Philipp Strack, 2024 Clark Medalist*

Drew Fudenberg

June 2024

Philipp Strack is creative, insightful, and skillful, which has allowed him to make major contributions to many areas of microeconomic theory, including behavioral economics, information acquisition and learning, and mechanism design. He is also extraordinarily productive: As of June 2024 he had 37 published papers and several others that were forthcoming. Some of these papers provide a new understanding of important economics phenomena, others introduce results and techniques that will be used for years to come, and some of the papers do both.

Philipp was born in Bonn, Germany to a Greek mother (Nora Andrikopoulou, an archeologist) and a German father (Ulrich Strack, a lawyer.) In high school, he took college-level math classes at the University of Bonn. He then matriculated there, where he was greatly influenced by Paul Heidhues and Benny Moldovanu. Philipp earned *diplom*s (roughly equivalent to a master's degree) in both economics and mathematics at Bonn and then started Ph.D. work there in both disciplines simultaneously. Luckily for us, he subsequently decided to focus on economics and completed his economics Ph.D. in 2013.

I was very impressed by Philipp when he was on the junior job market after his Ph.D., and I began working with him then. He started his career at U.C. Berkeley, where he was promoted to tenure before moving to Yale in 2019. In 2023, Philipp won the Bodossaki Distinguished Young Scientist Award, which is given every two years

*I thank Roberto Corrao, Yifan Dai, David Kreps, Annie Liang, Giacomo Lanzani, Benny Moldovanu, Xiaosheng Mu, Luciano Pomato, Larry Samuelson, and Omer Tamuz for helpful comments.

to young Greek scientists whose "outstanding performance has already significantly contributed to the advancement of science." As I show below, this accolade was well-deserved, as was the Clark medal.

Mechanism Design

Mechanism design and its applications have been one of the great successes of mordern economic theory. Its initial wave focused on static settings and "rational" agents with expected utility preferences, but dynamic issues and/or other forms of preferences are relevant in many settings of interest. Philipp's work here has provided useful new tools for mechanism design, and applied them to develop new insights into a number of important applications.

Inducing Optimal Stopping Philipp's job market paper, Kruse and Strack [2015], studies the optimal contract for a principal to give an agent who privately observes the value of a Markov chain until they stop and obtain a reward that depends on the current level of the chain and on the stopping time.[1] The principal can give the agent a transfer that depends only on the time the agents stops. The paper develops a very nice dynamic generalization of the usual single-crossing condition, and shows it implies that a stopping rule is implementable if and only if it takes the form of a cutoff rule, i.e. "Stop if the current value is at least the cutoff value."

Majorization Kleiner, Moldovanu, and Strack [2021] provides two powerful technical results that both highlight a connection between many already-studied problems of mechanism and/or information design and can be used to analyze new ones. The key observation is that many such problems reduce to selecting a best element from

[1] Much of Strack's early work was on related topics in dynamic mechanism design, e.g., Bergemann and Strack [2015] which studies a model where the preferences of consumers follow a mean-reverting Markov process instead of the usual assumptions of fixed or i.i.d. preferences, and Gershkov, Moldovanu, and Strack [2018] which supposes that consumers can strategically time their purchases and that neither the time that they arrive in the marketplace nor the process that generates arrivals is known by the seller.

a constraint set that is given by a "majorization" condition. Hardy, Littlewood, and Polya [1929] said that real-valued function f *majorizes* another real-valued function g if they both have the same domain and same mean and the integral of f from any point in the domain to its upper limit exceeds that of g. For example, one probability distribution second-order stochastically dominates another if and only if its cumulative distribution majorizes the other one.[2] In many economic applications, the functions in question are monotone, which leads Kleiner, Moldovanu, and Strack [2021] to study the sets of monotone functions that either majorize or are majorized by a given function. They characterize the extreme points of these sets with a generalization of Myerson [1981]'s ironing procedure: Any extreme point is characterized by a collection of intervals on which the extreme point is piecewise constant, while outside the collection it coincides with the given function.

One application is to feasible symmetric allocations in independent private value allocation problems, where all agents agree on the rankings of the objects but some agents value every object more. The paper shows that the feasible allocations are described by a majorization condition, as in Border [1991]'s analysis of the case of 1 object. The paper then notes that the extreme points of the feasible set are all dominant-strategy incentive compatible and weakly monotone, as they involve either an efficient or a random allocation. Thus all Bayesian incentive compatible allocations, which must be feasible and so are integrals of the extreme points, can be implemented in dominant strategies, as in Gershkov, Goeree, Kushnir, Moldovanu, and Shi [2013].

Next, the paper analyzes the maximization of linear functionals given a majorization constraint, and applies these results to study revenue maximization in ranked-value auctions and Bayesian persuasion, as well as to optimal delegation and matching. This approach sheds new light on these widely-studied applications, while providing shorter and more revealing proofs.

Applications of Majorization In work with Alex Gershkov, Benny Moldovanu, and Mengxi Zhang, Philipp used majorization and the extreme points approach to

[2] Blackwell [1951] and Strassen [1965] show that any feasible posterior belief majorizes the prior.

extend classical methods of mechanism design to frameworks that allow agents to be risk aversion and/or have non-expected utility preferences. Gershkov, Moldovanu, Strack, and Zhang [2022] derives the optimal auction when bidders have non-expected utility preferences that exhibit constant risk aversion, which covers loss aversion (Kőszegi and Rabin [2006]), disappointment aversion (Gul [1991]), and dual preferences (Yaari [1987]). The paper shows that in the optimal mechanism, each agent's utility is independent of other agents' reports, even though Maskin and Riley [1984] shows is not generally the case with expected utility bidders who are strictly risk averse. In another application of majorization techniques, Gershkov, Moldovanu, Strack, and Zhang [2023] solves the optimal insurance design problem for agents that face a loss of random magnitude and have dual preferences a la Yaari. The paper shows that the optimal contracts are menus of deductibles when agents have private information about the probability of losses, and menus of coverage limits when the private information is about the magnitude of losses. Coverage limits are often observed in practice, yet they are not explained by the classic insurance model. Gershkov, Moldovanu, Strack, and Zhang [2021] shows that auctions where ex-ante symmetric bidders with i.i.d. valuations can make an investment after learning their type but before bidding in the auction, can be viewed as auctions with a form of non-expected utility that lets them apply majorization techniques. Here, unlike in standard auctions with symmetric bidders, asymmetric auctions may maximize the seller's expected revenue. The paper shows when symmetry is revenue-maximizing and that when it is, the optimal mechanism can be implemented as either a uniform-price auction or a discriminatory pay-your-bid auction with reserve prices that depend on the number of bidders and the number of units for sale. In addition to these applications, the paper's extreme-points approach provides a unified perspective on other analyses of mechanism design problems, such as mechanism design with redistributive concerns (Akbarpour, Dworczak, and Kominers [2024]), selling hard information (Ali, Haghpanah, Lin, and Siegel [2022]), and non-linear pricing by a seller who maximizes their worst-case profit (Bergemann, Heumann, and Morris [2023]).

Information Economics

The study of asymmetric information and adverse selection has shed light on many economic phenomena and institutions since it took off in the 1960s, but much of this work takes the information the agents have as given, despite Arrow [1996]'s complaint that "Not enough weight has been given to the possibility that information can be... altered by economic decisions." Philipp's work on sequential sampling, the comparison of experiments, and social learning has illuminated some important effects of endogenous information, and rekindled interest in these topics He has also made important contributions to the study of how to reveal useful information while respecting privacy constraints.

Of course, studying the choice of information structure requires specifying what sort of information can be obtained and what the cost of obtaining it is. Arrow [1985] proposes a few examples of information costs, including the Shannon entropy that has become popular following the work of Sims [2003]. However, the optimal allocation of attention can depend crucially on the specification of its cost, and the question of which cost functions are reasonable has remained underexplored. Philipp's work has shed important light on this.

The drift-diffusion model A large literature in neuroscience and psychology uses the *drift-diffusion model* ("DDM") to explain the joint distribution of choices and decision times in laboratory experiments. This model is a continuous-time version of the hypothesis-testing problem in Wald [1947], where an agent gathers costly samples before stopping and either accepting or rejecting the null hypothesis. The classic DDM also assumes there are only two states, but now the agent observes a diffusion process with known volatility and uncertain drift until they take one of two possible actions, where the payoff and the drift of the signal are both determined by the unknown state of the world.[3]

[3] The DDM was first proposed as a model of choice processes in perception tasks, where the subjects are asked to correctly identify visual or auditory stimuli. DDM-style models were subsequently applied to choice experiments, where subjects are choosing from a set of consumption goods presented to them. e.g. Roe, Busemeyer, and Townsend [2001]; Krajbich, Armel, and Rangel

The two-state model is analytically tractable (unlike the three-state model in Arrow, Blackwell, and Girshick [1949]), but it predicts that the agent stops and makes a decision whenever the state of the diffusion process reaches a time-invariant stopping boundary, which implies that the accuracy of the agent's choices is uncorrelated with when they stop. However, in a fixed experiment, earlier decisions are more likely to be correct.[4] To better understand this regularity, Fudenberg, Strack, and Strzalecki [2018] proposes a variant of the DDM where the agent has normally distributed independent priors on the utility of each choice, and the drift of the signal depends on the difference in the utilities. In this model, unlike the classic DDM, signal strength depends on the utility difference or on the ease of the perceptual task. Here, unlike with a two-state prior, beliefs update more slowly as data accumulates. Thus if the agent has been observing the process for a long time and the posterior means of the two choices are close, there are two reasons to stop sampling that aren't present in the binomial model: The agent doesn't expect to learn much more, and the two actions are probably about as good.

The paper first shows that an arbitrary stopping rule in a DDM with a fixed drift leads to a negative correlation between accuracy and stopping time if the value of the diffusion that makes the agent want to stop, i.e. the "stopping boundary," is decreasing in time. Intuitively, in this case when the agent stops quickly, it is because they received a very strong signal. Moreover, accuracy can be decreasing in decision time even if the stopping boundary on each trial is not monotone decreasing when the data consists of trials with randomly drawn values of the drift, because averaging over trials creates an additional selection effect beyond the one that occurs when the value of the drift is fixed: When the boundary does not increase "too quickly," the selection effect implies that in trials where the agent chose quickly, the drift was probably larger than usual, so the agent received sharper signals and is more likely to be correct. Finally, the paper shows that the aggregate data has a speed-accuracy complementarity when the agent uses an *optimal* stopping rule and their prior over

[2010]; Krajbich and Rangel. [2011]; Krajbich, Lu, Camerer, and Rangel [2012]; Milica, Jonathan, Huth, Koch, and Antonio [2010]; Reutskaja, Nagel, Camerer, and Rangel [2011].

[4]See e.g. Churchland, Kiani, and Shadlen [2008]; Ditterich [2006].

the states is correct.[5] The paper also considers the case where the agent has a fixed amount of attention and decides how much to focus on each signal. The optimum is to devote equal attention to both signals at every point in time, which reduces to the main mode.

Many experimental studies of accuracy and decision time fit the data with arbitrary stopping boundaries instead of boundaries that are derived with a theoretical foundation. Fudenberg, Newey, Strack, and Strzalecki [2020] characterizes the choice probabilities that can be generated by any stopping boundary and then shows how to estimate the drift and the boundary and construct a test statistic.[6]

The comparison of experiments One information structure (a map from states to signals) is a *garbling* of another if it can be generated by adding noise to it. Blackwell [1951] proves that signal structure K is (weakly) preferred to signal structure K' for all decision problems if and only if K' is a garbling of K; in this case, we say that K *Blackwell dominates* K'. This dominance notion has been very influential and has led to insights into mechanism design (e.g. Grossman and Hart [1983]), accounting (Demski [1973]), game theory (Kandori [1992]), and "Bayesian persuasion" (Kamenica and Gentzkow [2011]).

Experiments are often not done in isolation. Blackwell [1953] asks whether there are experiments that aren't ranked by garbling, but become ranked when many samples are taken from each, and Torgersen [1970] gives an example. Mu, Pomatto, Strack, and Tamuz [2021] characterizes when this occurs in the case of binary states. Their very clever proof uses the idea of the Renyi divergence, which is an easily-computed parametric measure of informativeness that extends the Blackwell order: If K Blackwell dominates K', it dominates K'' in the Renyi order for any parameter t.[7] Importantly, the Renyi divergence of n independent trials is simply n times the

[5] Alternative explanations for this complementarity include time-varying costs Drugowitsch, Moreno-Bote, Churchland, Shadlen, and Pouget [2012] and endogenous time variation in signal intensity Woodford [2014].

[6] Subsequent work on optimal stopping and stochastic choice includes Liang, Mu, and Syrgkanis [2022], Hébert and Woodford [2023], Gonçalves [2024], and Che and Mierendorff [2019].

[7] This Renyi order is also used in Frick, Iijima, and Ishii [2023].

t-divergence of one trial, so if n draws from K Blackwell-dominates n draws from K', K Reny-dominates K'. Mu, Pomatto, Strack, and Tamuz [2021] shows that under mild technical complications, the converse holds as well when n is sufficiently large, so if K Renyi-dominates K', every decision maker prefers a large sample from K to a large sample from K. The proof starts from the fact that the posterior belief after a sequence of i.i.d. experiments is a function of the prior and the product of the likelihood ratios for each observation, or equivalently the sum of the log-likelihoods. It then shows that two experiments are Blackwell ordered if and only if one of the associated "perfected log-likelihood ratios" first-order stochastically dominates the other.[8] To complete the proof, the paper uses a sharp large deviations bound to relate the Renyi order to the perfected log-likelihood ratios.

This beautiful result raises interesting open questions such as when will n draws from K and m draws from L be preferred to n draws from K' and m draws from L'? And what can be said about comparing large experiments when there are more than two states, and the Blackwell order is not the same as first-order stochastic dominance?

Social Learning Most of the social learning literature assumes either that each agent only acts once, as in Banerjee [1992] and Bikhchandani, Hirshleifer, and Welch [1992], or that the agents use a boundedly rational heuristic, as in DeGroot [1974] and Ellison and Fudenberg [1993]. Harel, Mossel, Strack, and Tamuz [2021] and Huang, Strack, and Tamuz [2024] study social learning by long-lived Bayesian agents in a setting where each agent receives an exogenous i.i.d. signal and so will eventually learn the state whether or not they observe the actions of others. Both papers focus on the rate of social learning and show that the speed is bounded above regardless of how many players there are, even though the rate is proportional to the number of players when all signals are public. The intuition is reminiscent of the Grossman-Stiglitz paradox in the rational expectations literature: If individuals learned quickly they would quickly reduce the amount they update their beliefs based on their private

[8]The perfected ratios are the sum of the log-likelihood ratios of the two experiments and a carefully chosen independent random variable.

signals, so that little or no new social information could be gained.

Harel, Mossel, Strack, and Tamuz [2021] studies a model with two states, myopic agents, and a fully connected social network, i.e. each agent observes the actions of all others. Huang, Strack, and Tamuz [2024] generalizes this to allow any finite number of states, non-myopic agents, and more general networks, though the main result is for the strongly connected case where there is an observational path between every pair of agents. As in most of the social learning literature, both papers assume that the agents' payoff functions do not depend on the actions of the other agents. Both papers also assume that there is a unique optimal action in each state and an upper bound on the informativeness of any signal.[9]

The first steps in Huang, Strack, and Tamuz [2024] are to show that after some random time T all agents behave myopically, and that in every strongly connected network all agents learn at the same rate, so the asymptotic learning rate is the same as with myopic agents and a fully connected network.[10] Lemma 3 then shows that this outside observer must learn at least as quickly as the society does, and Lemma 4 upper-bounds the private likelihoods by twice the maximum informativeness of any agent's signal. The main theorem then follows from the fact that if the equilibrium speed of learning were higher than this maximal private rate, an outside observer would know what action an agent would take before seeing it and so not learn anything more from that agent's actions. This would imply that the agents also stop learning from their private signals, a contradiction.[11]

Privacy constraints In some settings, there are legal or ethical constraints on revealing information such as a person's race or gender. Strack and Yang [2024]

[9] I.e. a bound on the absolute value of the log-likelihood ratio any signal gives to any pair of states. Harel, Mossel, Strack, and Tamuz [2021] allows for a continuum of signals, Huang, Strack, and Tamuz [2024] assumes the set of signals is finite.

[10] To prove this main result, the paper writes the log-likelihood ratios for each player and pair of states as the sum of the player's log-likelihood of the states and a "public" term that sums the prior log-likelihood and the history of actions the player has observed, where the public term describes the posterior log-likelihood of an observer who has the same prior and observes the same actions.

[11] Interested readers should know that the proofs in this paper are particularly short and clear, perhaps because they proceed from first principles. The "heaviest" result used is the classic large-deviations argument that the learning rate for a single agent is well-defined.

characterizes the experiments that are *privacy preserving* with respect to a collection of "privacy sets," meaning that after every realization of the experiment's signal, the posterior probability of each privacy set is the same as its prior. For example, if there are three possible states $\{1, 2, 3\}$ and a uniform prior, then a general experiment l could lead to any distribution over posteriors that averages out to the prior, but if information about state 1 is protected, then every signal the experiment can generate must also satisfy the constraint that the posterior probability of state 1 is 1/3. In general, privacy sets need not simply be a collection of states, but Strack and Yang [2024] shows it is without loss of generality to assume that they are, which greatly simplifies the analysis. The paper shows that when the state is one-dimensional, a signal is privacy-preserving if and only if it is Blackwell-dominated by some "reordered quantile" signal[12] and that every reordered quantile signal is Blackwell-undominated among privacy-preserving signals. It then proves a version of the same result for general state spaces, e.g. to cases where the state includes both a measure of ability or fitness for a position and information about protected characteristics. This is an important generalization of the two-state case analyzed in He, Sandomirskiy, and Tamuz [2022], both conceptually and for applications to "algorithmic fairness."

The cost of information Pomatto, Strack, and Tamuz [2023] provides an axiomatic characterization of information cost functions in a setting with a finite number of possible states of the world where the likelihood ratios for each pair of states are finite almost surely. It shows that a cost function that satisfies four axioms must be a *LLR cost*, which means that it is a weighted sum of the expected log-likelihood ratios between pairs of states, where the weight on the log-likelihood between state i and state j can depend on the prior probability of state i.

The axioms are that (1) Two experiments that are Blackwell equivalent have the same cost; (2) The cost of running two independent experiments is the sum of their costs; (3) If an experiment either succeeds (and produces an informative signal) or

[12]The reordered quantile reduces to the usual quantile function when the cumulative distribution function is continuous.

fails and yields no information, the marginal cost of increasing the probability of success is constant, and (4) The cost function is continuous in the total variation norm. In addition, the setup of the model assumes that the cost is always finite. Axiom 1 is automatically satisfied when experiments are chosen optimally, as all decision-makers prefer the Blackwell-equivalent experiment with the lowest cost. Axiom 2 is implicit in the sequential-sampling models discussed above; in that context, it is important that the time taken to gather information has no cost beyond the number of samples. In the presence of Axiom 2, Axiom 3 is satisfied if the agent can create "diluted" experiments that randomize between a fixed signal structure and an uninformative signal, and the cost of an experiment is not greater than the expected cost of performing its diluted version until an informative signal is generated. Moreover, Axiom 3 is satisfied by any cost function that is *posterior separable* in the sense of Caplin and Dean [2013], including the widely-used entropy cost function, although LLR costs are not posterior separable.

To illustrate the differences between LLR cost functions and entropy, consider an agent who sees red and blue dots on a video screen and is uncertain whether there are x red dots and $100 - x$ blue dots or $100 - x$ red dots and x blue dots. Here the entropy cost of a symmetric binary signal that either says "there are more red dots" or "there are more blue dots" is independent of x, while the LLR cost need not be. Next, suppose the agent thinks both states are equally likely, and consider the two costs as functions of the binary signal's precision, i.e. the probability that it reports the correct state. Both costs are 0 for the completely uninformative signal that reports "more red" with probability 1/2 regardless of the state, and both are increasing and convex, but as the probability the signal reports the true state goes to 1, the LLR cost goes to infinity, while the entropy cost converges to 1/2: LLR's use of log likelihood ratios instead of the log-likelihood used in entropy places a greater penalty on very accurate signals.

Finally, with entropy, the marginal cost of a more accurate binary signal does not depend on the agent's prior, while with the LLR cost it does, because the weights in LLR cost are inversely proportional to the prior probabilities.[13]

[13] See Denti [2022] for reasons posterior separability may be too strong a condition.

Imperfect Learning and Behavioral Economics

A standard approach in behavioral economics is to modify one or two assumptions of a standard theory in the direction of greater psychological realism (Camerer, Loewenstein, and Rabin [2004]). However, as I noted in Fudenberg [2006], factors that support one modification from the standard model may also argue for additional modifications. Moreover, the great multiplicity of behavioral theories can make it difficult to know which ones to apply in a given situation or how they interact. As I argued, both of these problems could be mitigated if the many assumptions of behavioral economics were integrated into a few more general models. Philipp's work on imperfect learning has provided a major step in this direction. It has had two main focuses: learning by agents who have perfect memory but whose prior prevents them from learning the truth, and agents whose main difficulties in learning come from their imperfect memory. In both cases, Philipp develops models that link learning directly with belief and behavior, and generate realistic and interesting behavioral biases.

Misspecification and Learning Economists typically assume that agents are *correctly specified*, meaning that their prior beliefs do not rule out the truth. Such agents will eventually learn the consequences of any action they play infinitely often although, although they need not learn the consequences of actions they rarely or never play. Yet people sometimes seem to have incorrect beliefs despite having an abundance of relevant data, such as a belief that taxes are linear in income when they are not, or "causation neglect" about the impact of actions on outcomes. One explanation for this is that they may be *misspecified*, and have prior beliefs that rule out the true state of the world. Such misspecification is especially relevant for parametric learning models, since any parametric prior (such as the assumption of a linear demand curve with unknown slope and intercept) assigns probability 0 to "most" data-generating processes.

To understand the effect of misspecified priors, recall that the posterior likelihood of two states under Bayesian updating is their prior probability multiplied by the likelihood ratio they assign to the empirical distribution. Thus when agents are correctly specified and their data identifies the true data-generating process, their beliefs converge to the truth, as it is the only model that maximizes the empirical likelihood. Berk [1966] shows that the beliefs of a misspecified agent asymptotically concentrate on the set of models that maximize the likelihood of the observed data when the data-generating process is exogenous and sufficient to identify the true data generating process. However, in many economic applications, actions and associated signal distributions depend on the action taken by the agent, and knowing the probability distribution over outcomes that one action generates may not identify the probability distributions generated by others. This interdependence between the agent's misspecification, beliefs, and long-run actions makes the dynamics of misspecified learning much richer. Arrow and Green [1973] develops a general model of learning by misspecified agents who learn "passively," and always play the myopic best response to their current beliefs.[14] Esponda and Pouzo [2016] defines *Berk–Nash equilibrium* and relates it to the asymptotic behavior of passive misspecified learning by combining ideas from Berk [1966] and Fudenberg and Kreps [1993]. This led to a resurgence of interest in misspecified learning and helped inspire several of Philipp's papers.

Learning by Overconfident Agents Heidhues, Kőszegi, and Strack [2018] studies passive learning by an agent who is uncertain about a "fundamental" parameter but is certain that they know their own ability even though their true ability is lower than they think. Given the observed relatively low performance, the agent's overconfidence in their own ability leads them to underestimate the level of the fundamental and draw incorrect inferences regarding the fundamental. Over time the agent's beliefs become monotonically farther from the truth, and their action gets farther from optimal. As the paper notes, this provides another illustration of how assuming that the overconfident agent is otherwise "classical" can be misleading. To prove that beliefs and actions converge, the paper develops a version of the martingale law of

[14] Passive learning is optimal if agents do not care about their future payoffs or if the information they observe does not depend on their actions.

large numbers for Markov chains that replaces the usual square-integrability condition with a restriction to real-valued chains that grow sub-linearly almost surely, a result that will be useful more generally.

Convergence in One-Dimensional Models Heidhues, Kőszegi, and Strack [2021] modifies the observation structure of Heidhues, Kőszegi, and Strack [2018] so that the agent's inferences do not depend on their actions, which ensures that passive learning is optimal even if the agent is not myopic.[15] The paper then uses stochastic approximation to show that beliefs and actions converge. Unlike Heidhues, Kőszegi, and Strack [2018], it allows for multiple steady states, so even if the system almost surely converges, what it converges to may be stochastic.

Excessive experimentation Fudenberg, Romanyuk, and Strack [2017] was the first study of the dynamics of learning by a misspecified agent who is patient and thus is willing to engage in "active learning" by experimenting with actions that do not maximize their expected current payoff. When agents are correctly specified and very patient, with high probability they eventually stop "experimenting" and play a myopic best response, because additional information will have little impact on their beliefs and future play. Fudenberg, Romanyuk, and Strack [2017] shows this need not be the case when players are misspecified, which can lead actions to cycle instead of converge in a setting where they would converge with either correctly specified priors or with myopic optimization.

Convergence Fudenberg, Lanzani, and Strack [2021] provides a sharp general characterization of the limit outcomes of misspecified learning in single-player decision problems when models in the support of the agent's prior are mutually absolutely continuous with respect to the true data-generating process and the prior has *sub-exponential decay* in the sense of Fudenberg, He, and Imhof [2017].[16] It

[15] This is technically not a generalization as the models aren't quite nested.

[16] Other results on convergence to Berk-Nash equilibria require that the agents are myopic and assume either i.i.d. payoff shocks as in Fudenberg and Kreps [1993], a finite-support prior (Esponda and Pouzo [2016], Frick, Iijima, and Ishii [2020], and Bohren and Hauser [2021]), which rules out e.g.

first shows that any limit point must be a *uniform Berk-Nash equilibrium*, which strengthens Berk-Nash equilibrium by requiring that the action is a best reply to *any* mixture over the models that maximize the empirical likelihood over the models in the support of the agent's prior. To do this, the paper shows that the agent's beliefs concentrate on the distributions that maximize the empirical likelihood at an exponential rate that is uniform over the sample realizations. It then shows that if play converges to a particular action, the empirical outcome frequencies converge to the corresponding data-generating process with oscillations that die out more slowly than the exponential concentration of beliefs so that beliefs concentrate around every maximizer infinitely often. Thus if the limit action were not a best reply to one of the maximizers, the agent would switch to another action when the beliefs concentrate on that maximizer.

The paper then establishes that an action has a high probability of being the limit outcome for "nearby" beliefs if and only if it is a *uniformly strict Berk-Nash equilibrium*, meaning that is a strict best reply to every mixture over the relevant maximizers. The key step here is showing that if beliefs initially assign a very high probability to models that make the limit action optimal, they are unlikely to drop below the threshold that makes it suboptimal.[17]

Finally, equilibria are *positively attractive* if they have positive probability of being the long-run outcome from any starting belief. Uniformly strict Berk-Nash equilibria are positively attractive under various types of misspecification, including various forms of incorrect perceptions of how actions influence the distribution of outcomes, such as a belief that outcomes are exogenous when they are not, or that outcome distribution under one action is uncorrelated with that under another. The paper uses this to show that there can be a positive probability that a patient but misspecified agent will fall into an "underinvestment trap" when a correctly specified agent would not.

a uniform prior over the probability that a coin lands on heads. Esponda, Pouzo, and Yamamoto [2021] proves a related result about the convergence of action frequencies.

[17] The proof generalizes an upcrossing argument of Fudenberg and Levine [1992] to misspecified beliefs and then strengthens a bound from Frick, Iijima, and Ishii [2020] to hold uniformly.

Selective Memory Most people's memory is both limited and selective: They don't remember everything, and are more likely to remember some things than others.[18] Fudenberg, Lanzani, and Strack [2024] studies the long implications of selective memory when a myopic agent chooses actions that maximize their expected utility. It assumes that agents compute their beliefs from their prior and remembered experiences, as opposed to updating their posterior belief, and that agents are unaware of their selective memory. Thus they update their beliefs as if the experiences they remember are the only ones that occurred.[19] The paper also assumes that agents choose actions that maximize the current period's expected payoff.

Fudenberg, Lanzani, and Strack [2024] shows that if the agent eventually always plays a given action a, the action is a *selective memory equilibrium*, meaning that the action is a best response to a belief that maximizes the likelihood of a version of the outcome distribution that gives more weight to the outcomes the agent is more likely to remember. Using stochastic approximation results from Benaim, Hofbauer, and Sorin [2005], Esponda, Pouzo, and Yamamoto [2021], and a result on the rate that beliefs converge from Fudenberg, Lanzani, and Strack [2023], the paper shows that every limit strategy is a selective memory equilibrium. It also provides a sufficient condition for beliefs and actions to converge. Selective memory equilibrium resembles Berk-Nash equilibrium. Indeed, the paper shows that every long-run outcome that can be achieved as the result of misspecified learning can also be derived as a consequence of a selective memory and vice-versa.[20] Importantly, the form of misspecification that would lead to the same behavior as a given form of selective memory depends on the environment. That is, particular forms of misspecification and selective memory that coincide under one information structure can respond very differently to changes in what the agent observes. For example, combining positive and negative feedback has different effects on agents with ego-boosting memory than

[18]See e.g. Shadlen and Shohamy [2016], Schacter [2008], and Kahana [2012].

[19]d'Acremont, Schultz, and Bossaerts [2013] and Sial, Sydnor, and Taubinsky [2023] provide evidence that agents access their accumulated evidence each period when updating beliefs. Reder [2014], Zimmermann [2020], and Gödker, Jiao, and Smeets [2022] provide evidence of partial or complete unawareness of memory biases. Fudenberg, Lanzani, and Strack [2024]'s results extend to this partial unawareness.

[20]There are some subtleties involved when the equilibria in question are not uniformly strict.

on agents who are mistakenly sure that they have high ability. The paper shows that "associativeness" has no effect on long-run outcomes and that ego-boosting memory bias leads agents to underestimate the ability of their coworkers.

Risk Preferences

Risk preferences are a key factor in the insurance and investment markets and also matter for decisions related to health, job choice, and retirement. The risk attitudes of expected utility agents depend on the totality of risks they face as opposed to the properties of each risk in isolation, so it is important to understand how risk preferences change as additional risks are added. It is also important to understand the behavior of agents who do not maximize expected utility, but instead evaluate risks using non-linear functions of their probabilities. Philipp has done impressive work on both topics.

Adding Noise The concepts of first-order and second-order stochastic dominance play a key role in the economics of risk and insurance.[21] Adding noise to lotteries preserves stochastic dominance: If p first or second order stochastically dominates q, the same is true for the compound lotteries $p+r$ and $q+r$ for any lottery r that is independent of p and q. However, the converse is not true: Even if p does not first or second order dominate q, there can be an independent lottery r such that $p+r$ does dominate $q+r$. More strongly, Pomatto, Strack, and Tamuz [2020] shows that if the expected value of a random variable x is higher than that of y, there is an independent random variable r such that $x+r$ first-order stochastically dominates $y+r$, and that if p and q have the same mean and the variance of p is strictly less than that of q then there is an independent lottery r such that $p+r$ second-order

[21]Recall that real-valued lottery p first-order stochastically dominates lottery q if p is more likely than q to be above any fixed z, and that this is equivalent to the statement that the expected utility of p is (weakly) higher than that of q for all weakly increasing utility functions (provided these expectations exist.) Similarly, lottery p second-order stochastically dominates lottery q if and only if it is weakly preferred by all agents with weakly increasing and weakly concave preferences. When p and q in addition have the same mean, q is a mean-preserving spread of p (Rothschild and Stiglitz [1978]) and an example of the majorization relation discussed above.

stochastically dominates $q + r$, [22] It is important here that r can take on arbitrarily large and small values, but they argue that the required standard deviations need not be implausibly large. For example, if p pays 12 and -10 with probability $1/2$ each, then it does not first-order stochastically dominate 0, but there is an r with large standard deviation $(3, 525!)$ such that $p + r$ first order stochastically dominates $0 + r$.

The idea of the proof for first-order dominance is as follows. The assumption that the expected value of x is higher than that of y implies that the cumulative distribution function (cdf) of y is on average above that of x. This suggests that adding a lottery r with a "sufficiently diffuse" distribution will lead to the cdf of $y + r$ being above that of $x + r$ at least at most points. To show there is a lottery r that makes the cdf of $y + r$ *everywhere* above that of $x + r$ they make clever use of a result of Ruzsa and Szekely [1988], which shows that a signed measure that assigns total mass 1 to R can be "smoothed" into a probability measure by convolving it with an appropriately chosen probability measure.

Background Risk An immediate corollary of Pomatto, Strack, and Tamuz [2020] is that for each finite set of gambles, there is a background risk such that any agent whose preferences are monotone with respect to first order stochastic dominance will rank gambles by their expected value. Mu, Pomatto, Strack, and Tamuz [2024a] proves a related result: an agent who has background risk with full support on the reals (and in particular can be arbitrarily large or small) will accept any gamble that has positive expectation and whose riskiness in the sense of Aumann and Serrano [2008] is less than a measure of the "size" of the background risk. It uses this to argue that under plausible levels of background risk, expected utility theory, prospect theory, and rank-dependent utility all imply that agents who account for background risk and respect stochastic dominance are risk-neutral over small gambles.

[22] Tarsney [2018] independently found an example of this for first-order dominance. Mu, Pomatto, Strack, and Tamuz [2024b] generalizes Pomatto, Strack, and Tamuz [2020] to other statistics than moments, and provides a sharper result for the case where the added random variable is required to be bounded.

Cumulative Prospect Theory in A Dynamic Setting Cumulative prospect theory (Tversky and Kahneman [1992]) is the most widely used generalization of expected utility theory. It supposes that outcomes are evaluated relative to a "reference point" that separates outcomes into gains and losses, and the agent distorts cumulative probabilities as in rank-dependent utility (Quiggin [1982]). Most applications of cumulative prospect theory further assume that the agent overweights extremely unlikely gains and underweights extremely likely losses.

Ebert and Strack [2015] points out that under fairly general conditions, an agent who distorts probabilities in this way will accept some small binomial gambles with negative expected payoff provided their loss aversion is not too strong.[23] Ebert and Strack [2015] uses this fact about static choice to study an agent who needs to decide when to stop a diffusion process with a known negative drift and claim the current state as their payoff. A risk averse expected utility maximizer would stop immediately, but Ebert and Strack [2015] shows that a "naïve" agent with cumulative prospect theory preferences will never stop, because a simple stopping rule with asymmetric upper and lower boundaries can generate negative-expected-value binary lotteries that the agent finds appealing.[24] Conversely, Ebert and Strack [2018] shows that a "sophisticated" cumulative prospect theory agent who is aware their future selves will have different preferences but cannot commit to future behavior, may never start a diffusion with high positive drift because they are afraid that the future selves will stop too soon. It concludes that in dynamic settings, prospect theory needs to be modified to yield more realistic predictions.

[23] Azevedo and Gottlieb [2012] shows this is true for sufficiently large gambles.

[24] Here "naïve" means that the agent chooses their stopping rule believing it will be carried out, when in fact at subsequent periods they will choose to continue. This makes the agent dynamically inconsistent in the sense of Strzalecki [2024]. The main model uses continuous time to ensure there is a rich enough set of possible gambling strategies, but the paper shows the result holds in many discrete-time settings.

Conclusion

Philipp has developed substantial extensions of classic results in mechanism design, and helped spark what his Clark citation calls a "new wave of information economics." This work on its own is arguably Clark-worthy, but he is also one of the leaders in the push to provide foundations for various behavioral models and build bridges between mordern economic theory and fields such as behavioral economics, computer science, neuroscience, and psychology. In addition, Philipp has made contributions to the study of market design (Leshno and Strack [2020], Nikzad and Strack [2023]), contests (Fang, Noe, and Strack [2020], Seel and Strack [2013]), present bias (Heidhues and Strack [2019]), and epidemics (Kruse and Strack [2020]). All of his work is clear and insightful, and its influence seems sure to expand.

Philipp is also one of the most cheerful and friendly people I know. I have greatly enjoyed working with and learning from him.

References

Akbarpour, Mohammad, Piotr Dworczak, and Scott Duke Kominers (2024). "Redistributive allocation mechanisms." *Journal of Political Economy*, 132.

Ali, S Nageeb, Nima Haghpanah, Xiao Lin, and Ron Siegel (2022). "How to sell hard information." *The Quarterly Journal of Economics*, 137, 619–678.

Arrow, K. J., D. Blackwell, and M. A. Girshick (1949). "Bayes and Minimax Solutions of Sequential Decision Problems." *Econometrica*, 17, 213–244.

Arrow, Kenneth J (1985). "Informational structure of the firm." *The American Economic Review*, 75, 303–307.

— (1996). "The economics of information: An exposition." *Empirica*, 23, 119–128.

Arrow, KJ and JR Green (1973). "Notes on Expectations Equilibria in Bayesian Settings." Working Paper No. 33, Stanford University.

Aumann, Robert J and Roberto Serrano (2008). "An economic index of riskiness." *Journal of Political Economy*, 116, 810–836.

Azevedo, Eduardo M and Daniel Gottlieb (2012). "Risk-neutral firms can extract unbounded profits from consumers with prospect theory preferences." *Journal of Economic Theory*, 147, 1291–1299.

Banerjee, Abhijit V (1992). "A simple model of herd behavior." *The Quarterly Journal of Economics*, 107, 797–817.

Benaim, Michel, Josef Hofbauer, and Sylvain Sorin (2005). "Stochastic approximations and differential inclusions." *SIAM Journal on Control and Optimization*, 44, 328–348.

Bergemann, Dirk, Tibor Heumann, and Stephen Morris (2023). "Cost Based Nonlinear Pricing." *Working paper.*

Bergemann, Dirk and Philipp Strack (2015). "Dynamic revenue maximization: A continuous time approach." *Journal of Economic Theory*, 159, 819–853.

Berk, Robert H (1966). "Limiting behavior of posterior distributions when the model is incorrect." *The Annals of Mathematical Statistics*, 37, 51–58.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch (1992). "A theory of fads, fashion, custom, and cultural change as informational cascades." *Journal of political Economy*, 100, 992–1026.

Blackwell, David (1951). "Comparison of experiments." *Proceedings of the second Berkeley symposium on mathematical statistics and probability.* Vol. 1, 26.

— (1953). "Equivalent Comparisons of Experiments." *The Annals of Mathematical Statistics*, 24, 265–272.

Bohren, J Aislinn and Daniel N Hauser (2021). "Learning with heterogeneous misspecified models: Characterization and robustness." *Econometrica*, 89, 3025–3077.

Border, Kim C (1991). "Implementation of reduced form auctions: A geometric approach." *Econometrica: Journal of the Econometric Society*, 1175–1187.

Camerer, Colin F, George Loewenstein, and Matthew Rabin (2004). *Advances in behavioral economics.* Princeton University Press.

Caplin, Andrew and Mark Dean (2013). *Behavioral implications of rational inattention with shannon entropy.*

Che, Yeon-Koo and Konrad Mierendorff (Aug. 2019). "Optimal Dynamic Allocation of Attention." *American Economic Review*, 109, 2993–3029.

Churchland, Anne K, Roozbech Kiani, and Michael N Shadlen (2008). "Decision-Making with Multiple Alternatives." *Nature Neuroscience*, 11, 693–702.

d'Acremont, Mathieu, Wolfram Schultz, and Peter Bossaerts (2013). "The human brain encodes event frequencies while forming subjective beliefs." *Journal of Neuroscience*, 33, 10887–10897.

DeGroot, Morris H (1974). “Reaching a consensus.” *Journal of the American Statistical association*, 69, 118–121.

Demski, Joel S (1973). “The general impossibility of normative accounting standards.” *The Accounting Review*, 48, 718–723.

Denti, Tommaso (2022). “Posterior separable cost of information.” *American Economic Review*, 112, 3215–3259.

Ditterich, Jochen (2006). “Stochastic Models of Decisions about Motion Direction: Behavior and Physiology.” *Neural Networks*, 19, 981–1012.

Drugowitsch, Jan, Rubén Moreno-Bote, Anne K Churchland, Michael N Shadlen, and Alexandre Pouget (2012). “The cost of accumulating evidence in perceptual decision making.” *Journal of Neuroscience*, 32, 3612–3628.

Ebert, Sebastian and Philipp Strack (2015). “Until the bitter end: On prospect theory in a dynamic context.” *American Economic Review*, 105, 1618–1633.

— (2018). “Never, ever getting started: On prospect theory without commitment.” *Working Paper*.

Ellison, Glenn and Drew Fudenberg (1993). “Rules of thumb for social learning.” *Journal of political Economy*, 101, 612–643.

Esponda, Ignacio and Demian Pouzo (2016). “Berk–Nash equilibrium: A framework for modeling agents with misspecified models.” *Econometrica*, 84, 1093–1130.

Esponda, Ignacio, Demian Pouzo, and Yuichi Yamamoto (2021). “Asymptotic behavior of Bayesian learners with misspecified models.” *Journal of Economic Theory*, 195, 105260.

Fang, Dawei, Thomas Noe, and Philipp Strack (2020). “Turning up the heat: The demoralizing effect of competition in contests.” *Journal of Political Economy*, 128, 1940–1975.

Frick, Mira, Ryota Iijima, and Yuhta Ishii (2020). “Stability and robustness in misspecified learning models.”

— (July 2023). “Belief Convergence under Misspecified Learning: A Martingale Approach.” *The Review of Economic Studies*, 90, 781–814.

Fudenberg, Drew (2006). “Advancing beyond advances in behavioral economics.” *Journal of Economic Literature*, 44, 694–711.

Fudenberg, Drew, Kevin He, and Lorens A Imhof (2017). “Bayesian posteriors for arbitrarily rare events.” *Proceedings of the National Academy of Sciences*, 114, 4925–4929.

Fudenberg, Drew and David M Kreps (1993). "Learning mixed equilibria." *Games and Economic Behavior*, 5, 320–367.

Fudenberg, Drew, Giacomo Lanzani, and Philipp Strack (2021). "Limit points of endogenous misspecified learning." *Econometrica*, 89, 1065–1098.

— (2023). "Pathwise concentration bounds for misspecified Bayesian beliefs." *Theoretical Economics*, 18, 1585–1622.

— (2024). "Selective memory equilibrium." *Journal of Political Economy, forthcoming.*

Fudenberg, Drew and David K Levine (1992). "Maintaining a reputation when strategies are imperfectly observed." *The Review of Economic Studies*, 59, 561–579.

Fudenberg, Drew, Whitney Newey, Philipp Strack, and Tomasz Strzalecki (2020). "Testing the drift-diffusion model." *Proceedings of the National Academy of Sciences*, 117, 33141–33148.

Fudenberg, Drew, Gleb Romanyuk, and Philipp Strack (2017). "Active learning with a misspecified prior." *Theoretical Economics*, 12, 1155–1189.

Fudenberg, Drew, Philipp Strack, and Tomasz Strzalecki (2018). "Speed, accuracy, and the optimal timing of choices." *American Economic Review*, 108, 3651–3684.

Gershkov, Alex, Jacob K Goeree, Alexey Kushnir, Benny Moldovanu, and Xianwen Shi (2013). "On the equivalence of Bayesian and dominant strategy implementation." *Econometrica*, 81, 197–220.

Gershkov, Alex, Benny Moldovanu, and Philipp Strack (2018). "Revenue-Maximizing Mechanisms with Strategic Customers and Unknown, Markovian Demand." *Management Science*, 64, 2031–2046.

Gershkov, Alex, Benny Moldovanu, Philipp Strack, and Mengxi Zhang (2021). "A theory of auctions with endogenous valuations." *Journal of Political Economy*, 129, 1011–1051.

— (2022). "Optimal auctions: Non-expected utility and constant risk aversion." *The review of economic studies*, 89, 2630–2662.

— (2023). "Optimal Insurance: Dual Utility, Random Losses, and Adverse Selection." *American Economic Review*, 113, 2581–2614.

Gödker, Katrin, Peiran Jiao, and Paul Smeets (2022). "Investor memory."

Gonçalves, Duarte (2024). "Speed, Accuracy, and Complexity." *arXiv preprint arXiv:2403.11240.*

Grossman, Sanford J and Oliver D Hart (1983). "An Analysis of the Principal-Agent Problem." *Econometrica*, 51, 7–46.

Gul, Faruk (1991). “A theory of disappointment aversion.” *Econometrica: Journal of the Econometric Society*, 667–686.

Hardy, G. H., J. E. Littlewood, and G. Polya (1929). “Some simple inequalities satisfied by convex functions.” *Messenger Math*, 58. Reprinted in American Mathematical Monthly, 35, 85-93 (1928), 310, 145–152.

Harel, Matan, Elchanan Mossel, Philipp Strack, and Omer Tamuz (2021). “Rational groupthink.” *The Quarterly Journal of Economics*, 136, 621–668.

He, Kevin, Fedor Sandomirskiy, and Omer Tamuz (2022). “Private Private Information.” *working paper*.

Hébert, Benjamin and Michael Woodford (2023). “Rational inattention when decisions take time.” *Journal of Economic Theory*, 208, 105612.

Heidhues, Paul, Botond Kőszegi, and Philipp Strack (2018). “Unrealistic expectations and misguided learning.” *Econometrica*, 86, 1159–1214.

— (2021). “Convergence in models of misspecified learning.” *Theoretical Economics*, 16, 73–99.

Heidhues, Paul and Philipp Strack (2019). *Identifying procrastination from the timing of choices*. Tech. rep. Working Paper.

Huang, Wanying, Philipp Strack, and Omer Tamuz (2024). “Learning in repeated interactions on networks.” *Econometrica*, 92, 1–27.

Kahana, Michael Jacob (2012). *Foundations of human memory*. Oxford University Press.

Kamenica, Emir and Matthew Gentzkow (2011). “Bayesian persuasion.” *American Economic Review*, 101, 2590–2615.

Kandori, Michihiro (Jan. 1992). “Social Norms and Community Enforcement.” *The Review of Economic Studies*, 59, 63–80.

Kleiner, Andreas, Benny Moldovanu, and Philipp Strack (2021). “Extreme points and majorization: Economic applications.” *Econometrica*, 89, 1557–1593.

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

Krajbich, I., C. Armel, and A. Rangel (2010). “Visual fixations and the computation and comparison of value in simple choice.” *Nat Neurosci*, 13, 1292–1298.

Krajbich, Ian, Dingchao Lu, Colin Camerer, and Antonio Rangel (2012). “The Attentional Drift-Diffusion Model Extends to Simple Purchasing Decisions.” *Frontiers in Psychology*, 3.

Krajbich, Ian and Antonio Rangel. (2011). "Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions." *PNAS*, 108, 13852–13857.

Kruse, Thomas and Philipp Strack (2015). "Optimal stopping with private information." *Journal of Economic Theory*, 159, 702–727.

— (2020). "Optimal control of an epidemic through social distancing." *Available at SSRN 3581295*.

Leshno, Jacob D and Philipp Strack (2020). "Bitcoin: An axiomatic approach and an impossibility theorem." *American Economic Review: Insights*, 2, 269–286.

Liang, Annie, Xiaosheng Mu, and Vasilis Syrgkanis (2022). "Dynamically Aggregating Diverse Information." *Econometrica*, 90, 47–80.

Maskin, Eric and John Riley (1984). "Optimal auctions with risk averse buyers." *Econometrica: Journal of the Econometric Society*, 1473–1518.

Milica, Milosavljevic, Malmaud Jonathan, Alexander Huth, Christof Koch, and Rangel Antonio (2010). "The Drift Diffusion Model Can Account for the Accuracy and Reaction Time of Value-Based Choices Under High and Low Time Pressure." *Judgment and Decision Making*, 5.

Mu, Xiaosheng, Luciano Pomatto, Philipp Strack, and Omer Tamuz (2021). "From Blackwell dominance in large samples to Renyi divergences and back again." *Econometrica*, 89, 475–506.

— (2024a). "Background risk and small-stakes risk aversion." *American Economic Review: Insights*, 6, 262–276.

— (2024b). "Monotone additive statistics." *Econometrica*, 94, 995–1031.

Myerson, Roger B (1981). "Optimal auction design." *Mathematics of operations research*, 6, 58–73.

Nikzad, Afshin and Philipp Strack (2023). "Equity in Dynamic Matching: Extreme Waitlist Policies." *Management Science*.

Pomatto, Luciano, Philipp Strack, and Omer Tamuz (2020). "Stochastic dominance under independent noise." *Journal of Political Economy*, 128, 1877–1900.

— (2023). "The cost of information: The case of constant marginal costs." *American Economic Review*, 113, 1360–1393.

Quiggin, John (1982). "A theory of anticipated utility." *Journal of economic behavior & organization*, 3, 323–343.

Reder, Lynne M (2014). *Implicit memory and metacognition*. Psychology Press.

Reutskaja, Elena, Rosemarie Nagel, Colin F. Camerer, and Antonio Rangel (Apr. 2011). "Search Dynamics in Consumer Choice under Time Pressure: An Eye-Tracking Study." *American Economic Review*, 101, 900–926.

Roe, R. M., J. R. Busemeyer, and J. T. Townsend (2001). "Multialternative decision field theory: A dynamic connectionst model of decision making." *Psychological Review*, 108, 370–392.

Rothschild, Michael and Joseph E. Stiglitz (1978). "6 - Increasing Risk: I. A Definition." *Uncertainty in Economics*. Academic Press, 99–121.

Ruzsa, Imre Z. and Gabor J. Szekely (1988). *Algebraic Probability Theory*. Wiley Series in Probability and Statistics.

Schacter, Daniel L (2008). *Searching for memory: The brain, the mind, and the past*. Basic Books.

Seel, Christian and Philipp Strack (2013). "Gambling in contests." *Journal of Economic Theory*, 148, 2033–2048.

Shadlen, Michael N and Daphna Shohamy (2016). "Decision making and sequential sampling from memory." *Neuron*, 90, 927–939.

Sial, Afras Y, Justin R Sydnor, and Dmitry Taubinsky (2023). "Biased Memory and Perceptions of Self-Control."

Sims, Christopher A (2003). "Implications of rational inattention." *Journal of monetary Economics*, 50, 665–690.

Strack, Philipp and Kai Hao Yang (2024). "Privacy Preserving Signals." *working paper*.

Strassen, Volker (1965). "The existence of probability measures with given marginals." *The Annals of Mathematical Statistics*, 36, 423–439.

Strzalecki, Tomasz (2024). *Decision Theory*.

Tarsney, Christian (2018). "Exceeding expectations: Stochastic dominance as a general decision theory." *arXiv preprint arXiv:1807.10895*.

Torgersen, Erik Nikolai (1970). "Comparison of experiments when the parameter space is finite." *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 16, 219–249.

Tversky, Amos and Daniel Kahneman (1992). "Advances in prospect theory: Cumulative representation of uncertainty." *Journal of Risk and Uncertainty*, 5, 297–323.

Wald, Abraham (1947). "Foundations of a General Theory of Sequential Decision Functions." *Econometrica*, 15, 279–313.

Woodford, Michael (2014). "An optimizing neuroeconomic model of discrete choice." *Working paper*.

Yaari, Menahem E (1987). "The dual theory of choice under risk." *Econometrica: Journal of the Econometric Society*, 95–115.

Zimmermann, Florian (2020). "The dynamics of motivated beliefs." *American Economic Review*, 110, 337–61.