Online Appendix

for "Strategic Mistakes" by Flynn and Sastry

B State-Separable *vs.* Mutual Information Costs

In this Appendix, we compare the strategic mistakes model with the rational inattention model of Sims (2003). In Sims' rational inattention model, agents flexibly collect signals about an unknown state subject to a continuous cost or hard constraint monotone in the Shannon mutual information between the signal and the state, and then take actions measurable in this signal. Commonly, researchers assume that agents' information choice is unobserved and restrict focus to testing the model's predictions for behavior. This perspective is apparent in the early applications of Sims (2003, 2006), in the decision-theoretic analysis of Caplin et al. (2019, 2022), and in many of the applications surveyed by Maćkowiak et al. (2020). From this perspective, despite their very different motivations—ours from the perspective of costly planning, and Sims (2003)'s from the perspective of costly information acquisition—the strategic mistakes and mutual information models may be each be comparable "candidates" for studying imperfect optimization in a specific equilibrium setting.

We study the similarities and differences between the two models both in theory and practice. We first present an abstract equivalence result which underscores how the models may be equivalent for matching observed data (aggregate and cross-sectional) when the prior distribution is unknown. We then exemplify these differences in a numerical example of a beauty contest, in which the strategic mistakes model has unique predictions and monotone comparative statics while the rational inattention model does not.

B.1 Definitions and an Equivalence Result

We first provide abstract conditions under which a version of the strategic mistakes model makes identical equilibrium predictions to the mutual information model, to build intuition about the comparability and differences of the two approaches.

All information acquisition models that have a posterior separable representation, including mutual information, can be recast as a choice over stochastic choice rules in $\mathcal{P}$ subject to some convex cost functional c (Denti, 2022). The mutual information cost of a stochastic choice rule $P \in \mathcal{P}$ can be decomposed into two terms which we label below:[17]

$$c^{MI}(P) = \underbrace{\sum_{\Theta} \int_{\mathcal{X}} p(x|\theta) \log p(x|\theta) \, \mathrm{d}x \;\, \pi(\theta)}_{\text{State-Separable Term}} - \underbrace{\int_{\mathcal{X}} p(x) \log p(x) \, \mathrm{d}x}_{\text{Cross-State Interactions}} \tag{142}$$

The first term is in fact identical to the state-separable representation (2) with the (quasi-MLRP) kernel $\phi(p) = p \log p$. We label the resulting cost function c^{LSM}, or *logit strategic mistakes*. In a stochastic choice interpretation, this term encodes the agent's desire to *increase* the entropy of the conditional action distributions or play randomly. The second term equals the entropy of the unconditional action distribution and encodes the agents' desire to, on average, anchor toward commonly played actions. This force is absent in the logit strategic mistakes model, and therefore characterizes c^{MI} model compared to its "strategic mistakes cousin" c^{LSM}. Moreover, this decomposition makes clear that there is no conceptual difference in modelling any stochastic choice game with mutual information versus entropic stochastic choice other than that agents have different cost functions, and therefore preferences.

Matějka and McKay (2015) show that the second term ("anchoring") has marginally zero influence on actions when agents' actions are *ex ante* exchangeable, or agents play each action x with equal unconditional probability. From the analyst's perspective, the key free parameter for engineering such exchangeability is the prior $\pi(\cdot)$. We extend this result to show, constructively, that an analyst free to specify the prior can re-construct the equilibrium of a logit strategic mistakes model as an equilibrium of an equivalent game with a mutual information friction provided that a technical condition on payoffs, which ensures that all actions can be made *ex ante* equally attractive, holds:

Lemma 9 (Equilibrium Equivalence)**.** *Suppose that the action space $\mathcal{X}$ is finite. Let $\Omega = (P^*, \hat{X})$ be a symmetric equilibrium for the game $\mathcal{G}^{LSM} = (u(\cdot), \lambda c^{LSM}(\cdot), X(\cdot), \pi'(\cdot), \Theta, \mathcal{X})$. There exists some $\pi'(\cdot) \in \triangle(\Theta)$ such that Ω is an equilibrium of $\mathcal{G}^{LSM}$ and $\mathcal{G}^{MI} = (u(\cdot), \lambda c^{MI}(\cdot), X(\cdot), \pi'(\cdot), \Theta, \mathcal{X})$ if and only if the following linear system has a solution for $\pi' \in \Delta(\Theta)$:*

$$\tilde{U}\pi' = \frac{1}{|\mathcal{X}|} \tag{143}$$

where 1 is a $|\Theta|$ length vector, and $\tilde{U}$ is a $|\mathcal{X}| \times |\Theta|$ matrix with entries:

$$\tilde{u}_{x_i,\theta_j} = \frac{\exp\{u(x_i, \hat{X}(\theta_j), \theta_j)/\lambda\}}{\sum_{x_k \in \mathcal{X}} \exp\{u(x_k, \hat{X}(\theta_j), \theta_j)/\lambda\}} \tag{144}$$

[17] In this expression, we use the definition of the marginal distribution $p(x) = \sum_{\Theta} p(x|\theta)\pi(\theta)$.

Proof. To establish that Ω is an equilibrium of the mutual information model, it is sufficient to establish that P^* solves each individual's optimization problem when they take $\hat{X}$ as given. By Corollary 2 in Matějka and McKay (2015), all interior unconditional choice probabilities $p(x) = \sum_{\theta \in \Theta} p(x|\theta)\pi(\theta)$ in the mutual information model satisfy the following first-order condition:

$$p(x \mid \theta) = \frac{p(x) \exp\{u(x, \hat{X}(\theta), \theta)/\lambda\}}{\sum_{\tilde{x} \in \mathcal{X}} p(\tilde{x}) \exp\{u(\tilde{x}, \hat{X}(\theta), \theta)/\lambda\}} \tag{145}$$

and the following additional constraint:

$$\sum_{\theta \in \Theta} \frac{\exp\{u(x, \hat{X}(\theta), \theta)/\lambda\}}{\sum_{\tilde{x} \in \mathcal{X}} p(\tilde{x}) \exp\{u(\tilde{x}, \hat{X}(\theta), \theta)/\lambda\}} \pi'(\theta) = 1 \tag{146}$$

Observe that, if and only if $p(x) = p(x')$ for all $x, x' \in \mathcal{X}$, then the choice probabilities that solve (145) are

$$p(x \mid \theta) = \frac{\exp\{u(x, \hat{X}(\theta), \theta)/\lambda\}}{\sum_{\tilde{x} \in \mathcal{X}} \exp\{u(\tilde{x}, \hat{X}(\theta), \theta)/\lambda\}} \tag{147}$$

This would verify that the stochastic choice rule P^* is a unique, interior solution to agents' choice problem. Hence it remains only to verify that $p(x) = p(x')$ for all $x, x' \in \mathcal{X}$, or exchangeability, in the agent's optimal program.

It is straightforward to derive such a condition using (146). Stacking equation (146) over all interior $x \in \mathcal{X}$, we obtain the system:

$$\tilde{U}(\{p(x)\}_{x \in \mathcal{X}})\pi' = 1 \tag{148}$$

where:

$$\tilde{u}_{x_i, \theta_j}(\{p(x)\}_{x \in \mathcal{X}}) = \frac{\exp\{u(x_i, \hat{X}(\theta_j), \theta_j)/\lambda\}}{\sum_{x_k \in \mathcal{X}} p(x_k) \exp\{u(x_k, \hat{X}(\theta_j), \theta_j)/\lambda\}} \tag{149}$$

and 1 is a $|\Theta|$ length vector. Thus, there exists a prior consistent with uniform unconditional choice $p(x) = \frac{1}{|\mathcal{X}|}$ if and only if the following linear system has a solution probability vector $\pi' \in \Delta(\Theta)$:

$$\tilde{U}\pi' = |\mathcal{X}|^{-1} 1 \tag{150}$$

where 1 is a $|\Theta|$ length vector, and $\tilde{U}$ is as stated in the result. This completes the proof, with π' solving the given system supporting the equilibrium under the mutual information model. □

The proof establishes from first-order conditions that (143) corresponds with a flat unconditional distribution over actions. The condition ensures that there exists a prior such that all actions yield ex-ante equal payoffs. Heuristically, it is likely to fail if some actions in $\mathcal{X}$ are unappealing regardless of the state or the state space does not have many realizations. The intuition for the first idea is clearest in the extreme case in which some actions are dominated by others for all values of X and θ. In this case, there is nothing that an agent could believe that would ever rationalize playing these actions; and the bridge between the control-cost model and the rational-inattention model cannot be crossed. The intuition for the second relates to the fact that our construction varies the prior to make all actions *ex ante* equally plausible. If, for instance, there are only two states but $N > 2$ actions with very different payoffs from one another in each state, then there is likely no belief that will make all of the actions seem equally appealing.

This result has two practical implications. First, an analyst who is unsure of the physical prior distribution can think of the logit strategic mistakes model as a selection criterion for the mutual information model, across *games indexed by different priors* and, within each prior, a *potentially non-singleton set of equilibria*. This is a general-equilibrium analogue of Matějka and McKay's (2015) insight about the relationship between logit and mutual-information models for individual choice: the former approximates the latter when the analyst does not take a specific stand on anchoring toward defaults. Second, comparative statics in the strategic mistakes model which perturb payoffs $u(\cdot)$ or compare across states $\theta \in \Theta$ may be interpreted, under the conditions of Lemma 9, as comparative statics in a mutual information model *jointly* across the aforementioned features and the physical prior and given a specific equilibrium selection rule.

B.2 Numerically Revisiting The Beauty Contest

We now return to the beauty contest model to illustrate the differences between the strategic mistakes and mutual information models in a practical scenario that maps to the applications of Section 4. Because closed-form solutions are not available for equilibrium action profiles under the mutual information cost, we instead make a feasible approximation of the model on a gridded action space.[18] We will show in this context sharp differences between the predictions of the logit strategic mistakes and mutual information models regarding equilibrium multiplicity and comparative statics, and that these stem from the cross-state interactions embedded in the mutual information cost functional.[19]

[18]This is due to two reasons, in our application with quadratic preferences: the lack of a Gaussian prior and the bounded action space. Moreover, if we had numerically solved a generalized beauty contest with state-dependent costs of mis-optimization, the non-quadratic payoffs would preclude a closed-form mutual-information solution even with a Gaussian prior and unbounded state space.

[19]Note that using logit strategic mistakes will imply that all actions are played with positive probability. To obtain endogenous consideration sets in the strategic mistakes model, we could have instead used a

B.2.1 Environment and Solution Method

For the simplest exposition and comparison to existing work, we use a version of our model that reduces to the linear beauty contest. We study quadratic payoffs of the form,

$$u(x, X, \theta) = \alpha(X, \theta) - \beta(X, \theta)(x - \gamma(X, \theta))^2 \tag{151}$$

and set $\alpha(X, \theta) \equiv 0$, eliminating the pure externality; $\beta(X, \theta) \equiv 1$, giving constant costs of misoptimization; and $\gamma(X, \theta) = (1 - r)\theta + rX$ with $r = 0.85$.[20] The aggregator is the mean. The state space has two points of support, $\Theta = \{\theta_0, \theta_1\} = \{1.0, 2.0\}$. The action space $\mathcal{X}$ is approximated with a 40-point grid between lower endpoint $\underline{x} = 0$ and upper endpoint $\overline{x} = 3$. We use a flat prior with $\pi(\theta_0) = \pi(\theta_1) = \frac{1}{2}$. And we scale both logit and mutual information costs by $\lambda = 0.25$.

Let $p^*(\hat{X}) \in \Delta(\mathcal{X})^2$ return each agent's (unique) optimal stochastic choice rule, expressed as pair of probability mass functions, when they conjecture the equilibrium law of motion $\hat{X} = \left(\hat{X}(\theta_0), \hat{X}(\theta_1)\right)$.[21] As in the proof of our main results, let us define the operator $T : \mathbb{R}^2 \to \mathbb{R}^2$ which constructs essentially the "best response" of aggregates to aggregates by composing the best response with the aggregator:

$$T\hat{X} = \left(X \circ p^*(\theta_0; \hat{X}), X \circ p^*(\theta_1; \hat{X})\right) \tag{152}$$

We define equilibria by first searching over a grid covering $[\underline{x}, \overline{x}]^2$ for approximate fixed points $\hat{X}$, or low $||T\hat{X} - \hat{X}||$, and then using a numerical fixed-point solving algorithm with fine tolerance to confirm equilibria.

B.2.2 Equilibrium Uniqueness and the Contraction Map

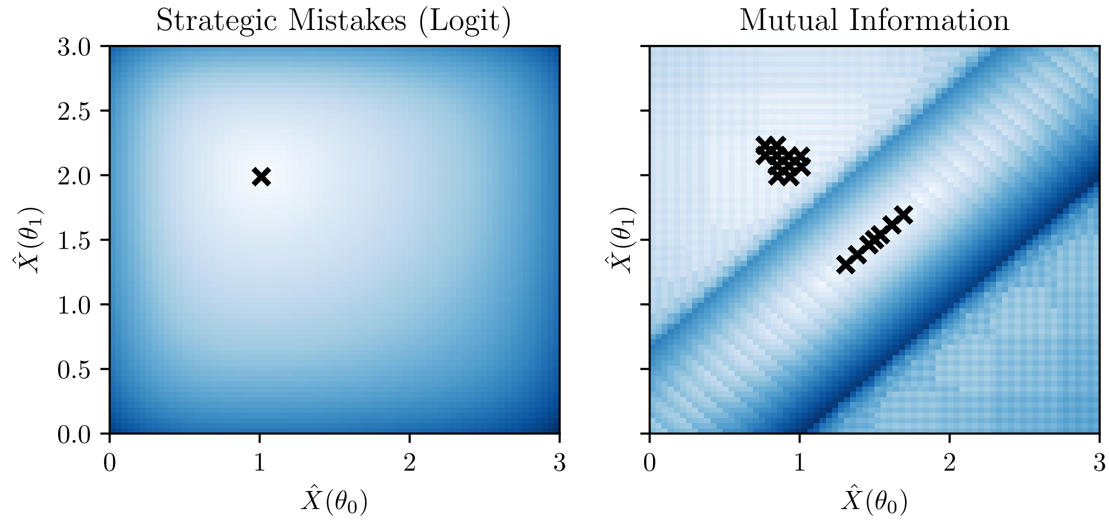
Figure 1 plots the accuracy of the equilibrium conjecture, $||T\hat{X} - \hat{X}||$, in a heat map or two-dimensional histogram over the grid of candidate conjectures. Whiter areas denote that the equilibrium conjecture is closer to the aggregate best response, bluer areas indicate the opposite, and crosses identify equilibria. The strategic mistakes model, on the left, features a

quadratic kernel.

[20] Hellwig and Veldkamp (2009) remark that, for dynamic beauty contests meant to mimic price-setting in New Keynesian models, that $r = 0.85$ is "commonly used." Finally, observe that these payoffs are jointly supermodular in (x, X, θ) but feature bounded complementarity based on the conditions established in the previous section, provided that $r \in (0, 1)$.

[21] For the logit strategic mistakes model, the optimal action profile is known in closed form. For the mutual information model, we apply the Blahut-Arimoto algorithm as described in Caplin et al. (2019) which iterates over the first-order condition for optimal stochastic choice and updates the marginal distribution over actions until convergence.

Figure 1: Equilibria in the Beauty Contest



Note: Each plot is a 2-D histogram of $||T\hat{X} - \hat{X}||$, where $||\cdot||$ indicates the Euclidean norm. Whiter colors indicate smaller values, and hence "closeness to equilibrium." The cross marks represent equilibria, defined such that $||T\hat{X} - \hat{X}|| < 10^{-6}$.

single-peaked surface and a single equilibrium. This is consistent with our theoretical results, and with the fixed-point condition (152) being a contraction. The mutual information model, on the right, features a non-monotone surface and 18 confirmed equilibria.

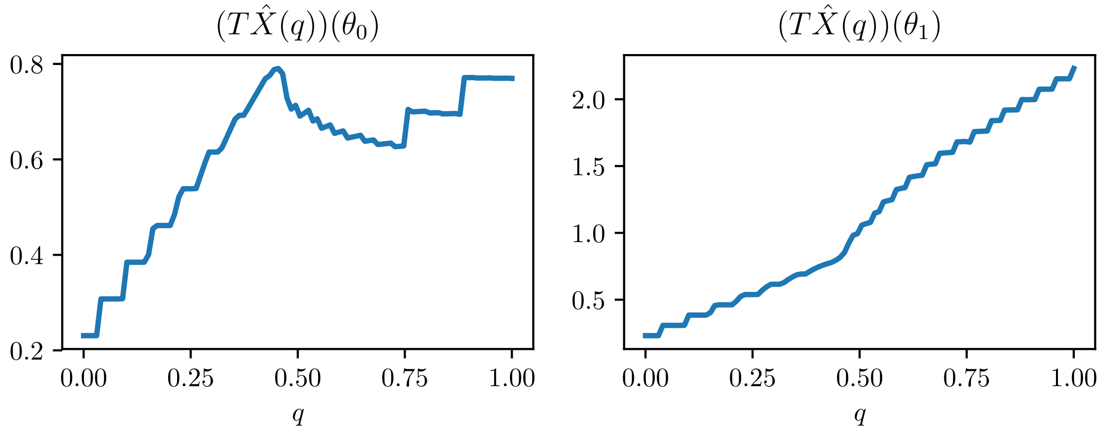
We now deconstruct further the failure of the contraction map argument for the mutual information model. Recall, in our proof of Theorem 1, that establishing monotonicity and discounting for the equilibrium operator T required first showing monotone and smooth comparative statics for the single-agent decision problem. To "test" this in the mutual information model, we parameterize a path that increases the equilibrium conjecture of $\hat{X}$ from $(0, 0)$ to one of its equilibrium values.[22] Formally, if we label this chosen equilibrium as $X^*_{MI} = (X^*_{\text{MI}}(\theta_0), X^*_{\text{MI}}(\theta_1))$, we consider points indexed by $q \in [0, 1]$:

$$\hat{X}(q) = (q \cdot X^*_{\text{MI}}(\theta_0), q \cdot X^*_{\text{MI}}(\theta_1)) \tag{153}$$

and the aggregate best response $T\hat{X}(q)$. Figure 2 shows each element of $T\hat{X}(q)$ as a function of q. The first element, plotted in the left panel, is (i) non-monotone and (ii) discontinuous in the equilibrium conjecture. In the language of the price-setting application, the mutual information model does not predict that expecting a higher price level increases one's own

[22] We pick the equilibrium with the largest value of $\hat{X}(\theta_1) - \hat{X}(\theta_0)$.

Figure 2: Partial Equilibrium Comparative Statics With Mutual Information



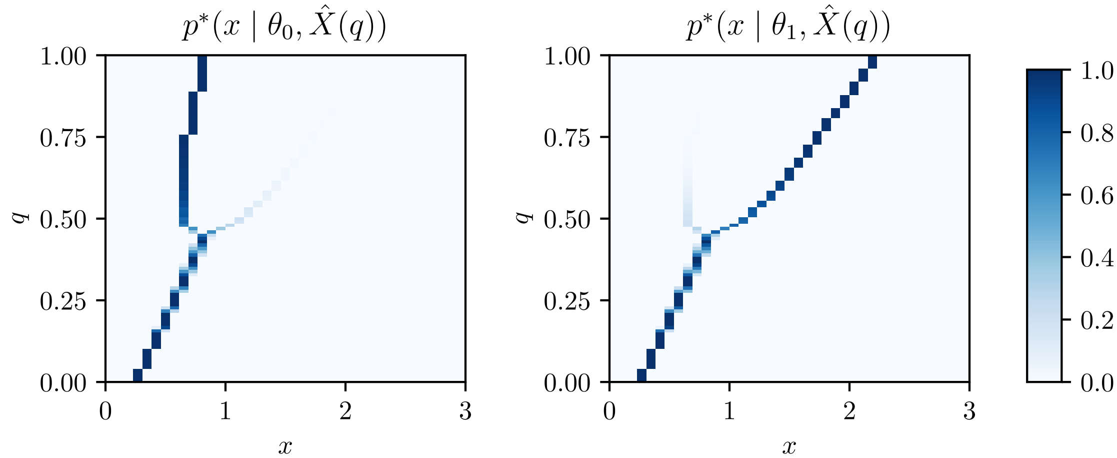
Note: These plots show aggregate best response $T\hat{X}$ in state θ_0 (left pane) and θ_1 (right pane) along the path (153) for the equilibrium conjecture.

price, even though the payoff to setting a higher price has globally increased; and when prices increase, they may jump suddenly.

To better understand the agent's behavior along this path, we show in Figure 3 a two-dimensional histogram of the stochastic choice patterns conditional on each conjecture indexed by q. Equilibrium strategies are mostly supported on either one or two points. This sparsity of support is formally described by Jung et al. (2019) and Caplin et al. (2019) in discrete- and continuous-action variants of the mutual information model as a natural consequence of the lowered marginal costs (or, more loosely, "increasing returns to scale") of allocating probability mass to frequently played actions. Sparse behavior is a characteristic feature of the optimal policy in price-setting applications studied by Matějka (2015) and Stevens (2019). In our example, the optimal policy switches between one and two support points around $q = 0.45$. Matějka (2015) refers to such behavior as a *bifurcation* in the optimal policy. As q increases after the bifurcation point, the optimal policy in Figure 3 pushes the larger and smaller support points away from one another. This violates monotonicity in the sense of first-order stochastic dominance, and therefore can lead to a non-monotone aggregate with respect to some admissible aggregators. Under our chosen aggregator, this behavior causes $X(\theta_0)$ to decrease, as evident in the left panel of Figure 2.

Our observation is that this force can support multiple equilibria in coordination games because it breaks the contractive properties of the equilibrium map. These multiple equilibria are not, in our reading, very easily interpretable given that choices have an ordinal interpretation, payoffs leverage this interpretation in their definition of complementarity, and

Figure 3: Stochastic Choice Strategies With Mutual Information



Note: Each slice on the the vertical axis (q) gives the probability distribution of actions in state 0 (left) or 1 (right), represented via a "heat map" (scale on the right). The path of the equilibrium conjecture corresponds to the same in Figure 2.

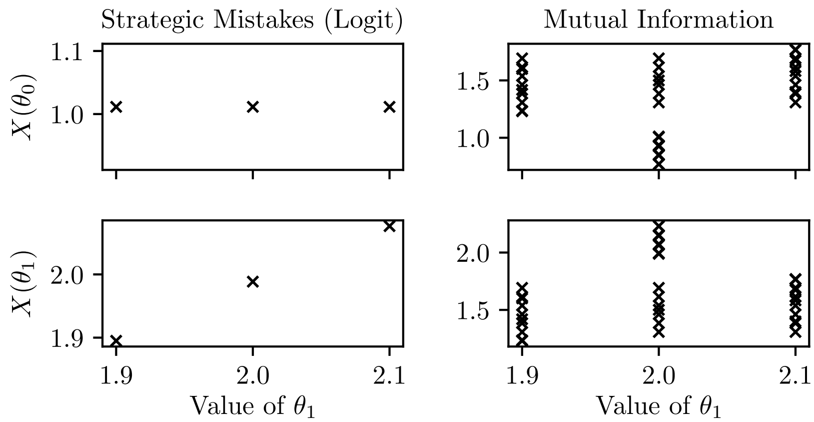
agents have a continuum of possible options. This reasoning is quite stark in the price-setting application which Matějka (2015) and Stevens (2019) study with mutual information. While it is quite reasonable that a single firm wavers between charging \$1.99 and \$2.99 for its product, and indeed Stevens (2019) provides direct evidence for such behavior, it is a much stronger prediction that an entire (symmetric) economy of firms switches between a coordinated equilibrium of charging (\$1.99, \$2.99), respectively in each of two states of nature, to a different equilibrium of charging (\$1.98, \$3.00).

B.2.3 Equilibrium Comparative Statics

A point emphasized in our theory, and in particular the transition from Theorem 1 (existence and uniqueness) to Theorem 2 and Theorem 3 (monotone aggregates and precision), was that the contraction map structure goes hand-in-hand with proving equilibrium comparative statics. We now illustrate the contrast between comparative statics with strategic mistakes and information acquisition in our model. We vary the value of the higher state θ_1 on the grid $\{1.90, 2.00, 2.10\}$ and re-solve for all equilibria of each model. Our main results for the strategic mistakes model suggest that $X^*(\theta_1)$ should monotonically increase in that model while $X^*(\theta_0)$ stays constant, owing to the separability of decisions by state. For the mutual information model, there are no equivalent theoretical results.

Figure 4 plots the equilibria of each model as a function of the chosen θ_1. In the strategic mistakes model, we verify the predicted comparative statics across unique equilibria. In

Figure 4: Equilibrium Comparative Statics in the Beauty Contest



Note: Each cross mark is an equilibrium, under the strategic mistakes (left) and mutual information (right) models, for different values of θ_1. Note the different axis scales in each figure.

the mutual information model, we observe non-monotone comparative statics as equilibria move in and out of the set. Thus, while a mutual-information model may be an appealing laboratory to study specific behaviors like discrete pricing, it may not lend itself to straightforward comparative statics analysis conditional on this feature outside of specific numerical calibrations.[23]

C State-Separable Costs in Binary-Action Games

In this Appendix, we adapt our analysis to study binary-action games, which are also common for modeling coordination phenomena in macroeconomics and finance. We first provide results ensuring existence, uniqueness and monotone comparative statics. We next apply our results to study the "investment game," introduced by Carlsson and Van Damme (1993) and studied recently by Yang (2015) and Morris and Yang (2022). Bridging our continuous-action and binary-action analyses, we finally discuss how the the action space can have a large bearing on our model's uniqueness predictions. This may be an important consideration for researchers when the choice of action space is primarily based on analytical convenience and

[23] Of course whether this is a "bug" or instead a "feature," reflecting the unstable coordinational nature of activities like price-setting, is an open question that merits additional research. Stevens (2019), for instance, uses a model of coarse pricing with mutual-information costs to match micro-level evidence on pricing strategies and macroeconomic dynamics for aggregates. The micro-economic calibration builds the case that non-uniqueness and ambiguous comparative statics may indeed be features of the "correct" descriptive model of this setting.

not descriptive realism regarding adjustment on an extensive margin.

C.1 Existence, Uniqueness, and Comparative Statics

We now study the same environment as Section 2 with the sole change that agents now have a binary action set $\mathcal{X} = \{0, 1\}$.[24] Let $p(\theta)$ denote the probability that a given agent plays action 1 in state θ. It is without loss of generality to restrict to the aggregator $X(p(\theta)) = p(\theta)$, since transformations of this aggregate can be applied within payoffs, and we adopt this convention throughout. Given a conjecture for the law of motion of the aggregate $\hat{p}$ and state $\theta \in \Theta$, we define the cost-adjusted benefit of playing action 1 over action 0 as:

$$\Delta\tilde{u}(\hat{p}(\theta), \theta) \equiv \frac{u(1, \hat{p}(\theta), \theta) - u(0, \hat{p}(\theta), \theta)}{\lambda(\hat{p}(\theta), \theta)} \tag{154}$$

We let $\Delta\tilde{u}_X$ denote this function's derivative in the first argument.

We now provide an existence and uniqueness result. To do so, we place the following regularity condition on the stochastic choice functional:[25]

Assumption 7. *The kernel of the cost functional satisfies the Inada condition* $\lim_{x\to 0} \phi'(x) = -\infty$. *Moreover,* ϕ'' *is globally strictly convex.*[26]

This rules out stochastic choice rule's being concentrated on only one of the two actions in any state. The result follows:[27]

Proposition 2. *Suppose that* ϕ *satisfies assumption 7 and* $\Delta u(p, \theta)$ *is continuously differentiable in its first argument. There exists an equilibrium. All equilibria are symmetric. A sufficient condition for there to be a unique* $p^*(\theta)$ *is that:*

$$\max_{p\in[0,1]} \Delta\tilde{u}_X(p, \theta) < 2\phi''\left(\frac{1}{2}\right) \tag{155}$$

A sufficient condition for there to be a unique p^* *is that (155) holds for all* $\theta \in \Theta$.

Proof. Under Assumption 7, for any θ, we have that $p^*(\theta) \in (0, 1)$. Thus equilibrium is characterized by the first-order condition obtained by moving probability of playing zero to

[24] Naturally, all integrals are now replaced with summations and density functions by mass functions.

[25] For existence, this can be weakened in the obvious way: the objective need only be continuous. We present results with this stronger assumption for brevity.

[26] Note that, in view of the Inada condition, it is impossible for ϕ'' to be globally strictly concave.

[27] One can extend this result in the obvious way beyond the differentiability assumption to allow for Lipschitz continuous $\Delta u(p, \theta)$. Naturally, the key property being ruled out is a sudden threshold around which the gains from playing action 1 change discontinuously.

playing one. Thus, the condition characterizing equilibrium is given by:

$$\Delta u(p^*(\theta), \theta) = \lambda(p^*(\theta)) \left(\phi'(p^*(\theta)) - \phi'(1 - p^*(\theta))\right) \tag{156}$$

To prove uniqueness for a given θ it is sufficient to prove that the minimal slope of the RHS exceeds the maximal slope of the LHS:

$$\max_{p \in [0,1]} \Delta\tilde{u}_X(p, \theta) < \min_{p \in [0,1]} \phi''(p) + \phi''(1 - p) \tag{157}$$

If ϕ'' is strictly convex, then the problem is solved by solving the FOC:

$$\phi'''(p) = \phi'''(1 - p) \tag{158}$$

As ϕ'' is strictly convex, ϕ''' is strictly increasing and is therefore invertible. Thus the unique solution is $p = \frac{1}{2}$ and the minimized value is $2\phi''(\frac{1}{2})$. Applying this argument state by state yields the global condition. □

Condition (155) checks the maximum value of complementarity (left-hand-side) against the lowest value for the slope of the marginal cognitive cost of investing (right-hand-side), which is realized at $p = \frac{1}{2}$.[28] We will provide a simple graphical intuition for this condition in the upcoming example.

It is moreover simple to establish when the aggregate $p^*(\theta)$ increases in θ. As in our main analysis, this simply requires supermodularity of payoffs in (x, p, θ), or that higher actions by others and states are complementary with playing $x = 1$:

Assumption 8 (Joint Supermodularity)**.** *The cost-adjusted benefit of playing action 1 over action 0 satisfies, for all $p' \geq p, \theta' \geq \theta$:*

$$\Delta\tilde{u}(p', \theta') \geq \Delta\tilde{u}(p, \theta) \tag{159}$$

Proposition 3. *Suppose that Assumptions 7 and 8 hold, and that the inequality in Equation 155 holds for all $\theta \in \Theta$ so that there is a unique equilibrium p^*. The unique equilibrium $p^*(\theta)$ is monotone increasing in θ.*

Proof. Under Assumption 7, the equilibrium is characterized by Equation 156. Under the assumption that the inequality in Equation 155 holds, there is a unique solution $p^*(\theta)$ for all $\theta \in \Theta$. Note that that this unique equilibrium occurs when $\Delta\tilde{u}(p, \theta)$ intersects $\phi'(p) - \phi'(1-p)$

[28] That $p = \frac{1}{2}$ is such a point can be derived by noting the symmetry of the state-separable cost around $p = \frac{1}{2}$ and the convexity of ϕ.

from above. Moreover, by Assumption 8 we know that $\Delta\tilde{u}(p,\theta)$ is increasing in (p,θ). Thus, when we take $\theta' \geq \theta$, we know that the unique intersection occurs for $p^*(\theta') \geq p^*(\theta)$. $\square$

Analogous results with general information acquisition or stochastic choice, by contrast, require more extensive analysis (see, *e.g.,* Yang, 2015; Morris and Yang, 2022).

C.2 Application: The Investment Game

We now apply these results in a variant of the binary-action investment game introduced by Carlsson and Van Damme (1993), which models coordination motives in financial speculation. Each agent chooses an action $x \in \{0,1\}$, or "not invest" and "invest." The state of nature $\theta \in \Theta \subseteq \mathbb{R}$ scales the desirability of investing independent of other conditions. Agents' payoffs depend on the action, the total fraction of investing agents, and the state of nature separably and linearly:

$$u(x,p,\theta) = x(\theta - r(1-p)) \tag{160}$$

where $r \geq 0$ scales the degree of strategic complementarity between investment decisions.

It is straightforward to derive the following fixed-point equation that describes the equilibria of the model when ϕ satisfies the Inada condition in Assumption 7 and $\lambda(p,\theta) \equiv 1$:

$$\theta + rp(\theta) - r = \phi'(p(\theta)) - \phi'(1-p(\theta)) \tag{161}$$

Equilibrium is guaranteed to be unique by Proposition 3 provided that the following condition holds relating strategic complementarity r with the second derivative of the kernel ϕ:

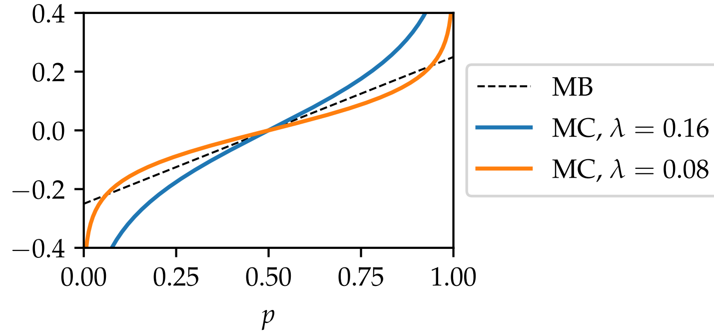
$$r < 2\phi''\left(\frac{1}{2}\right) \tag{162}$$

This condition is independent of the state space Θ or the prior. But it does depend on the scale and character of cognitive costs through $\phi''\left(\frac{1}{2}\right)$.

Condition (162) admits the following interpretation about uniqueness with vanishing costs under arbitrary functional forms. For any positive (but arbitrarily small) level of strategic complementarity, and with a sufficiently rich state space, there will be multiple equilibria for a sufficiently small cost of stochastic choice:

Corollary 5. *Consider a family of investment games $\{\mathcal{G}_\lambda : \lambda \in (0,L]\}$ with fixed payoffs, action space, and state space, each with the re-scaled cost functional for some common $\hat{\phi}$*

Figure 5: Multiplicity in the Investment Game



Note: The dotted line is the marginal benefits of investing more often as a function of others' investment probability, or the right-hand side of (161). The blue and orange lines are the marginal costs of investing more often under respectively more and less severe costs of stochastic choice. Each intersection is an equilibrium.

that satisfies Assumption 7, i.e., $\phi_\lambda = \lambda\hat{\phi}$. Then, for all

$$\lambda > L^* := \frac{r}{2\hat{\phi}''\left(\frac{1}{2}\right)} \tag{163}$$

game $\mathcal{G}_\lambda$ has a unique action profile $(p^*(\theta))_{\theta\in\Theta}$. Conversely, when $\lambda < L^*$, there exists at least some $\theta^* \in \mathbb{R}$ such that the equilibrium of $\mathcal{G}_\lambda$ is not unique if $\theta^* \in \Theta$.

Proof. Recall that for any ϕ_λ (owing to $\hat{\phi}$ satisfying Assumption 7), we have that:

$$\theta + rp^*(\theta) - r = \phi_\lambda'(p^*(\theta)) - \phi_\lambda'(1 - p^*(\theta)) \tag{164}$$

Consider state $\theta^* = \frac{r}{2}$. In this state, we have that $p^*(\theta^*) = \frac{1}{2}$ is an equilibrium. Moreover, see that the slope of the LHS in p is given by r and the slope of the RHS in p at $p = \frac{1}{2}$ is given by $2\lambda\hat{\phi}''(\frac{1}{2})$. Hence, when $\lambda < \frac{r}{2\hat{\phi}''(\frac{1}{2})}$, we have that the slope of the LHS exceeds the slope of the RHS. But we know that the RHS is continuous on $(0, 1)$ and that $\lim_{p\to1} \phi'(p) - \phi'(p) = \infty$. Thus, the RHS must intersect the LHS from below for some other $p \in (\frac{1}{2}, 1)$. Thus, in state θ^*, if $\lambda < \frac{r}{2\hat{\phi}''(\frac{1}{2})}$ there are multiple $p^*(\theta)$ that can arise in equilibrium. Consequently, if $\theta^* \in \Theta$ and $\lambda < \frac{r}{2\hat{\phi}''(\frac{1}{2})}$, we have that equilibrium is not globally unique. The final claim that we have global uniqueness for $\lambda > \frac{r}{2\hat{\phi}''(\frac{1}{2})}$ follows immediately from Theorem 2. □

The result contrasts with Corollary 1 which showed limit uniqueness in the generalized beauty contest. We will further discuss this issue in Section C.4.

To illustrate the uniqueness result, we consider a specialization of the model in which the kernel function is $\hat{\phi}(x) = x \log x$. In this case, $\phi''(0.5) = 2$ and the cost threshold for uniqueness is $L^* = \frac{r}{4}$.[29] Figure 5 illustrates the scope for multiplicity in a benchmark parameter case of this logit model. We fix $r = 0.50$, and $\theta = 0.25$, the state such that a 50% aggregate investment corresponds with zero payoff. The dotted black line is the "Marginal Benefit," which corresponds with the left-hand-side of (161). The blue and orange lines are the the "Marginal Cost" of increasing the investing probability, or the right-hand-side of (161), with respectively higher and lower values of λ or costs of attention. By construction, there is an equilibrium with $p = \frac{1}{2}$ for any value of λ. Whether or not there are additional equilibria corresponding to more "confident" play, or p closer to 0 or 1, depends on the *slope* of these marginal costs. When λ is high (blue line), it is costly to play more certainly and hence there is only one intersection with the dotted line. When λ is low (orange line), marginal costs cross marginal benefits from above at $p = 0.5$. This visualizes a violation of the condition in Proposition 2. As a result there are two more "confident" equilibria near $p = 0$ and $p = 1$.

The right-hand-side of the confident-wavering condition (162) is a well-defined moment which researchers may try to calibrate via laboratory experiments and could interpret in our model without taking a stand on the entire ϕ function. In this way, (162) can be read as a sufficient statistic gauge of the potential for multiplicity and fragility that relies only on one informative aspect of the underlying stochastic choice model.

C.3 State-Separable *vs.* Mutual Information Costs

In the vein of our main analysis' comparison of beauty contests with strategic mistakes and mutual information, we now compare the investment game under logit strategic mistakes with the equivalent game under mutual information, as studied by Yang (2015). Observe first that the mutual information model does not always admit an interior solution. Intuitively, if agents place an arbitrarily high prior weight on fundamentals always being very high or very low, they may decide to unconditionally invest or dis-invest without learning anything. These scenarios are ruled out by respectively assuming $\mathbb{E}_\pi \left[\exp\{\lambda^{-1}\theta\}\right] > \exp\{\lambda^{-1} r\}$ and $\mathbb{E}_\pi \left[\exp\{-\lambda^{-1}\theta\}\right] > 1$. No analogue of either is possible in the strategic mistakes model with logistic choice which always features positive probability of playing both actions in all states, so these conditions *a fortiori* rule out an application of Lemma 9. Nonetheless, after ruling out these cases, we can show the following:

[29]This is exactly the condition obtained by Yang (2015) for this game with information acquisition costs proportional to mutual information. This foreshadows a deeper connection which we will explore in the next subsection.

Corollary 6. *Compare identical investment games* $\mathcal{G}^{LSM}$ *and* $\mathcal{G}^{MI}$*, distinguished by their costs of stochastic choice, scaled by a common scalar* λ*. Assume*

1. *(Interiority)* $\mathbb{E}_\pi\left[\exp\{\lambda^{-1}\theta\}\right] > \exp\{\lambda^{-1}r\}$ *and* $\mathbb{E}_\pi\left[\exp\{-\lambda^{-1}\theta\}\right] > 1$
2. *(Global uniqueness)* $r < 4\lambda$

Each game has a unique equilibrium $(p^{LSM}(\cdot), p^{MI}(\cdot))$*. Moreover,*

$$\begin{cases} p^{LSM}(\theta) = p^{MI}(\theta), \forall\theta & \text{if } \sum_\Theta p^{MI}(\theta)\,\pi(\theta) = 1/2, \\ p^{LSM}(\theta) < p^{MI}(\theta), \forall\theta & \text{if } \sum_\Theta p^{MI}(\theta)\,\pi(\theta) > 1/2, \\ p^{LSM}(\theta) > p^{MI}(\theta), \forall\theta & \text{if } \sum_\Theta p^{MI}(\theta)\,\pi(\theta) < 1/2. \end{cases} \tag{165}$$

Proof. It follows from Proposition 2 of Yang (2015), that when $\mathbb{E}_\pi\left[\exp\{\lambda^{-1}\theta\}\right] > \exp\{\lambda^{-1}r\}$ and $\mathbb{E}_\pi\left[\exp\{-\lambda^{-1}\theta\}\right] > 1$, the equilibria of the game with mutual information cost are characterized by:

$$\theta + rp^{\text{MI}}(\theta) - r = \lambda\left[\ln\left(\frac{p^{\text{MI}}(\theta)}{1-p^{\text{MI}}(\theta)}\right) - \ln\left(\frac{\bar{p}^{\text{MI}}}{1-\bar{p}^{\text{MI}}}\right)\right] \tag{166}$$

for all $\theta \in \Theta$ where $\bar{p}^{\text{MI}} = \sum_\Theta p^{\text{MI}}(\theta)\pi(\theta)$. It moreover follows from Proposition 3 of Yang (2015) that when $r < 4\lambda$, this model features a unique equilibrium. Recall that when $r < 4\lambda$ our model with entropic stochastic choice also features a unique equilibrium and this is characterized by:

$$\theta + rp^{L}(\theta) - r = \lambda\left[\ln\left(\frac{p^{L}(\theta)}{1-p^{L}(\theta)}\right)\right] \tag{167}$$

Moreover, when $\bar{p}^{\text{MI}} > \frac{1}{2}$, we have that $\ln\left(\frac{\bar{p}^{\text{MI}}}{1-\bar{p}^{\text{MI}}}\right) > 0$, when $\bar{p}^{\text{MI}} = \frac{1}{2}$, we have that $\ln\left(\frac{\bar{p}^{\text{MI}}}{1-\bar{p}^{\text{MI}}}\right) = 0$ and when $\bar{p}^{\text{MI}} < \frac{1}{2}$, we have that $\ln\left(\frac{\bar{p}^{\text{MI}}}{1-\bar{p}^{\text{MI}}}\right) < 0$. It is then immediate that $p^L(\theta) < p^{\text{MI}}(\theta)$ when $\bar{p}^{\text{MI}} > \frac{1}{2}$, $p^L(\theta) = p^{\text{MI}}(\theta)$ when $\bar{p}^{\text{MI}} = \frac{1}{2}$, and $p^L(\theta) > p^{\text{MI}}(\theta)$ when $\bar{p}^{\text{MI}} < \frac{1}{2}$. □

Conditional on interiority, anchoring in the mutual information model distorts the choice probabilities but perhaps more surprisingly is completely separable from the game's uniqueness properties. More formally, in binary-action games with mutual information, the only difference between the strategic mistakes model with entropy is that log-odds ratio $\log\left(\frac{p(\theta)}{1-p(\theta)}\right)$ in state $\theta \in \Theta$ differs across the models by a state-independent additive constant. In our earlier graphical analysis, this can be seen as a vertical shift of the marginal cost curve. Thus, our confident wavering argument applies directly to the mutual information model and offers an alternative window into the main result of Yang (2015). This separability of anchoring from uniqueness properties with binary actions may be an independently useful insight in other models with mutual information cost.

C.4 Discussion: Global *vs.* Local Mistakes

Binary-action settings are sometimes used as a convenient metaphor for underlying environments with many possible actions—for instance, simplifying financial speculation as the choice between extremes of investing and dis-investing instead of a continuous portfolio choice. Our analysis reveals that, in models of stochastic choice, the restriction to two extreme actions may significantly change the character of the game because it removes the possibility of local substitution of actions. The binary-action game allows for "global mistakes," like fully investing when fully disinvesting is instead optimal, that impose discontinuously different externalities and can support multiple equilibria. Our benchmark continuous-action model implies by contrast that agents make "local mistakes" like substituting an optimal action with an alternative that is sub-optimal but nearby in the action space. Whether an analyst should use the binary-action or continuous-action model then depends on the problem at hand and how seriously they take the prediction of global substitution relative to the potential loss in tractability.

Our results also contrast with those in the global games literature in which there is, instead of stochastic choice, vanishing private measurement error in observing the fundamental (Carlsson and Van Damme, 1993; Morris et al., 1995; Frankel et al., 2003). When combined with the earlier observation linking strategic mistakes with cross-sectional heterogeneity in payoff functions (Section 2.1), our results draw a sharp distinction between measurement errors for payoffs (studied here, which do not yield limit uniqueness) and measurement errors for fundamentals (studied in the aforementioned literature, which do yield limit uniqueness). One way of thinking about the difference is that the "contagion" argument formalized in the above references, which shows that having dominant actions in specific states iteratively implies unique rationalizable actions in neighboring states, has no analogue in the present model with no interim beliefs or cross-state reasoning. A different interpretation is that the mere observation that agents have trembling hands is not sufficient to imply the sharp and specific predictions of canonical global games, a point also made by Yang (2015) and Morris and Yang (2022).

References

Caplin, A., Dean, M., and Leahy, J. (2019). Rational inattention, optimal consideration sets, and stochastic choice. *The Review of Economic Studies*, 86(3):1061–1094.

Caplin, A., Dean, M., and Leahy, J. (2022). Rationally inattentive behavior: Characterizing and generalizing shannon entropy. *Journal of Political Economy*, 130(6):1676–1715.

Carlsson, H. and Van Damme, E. (1993). Global games and equilibrium selection. *Econometrica*, pages 989–1018.

Hellwig, C. and Veldkamp, L. (2009). Knowing what others know: Coordination motives in information acquisition. *The Review of Economic Studies*, 76(1):223–251.

Jung, J., Kim, J. H. J., Matějka, F., and Sims, C. A. (2019). Discrete Actions in Information-Constrained Decision Problems. *The Review of Economic Studies*, 86(6):2643–2667.

Maćkowiak, B., Matejka, F., and Wiederholt, M. (2020). Rational inattention: A review. Discussion Paper 15408, CEPR.

Matějka, F. (2015). Rationally Inattentive Seller: Sales and Discrete Pricing. *The Review of Economic Studies*, 83(3):1125–1155.

Matějka, F. and McKay, A. (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105(1):272–98.

Morris, S. and Yang, M. (2022). Coordination and continuous stochastic choice. *The Review of Economic Studies*, 89(5):2687–2722.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690.

Sims, C. A. (2006). Rational inattention: Beyond the linear-quadratic case. *American Economic Review*, 96(2):158–163.

Stevens, L. (2019). Coarse Pricing Policies. *The Review of Economic Studies*, 87(1):420–453.

Yang, M. (2015). Coordination with flexible information acquisition. *Journal of Economic Theory*, 158:721–738.