Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 212 (2023) 105704

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Strategic mistakes ☆

Joel P. Flynn [a,*,1], Karthik A. Sastry [b,*,2]

[a] *Yale, United States of America*
[b] *Princeton, United States of America*

Received 14 August 2021; final version received 18 June 2023; accepted 14 July 2023
Available online 1 August 2023

Abstract

To study the equilibrium implications of decision frictions, we introduce a new class of control costs in continuum-player, continuum-action games in which agents interact via an aggregate of the actions of others. The costs that we study accommodate a rich class of decision frictions, including *ex post* misoptimization, imperfect *ex ante* planning, cognitive constraints that depend endogenously on the behavior of others, and consideration sets. We provide primitive conditions such that equilibria exist, are unique, are efficient, and feature monotone comparative statics for action distributions, aggregates, and the size of agents' mistakes. We apply the model to make robust equilibrium predictions in a monetary business-cycle model of price-setting with planning frictions and a model of consumption and savings during a liquidity trap when endogenous stress worsens decisions.

JEL classification: D5; D8; E1; E7

Keywords: Decision frictions; Equilibrium; Monotone comparative statics; Aggregative games; Stochastic choice; Behavioral macroeconomics

☆ An earlier version of this paper was entitled "Strategic Mistakes in Large Games." We thank Daron Acemoglu, George-Marios Angeletos, Oğuzhan Çelebi, Roberto Corrao, Drew Fudenberg, Chen Lian, Nicolas Minudri, Stephen Morris, George Nikolakoudis, Chris Sims, Alp Simsek, Tomasz Strzalecki, John Sturm, Iván Werning, Alexander Wolitzky, the editor Guillermo Ordoñez, two anonymous referees, and participants in the MIT Macro Lunch, MIT Theory Lunch, MIT Behavioral Economics Lunch, and MIT Sloan Finance Lunch for helpful comments. First posted version: July 29, 2020.

* Corresponding authors.
E-mail addresses: joel.flynn@yale.edu (J.P. Flynn), ksastry@princeton.edu (K.A. Sastry).

1 Yale Department of Economics, 30 Hillhouse Avenue, New Haven, CT, 06511.

2 Princeton Department of Economics, Julis Romo Rabinowitz Building, Princeton, NJ 08544.

https://doi.org/10.1016/j.jet.2023.105704

1. Introduction

People commonly make mistakes that affect others. Consider a monopolistically competitive firm choosing its price to maximize profits, taking into account projected demand and competitors' prices. The complexity of firms' decision-making processes makes clear that even though the problem is well-defined and an ideal solution surely exists, determining that solution is difficult. Thus, firms may fail to set the optimal price. Such a deviation from the ideal price may affect all other competitors' *benefits* from setting the right price—for instance, by altering the residual demand that they face. Moreover, the pricing of other firms may directly influence the *costs* of setting the right price—for instance, if tough competition induces managerial stress that contributes to worse decision-making. Thus, observed pricing arises from a process of *strategic mistakes*: the combination of imperfect optimization and strategic interaction that may affect both the benefits and the costs of precise decision-making.

To study such strategic mistakes, this paper introduces a model of non-parametric, state-dependent stochastic choice in continuum-player games with a continuum of actions. Agents' payoffs depend on their own action, an exogenous state, and a one-dimensional aggregate of the cross-sectional distribution of others' actions. Such a setting is ubiquitous in macroeconomic models of price-setting (Woodford, 2003; Maćkowiak and Wiederholt, 2009; Costain and Nakov, 2019), production (Angeletos and La'O, 2010, 2013; Benhabib et al., 2015; Chahrour and Ulbricht, 2023), and beauty-contest games more generally (Morris and Shin, 2002; Angeletos and Pavan, 2007; Bergemann and Morris, 2013; Huo and Pedroni, 2020).

Agents face a problem of *costly control*: conditional on their conjecture for fundamentals and others' actions, they pick a stochastic choice pattern that trades off playing the best actions with a cost that penalizes playing too precisely. We introduce a new family of control cost functionals that are *state-separable*, *i.e.,* total control costs are additive over states. These costs allow us to model several kinds of decision frictions that have not previously been jointly studied. The first is *ex post* misoptimization, as in the literatures on control costs (Stahl, 1990; Van Damme, 1991) and quantal response equilibrium (McKelvey and Palfrey, 1995; Goeree et al., 2016), in which agents' imprecise play responds to strategic incentives within a given state of the world. The second is *ex ante* planning frictions, as in the literature on costly information acquisition in games (see *e.g.,* Yang, 2015; Morris and Yang, 2022; Hébert and La'O, 2022; Denti, 2023), whereby agents must weigh the benefits of precise planning for a state with the cost of that state never being realized. The third is exogenous and endogenous state-dependence in control costs, as in Hébert and La'O (2022) and Angeletos and Sastry (2023). The fourth is equilibrium determination of agents' consideration sets, *i.e.,* the subset of actions that they play, as in Matějka (2015) and Stevens (2019).

We show that, despite the rich behavioral patterns that our model accommodates, equilibrium analysis remains tractable. Concretely, we provide four theoretical results that provide conditions for equilibrium existence, uniqueness, efficiency, and monotone comparative statics for actions, aggregates, and the size of agents' mistakes.

Toward establishing the existence and uniqueness of equilibrium, we first characterize equilibrium as a functional fixed-point equation for the cross-sectional distribution of actions and provide primitive conditions under which this equilibrium fixed-point operator is a contraction. This result follows from showing first that agents' state-dependent stochastic choice rules are *monotone*, *i.e.,* they are increasing in the sense of first-order stochastic dominance when aggregate actions are higher, and *discounted*, *i.e.,* increasing aggregate outcomes has a less than one-for-one effect on agents' stochastic best replies. This requires three primitive conditions:

(i) that agents' actions and aggregates are jointly complementary for physical payoffs and the psychological costs of precise optimization; (ii) that this complementarity is dominated by the concavity of agents' physical payoffs relative to their psychological costs; and (iii) a technical restriction on the shape of agents' cost functionals that allows us to translate dominance in payoff units into first-order stochastic dominance in the space of stochastic choice rules. Moreover, we show that the last of these assumptions is satisfied under the two leading cost functions in the control costs literature: entropic and quadratic costs. Second, we show that, if the equilibrium aggregator is (i) increasing in agents' actions and (ii) such that level shifts of the action distribution have less than one-for-one effects on aggregates, then the equilibrium fixed-point operator is a contraction. These assumptions on aggregation are satisfied under common aggregators, such as those that take the mean or the median of the cross-sectional action distribution. Finally, since the equilibrium operator is a contraction, the existence and uniqueness of equilibrium (Theorem 1) follows.

We next study equilibrium comparative statics. First, if actions, aggregates, and the state are jointly complementary for agents' physical payoffs and psychological costs, then the unique equilibrium action distribution and aggregate are monotone in the state (Theorem 2). Under a further condition that payoffs depend only on the distance between one's own action and some optimal action, we show that the size of agents' mistakes is monotone in the state when the ratio between the stakes of misoptimization and the cost of precise optimization is monotone in endogenous and exogenous states (Theorem 3).

Turning to normative analysis, we provide a necessary condition for the efficiency of the unique equilibrium: the average marginal physical benefit of increasing the aggregate action must equal the average marginal psychological cost of so doing (Theorem 4).

We finally employ our results in two macroeconomic applications. The first application is to price-setting in a monetary economy à la Woodford (2003) and Hellwig and Venkateswaran (2009), but where firms face *ex ante* planning frictions: firms must plan for what prices to set across contingencies for the realized level of the money supply and inflation. We derive and interpret conditions under which the aggregate price level, the distribution of prices, and the dispersion of prices are monotone increasing in an exogenous shock to the money supply in the unique equilibrium. We use these results to give a costly-planning explanation for the empirical finding that there is a positive relationship between price dispersion and aggregate inflation at rare and high, but not common and low, levels of inflation (Alvarez et al., 2019; Nakamura et al., 2018). The key mechanism is that firms set more dispersed prices in rare, highly inflationary states, because they did not invest many resources into forming precise plans for these unlikely states.

The second application is to consumption and savings in a liquidity trap, in which agents' incomes directly influence cognitive function. This is motivated by the experimental finding that individuals make worse decisions when they are poor (Mani et al., 2013) and the survey finding that individuals report being significantly distracted when near financial constraints (Sergeyev et al., 2022). We derive and interpret conditions under which the unique equilibrium features aggregate output and a consumption distribution that is monotone increasing in aggregate demand, while consumption dispersion decreases in aggregate demand as agents become less cognitively constrained. We show that this economy features a novel externality: when aggregate output is lower, agents' decision costs are higher, and they make larger consumption and savings mistakes. This mechanism provides a new explanation for the finding that consumption dispersion rises in downturns (Berger et al., 2023), in our case because of the equilibrium effect of low income causing stress that worsens decisionmaking.

We discuss two extensions of our analysis in the Appendix. First, in Appendix B, we provide a detailed comparison of our model with the mutual information model of Sims (2003). Using a numerical example of a linear beauty contest (Morris and Shin, 2002), we observe that the mutual information model does *not* imply monotone and discounted stochastic choice rules and therefore opens the door to multiple equilibria defined by coordination on specific support points for the action distribution. This analysis provides a direct counter-example to the possibility that equilibrium analysis similar to ours is possible in workhorse models of unrestricted information acquisition and illustrates the tractability advantage that our model may have for specific applications. Second, in Appendix C, we study strategic mistakes in binary-action coordination games, which are also ubiquitous in macroeconomics and finance (Angeletos and Lian, 2016). We derive sufficient conditions on cognitive costs and payoffs to ensure unique and monotone equilibria and illustrate our results in a canonical investment game with linear payoffs (as in Yang, 2015).

Related literature. The main contribution of our paper is to provide a unified equilibrium analysis of a wide variety of decision frictions—including *ex post* misoptimization, imperfect *ex ante* planning, endogenous cognitive constraints, and endogenous consideration sets—in aggregative games of the kind that are common in macroeconomics and finance (see Angeletos and Lian, 2016, for a review). To our knowledge, comparable results on uniqueness, efficiency, and monotone comparative statics for these games do not exist in the literatures on the two most comparable decision frictions, random utility and costly information acquisition. We detail our connection to these literatures below.

An influential model of equilibrium with non-vanishing “mistakes” induced by random utility is the Quantal Response Equilibrium (QRE) of McKelvey and Palfrey (1995). These authors add type-I extreme value noise to agents’ utility functions to smooth best responses into “better responses” (see the review by Goeree et al., 2016). Subsequent work generalizes this analysis by allowing for different noise distributions that imply different shapes of best-replies (see *e.g.,* Melo, 2022; Fosgerau et al., 2020; Allen and Rehbeck, 2021). Most related to us, Melo (2022) studies games with a finite number of players and actions and general noise distributions and, using convex analysis techniques, shows that QRE are unique if agents’ payoffs are sufficiently concave relative to the extent of strategic complementarity. Our analysis differs from this literature’s in four important ways. First, we consider games with a continuum of agents and actions. This is important because such games are common in macroeconomics and finance and, outside of Melo’s (2022) analysis with discrete actions and players, little remains understood about the uniqueness of QRE in games with a large number of players and/or actions (Goeree et al., 2016). Second, we provide monotone comparative statics results for action distributions, in terms of both first-order stochastic dominance and dispersion. We are not aware of any analogous results in the random-utility literature for the class of games that we study. Third, we can accommodate additional decision frictions which are not well captured by fixed payoff noise—for instance, costly *ex ante* planning and endogenous cognitive constraints. Finally, our analysis has meaningfully different normative properties because we model control costs.

With unrestricted costly information acquisition, we are aware of few equilibrium results that apply to our setting. Hébert and La’O (2022) provide sufficient conditions for equilibrium existence and efficiency in a setting with costly information acquisition under restrictions, relative to our set-up, to consider only mean-critical payoff functions and only the mean aggregator. Hébert and La’O (2022) provide an equilibrium uniqueness result *only* when equilibria are efficient, while our result applies to both efficient and inefficient equilibria under appropriate restrictions on complementarity arising from both payoffs and endogenous cognitive costs. Yang (2015) and

Morris and Yang (2022) study equilibrium existence and uniqueness in *binary-action* settings, to which we extend our analysis in Appendix C. To our knowledge, no references study monotone comparative statics at our level of generality.

Our paper contributes to the theoretical literature on aggregative games (see Jensen, 2018, for a review) by studying these games under general decision frictions. Our analysis also relates to a large literature on uniqueness in games with strategic complementarity (*e.g.,* Morris and Shin, 1998, 2002; Weinstein and Yildiz, 2007; Yang, 2015). Our proof strategy is most closely related to Frankel et al. (2003) and Mathevet (2010), in that we use contraction-mapping techniques, but differs in our use of variational techniques to derive necessary conditions for best responses that imply monotonicity and discounting. Our results on comparative statics are similar in spirit to those of Van Zandt and Vives (2007), but differ in that we study different games, with decision frictions, and provide comparative statics for action distributions.

Finally, our paper contributes to the literature on control costs and stochastic choice by proposing a new class of state-separable cost functionals and applying them in games. This builds upon the analysis of Harsanyi (1973), Stahl (1990), and Van Damme (1991) who introduce specific control cost functionals that penalize the playing of sharply peaked stochastic choice rules, and Mattsson and Weibull (2002), who axiomatize entropic costs. Most relatedly, in decision problems, Fudenberg et al. (2015) axiomatize the class of additive perturbed utility cost (APU) functionals which penalize the expected utility of a mixed action with *any* convex function of the distribution of the mixed action. Our cost function is a weighted sum of APU cost functionals over states, with a weighting function that can depend arbitrarily on both exogenous and exogenous states. Concretely, with weights given by the agents' priors, our cost functional reduces to a state-by-state APU control cost functional that models *ex post* misoptimization. With uniform weights across states, our cost functional captures *ex ante* planning, as control costs must be incurred *ex ante*, while the benefits of plans only realize with probabilities given by the agents' priors. With state-dependent weights, our cost functional allows for exogenous and endogenous state-dependence in the difficulty of choosing precise stochastic choice rules. As we later argue (see Section 2.3), capturing this broad range of behavior enables our class of cost functions to be consistent with the empirical regularities from the psychometrics literature (see Woodford, 2020, for a review), the literature on stress and decision-making (Mani et al., 2013; Sergeyev et al., 2022), and the perceptual tests performed by Dean and Neligh (2022).

Outline. Section 2 introduces the model. Section 3 presents our main results on equilibrium properties. Section 4 discusses applications of our main results. Section 5 briefly discusses two extensions, a detailed comparison of state-separable and mutual information costs and an analysis of binary-action games. Section 6 concludes.

2. Model

2.1. Basic set-up: aggregative games with stochastic choice

A continuum of identical agents is indexed by $i \in [0, 1]$. They take actions $x_i \in \mathcal{X} = [\underline{x}, \overline{x}] \subset \mathbb{R}$. Cross-sectional distributions of actions are aggregated by an aggregator functional $X : \Delta(\mathcal{X}) \to \mathbb{R}$. There is an underlying and payoff-relevant state of the world $\theta \in \Theta \subset \mathbb{R}$. The state space Θ is a finite set, over which the agent has full-support prior $\pi \in \Delta(\Theta)$. Agents have identical utility functions $u : \mathcal{X} \times \mathbb{R} \times \Theta \to \mathbb{R}$, where $u(x, X, \theta)$ is an agent's utility from playing x when the aggregate is X and the state is θ. We assume that u and X are continuous and bounded.[3]

Given a conjecture that the aggregate follows a law of motion $\tilde{X} : \Theta \to \mathbb{R}$, which lies in the space of bounded functions $\mathcal{B} = \{\tilde{X} | \tilde{X} : \Theta \to \mathbb{R}\}$, each agent chooses a stochastic choice rule $P : \Theta \to \Delta(\mathcal{X})$ with $P(x|\theta)$ describing the cumulative distribution of actions x taken by the agent in state θ. When this admits a density function, we denote a stochastic choice rule by $p(x|\theta)$. We call the set of measurable stochastic choice rules $\mathcal{P}$. We model the cost of "controlling mistakes" via a cost functional $c : \mathcal{P} \times \mathcal{B} \to \overline{\mathbb{R}}$. This cost may depend on both the conjectured mapping from states to aggregates (as indicated) and on the prior for the state of nature (suppressed, as this prior is fixed in our analysis). In the next subsection, we specialize these costs to a specific class for our analysis.

The agent maximizes expected utility net of the control cost given their conjecture for how aggregate outcomes depend on the state. This is summarized in the following program:

$$\max_{P \in \mathcal{P}} \sum_{\Theta} \int_{\mathcal{X}} u(x, \tilde{X}(\theta), \theta) \, \mathrm{d}P(x|\theta) \, \pi(\theta) - c(P, \tilde{X}) \tag{1}$$

An equilibrium in this context is a Nash equilibrium: agents' play is optimal given aggregate outcomes, and aggregate outcomes are those that are implied by agents' play.[4]

Definition 1 *(Equilibrium).* An equilibrium is a collection of stochastic choice rules $\{P_i^*\}_{i \in [0,1]}$ and an equilibrium law of motion for aggregates $\hat{X} : \Theta \to \mathbb{R}$ such that:

1. All agents solve Program (1) under the conjecture that $\tilde{X}(\theta) = \hat{X}(\theta)$ for all $\theta \in \Theta$
2. The equilibrium law of motion is consistent with agents' play, or $\hat{X} = X \circ \int_{[0,1]} P_i^* \, \mathrm{d}i$

An equilibrium is *symmetric* if $P_i^* = P^*$ for all $i \in [0, 1]$.

2.2. State-separable cost functionals

We now specialize to a new class of cost functionals that we introduce:

Definition 2 *(State-separable cost functional).* A cost functional c has a state-separable representation if there exists a strictly convex function $\phi : \mathbb{R}_+ \to \mathbb{R}$ and a weighting function $\lambda : \mathbb{R} \times \Theta \to \mathbb{R}_{++}$ such that for any stochastic choice rule P with density p:

$$c(P, \tilde{X}) = \sum_{\Theta} \lambda(\tilde{X}(\theta), \theta) \pi(\theta) \int_{\mathcal{X}} \phi(p(x|\theta)) \, \mathrm{d}x \tag{2}$$

with the convention that the cost is ∞ if P does not have a density.

Formally, state-separable cost functionals are a weighted sum across states of the APU cost functionals of Fudenberg et al. (2015). Informally, state-separable cost functionals capture the

[3] Throughout, our notion of continuity for functionals is the sup norm.

[4] Methodologically, our setup recasts the game with incomplete information in the interim as an *ex ante* game with complete information and a strategy space sufficiently rich to embed all profiles of state-dependent mixed strategies. Morris and Yang (2022) use this approach to study *binary-action* games, as we also do in Appendix C.

idea that it is costly for agents to control "mistakes." The costs of controlling mistakes in different states potentially depend on both the identity of the states and the endogenous outcomes predicted for those states via the weighting function $\lambda(X, \theta)$.

In the remainder of this subsection, we give four specific examples of state-separable costs that capture *ex post* misoptimization, *ex ante* planning frictions, endogenous cognitive constraints, and endogenous consideration sets. In the next subsection, we discuss how these costs are consistent with empirical evidence on decision frictions.

Ex post misoptimization with entropy costs. As a first example, we consider a case in which $\phi(p) = p \log p$ and $\lambda(X, \theta) \equiv \bar{\lambda} > 0$. These costs equal the expectation of the negative entropy of the conditional action distributions. Expected entropy costs encode that precise choice is costly and, therefore, that agents will *ex post* misoptimize. The expected entropy cost model is often applied in macroeconomics to study *ex post* misoptimization (*e.g.,* Costain and Nakov, 2019; Macaulay, 2020; Flynn and Sastry, 2022). Expected entropy costs imply optimal action distributions of the following "logit" form:

$$p(x|\theta) = \frac{\exp\left(\bar{\lambda}^{-1} u(x, \tilde{X}(\theta), \theta)\right)}{\int_{\mathcal{X}} \exp\left(\bar{\lambda}^{-1} u(z, \tilde{X}(\theta), \theta)\right) \mathrm{d}z} \tag{3}$$

When the set of actions is discrete, these choice patterns are identical to those generated in the model of McFadden (1973) in which agents perceive the perturbed utility function $\tilde{u}(x, X, \theta) = u(x, X, \theta) + \varepsilon_x$, where ε_x is distributed type-I extreme value and IID across agents and actions. This model is ubiquitous for modeling consumer demand in industrial organization (see, *e.g.*, Berry and Haile, 2021). The same model for choice is applied in game theory by McKelvey and Palfrey (1995) to define Quantal Response Equilibrium. However, our entropy-cost case differs from what is studied in these references in two key ways. First, actions in our model are continuous. Second, our model's normative analysis is much different. Control costs model the fact that avoiding mistakes has real costs, while random utility treats random choices as *ex post* optimal.

Finally, Matějka and McKay (2015) show that logit choice can be obtained as a limit case of a model of information acquisition with mutual information costs plus a restriction that all actions are *ex ante* exchangeable. We revisit this last connection in Online Appendix B, which studies the difference between our model and the mutual information model.

Prior-dependence and imperfect ex ante planning. As a second example, consider an arbitrary kernel, but now set $\lambda(\theta) = \pi(\theta)^{-1}\bar{\lambda}$. In this case, the agent's cost functional is given by:

$$c(P) = \bar{\lambda} \sum_{\Theta} \int_{\mathcal{X}} \phi(p(x|\theta))\, \mathrm{d}x \tag{4}$$

This captures costly planning, where the agent plans for each state in advance, and then implements these plans when states realize. Thus, costs of planning actions are incurred *ex ante*, and are therefore proportional to the *number of required plans* (*i.e.,* states contemplated) and not their likelihood of occurring. Under the entropy kernel, this process generates the following choice probabilities:

$$p(x|\theta) = \frac{\exp\left(\bar{\lambda}^{-1}\pi(\theta)u(x, \tilde{X}(\theta), \theta)\right)}{\int_{\mathcal{X}} \exp\left(\bar{\lambda}^{-1}\pi(\theta)u(z, \tilde{X}(\theta), \theta)\right) \mathrm{d}z} \tag{5}$$

The agent optimally chooses to form better plans in states that they believe to be more likely. Concretely, the agent trades off the benefits of precise planning in a state against the cost that the state will not be realized and the plan will be useless. This allows us to capture the idea that agents rationally may prepare for very rare events, even if actions during those events are very important (an idea also proposed by Maćkowiak and Wiederholt, 2018). In Section 4.1, we apply this model to study equilibrium price-setting by monopolistically competitive firms in a monetary macroeconomic model.

Endogenous cognitive constraints. We next consider an example that sets $\lambda(X, \theta) = \tilde{\lambda}(X)$, for some decreasing function $\tilde{\lambda}$. Combined with the normalization that u is monotone in X, this embodies the possibility that more favorable aggregate outcomes decrease decision costs while less favorable aggregate outcomes increase decision costs. A leading example studied by Mani et al. (2013) and Mullainathan and Shafir (2013) is that poverty impedes cognitive ability and induces mistakes in decisions. Our framework can model the possibility that this force is endogenous to others' actions and/or mistakes, insofar as income is determined in equilibrium. Under the entropy kernel, choice probabilities follow:

$$p(x|\theta) = \frac{\exp\left(\tilde{\lambda}(\tilde{X}(\theta))^{-1}u(x, \tilde{X}(\theta), \theta)\right)}{\int_{\mathcal{X}} \exp\left(\tilde{\lambda}(\tilde{X}(\theta))^{-1}u(z, \tilde{X}(\theta), \theta)\right) \mathrm{d}z} \tag{6}$$

In states with low weights $\tilde{\lambda}(X)$, when aggregate outcomes are good and stress is low, choices are more precisely concentrated on high-payoff choices; in states with high weights $\tilde{\lambda}(X)$, when aggregate outcomes are bad and stress is high, the opposite is true. Thus, in both cases, the characteristics of aggregate states and their psychological effects shape choice in ways that are not summarized by physical payoffs. In Section 4.2, we study a macroeconomic model in which endogenous stress shapes the determination of aggregate demand and income.

Consideration sets with quadratic costs. We now consider the quadratic kernel $\phi(p) = \bar{\lambda}\frac{p^2}{2}$ studied by Rosenthal (1989). Like the entropy kernel, the quadratic kernel penalizes action distributions that are more sharply peaked and rewards those that are more thinly spread. Unlike the entropy kernel, the quadratic kernel allows for agents to put exactly zero probability on certain actions. In the marketing literature, this phenomenon of agents playing only a strict subset of possible actions is sometimes referred to as a "consideration set" (*e.g.,* Hauser and Wernerfelt, 1990). In the context of rational inattention models, Jung et al. (2019), Caplin et al. (2019), and Fosgerau et al. (2020) study this phenomenon. Stevens (2019) shows evidence of sparse price-setting choices in micro-data and argues that these patterns are consistent with a model of mutual-information costs.

We now illustrate how consideration sets emerge. Choice probabilities follow:

$$p(x|\theta) = \frac{1}{\bar{\lambda}}(u(x, \tilde{X}(\theta), \theta) - \bar{u}(\tilde{X}(\theta), \theta)) \cdot \mathbb{I}[u(x, \tilde{X}(\theta), \theta) \geq \bar{u}(\tilde{X}(\theta), \theta)] \tag{7}$$

where $\mathbb{I}[\cdot]$ is the indicator function and $\bar{u}(\tilde{X}(\theta), \theta)$ is defined such that $\int_{\mathcal{X}} p(x|\theta)\,\mathrm{d}x = 1$. The consideration set of actions in state θ is therefore given by:

$$\mathcal{X}(\theta, \tilde{X}) = \{x \in \mathcal{X} : u(x, \tilde{X}(\theta), \theta) \geq \bar{u}(\tilde{X}(\theta), \theta)\} \tag{8}$$

If $\bar{u}(\tilde{X}(\theta), \theta) > \min_{\mathcal{X}} u(x, \hat{X}(\theta), \theta)$, then a strictly positive (Lebesgue) measure of actions is chosen with zero probability in state θ. In general, without further assumptions, this set can contain many disjoint intervals. However, if u is quasiconcave in x, then $\mathcal{X}(\theta; \hat{X})$ is a closed interval. Finally, observe that these consideration sets are endogenous to equilibrium outcomes as they depend on the equilibrium aggregate.

More generally, away from the quadratic kernel, consideration sets can obtain when ϕ does not satisfy an Inada condition, *i.e.,* when $\lim_{p \to 0} \phi'(p) > -\infty$.

2.3. Experimental evidence and comparisons to the literature

Having illustrated the model's capacity to generate a rich set of decision frictions, we now assess the model's ability to match experimental evidence. We compare and contrast this with the ability of random utility and costly information acquisition models to do the same. We organize this discussion around five key stylized facts that emerge from the classic literature in experimental economics and experimental psychology surveyed by Woodford (2020), the state-of-the-art perceptual study by Dean and Neligh (2022), and the cognitive experiments of Mani et al. (2013). We show that specializations of the state-separable model can capture combinations of these facts.

Fact 1: choice is random. People make inaccurate and random judgments in decision problems. These imperfect random choices are often measured in experiments that ask participants to pick which of two stimuli is larger (*i.e.,* which noise is louder) and summarized as *psychometric functions* that plot the probability of choosing the correct option against objective differences in the stimuli that are varied across experiments. These typically reveal a smooth, monotone relationship that is interior to (0, 1) (see *e.g.,* Figure 1 of Woodford, 2020, and each of the experiments in Dean and Neligh, 2022). The state-separable, random utility, and information acquisition models all rationalize random choice. The state-separable model does so by making precise optimization costly.

Fact 2: choice responds to incentives. People make more accurate and precise choices when the payoffs from doing so are higher. In perceptual tasks, error rates decrease in rewards (see, *e.g.,* Figure 2b of Woodford, 2020, and Experiment 2 in Dean and Neligh, 2022). The state-separable and rational inattention model this as a rational response to higher returns to cognitive effort; the random utility model generates a similar prediction because larger payoff differences drown out fixed payoff noise.

Fact 3: choice depends on prior beliefs. People's random choice responds to the probabilities of states, in repeated experiments where it may be reasonable to interpret these as prior beliefs. In repeated perceptual tasks, average error rates are lower in states that recur more often (see, *e.g.*, Figure 2a of Woodford, 2020, and Experiment 3 of Dean and Neligh, 2022). This is consistent with state-separable costs that capture *ex ante* planning, as agents have incentives to exert more effort to prepare for more likely states. This result is also natural in many models of costly information acquisition. However, this result cannot be understood through the lens of random utility models (or, in games, QRE), as they embody no notion of *ex ante* planning and agents' priors are irrelevant.

Fact 4: choice depends on decision context. The accuracy and precision of choice vary with the "context" of decision problems, such as the action space and the state space.

First, Dean and Neligh (2022) show the importance of the action space. In Experiment 1, the authors first ask participants to pick between two options. The authors then introduce a third choice (*i.e.,* expand the action space). They find that this *increases* the probability of one of the initial actions. This is consistent with state-separable costs exactly when different action spaces affect the difficulty of making choices (modeled through a change in the value of the weighting function). As observed by Dean and Neligh (2022), this is also consistent with models of costly information acquisition, but *inconsistent* with models of random utility, which predict that larger action spaces decrease the probabilities that all actions are played.

Second, three examples demonstrate the impact of changes in the *state space*. Experiment 4 in Dean and Neligh (2022) shows that choice probabilities are more inaccurate when participants are asked to distinguish states that look more similar. Woodford (2020) surveys two related results in the psychometric literature. First, when laboratory participants are asked to reproduce a set of unknown distances from memory, they overestimate the shorter distances and underestimate the longer distances on average (Figure 4 of Woodford, 2020). Second, the extent of bias can depend systematically on the scale of stimuli (Figure 5 of Woodford, 2020). All of these results are consistent with the state-separable model where the weighting function depends on the state space, capturing the idea that some problems are easier to solve than others. These results are also consistent with information acquisition models that emphasize that the topology of the state space matters (e.g., Hébert and Woodford, 2020). However, they are inconsistent with information acquisition models that satisfy the Invariance Under Compression Axiom (Caplin et al., 2022), such as the canonical mutual information cost proposed by Sims (2003).

Fact 5: choice depends on decision-irrelevant context. The accuracy and precision of decisions can also depend on context that is *not* decision-relevant. For example, Mani et al. (2013) show that performance on abstract cognitive tasks declines when individuals are reminded of the difficulty of making financial decisions under poverty or, for predictable reasons, have higher or lower income from a seasonal cycle. In each case, except for the interaction with the (small) financial incentives, income could be viewed as irrelevant for the decision problem solved.

As mentioned earlier, our state-separable model can embody this property directly via appropriate specification of how the weighting function depends on endogenous states in a game (see Equation (6)). This directly embodies the idea expressed in the title of the Mani et al. (2013) that "Poverty Impedes Cognitive Function," no matter what decision problem agents are solving (*i.e.,* what are their payoffs, action space, or state space).

A model of costly information acquisition has the flexibility to explain this sort of finding, mathematically speaking. But this has an important caveat. The ability of this model to generate more "mistakes" in a poverty state relies on the premise of imperfect *observation* of income, or a heightened inability to determine income when it is low overall.[5] But this would be hard to square with the findings of Mani et al. (2013) in tasks for which income is (almost) decision-irrelevant. More broadly, the notion that imprecise choice *must* arise through imperfect learning places significant restrictions on how decision frictions vary across contexts.

[5] Concretely, one could apply a variant of the Hébert and Woodford (2020) neighborhood-based cost or the Pomatto et al. (2023) log-likelihood-ratio cost in which states corresponding to poverty are harder to distinguish from others.

Summary. State-separable costs are consistent with Facts 1 to 5. Random utility models can explain only Facts 1 and 2. All models of costly information acquisition can explain Facts 1, 2, and 3; some models can be consistent with Fact 4 (but not mutual information); but none could easily explain Fact 5. On the basis of this, we argue that state-separable costs provide a flexible way of modeling a variety of decision frictions in a way that is consistent with our best experimental evidence. An example "workhorse" model that could capture all five facts is a model with the quadratic kernel $\phi(p) = \frac{p^2}{2}$ with weighting function $\lambda(\theta) = \pi(\theta)^{-1}\tilde{\lambda}(X)$, for some decreasing function of X.

However, there are potentially testable implications of information acquisition models with which state-separable costs would not be consistent. In particular, information acquisition models make predictions about the joint properties of beliefs and actions. This notwithstanding, it has been customary in the decision-theoretic literature to ignore these predictions, and instead to focus entirely on predictions for choice, under the premise that internal mental states are unobservable (*e.g.,* Caplin and Dean, 2015; Caplin et al., 2022). Moreover, existing tests of information acquisition models derived from the analysis of Caplin and Dean (2015) and Caplin and Martin (2015) and performed by Dean and Neligh (2022) are one-sided: they reveal that information acquisition is consistent with the data, but not that non-informational models are inconsistent with the data.

3. Main results

We now prove existence, uniqueness, efficiency, and equilibrium monotone comparative statics for both the aggregate and the cross-sectional action distribution. Our approach will be to establish that the correct notion of a "best response function" for the aggregate action X is a contraction map that satisfies certain properties.

3.1. Assumptions: payoffs and aggregator

We first identify conditions on payoffs, aggregators, and stochastic choice functionals sufficient to guarantee uniqueness. For payoffs, we first require complementarities in the underlying game in the form of supermodularity in *cost-normalized payoffs* between an agent's own action and the aggregate. Second, we require that these complementarities are not too strong in the sense that payoffs are sufficiently concave to outweigh them:

Assumption 1 *(Supermodularity and sufficient concavity).* The payoff function u and weighting function λ are such that the following holds for all $x' \geq x$, $X' \geq X$, and θ:

$$\frac{u(x', X', \theta) - u(x, X', \theta)}{\lambda(X', \theta)} \geq \frac{u(x', X, \theta) - u(x, X, \theta)}{\lambda(X, \theta)} \tag{9}$$

Moreover, for all $\alpha \in \mathbb{R}_+$, $x' \geq x$, X, and θ, the following holds[6]:

$$\frac{u(x' - \alpha, X, \theta) - u(x - \alpha, X, \theta)}{\lambda(X, \theta)} \geq \frac{u(x', X + \alpha, \theta) - u(x, X + \alpha, \theta)}{\lambda(X + \alpha, \theta)} \tag{10}$$

6 In stating this assumption, we are implicitly extending the domain of u so that it is well-defined under such translations.

Informally, the former part of the assumption ensures that when aggregate actions go up, agents have an incentive to increase their own action. The latter part of the assumption ensures that agents' actions are less than one-for-one sensitive to the aggregate.

To gain a stronger intuition for the role of this assumption, and to provide easily verifiable conditions under which it holds, we characterize it with twice continuously differentiable payoffs u and weighting functions λ:

Lemma 1. *When $u(\cdot,\theta)$ is twice continuously differentiable in (x,X) and $\lambda(\cdot,\theta)$ is twice continuously differentiable in X for all θ, Assumption 1 is equivalent to the following:*

$$0 \le u_{xX}(x,X,\theta) - u_x(x,X,\theta)\frac{\lambda_X(X,\theta)}{\lambda(X,\theta)} \le -u_{xx}(x,X,\theta) \tag{11}$$

for all x, X and θ.

Proof. See Appendix A.1. □

When cognitive constraints are exogenous, this condition reduces to the requirement that $0 \le u_{xX} \le -u_{xx}$, which is a standard condition for unique equilibrium in supermodular games (see *e.g.,* Weinstein and Yildiz, 2007). Intuitively, this condition requires that the slope of agents' optimal actions to changes in aggregate actions are bounded between zero and one.

When cognitive costs are endogenous, strategic complementarity now has both a physical payoff complementarities component u_{xX} and a *cognitive complementarities* component $-u_x\frac{\lambda_X}{\lambda}$. To understand why cognitive complementarities take this form, suppose that aggregate actions increase and this raises cognitive costs by $\frac{\lambda_X}{\lambda}$ percent. This gives the agent an incentive to spread out their actions around any locally optimal action. If the agent is playing an action greater than any locally optimal action, their marginal utility from increasing their own action is negative ($u_x < 0$). However, as cognitive costs have gone up, the agent is now more willing to accept such a negative marginal payoff, and so has incentives to further increase the likelihood that their action lies further from the locally optimal point. Thus, when u_x is negative, when cognitive costs increase, the agent has an incentive to play higher actions and there is strategic complementarity, *i.e.,* $-u_x\frac{\lambda_X}{\lambda} > 0$. When u_x is positive, the reverse logic is true, and increased cognitive costs make actions strategic substitutes. Thus, with cognitive strategic externalities, we require that (i) any strategic substitutability through cognitive costs never outweighs strategic complementarities in physical payoffs, and (ii) agents' payoff functions are sufficiently concave to outweigh both physical payoff complementarities and cognitive complementarities.

Having identified conditions on payoffs, we now turn to the aggregator. To retain the ordering between actions and aggregates, we assume that the aggregator is monotone in the sense of first-order stochastic dominance. We further assume that the aggregator satisfies discounting, which is to say that it is sub-linear in level shifts of the cross-sectional action distribution (see Cerreia-Vioglio et al., 2020, for a discussion of monotone and (sub-)linear aggregators):

Assumption 2 *(Monotone and discounted aggregator).* For all $g, g' \in \Delta(\mathcal{X})$:

$$g' \succeq_{FOSD} g \implies X(g') \ge X(g) \tag{12}$$

Moreover, there exists $\beta \in (0,1)$ such that for any distribution $g \in \Delta(\mathcal{X})$ and any $\alpha \in \mathbb{R}_+$:

$$X(\{g(x-\alpha)\}_{x\in\mathcal{X}}) \le X(\{g(x)\}_{x\in\mathcal{X}}) + \beta\alpha \tag{13}$$

We moreover, note that the assumption that $\beta < 1$ can be relaxed to allow $\beta = 1$ if the second inequality in Assumption 1 (Equation (10)) is made strict. In the interests of concreteness, the following Lemma (the proof of which is immediate, and therefore omitted) provides several important and natural aggregator functions that satisfy Assumption 2.

Lemma 2. *The following aggregators satisfy Assumption 2:*

1. *Linear aggregators:*

$$X(g) = \beta \int_{\mathcal{X}} f(x) g(x) \, \mathrm{d}x \tag{14}$$

   where $\beta \in [0, 1)$ *is a parameter controlling discounting and* $f : X \to \mathbb{R}$ *is a differentiable function such that* $f' \in [0, 1]$.
2. *Quantile aggregators:*

$$X(g) = \beta G^{-1}(l) \tag{15}$$

   where $\beta \in [0, 1)$ *is a parameter controlling discounting,* $G(x) = \int_{\underline{x}}^{x} g(\tilde{x}) \, \mathrm{d}\tilde{x}$ *is the CDF of the cross-sectional action distribution,* G^{-1} *is its left-inverse, and* $l \in (0, 1)$.

Linear aggregators with polynomial kernels $f(x) = a_0 + a_1 x + \ldots + a_l x^l$ (subject to the monotonicity and discounting constraints that $f'(x) \in [0, 1]$ on $[\underline{x}, \overline{x}]$) allow the aggregator to depend on *all* moments of the cross-sectional distribution of actions. The mean aggregator, $X(g) = \beta \int_{\mathcal{X}} x g(x) \, \mathrm{d}x$, is a special case of this class when $l = 1$. Thus, our analysis nests the common assumption in macroeconomics that interactions take place through the mean action (see Angeletos and Lian, 2016, for a review). Moreover, the polynomial sub-class allows for higher moments of the action distribution to enter agents' payoffs. This allows the dispersion $l = 2$, skewness $l = 3$, and kurtosis $l = 4$ of other agents' actions to matter for agents' strategic incentives. Such aggregators also have natural macroeconomic applications. For example, in Flynn and Sastry (2022), the fact that dispersion reduces aggregate outcomes generates important general equilibrium forces. Quantile aggregators include the median when $l = \frac{1}{2}$. Such aggregators are relevant when agents care about what an average agent does, rather than what other agents do on average.

Assumption 2 rules out aggregators that do not preserve the monotonicity of actions, *e.g.,* linear aggregators with a negative slope, or those that are more than one-for-one sensitive to translations of actions, *e.g.,* linear aggregators with a slope greater than one. Intuitively, such aggregators break either strategic complementarity or sufficient concavity.

3.2. Intermediate result: properties of stochastic choice

Assumption 2 suggests a path toward ensuring that equilibrium is described by a contraction map if, in response to level shifts in the aggregate, the optimal stochastic choice pattern increases in the sense of first-order stochastic dominance (monotonicity) but remains dominated by the level shift itself (discounting). These are intuitive properties given the supermodularity and concavity of payoffs, which encode that level shifts in the (conjectured) aggregate globally increase the attractiveness of playing higher x, but in a way that is less than one-for-one. We now show an interpretable sufficient condition within the state-separable class which guarantees that monotonicity and discounting translate appropriately to stochastic choice.

We first define a new property of a function that we label the quasi-monotone-likelihood-ratio-property (quasi-MLRP). This condition allows us to relate the underlying cost functional to the distribution of actions induced by optimality.

Definition 3 *(Quasi-MLRP).* A function $f : \mathbb{R}_+ \to \mathbb{R}$ satisfies quasi-MLRP if for any two distributions $g', g \in \Delta(\mathcal{X})$:

$$\left(f(g'(x')) - f(g'(x)) \geq f(g(x')) - f(g(x)) \; \forall x' \geq x \right) \implies g' \succeq_{FOSD} g \tag{16}$$

With this definition in hand, we can now state our final technical assumption on stochastic choice functionals, which ensures that we can always translate dominance in payoff units to dominance in terms of distributions:

Assumption 3 *(Quasi-MLRP kernel).* Costs have a differentiable kernel ϕ such that ϕ' satisfies quasi-MLRP.

It is important to note that the two workhouse kernels in the literature on control costs satisfy this assumption:

Lemma 3. *The entropy kernel $\phi(p) = p \log p$ and the quadratic kernel $\phi(p) = \frac{1}{2}p^2$ satisfy Assumption 3.*

Proof. See Appendix A.2. □

We can now state a Proposition using this assumption and our earlier assumptions on payoffs to establish monotonicity and discounting of the solution of the stochastic choice problem:

Proposition 1 *(Monotone and discounted stochastic choice). Consider the stochastic choice program with payoffs satisfying Assumption 1 and cost functional satisfying Assumption 3. Then,*

1. *The optimal stochastic choice rule p^* is weakly increasing in the sense that if $\hat{X}' \geq \hat{X}$ then $p^*(\theta; \hat{X}') \succeq_{FOSD} p^*(\theta; \hat{X})$ for all $\theta \in \Theta$.*
2. *The optimal choice profile is discounted in the sense that when $\hat{X}$ and $\hat{X}' = \hat{X} + \alpha$ for $\alpha \in \mathbb{R}_+$, we have that $p^*_{-\alpha}(\theta; \hat{X}) \succeq_{FOSD} p^*(\theta; \hat{X}')$ for all $\theta \in \Theta$, where $p^*_{-\alpha}$ denotes the translation of p^* to the right by α.*

Proof. See Appendix A.3. □

The key to both parts is that quasi-MLRP allows us to "invert" dominance relationships in payoffs to obtain dominance relationships between distributions. For the first part, we show that the dominance of payoffs for playing higher x from supermodularity implies dominance of distributions under quasi-MLRP. For the second part, we use the property off payoffs from (10) that concavity of utility exceeds strategic complementarity, to show the optimal stochastic choice rule is dominated by the claimed level shift in the rule.

Proposition 1 is the core of our environment's tractability. It is in principle the ingredient that might be replaced in an alternative model of stochastic choice, like a form of unrestricted information acquisition. But, to our knowledge, such monotonicity and discounting results do not exist for any form of information acquisition in general environments. Moreover, this is not merely a technical glitch. A very relevant mechanism, anchoring toward frequently played actions, fights such monotonicity and discounting in information acquisition models. In a numerical example with the mutual-information cost (Sims, 2003) in Appendix B, we show that violations of monotonicity and discounting obtain in the single-agent problem and how this leads to non-uniqueness and non-monotone comparative statics in the equilibrium of an example game.

3.3. Existence and uniqueness

We can now state our main existence and uniqueness result:

Theorem 1 *(Existence, uniqueness, and symmetry). Under Assumptions 1, 2 and 3, there exists a unique equilibrium. The unique equilibrium is symmetric.*

Proof. See Appendix A.4. □

As alluded to above, we show this result by defining an equilibrium operator that maps the law of motion of the aggregate in the state to the resulting optimal stochastic choice rule and then maps this back to a law of motion of the aggregate, and then determining that said operator is a contraction map. More formally, let $\mathcal{B} = \{\hat{X} | \hat{X} : \Theta \to \mathbb{R}\}$ be the space of (bounded) functions endowed with the sup norm. We define the operator $T : \mathcal{B} \to \mathcal{B}$:

$$T\hat{X} = X \circ p^*(\hat{X}) \tag{17}$$

To show uniqueness of the equilibrium law of motion of aggregates, it then suffices to prove that T is a contraction map. We prove this by showing that, under the given assumptions, T satisfies both of Blackwell's conditions of monotonicity and discounting. Given the unique equilibrium-consistent law of motion which satisfies $T\hat{X} = \hat{X}$, the equilibrium stochastic choice rule is then the unique solution of the stochastic choice problem given that law of motion, or $p^*(\hat{X})$. This extends classic uniqueness results to the realm of stochastic choice.[7]

3.4. Monotone comparative statics

Once we lie in the realm of unique equilibria, it is well-posed to consider comparative statics in equilibrium. We provide two such results, showing when the action distribution and aggregate action are monotone in the state and when the precision of agents' actions is monotone in the state.

3.4.1. Monotonicity of action distributions

To show monotonicity of distributions and aggregates, we require a stronger supermodularity assumption that not only are individual actions and aggregate actions complements, but so too

[7] One could dispense with Assumptions 1, 2, and 3 and prove existence in our setting only by applying the Schauder fixed-point theorem. We omit this result as it is simple, and because our analysis will proceed afterward under Assumptions 1, 2, and 3.

is the underlying state itself a complement to both individual actions and aggregates in cost-adjusted payoffs:

Assumption 4. The payoff function u and weighting function λ are such that the following holds for all $\theta' \geq \theta$, $X' \geq X$, $x' \geq x$:

$$\frac{u(x', X', \theta') - u(x, X', \theta')}{\lambda(X', \theta')} \geq \frac{u(x', X, \theta) - u(x, X, \theta)}{\lambda(X, \theta)} \tag{18}$$

As before, to gain a stronger intuition and provide an easily verifiable condition, we characterize this assumption when u and λ are twice continuously differentiable:

Lemma 4. *When u is twice continuously differentiable in (x, X, θ) and λ is twice continuously differentiable in (X, θ), Assumption 4 is equivalent to*

$$u_{xX}(x, X, \theta) - u_x(x, X, \theta)\frac{\lambda_X(X, \theta)}{\lambda(X, \theta)} \geq 0 \quad \text{and} \quad u_{x\theta}(x, X, \theta) - u_x(x, X, \theta)\frac{\lambda_\theta(X, \theta)}{\lambda(X, \theta)} \geq 0 \tag{19}$$

for all x, X and θ.

The proof follows from the same steps as in the proof of Lemma 1, simply relabeling X as θ, and is therefore omitted. The first inequality ("complementarity with X") is identical to that in Lemma 1, and the second is its mirror image for "complementarity with θ". When cognitive costs do not depend on exogenous states $\lambda_\theta = 0$, this second condition reduces to $u_{x\theta} \geq 0$. When cognitive costs depend on exogenous states, the intuition for the additional term echoes the discussion of complementarity with X. The presence of this additional term underscores the fact that state-varying control costs affect agents' incentives to shift their entire distribution of stochastic choice upward in higher states.

Under this assumption, we show the following result:

Theorem 2 *(Monotone actions and aggregates). Under Assumptions 1, 2, 3, and 4, the unique equilibrium action distribution is monotone increasing in the sense of FOSD and the law of motion of the aggregate is increasing in the underlying state.*

Proof. See Appendix A.5. □

The intuition for this result is simple: higher θ makes higher actions more desirable, so the distribution of actions in higher states dominates the distribution in lower states. This is complicated by the fact that agents may face higher cognitive costs in higher states. Hence, the relevant notion of complementarity is complementarity in cost-adjusted payoffs. The proof strategy makes use of the contraction mapping property used in the uniqueness proof. In particular, it shows that monotonicity is preserved by the fixed point operator and therefore that the fixed point must itself be monotone.

This result has the following immediate implication for the supports of action distributions:

Corollary 1 *(Monotone consideration sets). Under the conditions of Theorem 2, in the unique equilibrium agents' consideration sets* $\mathcal{X}(\theta) = cl_{\mathcal{X}}\{x \in \mathcal{X} : p^*(x|\theta, \hat{X}(\theta)) > 0\}$ *are increasing in the strong set order.*[8]

As optimal distributions increase in the sense of first-order stochastic dominance, the supports must move in the sense of the strong set order. The result is vacuous if the cost kernel satisfies an Inada condition and $\mathcal{X}(\theta) = \mathcal{X}$ for all θ. The result has bite if agents, for example, have costs with the quadratic kernel, which does not satisfy an Inada condition and may result in agents' optimally playing only a subset of available actions. In this case, the result puts structure on the endogeneity of consideration sets—agents consider larger actions in higher states in equilibrium.

3.4.2. *Monotonicity of action precision*

We now turn to establish when the precision of, or extent of mistakes in, agents' actions is monotone in the state in equilibrium. To this end, in our context with flexible stochastic choice, we first need a non-parametric notion of precision:

Definition 4 *(Precision).* Fix an $h : \mathbb{R} \to \mathbb{R}$. A symmetric distribution g is more precise about a point x^* than g' about $x^{*'}$ under h if $h \circ g(|x - x^*|)$ is faster decreasing in $|x - x^*|$ than is $h \circ g'(|x' - x^{*'}|)$ in $|x' - x^{*'}|$.[9]

Informally, this definition requires that a distribution is more precise than another if its density is more rapidly decreasing away from the point about which precision is being considered. This definition generalizes the property that Gaussian distributions are more precise about their mean when they have a lower standard deviation to cases with non-Gaussian densities by exactly capturing the idea that a distribution is more precise if its tails decay faster from the point about which a distribution is centered.[10]

Having defined precision, we now state sufficient assumptions on payoffs for precision to be monotone. To show this result, we specialize to a distance-based payoff environment, which we refer to throughout as *generalized beauty contest* payoffs:

Assumption 5 *(Generalized beauty contests).* The utility function is given by:

$$u(x, X, \theta) = \alpha(X, \theta) - \beta(X, \theta)\Gamma(|x - \gamma(X, \theta)|) \tag{21}$$

where Γ is monotone increasing and such that $\Gamma(0) = 0$, $\gamma(X, \theta)$ is monotonically increasing in (X, θ), and $\beta(X, \theta)$ is positive, for every (X, θ).

8 Where $\mathrm{cl}_{\mathcal{X}} A$ denotes the closure of set A within $\mathcal{X}$.

9 On an asymmetric support, we call a distribution g symmetric if $g(x) = g(-x)$ whenever both $g(x)$ and $g(-x)$ are defined. For any symmetric functions $\xi, \hat{\xi} : A \to \mathbb{R}$, we say that ξ is faster decreasing than $\hat{\xi}$ in their arguments if $\xi(0) - \xi(|x|) \geq \hat{\xi}(0) - \hat{\xi}(|x|)$ for all $x \in A$.

10 To see this, recall that a Gaussian random variable with mean μ and standard deviation σ has pdf:

$$g(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2}\left(\frac{x - \mu}{\sigma}\right)^2\right\} \tag{20}$$

Thus, for two Gaussian distributions with means μ, μ' and standard deviations σ, σ' such that $\sigma < \sigma'$, we have that $h \circ g(|x - \mu|)$ is faster decreasing than $h \circ g'(|x - \mu'|)$ whenever h is monotone. Thus, under monotone h, we have that Gaussian distributions with lower standard deviations are more precise about their mean under our definition of precision.

Under distance-based payoffs with distance function Γ, an agent cares only about how far their action is from an optimal action conditional on others' play X and the state θ, $\gamma(X,\theta)$. The extent to which they care is governed by $\beta(X,\theta)$, with larger values inducing greater losses from failing to match the optimal action.

We note that this formulation nests the quadratic payoff functions, which can be justified via a second-order approximation of any smooth, concave utility function around its maximum value $\gamma(X,\theta)$:

Lemma 5. *Consider a payoff function $u : \mathcal{X} \times \mathbb{R} \times \Theta$ that is twice differentiable, strictly concave in its first argument, and maximized for every $(X,\theta) \in \mathbb{R} \times \Theta$ at some $x^*(X,\theta) \in int(\mathcal{X})$. Then, up to a term that is on the order of $|x - x^*(X,\theta)|^3$, payoffs conditional on each (X,θ) take the form of Equation (21) with $\alpha(X,\theta) = u(x^*(X,\theta), X, \theta)$, $\beta(X,\theta) = \frac{1}{2}|u_{xx}(x^*(X,\theta), X, \theta)|$, $\gamma(X,\theta) = x^*(X,\theta)$, and $\Gamma(x) = x^2$.*

This result follows immediately from taking a Taylor expansion of u around its optimal value in each state, observing that the first-order term is zero because of the first-order condition for optimality, and using Taylor's Theorem to describe the residual error. In this interpretation, $\gamma(X,\theta)$ is the optimal action conditional on (X,θ) and $\beta(X,\theta)$ measures the curvature of payoffs, or second-order loss of mis-optimization, around that point.

We now state the result, which encapsulates the idea that precision is higher when the losses from mis-optimization are higher for endogenous or exogenous reasons:

Theorem 3 *(Monotone precision). Under Assumptions 1, 2, 3, 4, and 5, $p^*(\theta) \in \Delta(\mathcal{X})$ is more precise about $\gamma(\hat{X}(\theta),\theta)$ than $p^*(\theta')$ about $\gamma(\hat{X}(\theta'),\theta')$ under ϕ'*

1. *For any $\theta \le \theta'$ if $\frac{\beta(X,\theta)}{\lambda(X,\theta)}$ is monotone decreasing in both arguments.*
2. *For any $\theta \ge \theta'$ if $\frac{\beta(X,\theta)}{\lambda(X,\theta)}$ is monotone increasing in both arguments.*

Proof. See Appendix A.6. □

The proof of this result shows that the agents' incentives to transfer probability mass from the ideal point $\gamma(\hat{X}(\theta),\theta)$ to any other $x \in \mathcal{X}$ are strictly lower when $\frac{\beta(X,\theta)}{\lambda(X,\theta)}$ is larger, which translates directly to our notion of precision. Note that this combines the incentives for precision from the curvature in the utility function, β, and from the scaling of the cost function, λ. This calculation relies on the symmetry of distance-based payoffs around $\gamma(\hat{X}(\theta),\theta)$. It then leverages the fact that $\hat{X}$ is monotone in θ in equilibrium, because of Theorem 2, which in turn implies monotonicity of the mapping $\theta \mapsto \frac{\beta(\hat{X}(\theta),\theta)}{\lambda(\hat{X}(\theta),\theta)}$, decreasing in case 1 and increasing in case 2. Put differently, the "endogenous" and "exogenous" stakes of making good choices both move in the same direction in equilibrium. Thus, precision is monotone in the state.[11]

This result has the following immediate implication for the size of agents' equilibrium consideration sets:

[11] Unsurprisingly, we cannot state a general result when $\frac{\beta(X,\theta)}{\lambda(X,\theta)}$ is not strictly monotone in its two arguments; but we could of course still use part of the previous argument to compare precision in any two states $(\theta, \hat{X}(\theta))$, $(\theta', \hat{X}(\theta'))$ after solving for equilibrium.

Corollary 2 *(Monotone size of consideration sets). Under the conditions of Theorem 3, if 1. (resp. 2.) holds, then the Lebesgue measure of $\mathcal{X}(\theta)$ is increasing (resp. decreasing) in θ.*

Thus, as is intuitive, in states where agents' cost-adjusted states are higher, agents choose from smaller consideration sets.

3.5. Efficiency

A further question of interest is when equilibria of our model are efficient. As our agents are symmetric, *ex-ante* Pareto efficiency and utilitarian efficiency are equivalent. We therefore say that a stochastic choice rule is efficient if it maximizes utilitarian welfare:

Definition 5. A stochastic choice rule $P^E \in \mathcal{P}$ is efficient if it solves the following program:

$$P^E \in \arg\max_{P \in \mathcal{P}} \sum_{\Theta} \int_{\mathcal{X}} u\left(x, X(P(\theta)), \theta\right) \, \mathrm{d}P(x|\theta)\, \pi(\theta) - c(P, X(P)) \tag{22}$$

An efficient stochastic choice rule both fully internalizes the effect choices have on aggregates and the costs of stochastic choice. Moreover, this notion of efficiency takes seriously that agents do incur the cognitive cost of constraining their mistakes. We now ask, when will equilibrium be efficient? The following result relates the answer to this question to the *balancing of aggregate externalities* in physical and payoffs. To derive a variational necessary condition, we make technical assumptions sufficient to guarantee differentiability:

Assumption 6 *(Regularity conditions for efficient program).* Suppose that the planner's objective in Equation (22) is strictly concave in P, u is differentiable in its second argument X, λ is differentiable in its first argument X, and the aggregator is linear:

$$X(g) = \int_{\mathcal{X}} f(x) \, \mathrm{d}G(x) \tag{23}$$

for some nowhere-constant function f.

Theorem 4. *Under Assumption 6, a necessary condition for efficiency of an equilibrium stochastic choice rule p^* is that:*

$$\int_{\mathcal{X}} u_X(\tilde{x}, X(p^*(\theta)), \theta)\, p^*(\tilde{x}|\theta) \, \mathrm{d}\tilde{x} = \lambda_X(X(p^*(\theta)), \theta) \int_{\mathcal{X}} \phi(p^*(\tilde{x}|\theta)) \, \mathrm{d}\tilde{x} \tag{24}$$

for all $\theta \in \Theta$.

Proof. See Appendix A.7. □

To understand this result, we first consider the case in which there are no payoff externalities in costs or $\lambda_X = 0$. In this case, the condition requires that the *average externality of increasing the aggregate is zero*. This condition is *evaluated* at the equilibrium stochastic choice pattern, but does not depend directly on the structure of cognitive costs. Thus, to evaluate such a condition (under the assumption that $\lambda_X = 0$), an observer needs only to know about payoff externalities and the observed distribution of choices.

We next consider the case in which $\lambda_X \neq 0$. In this case, efficiency obtains when the aforementioned payoff externality balances with a *cognitive externality*, to use the language of Angeletos and Sastry (2023), operating directly through costs. Consider our recurring example of cognitive costs that decrease with the value of X because of poverty-induced stress ($\lambda_X < 0$) and assume that the utility costs of cognition are positive in all states.[12] The *cognitive externality* is that increasing X directly decreases every agent's cognitive cost. A non-paternalistic planner, who takes cognitive costs into account, considers also this externality. Thus, an optimal allocation (if it exists) tolerates a *negative* marginal payoff externality to achieve a *positive* marginal cognitive externality. We return to this specific point in a concrete example in Section 4.2 and Corollary 4.

Relative to the literature, our analysis therefore identifies a new channel through which rational decision frictions can create equilibrium externalities and induce inefficiency. This supplements the findings of Hébert and La'O (2022) for aggregative games with information acquisition and Angeletos and Sastry (2023) for Arrow-Debreu economies with information acquisition. Relative to the related results in those papers, our Theorem 4 has three substantial differences. First, it clarifies how cognitive externalities can operate outside of information acquisition models. Second, it sheds light on the *nature* of inefficiency—in particular, the direction in which a social planner would want to perturb aggregates—in inefficient equilibrium. By contrast, due to the intractable structure of general cognitive externalities in information-acquisition models, Hébert and La'O (2022) and Angeletos and Sastry (2023) can say relatively little about the same in their settings. Third, leveraging our state-separable structure, it provides a testable condition to compare the extent of these externalities with "standard" payoff externalities to gauge efficiency.

4. Applications

We now apply our model to make equilibrium predictions in two macroeconomic settings. We first study price-setting by monopolistic firms, a cornerstone of the "supply block" of modern macroeconomic models. In our model, firms imperfectly price their goods because of *ex ante* planning frictions. We show how to make equilibrium predictions for the aggregate price level and price dispersion that take into account the aggregate consequences of "pricing mistakes" and firms' differential incentives to rein in these mistakes in different aggregate states. We next study consumption and savings decisions in a liquidity trap, a cornerstone of the "demand block" of modern macroeconomic models. In our model, consumption plans are imperfect because of costly control. Moreover, these costs increase when households have low income, capturing the possibility that psychological stress impairs decisionmaking in these states. We show how to make predictions for aggregate income and consumption inequality and characterize a novel equilibrium externality that arises because one agent's lack of consumption increases others' costly stress.

[12] Note, of course, that in our model these costs need not be positive. The following more perverse model would also be consistent with empirical evidence that scarcity reduces *decision quality* (*e.g.,* Mani et al., 2013): low X increases the *scale* of cognitive costs, reducing *relative* incentives for precise optimization, but has a positive *level* effect on welfare. In this case, our intuition for why there is a role of cognitive externalities would be the same, and Theorem 4 would still hold; but the intuition for the sign of effects would flip.

4.1. Price-setting with planning frictions

Set-up. Each agent $i \in [0, 1]$ is a firm that produces a differentiated variety in quantity q_i at price $p_i \in [\underline{p}, \overline{p}]$ with $\underline{p} > 0$. These firms use intermediate goods z_i, with marginal cost k, to produce according to the production technology $q_i = z_i$. The outputs of these firms are consumed by a representative household, with constant elasticity of substitution (CES) consumption bundle:

$$C = \left(\int_0^1 q_i^{\frac{\eta-1}{\eta}} \,\mathrm{d}i \right)^{\frac{\eta}{\eta-1}} \tag{25}$$

where $\eta > 1$. As is standard (see *e.g.,* Hellwig and Venkateswaran, 2009), the household's preferences over consumption and real money balances $\frac{M}{P}$ are given by:

$$V\left(C, \frac{M}{P}\right) = \frac{C^{1-\sigma}}{1-\sigma} + \ln \frac{M}{P} \tag{26}$$

where $\sigma \geq 0$. The money supply is an exogenous shock in the discrete set $\mathcal{M}$ with minimal and maximal elements $\underline{M}$ and $\overline{M}$, such that $\underline{M} > \underline{p}$ and $\overline{M} < \overline{p}$. Moreover, we make the standard simplifying assumption (Alves et al., 2020; Flynn and Sastry, 2022) that real marginal costs are a log-linear function of aggregate output:

$$\frac{k}{P} = C^{\chi} \tag{27}$$

where $\chi > 0$ represents "factor price pressure", *i.e.,* the extent to which real marginal costs are increasing in the level of output in the economy.

To study how planning frictions matter, we subject the firm to a state-separable cost function with any kernel satisfying Assumption 3 (*e.g.,* the entropy kernel $\phi(p) = p \log p$) and a weighting function inversely proportional to how likely the firm thinks each realization of the money supply $\pi(M)$ is, *i.e.,* $\lambda(M) = \frac{1}{\pi(M)}$. This captures a situation in which firms must plan for contingencies (realizations of the money supply) in advance and then implement these plans when the state is realized. This premise is shared by the analysis of Maćkowiak and Wiederholt (2018), who also study *ex ante* planning with mutual information costs. However, our analysis differs in the specific monetary business-cycle model that we study, the cost functions we consider, our analysis of general-equilibrium implications, and our predictions for the price distribution. Moreover, we assume that $\pi(M) \propto M^{\delta}$, where $\delta > 0$ corresponds to high money supply states being more likely and $\delta < 0$ means that low money supply states are more likely.

Recasting the economy as a game. Given the CES aggregator, the firm faces an isoelastic demand curve:

$$q_i = \left(\frac{p_i}{P}\right)^{-\eta} C \tag{28}$$

where P is the ideal price index under CES production:

$$P = \left(\int_0^1 p_i^{1-\eta} \,\mathrm{d}i \right)^{\frac{1}{1-\eta}} \tag{29}$$

The firms' profits are moreover priced according to the real stochastic discount factor (the household's marginal utility from consumption) $C^{-\sigma}$. Thus, the firm's objective function is:

$$\pi(p_i, P, C, k) = C^{-\sigma} \frac{p_i - k}{P} \left(\frac{p_i}{P}\right)^{-\eta} C = C^{1-\sigma} P^{\eta-1} (p_i - k) p_i^{-\eta} \tag{30}$$

Substituting in the equilibrium conditions that $k = PC^{\chi}$ (factor supply) and $C = \left(\frac{M}{P}\right)^{\frac{1}{\sigma}}$ (money demand), we obtain that the firm's payoff function is:

$$u(p_i, P, M) = M^{\frac{1-\sigma}{\sigma}} P^{\eta - \frac{1}{\sigma}} \left(p_i - M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}}\right) p_i^{-\eta} \tag{31}$$

To apply all of our results, we perform the standard approximation (as per Lemma 5) of the firm's objective function to second-order around the optimal price in each state. This yields the payoff function:

$$u(p_i, P, M) = \alpha(P, M) - \beta(P, M)(p_i - \gamma(P, M))^2 \tag{32}$$

where:

$$\begin{aligned}
\alpha(P, M) &= \frac{1}{\eta - 1} \left(\frac{\eta}{\eta - 1}\right)^{-\eta} M^{\frac{1-\sigma+\chi(1-\eta)}{\sigma}} P^{\eta - \frac{1}{\sigma} + (1-\eta)(1-\frac{\chi}{\sigma})} \\
\beta(P, M) &= \frac{\eta}{2} \left(\frac{\eta}{\eta - 1}\right)^{-(\eta+2)} M^{\frac{1-\sigma-\chi(\eta+1)}{\sigma}} P^{(\eta+1)\frac{\chi}{\sigma} - \frac{1}{\sigma} - 1} \\
\gamma(P, M) &= \frac{\eta}{\eta - 1} M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}}
\end{aligned} \tag{33}$$

and we impose that this game has complementarity in optimal actions by assuming factor price pressure is weaker than income effects in money demand, or $\chi < \sigma$. Finally, as is also standard, we approximate the aggregator to first order as:

$$P = \int_0^1 p_i \, \mathrm{d}i \tag{34}$$

which simply says that the aggregate price level is the average price set by firms.

Results and interpretation. To build intuition, we first characterize equilibrium in this model in the absence of *ex ante* planning frictions. In this case, the optimal price that a firm sets is given by:

$$p = \frac{\eta}{\eta - 1} M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}} \tag{35}$$

which is a constant markup over marginal cost. Thus, we observe that there is a unique equilibrium in which all firms set the same price (there is no price dispersion) and the aggregate price level is given by:

$$P = \left(\frac{\eta}{\eta - 1}\right)^{\frac{\sigma}{\chi}} M \tag{36}$$

In this equilibrium, the elasticity of prices to the money supply is one, *i.e.,* a 1% increase in the money supply leads to a 1% increase in the price level.

We now apply our general results to this economy when firms face *ex ante* planning frictions. Specifically, we ask when this price-setting economy has a unique equilibrium, when the aggregate price level is increasing in money, when the distribution of prices is increasing (in the sense of FOSD) in money, and when the dispersion in prices is highest when inflation is high.

Corollary 3. *There is a unique equilibrium if, for all $p, P \in [\underline{p}, \overline{p}]$ and $M \in \mathcal{M}$:*

$$-\left(\frac{P}{\gamma(P,M)\left(1-\frac{\chi}{\sigma}\right)}-1\right) < \frac{-1-\frac{1}{\sigma}+(\eta+1)\frac{\chi}{\sigma}}{\left(1-\frac{\chi}{\sigma}\right)}\left(\frac{p}{\gamma(P,M)}-1\right) < 1 \tag{37}$$

The unique aggregate price level and distribution of prices are both increasing in the money supply if, in addition:

$$\frac{1-\sigma-\chi(\eta+1)+\delta\sigma}{\chi}\left(\frac{p}{\gamma(P,M)}-1\right) < 1 \tag{38}$$

Moreover, price precision is decreasing in the money supply and the price level if, in addition:

$$\chi(\eta+1) \in (1+\sigma(\delta-1), 1+\sigma) \tag{39}$$

Proof. See Appendix A.9. □

To understand this result, we go through each condition in turn. The uniqueness condition (Equation (37)) comprises two inequalities. The inequality on the right ensures that the game is one of complementarities. The inequality on the left ensures that utility has sufficient concavity relative to complementarity. In the special case where the losses from mispricing do not depend on the aggregate price level ($(1+\eta)\frac{\chi}{\sigma}-\frac{1}{\sigma}-1=0$), the middle term is equal to zero and the complementarity condition always holds. This is because we have assumed that factor price pressure is weaker than income effects ($\chi < \sigma$), which makes the optimal price increase in the aggregate price. When the losses from mispricing are instead endogenous ($(1+\eta)\frac{\chi}{\sigma}-\frac{1}{\sigma}-1 \neq 0$), there is a new effect that must be accounted for. Intuitively, suppose without loss of generality that an agent is setting a price that is less than the optimal price and an increase in the aggregate price level increases (decreases) the losses from mispricing, then the agent now has a greater (lesser) incentive to reduce the magnitude of this mistake and increase (decrease) their price. In the former case, this endogeneity of the costs of mispricing induces greater strategic complementarity. In the latter case, it induces strategic substitutability. The exact inequality we derive disciplines the magnitudes of these effects in a verifiable way that ensures that strategic complementarity always obtains.

The sufficient concavity condition similarly has a "simple" and "complex" interpretation. When the losses from mispricing are exogenous, the condition requires that the optimal price has a slope less than one in the aggregate price level. This condition that "best responses have a slope less than one" is familiar from games without decision frictions. More generally, when the aggregate price matters for the losses from mispricing, the "slope" that needs to be bounded depends directly on the deviation of the price from the optimal price and the considerations described above.

The monotonicity condition (Equation (38)) requires that higher levels of the money supply are complementary with higher prices for firms. When the losses from mispricing do not depend on the money supply $\frac{1-\sigma-\chi(\eta+1)}{\sigma}+\delta=0$, this condition always holds as factor price pressure from higher money supply (which increases demand, which increases production, which increases marginal costs) makes optimal prices higher. More generally, as above for the endogenous price level, the monotonicity condition ensures complementarity between pricing and the exogenous money supply when the losses from mispricing depend on the money supply.

Finally, the condition for monotone precision (Equation (39)) conveys, in terms of deep parameters, when price-setters optimally respond to lower money and prices by making more precise decisions. To understand this, it is useful to first turn off factor price pressure or set $\chi = 0$. In this case, the condition corresponds to $\delta + \left(\frac{1}{\sigma} - 1\right) < 0$. The first term isolates the role of costly planning—when high-money states are less likely ($\delta < 0$), firms optimally put in less effort to plan for them, and their pricing decisions in these states are less precise. The second term conveys the roles of aggregate demand externalities, which have elasticity $1/\sigma$ with respect to the money supply, and the stochastic discount factor, which has elasticity -1 with respect to the money supply. The demand externality pushes toward high precision in high-demand states, because any price mistake leads to more lost sales. The stochastic discount factor pushes toward high precision in low-demand states, since profits are more valuable in these states (Flynn and Sastry, 2022). Finally, when factor price pressures are re-introduced, they loosen the constraint corresponding to incentives from the money supply and tighten the constraint corresponding to incentives from aggregate prices.

Economic lessons. Our finding can be used to rationalize empirical findings on the cyclicality of price dispersion. Empirically, Alvarez et al. (2019) find that price dispersion among firms in Argentina has an elasticity of about 1/3 to the inflation rate in high-inflation periods (*e.g.,* an annual rate above 50%) and an elasticity that is positive, but close to zero, in low-inflation periods. Nakamura et al. (2018) find limited evidence that the dispersion of US prices increased in the "Great Inflation" of the 1970s and 1980s, during which annual US inflation was regularly between 5% and 10%. This evidence is consistent with the costly planning mechanism, if firms believe that hyperinflation states (in Argentina and in the US) are relatively unlikely ($\delta < 0$) and are far in the tail of the distribution for M. Intuitively, this allows the possibility that price dispersion is especially high in hyperinflations precisely because firms have not precisely formulated plans for these unlikely states. This prediction would not be obtained in standard analysis of this model with a state-invariant decision friction, like exogenous information (Hellwig and Venkateswaran, 2009).

Moreover, empirical evidence from Brunnermeier et al. (2022) regarding the German hyperinflation of the 1920s demonstrates both firms' unpreparedness and the concrete organizational difficulties and "mistakes" that resulted from this. First, Brunnermeier et al. (2022) show how markets retained low inflation expectations until the Summer of 1922. This is despite the fact that inflation had been abnormally high since 1919. Thus, over this three-year window, people seem to remain persistently incorrect regarding inflation. Second, Brunnermeier et al. (2022) provide an historical account of how the rare, highly inflationary state led to acute organizational difficulties in making decisions because of a lack of preparedness. Indeed, a system for "inflation accounting" did not exist at that time and this led firms to even make mistakes in constructing their own balance sheets. Brunnermeier et al. (2022) write that: "For example, the 1923 financial report of Darmstadter und Nationalbank stated that 'the figures in our balance sheet and profit-and-loss statement are, as in those of all German companies, unfit for any serious scrutiny, and to examine them in detail is folly.' Similarly, Hoffmann and Walker (2020) provide examples of firms noting that the calculation of balance sheets and income in paper marks 'lost its economic meaning' and that firms only reported financial statements out of legal obligation."

4.2. Consumption and savings with a stress externality

Set-up. Each agent is a consumer that lives for infinite periods, indexed by $t \in \mathbb{N}$. They choose consumption levels $c_{it} \in \mathbb{R}$ and labor levels $n_{it} \in \mathbb{R}$ and have quadratic payoffs. They maximize expected discounted utility:

$$U(\{c_{it}, n_{it}\}_{t=0}^{\infty}, \theta_d) = (1+\theta_d)c_{it} - \frac{c_{it}^2}{2} - \chi\frac{n_{it}^2}{2} + \sum_{t=1}^{\infty}\delta^t\left(c_{it} - \frac{c_{it}^2}{2} - \chi\frac{n_{it}^2}{2}\right) \tag{40}$$

where $\delta \in (0, 1)$ is a discount factor, χ is a parameter controlling the labor-leisure trade-off, and θ_d is a *demand shock* in a discrete set Θ_d, with maximum element $\overline{\theta}_d \leq 0$ and minimal element $\underline{\theta}_d > -1$, that reduces the household's relative preference to consume in period 0. Each agent can save in a risk-free bond with interest rate $R = 1/\delta$, fixed and unresponsive to demand as in a small open economy. Each agent receives income $w_t n_{it}$ in each period, where $w_t \in \mathbb{R}_+$ is a wage and $n_{it} \in \mathbb{R}$ is the amount that agent i works. Therefore, for each period t, the agent faces a budget constraint $c_{it} + b_{it} \leq w_t n_{it} + Rb_{i,t-1}$, where b_{it} is savings and $b_{i,-1} = 0$ for all agents.

Goods are produced by a representative firm, at which all of the agents work. The firm produces output via a linear production technology, $y_t = \int_{[0,1]} n_{it}\,\mathrm{d}i$. In all periods, the output market clears as $y_t = \int_{[0,1]} c_{it}\,\mathrm{d}i$ and the bond market clears as $0 = \int_{[0,1]} b_{it}\,\mathrm{d}i$. At $t = 0$, given the fixed interest rate, these conditions would be incompatible with equilibrium in the labor market. We therefore make the conventional assumption that, in this period, the firm commits to satisfying demand at the (fixed) price, households lie off their labor supply curve, and households all work an equal amount. We refer to this period as a *liquidity trap*, since the market failure is caused by the inability of interest rates to adjust downward to accommodate the negative demand shock.

We are interested in how equilibrium at $t = 0$ is affected by demand shocks, under the assumptions that households imperfectly optimize and that their cost is affected by financial stress. To simplify our analysis, we assume that all choices for $t \geq 1$, after the economy exits the liquidity trap, are made frictionlessly. At $t = 0$, households choose $c_{i0} \in [\underline{c}, \overline{c}]$, where $\overline{c} < \frac{\delta}{1-\delta}$, to maximize expected utility net of cognitive costs, given rational expectations about future aggregates and their future behavior.[13]

We introduce the idea that stress may lead to lower-quality decisions in low-income states via the cost functional. This idea is motivated by the experimental findings of Mani et al. (2013) suggesting that poverty, transitory or persistent, reduces performance in cognitive tasks. Mullainathan and Shafir (2013) hypothesize that involuntary capture of attention toward contemplating negative outcomes in these states reduces the available *bandwidth* to make decisions, and therefore makes people more prone to "forgetfulness" and "cognitive slips" (p. 14). We model this by letting $\lambda(y, \theta_d) = y^{-\tau}$, where y is consumers' period-0 income and $\tau \geq 0$ is a parameter controlling how quickly decision costs increase when income is low. We let the cost-functional kernel be any ϕ that satisfies Assumption 3.

Our model captures in reduced form the diversion of cognitive resources away from the decision of interest for consumption and savings, and hence the *scarcity* of attention available for

[13] The condition $\overline{c} < \frac{\delta}{1-\delta}$ ensures that consumption in periods $t \geq 1$ does not exceed the bliss point.

the decision problem of interest.[14] In equilibrium, this diversion will depend on the actions (consumption) of others, because this will determine aggregate demand and, therefore, individuals' income. In this way, our model is motivated by the combination of the experimental and survey evidence of Mani et al. (2013) and Sergeyev et al. (2022) about individual-level stress and decisions; the fact that business cycles shift aggregate poverty rates (see, *e.g.*, Meyer and Sullivan, 2011); and the fact that adverse mental health outcomes and anxiety increase in aggregate during economic downturns (see, *e.g.,* Frasquilho et al., 2015).

Recasting the economy as a game. We now analyze the consumer's problem to reduce the equilibrium determination of first-period consumption to a game to which our results can be applied. It is simple to show that, for $t \geq 1$, aggregate output is fixed at a level $\bar{y} > 0$ and each agents' consumption is fixed at a specific level which depends on their period 0 savings. This exact consumption-smoothing result follows from the intertemporal Euler equation and the simplifying assumption that $\delta R = 1$ (see, *e.g.,* Hall, 1978). Next, because the payoff for $t \geq 1$ is always increasing in period 0 savings, the agent saves all unspent income at $t = 0$: $b_{i0} = y_0 - c_{i0}$. Using these observations, and defining $c_i = c_{i0}$ and $y = \int_{[0,1]} c_i \, \mathrm{d}i$, we can re-write the objective as

$$u(c, y, \theta_d) = \alpha(y, \theta_d) - \beta(y, \theta_d)(c - \gamma(y, \theta_d))^2 \tag{41}$$

where[15]

$$\begin{aligned} \gamma(y, \theta_d) &= (1 - m)(\theta_d + \bar{y}) + my \\ \beta(y, \theta_d) &= \frac{1}{2(1 - m)} \end{aligned} \tag{42}$$

and $m = \frac{\chi(1-\delta)}{\chi+\delta} \in (0, 1)$ is the agent's marginal propensity to consume (MPC), which itself depends positively on labor disutility χ and negatively on the discount factor δ. In the limit where labor supply is inelastic, or $\chi \to \infty$, then $m \to 1 - \delta$ as is familiar from the permanent income hypothesis. The payoff representation is exact, not approximate, since the original payoffs were quadratic. We finally observe that the cost shifter can be written as $\lambda(y) = y^{-\tau}$, because of the aforementioned fact that each agent's income equals aggregate demand in equilibrium in the liquidity trap. Our model therefore captures cognitive stress induced by the aggregate business cycle; since our model has no individual heterogeneity in income during the liquidity trap, it abstracts from this dimension of cognitive stress.

Results and interpretation. Applying our general results, we can provide conditions under which a generalized beauty contest has a unique equilibrium with a number of economically relevant properties:

Corollary 4. *If the following condition holds for all $c, y \in [\underline{c}, \overline{c}]$ and $\theta_d \in \Theta_d$:*

$$0 < m - \frac{\tau}{y} \left(c - (1 - m)\left(\bar{y} + \theta_d\right) - my\right) < 1 \tag{43}$$

[14] As a different, and complementary formalization that is consistent with their novel survey evidence, Sergeyev et al. (2022) formalize the Mullainathan and Shafir (2013) hypothesis as an involuntary use of time that could otherwise be allocated to labor or leisure.

[15] A more cumbersome expression for $\alpha(y, \theta_d)$ is given in Appendix A.10.

then there exists a unique equilibrium in which (i) the distribution of consumption and aggregate output are monotone increasing in the demand shock θ_d and (ii) the precision of consumption is monotone decreasing in the demand shock θ_d and in aggregate output y. Moreover, if the planner's problem is strictly concave, a necessary condition for the efficiency of the unique equilibrium is that:

$$y = \bar{y} + \frac{\tau}{\chi} y^{-\tau-1} \int_{\underline{c}}^{\bar{c}} \phi(p^*(c \mid \theta_d))\,\mathrm{d}c \tag{44}$$

Thus, whenever cognitive costs are positive, an efficient allocation in an economy with $\tau > 0$ has higher output than an efficient allocation in an economy with $\tau = 0$.

Proof. See Appendix A.10. □

The conditions in Equation (43) follow from the calculation in Lemmas 1 and 4. These conditions are trivially satisfied if $\tau = 0$ as higher demand increases income which increases consumption (as $m > 0$), but less than one-for-one since the household discounts the future and therefore has an MPC strictly less than one (as $m < 1$). If $\tau > 0$, then there are potentially countervailing forces that affect strategic complementarity. Concretely, when aggregate output increases, stress decreases, and agents' costs of precise optimization fall. If an agent is consuming more than the optimal level, this makes them prone to consume closer to the optimal level and lower their consumption, inducing strategic substitutability. Conversely, if an agent is consuming less than the optimal level, this makes them prone to consume more and induces additional strategic complementarity. The condition provides the precise conditions under which these concerns do not upset total strategic complementarity (consumption increases when income increases) and sufficient concavity (consumption increases less than one-for-one). Under these conditions, we know that higher demand increases aggregate output (point (i)); that higher demand shifts the entire distribution of consumption upward of first-order stochastic dominance (point (ii)); and that agents' actions are more precise in high states, due to their experiencing lower stress and, therefore, (endogenously) lower costs of attention in these states.

The second result, the necessary condition for efficiency, conveys that the introduction of the stress mechanism increases the optimal level of output. The reason is that the stress mechanism creates an externality operating through cognitive costs: if one agent consumes more, increasing aggregate demand and output, they reduce stress (cognitive costs) for all other agents. The extent of this externality in state θ_d is proportional to the cognitive cost paid *ex post* in that state. Thus, the externality would disappear were there no cost of cognition. And the extent of cognitive costs would not affect the optimal allocation were there no stress and, by implication, no externality operating purely through cognition.

Economic lessons. Our prediction for endogenous precision, or higher consumption "mistakes" in low output states, is consistent with the evidence from Berger et al. (2023) that the cross-sectional distribution of US consumption becomes more dispersed in recessions. In our case, this result arises because of the equilibrium effect of low income causing stress that worsens decisionmaking. Our psychological explanation complements mechanisms studied in the literature related to the cyclicality of income risk and the role of financial constraints.[16]

In the emerging literature on how household stress affects decisionmaking, our results complement those in the study of Sergeyev et al. (2022), who use original survey evidence to measure the extent of financial stress among US households and to calibrate a macroeconomic model in which financial stress distracts from productive labor supply. Our mechanism is different (stress reducing decision *quality*) and makes a different prediction, potentially in line with the data, about consumption dispersion.

Our normative results clarify how the stress channel may translate into inefficiency at the macro level. In particular, our results rationalize a "paradox of scarcity" logic: by not spending, households contribute toward lower overall output, which induces further financial stress for others and has psychological costs. This mechanism relies crucially on the *endogeneity* of income, and hence stress, to others' decisions.

Finally, we note that our analysis contrasts with abstract results in examples studied by Hébert and La'O (2022) and Angeletos and Sastry (2023) in two ways. First, we isolate a cognitive externality that may be difficult to formalize in a model of costly information acquisition (see the discussion of Fact 5 in Section 2.3). Second, we can precisely characterize equilibrium, its comparative statics properties, the equilibrium externality, and the optimal direction of policy response in an *inefficient* setting.

5. Extensions

5.1. State-separable vs. mutual information costs

Although its foundations are in information theory, the mutual information model of Sims (2003) also makes predictions for stochastic choice or "imperfect optimization." Decision-theoretic work by Caplin et al. (2022) characterizes these behavioral predictions, and Woodford (2012) and Dean and Neligh (2022) discuss how they match some, but not all, features of imperfect perception and choice in the lab. Moreover, in many applications in macroeconomics and finance, information choice is unobserved and/or not the focus of predictions *per se*. Instead, the focus is on the aforementioned predictions for imperfect optimization and how they play out in equilibrium.

In Appendix B, we contrast the predictions of state-separable and mutual information costs as alternative models of stochastic choice in large games. First, extending a result in Matějka and McKay (2015), we give abstract conditions under which the predictions of a version of the strategic mistakes model with logit costs gives identical predictions to a twin model with mutual information costs and a restriction of agents' (subjective) priors. Relaxing this condition isolates the key difference between the models—the mutual information model naturally allows agents to anchor toward commonly played actions as if they were "default points."

Next, we numerically explore a linear beauty contest game (Morris and Shin, 2002; Angeletos and Pavan, 2007) under both state-separable and mutual information costs. The model with state-separable costs predicts a unique equilibrium in which aggregate quantities are monotone in a driving shock, consistent with our abstract results. The mutual information model opens the door

[16] Moreover, Sergeyev et al. (2022) find in their original survey that liquidity constraints exacerbate reported psychological stress related to making economic decisions. Therefore, in practice, the psychological and liquidity-constraint channels may reinforce one another in a richer model that accommodates both.

to multiple equilibria, via coordination on specific support points of action distributions. We show how the equilibrium operator in the mutual information model is not a contraction map, thus providing an explicit counterexample to the possibility of using this paper's analytical tools to show similar results in a mutual-information setting.

We conclude that, while the information-acquisition underpinning and "anchoring" observation may be realistic for individual behavior in some applications, these components of the mutual information model open up the door to somewhat pathological equilibrium predictions and preclude sensible comparative statics analysis. Thus, in situations where researchers are concerned primarily with stochastic choice, the strategic mistakes model may be a tractable alternative that is still behaviorally rich enough to capture important, experimentally verified features of behavior (see Section 2.3).

5.2. *State-separable costs in binary-action games*

In Online Appendix C, we study strategic mistakes in binary-action games, which are used in many applications to capture an extensive margin of adjustment and/or to simplify analysis.[17] We derive sufficient conditions on cognitive costs and payoffs to ensure unique and monotone equilibria and illustrate our results in the context of a simple investment game with linear payoffs (as in Yang, 2015). Unlike the continuous-action games studied in our main analysis, binary-action supermodular games may have multiple equilibria with small stochastic choice frictions. This result hinges on agents' ability to waver between options that have similar payoffs, but are far apart in the action space and induce very different equilibrium externalities. This result offers the following insight for researchers interested in well-posed comparative statics and not multiplicity *per se*: a "more complex" continuous-action model, by smoothing out aggregate best-response functions, may admit simpler analysis than a comparable binary-action model.

6. Conclusion

This paper introduces a new class of *state-separable* control costs in large games. We show how these costs accommodate a rich class of decision frictions. We provide results on equilibrium existence, uniqueness, efficiency, and monotonicity of equilibrium distributions, aggregates, and mistakes. We apply these results to make robust equilibrium predictions in two macroeconomic applications, respectively to price-setting in a monetary economy and consumption and savings in a liquidity trap.

This paper's analysis of decision frictions in large games may be applicable to many additional settings in macroeconomics and finance. In Section 4, we show how to recast price-setting in a monetary economy and consumption-savings choice in a liquidity trap as games with common payoff-relevant states (the money supply or aggregate demand shock) and strategic complementarity summarized in payoffs by an aggregator (the price level or real GDP). Angeletos and Lian (2016) surveys other settings in macroeconomics and finance with similar characteristics, including asset pricing and strategic firm investment.

The following "practical guide" generalizes the steps of Section 4 and may be useful to researchers in macroeconomics and finance who want to make general equilibrium statements about the properties of economies that feature decision frictions. First, micro-found payoffs and

[17] See Angeletos and Lian (2016) (in particular, Section 5) for a review of this literature.

aggregation in the setting of interest. Second, based on an understanding of how imperfect optimization varies across states, specify an appropriate weighting function λ (or a class of plausible candidates, whose predictions one wants to contrast). Third, algebraically verify the conditions underlying our main results for equilibrium existence, equilibrium uniqueness, monotone comparative statics, and equilibrium efficiency. Fourth, use these conditions to generate *theoretically robust* and, potentially, empirically testable predictions.

Declaration of competing interest

We declare that we have no relevant financial or material interests that relate to the research described in the submitted paper.

Appendix A. Omitted proofs

A.1. Proof of Lemma 1

Proof. Define $\tilde{u} = \frac{u}{\lambda}$. Inequality (9) can be re-expressed as:

$$\tilde{u}(x', X', \theta) + \tilde{u}(x, X, \theta) \geq \tilde{u}(x', X, \theta) + \tilde{u}(x, X', \theta) \tag{45}$$

which is the statement that $\tilde{u}$ is a supermodular function in (x, X). By Topkis' Characterization Theorem (see *e.g,* Milgrom and Roberts, 1990), when $\tilde{u}$ is twice continuously differentiable in (x, X), this is equivalent to the statement that $\tilde{u}_{xX}(x, X, \theta) \geq 0$. As we have assumed that u and λ are both twice continuously differentiable in (x, X), Inequality (9) is equivalent to:

$$\tilde{u}_{xX}(x, X, \theta) = \frac{u_{xX}(x, X, \theta) - u_x(x, X, \theta)\frac{\lambda_X(X,\theta)}{\lambda(X,\theta)}}{\lambda(X, \theta)} \geq 0 \tag{46}$$

Inequality (10) can be re-expressed as:

$$\tilde{u}(x', X + \alpha, \theta) + \tilde{u}(x - \alpha, X, \theta) \leq \tilde{u}(x, X + \alpha, \theta) + \tilde{u}(x' - \alpha, X, \theta) \tag{47}$$

Define $f(y, \gamma; \theta, X) = \tilde{u}(y + \gamma, X + \gamma, \theta)$. Set $y' = x' - \alpha$, $y = x - \alpha$, $\gamma' = \alpha$ and $\gamma = 0$ and observe that $y' \geq y$ and $\gamma' \geq \gamma$. Inequality (47) is equivalent to:

$$f(y', \gamma'; \theta, X) + f(y, \gamma; \theta, X) \leq f(y, \gamma'; \theta, X) + f(y', \gamma; \theta, X) \tag{48}$$

Which is equivalent to submodularity of $f(\cdot; \theta, X)$ in (y, γ). Again by Topkis' Characterization Theorem, and by twice continuous differentiability of f in (y, γ), this is equivalent to:

$$f_{y\gamma}(y, \gamma; \theta, X) = \tilde{u}_{xx}(y + \gamma, X + \gamma, \theta) + \tilde{u}_{xX}(y + \gamma, X + \gamma, \theta) \leq 0 \tag{49}$$

Moreover, $\tilde{u}_{xx} = \frac{u_{xx}}{\lambda}$. Thus, Inequality (10) is equivalent to:

$$-\frac{u_{xx}(x, X, \theta)}{\lambda(X, \theta)} \geq \tilde{u}_{xX}(x, X, \theta) = \frac{u_{xX}(x, X, \theta) - u_x(x, X, \theta)\frac{\lambda_X(X,\theta)}{\lambda(X,\theta)}}{\lambda(X, \theta)} \tag{50}$$

Combining Inequalities (46) and (50) and multiplying by $\lambda > 0$, we obtain the claimed result. □

A.2. Proof of Lemma 3

To establish the result, as the entropy kernel has derivative $\phi'(x) = 1 + \log x$ and $\phi'(x) = x$, it is sufficient to show the following:

Lemma 6. *$\mathcal{F}$, the class of functions satisfying quasi-MLRP, contains* $\{\log(\cdot), Id(\cdot)\}$.

Proof. To see that quasi-MLRP is satisfied for $f(x) = \log x$ (and $1 + \log x$), the required condition (Equation (16)) becomes:

$$\left(\frac{g'(x')}{g'(x)} \geq \frac{g(x')}{g(x)} \; \forall x' \geq x \right) \implies g' \succeq_{FOSD} g \tag{51}$$

The left-hand side of this implication is simply the MLRP property. Moreover, MLRP implies FOSD. We now prove that $f(x) = x$ satisfies quasi-MLRP. This requires us to prove that for any two distributions $g', g \in \Delta(\mathcal{X})$:

$$\left(g'(x') - g'(x) \geq g(x') - g(x) \; \forall x' \geq x \right) \implies g' \succeq_{FOSD} g \tag{52}$$

To do this, we first prove a technical lemma, which may be of future use for characterizing other functions that satisfy quasi-MLRP:

Lemma 7. *For any two distributions $g', g \in \Delta(\mathcal{X})$, the following holds:*

$$\begin{gathered} \left(f(g'(x')) - f(g'(x)) \geq f(g(x')) - f(g(x)) \; \forall x' \geq x \right) \implies \\ \left(\frac{\int_x^{\overline{x}} \left[f(g'(\tilde{x})) - f(g(\tilde{x})) \right] \mathrm{d}\tilde{x}}{\overline{x} - x} \geq \frac{\int_{\underline{x}}^{x} \left[f(g'(\tilde{x})) - f(g(\tilde{x})) \right] \mathrm{d}\tilde{x}}{x - \underline{x}} \quad \forall x \in \mathcal{X} \right) \end{gathered} \tag{53}$$

Proof. To prove the required implication, we begin with the hypothesis:

$$f(g'(x')) - f(g'(x)) \geq f(g(x')) - f(g(x)), \; \forall x' \geq x \tag{54}$$

Which can be rewritten as:

$$f(g'(x')) + f(g(x)) \geq f(g(x')) + f(g'(x)), \; \forall x' \geq x \tag{55}$$

We now integrate from $\underline{x}$ to x' with respect to x to obtain the inequality:

$$(x' - \underline{x}) f(g'(x')) + \int_{\underline{x}}^{x'} f(g(x)) \, \mathrm{d}x \geq (x' - \underline{x}) f(g(x')) + \int_{\underline{x}}^{x'} f(g'(x)) \, \mathrm{d}x \tag{56}$$

Imposing $x' = x$ we obtain:

$$(x - \underline{x}) \left[f(g'(x)) - f(g(x)) \right] \geq \int_{\underline{x}}^{x} \left[f(g'(\tilde{x})) - f(g(\tilde{x})) \right] \mathrm{d}\tilde{x} \tag{57}$$

Applying the same procedure but this time integrating from x to $\overline{x}$ with respect to x' and evaluate at $x' = x$ to obtain this inequality:

$$\int_{x}^{\overline{x}} \left[f(g'(\tilde{x})) - f(g(\tilde{x})) \right] \mathrm{d}\tilde{x} \geq (\overline{x} - x) \left[f(g'(x)) - f(g(x)) \right] \tag{58}$$

Combining our two inequalities we obtain the required one:

$$\frac{\int_{x}^{\overline{x}} \left[f(g'(\tilde{x})) - f(g(\tilde{x})) \right] \mathrm{d}\tilde{x}}{\overline{x} - x} \geq \frac{\int_{\underline{x}}^{x} \left[f(g'(\tilde{x})) - f(g(\tilde{x})) \right] \mathrm{d}\tilde{x}}{x - \underline{x}} \quad \forall x \in \mathcal{X} \tag{59}$$

Which completes the proof. □

Thus, if it can be established that:

$$\left(\frac{\int_{x}^{\overline{x}} \left[f(g'(\tilde{x})) - f(g(\tilde{x})) \right] \mathrm{d}\tilde{x}}{\overline{x} - x} \geq \frac{\int_{\underline{x}}^{x} \left[f(g'(\tilde{x})) - f(g(\tilde{x})) \right] \mathrm{d}\tilde{x}}{x - \underline{x}} \quad \forall x \in \mathcal{X} \right) \implies g' \succeq_{FOSD} g \tag{60}$$

then we will have established that function f satisfies quasi-MLRP.

We now use this to prove that $f(x) = x$ satisfies quasi-MLRP. Plugging in to the derived integral condition, we obtain:

$$\frac{G(x) - G'(x)}{\overline{x} - x} \geq \frac{G'(x) - G(x)}{x - \underline{x}} \quad \forall x \in \mathcal{X} \tag{61}$$

Re-arranging this:

$$G(x) \geq G'(x) \quad \forall x \in \mathcal{X} \tag{62}$$

which is the definition that $g' \succeq_{FOSD} g$. This completes the proof and establishes that quasi-MLRP is a strict weakening of MLRP. □

A.3. *Proof of Proposition 1*

Proof. This follows immediately from the step proving the monotonicity and discounting conditions in Theorem 1. Note that this invokes only Assumptions 1 and 3. □

A.4. *Proof of Theorem 1*

Proof. We first study the problem of a single agent i who is best replying to the conjecture that the law of motion of the aggregate is $\hat{X} : \Theta \to \mathbb{R}$. See that this agent faces the problem:

$$\mathcal{P}^*(\hat{X}) = \arg\max_{P \in \mathcal{P}} \sum_{\Theta} \int_{\mathcal{X}} u(x, \hat{X}(\theta), \theta) \, \mathrm{d}P(x|\theta) \, \pi(\theta) - c(P, \hat{X}) \tag{63}$$

First, let us examine the set of stochastic choice rules:

$$\mathcal{P} = \{P : \Theta \to \Delta(\mathcal{X})\} = \prod_{\theta \in \Theta} \Delta(\mathcal{X}) \tag{64}$$

See that $\Delta(\mathcal{X})$ is compact as $\mathcal{X}$ is compact. It therefore follows by finiteness of Θ that $\mathcal{P}$ is compact.

Define $k : \mathcal{P} \times \mathcal{B} \to \bar{\mathbb{R}}$, where $\mathcal{B} = \{\hat{X} : \Theta \to \mathbb{R}\}$ as:

$$k(P, \hat{X}) = \sum_{\Theta} \int_{\mathcal{X}} u(x, \hat{X}(\theta), \theta)\, \mathrm{d}P(x|\theta)\, \pi(\theta) - c(P, \hat{X}) \tag{65}$$

As ϕ is strictly convex and u is bounded, it is without loss of optimality to restrict to optimizing over the set of stochastic choice rules with density bounded above by some $M \in \mathbb{R}$, $\mathcal{P}_M$. This is a closed set, which is a subset of a compact set $\mathcal{P}$, and therefore also compact. Moreover, k is continuous in P, by continuity of u and continuity of c over $\mathcal{P}_M$ for any M. Thus, by Weierstrass' theorem, there exists a maximum. Moreover, by strict convexity of $k(\cdot, \hat{X})$, it is unique. It immediately follows that in any equilibrium $P_i^* = P^* = \mathcal{P}^*(\hat{X})$ for all i and thus that there cannot exist asymmetric equilibria.

To show existence of an equilibrium it suffices to show that there exists a $\hat{X}$ such that:

$$\hat{X} = X \circ \mathcal{P}^*(\hat{X}) \tag{66}$$

To this end define the operator $T : \mathcal{B} \to \mathcal{B}$ such that:

$$T(\hat{X}) = X \circ \mathcal{P}^*(\hat{X}) \tag{67}$$

We wish to show that T has a fixed point. We will moreover prove that this fixed point is unique as under the stated assumptions we can prove that T is a contraction map. To this end, we wish to apply Blackwell's sufficient conditions for an operator to be a contraction. More specifically, if T operates on the space of bounded functions and is endowed with the sup norm, then the following are sufficient for T to be a contraction:

1. Monotonicity: $\hat{X}' \geq \hat{X} \implies T(\hat{X}') \geq T(\hat{X})$ for any $\hat{X}', \hat{X} \in \mathcal{B}$
2. Discounting: there exists $\beta \in (0, 1)$ such that $T(\hat{X} + \alpha) \leq T(\hat{X}) + \beta\alpha$ for all $\alpha \in \mathbb{R}_+$ and any $\hat{X} \in \mathcal{B}$

Toward proving these properties, we first derive some necessary conditions for optimal stochastic choice. To this end, see that the stochastic choice program under an equilibrium conjecture $\hat{X}$ is given by:

$$\max_{p \in \mathcal{P}} \sum_{\Theta} \int_{\mathcal{X}} u(x, \hat{X}(\theta), \theta)\, \mathrm{d}P(x|\theta)\, \pi(\theta) - \sum_{\Theta} \int_{\mathcal{X}} \phi(p(x|\theta))\, \mathrm{d}x\, \pi(\theta) \lambda(\hat{X}(\theta), \theta) \tag{68}$$

Take the optimal policy p and now consider a family or perturbations of p around actions $x, x' \in \mathcal{X}$ in state $\theta \in \Theta$ such that $p(x|\theta; \hat{X}), p(x'|\theta; \hat{X}) > 0$ by $\varepsilon > 0$ and $\delta \geq 0$ such that:

$$\begin{aligned} \tilde{p}(\tilde{x}|\theta; \hat{X}) &= p(\tilde{x}|\theta; \hat{X}) + \varepsilon, \ \tilde{x} \in [x', x' - \delta] \\ \tilde{p}(\tilde{x}|\theta; \hat{X}) &= p(\tilde{x}|\theta; \hat{X}) - \varepsilon, \ \tilde{x} \in [x, x - \delta] \end{aligned} \tag{69}$$

For p that has full support on $[x', x' - \delta]$, $[x, x - \delta]$, we have that $\tilde{p} \in \mathcal{P}$. Moreover, as u is continuous, if δ is sufficiently small, such a full-support perturbation is possible by the property that $p(x|\theta; \hat{X}), p(x'|\theta; \hat{X}) > 0$ and the fact that p is optimal.

Taking the derivative of the value of $\tilde{p}$ in ε and evaluating at $\varepsilon = 0$, we obtain the necessary optimality condition:

$$\int_{x'-\delta}^{x'} \left[u(\tilde{x}, \hat{X}(\theta), \theta)\pi(\theta) - \phi'(p(\tilde{x}|\theta; \hat{X}))\pi(\theta)\lambda(\hat{X}(\theta), \theta)\right] \mathrm{d}\tilde{x} = \int_{x-\delta}^{x} \left[u(\tilde{x}, \hat{X}(\theta), \theta)\pi(\theta) - \phi'(p(\tilde{x}|\theta; \hat{X}))\pi(\theta)\lambda(\hat{X}(\theta), \theta)\right] \mathrm{d}\tilde{x} \tag{70}$$

Normalizing both sides by $\delta > 0$, we obtain:

$$\frac{\int_{x'-\delta}^{x'} \left[u(\tilde{x}, \hat{X}(\theta), \theta)\pi(\theta) - \phi'(p(\tilde{x}|\theta; \hat{X}))\pi(\theta)\lambda(\hat{X}(\theta), \theta)\right] \mathrm{d}\tilde{x}}{\delta} = \frac{\int_{x-\delta}^{x} \left[u(\tilde{x}, \hat{X}(\theta), \theta)\pi(\theta) - \phi'(p(\tilde{x}|\theta; \hat{X}))\pi(\theta)\lambda(\hat{X}(\theta), \theta)\right] \mathrm{d}\tilde{x}}{\delta} \tag{71}$$

Taking the limit of both sides as $\delta \to 0$, applying L'Hôpital's rule and Leibniz's rule we obtain:

$$u(x', \hat{X}(\theta), \theta) - \lambda(\hat{X}(\theta), \theta)\phi'(p(x'|\theta; \hat{X})) = u(x, \hat{X}(\theta), \theta) - \lambda(\hat{X}(\theta), \theta)\phi'(p(x|\theta; \hat{X})) \tag{72}$$

This condition is necessary for all $x, x' \in \mathcal{X}$ that have a positive density in state θ.

By the previous necessary condition and the supermodularity assumption (Assumption 1) we have that, for all $x' \geq x$ in the support of both stochastic choice rules, all θ, and any conjectures $\hat{X}$ and $\hat{X}'$ such that $\hat{X}' \geq \hat{X}$:

$$\begin{aligned}\phi'(p(x'|\theta; \hat{X}')) - \phi'(p(x|\theta; \hat{X}')) &= \frac{u(x', \hat{X}'(\theta), \theta)}{\lambda(\hat{X}'(\theta), \theta)} - \frac{u(x, \hat{X}'(\theta), \theta)}{\lambda(\hat{X}'(\theta), \theta)} \\ &\geq \frac{u(x', \hat{X}(\theta), \theta)}{\lambda(\hat{X}(\theta), \theta)} - \frac{u(x, \hat{X}(\theta), \theta)}{\lambda(\hat{X}(\theta), \theta)} \\ &= \phi'(p(x'|\theta; \hat{X})) - \phi'(p(x|\theta; \hat{X}))\end{aligned} \tag{73}$$

We now need to check the cases where the stochastic choice rules do not have full support. Define the support in state θ under law of motion $\hat{X}$ as $\mathcal{X}(\theta, \hat{X}) = \mathrm{cl}_{\mathcal{X}}\{x \in \mathcal{X} : p^*(x|\theta, \mathcal{X}) > 0\}$. Let $\hat{x} \in \mathcal{X}(\theta, \hat{X})$, $\tilde{x} \in \mathcal{X}(\theta, \hat{X}')$ and define $x' = \max\{\hat{x}, \tilde{x}\}$, $x = \min\{\hat{x}, \tilde{x}\}$. We proceed to show that $\mathcal{X}(\theta, \hat{X})$ is monotone in the strong set order in $\hat{X}$. That is, for $\hat{X}' \geq \hat{X}$, we have that $x' \in \mathcal{X}(\theta, \hat{X}')$ and $x \in \mathcal{X}(\theta, \hat{X})$. By Assumption 1, u is a concave function of x. This implies that $\mathcal{X}(\theta, \hat{X})$ is an interval. We will denote its lower end-point by $\underline{x}(\theta, \hat{X})$ and its upper end-point by $\overline{x}(\theta, \hat{X})$. We also note that $p(\underline{x}(\theta, \hat{X})|\theta, \hat{X}) = p(\overline{x}(\theta, \hat{X})|\theta, \hat{X}) = 0$. Showing that $\mathcal{X}(\theta, \hat{X})$ is increasing in the strong set order therefore reduces to showing that $\underline{x}(\theta, \hat{X}) \leq \underline{x}(\theta, \hat{X}')$ and $\overline{x}(\theta, \hat{X}) \leq \overline{x}(\theta, \hat{X}')$. Without loss of generality (the other case follows by identical arguments), we will show that $\overline{x}(\theta, \hat{X}) \leq \overline{x}(\theta, \hat{X}')$.

Toward a contradiction, suppose that $\overline{x}(\theta, \hat{X}) > \overline{x}(\theta, \hat{X}')$. There are two cases to consider: the case in which the supports strictly overlap $\underline{x}(\theta, \hat{X}) < \overline{x}(\theta, \hat{X}')$, and the case in which they do not $\underline{x}(\theta, \hat{X}) \geq \overline{x}(\theta, \hat{X}')$. First, consider the case in which $\underline{x}(\theta, \hat{X}) < \overline{x}(\theta, \hat{X}')$. By continuity of $p(\cdot|\theta, \hat{X}')$ and $p(\cdot|\theta, \hat{X})$, the fact that $0 = p(\underline{x}(\theta, \hat{X})|\theta, \hat{X}) < p(\underline{x}(\theta, \hat{X})|\theta, \hat{X}')$, and the fact that $p(x|\theta, \hat{X}) > 0$ for $x \in (\underline{x}(\theta, \hat{X}), \overline{x}(\theta, \hat{X}'))$, there exists an $x \in (\underline{x}(\theta, \hat{X}), \overline{x}(\theta, \hat{X}'))$ such that $p(x|\theta, \hat{X}') > p(x|\theta, \hat{X})$. Fix also any $x' \in (\overline{x}(\theta, \hat{X}'), \overline{x}(\theta, \hat{X}))$. Consider a perturbation, as per

Equation (69) that moves density from (a neighborhood of) x to (a neighborhood of) x' in state θ under conjecture $\hat{X}'$. We have that the following holds:

$$\begin{aligned}\phi'(0) - \phi'(p(x|\theta,\hat{X}')) &< \phi'(p(x'|\theta,\hat{X})) - \phi'(p(x|\theta,\hat{X})) \\ &= \frac{u(x',\theta,\hat{X}(\theta))}{\lambda(\hat{X}(\theta),\theta)} - \frac{u(x,\theta,\hat{X}(\theta))}{\lambda(\hat{X}(\theta),\theta)} \\ &\leq \frac{u(x',\theta,\hat{X}'(\theta))}{\lambda(\hat{X}'(\theta),\theta)} - \frac{u(x,\theta,\hat{X}'(\theta))}{\lambda(\hat{X}'(\theta),\theta)}\end{aligned} \tag{74}$$

where the first line follows from the strict convexity of ϕ, the fact that $p(x'|\theta,\hat{X}) > 0$, and the fact that $p(x|\theta,\hat{X}') > p(x|\theta,\hat{X})$. The second line follows from the optimality of $p(\cdot|\theta,\hat{X})$. The third line follows by Assumption 1. However, Equation (74) implies that the considered perturbation provides a strict gain relative to $p(\cdot|\theta,\hat{X}')$, which contradicts the optimality of $p(\cdot|\theta,\hat{X}')$.

Consider now the case in which $\underline{x}(\theta,\hat{X}) \geq \overline{x}(\theta,\hat{X}')$. By the fact that $p(\cdot|\theta,\hat{X})$ is strictly preferred to $p(\cdot|\theta,\hat{X}')$ when the aggregate follows $\hat{X}$, we have that:

$$\begin{aligned}&\int_{\underline{x}(\theta;\hat{X})}^{\overline{x}(\theta;\hat{X})} \frac{u(x,\hat{X}(\theta),\theta)}{\lambda(\hat{X}(\theta),\theta)} p(x|\theta,\hat{X})\,\mathrm{d}x - \int_{\underline{x}(\theta;\hat{X}')}^{\overline{x}(\theta;\hat{X}')} \frac{u(x,\hat{X}(\theta),\theta)}{\lambda(\hat{X}(\theta),\theta)} p(x|\theta,\hat{X}')\,\mathrm{d}x \\ &\quad > \int_{\underline{x}(\theta;\hat{X})}^{\overline{x}(\theta;\hat{X})} \phi(p(x|\theta,\hat{X}))\,\mathrm{d}x - \int_{\underline{x}(\theta;\hat{X}')}^{\overline{x}(\theta;\hat{X}')} \phi(p(x|\theta,\hat{X}'))\,\mathrm{d}x\end{aligned} \tag{75}$$

Moreover, as $\underline{x}(\theta,\hat{X}) > \overline{x}(\theta,\hat{X}')$, we have by Assumption 1 that:

$$\begin{aligned}&\int_{\underline{x}(\theta;\hat{X})}^{\overline{x}(\theta;\hat{X})} \frac{u(x,\hat{X}'(\theta),\theta)}{\lambda(\hat{X}'(\theta),\theta)} p(x|\theta,\hat{X})\,\mathrm{d}x - \int_{\underline{x}(\theta;\hat{X}')}^{\overline{x}(\theta;\hat{X}')} \frac{u(x,\hat{X}'(\theta),\theta)}{\lambda(\hat{X}'(\theta),\theta)} p(x|\theta,\hat{X}')\,\mathrm{d}x \\ &\quad \geq \int_{\underline{x}(\theta;\hat{X})}^{\overline{x}(\theta;\hat{X})} \frac{u(x,\hat{X}(\theta),\theta)}{\lambda(\hat{X}(\theta),\theta)} p(x|\theta,\hat{X})\,\mathrm{d}x - \int_{\underline{x}(\theta;\hat{X}')}^{\overline{x}(\theta;\hat{X}')} \frac{u(x,\hat{X}(\theta),\theta)}{\lambda(\hat{X}(\theta),\theta)} p(x|\theta,\hat{X}')\,\mathrm{d}x\end{aligned} \tag{76}$$

Combining these inequalities, we obtain that:

$$\begin{aligned}&\int_{\underline{x}(\theta;\hat{X})}^{\overline{x}(\theta;\hat{X})} u(x,\hat{X}'(\theta),\theta) p(x|\theta,\hat{X})\,\mathrm{d}x - \lambda(\hat{X}'(\theta),\theta) \int_{\underline{x}(\theta;\hat{X})}^{\overline{x}(\theta;\hat{X})} \phi(p(x|\theta,\hat{X}))\,\mathrm{d}x > \\ &\quad \int_{\underline{x}(\theta;\hat{X}')}^{\overline{x}(\theta;\hat{X}')} u(x,\hat{X}'(\theta),\theta) p(x|\theta,\hat{X}')\,\mathrm{d}x - \lambda(\hat{X}'(\theta),\theta) \int_{\underline{x}(\theta;\hat{X}')}^{\overline{x}(\theta;\hat{X}')} \phi(p(x|\theta,\hat{X}'))\,\mathrm{d}x\end{aligned} \tag{77}$$

which implies that $p(\cdot|\theta,\hat{X})$ is strictly better than $p(\cdot|\theta,\hat{X}')$ when the aggregate follows $\hat{X}'$, which contradicts the optimality of $p(\cdot|\theta,\hat{X}')$.

Thus, we have shown that $\mathcal{X}(\theta, \hat{X})$ is monotone in the strong set order in $\hat{X}$ and we have derived (by Equation (72)) that:

$$\phi'(p(x'|\theta; \hat{X}')) - \phi'(p(x|\theta; \hat{X}')) \geq \phi'(p(x'|\theta; \hat{X})) - \phi'(p(x|\theta; \hat{X})) \tag{78}$$

for all $x' \geq x$ such that $x', x \in \mathcal{X}(\theta, \hat{X}) \cap \mathcal{X}(\theta, \hat{X}')$. Thus, if ϕ' satisfies quasi-MLRP (Assumption 3), then we have that $p(\theta; \hat{X}') \succeq_{FOSD} p(\theta; \hat{X})$ for all θ. It then follows by the monotonicity property of the aggregator (Assumption 2) that $X(p(\theta; \hat{X}')) \geq X(p(\theta; \hat{X}))$ for all θ and therefore that $T(\hat{X}') \geq T(\hat{X})$, which establishes the required monotonicity property of the equilibrium operator.

We now prove discounting. To this end, we will show that when we take $\hat{X}' = \hat{X} + \alpha$ for $\alpha \in \mathbb{R}_+$ that the resulting stochastic choice is dominated by an α right translation of the original stochastic choice under $\hat{X}$. Under this transformation, observe by the necessary condition for optimality and the sufficient concavity condition on utility (Assumption 1), we can apply the same arguments as above to derive that for all $x' \geq x$ such that $x', x \in \mathcal{X}(\theta, \hat{X}) \cap \mathcal{X}(\theta, \hat{X} + \alpha)$:

$$\begin{aligned}
\phi'(p_{-\alpha}(x'|\theta; \hat{X})) - \phi'(p_{-\alpha}(x|\theta, \hat{X})) &= \frac{u(x' - \alpha, \hat{X}(\theta), \theta)}{\lambda(\hat{X}(\theta), \theta)} - \frac{u(x - \alpha, \hat{X}(\theta), \theta)}{\lambda(\hat{X}(\theta), \theta)} \\
&\geq \frac{u(x', \hat{X}(\theta) + \alpha, \theta)}{\lambda(\hat{X}(\theta) + \alpha, \theta)} - \frac{u(x, \hat{X}(\theta) + \alpha, \theta)}{\lambda(\hat{X}(\theta) + \alpha, \theta)} \\
&= \phi'(p(x'|\theta; \hat{X} + \alpha)) - \phi'(p(x|\theta; \hat{X} + \alpha))
\end{aligned} \tag{79}$$

We now show that $\underline{x}(\theta, \hat{X}) + \alpha \geq \underline{x}(\theta, \hat{X} + \alpha)$ and $\overline{x}(\theta, \hat{X}) + \alpha \geq \overline{x}(\theta, \hat{X} + \alpha)$. Without loss of generality (the other case follows by identical arguments), we will show that $\overline{x}(\theta, \hat{X}) + \alpha \geq \overline{x}(\theta, \hat{X} + \alpha)$. Toward a contradiction, suppose that $\overline{x}(\theta, \hat{X}) + \alpha < \overline{x}(\theta, \hat{X} + \alpha)$. As in the previous argument, there are two cases to consider, the case in which the supports strictly overlap $\overline{x}(\theta, \hat{X}) + \alpha > \underline{x}(\theta, \hat{X} + \alpha)$ and the case in which the supports are disjoint $\overline{x}(\theta, \hat{X}) + \alpha \leq \underline{x}(\theta, \hat{X} + \alpha)$. In the overlapping support case, fix an $x \in (\underline{x}(\theta, \hat{X} + \alpha), \overline{x}(\theta, \hat{X}) + \alpha)$ such that $p_{-\alpha}(x|\theta, \hat{X}) > p(x|\theta, \hat{X} + \alpha) > 0$ (which is possible by the same arguments used in the first part of the proof). Fix next a point $x' \in (\overline{x}(\theta, \hat{X}) + \alpha, \overline{x}(\theta, \hat{X} + \alpha))$. And consider a perturbation that moves density from $x - \alpha$ to $x' - \alpha$ in state θ under conjecture $\hat{X}$. The following holds:

$$\begin{aligned}
\phi'(0) - \phi'(p_{-\alpha}(x|\theta, \hat{X})) &< \phi'(p(x'|\theta, \hat{X} + \alpha)) - \phi'(p(x|\theta, \hat{X} + \alpha)) \\
&= \frac{u(x', \hat{X}(\theta) + \alpha, \theta)}{\lambda(\hat{X}(\theta) + \alpha, \theta)} - \frac{u(x, \hat{X}(\theta) + \alpha, \theta)}{\lambda(\hat{X}(\theta) + \alpha, \theta)} \\
&\leq \frac{u(x' - \alpha, \hat{X}(\theta), \theta)}{\lambda(\hat{X}(\theta), \theta)} - \frac{u(x - \alpha, \hat{X}(\theta), \theta)}{\lambda(\hat{X}(\theta), \theta)}
\end{aligned} \tag{80}$$

where the first inequality follows by construction, the second by optimality, and the third by Assumption 1. This contradicts the optimality of $p(\cdot|\theta, \hat{X})$. In the non-overlapping case, we can follow the same steps as the first part of the proof, adapted in the obvious way (using the sufficient concavity inequality in place of the supermodularity inequality).

Thus, by quasi-MLRP of ϕ' (Assumption 3), we have that $p_{-\alpha}(\theta, \hat{X}) \succeq_{FOSD} p(\theta, \hat{X} + \alpha)$ where $p_{-\alpha}$ is the described right translation by α of p. Moreover, by the discounting property of the aggregator (Assumption 2), we then have that:

$$T(\hat{X} + \alpha) \leq X \circ p_{-\alpha}(\hat{X}) \leq T(\hat{X}) + \beta\alpha \tag{81}$$

which establishes the discounting property of T. We have now shown that T satisfies Blackwell's sufficient conditions and is a contraction map. By the Banach fixed point theorem, there then exists a unique equilibrium Ω. □

A.5. Proof of Theorem 2

Proof. To show that the unique equilibrium aggregate law of motion of monotone in θ, we use Corollary 1 from Chapter 3 of Stokey et al. (1989).

Define the set of monotone increasing and bounded functions $\mathcal{M} = \{\hat{X} \in \mathcal{B} | \hat{X}(\theta') \geq \hat{X}(\theta) \quad \forall \theta, \theta' \in \Theta : \theta' \geq \theta\}$. See that this set is closed. If we can show that $T(\hat{X}) \in \mathcal{M}$ for any $\hat{X} \in \mathcal{M}$, then we know that the unique fixed point of T is in $\mathcal{M}$ and therefore that the unique equilibrium law of motion is in $\mathcal{M}$ according to Corollary 1 of Stokey et al. (1989). To this end, we wish to show that:

$$\hat{X}(\theta') \geq \hat{X}(\theta) \quad \forall \theta, \theta' \in \Theta : \theta' \geq \theta \implies T(\hat{X})(\theta') \geq T(\hat{X})(\theta) \quad \forall \theta, \theta' \in \Theta : \theta' \geq \theta \tag{82}$$

This follows immediately from the necessary condition used in the proof of Theorem 1. More precisely, by the necessary optimality condition (Equation (72)) from the proof of Theorem 1 and Assumption 4, we have that for all $x' \geq x$ such that $x', x \in \mathcal{X}(\theta) \cap \mathcal{X}(\theta')$

$$\begin{aligned} \phi'(p(x'|\theta', \hat{X})) - \phi'(p(x|\theta', \hat{X})) &\geq \frac{u(x', \hat{X}(\theta'), \theta')}{\lambda(\hat{X}(\theta'), \theta')} - \frac{u(x, \hat{X}(\theta'), \theta')}{\lambda(\hat{X}(\theta'), \theta')} \\ &\geq \frac{u(x', \hat{X}(\theta), \theta)}{\lambda(\hat{X}(\theta), \theta)} - \frac{u(x, \hat{X}(\theta), \theta)}{\lambda(\hat{X}(\theta), \theta)} \\ &= \phi'(p(x'|\theta, \hat{X})) - \phi'(p(x|\theta, \hat{X})) \end{aligned} \tag{83}$$

In the case where optimal action distributions do not have full support, the same arguments for the monotonicity of $\mathcal{X}(\theta, \hat{X})$ imply monotonicity of $\mathcal{X}(\theta) = \mathcal{X}(\theta, \hat{X}(\theta))$ in the strong set order when $\hat{X}$ is monotone increasing. Thus, by the quasi-MLRP property of ϕ' (Assumption 3) we then have that $p(\theta'; \hat{X}) \succeq_{FOSD} p(\theta; \hat{X})$ and thus by the monotonicity of the aggregator (Assumption 2) that $T(\hat{X})(\theta') \geq T(\hat{X})(\theta)$. □

A.6. Proof of Theorem 3

Proof. Recall also by Theorem 1, that the unique symmetric stochastic choice rule consistent with the unique equilibrium $\hat{X}$ solves the following program:

$$p \in \arg\max_{p \in \mathcal{P}} \sum_{\Theta} \int_{\mathcal{X}} u(x, \hat{X}(\theta), \theta) \, \mathrm{d}P(x|\theta) \, \pi(\theta) - \sum_{\Theta} \int_{\mathcal{X}} \phi(p(x|\theta)) \mathrm{d}x \, \pi(\theta) \lambda(\hat{X}(\theta), \theta) \tag{84}$$

where we will suppress the dependence of the optimal policy on $\hat{X}$ as it is unique. Applying the necessary optimal condition from the proof of Theorem 1 (Equation (72)), for a given x such that $p(x|\theta) > 0$, we have that:

$$\begin{aligned} &u(\gamma(\hat{X}(\theta), \theta), \hat{X}(\theta), \theta) - u(x, \hat{X}(\theta), \theta) \\ &\quad = \lambda(\hat{X}(\theta), \theta) \left(\phi'(p(\gamma(\hat{X}(\theta), \theta)|\theta)) - \phi'(p(x|\theta)) \right) \end{aligned} \tag{85}$$

Under Assumption 5, we moreover have that

$$u(x, X, \theta) = \alpha(X, \theta) - \beta(X, \theta)\Gamma(|x - \gamma(X, \theta)|) \tag{86}$$

Thus our necessary condition simplifies to:

$$\frac{\beta(\hat{X}(\theta), \theta)}{\lambda(\hat{X}(\theta), \theta)}\Gamma(|x - \gamma(\hat{X}(\theta), \theta)|) = \phi'(p(\gamma(\hat{X}(\theta), \theta)|\theta)) - \phi'(p(x|\theta)) \tag{87}$$

Now consider any θ, θ' such that $\tilde{\beta}(\theta', \hat{X}(\theta')) \geq \tilde{\beta}(\theta, \hat{X}(\theta))$ (where $\tilde{\beta} = \frac{\beta}{\lambda}$). Note that, by Theorem 2, the aggregate $\hat{X}$ is monotone increasing in the state θ. Thus if $\tilde{\beta}(\theta, X)$ is decreasing in both arguments, the stated case corresponds to $\theta' \leq \theta$. If instead $\tilde{\beta}(\theta, X)$ is increasing in both arguments, the stated case corresponds to $\theta' \geq \theta$. Therefore, to verify the desired result, we now prove that the action distribution in state θ' is more precise about $\gamma(\theta', \hat{X}(\theta'))$ than the action distribution in state θ is about $\gamma(\theta, \hat{X}(\theta))$, with respect to ϕ'.

To that end, we take x, x' such that:

$$|x - \gamma(\hat{X}(\theta), \theta)| = |x' - \gamma(\hat{X}(\theta'), \theta')| \tag{88}$$

It follows that:

$$\begin{aligned}\phi'(p(\gamma(\hat{X}(\theta), \theta)|\theta)) - \phi'(p(x|\theta)) &= \tilde{\beta}(\hat{X}(\theta), \theta)\Gamma(|x - \gamma(\hat{X}(\theta), \theta)|) \\ &\geq \tilde{\beta}(\hat{X}(\theta'), \theta')\Gamma(|x' - \gamma(\hat{X}(\theta'), \theta')|) \\ &= \phi'(p(\gamma(\hat{X}(\theta'), \theta')|\theta')) - \phi'(p(x'|\theta))\end{aligned} \tag{89}$$

We now take care of those points that have no density. To this end consider the first-order condition for $p(x|\theta)$:

$$u(x, \hat{X}(\theta), \theta) - \phi'(p(x|\theta)) - \lambda(\theta) - \kappa(x, \theta) = 0 \tag{90}$$

where $\lambda(\theta)$ is the Lagrange multiplier on the constraint that $\int_{\mathcal{X}} p(x|\theta) = 1$ and $\kappa(x, \theta)$ is the Lagrange multiplier on the constraint that $p(x|\theta) \geq 0$. When $p(x|\theta) = 0$, we have that $\kappa(x, \theta) \leq 0$. Given our assumption on utility, this is given by:

$$\kappa(x, \theta) = -\beta(\hat{X}(\theta), \theta)\Gamma(|x - \gamma(\hat{X}(\theta), \theta)|) + \alpha(\hat{X}(\theta), \theta) - \lambda(\theta) \tag{91}$$

which is monotonically decreasing in $|x - \gamma(\hat{X}(\theta), \theta)|$. Thus, if there is an x such that $p(x|\theta) = 0$, then there exists an $\bar{x}(\theta)$ such that $p(x|\theta) = 0$ if and only if $|x - \gamma(\hat{X}(\theta), \theta)| \geq |\bar{x}(\theta) - \gamma(\hat{X}(\theta), \theta)|$. Moreover, by monotonicity of $\tilde{\beta}(\hat{X}(\theta), \theta)$ in θ, we have that $|\bar{x}(\theta) - \gamma(\hat{X}(\theta), \theta)| \leq |\bar{x}(\theta') - \gamma(\hat{X}(\theta'), \theta')|$. Hence, so long as $x \in [\gamma(\hat{X}(\theta), \theta) - \bar{x}(\theta), \gamma(\hat{X}(\theta), \theta) + \bar{x}(\theta)]$, we always have that $x' \in [\gamma(\hat{X}(\theta'), \theta') - \bar{x}(\theta'), \gamma(\hat{X}(\theta'), \theta') + \bar{x}(\theta')]$. Thus, every element of the support of $p(\theta)$ satisfies:

$$\phi'(p(\gamma(\hat{X}(\theta), \theta)|\theta)) - \phi'(p(x|\theta)) \geq \phi'(p(\gamma(\hat{X}(\theta'), \theta')|\theta')) - \phi'(p(x'|\theta)) \tag{92}$$

It follows then by the definition of precision that $p(\theta)$ is more precise about $\gamma(\hat{X}(\theta), \theta)$ than $p(\theta')$ about $\gamma(\hat{X}(\theta'), \theta')$ under ϕ'. □

A.7. Proof of Theorem 4

Proof. By Assumption 6, there is a unique efficient stochastic choice rule P^E. Moreover, for any $x, x' \in \mathcal{X}$ and $\theta \in \Theta$ such that $p^E(x|\theta) > 0$ and $p^E(x'|\theta) > 0$, by the same variational arguments used in the Proof of Theorem 1, and exploiting linearity of the aggregator we have that:

$$\begin{aligned} & u(x', X(p^E(\theta)), \theta) - u(x, X(p^E(\theta)), \theta) \\ & \quad + [f(x') - f(x)] \int_{\mathcal{X}} u_X(\tilde{x}, X(p^E(\theta)), \theta) p^E(\tilde{x}|\theta) \, \mathrm{d}\tilde{x} \\ & = \lambda(X(p^E(\theta)), \theta) \left(\phi'(p^E(x'|\theta)) - \phi'(p^E(x|\theta)) \right) \\ & \quad + [f(x') - f(x)] \lambda_X(X(p^E(\theta)), \theta) \int_{\mathcal{X}} \phi(p^E(\tilde{x}|\theta)) \, \mathrm{d}\tilde{x} \end{aligned} \tag{93}$$

is necessary for optimality of p^E. Moreover, if the efficient stochastic choice rule obtains in equilibrium, we have that (by Equation (72)):

$$\begin{aligned} & u(x', X(p^E(\theta)), \theta) - u(x, X(p^E(\theta)), \theta) \\ & \quad = \lambda(X(p^E(\theta)), \theta) \left(\phi'(p^E(x'|\theta)) - \phi'(p^E(x|\theta)) \right) \end{aligned} \tag{94}$$

These conditions coincide if and only if:

$$\begin{aligned} & [f(x') - f(x)] \int_{\mathcal{X}} u_X(\tilde{x}, X(p^E(\theta)), \theta) p^E(\tilde{x}|\theta) \, \mathrm{d}\tilde{x} \\ & \quad = [f(x') - f(x)] \lambda_X(X(p^E(\theta)), \theta) \int_{\mathcal{X}} \phi(p^E(\tilde{x}|\theta)) \, \mathrm{d}\tilde{x} \end{aligned} \tag{95}$$

As f is nowhere-constant, $f(x') \neq f(x)$, and this condition reduces to:

$$\int_{\mathcal{X}} u_X(\tilde{x}, X(p^E(\theta)), \theta) p^E(\tilde{x}|\theta) \, \mathrm{d}\tilde{x} = \lambda_X(X(p^E(\theta)), \theta) \int_{\mathcal{X}} \phi(p^E(\tilde{x}|\theta)) \, \mathrm{d}\tilde{x} \tag{96}$$

Substituting in $p^E = p^*$, we obtain the statement in the claim. □

A.8. Statement and Proof of Lemma 8

In this appendix, we state and prove a Lemma that specializes several of main results to the case with quadratic payoffs of the form

$$u(x, X, \theta) = \alpha(X, \theta) - \beta(X, \theta)(x - \gamma(X, \theta))^2 \tag{97}$$

In the statement below, we use the definition $\tilde{\beta}(X, \theta) = \beta(X, \theta)/\lambda(X, \theta)$. We also define the bias and dispersion of a stochastic choice rule P in state θ around optimal point $\gamma(X(P), \theta)$ as

$$\begin{aligned} & \text{Bias}[P, \theta] \equiv \int_{\mathcal{X}} (x - \gamma(X(P), \theta)) \ \mathrm{d}P(x|\theta) \\ & \text{Disp}[P, \theta] \equiv \left(\int_{\mathcal{X}} (x - \gamma(X(P), \theta))^2 \ \mathrm{d}P(x|\theta) \right)^{\frac{1}{2}} \end{aligned} \tag{98}$$

Lemma 8. *Suppose that Assumptions 2 and 3 hold and that payoffs are given by Equation (97). The following properties hold under the additional stated conditions.*

1. ***Uniqueness.*** *There exists a unique equilibrium if the following holds for all $x \in \mathcal{X}$, $X \in \mathcal{X}$ and $\theta \in \Theta$:*

$$-(1-\gamma_X(X,\theta)) < \frac{\tilde{\beta}_X(X,\theta)}{\tilde{\beta}(X,\theta)}(x-\gamma(X,\theta)) < \gamma_X(X,\theta) \tag{99}$$

2. ***Monotone actions.*** *The cross-sectional distribution of actions and the aggregate action X are monotone in the fundamental if, in addition to the condition (99), the following holds for all $X \in \mathcal{X}$, $x \in \mathcal{X}$, and $\theta \in \Theta$*[18]*:*

$$\frac{\tilde{\beta}_\theta(X,\theta)}{\tilde{\beta}(X,\theta)}(x-\gamma(X,\theta)) < \gamma_\theta(X,\theta) \tag{100}$$

3. ***Monotone precision.*** *The precision of actions about the optimal action γ under ϕ' is decreasing (increasing) in the strength of fundamentals if, in addition to (99) and (100), $\tilde{\beta}$ is monotone decreasing (increasing) in both arguments.*
4. ***Efficiency.*** *A necessary condition for efficiency of the stochastic choice rule P^* under Assumption 6 is that, for all θ,*

$$\begin{aligned}&\lambda_X(X(P^*(\theta)),\theta)\int_{\mathcal{X}}\phi(p^*(x\mid\theta))\,\mathrm{d}x\\ =&\alpha_X(X(P^*(\theta)),\theta)-\beta_X(X(P^*(\theta)),\theta)(Disp[P^*(\theta),\theta])^2\\ &+2\,\gamma_X(X(P^*(\theta)),\theta)\,\beta(X(P^*(\theta)),\theta)\,Bias[P^*(\theta),\theta]\end{aligned} \tag{101}$$

Proof. We have directly assumed that Assumptions 2, 3 and 5 hold. The first claim follows so long as condition (99) implies Assumption 1, Supermodularity and Sufficient Concavity, for the outcome-equivalent game with payoff curvature $\tilde{\beta}$ and associated payoff $\tilde{u}$.

For supermodularity, it is sufficient to show that $\tilde{u}_{xX}(x,X,\theta) > 0$. We observe that $\tilde{u}_{xX} = -2\tilde{\beta}_X(X,\theta)(x-\gamma(X,\theta)) + 2\gamma_X(X,\theta)\tilde{\beta}(X,\theta)$ This condition simplifies to $\gamma_X(X,\theta) > \frac{\tilde{\beta}_X(X,\theta)}{\tilde{\beta}(X,\theta)}(x-\gamma(X,\theta))$, which is the second inequality of Equation (99).

For sufficient concavity, it is sufficient to show that $|\tilde{u}_{xx}(X,\theta)| > \tilde{u}_{xX}(x,X,\theta)$. Observe that $|\tilde{u}_{xx}(X,\theta)| = 2\tilde{\beta}(X,\theta)$. The condition

$$2\tilde{\beta}(X,\theta) > \tilde{u}_{xX} = -2\tilde{\beta}_X(X,\theta)(x-\gamma(X,\theta)) + 2\gamma_X(X,\theta)\tilde{\beta}(X,\theta) \tag{102}$$

simplifies to the first inequality of Equation (99): $-(1-\gamma_X(X,\theta)) < \frac{\tilde{\beta}_X(X,\theta)}{\tilde{\beta}(X,\theta)}(x-\gamma(X,\theta))$.

The second claim of the Lemma follows so long as condition (100) implies Assumption 4. To see this, as we have already that $\tilde{u}_{xX}(x,X,\theta) > 0$ for all x, X, θ, it is sufficient to check that $\tilde{u}_{x\theta}(x,X,\theta) > 0$ for all x, X, θ. We note that $\tilde{u}_{x\theta}(x,X,\theta) = -2\tilde{\beta}_\theta(X,\theta)(x-\gamma(X,\theta)) + 2\gamma_\theta(X,\theta)\tilde{\beta}(X,\theta)$ and re-arrange to the desired expression.

The third claim follows directly by Theorem 3 as the payoffs in Equation (97) satisfy Assumption 5.

The fourth claims follow by Theorem 4. Recall from Theorem 4 that a necessary condition for efficiency of an equilibrium P^* under Assumption 6 is that:

[18] Where, for simplicity, we allow β to be defined for all states in a closed interval that contains Θ, and assume it is differentiable in its second argument.

$$\int_{\mathcal{X}} u_X(\tilde{x}, X(P^*(\theta)), \theta)\, \mathrm{d}P^*(\tilde{x}|\theta) = \lambda_X(X, \theta) \int_{\mathcal{X}} \phi(p^*(x \mid \theta))\, \mathrm{d}x \tag{103}$$

for all $\theta \in \Theta$. Using the payoff function, we calculate:

$$u_X(x, X, \theta) = \alpha_X(X, \theta) - \beta_X(X, \theta)(x - \gamma(X, \theta))^2 + 2\gamma_X(X, \theta)\beta(X, \theta)(x - \gamma(X, \theta)) \tag{104}$$

Plugging this into the necessary condition and evaluating at the equilibrium aggregate $\hat{X}(\theta) = X(P^*(\theta))$, we obtain:

$$\begin{aligned} &\int_{\mathcal{X}} u_X(\tilde{x}, X(P^*(\theta)), \theta)\, \mathrm{d}P^*(\tilde{x}|\theta) \\ =& \int_{\mathcal{X}} \Big[\alpha_X(X(P^*(\theta)), \theta) - \beta_X(X(P^*(\theta)), \theta)(\tilde{x} - \gamma(X(P^*(\theta)), \theta))^2 \\ &+ 2\gamma_X(X(P^*(\theta)), \theta)\beta(X(P^*(\theta)), \theta)(\tilde{x} - \gamma(X(P^*(\theta)), \theta))\Big]\, \mathrm{d}P^*(\tilde{x}|\theta) \end{aligned} \tag{105}$$

Which can be rewritten in terms of the equilibrium bias and variance with respect to γ as:

$$\begin{aligned} &\int_{\mathcal{X}} u_X(\tilde{x}, X(P^*(\theta)), \theta)\, \mathrm{d}P^*(\tilde{x}|\theta) \\ =&\alpha_X(X(P^*(\theta)), \theta) - \beta_X(X(P^*(\theta)), \theta) \left(\mathrm{Disp}[P^*(\theta), \theta]\right)^2 \\ &+ 2\gamma_X(X(P^*(\theta)), \theta)\beta(X(P^*(\theta)), \theta)\mathrm{Bias}[P^*(\theta), \theta] \end{aligned} \tag{106}$$

as desired. □

A.9. *Proof of Corollary 3*

We first derive the payoff representation of Equation (32). This is a second-order approximation of the payoff function in Equation (31), reprinted here:

$$u(p_i, P, M) = M^{\frac{1-\sigma}{\sigma}} P^{\eta - \frac{1}{\sigma}} \left(p_i - M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}}\right) p_i^{-\eta} \tag{107}$$

We first calculate

$$\begin{aligned} u_p(p_i, P, M) &= M^{\frac{1-\sigma}{\sigma}} P^{\eta - \frac{1}{\sigma}} \left((-\eta + 1)p_i^{-\eta} + \eta M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}} p_i^{-\eta-1}\right) \\ u_{pp}(p_i, P, M) &= M^{\frac{1-\sigma}{\sigma}} P^{\eta - \frac{1}{\sigma}} \left(\eta(\eta - 1)p_i^{-\eta-1} - \eta(\eta + 1)M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}} p_i^{-\eta-2}\right) \end{aligned} \tag{108}$$

We define $\gamma(P, M)$ as the (unique) solution to $u_p(p_i, P, M)|_{p_i=\gamma(P,M)} = 0$. Re-arranging:

$$\begin{aligned} 0 &= M^{\frac{1-\sigma}{\sigma}} P^{\eta - \frac{1}{\sigma}} \left((-\eta + 1)\gamma(P, M)^{-\eta} + \eta M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}} \gamma(P, M)^{-\eta-1}\right) \\ 0 &= \left((-\eta + 1) + \eta M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}} \gamma(P, M)^{-1}\right) \\ \gamma(P, M) &= \frac{\eta}{\eta - 1} M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}} \end{aligned} \tag{109}$$

We define $\alpha(P, M) = u(\gamma(P, M), P, M)$. We first observe that

$$\gamma(P, M) - M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}} = \frac{1}{\eta - 1} M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}} \tag{110}$$

Then, by direct calculation,

$$\begin{aligned}\alpha(P, M) &= M^{\frac{1-\sigma}{\sigma}} P^{\eta-\frac{1}{\sigma}} \left(\frac{1}{\eta-1} M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}}\right)\left(\frac{\eta}{\eta-1}\right)^{-\eta} M^{-\eta\frac{\chi}{\sigma}} P^{-\eta+\eta\frac{\chi}{\sigma}} \\ &= \frac{1}{\eta-1}\left(\frac{\eta}{\eta-1}\right)^{-\eta} M^{\frac{1-\sigma+\chi(1-\eta)}{\sigma}} P^{\eta-\frac{1}{\sigma}+(1-\eta)(1-\frac{\chi}{\sigma})}\end{aligned} \tag{111}$$

as desired

We define $\beta(P, M) = -\frac{1}{2} u_{pp}(p_i, P, M)|_{p=\gamma(P,M)}$. We first observe that

$$(\eta - 1)\gamma(P, M) - (\eta + 1)M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}} = -M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}} \tag{112}$$

Then, by direct calculation,

$$\begin{aligned}\beta(P, M) &= -\frac{1}{2}\left(M^{\frac{1-\sigma}{\sigma}} P^{\eta-\frac{1}{\sigma}} \eta\gamma(P, M)^{-\eta-2}\left((\eta-1)\gamma(M, P) - (\eta+1)M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}}\right)\right) \\ &= \frac{1}{2}\left(M^{\frac{1-\sigma}{\sigma}} P^{\eta-\frac{1}{\sigma}} \eta\gamma(P, M)^{-\eta-2} M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}}\right) \\ &= \frac{1}{2}\left(M^{\frac{1-\sigma}{\sigma}} P^{\eta-\frac{1}{\sigma}} \eta\left(\frac{\eta}{\eta-1} M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}}\right)^{-\eta-2} M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}}\right) \\ &= \frac{\eta}{2}\left(\frac{\eta}{\eta-1}\right)^{-(\eta+2)} M^{\frac{1-\sigma-\chi(\eta+1)}{\sigma}} P^{-1-\frac{1}{\sigma}+(\eta+1)\frac{\chi}{\sigma}}\end{aligned} \tag{113}$$

as required. We finally observe that, since $\lambda(M) = \frac{1}{\pi(M)} = KM^{-\delta}$ (where K is a normalizing constant), we have

$$\tilde{\beta}(P, M) = \frac{\beta(P, M)}{\lambda(P, M)} = \frac{\eta}{2K}\left(\frac{\eta}{\eta-1}\right)^{-(\eta+2)} M^{\frac{1-\sigma-\chi(\eta+1)}{\sigma}+\delta} P^{-1-\frac{1}{\sigma}+(\eta+1)\frac{\chi}{\sigma}} \tag{114}$$

We now apply the conditions of Lemma 8 to prove the stated result. We first calculate that

$$\gamma_P(P, M) = \left(1 - \frac{\chi}{\sigma}\right)\frac{\gamma(P, M)}{P} \qquad \frac{\tilde{\beta}_P(P, M)}{\tilde{\beta}(P, M)} = P^{-1}\left(-1 - \frac{1}{\sigma} + (\eta+1)\frac{\chi}{\sigma}\right) \tag{115}$$

Applying the condition of Equation (99), we get

$$\begin{aligned}-\left(1 - \left(1 - \frac{\chi}{\sigma}\right)\frac{\gamma(P, M)}{P}\right) &< P^{-1}\left(-1 - \frac{1}{\sigma} + (\eta+1)\frac{\chi}{\sigma}\right)(p - \gamma(P, M)) \\ &< \left(1 - \frac{\chi}{\sigma}\right)\frac{\gamma(P, M)}{P}\end{aligned} \tag{116}$$

Since $\chi/\sigma < 1$, we divide all three expressions by $\left(1 - \frac{\chi}{\sigma}\right)\gamma(P, M)/P$ to get

$$-\left(\frac{P}{\gamma(P, M)\left(1 - \frac{\chi}{\sigma}\right)} - 1\right) < \frac{-1 - \frac{1}{\sigma} + (\eta+1)\frac{\chi}{\sigma}}{\left(1 - \frac{\chi}{\sigma}\right)}\left(\frac{p}{\gamma(P, M)} - 1\right) < 1 \tag{117}$$

as desired.

We next verify a condition for monotone aggregates. We first calculate:

$$\gamma_M(P, M) = \frac{\chi}{\sigma}\frac{\gamma(P, M)}{M}, \qquad \frac{\tilde{\beta}_M(P, M)}{\tilde{\beta}(P, M)} = M^{-1}\left(\frac{1-\sigma-\chi(\eta+1)}{\sigma} + \delta\right) \tag{118}$$

We then apply Equation (100):

$$M^{-1}\left(\frac{1-\sigma-\chi(\eta+1)}{\sigma} + \delta\right)(p - \gamma(P, M)) < \frac{\chi}{\sigma}\frac{\gamma(P, M)}{M} \tag{119}$$

Dividing both sides by $\gamma(P, M)\chi/(M\sigma)$, this becomes

$$\frac{1-\sigma-\chi(\eta+1)+\delta\sigma}{\chi}\left(\frac{p}{\gamma(P, M)} - 1\right) < 1 \tag{120}$$

We finally derive a condition for monotone precision. For this, we need $\tilde{\beta}$ to decrease in both M and P. This respectively requires:

$$\begin{aligned} 0 &> \frac{1-\sigma-\chi(\eta+1)}{\sigma} + \delta \\ 0 &> -1 - \frac{1}{\sigma} + (\eta+1)\frac{\chi}{\sigma} \end{aligned} \tag{121}$$

Re-arranging these inequalities gives the desired condition,

$$\chi(\eta+1) \in (1+\sigma(\delta-1), 1+\sigma) \tag{122}$$

A.10. *Proof of Corollary 4*

Proof. We first derive the payoff representation of Equations (41) and (42). To this end, we begin by deriving the consumer's choices at $t \geq 1$. At $t = 1$, given savings b_{i0} from the first period, each consumer i solves the following program at $t = 1$:

$$\begin{aligned} &\max_{\{c_{it}, n_{it}\}_{t=1}^{\infty}} \sum_{t=1}^{\infty} \delta^t \left(c_{it} - \frac{c_{it}^2}{2} - \chi\frac{n_{it}^2}{2}\right) \\ &\sum_{t=1}^{\infty}\frac{c_{it}}{R^t} \leq b_{i0} + \sum_{t=1}^{\infty}\frac{w_t n_{it}}{R^t} \end{aligned} \tag{123}$$

where $b_{i0} = y_0 - c_{i0}$ is the agent's savings from $t = 0$. This problem is concave in all arguments. Letting κ denote the Lagrange multiplier on the constraint, we find first-order conditions $\delta^t(1 - c_{it}) = \kappa R^{-t}$ for each c_{it} and $\delta^t \chi n_{it} = w_t \kappa R^{-t}$ for each n_{it}. Using $\delta R = 1$, we transform the former into $\kappa = 1 - c_{it}$ for all t. This implies that consumption is constant. Plugging this into the labor-supply condition, we derive $n_{it} = \frac{1}{\chi} w_t(1 - c_{it})$. This is also constant, if consumption is constant.

We next prove that output is identically equal to $y_t = \bar{y}$ for $t \geq 1$ and solve for $\bar{y}$. Profit maximization for the firm implies that the firm elastically demands labor at the wage $w_t = 1$. Evaluated at this wage, labor supply for each agent i is $n_{it} = \frac{1}{\chi}(1 - c_{it})$. Integrating both sides over i, we get $n_t = \frac{1}{\chi}(1 - c_t)$. Substituting in the production function and market clearing, this becomes $y_t = \frac{1}{\chi}(1 - y_t)$. Therefore, $y_t = \bar{y} = \frac{1}{1+\chi}$.

To derive the household's consumption and labor supply, we return to the budget constraint and simplify it by plugging in constant consumption $c_{it} = c_{i1}$, labor demand, and $w_t = 1$, and by simplifying the sums:

$$\frac{1}{1-R^{-1}} c_{i1} \leq R b_{i0} + \frac{1}{1-R^{-1}} \frac{1}{\chi}(1 - c_{i1}) \tag{124}$$

Rearranging, we write

$$c_{i1} \leq \frac{\chi}{1+\chi} \frac{1-\delta}{\delta} b_{i0} + \bar{y} \tag{125}$$

This holds at equality if the right-hand-side is less than 1 (the agent's bliss point). This is guaranteed under the maintained assumption that $b_{i0} \leq \overline{c} < \delta/(1-\delta)$. We finally write the value function from Equation (123) as $V(b_{i0})$. And we observe from the envelope theorem that

$$V'(b_{i0}) = \kappa = 1 - \frac{\chi}{1+\chi} \frac{1-\delta}{\delta} b_{i0} - \bar{y} \tag{126}$$

We now return to the payoff of the consumer at time 0, who chooses consumption given rational expectations about this future equilibrium path and their future choices. For notational simplicity, we let $c_{i0} = c$ and $y_0 = y$. The agent's payoff is

$$U(c, y, \theta_d) = (1+\theta_d)c - \frac{c^2}{2} - \chi \frac{y^2}{2} + V(y - c) \tag{127}$$

Note that all agents are off their labor supply curve and work y labor hours.

We now derive the form in Equation (41). We first observe that $U_c(c, y, \theta_d)|_{c=\gamma(y,\theta_d)} = 0$. Taking the first derivative,

$$\begin{aligned} U_c(c, y, \theta_d) &= (1+\theta_d) - c - V'(y-c) \\ &= (1+\theta_d) - c - \left(1 - \frac{\chi}{1+\chi} \frac{1-\delta}{\delta}(y-c) - \bar{y}\right) \end{aligned} \tag{128}$$

We next use $U_c(c, y, \theta_d)|_{c=\gamma(y,\theta_d)} = 0$ and rearrange to write

$$\gamma(y, \theta_d) = (1-m)(\theta_d + \bar{y}) + my \tag{129}$$

where $m = \frac{\chi(1-\delta)}{\chi+\delta}$ is the marginal propensity to consume.

We next observe that $-U_{cc}(c, y, \theta_d) = 2\beta(y, \theta_d)$. We calculate, from above,

$$U_{cc} = -\frac{1}{1-m}, \qquad \beta(y, \theta_d) = \frac{1}{2(1-m)} \tag{130}$$

Moreover, we have $\tilde{\beta}(y, \theta_d) = \beta(y, \theta_d)/\lambda(y, \theta_d) = \frac{1}{2(1-m)} y^{\tau}$.

We finally observe that $U(c, y, \theta_d)|_{c=\gamma(y,\theta_d)} = \alpha(y, \theta_d)$. We therefore define

$$\alpha(y, \theta_d) = (1+\theta_d)\gamma(y, \theta_d) - \frac{\gamma(y, \theta_d)^2}{2} - \chi \frac{y^2}{2} + V(y - \gamma(y, \theta_d)) \tag{131}$$

Having verified the payoff representation, we now prove the claim by applying Lemma 8. First, to show uniqueness, we specialize the condition in Equation (99). We note that

$$\begin{aligned} \gamma_y(y, \theta_d) &= m \\ \frac{\tilde{\beta}_y(y, \theta_d)}{\tilde{\beta}(y, \theta_d)} &= \tau y^{-1} \end{aligned} \tag{132}$$

Using these expressions, we derive the condition that, for all $y, c \in [\underline{c}, \overline{c}]$ and $\theta_d \in \Theta_d$,

$$-(1-m) < \frac{\tau}{y}\left(c - (1-m)\left(\bar{y}+\theta_d\right) - my\right) < m \tag{133}$$

We re-arrange this algebraically to

$$0 < m - \frac{\tau}{y}\left(c - (1-m)\left(\bar{y}+\theta_d\right) - my\right) < 1 \tag{134}$$

Next, to show monotonicity, we observe that $\tilde{\beta}_{\theta_d}(y, \theta_d) = 0$, and hence Equation (100) reduces to $\gamma_{\theta_d}(y, \theta_d) > 0$, which is by assumption when $\delta > 0$ and $\chi > 0$.

Next, to show monotone precision, we observe that $\tilde{\beta}(y, \theta_d) = \frac{y^\tau}{\delta}$ is monotone increasing in X, which from Lemma 8 implies that precision in increasing in fundamentals.

Finally, to show efficiency, we plug directly into Equation (101). We first use the definition of α and the envelope theorem to observe that

$$\begin{aligned}\alpha_y(y, \theta_d) &= U_y(c, y, \theta_d)|_{c=\gamma(y,\theta_d)} + \gamma_c(y, \theta_d) U_c(c, y, \theta_d)|_{c=\gamma(y,\theta_d)} \\ &= U_y(c, y, \theta_d)|_{c=\gamma(y,\theta_d)}\end{aligned} \tag{135}$$

We then calculate

$$\begin{aligned}U_y(c, y, \theta_d) &= -\chi y + V'(y-c) \\ &= -\chi y + 1 - \frac{\chi}{1+\chi}\frac{1-\delta}{\delta}(y-c) - \bar{y} \\ &= -\chi y + 1 - \frac{m}{1-m}(y-c) - \bar{y}\end{aligned} \tag{136}$$

where in the last line we use the definition of m.

We next observe that

$$\lambda_y(y, \theta_d) = -\tau y^{-\tau-1} \qquad \gamma_y(y, \theta_d) = m \qquad \beta(y, \theta_d) = \frac{1}{2(1-m)} \tag{137}$$

Finally, because of linear aggregation,

$$\text{Bias}[P, \theta_d] = \int_{\mathcal{X}} \left(x - \gamma(y(P), \theta_d)\right) \, \mathrm{d}P\left(x|\theta_d\right) = y - \gamma(y, \theta_d) \tag{138}$$

Using all of this, we re-write the condition for efficiency (Equation (101)) as

$$\begin{aligned}-\tau y^{-\tau-1}\int_{\mathcal{X}} \phi(p^*(x \mid \theta_d))\,\mathrm{d}x &= -\chi y + 1 - \frac{m}{1-m}(y - \gamma(y, \theta_d)) - \bar{y} \\ &\quad + \frac{m}{1-m}\left(y - \gamma(y, \theta_d)\right) \\ &= -\chi y + 1 - \bar{y}\end{aligned} \tag{139}$$

This re-arranges to

$$y = \bar{y} + \frac{\tau}{\chi} y^{-\tau-1} \int_{\mathcal{X}} \phi(p^*(x \mid \theta_d))\,\mathrm{d}x \tag{140}$$

When $\tau = 0$, the solution to the fixed-point equation is $y = \bar{y}$. When $\tau > 0$, then

$$y - \bar{y} = \frac{\tau}{\chi} y^{-\tau-1} \int_{\mathcal{X}} \phi(p^*(x \mid \theta_d))\, \mathrm{d}x \tag{141}$$

The right-hand side is weakly positive under the assumption that cognitive costs in each state are positive. Thus, a necessary condition for efficiency is that $y > \bar{y}$. □

Online Appendix. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jet.2023.105704.

References

Allen, R., Rehbeck, J., 2021. A generalization of quantal response equilibrium via perturbed utility. Games 12 (1), 20.
Alvarez, F., Beraja, M., Gonzalez-Rozada, M., Neumeyer, P.A., 2019. From hyperinflation to stable prices: Argentina's evidence on menu cost models. Q. J. Econ. 134 (1), 451–505.
Alves, F., Kaplan, G., Moll, B., Violante, G.L., 2020. A further look at the propagation of monetary policy shocks in HANK. J. Money Credit Bank. 52 (S2), 521–559.
Angeletos, G.-M., La'O, J., 2010. Noisy business cycles. NBER Macroecon. Annu. 24 (1), 319–378.
Angeletos, G.-M., La'O, J., 2013. Sentiments. Econometrica 81 (2), 739–779.
Angeletos, G.-M., Lian, C., 2016. Incomplete information in macroeconomics: accommodating frictions in coordination. In: Handbook of Macroeconomics, vol. 2. Elsevier, pp. 1065–1240.
Angeletos, G.-M., Pavan, A., 2007. Efficient use of information and social value of information. Econometrica 75 (4), 1103–1142.
Angeletos, G.-M., Sastry, K.A., 2023. Inattentive economies. Working paper. Northwestern University.
Benhabib, J., Wang, P., Wen, Y., 2015. Sentiments and aggregate demand fluctuations. Econometrica 83 (2), 549–585.
Bergemann, D., Morris, S., 2013. Robust predictions in games with incomplete information. Econometrica 81 (4), 1251–1308.
Berger, D.W., Bocola, L., Dovis, A., 2023. Imperfect risk-sharing and the business cycle. Q. J. Econ. 138 (3), 1765–1815. https://doi.org/10.1093/qje/qjad013.
Berry, S.T., Haile, P.A., 2021. Foundations of demand estimation. In: Handbook of Industrial Organization, vol. 4. Elsevier, pp. 1–62.
Brunnermeier, M.K., Correia, S., Luck, S., Verner, E., Zimmermann, T., 2022. The debt-inflation channel of the German hyperinflation. Working Paper 4303537, SSRN.
Caplin, A., Dean, M., 2015. Revealed preference, rational inattention, and costly information acquisition. Am. Econ. Rev. 105 (7), 2183–2203.
Caplin, A., Dean, M., Leahy, J., 2019. Rational inattention, optimal consideration sets, and stochastic choice. Rev. Econ. Stud. 86 (3), 1061–1094.
Caplin, A., Dean, M., Leahy, J., 2022. Rationally inattentive behavior: characterizing and generalizing Shannon entropy. J. Polit. Econ. 130 (6), 1676–1715.
Caplin, A., Martin, D., 2015. A testable theory of imperfect perception. Econ. J. 125 (582), 184–202.
Cerreia-Vioglio, S., Corrao, R., Lanzani, G., 2020. Robust opinion aggregation and its dynamics. Working paper, IGIER. Università Bocconi.
Chahrour, R., Ulbricht, R., 2023. Robust predictions for DSGE models with incomplete information. Am. Econ. J. Macroecon. 15 (1), 173–208.
Costain, J., Nakov, A., 2019. Logit price dynamics. J. Money Credit Bank. 51 (1), 43–78.
Dean, M., Neligh, N.L., 2022. Experimental tests of rational inattention. Working paper. Columbia University.
Denti, T., 2023. Unrestricted information acquisition. Theor. Econ. 18, 1101–1140.
Flynn, J.P., Sastry, K.A., 2022. Attention cycles. Working Paper 3592107, SSRN.
Fosgerau, M., Melo, E., De Palma, A., Shum, M., 2020. Discrete choice and rational inattention: a general equivalence result. Int. Econ. Rev. 61 (4), 1569–1589.
Frankel, D.M., Morris, S., Pauzner, A., 2003. Equilibrium selection in global games with strategic complementarities. J. Econ. Theory 108 (1), 1–44.
Frasquilho, D., Matos, M.G., Salonna, F., Guerreiro, D., Storti, C.C., Gaspar, T., Caldas-de Almeida, J.M., 2015. Mental health outcomes in times of economic recession: a systematic literature review. BMC Public Health 16 (1), 1–40.

Fudenberg, D., Iijima, R., Strzalecki, T., 2015. Stochastic choice and revealed perturbed utility. Econometrica 83 (6), 2371–2409.
Goeree, J., Palfrey, T.R., Holt, C.A., 2016. Quantal Response Equilibrium: A Stochastic Theory of Games. Princeton University Press.
Hall, R.E., 1978. Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence. J. Polit. Econ. 86 (6), 971–987.
Harsanyi, J.C., 1973. Oddness of the number of equilibrium points: a new proof. Int. J. Game Theory 2 (1), 235–250.
Hauser, J.R., Wernerfelt, B., 1990. An evaluation cost model of consideration sets. J. Consum. Res. 16 (4), 393–408.
Hébert, B.M., La'O, J., 2022. Information acquisition, efficiency, and non-fundamental volatility. Technical report. National Bureau of Economic Research.
Hébert, B.M., Woodford, M., 2020. Neighborhood-based information costs. Technical report. National Bureau of Economic Research.
Hellwig, C., Venkateswaran, V., 2009. Setting the right prices for the wrong reasons. J. Monet. Econ. 56, S57–S77.
Hoffmann, S., Walker, S.P., 2020. Adapting to crisis: accounting information systems during the Weimar hyperinflation. Bus. Hist. Rev. 94 (3), 593–625.
Huo, Z., Pedroni, M., 2020. A single-judge solution to beauty contests. Am. Econ. Rev. 110 (2), 526–568.
Jensen, M.K., 2018. Aggregative games. In: Handbook of Game Theory and Industrial Organization, vol. I. Edward Elgar Publishing.
Jung, J., Kim, J.H.J., Matějka, F., Sims, C.A., 2019. Discrete actions in information-constrained decision problems. Rev. Econ. Stud. 86 (6), 2643–2667.
Macaulay, A., 2020. Cyclical attention to saving. Working paper. Oxford University.
Maćkowiak, B., Wiederholt, M., 2009. Optimal sticky prices under rational inattention. Am. Econ. Rev. 99 (3), 769–803.
Maćkowiak, B., Wiederholt, M., 2018. Lack of preparation for rare events. J. Monet. Econ. 100, 35–47.
Mani, A., Mullainathan, S., Shafir, E., Zhao, J., 2013. Poverty impedes cognitive function. Science 341 (6149), 976–980.
Matějka, F., 2015. Rationally inattentive seller: sales and discrete pricing. Rev. Econ. Stud. 83 (3), 1125–1155.
Matějka, F., McKay, A., 2015. Rational inattention to discrete choices: a new foundation for the multinomial logit model. Am. Econ. Rev. 105 (1), 272–298.
Mathevet, L., 2010. A contraction principle for finite global games. Econ. Theory 42 (3), 539.
Mattsson, L.-G., Weibull, J.W., 2002. Probabilistic choice and procedurally bounded rationality. Games Econ. Behav. 41 (1), 61–78.
McFadden, D.L., 1973. Conditional logit analysis of qualitative choice behavior. In: Zarembka, P. (Ed.), Frontiers in Econometrics. Academic Press, New York, pp. 105–142.
McKelvey, R.D., Palfrey, T.R., 1995. Quantal response equilibria for normal form games. Games Econ. Behav. 10 (1), 6–38.
Melo, E., 2022. On the uniqueness of quantal response equilibria and its application to network games. Econ. Theory 74 (3), 681–725.
Meyer, B.D., Sullivan, J.X., 2011. Consumption and income poverty over the business cycle. Res. Labor Econ. 32, 51–82.
Milgrom, P., Roberts, J., 1990. Rationalizability, learning, and equilibrium in games with strategic complementarities. Econometrica, J. Econ. Soc., 1255–1277.
Morris, S., Shin, H.S., 1998. Unique equilibrium in a model of self-fulfilling currency attacks. Am. Econ. Rev., 587–597.
Morris, S., Shin, H.S., 2002. Social value of public information. Am. Econ. Rev. 92 (5), 1521–1534.
Morris, S., Yang, M., 2022. Coordination and continuous stochastic choice. Rev. Econ. Stud. 89 (5), 2687–2722.
Mullainathan, S., Shafir, E., 2013. Scarcity: Why Having Too Little Means so Much. Macmillan.
Nakamura, E., Steinsson, J., Sun, P., Villar, D., 2018. The elusive costs of inflation: price dispersion during the US great inflation. Q. J. Econ. 133 (4), 1933–1980.
Pomatto, L., Strack, P., Tamuz, O., 2023. The cost of information: the case of constant marginal costs. Am. Econ. Rev. 113 (5), 1360–1393.
Rosenthal, R.W., 1989. A bounded-rationality approach to the study of noncooperative games. Int. J. Game Theory 18 (3), 273–292.
Sergeyev, D., Lian, C., Gorodnichenko, Y., 2022. The economics of financial stress. Working paper. UC Berkeley.
Sims, C.A., 2003. Implications of rational inattention. J. Monet. Econ. 50 (3), 665–690.
Stahl, D.O., 1990. Entropy control costs and entropic equilibria. Int. J. Game Theory 19 (2), 129–138.
Stevens, L., 2019. Coarse pricing policies. Rev. Econ. Stud. 87 (1), 420–453.
Stokey, N.L., Lucas, R.E., Prescott, E.C., 1989. Recursive Methods in Economic Dynamics. Harvard University Press.
Van Damme, E., 1991. Stability and Perfection of Nash Equilibria, vol. 339. Springer.
Van Zandt, T., Vives, X., 2007. Monotone equilibria in Bayesian games of strategic complementarities. J. Econ. Theory 134 (1), 339–360.

Weinstein, J., Yildiz, M., 2007. Impact of higher-order uncertainty. Games Econ. Behav. 60 (1), 200–212.
Woodford, M., 2003. Imperfect common knowledge and the effects of monetary policy. In: Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps.
Woodford, M., 2012. Inattentive valuation and reference-dependent choice. Working paper. Columbia University.
Woodford, M., 2020. Modeling imprecision in perception, valuation, and choice. Annu. Rev. Econ. 12, 579–601.
Yang, M., 2015. Coordination with flexible information acquisition. J. Econ. Theory 158, 721–738.