed-previous nonadjacency, real-time play, a supported opponent score, and complete recorded recipient history over those 90 days, and the frame is not restricted to accounts that eventually enter the exposure sample.

Because the primary sample has no spacing, no earlier exposure closes a game's eligibility: a game is eligible on its own history coverage, and a pregame calendar cut requires game start plus 90 days to precede the panel end, so that current game endings and future exposures do not determine the eligibility flag.

Table D.4 reports three outcomes: any active timeout; a recipient-side disconnection in the main panel; and final analysis entry. In this frame entry coincides with a recipient-side competitive-position grant-or-claim event and with selection into the primary sample, 146,524 games each, because every supported such event is an exposure; the table reports the one that defines the sample.

With the all-time score and the primary controls (Panel A), the coefficients on the score are -0.0135 (standard error 0.0011) for entry into the exposure sample, on a base of 0.46 percent, 2.9 percent of entrants per score unit; -0.079 (0.003) for any active timeout, 3.3 percent of its rate of 2.43 percent; and -0.031 (0.002) for a recipient-side disconnection, 2.5 percent of its rate of 1.27 percent. Without the recipient-history controls (Panel B) the three are -0.0138, -0.080, and -0.032, so the recipient's own history accounts for little of them. The association is largest for timeouts of any kind, smaller for the recipient's own disconnections, and smaller again for entry.

The calibration is as follows. Removing 2.9 percent of entrants per score unit changes the sample mean of later granting by 0.029 times the gap between the marginal entrants and the rest; to generate the reduced form of 0.024 points, the marginal entrants would have to grant about 0.8 points less than the rest, nearly half the sample's rate of 1.72 percent. Whether such a group exists is what the recipient-history balance tests speak to; in this frame the grant rate moves 0.0001 trait standard deviations per score standard deviation (0.0002), whether the recipient granted at all -0.0001 (0.0002), and the joint test on the eight traits rejects on the recipient's disconnections and claims ($\chi^2(8) = 33.8$, $p < 0.001$; Table F.5), as in the exposure sample, with the limitation that prior behavior is a noisy proxy and that many recipients have no prior competitive choice in the window.

The ordinary-ending diagnostic asks whether the instrument predicts later granting when the index game did not end in a disconnection. It restricts the 90-day eligible sampled frame to index games ending in checkmate, resignation, or stalemate, with complete follow-up from the index game's recorded last-move anchor. Its outcome is a grant at the recipient's first later competitive-position choice against a different opponent within 90 days, with zero assigned when a complete window has no qualifying choice. The ordinary ending therefore characterizes the index game, while granting is measured at a subsequent disconnection choice.

Table D.5 reports the test with the all-time score and the primary controls on a frame of 21,025,393 ordinary endings, beside the reduced form in the exposure sample and their

Table D.4: Opponent grant score and entry into the exposure sample

Outcome	Estimate	SE	p-value	95% interval
A. 90-day eligibility, primary controls, 31,836,314 games				
2,678,988 recipients; entry rate 0.460 percent.				
Any active timeout	−0.0791	0.0027	< 0.001	$[-0.0845, -0.0737]$
Recipient-side main disconnection	−0.0312	0.0019	< 0.001	$[-0.0349, -0.0275]$
Entry into the exposure sample	−0.0135	0.0011	< 0.001	$[-0.0157, -0.0112]$
B. Same frame, without the recipient-history controls				
Any active timeout	−0.0801	0.0028	< 0.001	$[-0.0855, -0.0747]$
Recipient-side main disconnection	−0.0321	0.0019	< 0.001	$[-0.0358, -0.0284]$
Entry into the exposure sample	−0.0138	0.0011	< 0.001	$[-0.0160, -0.0115]$

Notes: Percentage points per 0.0498 of the score, the paper's common unit in the sampled all-ending frame under the 90-day eligibility rule. Panel A uses the primary controls, including the environment and recipient-history controls; Panel B drops the recipient-history controls. Shared-account-or-date covariance; p-values unadjusted.

difference. The association on all ordinary-ending games must exclude a direct effect of 0.024 points per score unit, the value that would account for the whole reduced form, and it does: the 95 percent interval reaches 0.004. The test covers channels that operate in games without a disconnection; a direct effect confined to disconnection games, such as a message left in the game's chat at the disconnection, is not testable with ordinary endings.

Table D.5: Later granting and reach following ordinary-ending index games

Later outcome	Estimate	SE	Games	95% interval
Frame of 21,025,393 ordinary endings, 2,424,236 recipients, all-time score, primary controls				
Grant at first later competitive choice, no choice coded zero	−0.0014	0.0028	21,025,393	$[-0.0069, 0.0041]$
same, without the recipient-history controls	−0.0001	0.0029	21,025,393	$[-0.0058, 0.0057]$
Reach a later competitive choice	−0.0023	0.0086	21,025,393	$[-0.0192, 0.0146]$
Matching receipt reduced form, primary sample (Table 6)	0.0236	0.0041	9,736,182	$[0.0155, 0.0316]$
Receipt reduced form minus the ordinary-ending association	0.0250	0.0050	—	$[0.0152, 0.0348]$

Notes: Outcomes over the subsequent 90 days, in percentage points per 0.0498 of the score, the paper's common unit, with the primary controls, cells, and time effects and the shared-account-or-date covariance. The later-grant mean is 1.94 percent. The first row is the Table 6 estimate; its 90 percent interval ends at 0.0032, less than a sixth of the reduced form, and the difference row lets the two samples share accounts and dates ($p < 0.001$).

The balance diagnostics regress each recipient history measure on the residualized opponent score under each specification's own adjustment. The coefficients are measured in standard deviations of the trait per native residual-score standard deviation, a

different scale from the percentage points per common score standard deviation used for first stages, reduced forms, and entry diagnostics, and the two should not be compared directly. In the primary design the eight traits are the seven history measures and the volume measure over the 90 days before the index game's start; in the spaced design the 90-day history measures cover the interval from 97 to 7 days before the start, the recent history measures the interval from 37 to 7 days before it, and the volume measure counts rated games in the immediately preceding 30 days. All logged counts use $\log(1 + \text{count})$, and grant rates retain their history-availability indicators when no prior competitive choice is observed.

Traits already included directly as controls are not independent balance evidence: the primary design is therefore tested on its eight traits with the recipient-history controls omitted, and the spaced design's specifications that control for the 90-day traits are tested on the seven recent-history traits. The trait-by-trait coefficients are reported in Appendix F.5.

Tables F.5 and F.6 report the joint tests with the trait-by-trait coefficients. In the primary sample, with the environment controls retained and the recipient-history controls omitted, the eight-trait test rejects ($\chi^2(8) = 39.7$, $p < 0.001$), carried by the recipient's disconnections and claims, about -0.0008 trait standard deviations per score standard deviation each, and by its receipts of a draw, $+0.0009$ (standard errors 0.0003); the grant traits are within 0.0001 of zero. The entry and ordinary-ending frames, tested the same way, reject as well ($\chi^2(8) = 33.8$ and 27.1, $p < 0.001$), on the same two traits: the recipient's disconnections and claims move about -0.0007 in the entry frame and -0.0006 in the ordinary-ending frame (standard errors 0.0002), while receipts do not lean in either frame (-0.0003 and -0.0002, standard errors 0.0002) and the grant traits are within 0.0003 of zero. On the spaced rows, with the primary controls of that design, the seven-recent-trait test has $p = 0.067$, and without the environment controls $p = 0.017$, carried by the recipient's recent disconnections and claims; on those rows the environment measures remove the associations while moving the receipt estimate by two percent (Appendix D.3). In the primary sample, which has three times the rows, and in its frames, the same two traits remain detectable with the environment measures in place, and the recipient's own history absorbs them as a control, moving the receipt estimate by three percent (Table D.2) and the entry associations by less.

The disconnection and claim imbalance is the opponent's environment showing through the matchmaker's cells. The grant-rate traits, the strongest predictors of later granting in the panel (Table 3), lean slightly toward the score on the spaced rows and not at all in the primary sample or its frames: whether the recipient granted in the preceding month moves 0.0009 trait standard deviations per score standard deviation (0.0004) on

the spaced rows, and whether it granted in the 90 days before the game moves 0.0001 (0.0003) in the primary sample, −0.0001 (0.0002) in the entry frame, and −0.0001 (0.0002) in the ordinary-ending frame.

The implied-bias calibration, on the spaced rows, turns the fifteen recorded traits into one number: later granting is projected on all fifteen, with the fixed effects and the non-history controls, and the fitted trait component is regressed on the residualized score under the primary model. The implied association is 0.0005 points per score unit (standard error 0.0003, which includes the uncertainty in the projection), 1.9 percent of the reduced form of 0.026; the same calculation for receipt gives 0.0001 against a first stage of 7.12. It is a calibration on recorded traits, not a bound on selection on unrecorded ones.

On the spaced rows, two variants complete the balance evidence (Table F.6, columns 5 and 6). Dropping the recipient history and volume controls, with the environment controls retained, gives a fifteen-trait test at the margin ($\chi^2(15) = 26.7$, $p = 0.031$; the seven recent traits alone give $\chi^2(7) = 16.4$, $p = 0.021$), carried by the recipient's disconnections and claims at about −0.0013 each (standard errors 0.0005), whereas without the environment controls the same fifteen-trait test rejects outright ($\chi^2(15) = 83.1$). With recipient fixed effects on the rows of recipients with at least two retained exposures, the joint test rejects ($\chi^2(7) = 92.2$, $p < 0.001$), and one trait carries it: within recipient, the score moves with the recipient's recent receipts, 0.0056 trait standard deviations per score standard deviation (0.0006), while the other six traits lie within 0.0014 of zero. Recent receipts predict later granting (Section 6.1), so that association would push the fixed-effects variant upward; its estimate, 0.264 (0.140), is below the primary one.

D.2.1 Monotonicity

On the spaced rows, the first stage is positive in all 30 supported subgroups of game speed, recipient rating, prior competitive choices, and exposure month, and every 95 percent interval lies above zero (Appendix F.6). In percentage points of receipt per score unit, it rises with the recipient's rating, from 6.31 below 1200 to 8.30 at 2000 and above; it is lower in rapid and classical games, 6.62 and 6.40 against 7.39 in bullet and 7.45 in blitz; it is 6.88 among recipients with no competitive choice in the prior 90 days, 7.43 among those with choices and no grant, and 7.33 among those who had granted; and it lies between 6.85 and 7.28 in every exposure month. No subgroup changes sign between the raw and adjusted fits. The subsamples of the paper's own splits agree, with the primary controls: 7.13 (standard error 0.03) among the 2,832,568 exposures without a large material imbalance; 7.32 (0.03) and 6.76 (0.04) among the 2,127,811 and 975,469 exposures of recipients who had and had not faced a chooser-side competitive decision earlier in the panel; 7.42

(0.04) and 6.88 (0.04) among the 1,513,306 and 1,589,974 exposures of recipients who had and had not faced one in the 90-day window; and 8.05 (0.05) on the 2,245,874 exposures whose opponent has decisions on both sides of the exposure.

On the primary rows, receipt rises at every step across the twenty equal-count bins of the residualized score (Figure 3), from -3.1 to $+28.7$ percentage points. These are testable implications of average monotonicity.[23]

Complier means describe the weighted margin of receipt shifted by the instrument, rather than an observed group of accounts (Table D.6). Under the maintained conditional IV assumptions, this margin resembles the sample on the reported characteristics. On the spaced rows its mean rating is 1,343, compared with 1,427 for all exposures; recipients at this margin also played fewer recent rated games (130 versus 147) and had fewer prior disconnections (4.29 versus 4.84). Prior granting and the game-speed mix are similar. On the primary rows, instrumenting the transformed outcomes $Y_e D_e$ and $Y_e(1 - D_e)$ with the same model gives the compliers' later-granting rates with and without a draw received, 2.19 percent (standard error 0.05) and 1.86 percent (0.04); their difference reproduces the two-stage estimate. These are diagnostics under the maintained assumptions, not a test of them.

D.3 The opponent's disconnection environment

The opponent's score has small negative associations with the recipient's recent disconnections and claims (Section 6.2). The score is built from the opponent's decisions at other players' disconnections, so it is natural to ask whether it carries information about the connection conditions the opponent plays in, which the matchmaker's cells would then share with the recipient. Three measures index those conditions: over the opponent's authenticated rated games starting in the 90 days ending seven days before the exposure month, excluding the current pair, the logged number of games, the share that ended in a certified active resolution with the opponent as the disconnected side, and the share that ended that way with the opponent's counterpart disconnected. They are rates of observed active-resolution events, not a census of connection failures. In the primary specification they enter with an indicator for exposures whose opponent lacks a supported history, whose measures are set to zero, so that every exposure is retained;

[23] Pairwise monotonicity is sufficient but stronger than necessary. A weaker average-monotonicity condition requires, for each opportunity, a nonnegative relationship across possible assignments between the assigned score and whether a draw would be granted, under the instrument and conditioning specification actually used [Goldsmith-Pinkham et al., 2026]. The formal test of Frandsen et al. [2023] is not applied: it uses mutually exclusive assignment instruments, and applying it to the continuous, time-varying, pair-excluded historical score would require a validated mapping and inference adaptation. The subsample check does not test exclusion.

Table D.6: Complier characteristics, spaced rows

	Complier margin		All	Draw
Characteristic	Mean	SE	exposures	recipients
Displayed rating	1,343	3.7	1,427	1,383
Rated games, preceding 30 days	130	1.2	147	141
Competitive choices, prior 90 days	1.56	0.016	1.61	1.59
Granted at least once, prior 90 days	0.027	0.0009	0.026	0.027
No competitive choice, prior 90 days	0.527	0.0028	0.512	0.519
Disconnections, prior 90 days	4.29	0.054	4.84	4.45
Draws received, prior 90 days	0.07	0.002	0.07	0.07
Bullet share	0.101	0.0024	0.112	0.111
Blitz share	0.544	0.0048	0.552	0.535
Rapid share	0.335	0.0049	0.320	0.335
Classical share	0.020	0.0013	0.016	0.019

Notes: Complier means are instrument-weighted under the maintained conditional IV assumptions, with the all-time score and primary controls; they describe a treatment margin, not an observed group. Negative weights, 2.5 percent of the signed total, are retained. Comparison columns are exposure-weighted means for the 3,103,280 spaced exposures and their 83,907 draw receipts. Standard errors use the shared-account-or-date covariance. Indicator rows are shares.

11.2 percent of the spaced exposures lack one.

The comparison in this subsection is on the spaced rows. Table D.7 compares the specifications with and without them on identical rows. The joint balance test on the seven recipient-history traits moves from $p = 0.017$ to $p = 0.067$ in the full exposure sample, while the two-stage estimate moves from 0.372 (0.096) to 0.363 (0.096), and from 0.326 (0.099) to 0.317 (0.099) among exposures without a large material imbalance. The entry associations shrink but do not vanish (Table D.4); Section 6.2 calibrates what remains. The recipient-history imbalance tracks the opponent's environment rather than the recipient's type, and absorbing it barely changes the receipt estimate.

D.4 What recipients respond to

Section 6.3 asks whether recipients respond to the opponent's act or to the result with two comparisons; Table D.8 reports them. Panel A compares players whose ordinary games, with no disconnection, ended in a draw with players whose ordinary games ended in a loss, on the same later-granting outcome as the instrumental-variables design. The main specification weights ordinary-draw and ordinary-loss games by overlap [Li et al., 2018] on the propensity of a draw given both players' ratings, the rating gap, the time control, prior activity, and the recipient's recent disconnection, claim, receipt, and grant history; a sensitivity matches exactly on time control, and a third compares games of the same

Table D.7: Adding the opponent's disconnection environment, spaced rows

Controls	First stage	2SLS (SE)	95% interval	Balance, joint p
A. All 3,103,280 spaced exposures				
Without the environment controls	7.128	0.372 (0.096)	[0.183, 0.561]	0.017
With the environment controls	7.125	0.363 (0.096)	[0.174, 0.552]	0.067
B. 2,832,568 exposures with a final material difference below four points				
Without the environment controls	7.131	0.326 (0.099)	[0.132, 0.521]	—
With the environment controls	7.128	0.317 (0.099)	[0.122, 0.512]	—

Notes: All-time score; first stages in percentage points per 0.0498 of the score, the paper's common unit. Balance is the joint test on the seven pre-exposure traits. The environment controls are the opponent's logged rated games and two disconnection shares, with an availability indicator. In Panel A the reduced form is 0.0265 (0.0069) without them and 0.0258 (0.0069) with them; the balance test is not refitted on the Panel B rows. Shared-account-or-date covariance.

recipient. The covariance allows dependence across games sharing an account in any role. The largest standardized mean difference after weighting is 0.08, on the rating gap.

Panel B estimates the receipt effect, on the spaced rows, separately by the recipient's board position at the disconnection, from the recipient's own side: slightly behind, $[-100, 0)$ centipawns; roughly equal, $[0, 100)$; ahead, $[100, 300)$; and well ahead, 300 or more. The model uses the headline cells, exposure-date and hour-of-week effects, and controls, adds the band terms and the two stakes, the opponent's forgone win premium and the recipient's half-win stake, each centered and scaled by five rating points, and instruments receipt interacted with the four bands and with the two stakes, using the score interacted with the same variables; the score is omitted as an ordinary control, and only the band estimates are reported. The well-ahead-minus-near-equal contrast compares the top band with the pooled near-equal bands using the joint covariance. Board position and the stakes are determined after opponent assignment, so this heterogeneity model carries assumptions of its own and does not inherit the headline model's identification, which deliberately omits those conditions; its intervals are nominal. The smallest conditional first-stage F among the six instrumented columns is 77.

Table D.8: The act and the result

Comparison	Estimate	SE	95% interval
A. Later granting after ordinary draws against ordinary losses, no choice coded zero (pp); overlap weighting unless stated; treated rows (later grants) in brackets			
Ordinary draws vs ordinary losses [111,983 (2,182)]	0.043	0.041	$[-0.037, 0.122]$
same, exact time-control matching [111,590 (2,167)]	0.004	0.039	$[-0.072, 0.081]$
same, within recipient [54,453 recipients]	0.034	0.068	$[-0.098, 0.166]$
B. Receipt effect by the recipient's position at the disconnection, instrumented (pp); 3,103,280 spaced exposures			
Slightly behind, $[-100, 0)$	0.441	0.191	$[0.066, 0.816]$
Roughly equal, $[0, 100)$	0.229	0.156	$[-0.076, 0.535]$
Ahead, $[100, 300)$	0.481	0.227	$[0.036, 0.927]$
Well ahead, 300 or more	0.431	0.219	$[0.001, 0.861]$
Well ahead minus near equal	0.105	0.251	$[-0.387, 0.597]$

Notes: Panel A: overlap-weighted comparisons of ordinary draws with ordinary losses, 1,290,237 games in all, positive meaning the first-named group grants more later; brackets give treated rows and, in parentheses, their later grants. Panel B: the 3,103,280 spaced exposures, instrumented with the all-time score interacted with the same variables; the model also includes receipt interacted with the opponent's forgone win premium and with the recipient's half-win stake, instrumented the same way, whose estimates are not reported. Spaced sample: 3,103,280 exposures from 1,835,628 recipients with 83,907 granted draws.

D.5 Receipt by kind of later granting

Section 3.6 says what receipt should raise. The two outcomes are a grant at the recipient's first later competitive-position opportunity and a grant at its first later clearly-worse opportunity, each within 90 days, with zero assigned when no such opportunity arrives; shares are relative to the untreated base on the same frame. Shares are the comparison the framework makes informative: at these base rates, a rise in construal or a fall in the weight on rating points scales both grant probabilities by about the same proportion, while a rise in the value of a deserved result moves merited granting alone.

Panel A of Table D.9 uses the primary frame. Panel B extends the instrument to the 14,676,322 exposures, from 2,199,975 recipients, in which the recipient's position was clearly lost, so that the draw received was an unmerited concession; there the first stage is 1.72 points per score unit (0.03; $F = 3{,}065$), and the opponent-environment measures are missing for most exposures and enter through their support indicator.

Receipt raises later merited granting by about a fifth of its base after a merited concession and, imprecisely, after an unmerited one. Its effect on unmerited granting is smaller as a share of its base in every design, but the instrumental-variables difference in shares is not significant ($p = 0.27$ and 0.32). Receipt also lowers the chance of reaching a later opportunity, by more for clearly-worse opportunities and by far more on the clearly-worse frame (the reach rows), so the zero-coded outcomes combine the arrival of an opportu-

nity with the decision at it; the arrival part is small, since at the untreated conditional grant rates a 0.70-point fall in reach moves the zero-coded outcome by about 0.005 points on the primary frame and the larger falls on the clearly-worse frame by about 0.03. A likely mechanism for the reach pattern is the conduct record: a claim marks the absent player's account and, when that player was down four or more points of material, moves the conduct counter on which the matchmaker pairs (Section 2.2), so claim recipients are paired with players who leave games more often, and more opportunities arrive, most of all clearly-worse ones and most of all on the frame where material deficits are the norm. Panel B is therefore reported for completeness and not read causally.

Table D.9: Receipt by kind of later granting

	Later merited granting	Later unmerited granting	Difference
Panel A. Merited receipt: primary frame, 9,736,182 exposures			
Untreated base (%)	1.706	0.498	
Two-stage least squares (pp per receipt)	0.325 (0.056)	0.057 (0.033)	−0.268 (0.061)
95% interval	[0.214, 0.435]	[−0.008, 0.121]	
Share of base (%)	19.0 (3.3)	11.4 (6.6)	−7.7 (6.9), $p = 0.27$
Matched (CEM)	0.308 (0.030)	0.039 (0.015)	
Least squares	0.270 (0.028)	0.034 (0.014)	
Reach an opportunity of that kind (IV, pp)	−0.161 (0.177)	−0.697 (0.163)	
Panel B. Unmerited receipt: clearly-worse frame, 14,676,322 exposures			
Untreated base (%)	1.100	0.402	
Two-stage least squares (pp per receipt)	0.297 (0.158)	−0.007 (0.106)	−0.304 (0.184)
95% interval	[−0.012, 0.606]	[−0.215, 0.202]	
Share of base (%)	27 (14)	−2 (26)	−29 (29), $p = 0.32$
Matched (CEM)	0.198 (0.045)	0.052 (0.030)	
Least squares	0.163 (0.044)	0.036 (0.030)	
Reach an opportunity of that kind (IV, pp)	−1.428 (0.659)	−5.284 (0.631)	

Notes: Standard errors in parentheses, shared-account-or-date covariance with the scores held fixed. Panel A: the primary rows, controls, and fixed effects of Table 6; first stage 7.25 points per score unit, $F = 91{,}156$. Panel B: exposures in which the recipient's position was clearly lost (evaluation at most −300 centipawns), built under the primary design's eligibility, calendar, and score rules, from 2,199,975 recipients; first stage 1.72 (0.03), $F = 3{,}065$. Outcomes are a grant at the recipient's first later opportunity of each kind within 90 days against a different opponent, zero when none arrives; shares divide the estimate by the untreated base on the same frame, and differences use the joint covariance. Matched rows are coarsened exact matching within cells of month, exact time control, mean rating in 200-point bins, signed rating gap in 100-point bins, evaluation bin, and the sign of the draw payoff, with treated weights of one and control weights capped at two. Reach is an indicator for any opportunity of that kind within 90 days.

D.6 Receipt and the Patron badge

The outcome is the recipient's Patron badge in the August 2026 profile snapshot of Appendix E: a stock observed 382 to 940 days after the exposure (median 655), not the adoption of support. Profiles exist for the accounts of the original acquisition cohort, so the sample is the 6,125,535 primary exposures whose recipients were acquired and returned a profile, from 794,035 recipients, 2,476 of whom hold the badge. Rows are weighted by the inverse of each account's original sampling probability: one for the accounts that ever granted, and the cell's selected-control share for the controls, with a maximum weight of 19.3. The weights restore the acquisition design within its support but cannot represent the 413,818 exposures, 4.3 percent, whose recipients were outside the original chooser universe.

Rows, controls, fixed effects and covariance are those of the primary model; on these rows the first stage is 7.25 points per score unit ($F = 55{,}504$), as in the full frame. The reduced form is -0.0015 points per score unit (standard error 0.0015, $p = 0.32$); per receipt it is -0.021 (0.021), with a 95 percent interval of -0.062 to 0.020 on a base of 0.165 percent (Table D.10). The score does not predict whether a recipient's profile was returned (-0.002, standard error 0.013) and predicts membership in the acquisition universe only slightly (-0.017, 0.006, on a base of 95.7 percent). No alternative weighting detects an effect; the recipient-weighted estimate leans negative (-0.076, 0.045). Under the primary weights the interval excludes a lasting increase above about a tenth of the base, but the design cannot date adoption or rule out a temporary effect.

Table D.10: Receipt and the Patron badge

	Exposure IPW (primary)	Unweighted	Recipient IPW	Prior choosers, IPW
Panel A. Estimates				
Exposures	6,125,535	6,125,535	6,125,535	4,999,144
Recipients (Patrons)	794,035 (2,476)	794,035 (2,476)	794,035 (2,476)	666,215 (2,090)
Patron base (%)	0.165	0.186	0.234	0.176
First stage (pp per score unit)	7.25 (0.03)	7.35 (0.03)	7.00 (0.05)	7.32 (0.03)
First-stage F	55,504	73,946	17,704	61,246
Reduced form (pp per score unit)	−0.0015 (0.0015)	−0.0004 (0.0017)	−0.0053 (0.0031)	0.0000 (0.0018)
Two-stage least squares (pp per receipt)	−0.021 (0.021)	−0.006 (0.022)	−0.076 (0.045)	0.000 (0.025)
95% interval	$[-0.062, 0.020]$	$[-0.050, 0.038]$	$[-0.164, 0.012]$	$[-0.049, 0.050]$
Panel B. Selection and coverage, full frame of 9,736,182 exposures				
Score and membership in the acquisition universe (base 95.7%)	−0.0165 (0.0061) pp per score unit			
Score and a returned profile (base 62.9%)	−0.0016 (0.0133) pp per score unit			
Observed exposures (recipients)	6,125,535 (794,035)			
Eligible, not sampled; represented by the weights	3,196,126 (814,105)			
Outside the acquisition universe; no support	413,818 (269,155)			
Selected, profile not returned	703 (100)			
Days from exposure to snapshot: minimum, median, maximum	382, 655, 940			

Notes: All columns use the primary model's rows, controls, fixed effects, and shared-account-or-date covariance, with the scores held fixed; units of the score as in Table 6. Weights are inverse original sampling fractions: accounts that ever granted have probability one and controls their cell's selected share, with a maximum weight of 19.3. Recipient IPW divides each row's weight by the recipient's exposure count in the full frame. The prior-chooser column keeps exposures whose recipient had a competitive choice in the preceding 90 days. The badge can be bought, gifted, or lapse, and the snapshot does not show when it began.

E Same-Platform Supporter Marker and Proofs

This appendix documents the outcome, the acquisition of the matched cohort, and the specifications behind Section 5.2, and closes with the proofs for Section 3. Two comparisons cited in the main text are reported here.

First, a separate tenure sensitivity keeps only opportunities faced after an account was at least 365 days old, with account age measured at each game's start, so an account can enter partway through the panel. Grant-history indicators and opportunity counts are rebuilt from the retained opportunities, and accounts need at least four retained competitive and four retained clearly-worse opportunities, which leaves 462,754 accounts in 24 matching cells. The model includes matching-cell fixed effects, mean rating, the logarithms of competitive and clearly-worse opportunity counts, and the logarithm of account age at baseline, a smaller control set than Table 4, Panel B. The associations are $+0.604$ percentage points for competitive-position granting (standard error 0.044; 95 percent interval 0.512 to 0.695) and -0.272 for clearly-worse granting (0.046; -0.366 to -0.177), with standard errors clustered by matching cell and intervals from the t distribution with 23 degrees of freedom. The signs match the main specification; because the controls differ, the difference in magnitudes does not isolate tenure.

Second, the recipient-side comparison: within chooser, and with the controls of Appendix C.2, opponents who hold the badge in the August 2026 snapshot are granted draws 0.319 percentage points more often in competitive positions (standard error 0.112; 95 percent interval 0.099 to 0.538) and -0.001 points in clearly-worse positions (0.055; -0.109 to 0.108), no detectable association. The model pools the 22,496,521 competitive- or clearly-worse-position opportunities whose disconnected opponent has an observed badge status in the August 2026 snapshot, with chooser fixed effects, and its covariance allows dependence among opportunities that share an account or a date; it is not restricted by account age. The two state coefficients are not compared formally. Badge possession is observed in the snapshot rather than at the game, so neither comparison identifies a response to the badge as displayed.

E.1 Outcome and acquisition

The outcome is possession of the Patron badge, read from public account profiles queried in August 2026. Status can be purchased, can be gifted by another account, and can lapse, so the profile records whether an account holds the badge at the moment of the query and shows neither who paid, nor how much, nor when support began. The comparison relates that later status to granting behavior during the November 2023 to October 2025

panel.

The acquisition cohort was fixed before any profile was queried. Every account with at least one granted draw in any board state in the analysis sample is a granting account, and controls are accounts with no granted draw in any board state in that sample, including the modestly-worse region.

Allocation proceeded within matching cells, which cross ten bins of competitive-position opportunities in the panel, 0, 1, 2, 3, 4, 5–9, 10–20, 21–50, 51–100, and 101 or more, with the account's modal game speed over all its opportunities in the panel, ties broken by the alphabetical order of the speed labels. Within each cell, granting and control accounts were separately ordered by a fixed-seed SHA-256 hash of their identifiers, and controls were allocated to granting accounts without replacement, up to three per granting account. Accounts that later proved unavailable are not replaced. August outcomes are unavailable for 30 granting accounts and 177 controls, and a further 90 controls belong to the 30 matching groups whose focal granting account has no observed profile, which the main analysis excludes; the analysis sample contains 212,828 granting accounts and 638,298 controls.

That cohort is broader than the comparison built on it. It includes the zero-opportunity bin, whereas the main comparison of Section 5.2 uses only the matching groups whose focal granting account granted in a competitive position, and those occupy the 36 cells that cross the nine positive-opportunity bins with the bullet, blitz, rapid, and classical/long speeds.

E.2 Specification and sensitivity

All Patron regressions are unweighted account-level linear probability models with matching-cell fixed effects. The fixed effects are at the level of the cell, not of the matched group, and no weights are applied to equalize control ratios across cells.

The full control set comprises two groups that differ in when they were measured. Panel-measured controls are the account's competitive and total opportunity counts; its board-state composition and mean evaluation; its rating-consequence composition and average draw payoff and forgone premium; its active opportunity months; its speed composition; and its mean rating, rating variability, and rating tier, all taken from the behavioral panel. Profile-based controls are account age and recency as of the profile query, reported playtime, current game counts by speed, and current speed-specific ratings including correspondence ratings. Because these are measured after the granting behavior, they provide descriptive adjustment and do not establish that the groups were comparable beforehand. Missing profile covariates are imputed with within-cell medians, with a

global-median fallback and supported missingness indicators, and constant or collinear terms are dropped.

Table E.1 reports the main estimates with two standard errors. HC1 is the principal inference. Clustering by matching cell is a sensitivity, and with 36 clusters it is not treated as resolving finite-cluster inference.

Table E.1: Main Patron comparison: estimates

	Estimate (pp)	HC1 SE	Cell SE
Raw gap in Patron rates	0.312		
Adjusted association, cell fixed effects and full controls	0.277	0.019	0.035

Notes: The main comparison's 851,126 accounts in 36 cells.

E.3 State-specific granting

The state-specific comparison of Section 5.2 uses a different sample from the main comparison. It is drawn from the broader available-profile acquisition cohort of accounts with grants in any state and their selected controls, 1,105,629 accounts in the August snapshot, and it does not impose the main comparison's requirement that the focal granting account's profile be available. Within that cohort it retains the 599,045 accounts with at least four competitive-position and at least four clearly-worse-position opportunities, which occupy 24 matching cells.

Indicators for ever granting in a competitive position and for ever granting in a clearly-worse position enter a single regression, together with matching-cell fixed effects, the full control set, and $\log(1 + \text{clearly-worse opportunities})$ alongside the competitive and total opportunity controls. An account that granted in both states satisfies both indicators. The reported difference is the competitive-position coefficient minus the clearly-worse-position coefficient; its uncertainty is computed from their joint covariance.

Table E.2 reports the estimates with both standard errors and the four raw groups. In those groups, "only" means at least one grant in the named state and none in the other, and "neither" means no grant in either named state. Grants in modestly-worse positions do not affect the classification, so an account in the "neither" group may have granted in the middle region and need not belong to the never-granter control group of the main comparison. The association therefore varies with the circumstances of granting as well as with its occurrence. This is a descriptive comparison: it establishes no causal direction, and it does not rule out attachment to the platform that is itself related to where an account grants.

Table E.2: Patron status and granting by board state

Panel A. Joint model, 599,045 accounts (percentage points)	Estimate	HC1 SE	Cell SE (24 cells)
Ever granted in a competitive position	0.388	0.028	0.032
Ever granted in a clearly-worse position	−0.195	0.031	0.033
Competitive minus clearly worse	0.583	0.046	0.059
Panel B. Raw Patron rates by granting history	Patrons	Accounts	Rate (%)
Granted only in competitive positions	908	100,671	0.902
Granted in both states	89	18,098	0.492
Granted in neither state	1,957	454,045	0.431
Granted only in clearly-worse positions	68	26,231	0.259

Notes: Accounts in the broader cohort with available profiles and at least four opportunities in each state. Panel A: both indicators in one model with cell fixed effects, the full controls, and $\log(1 +$ clearly-worse opportunities$)$; the difference uses the joint covariance. Panel B: unadjusted rates.

E.4 Proofs for Section 3

Lemma 1. Subtract the two lines of (2).

Proposition 1. $P_{io} = c_i F(x_{io})$, which is zero when $c_i = 0$, with $x_{io} = b_i m(\pi_o) - \gamma s_o - \lambda_i \kappa_o - a$ and F strictly increasing. Part (i): the three variables enter x_{io} only through $m(\pi_o)$; π is increasing in each and m is nondecreasing, strictly increasing above $\underline{\pi}$ and constant below. Part (ii) is immediate from the signs of the coefficients in x_{io}.

Proposition 2. Fix a κ cell and an investment level s, and write $e = \gamma s$. For an account with traits (c_i, b_i, λ_i), the probability of a grant at a lost opportunity is $c_i F_L(\lambda_i)$ with $F_L(\lambda) = F(-a - e - \lambda\kappa)$, decreasing in λ and constant in b, and at a competitive opportunity it is $c_i F_C(b_i, \lambda_i)$ with $F_C(b, \lambda) = F(-a - e - \lambda\kappa + b\mu)$, $\mu = m(\pi) > 0$, increasing in b and decreasing in λ. Write $q = \Pr(c_i = 1) > 0$. The badge probability is $H(b, \lambda) = \Pr(\nu_i \geq -p_0 - \tau_b b - \tau_\lambda \lambda)$, nondecreasing in each argument and strictly increasing in an argument with a positive loading. Because c_i is independent of (b_i, λ_i, ν_i) and absent from H, $\operatorname{Cov}(S_i, G_i^L) = q \operatorname{Cov}(H, F_L)$ and $\operatorname{Cov}(S_i, G_i^C) = q \operatorname{Cov}(H, F_C)$. Lost-only monotonicity, the condition of the second ranking in part (iii), is that $M(\lambda) = F_L(\lambda) \operatorname{E}_b[1 - F_C(b, \lambda)]$, the probability that a construing account grants at the lost opportunity and claims at the competitive one, averaged over b_i, is nonincreasing in λ; greater weight on rating points lowers both grant probabilities, and the condition says that its effect on the lost grant is not outweighed by its effect on the competitive claim.

Proof. Part (i). Decompose

$$\operatorname{Cov}\big(H(b_i, \lambda_i), F_L(\lambda_i)\big) = \operatorname{E}_b\big[\operatorname{Cov}_\lambda(H(b, \lambda_i), F_L(\lambda_i))\big] + \operatorname{Cov}_b\big(\operatorname{E}_\lambda H(b, \lambda_i), \operatorname{E}_\lambda F_L(\lambda_i)\big).$$

The second term is zero because $\mathrm{E}_\lambda F_L$ does not depend on b. In the first, $H(b,\cdot)$ is strictly increasing and F_L strictly decreasing in λ, so the covariance is negative.

Part (ii). With $\tau_\lambda = 0$, H depends on b alone and the decomposition over b within λ gives a positive sign, both functions being strictly increasing in b. With $\tau_b = 0$, the decomposition over λ within b gives a negative sign. With both loadings positive, the within-λ term is positive and the between-λ term negative.

Part (iii). Both grant probabilities lie strictly between zero and one because F has full support. First condition on $c_i = 1$, $G_i^C = r$, and $\lambda_i = \lambda$, for $r \in \{0,1\}$. The conditional density of b_i is proportional to $\phi_b(b)F_C(b,\lambda)$ when $r = 1$ and to $\phi_b(b)[1 - F_C(b,\lambda)]$ when $r = 0$. Both F and $1 - F$ are log-concave because the density is [Bagnoli and Bergstrom, 2005], and each weighting function is such a function evaluated at $b\mu - \lambda\kappa - a - e$, so its logarithm has increasing differences in (b,λ): $\partial^2 \log F_C / \partial b\, \partial\lambda = -\mu\kappa(\log F)'' \geq 0$, and likewise for $\log(1 - F_C)$. As λ rises, the conditional distribution of b_i therefore rises in the likelihood-ratio order, and since H is nondecreasing in b and strictly increasing in λ, $h_r(\lambda) = \mathrm{E}[H(b_i,\lambda) \mid c_i = 1, G_i^C = r, \lambda_i = \lambda]$ is strictly increasing. Given (c_i, G_i^C, λ_i), the lost grant G_i^L depends only on its own shock, so it is independent of b_i and ν_i and occurs with probability $F_L(\lambda_i)$, which is strictly decreasing. Hence

$$\mathrm{Cov}(S_i, G_i^L \mid c_i = 1, G_i^C = r) = \mathrm{Cov}\big(h_r(\lambda_i), F_L(\lambda_i) \mid c_i = 1, G_i^C = r\big) < 0,$$

the conditional distribution of λ_i being nondegenerate, so that

$$\mathrm{E}[S_i \mid c_i = 1, G_i^C = r, G_i^L = 0] \;>\; \mathrm{E}[S_i \mid c_i = 1, G_i^C = r, G_i^L = 1]. \tag{9}$$

For $r = 1$ the conditioning on $c_i = 1$ is redundant, since a competitive grant requires it, and (9) is the first ranking.

For the second ranking, accounts that grant only in lost positions construe the choice, and their traits are distributed as the population density times $W(b,\lambda) = F_L(\lambda)[1 - F_C(b,\lambda)]$, so

$$\mathrm{E}[S_i \mid G_i^C = 0, G_i^L = 1] - \mathrm{E}[H] = \frac{\mathrm{Cov}(H,W)}{\mathrm{E}[W]},$$
$$\mathrm{Cov}(H,W) = \mathrm{E}_\lambda\big[\mathrm{Cov}_b(H,W \mid \lambda)\big] + \mathrm{Cov}_\lambda\big(\mathrm{E}_b[H \mid \lambda],\, M(\lambda)\big),$$

with $M(\lambda) = \mathrm{E}_b[W \mid \lambda] = F_L(\lambda)\,\mathrm{E}_b[1 - F_C(b,\lambda)]$ and the moments taken under the independent population distribution of (b_i,λ_i). The first term is nonpositive because, at a given λ, H is nondecreasing and W nonincreasing in b; the second is nonpositive because $\mathrm{E}_b[H \mid \lambda]$ is increasing in λ and lost-only monotonicity makes M nonincreasing. Thus

$\mathrm{E}[S_i \mid G_i^C = 0, G_i^L = 1] \leq \mathrm{E}[H] = \mathrm{E}[S_i \mid c_i = 0]$, the equality because c_i is independent of the other traits and absent from the badge index. The accounts that never grant are a mixture of those with $c_i = 0$, whose badge rate is $\mathrm{E}[H]$, and construing accounts with $G_i^C = G_i^L = 0$, whose badge rate exceeds the lost-only rate strictly by (9) with $r = 0$; the second group has positive probability since $q > 0$ and both grant probabilities are below one. The mixture's badge rate therefore exceeds the lost-only rate strictly. When $q = 1$ the first group is empty and the condition is not used. □

Sufficient conditions for lost-only monotonicity. Write $x_L = -a - e - \lambda\kappa$, $x_C = x_L + b\mu$, and $f = F'$. Since $\partial_\lambda \log\{F_L(\lambda)[1 - F_C(b, \lambda)]\} = \kappa\,[\,f(x_C)/(1 - F(x_C)) - f(x_L)/F(x_L)\,]$, the condition holds whenever $f(x_L)/F(x_L) \geq f(x_C)/(1 - F(x_C))$ for every type. Because the reverse hazard f/F of a log-concave distribution is decreasing and $x_L \leq x_C$, it suffices that $F(x_C) \leq 1/2$ for every type: no construing account grants a competitive opportunity with probability above one half. For logistic F the pointwise condition is $F(x_L) + F(x_C) \leq 1$, and an averaged condition suffices: with $p_L = F(x_L)$ and $p_C = F(x_C)$, $\partial_\lambda \log M(\lambda) = \kappa\,\{\mathrm{E}_b[p_C(1 - p_C)]/\,\mathrm{E}_b[1 - p_C] - (1 - p_L)\}$, where the first term is at most $\mathrm{E}_b[p_C \mid \lambda]$ and $a > 0$ gives $p_L < 1/2$, so $\mathrm{E}_b[p_C \mid \lambda] \leq 1/2$ at every λ suffices even if some types grant almost every competitive opportunity.

Histories. The groups of Table 4, Panel C, classify accounts by whether they ever granted in each region. With n_C competitive and n_L lost opportunities at fixed conditions and independent shocks, the ever-grant probabilities of a construing account are $F_{n_C}(x_C)$ and $F_{n_L}(x_L)$, where $F_n(x) = 1 - [1 - F(x)]^n$ has density $nf(x)[1 - F(x)]^{n-1}$, log-concave because f and $1 - F$ are, so that F_n and $1 - F_n$ are log-concave as well. The proof goes through with $F_{n_C}(x_C)$ and $F_{n_L}(x_L)$ in place of F_C and F_L, and lost-only monotonicity is then a condition on $F_{n_L}(x_L)\,\mathrm{E}_b[1 - F_{n_C}(x_C)]$. The proposition compares accounts at fixed opportunity counts and conditions; the raw groups of Panel C pool both, so their ordering is consistent with the proposition rather than implied by it.

E.5 Model comparison: restrictions and derivations

This subsection states the restrictions behind Section 7 for explicit specifications. It compares theoretical restrictions with results already reported; no model is estimated. For a fixed chooser i, let z describe the board, the clocks, and the stage, and let X hold the pregame rating states and both players' rating changes under claiming and granting, which are fixed before z evolves. Let $I_i(X, z)$ be the utility of granting minus that of claiming, net of the shocks; with $\varepsilon = \eta^C - \eta^G$ and distribution function F, a construing account grants with probability $F(I_i(X, z))$, as in Section 3.

State invariance. Suppose $U_i(a) = V_i(x_a) + W_i(a) + \eta^a$, where x_a is the two players' realized rating payoffs under action a and neither V_i nor W_i depends on z. Then the index $V_i(x_G) - V_i(x_C) + W_i(G) - W_i(C)$ is constant in z, so at a fixed chooser and fixed payoffs the grant probability does not depend on the board, the clocks, or the stage, whatever the dimension of the preference heterogeneity. Linear altruism toward the opponent's rating change, $I = -\lambda_i\kappa + \alpha_i\kappa_j - a$; a state-independent warm glow, $I = \omega_i - \lambda_i\kappa - a$ [Andreoni, 1990]; and inequity aversion over the two players' rating changes [Fehr and Schmidt, 1999] are all of this form. Each can respond to the price κ; none can produce the position gradient. A self-image payoff is of this form only if the inference the act supports does not depend on the board; an image of fairness does, and is then a reading of b_i.

Loss aversion around the pre-game rating. With $v(x) = x$ for $x \geq 0$ and $\ell_i x$ for $x < 0$, the cost of granting is κ when the draw gains points and $\kappa + (\ell_i - 1)(-r_G)$ when it loses them, continuous at zero with a change of slope. It has no board argument, so it cannot produce the position gradient, and Section 4.5 finds neither the level difference nor the growing cost it implies.

A continuation reference. Let the chooser evaluate the result against the lottery anticipated before the disconnection, a win with probability $1 - \pi$, a draw or a loss otherwise, with consumption utility $\lambda_i r$ and gain–loss utility $\eta_i u$ for gains and $\eta_i \ell_i u$ for losses, $\eta_i \geq 0$, $\ell_i \geq 1$ [Kőszegi and Rabin, 2006]. Relative to the anticipated win, granting is a loss of $\lambda_i\kappa$ and claiming is not; relative to an anticipated draw or loss, both are gains and differ by $\lambda_i\kappa$. Averaging over the reference lottery, with a constant benefit ω_i from giving,

$$I_i^R = \omega_i - a - \lambda_i\kappa\big[1 + \eta_i\{1 + (\ell_i - 1)(1 - \pi)\}\big], \tag{10}$$

which rises in π when $\lambda_i\kappa\eta_i(\ell_i - 1) > 0$, has no plateau for small π, and has no direct dependence on investment at fixed π; since a given advantage is harder to overturn late in a game, π rises with the stage where the opponent is ahead, so this index alone predicts more granting late, and it needs the same separate investment term as the framework to produce the stage decline of Section 4.4. At fixed κ it is affine in π and observationally like a smooth desert index. The two differ in one cross-partial: $\partial^2 I_i^R / \partial\pi\,\partial\kappa = \lambda_i\eta_i(\ell_i - 1) > 0$, whereas the desert index $b_i m(\pi) - \gamma s - \lambda_i\kappa - a$ has none. The paper conditions on κ and does not observe π, so its regressions do not separate the two.

A common trait. Fix an opportunity history X. Among construing accounts, let $q_C(\theta)$ and $q_L(\theta)$ be the probabilities of ever granting in competitive and in clearly-lost positions and $h(\theta)$ the badge probability, all nondecreasing in a scalar trait θ, with the two grant histories and the badge independent conditional on θ and X, and let construal be a gate with any dependence on θ.

Proposition 3 (Monotone-factor ordering). *Among accounts that grant in competitive positions, at fixed X,* $\mathrm{E}[S \mid G^C = 1, G^L = 1, X] \geq \mathrm{E}[S \mid G^C = 1, G^L = 0, X]$.

Proof. A competitive grant reveals that the gate is open. Let Q be the posterior of θ given $G^C = 1$ and X. By conditional independence, the difference equals $\mathrm{Cov}_Q(h, q_L)/[\bar{q}(1-\bar{q})]$ with $\bar{q} = \mathrm{E}_Q[q_L]$, and for independent draws θ, θ' from Q, $2\,\mathrm{Cov}_Q(h, q_L) = \mathrm{E}_Q[(h(\theta) - h(\theta'))(q_L(\theta) - q_L(\theta'))] \geq 0$ because both functions are nondecreasing. □

A lost-position grant is further favorable evidence about the same trait, so the badge rate cannot fall. Table 4, Panel C, shows 0.49 percent among accounts that granted in both regions against 0.90 among competitive-only granters; the pooled groups do not hold X fixed, so the comparison is not the exact conditional test, but together with the opposite signs in Panel B it bears against a common trait. The persistence correlations do not supply this discrimination: with θ uniform on $[0,1]$, $q_C = \theta$ and $q_L = \theta^k$, the two rates correlate $\sqrt{3(2k+1)}/(k+2)$, which is 0.42 at $k = 30$, inside the paper's 0.40 to 0.45, and scaling both probabilities down to make granting rare leaves the correlation unchanged. The same monotone trait would also raise the badge rate of lost-only granters above that of accounts that never grant when every account construes the choice; Panel C shows 0.26 against 0.43. A single trait that raised both kinds of granting but lowered badge holding would put the both-kind group below the competitive-only group, as the data show, but when every account construes the choice it would also put competitive-only granters below accounts that never grant, and Panel C has them at 0.90 against 0.43.

Two heterogeneous weights. Proposition 2 gives the reverse ordering under the framework. It is not unique in that. The continuation-reference model with independently heterogeneous ω_i and λ_i, and a badge probability increasing in both, penalizes λ_i more in a lost position (where π is small) than in a competitive one, and so can also produce opposite badge signs and the full ordering of the four granting histories. Two illustrative populations make the point exactly; each has $c = 1$ with probability 0.2, two equally likely values of each trait drawn independently, standard logistic choice and badge shocks, one opportunity in each region, and $\kappa = 1$. In the desert population, $b \in \{0, 3\}$, $\lambda \in \{0, 2\}$, $a = 3$, $s = 0$, $m_C = 1$, $m_L = 0$, and the badge probability is $\Lambda(-8.5 + b + 0.5\lambda)$. In the continuation-reference population, $\omega \in \{0, 3\}$, $\lambda \in \{0.5, 1.5\}$, $a = 1$, $\eta = 1$, $\ell = 5$, $\pi_C = 0.75$, $\pi_L = 0$, and the badge probability is $\Lambda(-8 + \omega + \lambda)$. Table E.3 reports the implied rates, computed exactly by summing over the trait combinations and the gate; both populations order the four histories as the data do and give the badge regression opposite signs. The levels are illustrations, not calibrations.

What the comparison establishes. Models that value only the rating outcomes cannot produce the board patterns, and a common monotone trait cannot produce the badge

Table E.3: Two illustrative populations

	Desert	Continuation reference
Badge rate, competitive only (%)	0.503	1.193
Badge rate, both (%)	0.395	1.091
Badge rate, neither (%)	0.392	1.055
Badge rate, lost only (%)	0.214	0.966
Competitive grant rate (%)	3.367	3.891
Lost grant rate (%)	0.541	1.439
Badge on the competitive indicator (pp)	$+0.113$	$+0.137$
Badge on the lost indicator (pp)	-0.162	-0.096

Notes: The populations are specified in the text. The last two rows are the population linear projection of the badge on a constant and the two grant indicators.

ordering. Both the framework and the continuation-reference model have two traits and match the signs; the framework's hard threshold misses the rise in granting within the lost region, which the continuation reference allows. The comparison does not identify a unique motive, and the receipt estimates, which concern a change in behavior, lie outside these static models.

F Persistence, Receipt, and the Instrument: Additional Results

This appendix reports, for Sections 5 and 6 and Appendix D, the persistence cohort and the persistence comparisons by board state, the interval profile of the receipt association, the instrumental-variables results across exposure specifications, the trait-by-trait balance coefficients behind the joint tests, the first stages by subgroup, and the competitive-only score.

F.1 Persistence across twelve-month periods

The persistence comparisons of Section 5.1 use competitive-position opportunities at all game stages, divided between the panel's two twelve-month periods, November 2023 through October 2024 and November 2024 through October 2025. For account i, let n_{i1} and n_{i2} be its competitive opportunities in the two periods and k_{i1} and k_{i2} its granted draws among them. The cohort retains accounts with $n_{i1} \geq 1$, $n_{i2} \geq 1$, and $n_{i1} + n_{i2} \geq 4$, so four is the minimum total across both periods rather than a minimum in either, and accounts that never grant remain. It contains 666,131 accounts with 10,071,687 competitive opportunities and 245,374 granted draws across both periods. Each account's rate in period y is $p_{iy} = k_{iy}/n_{iy}$.

The prior-grant groups divide the cohort by whether $k_{i1} = 0$ or $k_{i1} \geq 1$, so "none" refers to competitive-position grants in the first period only. Each group's second-period rate pools grants and opportunities, $100 \times \sum_i k_{i2} / \sum_i n_{i2}$, and therefore weights accounts by n_{i2} rather than equally; the ratio of the two rates, computed from the unrounded rates, is 16.865. Because the classification depends on whether any grant occurred in the first period, it depends partly on how many opportunities an account had: at a given underlying rate, an account with more first-period opportunities is more likely to be classified as a granter. Standard errors for these correlations and rates come from the account bootstrap described in the notes to Table 3.

Three further comparisons calibrate the persistence estimates. Among accounts with at least one first-period competitive-position opportunity, 65.9 percent of those with a first-period competitive-position grant return for a competitive opportunity in the second period, against 58.3 percent of those without, so the cohort is selected on return and the correlations describe returning accounts. A random split of each account's competitive opportunities into two halves gives a split-half correlation of 0.600 unweighted and 0.690 with opportunity weights, a benchmark for persistence when the same observed opportunities are split at random; the year-to-year correlations of 0.545 and 0.618 are about nine tenths of these values. Adjusting each opportunity's outcome for its board

state, game stage, and speed with a leave-one-out cell model and correlating the adjusted rates across periods gives 0.527 among established returning accounts, the same as the unadjusted correlation in that subsample.

F.2 Persistence by board state

Table 3, Panels B and C, reports persistence separately by board state on the balanced cohort of accounts with at least four competitive-position and at least four clearly-worse-position opportunities in each twelve-month period. This subsection reports the same comparisons at thresholds of one and ten opportunities per state and period, in games ending after ply ten, and with all noncompetitive positions (evaluation below -100 centipawns) in place of the clearly-worse region, together with the retention of each first-year group and regressions of second-year rates on first-year granting. Grants in the modestly-worse region do not enter the four-group classification in the clearly-worse comparisons. Period definitions and the evaluation convention are as in Section 5.1; the correlations in this subsection are reported without standard errors.

Table F.1 reports the four cross-year correlations and the implied correlation of the two propensities, which divides each cross-state correlation by the geometric mean of the two own-state correlations. Raising the threshold raises every correlation, as the rates become less noisy, while the implied correlation stays between 0.40 and 0.46 in every sample; with all noncompetitive positions in place of the clearly-worse region, where the second state includes the modestly-worse positions in which granting is intermediate, it is 0.59. The persistence of selectivity itself, measured as each account's competitive-minus-clearly-worse rate difference, is 0.458, 0.689, and 0.766 at the three thresholds, against 0.533, 0.727, and 0.790 for the mean of the two rates.

Among the 266,399-account cohort, the share of each first-year group that is eligible in the second year, with at least four opportunities in each state, is 50.2 percent for first-year non-granters, 51.7 for competitive-only granters, 51.9 for lost-only granters, and 50.9 for accounts that granted in both states.

Table F.2 reports the regressions. In the indicator model, a first-year competitive-position grant raises the second-year competitive rate by 10.95 percentage points (on a mean of 1.94) and the second-year clearly-worse rate by 0.35 points (on a mean of 0.29), while a first-year clearly-worse grant raises them by 0.92 and 0.51 points; in units of each outcome's own mean, first-year competitive granting predicts second-year competitive granting more than second-year clearly-worse granting by 4.46 (standard error 0.08), and first-year clearly-worse granting predicts second-year clearly-worse granting more than second-year competitive granting by 1.26 (0.13). In the continuous model, a point of first-

year competitive rate carries 0.70 of a point into the second-year competitive rate and 0.02 into the clearly-worse rate, and a point of first-year clearly-worse rate carries 0.48 into the clearly-worse rate and 0.14 into the competitive rate.

Table F.1: Persistence by board state: cross-year correlations by threshold and sample

Sample	Accounts	C_1, C_2	L_1, L_2	C_1, L_2	L_1, C_2	Implied	Harmonic C_1, C_2
At least 1 per state and year	670,480	0.522	0.300	0.164	0.181	0.41–0.46	0.644
At least 4 (Table 3)	266,399	0.727	0.494	0.239	0.271	0.40–0.45	0.751
At least 10	82,311	0.793	0.616	0.295	0.321	0.42–0.46	0.800
At least 4, games after ply 10	265,901	0.661	0.493	0.236	0.265	0.41–0.46	0.702
At least 4, all noncompetitive positions	278,774	0.725	0.514	0.349	0.376	0.57–0.62	0.749

Notes: C is the account's competitive-position grant rate and L its clearly-worse rate (its rate over all noncompetitive positions in the last row), with subscripts for the two twelve-month periods. Entries are equal-account Pearson correlations; the last column weights accounts by the harmonic mean of their two competitive opportunity counts, and the harmonic-weighted values of the other three correlations are 0.215, 0.240, and 0.442 at threshold one, 0.261, 0.299, and 0.563 at threshold four, and 0.307, 0.349, and 0.657 at threshold ten. The implied column divides each cross-state correlation by $\sqrt{\rho(C_1, C_2)\,\rho(L_1, L_2)}$ and reports the two values; it treats each year as a repeated measurement of a fixed propensity. The after-ply-10 sample retains cohort accounts with positive exposure after ply ten in both states and periods.

Table F.2: Second-year grant rates by state regressed on first-year granting, 266,399 accounts

	Second-year competitive rate	Second-year clearly-worse rate
Panel A. First-year granting as indicators (percentage points)		
Any first-year competitive-position grant	10.953 (0.134)	0.351 (0.015)
Any first-year clearly-worse grant	0.918 (0.067)	0.508 (0.041)
Both, interaction	12.329 (0.634)	4.549 (0.345)
Second-year mean (pp)	1.935	0.293
Panel B. First-year rates (points of second-year rate per point of first-year rate)		
First-year competitive rate	0.702 (0.006)	0.024 (0.002)
First-year clearly-worse rate	0.138 (0.021)	0.485 (0.028)
Panel C. Own-state minus cross-state prediction, in units of each outcome's mean		
First-year competitive granting: competitive minus clearly-worse outcome		4.462 (0.077)
First-year clearly-worse granting: clearly-worse minus competitive outcome		1.262 (0.132)

Notes: Account-level regressions on the balanced threshold-four cohort, 3,285,482 second-year competitive and 4,616,580 second-year clearly-worse opportunities; standard errors in parentheses. Panel A enters indicators for any first-year grant in each state and their interaction; Panel B enters the two first-year rates in percentage points. Panel C divides each first-year indicator's effect on each second-year outcome by that outcome's mean and reports the difference between the own-state and cross-state values, with p-values below 0.001 for both. Both models include the account's standardized second-year speed shares as controls, with the classical share dropped as collinear.

F.3 The interval profile

Table F.3 reports the interval profile, one regression per interval. Each starts from the 2,110,399 exposures with complete follow-up and retains the accounts with a competitive-position choice in the indicated interval after the exposure anchor; its outcome is a grant at the account's first competitive choice within that interval, not at its first choice overall classified by elapsed time. Elapsed time is measured between the recorded last-move timestamps of the index game and the later choice game.

An exposure therefore contributes to every interval in which the account has a choice, whatever it chose or granted in an earlier interval, and accounts without a choice in an interval are excluded from that interval's regression rather than coded zero. The summed observations across intervals, 1,547,761, count overlapping account-interval observations rather than unique recipients. The receipt design's exposure cells, weights, and exposure controls are reused, and standard errors cluster by exposure-opponent account. The first interval is bounded by six hours of clock time and not by a playing session.

Table F.3: Receipt association by interval since the index exposure

Interval since exposure anchor	Estimate (pp)	SE	Observations	Clusters
More than 0 through 6 hours	1.494	0.674	33,431	32,614
More than 6 hours through 1 day	1.142	0.533	45,447	44,011
More than 1 through 7 days	1.077	0.236	256,230	219,694
More than 7 through 30 days	0.686	0.159	533,040	404,671
More than 30 through 90 days	0.501	0.139	679,613	488,195

Notes: Each row is a separate regression on the accounts with a competitive-position choice in the interval, with the exposure cells, weights, and controls of the receipt design (Appendix D); the outcome is a grant at the first such choice within the interval. Clusters are exposure-opponent accounts. Intervals overlap in accounts, and the point estimates do not identify a causal decay rate.

F.4 Instrumental-variables results across exposure specifications

Table F.4 reports the first stage, the reduced forms for later granting and for reaching a choice, and two-stage least squares under eight exposure specifications, of which the first is the spaced design of Appendix D. All use the all-time score with its history controls and the environment controls, retain the 90-day outcome horizon and the different-opponent requirement, and differ only in which exposures enter or which recipient controls are applied. The repeated-recipient subset contains recipients with at least two retained primary exposures, and its pooled and recipient-fixed-effects models use identical rows. The final two models omit the recipient history and volume controls from the spaced and first-exposure rows respectively, leaving the other controls and time effects in place.

The two-stage point estimates are positive throughout, between 0.26 and 0.51. The repeated-recipient estimate with recipient fixed effects, 0.264 (0.140), is the only one whose interval includes zero; its balance test is reported in Appendix F.5. The first-exposure estimates are the least precise. Omitting the recipient history and volume controls changes the spaced estimate from 0.363 to 0.371 and leaves the first-exposure estimate unchanged, so the recipient controls, which the balance tests treat as controlled traits, do not drive the estimate. The comparison under shorter spacing may reuse a later choice as the outcome of more than one exposure.

Table F.4: Instrumental-variables results across exposure specifications

Specification	Exposures	FS (SE)	RF: grant (SE)	RF: reach (SE)	2SLS (SE)	95% interval
Spaced baseline: 90-day spacing, all recipients	3,103,280	7.12 (0.03)	0.026 (0.007)	0.000 (0.023)	0.363 (0.096)	[0.174, 0.552]
Repeated recipients, pooled	1,967,871	7.34 (0.04)	0.025 (0.009)	-0.002 (0.028)	0.338 (0.121)	[0.102, 0.575]
Repeated recipients, recipient fixed effects	1,967,871	7.38 (0.04)	0.019 (0.010)	0.015 (0.032)	0.264 (0.140)	[−0.009, 0.538]
30-day spacing	4,537,915	7.20 (0.03)	0.029 (0.006)	-0.020 (0.019)	0.397 (0.080)	[0.240, 0.555]
No spacing	9,614,127	7.25 (0.02)	0.024 (0.004)	-0.017 (0.013)	0.330 (0.058)	[0.217, 0.443]
First observed exposure, common history controls	1,314,334	6.84 (0.04)	0.035 (0.010)	0.019 (0.036)	0.506 (0.149)	[0.213, 0.799]
Spaced rows, without the recipient history and volume controls	3,103,280	7.12 (0.03)	0.026 (0.007)	-0.000 (0.026)	0.371 (0.100)	[0.175, 0.566]
First observed exposure, without the recipient history and volume controls	1,314,334	6.84 (0.04)	0.035 (0.010)	-0.009 (0.039)	0.506 (0.153)	[0.207, 0.805]

Notes: All-time score with its three history controls in every row; the environment controls and the primary assignment cells, exposure-date and hour-of-week effects throughout. Recipients: 1,835,628 in the spaced rows, 700,219 in the repeated-recipient rows, 1,855,251 under 30-day spacing, 1,862,144 unspaced, and 1,314,334 in the first-exposure rows. First stages (FS) and reduced forms (RF) are percentage points per standard deviation of the residualized score in the primary sample (0.0498); two-stage estimates are percentage points per receipt of a draw. The grant outcome is zero when a complete window contains no qualifying choice; the reach outcome records whether a qualifying choice occurs. The repeated-recipient rows are the recipients with at least two retained primary exposures, and the pooled and fixed-effects models use identical rows. The 30-day and unspaced specifications change exposure spacing while retaining the 90-day outcome, so a later choice can serve as the outcome of more than one exposure. The first-exposure rows are each recipient's first observed candidate exposure, retained if it satisfies the score's support rule and the history and follow-up requirements; they are not the first supported exposure. The last two rows omit the recipient history and volume controls. Standard errors and Wald intervals use the shared-account-or-exposure-date CR0 covariance with the historical scores treated as fixed. First-stage F is between 18,926 and 48,923 across the rows.

F.5 Trait-by-trait balance

Appendix D reports the joint balance tests. Table F.5 reports the eight trait comparisons behind them for the primary sample and for its entry and ordinary-ending frames, with the recipient-history and volume controls omitted and the environment controls retained. Table F.6 reports the trait comparisons for the spaced design, with the all-time score,

under each specification's own adjustment, with the joint tests in the last rows of each table. Units in both tables are trait standard deviations per native residual-score standard deviation. In the spaced design, recent history is measured from 37 to 7 days before the index game. The seven 90-day history traits and the preceding-30-day volume measure are among the controls in every panel, so they are not independent balance evidence and are omitted from the numerical comparisons.

Prior granting leans slightly positive in every panel: the ever-granted indicator moves between 0.0002 and 0.0014 trait standard deviations per score standard deviation across the panels (standard errors 0.0002 to 0.0006), the grant rate between 0.0002 and 0.0010, and in the exposure sample the ever-granted indicator is the trait that brings the joint test to $p = 0.067$. Read by the bias it implies, together with the other recorded traits, the lean is 1.9 percent of the reduced form (Appendix D). The entry and ordinary-ending frames in Table F.6 are the spaced design's, 17,745,415 and 11,410,477 games, built under its eligibility and history rules. Columns (5) and (6) of Table F.6 report the two balance variants: the exposure sample without the recipient history and volume controls, where all fifteen traits are tested ($p = 0.031$, carried by the disconnection and claim counts), and the repeated-recipient rows with recipient fixed effects, where the joint test rejects on account of a single trait, the recipient's recent receipts (0.0056 trait standard deviations per score standard deviation, standard error 0.0006), with the other six traits within 0.0014 of zero.

Table F.5: Recipient history and the residualized opponent grant score: primary sample and its frames

Recipient trait	Primary exposures 9,736,182	Entry frame 31,836,314	Ordinary endings 21,025,393
Any competitive opportunity	−0.000392 (0.000315)	−0.000226 (0.000176)	−0.000149 (0.000221)
Log competitive opportunities	−0.000339 (0.000321)	−0.000114 (0.000180)	−0.000121 (0.000213)
Ever granted	0.000100 (0.000319)	−0.000082 (0.000195)	−0.000099 (0.000231)
Competitive grant rate	−0.000028 (0.000345)	0.000123 (0.000198)	0.000295 (0.000239)
Log disconnections	−0.000781 (0.000306)	−0.000735 (0.000172)	−0.000626 (0.000204)
Log receipts	0.000944 (0.000329)	−0.000261 (0.000194)	−0.000185 (0.000237)
Log claims	−0.000831 (0.000306)	−0.000737 (0.000172)	−0.000623 (0.000203)
Log rated games	0.000276 (0.000296)	0.000117 (0.000172)	0.000164 (0.000200)
Joint Wald $\chi^2(8)$	39.705	33.770	27.144
Joint p	< 0.001	< 0.001	< 0.001

Notes: Coefficients are trait standard deviations per native residual-score standard deviation; standard errors are in parentheses. Traits cover the 90 days before the index game start, with no seven-day gap. Logged counts use $\log(1 + \text{count})$. Each trait is regressed on the opponent score with assignment-cell, exposure-date, and hour-of-week effects, rating and opponent-history controls, and opponent-environment controls. All eight recipient-history and volume controls are omitted for these balance regressions. Covariance allows shared accounts in either role or a shared exposure date.

Table F.6: Recipient history and the residualized opponent grant score: spaced design

Recipient trait	(1) Primary controls	(2) No environment controls	(3) Entry frame	(4) Ordinary-ending frame	(5) No history or volume controls	(6) Recipient fixed effects
Panel A. Recent-history traits, 37 to 7 days before the index game						
Any competitive opportunity	−0.000360	−0.000580	−0.000072	−0.000264	−0.000309	−0.000627
	(0.000353)	(0.000353)	(0.000162)	(0.000200)	(0.000546)	(0.000504)
	[0.307]	[0.100]	[0.656]	[0.186]	[0.571]	[0.213]
Log competitive opportunities	0.000010	−0.000244	0.000050	−0.000093	0.000125	0.000139
	(0.000330)	(0.000330)	(0.000147)	(0.000185)	(0.000554)	(0.000458)
	[0.975]	[0.460]	[0.736]	[0.615]	[0.822]	[0.762]
Ever granted	0.000863	0.000818	0.000230	0.000436	0.000931	0.001390
	(0.000414)	(0.000414)	(0.000188)	(0.000228)	(0.000605)	(0.000544)
	[0.037]	[0.048]	[0.221]	[0.056]	[0.124]	[0.011]
Grant rate	0.000689	0.000651	0.000238	0.000450	0.000767	0.000998
	(0.000420)	(0.000419)	(0.000185)	(0.000234)	(0.000606)	(0.000562)
	[0.101]	[0.121]	[0.198]	[0.055]	[0.206]	[0.076]
Log disconnections	−0.000257	−0.000619	−0.000163	−0.000145	−0.001270	0.000081
	(0.000292)	(0.000292)	(0.000127)	(0.000164)	(0.000544)	(0.000379)
	[0.379]	[0.034]	[0.201]	[0.378]	[0.020]	[0.831]
Log receipts	−0.000242	−0.000278	−0.000027	−0.000250	0.000155	0.005571
	(0.000464)	(0.000462)	(0.000205)	(0.000264)	(0.000574)	(0.000647)
	[0.602]	[0.548]	[0.894]	[0.342]	[0.787]	[< 0.001]
Log claims	−0.000210	−0.000574	−0.000170	−0.000149	−0.001264	−0.000205
	(0.000293)	(0.000292)	(0.000127)	(0.000164)	(0.000542)	(0.000378)
	[0.473]	[0.050]	[0.180]	[0.363]	[0.020]	[0.587]
Joint Wald χ^2 (df)	13.218 (7)	17.086 (7)			26.686 (15)	92.221 (7)
Joint p	0.067	0.017			0.031	< 0.001
Panel B. Column (5): the eight 90-day and volume traits also tested (estimate, standard error, p)						
Any competitive opportunity, 90-day history	−0.000131	(0.000535)	[0.806]			
Log competitive opportunities, 90-day history	−0.000003	(0.000555)	[0.995]			
Ever granted, 90-day history	0.000046	(0.000579)	[0.937]			
Grant rate, 90-day history	0.000094	(0.000602)	[0.875]			
Log disconnections, 90-day history	−0.001209	(0.000535)	[0.024]			
Log receipts, 90-day history	0.000600	(0.000582)	[0.302]			
Log rated games, preceding 30 days	−0.000012	(0.000503)	[0.980]			
Log claims, 90-day history	−0.001263	(0.000532)	[0.018]			

Notes: Estimates, with standard errors in parentheses and p-values in brackets. All-time leave-pair-out score. Units are trait standard deviations per native residual-score standard deviation. Recent history is measured from 37 to 7 days before the index game; the 90-day history measures and the preceding-30-day volume measure are controls in every panel and have no independent balance coefficient. "Primary controls" are those of Table 6, Panel A, including the opponent-environment measures. Inference uses the shared-account-or-exposure-date CR0 covariance. Columns (1), (2), and (5) use the 3,103,280 spaced exposures; column (3) the spaced design's entry frame of 17,745,415 games and column (4) its ordinary-ending frame of 11,410,477 games; column (6) the 1,967,871 exposures of recipients with at least two retained exposures, with recipient fixed effects. Column (5) tests all fifteen traits jointly; its seven recent traits alone give $\chi^2(7) = 16.4$, $p = 0.021$, and without the environment controls as well the fifteen-trait test gives $\chi^2(15) = 83.1$, $p < 0.001$. Joint tests for the primary sample and its frames are in Table F.5.

F.6 First stages by subgroup

Table F.7 reports the first stage by subgroup of game speed, recipient rating, prior competitive choices, and exposure month, each refitting the primary assignment cells, time effects, and controls within the subgroup. The first stage is positive in all 30 supported subgroups and every 95 percent interval lies above zero. It rises with the recipient's rating, is lower in rapid and classical games than in bullet and blitz, and varies within a narrow band across exposure months; no subgroup changes sign between the raw and adjusted fits. These are the testable implications of average monotonicity discussed in Appendix D, and they do not test exclusion.

Table F.7: First stages by subgroup

Subgroup	Exposures	Receipts	First stage	SE	Exposure month	Exposures	Receipts	First stage	SE
All exposures	3,103,280	83,907	7.12	0.03	February 2024	151,087	4,126	7.21	0.13
Game speed					March 2024	176,996	4,756	7.23	0.09
Bullet	346,492	9,293	7.39	0.09	April 2024	157,270	4,305	7.05	0.09
Blitz	1,713,722	44,884	7.45	0.04	May 2024	176,029	4,826	7.12	0.11
Rapid	992,373	28,111	6.62	0.05	June 2024	168,889	4,494	7.12	0.11
Classical	50,693	1,619	6.40	0.16	July 2024	162,026	4,302	7.18	0.12
Recipient rating					August 2024	167,901	4,411	7.20	0.12
Below 1200	912,043	28,070	6.31	0.04	September 2024	168,059	4,504	7.19	0.12
1200–1599	1,161,244	31,328	7.50	0.05	October 2024	180,382	4,862	7.20	0.10
1600–1999	820,021	19,635	7.97	0.07	November 2024	166,027	4,354	7.07	0.11
2000 and above	209,972	4,874	8.30	0.13	December 2024	184,579	4,899	7.09	0.08
Prior 90-day competitive choices					January 2025	182,163	5,109	7.28	0.09
No competitive choice	1,589,974	43,520	6.88	0.04	February 2025	164,646	4,467	7.08	0.10
Choices, none granted	1,433,434	38,109	7.43	0.04	March 2025	180,281	4,959	7.12	0.09
Granted at least once	79,872	2,278	7.33	0.14	April 2025	174,061	4,844	7.20	0.10
					May 2025	178,011	4,755	6.86	0.11
					June 2025	177,086	4,842	7.14	0.09
					July 2025	177,564	4,814	6.96	0.10
					August 2025	10,223	278	6.85	0.33

Notes: Adjusted first stages are percentage points of receipt per standard deviation of the residualized score (0.0498). Each subgroup refits the primary assignment cells, exposure-date and hour-of-week effects, and controls within the subgroup, with the all-time score and the environment controls. Standard errors use the shared-account-or-exposure-date CR0 covariance. The author's accounts are excluded in every role. The raw first stages, without controls, lie within 0.03 points of the adjusted ones in every subgroup. The 30 supported subgroups are the 36 specified groups less the five exposure months that the pre-exposure history and 90-day follow-up requirements leave empty, and the overall row; every first stage is positive and every 95 percent interval lies above zero.

F.7 The competitive-only score

The headline score pools the opponent's decisions across board states. An opponent who grants when far ahead looks lenient without being generous toward a competitor, so a natural check rebuilds the score from competitive-position decisions only, those in which the disconnected side had an evaluation of at least -100 centipawns, with the same all-

time donor set, pair exclusion, and donor-month state adjustment, and its own three history controls. This is the leave-own-type-out check of Maestas et al. [2013] and the reverse-sample test of Bhuller et al. [2020]: if the pooled score captured something other than the opponent's generosity toward competitors, the two scores would give different answers.

Table F.8 reports the results. On the 2,649,065 exposures whose opponent has at least five competitive-position decisions, the pooled score gives 0.302 (0.102) and the competitive-only score 0.316 (0.100); at a threshold of three decisions the two are 0.336 (0.099) and 0.352 (0.095). The score built from non-competitive decisions alone also predicts receipt at competitive positions, 4.74 points per standard deviation of its own residual, the reverse-sample pattern that average monotonicity implies, and gives 0.264 (0.154). With both scores in one model on the 2,599,477 rows that support both, the competitive component carries the first stage (partial R^2 0.201 against 0.004), the two-instrument estimate is 0.279 (0.101), and the overidentification test does not reject (Hansen $J = 0.42$, $p = 0.52$). The non-competitive decisions extend the score's coverage and add almost no independent first-stage information, and nothing indicates that pooling the board states contaminated the instrument.

Table F.8: The competitive-only score

Score and rows	Exposures	FS (SE)	RF (SE)	2SLS (SE)	95% interval
A. One score at a time					
Pooled score, rows with at least five competitive decisions	2,649,065	7.84 (0.04)	0.024 (0.008)	0.302 (0.102)	$[0.102, 0.502]$
Competitive-only score, at least five decisions	2,649,065	8.15 (0.03)	0.026 (0.008)	0.316 (0.100)	$[0.120, 0.512]$
Pooled score, rows with at least three competitive decisions	2,917,877	7.86 (0.03)	0.026 (0.008)	0.336 (0.099)	$[0.143, 0.529]$
Competitive-only score, at least three decisions	2,917,877	8.14 (0.02)	0.029 (0.008)	0.352 (0.095)	$[0.166, 0.538]$
Pooled score, rows with at least five non-competitive decisions	2,931,938	7.78 (0.04)	0.023 (0.008)	0.296 (0.100)	$[0.101, 0.492]$
Non-competitive-only score, at least five decisions	2,931,938	4.74 (0.04)	0.012 (0.007)	0.264 (0.154)	$[-0.039, 0.566]$
Pooled score, rows supporting both scores	2,599,477	7.82 (0.04)	0.020 (0.008)	0.253 (0.104)	$[0.050, 0.456]$
B. Both scores in one model, 2,599,477 rows					
Conditional first stage, competitive score		3.97 (0.02)			
Conditional first stage, non-competitive score		1.58 (0.04)			
Two-instrument two-stage least squares				0.279 (0.101)	$[0.081, 0.477]$
Partial first-stage R^2: 0.201 (competitive) and 0.0042 (non-competitive); joint first-stage $F = 51{,}063$. Hansen $J = 0.42$, $p = 0.52$.					

Notes: All scores use the all-time donor set, the same pair exclusion and donor-month state reference, and each carries its own three history controls; every row uses the primary assignment cells, time effects, and the environment controls. In Panel A, first stages (FS) and reduced forms (RF) are percentage points per standard deviation of the named score's own residual in that sample: 0.0518 and 0.0539 for the pooled score on the competitive-supported rows, 0.0963 and 0.1008 for the competitive-only scores, 0.0332 for the non-competitive-only score, and 0.0509 and 0.0500 for the pooled score on the other two row sets. In Panel B the conditional first stages are per 0.0498 units of each score, the common unit. The competitive-only score uses the opponent's decisions at which the disconnected side had an evaluation of at least -100 centipawns; the non-competitive-only score uses the rest. Two-stage estimates are percentage points per receipt; intervals are Wald intervals under the shared-account-or-exposure-date CR0 covariance. Hansen J is the efficient two-step GMM overidentification statistic with one degree of freedom, with the weight from the two-stage residual moments under the same covariance.