Games and Economic Behavior 153 (2025) 408–429

Contents lists available at ScienceDirect

Games and Economic Behavior

journal homepage: www.elsevier.com/locate/geb

Proxy variables and feedback effects in decision making ☆

Alexander Clyde

University College London, United Kingdom

ARTICLE INFO

JEL classification:
D91

Keywords:
Proxy variables
Bounded rationality
Expectations
Misspecified models

ABSTRACT

When using data, an analyst often only has access to proxies of the true variables. I propose a framework that models decision makers who naively assume potentially noisy proxy variables are perfect measurements. Due to feedback from choices into data, a notion of equilibrium is required to close the model. I illustrate the concept with applications to policing/crime and market entry. In these applications, we see that very small imperfections in the proxy variable can lead to large distortions in beliefs. I show that the set of strategies that can arise as equilibria with arbitrarily close to perfect measurement coincides with a version of Self-Confirming Equilibrium.

1. Introduction

Analysis of quantitative data to inform decisions is increasingly important to organizations and firms. However, data used in economic decision making are often imperfect measurements or proxies of the underlying variables of interest. Examples of proxy variables that play an important role in driving allocation of economic resources abound. GDP per capita is used as a proxy for living standards,[1] and guides entrepreneurs and traders in assessing the relative economic vitality of countries in which they are considering investment. Academic institutions use various citation metrics as proxies for academic impact.

Whilst a growing literature studies the impact of economic decision makers having a misspecified model of their environment,[2] this work generally considers decision makers who interpret perfectly measured data in an incorrect way. In this paper, I propose a framework where decision makers have a correct model of the world, but naively use potentially mismeasured proxy variables to form beliefs. The decision makers (henceforth DMs) treat the proxy variables as if they were *exactly identical* to the true variables that affect their utility.

The structure of the DMs' problem is as follows. First they draw the realization of signal variables s. They then choose an action variable x, and both the action and the signals then affect the realization of an outcome variable y. There is a true joint distribution over the variables reflecting these causal relationships.

$$p(y, x, s) = p(y|x, s)p(x|s)p(s) \tag{1}$$

The proxy variables $(y^\bullet, x^\bullet, s^\bullet)$ are drawn from a distribution π that depends on (y, x, s), resulting in the following joint distribution over the proxies.

☆ Financial support for this project has come from an Economic and Social Research Council studentship (1923039) and European Research Council grant no. 692995, and is gratefully acknowledged. I thank my advisors, Ran Spiegler and Philippe Jehiel, for their support and guidance, the Theory group at UCL and audiences at the Royal Economic Society conference, an ENTER network student seminar at Tilburg University and the Econometric Society Summer Meeting in Barcelona.

E-mail address: alexander.clyde.econ@gmail.com.

[1] See Coyle (2015) for an outline of various arguments in the debate over what GDP measures and its quality as a proxy.

[2] See Bohren and Hauser (2024b) for a survey.

https://doi.org/10.1016/j.geb.2025.07.008
Received 17 June 2024

Available online 31 July 2025

$$p_\pi(y^\bullet, x^\bullet, s^\bullet) = \sum_{(y,x,s)\in Y\times X\times S} \pi(y^\bullet, x^\bullet, s^\bullet | y, x, s) p(y, x, s) \tag{2}$$

The DM has a vNM utility function defined over the true variables: $u(y, x, s)$. Therefore they need to form beliefs about the conditional distribution $y|x, s$. The DM forms these beliefs by taking the proxies at face value

$$p_\pi(y^\bullet = y | x^\bullet = x, s^\bullet = s) = \frac{p_\pi(y^\bullet = y, x^\bullet = x, s^\bullet = s)}{p_\pi(x^\bullet = x, s^\bullet = s)} \tag{3}$$

If the proxies were perfect measurements this would recover the true conditional distribution $p(y|x, s)$, but with mismeasured proxies the beliefs formed according to this procedure are distorted. As is often the case in the misspecified models literature, with mismeasured proxies the beliefs of the DM can endogenously depend on the choices the DM makes. I therefore define a notion of *Proxy Equilibrium* that ensures consistency between these choices and the proxy distribution used to form beliefs.

A preview of the first application in the paper can illustrate why these feedback effects are of interest. Suppose the decision maker is a municipal policymaker who must decide on the number of police officers to employ. Before deciding, they first learn the realization of a signal variable s that affects the cost of crime but not the level of crime. The police numbers x they choose then affect the crime level y. Assume that the proxy variable for police numbers exhibits classical measurement error $x^\bullet = x + \epsilon$ with $\epsilon \sim \mathcal{N}(0, Var(\epsilon))$, while the proxy for crime is a perfect measurement $y^\bullet = y$.[3] Suppose that the true relationship between crime and police numbers is linear: $y = -\beta x$ for some $\beta > 0$.

Under this form of classical measurement error with a linear Gaussian model, the expected level of crime conditional on police numbers estimated using the proxies is attenuated towards zero according to the following formula[4]

$$\mathbb{E}[y^\bullet | x^\bullet = x] = -\beta \frac{Var(x)}{Var(x) + Var(\epsilon)} x \tag{4}$$

The dependence of this formula on true variance in the number of police officers $Var(x)$ illustrates the endogenous effects of the DMs choices on their beliefs. When $Var(x)$ is small, there can be large attenuation bias even when the measurement error problem is also small and $Var(\epsilon)$ is close to zero. If $Var(x)$ were fixed, then a small measurement problem would mean the extent of attenuation bias is also small.

In this policing application the neglect of measurement error results in overly rigid policy. Due to the attenuation bias in the measured relationship between policing and crime, policymakers vary police numbers with the cost of crime less than they would if they knew the true relationship. This rigidity can be extreme: there exists a Proxy Equilibrium where municipalities do not vary police numbers at all. This is true regardless of how strong the true relationship between crime and policing is, or how close police numbers are to being perfectly measured. Feedback effects result in a multiplicity of equilibria and a stark discontinuity between the extent of imperfection in proxy variables and the extent of the bias in the beliefs of the DMs. These features would not be apparent in a non-equilibrium analysis of attenuation bias.

In the second application, noisy measurement and equilibrium selection effects result in endogenous overoptimism and thus over-entry by firms deciding whether to enter a market. Moreover, the impact of changes in measurement noise can differ in important ways for pivotal firms who are on the margin between different choices and firms who are not. Due to the impact on the pivotal firm who must be indifferent between entering or not, more proxy 'noise' results in a greater extent of excessive entry. Without this equilibrium feedback effect, the effect of noise on entry is ambiguous. This is in contrast to other work on the behavioral bias caused by selection effects, such as Jehiel (2018) and Esponda and Pouzo (2017), in which more noise has an ambiguous effect on entry both in and out of equilibrium.

I build on the insights in these applications and give a characterization of all strategies that could arise as Proxy Equilibria when the proxy variables are arbitrarily close to being perfect measurements. I show that a strategy can be implemented as a Proxy Equilibrium for some proxy mapping that is arbitrarily close to perfect measurement if and only if it can be implemented as a version of Self-Confirming Equilibrium that I call Self-Confirming Optimal.

The characterization result clarifies a theme arising in the two applications; that small measurement problems can have a large effect on beliefs when the equilibrium strategy only puts weight on particular actions. The result also demonstrates the differences between Proxy Equilibrium and Self-Confirming Optimality. For general noise, a Proxy Equilibrium is not necessarily Self-Confirming Optimal as it allows beliefs that can differ from rational expectations for any action-signal combination. For particular almost-perfectly measured noise structures the set of Proxy Equilibrium can significantly refine the set of Self-Confirming Optimal strategies.

Proxy Equilibrium draws a distinction between the fact that the DM knows the realization of the signals and the action they have chosen but does not know how these variables covary with the outcomes. The story I have in mind is that Proxy Equilibrium is the long run steady state of some learning process. The learning process does not feature a long lived agent who repeats the same decision problem enough times to generate an asymptotic sample, but instead a sequence of short-lived agent who have to rely on a large public dataset of potentially mismeasured proxies generated by DMs in similar situations. In the policing application we can imagine a sequence of short lived municipal leaders. The data generated from each municipal leader's tenure is too sparse to apply the law of large numbers, so they have to draw inference from a national dataset designed for social scientists researching crime.

[3] The DM is assumed to lack an informative proxy for the cost of crime s in this application.

[4] See Chapter 3 of Carroll et al. (2006) for a detailed review of this kind of measurement error.

The main contribution of the paper is to develop an equilibrium framework in which we can model economic decision makers who use proxy variables. The applications in the paper show that taking into account how calculated decision making feeds back into data can raise issues that would not be apparent from a purely statistical analysis of the use of proxy variables. The paper also contributes to the literature on solution concepts with boundedly rational expectations. The concept generally does not fall neatly into others in the literature, and I explore these connections in Section 6.

2. Modeling set up

The space of variables V can be divided into a set of m signals $S \subseteq \mathbb{R}^m$, a set of actions $X \subseteq \mathbb{R}$ and a set of outcomes $Y \subseteq \mathbb{R}$. The variable space can thus be described as the product $V \equiv Y \times X \times S$. Assume the variable space is finite. In Appendix A.1 I show how the definition of Proxy Equilibrium can be extended to allow for general non-finite variable spaces.[5]

The DM learns the realization of the signals $s \in S$, before choosing an action $x \in X$ resulting in a distribution over the outcome variable $y \in Y$. The payoff of the DM is defined over the signal, action and outcome variables by utility function $u : Y \times X \times S \to \mathbb{R}$. The DM wants to form the conditional distribution $p(y|x, s)$. Doing this, they can calculate their objective expected utility

$$U(x,s) = \sum_{y \in Y} u(y,x,s)p(y|x,s) \tag{5}$$

Throughout the paper we treat choosing actions to maximize this expression as the normative benchmark. Let $\sigma : S \to \Delta(X)$ denote a *strategy mapping*. Assume the signal distribution is full support; $p(s) > 0$ for all $s \in S$. The joint distribution over outcomes, actions and signals can then be written as follows.

$$p(y,x,s;\sigma) = p(y|x,s)\sigma(x|s)p(s) \tag{6}$$

2.1. Proxy variables

In order to form beliefs about the conditional distribution $y|x, s$, the DM needs to have information on the joint distribution of (y, x, s). We assume that the DM can only access the joint distribution over proxies for these variables. Each of the variables has a respective proxy that can take any of the values the variable it is a proxy for can take. We denote a realization of the proxy for the outcome, action and signal by $(y^\bullet, x^\bullet, s^\bullet) \in Y \times X \times S$ respectively. We define a *proxy mapping* $\pi : Y \times X \times S \to \Delta(Y \times X \times S)$ that induces a distribution over the proxies for any realization of the true variables. The induced distribution over proxy variables is

$$p_\pi(y^\bullet, x^\bullet, s^\bullet; \sigma) = \sum_{y,x,s \in Y \times X \times S} \pi(y^\bullet, x^\bullet, s^\bullet | y, x, s)p(y,x,s;\sigma) \tag{7}$$

The DM needs to form beliefs about how their action affects the distribution over outcomes. We assume the DM uses the distribution over proxies p_π to form conditional beliefs.

$$p_\pi(y^\bullet | x^\bullet, s^\bullet; \sigma) = \frac{p_\pi(y^\bullet, x^\bullet, s^\bullet; \sigma)}{p_\pi(x^\bullet, s^\bullet; \sigma)} \tag{8}$$

Given this distorted belief distribution, the agent with signal s chooses an action x to maximize the perceived utility given below.

$$V(x,s;\sigma) = \sum_{y \in Y} u(y = y^\bullet, x, s)p_\pi(y^\bullet | x^\bullet = x, s^\bullet = s; \sigma) \tag{9}$$

2.1.1. The proxy mapping

Consider that i is any of the three variables, and that we can denote a realization of the true variable by v_i and a realization of the proxy by $v_i^\bullet$. It is possible that a proxy is a perfect measure for the underlying true variable, $v_i^\bullet = v_i$. Indeed, for all the applications in this paper some of the variables are perfectly observed. In the case where $v_i^\bullet \neq v_i$ with nonzero probability, we say that $i^\bullet$ is a *mismeasurement* of i. In applications, for simplicity we generally avoid drawing a distinction between the true and proxy variable when the proxy is a perfect measurement.

A proxy mapping that induces an identical joint distribution over the proxy variables and the true variables —for any initial distribution of the true variables— is called the *perfect measurement* mapping. Throughout the paper, we say that the beliefs induced by the perfect measurement mapping comprise the *rational expectations benchmark* or induce *correct beliefs*. In the absence of a knowledge of the true relationship between variables a rational Bayesian DM would have to form a prior belief about how the proxies and the true variables relate. From a normative perspective it is unclear how such a belief should be formed.

The formalism allows the DM to have imperfect equilibrium knowledge of the distribution of actions. This is motivated by a population level interpretation, where the dataset the DM is using to form beliefs is generated by other DMs facing the same or similar decision problems. The DM's own record of past actions only makes up a negligible part of this dataset. For example, a DM in a particular city may draw inference from a national level dataset when inferring the effect of a policy action on an outcome.

[5] The applications in Sections 3 and 4 require this general definition.

This allowance differs from the requirement of knowledge of the distribution of own actions in other solution concepts in the literature, such as Berk-Nash Equilibrium (Esponda and Pouzo, 2016) or Self-Confirming Equilibrium (Fudenberg and Levine, 1993). We discuss the substantive difference this can make in Section 6.

2.2. Equilibrium

We illustrate the dependence the decision maker's beliefs on their strategy $\sigma(x|s)$ using the following binary example.

Example 1. A team sports coach learns whether the realization of a signal s determining the cost of player injuries is high $s = \bar{s}$ or low $s = \underline{s} < \bar{s}$, before choosing whether to adopt attacking tactics $x = 1$ or defensive tactics $x = 0$. The game tactics in turn affects whether the injury levels of the players post-game is high $y = 1$ or low $y = 0$ according to the relationship $p(y = 1|x) = \beta x + (1 - \beta)(1 - x)$, where $\beta \in (\frac{1}{2}, 1)$.

The prior distribution over the injury cost signal is $p(\bar{s}) = \frac{1}{2}$. We assume the injury variable y is perfectly measured, but that the coach has no reliable proxy data for the injury cost. We can therefore write the proxy mapping as

$$\pi(y^\bullet, x^\bullet, s^\bullet|y, x, s) = \frac{1}{2}\pi_x(x^\bullet|x)\mathbb{1}\{y = y^\bullet\} \tag{10}$$

for some mapping $\pi_x : X \to \Delta(X)$. The tactical stance of teams is hard to measure, and sports analysts have created a proxy measurement from a large dataset of past games. Thus tactical stance is subject to measurement error, which we can write using the proxy mapping $\pi_x(x^\bullet = x|x) = \lambda$ where $\lambda \in (\frac{1}{2}, 1]$. As $\lambda \to 1$ we have close to perfect measurement for this variable.

Denote the ex-ante strategy as $\sigma(1) = \frac{1}{2}\sigma(1|\bar{s}) + \frac{1}{2}\sigma(1|\underline{s})$. The perceived effect of attacking tactics on high injury levels can then be calculated as

$$p_\pi(y^\bullet = 1|x^\bullet = 1; \sigma) - p_\pi(y^\bullet = 1|x^\bullet = 0; \sigma) = (2\beta - 1)(\frac{\sigma(1)\lambda}{\sigma(1)\lambda + \sigma(0)(1 - \lambda)} - \frac{\sigma(1)(1 - \lambda)}{\sigma(1)(1 - \lambda) + \sigma(0)\lambda}) \tag{11}$$

When the probability of attacking tactics $\sigma(1)$ falls, a greater proportion of the observations of attacking tactics are actually misclassified cases where defensive tactics occurred. This attenuates the perceived effect of attacking tactics on injuries from the true effect $2\beta - 1 > 0$ towards 0.

Thus, to characterize the DM's choices in general we need to define an equilibrium concept in order to establish consistency between strategies and beliefs. To ensure that conditional distributions are well defined, we first define an equilibrium with a small trembling probability.

Definition 1. Let σ^*_ϵ be a strategy such that $p_\pi(x^\bullet, s^\bullet; \sigma^*_\epsilon) > 0$ for every $(x^\bullet, s^\bullet) \in X \times S$. Then σ^*_ϵ is an ϵ-**Proxy Equilibrium** if for every $s \in S$, if

$$x \notin \arg\max \sum_{y^\bullet \in Y^\bullet} u(y = y^\bullet, x, s) p_\pi(y^\bullet|x^\bullet = x, s^\bullet = s; \sigma^*_\epsilon)$$

we have that $\sigma^*_\epsilon(x|s) < \epsilon$.

We then define a Proxy Equilibrium as the limit where the tremble probability goes to zero.

Definition 2. A strategy σ^* is a **Proxy Equilibrium** if there exists a sequence $\{\sigma^*_l\}_{l=1}^\infty$ converging to σ^* as well as a sequence $\epsilon^l \to 0$, such that for every l, σ^*_l is an ϵ^l-Proxy Equilibrium.

When the proxy mapping satisfies a *minimal responsiveness* condition, we can show the existence of at least one Proxy Equilibrium using conventional methods. A proxy mapping π is minimally responsive if whenever strategy mapping is full support $supp(\sigma(.|s)) = X$ for all $s \in S$; then we have $p_\pi(x^\bullet, s^\bullet; \sigma) > 0$ for any realization $(x^\bullet, s^\bullet) \in X \times S$.

Proposition 1. *Assume the proxy mapping π is minimally responsive. Then a Proxy Equilibrium exists.*

Proof. In Appendix. □

3. An illustrative application: rigid policing

Assume the DM is a municipality that has to make a decision on the number of police officers to hire. The municipality wants to hire police officers in order to reduce crime. We assume that there is noise in the measured variable for police numbers. Concern about measurement error in police staffing figures is not unprecedented. It is argued by Chalfin and McCrary (2018) that based on discrepancies between official data and administrative and census information there is significant measurement error in police staffing numbers in the literature estimating the effect of police numbers on crime. For expositional purposes, we assume that the crime variable is measured perfectly.[6]

The municipal leaders first learn the realization of a variable affecting the cost of crime in their municipality s. This is assumed to be distributed normally in the population of municipalities used in the dataset under consideration, $s \sim \mathcal{N}(\mu_s, \sigma_s^2)$. The municipal leaders then choose the change in the number of police officers. This affects the change in crime observed under their leadership via the relationship $y = \alpha - \beta x + u$, where $u \sim \mathcal{N}(0, \sigma_u^2)$ and $\beta > 0$. In the available data, it is assumed that changes in police numbers are measured as $x^\bullet = x + \epsilon$, where ϵ is normally distributed measurement error $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$. The DM lacks an informative proxy for the signal s, meaning that $p_\pi(y^\bullet|x^\bullet, s^\bullet) = p_\pi(y^\bullet|x^\bullet)$ for all $(y^\bullet, x^\bullet, s^\bullet) \in Y \times X \times S$, and so beliefs do not depend on the signal.[7]

The utility function of the municipal leader trades-off crime and policing costs. Higher s is assumed to reflect higher costs of crime relative to altering police numbers. We assume a symmetric cost of hiring and firing police officers. An alternative interpretation is that all the variables are in logs of the levels.

$$u(y, x, s) = -s \cdot y - \frac{1}{2}x^2 \tag{12}$$

Denote the rational expectations benchmark for how policing affects crime levels in expectation by $\mathbb{E}[y|x] = f(x) = \alpha - \beta x$. We can see that by plugging this into the utility function and calculating the best response that the optimal strategy under rational expectations for the municipal leader is to set police numbers such that $x^*(s) = \beta s$. Thus the rational expectations benchmark is for police numbers to be increased by more when the costs of crime is larger (higher s) and the effect of police numbers on crime is greater (higher $|\beta|$). Define a *linear equilibrium* as an equilibrium in which the strategy of the policy maker can be expressed as a linear function of the cost variable, $x(s) = \theta_0 + \theta_1 s$ for some $(\theta_0, \theta_1) \in \mathbb{R}^2$. We can characterize all the linear equilibria of the model as follows.

Proposition 2. *There is always a linear Proxy Equilibrium in which the municipal leader never changes police numbers, with best response* $x^{nv}(s) = 0$.

In addition, if $|\beta| \geq 2\frac{\sigma_\epsilon}{\sigma_s}$*, then there exist two additional linear Proxy Equilibria, with best response* $x^-(s) = (\frac{1}{2}\beta - \frac{1}{2}\sqrt{\beta^2 - 4\frac{\sigma_\epsilon^2}{\sigma_s^2}})s$ *and* $x^+(s) = (\frac{1}{2}\beta + \frac{1}{2}\sqrt{\beta^2 - 4\frac{\sigma_\epsilon^2}{\sigma_s^2}})s$.

There are no other linear Proxy Equilibria.

Proof. In Appendix. □

As previewed in the introduction through equation (4), due to measurement error in the police numbers proxy there is downward attenuation bias in the municipal leader's estimate of the expected change in the level of crime for any given change in police numbers.

$$\frac{\partial \mathbb{E}[y^\bullet|x^\bullet = x]}{\partial x^\bullet} = -\beta \frac{Var(x)}{Var(x) + \sigma_\epsilon^2} \tag{13}$$

For fixed $Var(x)$, this is attenuation bias from classical measurement error of the sort we see explained in introductory econometrics or statistics textbooks (See Wooldridge (2020), Chapter 9.4). What Proxy Equilibrium adds to the analysis is the endogeneity of $Var(x)$, which is determined by the equilibrium strategy σ. Under linear strategy $x(s) = \theta_1 \cdot s$, we have that $Var(x) = \theta_1^2\sigma_s^2$. Combining this with our attenuated formula for the marginal perceived effect of policing on crime gives

$$\frac{\partial \mathbb{E}[y^\bullet|x^\bullet]}{\partial x^\bullet} = -\beta \frac{\theta_1^2\sigma_s^2}{\theta_1^2\sigma_s^2 + \sigma_\epsilon^2} \tag{14}$$

Given the preferences of the municipalities, their equilibrium best response has to be such that $\frac{\partial \mathbb{E}[y^\bullet|x^\bullet]}{\partial x^\bullet} = -\theta_1$, so that greater perceived marginal effect of policing results in greater responsiveness of policing to the cost of crime s. Solving this equation then gives the linear Proxy Equilibria in Proposition 2.

The endogeneity of the true policing variable under Proxy Equilibrium results in multiplicity of equilibria. For a fixed measurement structure π, we have both an equilibrium where municipalities do not vary police numbers at all and equilibria where police numbers vary proportionally to the effectiveness of policing $|\beta|$. We can see how this is starkly different from a purely statistical analysis without endogenous actions by considering the case when the measurement error is arbitrarily small and σ_ϵ is close to zero. Here we have the coexistence of an equilibrium where police numbers are at close to the same level as under perfect measurement with

[6] Adding normally distributed measurement error to the crime variable, so that $y^\bullet = y + v$ with $v \sim \mathcal{N}(0, \sigma_v^2)$, does not change either the set of Proxy Equilibria nor does it change the rational expectations benchmark. That the rational expectations benchmark is unchanged is easy to see due to linearity of expectations. The same is true in the Proxy Equilibrium case due to both the linearity of the conditional expectation and the fact that the additional variance in $y^\bullet$ does not affect the marginal perceived incentive over the policing variable.

[7] The conditions on the proxy mapping for an uninformative signal proxy in this application require the definition of Proxy Equilibrium for general spaces in Appendix A.1. The condition is that for any Borel sets $\mathcal{Y} \times \mathcal{X} \times \mathcal{S} \subset Y \times X \times S$, we can decompose the proxy mapping as $\pi(\mathcal{Y} \times \mathcal{X} \times \mathcal{S}|y, x, s) = \pi_s(\mathcal{S}|s)\pi_{y,x}(\mathcal{Y} \times \mathcal{X}|y, x)$. Note also that since for the true variables $p(y|x, s) = p(y|x)$, with no informative proxy for s perfect measurement of (y, x) still results in correct beliefs.

an equilibrium with massive attenuation bias and no variation in police numbers at all. A purely statistical analysis of measurement error in this situation would focus on either one or the other.

A simple dynamic story of how these equilibria emerge can also give an intuition for why there is multiplicity. Suppose that DMs initially believe that policing has little effect on crime. This leads to little variation in police numbers across the municipalities, as it is not worth hiring and firing police officers if they have no effect. In steady state this leads to the no variation equilibrium. In contrast, if the DM believes initially that police have a large effect on crime, then municipalities who face a higher cost of crime will higher relatively more police officers. This can generate enough variation to mitigate the attenuation bias somewhat and preserve the DMs' beliefs about the effectiveness of policing in equilibrium.

In all linear Proxy Equilibria, the policy chosen is more rigid than the rational benchmark. The rigidity results solely from bias beliefs due to measurement noise, and as discussed above extreme rigidity can coincide with very small noise. Apparent over-rigidity in choices has been discussed in relation to monetary policy Tetlow and Von zur Muehlen (2001) and firm pricing Nakamura and Steinsson (2013). It is possible that this application could be extended into a more complicated model that could provide an alternative explanation of these phenomena as resulting from use of mismeasured data from decision makers.

The simple parameterization in the policing application focuses on the case where we have normally distributed measurement error in a linear model with a single explanatory variable. Although this results in attenuation bias, it is not true that for general measurement error attenuation bias always arises. As a concept Proxy Equilibrium is able to capture alternative measurement error structures, such as those described in Schennach (2022), and future applications of Proxy Equilibrium could utilize some of these alternative structures.

4. Market entry: endogenous overoptimism

Businesses entering into new markets have high rates of failure. Using data from the US Census Bureau Haltiwanger (2015) calculates that half of new firms exit the market within 5 years. In UK data, 38 percent of enterprises newly born in 2016 survived 5 years.[8] A literature in business and economics attributes these seemingly excessive levels of market entry to overoptimism on the part of the potential market entrants, see Hayward et al. (2006), Cooper et al. (1988), Malmendier and Tate (2005).

We build an application of our solution concept that generates firms that have an upwardly biased assessment of the payoffs from entering new markets as a feature of equilibrium. Firms draw on noisily recorded data drawn from past entrants. There is a variable $s \in [0,1] \equiv S$, representing the location of markets in some space, which could be geographical or based on demographic information. After learning the realization of this variable, the potential market entrant has to make a binary decision on whether to enter $x = 1$ or not $x = 0$. The payoff of the entrant is measured via an outcome variable $y \in \mathbb{R} \equiv Y$ representing the profitability of the enterprise, so that $u(y, x, s) = y$. The outcome variable is determined by both the entry decision and the market location variable by the following relationship.

$$\mathbb{E}[y|x,s] = \int_Y yp(y|x,s)d\mu(y) = m(s)x \tag{15}$$

We assume that the function $m : S \to \mathbb{R}$ is strictly increasing, bounded and right-continuous, with a single point of crossing $\alpha \in [0,1]$ such that $m(s) < 0$ for all $s \in [0,\alpha)$ and $m(s) \geq 0$ for $s \in [\alpha, 1]$. Thus for high enough realizations of the market location variable, the expected profitability of entry is always greater than the payoff of zero from not entering. Under rational expectations the best response of the potential entrant is clear; when $s \in [0,\alpha)$ $x = 0$ is optimal while for $s \in [\alpha, 1]$ the payoff from entering is weakly above zero and therefore entry is optimal.

We assume that while the potential entrant has perfectly measured data on firm profitability y and past entry choices x, they do not have data on how the market location s varies with the outcome and action variable. Instead they have access to a noisy recorded proxy variable for the location $s^\bullet$. The idea is that each market has a very granular definition, and in data it can only be recorded in an imprecise fashion. This could be due to data protection reasons when s is demographic information, for example. Thus we assume the proxy variable is generated by a mapping that has the following 'window' form.[9] There is some parameter $h \in (0, \frac{1}{2})$ such that every firm location is recorded in the data as being uniformly distributed in the nearest window of size $2h$. This means for $s \in [h, 1-h]$ we have that $s^\bullet$ is uniformly distributed on the window $[s-h, s+h]$, while for locations close to the boundary when $s \in [0,h)$ we have that $s^\bullet$ is distributed uniformly on $[0, 2h)$ and for all $s \in (1-h, 1]$ we have that $s^\bullet$ is distributed uniformly on $s \in (1-2h, 1]$.

The following result shows both that Proxy Equilibria exist in this setting and that any Proxy Equilibrium with entry will take a cut-off form. We assume $h < \min\{\frac{\alpha}{2}, \frac{1-\alpha}{2}\}$.[10]

Proposition 3. *Assume $h < \min\{\frac{\alpha}{2}, \frac{1-\alpha}{2}\}$. Then there is a cut-off $\bar{s} \in [h, 1-h]$ such that there is a Proxy Equilibrium with strategy $\sigma(x = 1|s) = 0$ for $s \in [0, \bar{s}]$ and $\sigma(x = 1|\bar{s}) = 1$ for $s \in (\bar{s}, 1]$.*

This cut-off is always strictly less than the rational entry point; $\bar{s} < \alpha$.

---

[8] This statistic is from Office for National Statistics (2023).

[9] This window form of proxy noise is similar to the notion of similarity used in Steiner and Stewart's (2008) model of learning in games.

[10] This ensures h is small enough so that the firm will not enter at low values of s and there is always an equilibrium in which the firm will enter at high values of s. The absence of this requirement on bandwidth parameter h complicates the analysis by requiring us to consider additional cases, but does not create a fundamental difference in that we can regard these cases as cut-off equilibria with cut-off points at the boundary.

In addition, there is always a Proxy Equilibrium where $\sigma(x=1|s)=0$ for all $s \in [0,1]$.
There are no other Proxy Equilibria.

Proof. In Appendix. □

We see that in equilibrium, if there is entry there is over-entry. Since in the equilibrium data there is negligible observations of entrants below a certain market location, the proxy observations for these markets are disproportionately higher location markets that have been misclassified as lower ones. This leads to an overestimate of the payoff from entering at these lower levels. However, enough over-entry reduces the extent of this proxy bias and in equilibrium the DM is indifferent between entering or not at some cut-off $\bar{s}$ below the cut-off they would enter at under rational expectations. Note also that there is always an equilibrium with no-entry, even if the measurement problem is very small and h is close to zero. The discontinuity between the size of measurement error and the bias in beliefs which holds in the policing application also holds in the market entry model.

While this equilibrium distortion in beliefs resembles the selection bias studied by econometricians such as in Heckman (1979), the distortion is caused by the interaction of the proxy noise and the extreme difference in the number of entering firms above and below the cut-off rather than complete lack of observations that do not meet some selection criteria. The Proxy Equilibrium with entry is the limit of a trembling-hand sequence of equilibria which have entry and non-entry at all signals. The overestimation of the value of entry only systematically occurs at signals that are close to the entry threshold, and at signals far from the threshold the firm can underestimate the value of entry depending on the parameters. This is a distinguishing feature from Jehiel (2018)'s model of selection bias and firm entry, where firms over-estimate the value of entry at all signals.

The fact there is a systematic effect of the beliefs of the pivotal firm is important for the following result, which establishes comparative statics for how Proxy Equilibria vary with the noise parameter. We see that noisier proxies —in the form of higher h— always lead to greater levels of excessive entry.

Proposition 4. *Consider two noise parameters $\min\{\frac{\alpha}{2}, \frac{1-\alpha}{2}\} > h_2 > h_1 > 0$. We have that the cut-off $\bar{s}(h_2)$ under the positive entry Proxy Equilibrium with noise parameter h_2 is strictly less that the cut-off $\bar{s}(h_1)$ under h_1.*

Proof. In Appendix. □

The intuition for this result is as follows. In general —for a fixed belief distribution— it is ambiguous whether larger h increases or decreases the payoff from entry at any given s. However, in equilibrium what matters is the beliefs at the pivotal cut-off $\bar{s}$. At the cut-off the DM must be indifferent between entering and not. An increase in h will always lead to greater weight on the part of the function $m(s)$ that is above the cut-off, in particular greater weight on the positive part of $m(s)$. This pushes up the expected payoff from entering strictly above zero at this cut-off, and the new cut-off at the larger h must be below in order to restore indifference.

The equilibrium requirement is vital for Proposition 4. If beliefs are formed from a proxy distribution with the action distribution fixed, and are not required to satisfy equilibrium conditions, we can construct cases in which entry is both excessive and increasing h results in firms choosing to enter at fewer signals. An example where this is the case is presented in Appendix A.2. In general, increasing proxy noise h does not lead to a change in beliefs resembling the DM having less information in a Blackwell (1953) sense. Depending on the shape of $m(.)$ and the true signal distribution $p(s)$ increasing h from some baseline can actually result in beliefs that are closer to those that would occur with perfect measurement. Proposition 4 holds because the equilibrium requirements mean increasing h leads the beliefs for the pivotal firm to always over-value entry relative to the rational benchmark.

As a corollary to Proposition 3 and 4, we have that as $h \to 0$, the cut-off equilibrium converges to the rational entry strategy. This is because $\bar{s}(h) < \alpha$ and $\bar{s}(h)$ is strictly decreasing in h.

Corollary 4.1. *Consider any sequence of bandwidth parameters $\{h_l\}_{l=1}^{\infty}$, such that $0 < h_l < \min\{\frac{\alpha}{2}, \frac{1-\alpha}{2}\}$ for all l and $h_l \to 0$. Then the corresponding sequence of Proxy Equilibrium cut-offs $\{\bar{s}(h_l)\}_{l=1}^{\infty}$ is such that $\bar{s}(h_l) \to \alpha$ from below.*

5. Almost perfect proxies

In this section we characterize the set of all strategies that can arise as Proxy Equilibria even as the variables are arbitrarily close to being perfectly measured. We call these strategies *Self-Confirming Optimal*. Self-confirming optimality is an adaption of the Self-Confirming Equilibrium of Fudenberg and Levine (1993), Battigalli (1987) to our setting. These strategies are optimal against any perceived beliefs that are correct 'on-path'. That is to say, correct for actions-signal combinations that occur with positive probability under the strategy itself. If a strategy is in the set, then we can choose a particular proxy mapping that implements that strategy as an equilibrium. If a strategy does not meet the conditions to be Self-Confirming Optimal, then it cannot be implemented as a Proxy Equilibrium for some proxy mapping that is above a certain level of proximity to perfect measurement.

In Section 6 we discuss another variant of SCE that differs from Self-Confirming Optimality in that DMs obtain even more limited feedback. Proxy Equilibrium can be considered a refinement of this more permissive version of SCE but in general is not a refinement of Self-Confirming Optimality if we have a general proxy mapping that is not required to be almost perfectly measured.

5.1. Definition of almost perfect proxies

The following division of the signal space allows us to distinguish between private payoff shocks and *control signals*. The DM has access to informative data on how control signals covary with actions and outcomes. Our notion of perfect measurement then applies to the control signals along with the actions and outcomes.

Split the signal vector $s = (s_1, ..., s_m) \in S$ into two component parts. The last $1 \leq l \leq m$ dimensions are a set of *private signals* $s^p = (s_l, ..., s_m)$. Denote the remaining signal dimensions by $s_c = (s_1, ..., s_{l-1})$, so that $s = (s_c, s_p) \in S \equiv S_c \times S_p$. We call the remaining signals $s_c \in S_c$ *control signals*. The private signal dimensions do not affect the outcome variable; $p(y|x, s_c, s_p) = p(y|x, s_c)$ for all $(y, x, s_c, s_p) \in Y \times X \times S_c \times S_p$. In addition, the DM does not have informative proxies for the private signals. For some $\pi_{s_p} : S_p \to \Delta(S_p)$ with full support on S_p, $supp(\pi_{s_p}) = S_p$, we have that

$$\pi(y^\bullet, x^\bullet, s_c^\bullet, s_p^\bullet | y, x, s_c, s_p) = \pi_{s_p}(s_p^\bullet)\pi(y^\bullet, x^\bullet, s_c^\bullet | y, x, s_c.s_p) \tag{16}$$

for all $(y, x, s_c, s_p) \in Y \times X \times S_c \times S_p$. This means it is without loss to focus on beliefs condition on control signals $p_\pi(y|x, s_c)$ rather than conditional on all signals $p_\pi(y|x, s)$.

We define the *perfect measurement proxy mapping* π_δ as a proxy mapping which gives perfect measurement for the outcome, action and control signals

$$\pi_\delta(y^\bullet, x^\bullet, s_c^\bullet | y, x, s_c, s_p) = \sum_{s_p^\bullet \in S_p} \pi_\delta(y^\bullet, x^\bullet, s_c^\bullet, s_p^\bullet | y, x, s_c, s_p) = \mathbb{1}\{(y^\bullet, x^\bullet, s_c^\bullet) = (y, x, s_c)\} \tag{17}$$

Similarly, denote the proxy mapping over the outcome, action and control signals by

$$\pi_{Y,X,S_c}(y^\bullet, x^\bullet, s_c^\bullet | y, x, s_c, s_p) = \sum_{s_p^\bullet \in S_p} \pi(y^\bullet, x^\bullet, s_c^\bullet, s_p^\bullet | y, x, s_c, s_p) \tag{18}$$

The *Total Variation distance* between any two probability distributions p and q is

$$TV(p, q) = max_{\mathcal{A} \in 2^A} |\sum_{a \in \mathcal{A}} p(a) - \sum_{a \in \mathcal{A}} q(a)| \tag{19}$$

We use the Total Variation distance to define a notion of proximity of the proxy variables to perfect measurement.

Definition 3. Given $\eta > 0$, we say the proxy mapping π is η-close to perfect if

$$TV(\pi_{Y,X,S_c}(.|y, x, s), \pi_\delta(.|y, x, s)) < \eta \tag{20}$$

We state a result demonstrating why our notion of proximity is suitable for the Proxy Equilibrium setting. It implies that if the joint distribution over the true variables satisfies a full support requirement, then beliefs become arbitrarily close to rational expectations as the proxy variables become close to perfect measurements in the total variation distance. Since our notion of almost perfect measurement applies to the control signals but not the private signals, the result demonstrates that good measurement for the control signals, actions and outcomes is all that is required to have convergence to correct beliefs.

Given distribution over the true variables $p(y, x, s) = p(y|x, s)\sigma(x|s)p(s)$, let $V^+(\sigma) = \{(x, s_c) \in X \times S_c : p(x, s_c) > 0\}$ be the support on $X \times S_c$ induced by σ. We say a strategy σ induces full support if $V^+(\sigma) = X \times S_c$. We can show that a continuity property holds for the perceived conditional distribution for any (x, s) that is in the induced support of σ.

Proposition 5. *Fix strategy σ. For any $\epsilon > 0$, there exists an $\eta > 0$ such that if the proxy mapping π is **η-close to perfect** then*

$$|p_\pi(y^\bullet = y | x^\bullet = x, s_c^\bullet = s_c; \sigma) - p(y|x, s_c)| < \epsilon$$

for every $(x, s_c) \in V^+(\sigma)$ and $y \in Y$.

Proof. In Appendix. □

The result concerns both equilibrium and out of equilibrium beliefs, and can be used as a diagnostic when considering equilibria in which full-support does not hold. For example, in our policing application the full-support assumption does not always hold and therefore we can have large belief distortions even as the proxy noise is close to zero.

5.2. Equivalence with self-confirming optimality

We define the conditions required for a strategy to be *Self-Confirming Optimal* below. The definition requires that the strategy meets different conditions for actions that are in the support of the strategy and actions that are not. We define a system of beliefs as a collection of conditional distributions $q : X \times S_c \to \Delta(Y)$, $q = \{q(.|x, s_c)\}_{(x,s_c) \in X \times S_c} \in Q$. Given a strategy σ, define the strategy conditional on the control signals as $\sigma(x|s_c) = \sum_{s_p \in S_p} \sigma(x|s_c, s_p)p(s_p)$.

Definition 4. A strategy $\sigma^* : S \to \Delta(X)$ is **Self-Confirming Optimal** if there exists a system of beliefs $q \in Q$ such that

1. The beliefs are correct for actions and control signals that arise with positive probability under σ^*. At every (x, s_c) such that $x \in supp(\sigma^*(.|s_c))$, we have $q(y|x, s_c) = p(y|x, s_c)$ for all $y \in Y$.
2. The strategy σ^* is optimal given beliefs q. At every $s \in S$, for any $x \in \sigma^*(.|s)$ and $x' \in X$ we have that

$$\sum_{y \in Y} u(y, x, s) q(y|x, s_c) \geq \sum_{y \in Y} u(y, x', s) q(y|x', s_c) \tag{21}$$

If there exists $q \in Q$ such that the first condition holds and (21) always holds strictly, we say that σ^* is **strictly Self-Confirming Optimal**.

The strength of this condition varies with the utility function and the split between private and control signals. If payoffs are such every action is a best response for some private signal regardless of beliefs, then only the strategies that are optimal against the true conditional distribution are Self-Confirming Optimal. If any action can be sub-optimal at all signals for some beliefs, then the condition is very permissive. We have the following result.

Proposition 6. *If strategy $\sigma^* : S \to \Delta(X)$ is **Self-Confirming Optimal** then for any $\eta > 0$ it is a Proxy Equilibrium under some proxy mapping that is η-close to perfect.*

*Moreover, if σ^* is **strictly Self-Confirming Optimal** and $supp(p(y|x, s)) = Y$ for all $(x, s) \in X \times S$, then for any $\eta > 0$ we can always construct the Proxy Mapping under which σ^* is a Proxy Equilibrium such that the outcome variable is perfectly measured.*

*If $\sigma^* : S \to \Delta(X)$ is not a **Self-Confirming Optimal** strategy then there is an $\bar{\eta} > 0$ such that for all $\eta \in (0, \bar{\eta})$ σ^* is not a Proxy Equilibrium under any proxy mapping that is η-close to perfect.*

Proof. In Appendix. □

Note that this result can be restated to say that: a strategy is Self-Confirming Optimal if and only if for any $\eta > 0$ it can be implemented as a Proxy Equilibrium under some proxy mapping that is η-close to perfect. This can be seen by taking the contrapositive of the last part of the proposition statement.

The first part of the result is proven by constructing a proxy mapping that is close to perfect measurement but has a small probability of randomly allocating a particular realization of the true outcome variable to the proxy of an action-control signal combination that has zero probability under the proposed equilibrium strategy. This small mismeasurement is chosen in a particular way to ensure the DM is deterred from choosing actions that are not part of σ^* at that control signal. Under strict Self-Confirming Optimality, this can be done in a way that ensures the outcome variable is perfectly measured. Thus many Self-Confirming Optimal strategies can be implemented under conditions even stronger than close to perfect measurement.

The second part of the result follows from applying Proposition 5. The fact that the proxy mapping is close to perfect measurement but not exactly perfectly measured is vital, there are in general Self-Confirming Optimal strategies that are not Proxy Equilibria with exact perfect measurement.

We can apply Proposition 6 to our binary sports coach example from Section 2.2 to analyze what strategies can arise as Proxy Equilibria for arbitrarily small measurement noise. This example is also used to illustrate how Proxy Equilibria that do not satisfy Self-Confirming Optimality break down as measurement gets closer to perfect.

Example 1 *(Continued)*. Let the payoff function of the coach be

$$u(y, x, s) = s(1 - y) - x(1 - s)$$

This reflects that when the cost s is high, the coach has a greater payoff loss from player injuries. However, attacking tactics can only benefit the team at points when injury costs are high enough. Assume that $\underline{s} = 0$ and $\bar{s} > 0$. Then we have $\sigma(0|\underline{s}) = 1$ in any Proxy Equilibrium. This is because at $s = 0$, defensive tactics $x = 0$ are a best response regardless of the beliefs of the DM about how x covaries with y.

There are three cases at which different strategies are Self-Confirming Optimal.

1. If $\bar{s} \in [0, \frac{1}{2(1-\beta)})$ then we must have $\sigma(0|\bar{s}) = 1$ in any Self-Confirming Optimal strategy. This is because at $s = \bar{s}$ the following inequality must hold for $x = 1$ to be a best response.

$$\bar{s}(2 - p_\pi(y = 1|x^\bullet = 1)) - 1 \geq \bar{s}(1 - p_\pi(y = 1|x^\bullet = 0)) \tag{22}$$

A strategy in which $x = 1$ is chosen with positive probability at $\bar{s}$ is one in which both actions have some positive probability for some signal s; $\sigma(0) > 0$ and $\sigma(1) > 0$. Self-Confirming Optimality then requires that both actions are optimal at some signal against a system of beliefs that is correct. Since $x = 1$ is not optimal at any signal given correct beliefs if $\bar{s} \in [0, \frac{1}{2(1-\beta)})$, we have our claim.

2. If $\bar{s} \in (\frac{1}{1-\beta}, \infty)$ then it must be the case that $\sigma(1|\bar{s}) = 1$. The argument is as follows: since $\sigma(1|\underline{s}) = 0$ always, we must have Self-Confirming Optimality with respect to a system of beliefs for which $p_\pi(y = 1|x^\bullet = 0) = 1 - \beta$. Optimality of $x = 0$ at $s = \bar{s}$ then requires that for some beliefs $p_\pi(y = 1|x^\bullet = 1) \in [0, 1]$
$$\bar{s}\beta \geq \bar{s}(2 - p_\pi(y = 1|x^\bullet = 1)) - 1 \tag{23}$$
This is violated for any such beliefs if $\bar{s} > \frac{1}{1-\beta}$.
3. If $\bar{s} \in [\frac{1}{2(1-\beta)}, \frac{1}{1-\beta}]$, then strategies $(\sigma(1|\bar{s}), \sigma(1|\underline{s})) = (1, 0)$ and $(\sigma(1|\bar{s}), \sigma(1|\underline{s})) = (0, 0)$ are both Self-Confirming Optimal. The first type of strategy can be implemented by the perfect measurement mapping, which is trivially arbitrarily close to perfect measurement. When $\bar{s} \in (\frac{1}{2(1-\beta)}, \frac{1}{1-\beta})$, the second type of strategy can be implemented by the following proxy mapping, which has perfect measurement of the outcome
$$\begin{aligned} &\pi(y^\bullet = y, x^\bullet = x|y, x) = 1 - \eta + \eta(1 - y) \\ &\pi(y^\bullet = y, x^\bullet \neq x|y, x) = \eta y \end{aligned} \tag{24}$$
This results in beliefs such that $p_\pi(y = 1|x^\bullet = 1) = 1$ and $p_\pi(y = 1|x^\bullet = 0) = \frac{(1-\beta)(1-\eta)}{(1-\beta)(1-\eta)+\beta} \to 1 - \beta$ as $\eta > 0$ tends to zero. This can then sustain the proposed strategy as a proxy equilibrium as it satisfies inequality (23) for $\bar{s} \in [0, \frac{1}{1-\beta})$.
When $\bar{s} = \frac{1}{1-\beta}$ we can implement the strategy with a proxy mapping that has imperfect measurement of the outcome variable.
$$\pi(y^\bullet, x^\bullet|y, x) = (1 - \eta + \eta\frac{1}{2}(1 - x^\bullet))\mathbb{1}\{y^\bullet = y, x^\bullet = x\} + \eta\frac{1}{2}x^\bullet y^\bullet$$
This mapping induces the exact beliefs $p_\pi(y^\bullet = 1|x^\bullet = 1) = 1$ and $p_\pi(y^\bullet = 1|x^\bullet = 0) = 1 - \beta$ when $(\sigma(1|\bar{s}), \sigma(1|\underline{s})) = (0, 0)$.

Note how a Proxy Equilibrium that doesn't satisfy the conditions for Self-Confirming Optimality will break down if the proxy mapping is close enough to perfect measurement. For example, consider that $\bar{s} > \frac{1}{1-\beta}$. The proxy mapping (24) given in the third case can sustain $(\sigma(1|\bar{s}), \sigma(1|\underline{s})) = (0, 0)$ as a Proxy Equilibrium when
$$\begin{aligned} &\bar{s}(1 - p_\pi(y = 1|x^\bullet = 0)) \geq \bar{s}(2 - p_\pi(y = 1|x^\bullet = 1)) - 1 \\ &\Rightarrow \bar{s}(1 - \frac{(1-\beta)(1-\eta)}{(1-\beta)(1-\eta)+\beta}) \geq \bar{s} - 1 \\ &\Rightarrow \eta \geq \frac{\bar{s}(1-\beta) - 1}{(1-\beta)(\bar{s} - 1)} > 0 \end{aligned}$$
Thus for small enough $\eta > 0$ this strategy cannot be sustained as a Proxy Equilibrium by proxy mappings of this form.

6. Related literature and variants

6.1. Relationship to Berk-Nash equilibrium

The Berk-Nash Equilibrium of Esponda and Pouzo (2016) gives a general solution concept for games in which players have to form beliefs about a mapping between actions, a signal variable and outcome variables that may depend on the actions and signals of multiple players. Each player has a set of subjective models over this mapping. Under the solution concept the beliefs of the players have to be such that any subjective model that is in the support of belief of the player minimizes the Kullback-Leibler divergence between the true distribution over outcomes and that projected by the model, weighted by that player's signal and action probabilities. This is then founded as the limit of a Bayesian learning process where the players have a prior with their set of subjective models as the support. The true model that generates the mapping between the action, signal and outcome variables may not be in the set of subjective models and thus players may have misspecified beliefs.

Proxy Equilibrium cannot be nested as special case of the exact version of Berk-Nash Equilibrium outlined in Esponda and Pouzo (2016). However, it can fit as a variant of an extended version they discuss in the supplementary appendix of that paper. The version of Berk-Nash Equilibrium in the main body requires that players perfectly observe the joint distribution of their signals, actions and outcomes and that payoffs are only measurable with respect to these variables. If DM's feedback consists both of perfectly observed actions and signals and imperfectly measured proxies, the set of subjective models cannot contain models that put probability one on the proxies being identical to the true variables. This is because the Kullback-Leibler divergence would not be well defined for that model, as it would place zero probability on the event that the proxies and observed true variable realizations differ even though they differ with positive probability.

If the only feedback agents receive is the distribution over proxies, then the Kullback-Leibler divergence is minimized at zero by any model that implies the correct distribution over proxies. Since under Proxy Equilibrium agents are using such a model, their beliefs are thus consistent with Kullback-Leibler divergence minimization.

To illustrate this further we can write a variant of Berk-Nash where players are not required to observe the true signal and action distribution, and payoffs are not required to be measurable with respect to feedback. The space of all possible conditional beliefs

$\{q(y|x,s)\}_{(y,x,s)\in Y\times X\times S}$ is denoted by $\hat{Q}$ and denotes proxy variable realizations by $w^\bullet = (y^\bullet, x^\bullet, s^\bullet)$. Given an observed distribution over proxies $p_\pi(y^\bullet, x^\bullet, s^\bullet)$, the set of conditional beliefs that are consistent with this feedback; $\mathcal{M}(p_\pi)$, are those for which there exists distribution over actions and signals $\tilde{p}$, and a proxy mapping $\hat{\pi}$, that would generate the same distribution over proxies.

$$\mathcal{M}(p_\pi) = \{q \in \hat{Q} : \exists \tilde{p}, \hat{\pi} \text{ s.t } \forall w^\bullet, p_\pi(w^\bullet) = \sum_{y,x,s} q(y|x,s)\tilde{p}(x,s)\hat{\pi}(w^\bullet|y,x,s)\} \tag{25}$$

Any subjective model that puts probability one on the distribution over feedback being p_π is consistent with this version of Berk-Nash Equilibrium, as the Kullback-Leibler divergence is then zero. Proxy Equilibrium would then be a selection from the set of Berk-Nash equilibria where q is chosen from $\mathcal{M}(p_\pi)$ so that $q(y|x,s) = p_\pi(y^\bullet = y|x^\bullet = x, s^\bullet = s)$, $\tilde{p}(x,s) = p_\pi(x^\bullet = x, s^\bullet = s)$ and $\hat{\pi}$ is the identity function.

Note that this variant of Berk-Nash is very permissive. In fact, for fixed p_π any beliefs $q \in Q$ are also in the set of beliefs that can be held in equilibrium; $q \in \mathcal{M}(p_\pi)$. We can see this if we choose $\hat{\pi}(w^\bullet|y,x,s) = p_\pi(w^\bullet)$ for all $(y,x,s) \in Y\times X\times S$ and $w^\bullet \in Y\times X\times S$. Proxy Equilibrium can significantly refine the set of possible beliefs held in equilibrium if we consider particular noise structures that are neither implausibly far from correct measurement nor tailored to implement a particular strategy. For example, the symmetric normally distributed or uniform window forms of measurement error in the applications in this paper.

6.2. Relationship to self-confirming equilibrium

Self-Confirming Equilibrium (SCE) originates from work by Fudenberg and Levine (1993), Battigalli (1987). Under this concept players in a game only receive limited feedback on the equilibrium actions of other players and nature. Other work has that has generalized the notion of limited feedback from these original papers includes Dekel et al. (2004), Battigalli et al. (2015), Fudenberg and Kamada (2015), Fudenberg and Kamada (2018), Lipnowski and Sadler (2019) and Battigalli et al. (2019).

We have shown in Section 5 that when we consider only very small noise, Proxy Equilibrium is a refinement of a version of SCE we call Self-Confirming Optimality. If all signals are control signals, $s = s_c$, then Self-Confirming Optimality coincides with the definition of Self-Confirming Equilibrium defined in footnote 31 of Esponda and Pouzo (2016). However if there are private signal dimensions, the definitions may differ. If $x \notin \sigma(.|s_c, s_p)$ but $x \in \sigma(.|s_c)$, Self-Confirming Optimality requires that beliefs are correct for (x, s_c, s_p) while Esponda and Pouzo (2016)'s definition of Self-Confirming Equilibrium puts no restriction on beliefs these actions and signals.

For general noise, a Proxy Equilibrium may not be Self-Confirming Optimal, as the 'on path' beliefs for x, s such that $\sigma(x|s)p(s) > 0$ may be incorrect. The variant of Berk-Nash Equilibrium discussed above can also be considered as another variant of SCE. This variant differs from Self-Confirming Optimality as due to the more limited feedback it does not require correct beliefs on-path.

6.3. Imperfect control

To illustrate the importance of the DMs imperfect equilibrium knowledge of the action distribution in Proxy Equilibrium, we consider a variant in which actions are imperfectly executed and the intended action is used by the DM as the mismeasured proxy of the actually implemented action. Unlike in Proxy Equilibrium, the feedback effects that can arise in this variant depend only on the support of the equilibrium strategy.

In the context of the policing application, the interpretation would be that the municipality decides how many police officers to hire but the actual number hired differs. The municipality then uses data on how intended hiring is jointly distributed with crime to infer the effect of policing on crime.

Let $x^\bullet$ be the proxy action chosen by the DM. Given a strategy mapping $\sigma(x^\bullet|s)$ and a distribution over signals s, we write the distribution over proxy actions as $\sigma(x^\bullet) = \sum_{s\in S}\sigma(x^\bullet|s)p(s)$. The true action $x \in X$ is determined by proxy mapping $\pi(x|x^\bullet)$ and the distribution over perfectly measured outcomes is given by $p(y|x)$.

We assume the DM can only observe the joint distribution over the proxy actions and the outcomes.

$$p_\pi(y, x^\bullet) = \sigma(x^\bullet)\sum_x \pi(x|x^\bullet)p(y|x) \tag{26}$$

Given $\sigma(x^\bullet) > 0$, if the DM were to take the conditional distribution $p_\pi(y|x^\bullet) = \sum_x \pi(x|x^\bullet)p(y|x)$ as their beliefs, then if payoffs only depend on the intended action, the outcome and the signal the DM's beliefs coincide with rational expectations. Thus any two strategy distribution with the same support over actions then result in the same beliefs. This differs from Proxy Equilibrium; consider the beliefs obtained in Example 1, under which strategies with the same support over actions may result in different beliefs.

This carries over if we consider any selection from the set of beliefs consistent with observing the distribution of proxy actions $\sigma(x^\bullet)$ and how they vary with outcomes. Let $\hat{q}(y,x|x^\bullet)$ denote a conditional joint distribution over (y,x) and let $\hat{Q}$ be the set of all such distributions. Define the set of conditional joint distributions consistent with the observable distribution p_π as

$$\mathcal{M}^{ic}(p_\pi) = \{\hat{q} \in \hat{Q} : \sum_x \hat{q}(y,x|x^\bullet)\sigma(x^\bullet) = p_\pi(y,x^\bullet)\} \tag{27}$$

This set of beliefs is invariant to strategies with the same support; σ, σ' such that $\sigma(x) > 0$ if and only if $\sigma'(x) > 0$. We can see this by dividing both sides of $\sum_x \hat{q}(y,x|x^\bullet)\sigma(x^\bullet) = p_\pi(y,x^\bullet)$ by $\sigma(x^\bullet) > 0$ as then $\mathcal{M}^{ic}(p_\pi)$ depends only on $p_\pi(y|x^\bullet) = \sum_x \pi(x|x^\bullet)p(y|x)$.

6.4. Learning foundation

Since in Proxy Equilibrium agent's naive belief that proxies are identical to true variables is dogmatic, it is unaffected by learning. Thus the main requirement for any learning foundation for Proxy Equilibrium is that the sample distribution over the proxies eventually converges to the Proxy Equilibrium distribution.

If we have full support for the action and signal proxies for all strategies $p_\pi(x^\bullet, s^\bullet) > 0$ for all $(x^\bullet, s^\bullet) \in X \times S$ and σ, then a learning foundation is simple. Since then no action-signal combination occurs with zero probability, standard law of large numbers arguments would ensure convergence of beliefs.

This would hold for particular proxy mappings, or if we had a proportion of DMs forced to experiment with actions in the learning model. Without these assumptions then any learning foundation would have to consider agents incentives to experiment as in Fudenberg and Levine (2006), Fudenberg and He (2018). Such active experimentation could clash with the idea underlying the concept that agents are short-lived and dogmatic in assuming that observed variables are not noisy.

6.5. Other related literature

The strand of literature that this paper is most clearly related to is that on equilibrium solution concepts with bounded rational expectations. Work on using Bayesian Networks as a formalism to model causal misperceptions originating from Spiegler (2016) has been developed to explore interactive beliefs in games (Spiegler, 2021); political narratives (Eliaz and Spiegler, 2020), (Eliaz et al., 2024); persuasion (Eliaz et al., 2021); contract theory (Schumacher and Thysen, 2022) and deception (Spiegler, 2020). Other solution concepts in this tradition include the Cursed Equilibrium of Eyster and Rabin (2005), the Behavioural Equilibrium of Esponda (2008) and the Analogy Based Expectation Equilibrium of Jehiel (2005), Jehiel and Koessler (2008).

The Berk-Nash Equilibrium of Esponda and Pouzo (2016) nests many of these concepts and provides a foundation for the literature on dynamic misspecified learning. Papers in the broader misspecified learning literature have explored overconfidence about one's ability; Heidhues et al. (2018), social learning; Bohren and Hauser (2021) and connections to Berk-Nash Equilibrium; Fudenberg et al. (2021). In particular, the work of Frick et al. (2020) on fragile social learning has a similar flavor to our paper. They show that arbitrarily small misperceptions about the distribution of other player's types can generate large breakdowns in information aggregation, similar to our results on arbitrarily small imperfections in proxies leading to large distortions in beliefs. A recent paper by Bohren and Hauser (2024a) links the misspecified models approach to the literature on biased belief updating rules. An interesting task for future work is to try and use their results to further analyze Proxy Equilibrium as a concept.

We can see this solution concept literature as modeling players whose actions contribute to an long-run steady state distribution of the outcomes of past decisions in the same or similar situations. In contrast, there is a literature modeling players in games as extrapolating from small samples of the equilibrium behavior of other players, the seminal work being Osborne and Rubinstein (1998) and Osborne and Rubinstein (2003). Several recent papers developing similar ideas include Salant and Cherry (2020), Patil and Salant (2024) and Gonçalves (2023).

This paper also connects to a body of work on naive inference from selected observations as a form of decision making bias. This models of sampling investors in Jehiel (2018) and elections with retrospective voters in Esponda and Pouzo (2017). Spiegler (2017) explores a procedure in which an analyst extrapolates from a dataset with partially missing information. Fudenberg et al. (2024) presents an equilibrium concept in which agents have selective recollection of their past experience. In all of these works, agents are considering a partially missing distribution. Under Proxy Equilibrium, data is not fully missing but instead distorted by measurement error.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Alexander Clyde reports financial support was provided by Economic and Social Research Council studentship no. 1923039. Alexander Clyde reports financial support was provided by European Research Council grant no. 692995.

Appendix A. Appendices

A.1. Definition of proxy equilibrium for general spaces

Assume that each dimension of the variable space $V = Y \times X \times S \subseteq \mathbb{R}^{m+2}$ has an appropriate topology such that it is a Polish space.[11] Given the variable space V, denote the Borel σ-algebra of this space by $B(V) \equiv B(Y \otimes X \otimes S) = B(Y) \otimes B(X) \otimes B(S)$.[12] Denote any Borel subset by $\mathcal{V} \equiv \mathcal{Y} \times \mathcal{X} \times \mathcal{S} \in B(V)$.

For any two measures h and g on measure space (Z, Σ), h is *absolutely continuous* with respect to g if for every $A \in \Sigma$, $g(A) = 0$ implies $h(A) = 0$. The two measures are *mutually absolutely continuous* if in addition for any $A \in \Sigma$, $h(A) = 0$ implies $g(A) = 0$. If h is absolutely continuous with respect to υ then by the Radon-Nikodym Theorem there exists a Σ measurable function $f : Z \to [0, \infty)$

[11] E.g. the euclidean topology for a continuous space and the discrete topology for a finite space.

[12] Since the variable spaces are Polish and thus second countable, the Borel product σ-algebra is equal to the product of the Borel σ-algebra for each dimension.

such that for any $A \in \Sigma$, $h(A) = \int_A f\,dg$. We refer to this function as the density of h with respect to g.[13] Given measure spaces (Z_1, Σ_1) and (Z_2, Σ_2), we define a *Markov kernel* as a function $m : Z_1 \times \Sigma_2 \to [0,1]$ such that for every $A_2 \in \Sigma_2$, the map $z_1 \mapsto m(z_1, A_2)$ is measurable with respect to Σ_1 and for every $z_1 \in Z_1$ the map $A_2 \mapsto m(z_1, A_2)$ is a probability measure on (Z_2, Σ_2). For ease of notation from now on we denote such a Markov kernel by $m : Z_1 \to \Delta(Z_2)$.

Define the prior distribution over the signals by a probability measure P_S. Define the strategy as a Markov kernel $\sigma : S \to \Delta(X)$. The joint distribution of $X \times S$ is then given by

$$P_{X,S}(\mathcal{X}, \mathcal{S}) = \int_{\mathcal{S}} \sigma(\mathcal{X}|s) d P_S(s) \tag{24}$$

Define an outcome Markov kernel $P_{Y|X,S} : X \times S \to \Delta(Y)$. The joint distribution over $Y \times X \times S$ is then

$$P(\mathcal{Y}, \mathcal{X}, \mathcal{S}) = \int_{\mathcal{X}\otimes\mathcal{S}} P_{Y|X,S}(\mathcal{Y}|x,s) d P_{X,S}(x,s) \tag{25}$$

for any Borel set $\mathcal{Y} \otimes \mathcal{X} \otimes \mathcal{S} \in B(V)$. Now define the proxy mapping as a Markov kernel $\pi : Y \times X \times S \to \Delta(Y \times X \times S)$. Given the true joint distribution over the variables, the joint distribution over the proxies is

$$P_\pi(\mathcal{V}^\bullet) = \int_{Y\times X\times S} \pi(\mathcal{V}^\bullet|y,x,s) d P(y,x,s) \tag{26}$$

for any $\mathcal{V}^\bullet \in B(V)$. We assume that P_π is mutually absolutely continuous with respect to σ-finite measure μ. This means that we can define a density function p_π such that $p_\pi(y^\bullet, x^\bullet, s^\bullet) > 0$ for all $(y^\bullet, x^\bullet, s^\bullet) \in Y \times X \times S$ and $P_\pi(\mathcal{V}^\bullet) = \int_{\mathcal{V}^\bullet} p_\pi(y^\bullet, x^\bullet, s^\bullet) d\mu$.[14]

We can then define the condition beliefs of the DM under Proxy Equilibrium given the distribution over proxies p_π as before using (8). The perceived utility of an agent with signal s choosing action x is

$$V(x,s;\sigma) = \int_{Y^\bullet} u(y = y^\bullet, x, s) p_\pi(y^\bullet | x^\bullet = x, s^\bullet = s; \sigma) d\mu(y^\bullet) \tag{27}$$

Whereas the expected utility of an agent who formed beliefs using the true variables would be

$$U(x,s) = \int_Y u(y,x,s) d P_{Y|X,S}(y|x,s) \tag{28}$$

A.1.1. Equilibrium

We make the following technical definitions to facilitate the description of Proxy Equilibrium. We say a sequence of strategies $\{\sigma\}_{j=1}^\infty$ *converges* to strategy $\bar{\sigma}$ if for every $s \in S$ the sequence of probability measures $\{\sigma(.|s)\}_{j=1}^\infty$ converges in distribution to the probability measure $\bar{\sigma}(.|s)$. A strategy σ that induces a belief density $p_\pi(y^\bullet, x^\bullet, s^\bullet; \sigma)$ *induces a belief density that has full support* if $p_\pi(y^\bullet, x^\bullet, s^\bullet; \sigma) > 0$ for any realization $(y^\bullet, x^\bullet, s^\bullet) \in Y \times X \times S$.

Definition 5. Let σ_ϵ^* be a strategy mapping that induces a belief density that has full support. For every $s \in S$ define the following set

$$X(s;\sigma_\epsilon^*) \equiv \{x \in X : x \notin \arg\max \int_{Y^\bullet} u(y = y^\bullet, x, s) p_\pi(y^\bullet | x^\bullet = x, s^\bullet = s; \sigma_\epsilon^*) d\mu(y^\bullet)\}$$

Then σ_ϵ^* is an ϵ-**Proxy Equilibrium** if for every Borel subset $\mathcal{X} \in B(X)$, $\mathcal{X} \subseteq X(s;\sigma_\epsilon^*)$, we have that $\sigma_\epsilon^*(\mathcal{X}|s) < \epsilon$.

Definition 6. A strategy σ^* is a **Proxy Equilibrium** if there exists a sequence $\{\sigma_l^*\}_{l=1}^\infty$ converging to σ^* as well as a sequence $\epsilon^l \to 0$, such that for every l, σ_l^* is an ϵ^l-Proxy Equilibrium.

An analogous condition to requiring that π is minimally responsive for this general setting is to require that if the distribution over the true variables P is mutually absolutely continuous with respect to σ-finite measure μ then P_π is also mutually absolutely continuous with respect to μ. This then ensures that if P admits a full support density then so does P_π.

[13] Density is sometimes only used to refer to the Radon-Nikodym derivative when variables are continuous, here for convenience we use it for general variable spaces.

[14] That $\mu(\mathcal{V}) = 0$ implies $P_\pi(\mathcal{V}) = 0$ for any $\mathcal{V} \in B(V)$ means we can define a measurable function $\tilde{p}_\pi : V \to [0,\infty)$ such that $P_\pi(\mathcal{V}) = \int_{\mathcal{V}} \tilde{p}_\pi(v) d\mu(v)$. Let $\mathcal{A} = \{v \in V : \tilde{p}_\pi(v) = 0\}$. As $P_\pi(\mathcal{V}) = 0$ implies $\mu(\mathcal{V}) = 0$ and $P_\pi(\mathcal{A}) = \int_{\mathcal{A}} \tilde{p}_\pi(v) d\mu(v) = 0$ by definition, we have that $\mu(\mathcal{A}) = 0$. Thus we can replace $\tilde{p}_\pi(v)$ with a full support density $p_\pi(v) = \max\{\tilde{p}_\pi(v), \xi\}$ for any $\xi > 0$ such that $P_\pi(\mathcal{V}) = \int_{\mathcal{V}} p_\pi(v) d\mu(v)$ for any $\mathcal{V} \in B(V)$.

A.2. Market entry without equilibrium

The following example demonstrates that without the equilibrium requirement, against an exogenous distribution the extent of excessive entry can decrease with the bandwidth parameter.

Example 2. Let the function $m(.)$ take the following form.

$$m(s) = \begin{cases} \frac{\frac{1}{2}k}{1-k(x-\alpha)} - \frac{1}{2}k & \text{if } s \in [0,\alpha) \\ \frac{\frac{1}{2}k+(x-\alpha)}{1-k(x-\alpha)-(x-\alpha)^2} - \frac{1}{2}k & \text{if } s \in [\alpha,1] \end{cases} \tag{29}$$

This function is increasing in the signal s and continuous. Assume the distribution over signals $p(s)$ is uniform and the conditional distribution $p(.|s^\bullet, x=1)$ is set exogenously to that which would be induced by the strategy $\sigma(x=1|s)=1$ for all s. This means that $p(.|s^\bullet, x=1)$ is uniform in $s \in [s^\bullet - h, s^\bullet + h]$ for each $s^\bullet \in [h, 1-h]$ and uniform on $s \in [0,h)$ and $s \in (1-h,1]$ for $s^\bullet \in [0,h)$ and $s^\bullet \in (1-h,1]$ respectively.

The perceived utility of entry at $s \in [h, 1-h]$ is then $\int_{s-h}^{s+h} m(\tilde{s})d\tilde{s}$. This is increasing in s as $m(.)$ is increasing in s. The cut-off $\bar{s}(h)$ for which entry is a best response is given by $\int_{\bar{s}(h)-h}^{\bar{s}(h)+h} m(\tilde{s})d\tilde{s} = 0$. We can solve this to get

$$\bar{s}(h) = \alpha - h + \frac{1}{2}k\frac{1-exp(2kh)}{exp(2kh)} + \sqrt{(h - \frac{1}{2}k\frac{1-exp(2kh)}{exp(2kh)})^2 - \frac{1-exp(2kh)}{exp(2kh)} - kh\frac{1+exp(2kh)}{exp(2kh)} - h^2} \tag{30}$$

We can find parameters for which there is over-entry but a lesser extent of over-entry when the bandwidth parameter is higher, in contrast to Proposition 4. Let $h_1 = 0.1$, $h_2 = 0.125$, $\alpha = 0.5$, $k = -2$. Then we have $\bar{s}(h_1) \approx 0.497 < \alpha$, $\bar{s}(h_2) \approx 0.499 < \alpha$, and $\bar{s}(h_1) < \bar{s}(h_2)$.

A.3. Proofs

Proof of Proposition 1

Proof. Denote the set of all strategies conditional on signal s as $\Sigma(s)$. For any $\xi > 0$ we can define a strategy that has full support; $\sigma_\xi(x|s) > 0$ for all $(x,s) \in X \times S$. By the minimal responsiveness property of π, this means beliefs $p_\pi(.|x^\bullet = x, s^\bullet = s; \hat{\sigma}_\xi)$ are well defined for all $(x,s) \in X \times S$. We can then define the following best response correspondence, given strategy $\hat{\sigma}_\xi$ and $\xi > 0$:

$$BR_\xi(\hat{\sigma}_\xi, s) = \{ \underset{\sigma(.|s)\in\Sigma(s)}{\arg\max} \sum_{x\in X} \sigma(x|s) \sum_{y\in Y} u(y=y^\bullet, x, s) p_\pi(y^\bullet|x^\bullet = x, s^\bullet = s; \hat{\sigma}_\xi) \text{ s.t } \sigma(x'|s) \geq \xi \ \forall x' \in X \}$$

Stack the best response correspondences into $BR_\xi(\hat{\sigma}) = \prod_{s\in S} BR_\xi(\hat{\sigma}, s)$. Since $p_\pi(y^\bullet|x^\bullet = x, s^\bullet = s; \tilde{\sigma})$ is continuous in $\tilde{\sigma}$ and the best response correspondence is the set of maximizers over a compact set defined by a finite set of inequalities, $BR_\xi(\hat{\sigma})$ is nonempty for any $\hat{\sigma}$. Moreover due to linearity in $\sigma(x|s)$, $BR_\xi(.)$ convex valued and continuity of $p_\pi(y^\bullet|x^\bullet = x, s^\bullet = s; \tilde{\sigma})$ implies $BR_\xi(.)$ has closed graph. We therefore have met all the requirements of Kakutani's fixed point theorem and a fixed point exists for any $\xi > 0$, $\sigma^*_\xi \in BR_\xi(\sigma^*_\xi)$.

For any $\epsilon > 0$, we can choose $\xi > 0$ in such a way that ensures that our ξ-fixed point is an ϵ-Proxy Equilibrium. For any $s \in S$ and σ, define

$$X(s;\sigma) \equiv \{ x \notin \underset{x'\in X}{\arg\max} \sum_{y\in Y} u(y=y^\bullet, x', s) p_\pi(y^\bullet|x^\bullet = x', s^\bullet = s; \sigma) \}$$

We have that $\sigma^*_\xi(x|s) = \xi$ for all $x \in X(s;\sigma^*_\xi)$. Therefore, we can choose $\xi > 0$ to ensure that $\sigma^*_\xi(x|s) = \xi < \epsilon$ for all $s \in S$. This ensures our fixed point, which we denote σ^*_ϵ, meets the definition of ϵ-Proxy Equilibrium.

Since finiteness ensures the space of strategies Σ is compact, we can find a convergent sequence of ϵ-equilibria as $\epsilon \to 0$, $\sigma^*_\epsilon \to \sigma^*$. □

Proof of Proposition 2

Proof. We first propose a general linear solution $x(s) = \theta_0 + \theta_1 s$, which is then used to calculate the perceived expectation $\mathbb{E}[y|x^\bullet = x]$ using the properties of the normal distribution. Under the proposed best response function, the joint normal distribution of $(y, x^\bullet)$ is

$$\begin{pmatrix} y \\ x^\bullet \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \alpha - \beta(\theta_0 + \theta_1 \mu_s) \\ \theta_0 + \theta_1 \mu_s \end{pmatrix}, \begin{pmatrix} \beta^2\theta_1^2\sigma_s^2 + \sigma_u^2 & -\beta\theta_1^2\sigma_s^2 \\ -\beta\theta_1^2\sigma_s^2 & \theta_1^2\sigma_s^2 + \sigma_\epsilon^2 \end{pmatrix}\right)$$

Using this we can calculate the conditional expectation of y given $x^\bullet$.

$$\mathbb{E}[y|x^\bullet] = \alpha - \beta(\theta_0 + \theta_1 \mu_s) - \frac{\beta\theta_1^2\sigma_s^2}{\theta_1^2\sigma_s^2 + \sigma_\epsilon^2}(x^\bullet - \theta_0 - \theta_1\mu_s)$$

Using the utility function we then get perceived expected utility $V(x, s; \sigma) = -s\mathbb{E}[y|x^\bullet = x] - \frac{1}{2}x^2$. Solving for a maximum then gives us $x(s) = \frac{\beta\sigma_s^2\theta_1^2}{\theta_1^2\sigma_s^2+\sigma_\epsilon^2} \cdot s$. In order to have a linear equilibria, we must therefore have $\theta_0 = 0$ and $\theta_1 = \frac{\beta\sigma_s^2\theta_1^2}{\theta_1^2\sigma_s^2+\sigma_\epsilon^2}$. We can solve the latter cubic equation to get the equilibria in the statement of the proposition. □

Proof of Proposition 3

We use the version of Proxy Equilibrium defined for general spaces in Appendix A.1. The window form of proxy noise in the signal can be described using a signal specific proxy mapping.

$$\pi_s(s^\bullet|s) = \frac{1}{2h}\mathbb{1}\{[s^\bullet \in [\max\{\min\{s-h, 1-2h\}, 0\}, \min\{\max\{s+h, 2h\}, 1\}]]\} \tag{31}$$

Given y and x are perfectly measured, using the Dirac delta mapping π_δ for any Borel Set $\mathcal{Y}^\bullet \times \mathcal{X}^\bullet \times \mathcal{S}^\bullet \subseteq Y \times X \times S$ we can write the proxy mapping

$$\pi(\mathcal{Y}^\bullet \times \mathcal{X}^\bullet \times \mathcal{S}^\bullet|y, x, s) = \int_{\mathcal{S}^\bullet} \pi_\delta(\mathcal{Y}^\bullet \times \mathcal{X}^\bullet|y, x)\pi_s(s^\bullet|s)d\mu \tag{32}$$

The distribution over proxies then admits a density

$$p_\pi(y, x, s^\bullet) = \int_S \pi_s(s^\bullet|s)p(y|x, s)\sigma(x|s)p(s)d\mu(s)$$

as using Fubini's theorem and the definition of π_δ we have that for any $\mathcal{Y} \times \mathcal{X} \times \mathcal{S}^\bullet$

$$\begin{aligned}
&P_\pi(\mathcal{Y} \times \mathcal{X} \times \mathcal{S}^\bullet) \\
&= \int_{\mathcal{S}^\bullet}[\int_{Y\times X\times S} \pi_\delta(\mathcal{Y} \times \mathcal{X}|y, x)\pi_s(s^\bullet|s)p(y|x)\sigma(x|s)p(s)d\mu(y, x, s)]d\mu(s^\bullet) \\
&= \int_{\mathcal{S}^\bullet\times S}[\int_{Y\times X} \pi_\delta(\mathcal{Y} \times \mathcal{X}|y, x)p(y|x)\sigma(x|s)d\mu(y, x)]\pi_s(s^\bullet|s)p(s)d\mu(s^\bullet, s) \\
&= \int_{\mathcal{S}^\bullet\times S}[\int_{\mathcal{Y}\times\mathcal{X}} p(y|x, s)\sigma(x|s)d\mu(y, x)]\pi_s(s^\bullet|s)p(s)d\mu(s^\bullet, s) \\
&= \int_{\mathcal{Y}\times\mathcal{X}\times\mathcal{S}^\bullet}[\int_S \pi_s(s^\bullet|s)p(y|x, s)\sigma(x|s)p(s)d\mu(s)]d\mu(y, x, s^\bullet) \\
&= \int_{\mathcal{Y}\times\mathcal{X}\times\mathcal{S}^\bullet} p_\pi(y, x, s^\bullet)d\mu(y, x, s^\bullet)
\end{aligned}$$

Given an induced perceived distribution over the outcome variable $p_\pi(y|x, s^\bullet; \sigma)$, the DMs perceived expected utility is

$$\begin{aligned}
V(x = 1, s; \sigma) &= \int_Y yp_\pi(y|x = 1, s^\bullet = s; \sigma)d\mu(y) \\
&= \int_Y y[\int_S p(y|x = 1, \tilde{s})p_\pi(\tilde{s}|x = 1, s^\bullet = s; \sigma)d\mu(\tilde{s})]d\mu(y) \\
&= \int_S [\int_Y yp(y|x = 1, \tilde{s})d\mu(y)]p_\pi(\tilde{s}|x = 1, s^\bullet = s; \sigma)d\mu(\tilde{s})
\end{aligned}$$

$$= \int_S m(\tilde{s}) p_\pi(\tilde{s}|x=1, s^\bullet = s; \sigma) d\mu(\tilde{s}) \tag{33}$$

The perceived utility of $x=0$ at s is always zero; $V(x=0, s;\sigma)=0$. We can see that the perceived utility depends on the distribution $p_\pi(\tilde{s}|x=1, s^\bullet = s;\sigma)$ induced by the strategy σ. From the distribution over proxies, this can be calculated as follows.

$$p_\pi(s|x=1, s^\bullet;\sigma) = \frac{\pi_s(s^\bullet|s)\sigma(x=1|s)p(s)}{\int_S \pi_s(s^\bullet|\hat{s})\sigma(x=1|\hat{s})p(\hat{s})d\mu(\hat{s})} \tag{34}$$

We demonstrate the following fact which is used several times in the proof.

Lemma A.1. *Let $[a_1, b_1]$, $[a_2, b_2]$ be intervals in $[0,1]$, with $a_1 > a_2$ and $b_1 > b_2$. Then for any $s_1 > s_2$ we have that*

$$\mathbb{1}\{s_1 \in [a_1, b_1]\} \cdot \mathbb{1}\{s_2 \in [a_2, b_2]\} \geq \mathbb{1}\{s_2 \in [a_1, b_1]\} \cdot \mathbb{1}\{s_1 \in [a_2, b_2]\} \tag{35}$$

Moreover, this inequality holds strictly if $s_1 \in [a_1, b_1] \setminus [a_2, b_2]$ or $s_2 \in [a_2, b_2] \setminus [a_1, b_1]$.

These facts also hold for half-open intervals $[a_1, b_1)$, $[a_2, b_2)$, $(a_1, b_1]$ and $(a_2, b_2]$.

Proof. For the right hand side of the inequality to be equal to one requires $s_2 < s_1 \leq b_2$, $s_2 \geq a_1 > a_2$, $s_1 > s_2 \geq a_1$ and $s_1 \leq b_2 < b_1$ so $s_2 \in [a_2, b_2]$ and $s_1 \in [a_1, b_1]$ and the left hand side is also equal to one.

The second part of the result holds by definition and the third part is clear by applying the arguments above again. □

We can then show an increasing best response property.

Lemma A.2. *Given a perceived distribution over outcomes induced by a full-support strategy σ, we have that $V(x=1, s;\sigma)$ is strictly increasing in $s \in [h, 1-h]$.*

Proof. For a given induced distribution p_π, we can use integration by parts to write the perceived utility of the DM as follows.

$$V(x=1, s;\sigma) = \int_0^1 m(\tilde{s}) p_\pi(\tilde{s}|s^\bullet = s, x=1) d\mu(\tilde{s}) = m(1) - \int_0^1 P_\pi(\tilde{s}|s^\bullet = s, x=1) dM(\tilde{s})$$

Where $P_\pi(\tilde{s}|s^\bullet, x=1)$ is the cdf of the induced distribution and M is the Lebesgue-Stieltjes measure satisfying $M((s_l, s_h]) = m(s_h) - m(s_l)$ for any $0 \leq s_l < s_h \leq 1$. Since $m(.)$ is strictly increasing and right continuous, this measure exists. Therefore, to show the result it is enough to show $P_\pi(s|s_1^\bullet, x=1) \leq P_\pi(s|s_2^\bullet, x=1)$ for any $1-h \geq s_1^\bullet > s_2^\bullet \geq h$ and all s, with strict inequality for all s in some interval $[a,b] \subseteq [0,1]$. Our assumptions about the conditional distribution of the proxies give us the following sequence of claims.

By Lemma A.1, we have that for any $s_1^\bullet > s_2^\bullet$ and $1-h \geq s_1 > s_2 \geq h$.

$$\mathbb{1}\{s_1^\bullet \in [s_1 - h, s_1 + h]\} \cdot \mathbb{1}\{s_2^\bullet \in [s_2 - h, s_2 + h]\} \geq \mathbb{1}\{s_2^\bullet \in [s_1 - h, s_1 + h]\} \cdot \mathbb{1}\{s_1^\bullet \in [s_2 - h, s_2 + h]\}$$

With strict inequality if $s_1^\bullet \in [s_1 - h, s_1 + h] \setminus [s_2 - h, s_2 + h]$ or $s_2^\bullet \in [s_2 - h, s_2 + h] \setminus [s_1 - h, s_1 + h]$.

Multiplying both sides by $\frac{\frac{1}{2h} p(s_1)\sigma(x=1|s_1)}{\int_S p(\hat{s})\sigma(x=1|\hat{s})\pi(s_1^\bullet|\hat{s})d\mu(\hat{s})} \cdot \frac{\frac{1}{2h} p(s_2)\sigma(x=1|s_2)}{\int_S p(\hat{s})\sigma(x=1|\hat{s})\pi(s_2^\bullet|\hat{s})d\mu(\hat{s})}$, we can then write the following

$$p_\pi(s_1|s_1^\bullet, x=1) p_\pi(s_2|s_2^\bullet, x=1) \geq p_\pi(s_1|s_2^\bullet, x=1) p_\pi(s_2|s_1^\bullet, x=1)$$

$$\Rightarrow \int_{\tilde{s}}^1 p_\pi(s_1|s_1^\bullet, x=1) ds_1 \cdot \int_0^{\tilde{s}} p_\pi(s_2|s_2^\bullet, x=1) ds_2 \geq \int_{\tilde{s}}^1 p_\pi(s_1|s_2^\bullet, x=1) ds_1 \cdot \int_0^{\tilde{s}} p_\pi(s_2|s_1^\bullet, x=1) ds_2$$

$$\Rightarrow P_\pi(s|s_1^\bullet, x=1) \leq P_\pi(s|s_2^\bullet, x=1)$$

for any $s_1^\bullet > s_2^\bullet$ and s. By the strict inequality case above, we have that $P_\pi(s|s_1^\bullet, x=1) < P_\pi(s|s_2^\bullet, x=1)$ for any $s \in [s_1^\bullet - h, s_1^\bullet + h] \cup [s_2^\bullet - h, s_2^\bullet + h]$. This completes the proof. □

An outline of the proof is as follows. We first show why any Proxy Equilibrium with partial entry must have a cut-off structure, and give a condition that the cut-off must satisfy in terms of perceived expected utility. We then show how we can construct a sequence of ϵ-Proxy Equilibria that converge to this cut-off structure.

We consider partial entry Proxy Equilibrium in which $\sigma(x=1|s) > 0$ at a strict subset of $s \in S^> \subset [0,1]$. For such an equilibrium to exist, we must have a sequence of ϵ^l-Proxy Equilibria, $\{\sigma_{\epsilon^l}^*\}_{l=1}^\infty$ that converge in distribution to it. Given we are considering partial entry Proxy Equilibria, for large enough l we must have that the induced belief p_π^l in the ϵ^l-Proxy Equilibrium in the sequence is such that $U(s, x=1; p_\pi^l) \leq 0$ for all $s \in [0, h]$, and that $x=1$ is a best response to p_π^l for some s. By the fact perceived utility is increasing in $s \in [h, 1-h]$, from Lemma A.2, and that a partial entry equilibrium must have no entry at some signal, there must be a cut-off $\bar{s}^{\epsilon^l} \in [h, 1]$ such that a best response is $x=0$ for $s \in [0, \bar{s}^{\epsilon^l}]$ and $x=1$ for $s \in (\bar{s}^{\epsilon^l}, 1]$.

By the definition ϵ^l-Proxy Equilibrium, we must have that $\sigma^*_{\epsilon^l}(x=1|s)<\epsilon^l$ for all $s\in[0,\bar{s}^{\epsilon^l}]$ and $\sigma^*_{\epsilon^l}(x=1|s)\geq 1-\epsilon^l$ for all $s\in(\bar{s}^{\epsilon^l},1]$. The perceived utility of the DM at this cut-off $\bar{s}^{\epsilon^l}$ is then

$$\int\limits_{\bar{s}^{\epsilon^l}-h}^{\bar{s}^{\epsilon^l}} m(\tilde{s})\frac{\sigma_\epsilon(x=1|\tilde{s})p(\tilde{s})}{\int_{\bar{s}^{\epsilon^l}-h}^{\bar{s}^{\epsilon^l}+h}\sigma_\epsilon(x=1|\hat{s})p(\hat{s})d\mu(\hat{s})}d\mu(\tilde{s})+\int\limits_{\bar{s}^{\epsilon^l}}^{\bar{s}^{\epsilon^l}+h} m(\tilde{s})\frac{\sigma_\epsilon(x=1|\tilde{s})p(\tilde{s})}{\int_{\bar{s}^{\epsilon^l}-h}^{\bar{s}^{\epsilon^l}+h}\sigma_\epsilon(x=1|\hat{s})p(\hat{s})d\mu(\hat{s})}d\mu(\tilde{s})=0$$

Thus as $l\to\infty$, if our sequence of ϵ-Proxy Equilibria converges it will converge to a Proxy Equilibrium with cut-off $\bar{s}^*$ such that $\sigma^*(x=1|s)=0$ for $s\in[0,\bar{s}^*]$ and $\sigma^*(x=1|s)=1$ for $s\in(\bar{s}^*,1]$. The perceived utility at the cut-off $\bar{s}^*$ will then be

$$\int\limits_{\bar{s}}^{\bar{s}+h} m(\tilde{s})\frac{p(\tilde{s})}{\int_{\bar{s}}^{\bar{s}+h}p(\hat{s})d\mu(\hat{s})}d\mu(\tilde{s})=0 \tag{36}$$

We then construct strategies that can form a sequence of ϵ-Proxy Equilibria that converge to a partial entry Proxy Equilibria. These strategies have a cut-off form where $\sigma_\xi(x=1|s)=\xi\in(0,\frac{1}{2})$ for $s\in[0,\bar{s}]$ and $\sigma_\xi(x=1|s)=1-\xi\in(\frac{1}{2},1)$ for $s\in(\bar{s},1]$, with $\bar{s}\in[0,1]$ as the cut-off. We can then define the following conditional density over $s\in[h,1-h]$ given $s^\bullet=\bar{s}$.

$$g_\xi(s|s^\bullet=\bar{s})=\frac{(1-\xi)\mathbb{1}\{\tilde{s}\in(\bar{s},\bar{s}+h]\}+\xi\mathbb{1}\{\tilde{s}\in[\bar{s}-h,\bar{s}]\}}{(1-\xi)\int_{\bar{s}}^{\bar{s}+h}p(\hat{s})d\mu(\hat{s})+\xi\int_{\bar{s}-h}^{\bar{s}}p(\hat{s})d\mu(\hat{s})}p(\tilde{s}) \tag{37}$$

For any cut-off $\bar{s}\in[h,1-h]$ and k, we can choose:

$$\xi(\bar{s},k)=\frac{k\int_{\bar{s}}^{\bar{s}+h}p(\tilde{s})d\mu(\tilde{s})}{k\int_{\bar{s}}^{\bar{s}+h}p(\tilde{s})d\mu(\tilde{s})+(1-k)\int_{\bar{s}-h}^{\bar{s}}p(\tilde{s})d\mu(\tilde{s})} \tag{38}$$

Which is arbitrarily small for small enough $1>k>0$. This ensures that

$$\int\limits_{\bar{s}-h}^{\bar{s}} g_\xi(\tilde{s}|s^\bullet=\bar{s})d\mu(\tilde{s})=\frac{\xi(\bar{s},k)\int_{\bar{s}-h}^{\bar{s}}p(\tilde{s})d\mu(\tilde{s})}{(1-\xi(\bar{s},k))\int_{\bar{s}}^{\bar{s}+h}p(\tilde{s})d\mu(\tilde{s})+\xi(\bar{s},k)\int_{\bar{s}-h}^{\bar{s}}p(\tilde{s})d\mu(\tilde{s})}=k$$

We can then write the perceived utility at $\bar{s}\in[h,1-h]$ against the beliefs induced by strategy $\sigma_{\xi(\bar{s},k)}$ with cut-off $\bar{s}\in[h,1-h]$ in the following way

$$\int\limits_0^1 m(\tilde{s})g_{\xi(\bar{s},k)}(\tilde{s}|s^\bullet=\bar{s})d\mu(\tilde{s})=(1-k)\overline{U}(\bar{s},x=1;\bar{s})+k\underline{U}(\bar{s},x=1;\bar{s}) \tag{39}$$

Which is a linear combination of the terms

$$\overline{U}(\bar{s},x=1;\bar{s})=\int\limits_{\bar{s}}^{\bar{s}+h} m(\tilde{s})\frac{p(\tilde{s})}{\int_{\bar{s}}^{\bar{s}+h}p(\hat{s})d\mu(\hat{s})}d\mu(\tilde{s}) \tag{40}$$

$$\underline{U}(\bar{s},x=1;\bar{s})=\int\limits_{\bar{s}-h}^{\bar{s}} m(\tilde{s})\frac{p(\tilde{s})}{\int_{\bar{s}-h}^{\bar{s}}p(\hat{s})d\mu(\hat{s})}d\mu(\tilde{s}) \tag{41}$$

We show that (39) is strictly increasing in $\bar{s}\in[h,1-h]$ by showing (40) and (41) are strictly increasing in $\bar{s}$.

Lemma A.3. *The expressions $\overline{U}(\bar{s},x=1;\bar{s})$ and $\underline{U}(\bar{s},x=1;\bar{s})$ are strictly increasing for all $\bar{s}\in[h,1-h]$.*

Proof. We define densities $\overline{g}(s;\bar{s})=\frac{1_{[s\in[\bar{s},\bar{s}+h]]}p(s)}{\int_{\bar{s}}^{\bar{s}+h}p(\tilde{s})d\mu(\tilde{s})}$ and $\underline{g}(s;\bar{s})=\frac{1_{[s\in[\bar{s}-h,\bar{s})]}p(s)}{\int_{\bar{s}-h}^{\bar{s}}p(\hat{s})d\mu(\hat{s})}$. We can then use the same steps as in the proof of Lemma A.2, applied with indicator functions $\mathbb{1}\{s\in(\hat{s},\hat{s}+h]\}$ and $\mathbb{1}\{s\in[\hat{s}-h,\hat{s}]\}$ instead of $\mathbb{1}\{s\in[\hat{s}-h,\hat{s}+h]\}$, to prove the result. □

With these results in hand, we can then both show existence of and characterize the equilibria for this application.

Proposition 3

Proof. At any ϵ-Proxy Equilibrium, the perceived utility of the DM is increasing strictly for $s\in[h,1-h]$ by Lemma A.2. The structure of the window form of proxy mapping means that the beliefs of the DM are identical on $s\in[0,h]$. If the DM is mixing $\sigma_\epsilon(x=1|s)>\epsilon$ on $s\in[0,h]$, then due to increasing expected payoff on $s\in[h,1-h]$, they must be playing $\sigma_\epsilon(x=1|s)\geq 1-\epsilon$ on $s\in(h,1]$. As $\epsilon\to 0$ and $\sigma_\epsilon\to\sigma$ their perceived utility at any $s\in[0,h]$ given potential equilibrium σ is

$$\int_0^h m(s)\frac{\sigma(x=1|s)p(s)}{\int_0^1 \sigma(x=1|\tilde{s})p(\tilde{s})d\mu(\tilde{s})}d\mu(s)+\int_h^{2h} m(s)\frac{p(s)}{\int_0^1 \sigma(x=1|\tilde{s})p(\tilde{s})d\mu(\tilde{s})}d\mu(s)$$

By the assumption that $h < \min\{\frac{\alpha}{2}, \frac{1-\alpha}{2}\}$ and m is increasing, the above expression is strictly negative. Thus for small enough ϵ at any ϵ-Proxy Equilibrium we must have $\sigma(x=1|s) < \epsilon$ for $s \in [0,h]$.

From this argument and Lemma A.2, any ϵ-Proxy Equilibria in which $\sigma(x=1|s) \geq \epsilon$ for some s must have some cut-off $\bar{s} \in (h,1]$ such that $\sigma(x=1|s) \geq \epsilon$ only if $s > \bar{s}$. The assumption $h < \min\{\frac{\alpha}{2}, \frac{1-\alpha}{2}\}$ means $\int_{1-h}^1 m(s)p(s)d\mu(s) > 0$ which ensures that any potential equilibrium strategy with cut-off $\bar{s} \geq 1-h$ will have $\sigma(x=1|s)=1$ as a best response for all $s \in [1-h,1]$, and thus $\sigma(x=0|s) < \epsilon$ for $s \in [1-h,1]$ in any ϵ-Proxy Equilibrium.

We construct the following cut-off ϵ-Proxy Equilibrium strategy. For any cut-off $\bar{s} \in (h,1]$, $\epsilon > 0$ and $k_\epsilon \in (0,1)$, define $\xi(\bar{s},k_\epsilon)$ as in (38). Let $\sigma_\epsilon(x=1|s) = \xi(\bar{s},k_\epsilon)$ on $s \in [0,\bar{s}]$ and $\sigma_\epsilon(x=1|s) = 1-\xi(\bar{s},k_\epsilon)$ on $s \in (\bar{s},1]$. We choose $k_\epsilon > 0$ small enough such that $\epsilon > \sup_{\bar{s}\in[h,1-h]} \xi(\bar{s},k_\epsilon)$. Then if we can find a $\bar{s}^*$ such that a best response to the beliefs induced by σ_ϵ is $\sigma(x=1|s)=0$ on $s \in [0,\bar{s}^*]$ and $\sigma(x=1|s)=1$ on $s \in (\bar{s}^*,1]$ we have an ϵ- Proxy Equilibrium.

We have shown that the constructed strategy induces the beliefs at the cut-off $\bar{s}$ according to equation (39). We have also shown in Lemma A.3 that this expression is strictly increasing in the cut-off $\bar{s} \in [h,1-h]$. Moreover, we have that as $\epsilon \to 0$, $k_\epsilon \to 0$, so this expression converges to that in equation (36). We have (36) is strictly negative at $\bar{s}=h$ and strictly positive at $\bar{s}=1-h$. Thus we can find a small enough $\epsilon > 0$ and hence $k_\epsilon > 0$ such that $\int_0^1 m(\tilde{s})g_{\xi(\bar{s},k_\epsilon)}(\tilde{s}|s^\bullet = h)d\mu(\tilde{s}) < 0$ and $\int_0^1 m(\tilde{s})g_{\xi(\bar{s},k_\epsilon)}(\tilde{s}|s^\bullet = 1-h)d\mu(\tilde{s}) > 0$. Since $\int_0^1 m(\tilde{s})g_{\xi(\bar{s},k_\epsilon)}(\tilde{s}|s^\bullet = \bar{s})d\mu(\tilde{s})$ is continuous and increasing in $\bar{s} \in [h,1-h]$, we can find a $\bar{s}=\bar{s}^*$ at which it is equal to zero by the intermediate value theorem. This $\bar{s}^*$ then gives us our ϵ-Proxy Equilibrium cut-off as stated above. As $\epsilon \to 0$, we can find a sequence of ϵ-Proxy Equilibria of this form that converge to that in the statement of the proposition.

Next, for contradiction consider that $\bar{s}^* \geq \alpha$. We have shown the cut-off $\bar{s}^*$ must solve the following equation.

$$\overline{U}(\bar{s}^*, x=1;\bar{s}^*) = \int_{\bar{s}^*}^{\bar{s}^*+h} m(\tilde{s})\frac{p(\tilde{s})}{\int_{\bar{s}^*}^{\bar{s}^*+h} p(\hat{s})d\mu(\hat{s})}d\mu(\tilde{s}) = 0$$

Then we have $\overline{U}(\bar{s}^*, x=1;\bar{s}^*) > 0$ as all the probability weight in the distribution is in $s \in [\alpha,1]$, a contradiction. Thus the cut-off must be such that $\bar{s} < \alpha$.

For the final part of the proposition, we can always find a sequence of ϵ-Proxy Equilibria that converges to a Proxy Equilibrium with $x=0$ for all $s \in [0,1]$. For example, with small enough $\epsilon > 0$ we can have an ϵ-Proxy Equilibrium such that the DM plays $x=1$ with probability $\epsilon > 0$ on $[0,\alpha)$ and probability ϵ^2 on $[\alpha,1]$. This induces beliefs to which $x=0$ is a best response for all s. □

Proof of Proposition 4

Proof. We are comparing the cut-off equilibrium at h_1 with the cut-off equilibrium at h_2. Consider the perceived utility at the cut-off under the equilibrium with noise h_1.

$$\overline{U}(\bar{s}(h_1), x=1;h_1) = \int_{\bar{s}(h_1)}^{\bar{s}(h_1)+h_1} m(\tilde{s})\frac{p(\tilde{s})}{\int_{\bar{s}}^{\bar{s}(h_1)+h_1} p(\hat{s})d\mu(\hat{s})}d\mu(\tilde{s})$$

As this is an equilibrium cut-off, we must have that

$$\int_{\alpha}^{\bar{s}(h_1)+h_1} m(\tilde{s})p(\tilde{s})d\mu(\tilde{s}) + \int_{\bar{s}(h_1)}^{\alpha} m(\tilde{s})p(\tilde{s})d\mu(\tilde{s}) = 0$$

If $\bar{s}(h_1)$ is fixed, then as h_1 increases to h_2, the first part of this expression that has weight on the positive part of the function $m(.)$ increases while the second part stays fixed. Thus the perceived utility at cut-off $\bar{s}(h_1)$ when the perceived distribution is induced by a strategy with cut-off $\bar{s}(h_1)$, must become positive at noise parameter $h_2 > h_1$. We have that $\overline{U}(\bar{s}(h_1), x=1;h_2) > 0$, $\overline{U}(h_1, x=1;h_2) < 0$ and $\overline{U}(\bar{s}, x=1;h_2)$ is continuous in $\bar{s} \in [h_1, \bar{s}(h_1)]$. Therefore by the intermediate value theorem we can find a new cut-off $\bar{s}(h_2) < \bar{s}(h_1)$ that characterizes the positive entry equilibrium under noise parameter h_2. □

Proof of Proposition 5

Enumerate the control, action and the $l-1$ control signals as 1, 2 and {3, ...,l+1} respectively. Then we can denote any subset of the variable space $M \equiv \{1,2,...,l+1\}$ by $N \subseteq M$. We write the probability over the subset of true variables in N as p^N and the subset of the proxy variables in N as p_π^N. For ease of notation in this section we suppress dependence of these distributions on the strategy σ. We can relate the distribution over all variables and the distribution over a subset N. The variable space containing the variables in any subset N is denoted as V_N and the variables not in N by V_{-N}. Denote $\pi(w_N^\bullet \times \{V_{-N}\}|w) = \sum_{w_{-N}^\bullet \in V_{-N}} \pi(w_N^\bullet, w_{-N}^\bullet|w)$ for any $w \in Y \times X \times S$ and remember that the perfect measurement mapping is such that $\pi(w_M^\bullet \times \{V_{-N}\}|w) = \mathbb{1}\{w_M^\bullet = w_M\}$.

$$p_\pi^N(w_N^\bullet) = p_\pi(w_N^\bullet \times \{V_{-N}\}) = \sum_{w\in Y\times X\times Z} \pi(w_N^\bullet \times \{V_{-N}\}|w)p(w) \tag{42}$$

Lemma A.4. *For any $\eta > 0$ and for any subset of the variables $N \subseteq M$, if the proxy mapping is η-close to perfect then*

$$TV(p^N, p_\pi^N) \le TV(p^M, p_\pi^M) < \eta \tag{43}$$

Proof. We have that

$$\begin{aligned} TV(p^N, p_\pi^N) &= \max_{A\in 2^{V^N}} |\sum_{a\in A}(p^N(a) - p_\pi^N(a))| \\ &= \max_{A\in 2^{V^N}} |\sum_{a\in A}\sum_{w\in Y\times X\times S}(\pi_\delta(a\times\{V_{-N}\}|w) - \pi(a\times\{V_{-N}\}|w))p(w)| \\ &\le \max_{\tilde{A}\in 2^{V^M}} |\sum_{\tilde{a}\in \tilde{A}}\sum_{w\in Y\times X\times S}(\pi_\delta(\tilde{a}\times\{V_{-M}\}|w) - \pi(\tilde{a}\times\{V_{-M}\}|w))p(w)| \\ &= TV(p^M, p_\pi^M) \\ &\le \max_{\tilde{A}\in 2^V}\sum_{\tilde{a}\in\tilde{A}}\sum_{w\in Y\times X\times S}|\pi_\delta(\tilde{a}\times\{V_{-M}\}|w) - \pi(\tilde{a}\times\{V_{-M}\}|w)|p(w) \\ &< \sum_{w\in Y\times X\times S}\eta p(w) = \eta \quad \square \end{aligned}$$

Proposition 5

Proof. For any $(y, x, s_c) \in Y \times X \times S_c$, we have that $|p(y, x, s_c) - p_\pi(y, x, s_c)| \le TV(p^M, p_\pi^M)$ and for any $(x, s_c) \in X \times S_c$ we have that $|p^M(x, s) - p_\pi^M(x, s)| \le TV(p^M, p_\pi^M)$.

Therefore for any $\eta > 0$, by Lemma A.4 if π is η-close to perfect then $|p(y, x, s_c) - p_\pi(y, x, s_c)| < \eta$ for all $(y, x, s_c) \in Y \times X \times S_c$ and $|p(x, s_c) - p_\pi(x, s_c)| < \eta$ for all $(x, s_c) \in X \times S_c$. As $p_\pi(y|x, s_c) = \frac{p_\pi(y,x,s_c)}{p_\pi(x,s_c)}$ and $(x, s_c) \in V^+(\sigma)$, by the algebra of limits we have the result. $\square$

Proof of Proposition 6

Proof. Given a prospective strategy σ, $XS_c^+(\sigma) = \{(x, s_c) \in X \times S_c : \sigma(x|s_c)p(s_c) > 0\}$. We first prove the following Lemma.

Lemma A.5. *For any $\epsilon > 0$, we can find an $\eta > 0$ such that if π is η-close to perfect then*

$$\max_{(x,s_c,s_p)\in XS_c^+(\sigma)\times S_p} |\sum_{y\in Y} u(y, x, s)[p_\pi(y^\bullet = y|x^\bullet = x, s^\bullet = s; \sigma) - p(y|x, s_c)]| < \epsilon$$

Proof. Define $\bar{u} = max_{(x,s)\in X\times S}|\sum_{y\in Y} u(y, x, s)|$. If $\bar{u} = 0$ then $\sum_{y\in Y} u(y, x, s) = 0$ for all $(x, s) \in X \times S$. By Proposition 5, for any $\epsilon > 0$ we can find an $\eta > 0$ such that if π is η-close to perfect then $\max_{(y,x,s_c)\in Y\times XS_c^+(\sigma)} |p(y|x, s_c) - p_\pi(y|x, s; \sigma)| < \epsilon$. When $\bar{u} = 0$, this then implies

$$\begin{aligned} &\sum_{y\in Y} u(y, x, s)(p(y|x, s_c) - \epsilon) < \sum_{y\in Y} u(y, x, s)p_\pi(y^\bullet = y|x^\bullet = x, s^\bullet = s; \sigma) < \sum_{y\in Y} u(y, x, s)(p(y|x, s_c) + \epsilon) \\ &\Rightarrow \sum_{y\in Y} u(y, x, s)p(y|x, s_c) < \sum_{y\in Y} u(y, x, s)p_\pi(y^\bullet = y|x^\bullet = x, s^\bullet = s; \sigma) < \sum_{y\in Y} u(y, x, s)p(y|x, s_c) \\ &\Rightarrow \sum_{y\in Y} u(y, x, s)p(y|x, s_c) - \epsilon < \sum_{y\in Y} u(y, x, s)p_\pi(y^\bullet = y|x^\bullet = x, s^\bullet = s; \sigma) < \sum_{y\in Y} u(y, x, s)p(y|x, s_c) + \epsilon \end{aligned}$$

Therefore consider $\bar{u} > 0$. Again by Proposition 5, for any $\frac{\epsilon}{\bar{u}} > 0$ we can find an $\eta > 0$ such that if π is η-close to perfect then $\max_{(y,x,s_c)\in Y\times XS_c^+(\sigma)} |p(y|x, s_c) - p_\pi(y|x, s; \sigma)| < \frac{\epsilon}{\bar{u}}$. Then as required we have that for any $(x, s_c.s_p) \in XS_c^+(\sigma)\times S_p$, as $\sum_{y\in Y} u(y, x, s) \le \bar{u}$

$$\begin{aligned} &\sum_{y\in Y} u(y, x, s)p(y|x, s_c) - \epsilon\frac{\sum_{y\in Y} u(y, x, s)}{\bar{u}} < \sum_{y\in Y} u(y, x, s)p_\pi(y^\bullet = y|x^\bullet = x, s^\bullet = s; \sigma) < \sum_{y\in Y} u(y, x, s)p(y|x, s_c) + \epsilon\frac{\sum_{y\in Y} u(y, x, s)}{\bar{u}} \\ &\Rightarrow \sum_{y\in Y} u(y, x, s)p(y|x, s_c) - \epsilon < \sum_{y\in Y} u(y, x, s)p_\pi(y^\bullet = y|x^\bullet = x, s^\bullet = s; \sigma) < \sum_{y\in Y} u(y, x, s)p(y|x, s_c) + \epsilon \quad \square \end{aligned}$$

⇐ (Necessity): We take a strategy σ^* and show that if it does not satisfy the conditions to be Self-Confirming Optimal (SCO) then for some $\eta > 0$ it cannot be a Proxy Equilibrium if the proxy mapping is η-close to perfect.

Clearly the second condition in the definition of SCO must hold for some beliefs q if σ^* is ever a Proxy Equilibrium. If at strategy σ^* the first condition is violated for any system of beliefs q satisfying the second condition, then either we have that for some (s_c, s_p), there are $x \in supp(\sigma^*(.|s_c, s_p))$ and $x' \in supp(\sigma^*(.|s_c))$ such that

$$\sum_{y\in Y} u(y,x,s)p(y|x,s_c) < \sum_{y\in Y} u(y,x',s)p(y|x',s_c)$$

and/or there exists $(s_c, s_p) \in S_p \times S_c$, $x^{ns} \notin supp(\sigma^*(.|s_c))$ and $x^s \in supp(\sigma^*(.|s_c, s_p))$ such that for all possible conditional beliefs $q(.|x^{ns}, s_c)$

$$\sum_{y\in Y} u(y,x^s,s)p(y|x^s,s_c) < \sum_{y\in Y} u(y,x^{ns},s)q(y|x^{ns},s_c)$$

We show that neither of these cases can hold. To show the first case cannot hold, for given (x, x', s_c, s_p) define $\Delta = \sum_{y\in Y} u(y,x',s)p(y|x',s_c) - \sum_{y\in Y} u(y,x,s)p(y|x,s_c)$. We have that $\Delta > 0$ if we are in the first case. By Lemma A.5, we can find $\eta > 0$ such that for any π that is η-close to perfect

$$\sum_{y\in Y} u(y,x,s)p(y|x,s_c) + \frac{\Delta}{4} > \sum_{y\in Y} u(y,x,s)p_\pi(y = y^\bullet|x^\bullet = x, s^\bullet = s; \sigma^*)$$
$$\sum_{y\in Y} u(y,x',s)p(y|x',s_c) - \frac{\Delta}{4} < \sum_{y\in Y} u(y,x,s)p_\pi(y = y^\bullet|x^\bullet = x', s^\bullet = s; \sigma^*)$$

For σ^* to be implementable as a Proxy Equilibrium at π requires that

$$\sum_{y\in Y} u(y,x,s)p_\pi(y = y^\bullet|x^\bullet = x, s^\bullet = s; \sigma^*) \geq \sum_{y\in Y} u(y,x',s)p_\pi(y = y^\bullet|x^\bullet = x', s^\bullet = s; \sigma^*)$$

We combine this with the inequalities above to get a contradiction to the definition of Δ.

$$\frac{\Delta}{2} > \sum_{y\in Y} u(y,x',s)p(y|x',s_c) - \sum_{y\in Y} u(y,x,s)p(y|x,s_c) = \Delta > 0$$

For the second case, since $Q(x^{ns}, s_c)$ the space of all possible conditional beliefs given (x^{ns}, s_c) is compact, we have that

$$\bar{\xi} = \min_{q\in Q(x^{ns},s_c)} \sum_{y\in Y} u(y,x^{ns},s)q(y|x^{ns},s_c) - \sum_{y\in Y} u(y,x^s,s)p(y|x^s,s_c) > 0$$

is attained. We then make the same argument as in the first step with $\bar{\xi}$ replacing Δ.

⇒ (Sufficiency): For this part of the proof, we show that for any $\eta > 0$ we can find a proxy mapping that is η close to perfect such that under the SCO strategy σ^* the perceived beliefs ensure σ^* is a Proxy Equilibrium. This works because the perceived beliefs will be sufficiently close to those conditional beliefs that ensure σ^* satisfies the SCO conditions.

We first show the sufficiency of a proxy mapping where outcome variables are perfectly measured, (21) is strict and $supp(p(y|x,s_c)) = Y$ for all $(x, s_c) \in X \times S_c$. Denote the set of action and control signal combinations that are not in the support of σ^* by

$$XS_c^{ns}(\sigma^*) = \{(x, s_c) \in X \times S_c : x \notin supp(\sigma^*(.|s_c))\}$$

The set of action and control combinations in the support is then denoted $XS_c^s(\sigma^*) = (X \times S_c) \setminus XS_c^{ns}(\sigma^*)$. Then for any (x^{ns}, s_c) we find $q \in Q(x^{ns}, s_c)$, so as to satisfy (21) strictly. For any $r \in (0,1)$, we construct beliefs for all $(x, s_c) \in XS_c^s(\sigma^*)$

$$\hat{q}(y|x,s_c) = \frac{1}{r}p(y|x,s_c) - \frac{1-r}{r} \sum_{(x',s_c')\in XS_c^{ns}(\sigma^*)} q(y|x',s_c')\frac{1}{|XS_c^{ns}(\sigma^*)|}$$

For r close to 1, since $p(y|x,s_c) \in (0,1)$ for all $(y,x,s_c) \in Y \times X \times S_c$ we have that $\hat{q}(y|x^{ns}, s_c) \in (0,1)$ for all $(x^{ns}, s_c) \in X \times S_c$. To implement these beliefs, we define the proxy mapping[15]

$$\pi_c(y^\bullet = y, x^\bullet, s_c^\bullet|y,x,s_c) = \begin{cases} \dfrac{\hat{q}(y|x^\bullet,s_c^\bullet)r\sigma^*(x^\bullet|s_c^\bullet)p(s_c^\bullet)}{\sum_{(\tilde{x},\tilde{s}_c)\in X\times S_c} p(y|\tilde{x},\tilde{s}_c)\sigma^*(\tilde{x}|\tilde{s}_c)p(\tilde{s}_c)} & \text{if } (x^\bullet, s_c^\bullet) \in XS_c^s(\sigma^*) \\ \dfrac{q(y|x^\bullet,s_c^\bullet)\frac{1-r}{|XS_c^{ns}(\sigma^*)|}}{\sum_{(\tilde{x},\tilde{s}_c)\in X\times S_c} p(y|\tilde{x},\tilde{s}_c)\sigma^*(\tilde{x}|\tilde{s}_c)p(\tilde{s}_c)} & \text{if } (x^\bullet, s_c^\bullet) \in XS_c^{ns}(\sigma^*) \end{cases}$$

[15] Since the proxy mapping is uninformative about $s_p \in S_p$, we can restrict focus on the proxy mapping over $Y \times X \times S_c$.

with $\pi_c(y^\bullet \neq y, x^\bullet, s_c^\bullet | y, x, s_c) = 0$. We can the see that by our definition of $\hat{q}$ that for any $(y, x, s_c) \in Y \times X \times S_c$, $\sum_{y^\bullet \in Y} \pi_c(y^\bullet, x^\bullet, s_c^\bullet | y, x, s_c) > 0$ and

$$\begin{aligned}
\sum_{y^\bullet, x^\bullet, s_c^\bullet \in Y \times X \times S_c} \pi_c(y^\bullet, x^\bullet, s_c^\bullet | y, x, s_c) &= \sum_{(x^\bullet, s_c^\bullet) \in X \times S_c} \pi_c(y^\bullet = y, x^\bullet, s_c^\bullet | y, x, s_c) \\
&= \frac{r \sum_{(x^\bullet, s_c^\bullet) \in XS_c^s(\sigma^*)} \hat{q}(y|x^\bullet, s_c^\bullet)\sigma^*(x^\bullet|s_c^\bullet)p(s_c^\bullet)}{\sum_{(\tilde{x}, \tilde{s}_c) \in X \times S_c} p(y|\tilde{x}, \tilde{s}_c)\sigma^*(\tilde{x}|\tilde{s}_c)p(\tilde{s}_c)} + \frac{(1-r)\sum_{(x^\bullet, s_c^\bullet) \in XS_c^{ns}(\sigma^*)} q(y|x^\bullet, s_c^\bullet)\frac{1}{XS_c^{ns}(\sigma^*)}}{\sum_{(\tilde{x}, \tilde{s}_c) \in X \times S_c} p(y|\tilde{x}, \tilde{s}_c)\sigma^*(\tilde{x}|\tilde{s}_c)p(\tilde{s}_c)} \\
&= 1
\end{aligned}$$

Thus we have that π_c is a valid proxy mapping. For any $\eta > 0$ we can then define

$$\pi(y^\bullet, x^\bullet, s_c^\bullet | y, x, s_c) = (1-\eta)\pi_\delta(y^\bullet, x^\bullet, s_c^\bullet | y, x, s_c) + \eta \pi_c(y^\bullet, x^\bullet, s_c^\bullet | y, x, s_c)$$

Clearly this proxy mapping is η close to perfect. This mapping induces conditional beliefs according to

$$\begin{aligned}
&p_\pi(y|x^\bullet, s_c^\bullet; \sigma^*) \\
&= \frac{(1-\eta)p(y^\bullet|x^\bullet, s_c^\bullet)\sigma^*(x^\bullet|s_c^\bullet)p(s_c^\bullet)}{(1-\eta)\sigma^*(x^\bullet|s_c^\bullet)p(s_c^\bullet) + \eta\sum_y \pi_c(x^\bullet, s_c^\bullet|y)\sum_{(x, s_c) \in X \times S_c} p(y|x, s_c)\sigma^*(x|s_c)p(s_c)} \\
&+ \frac{\eta\pi_c(x^\bullet, s_c^\bullet|y)\sum_{(x, s_c) \in X \times S_c} p(y|x, s_c)\sigma^*(x|s_c)p(s_c)}{(1-\eta)\sigma^*(x^\bullet|s_c^\bullet)p(s_c^\bullet) + \eta\sum_y \pi_c(x^\bullet, s_c^\bullet|y)\sum_{(x, s_c) \in X \times S_c} p(y|x, s_c)\sigma^*(x|s_c)p(s_c)}
\end{aligned}$$

Which means that for $(x, s_c) \in XS_c^s(\sigma^*)$, $p_\pi(y|x, s_c; \sigma^*) = \frac{1-\eta}{(1-\eta)+r\eta}p(y|x, s_c) + \frac{r\eta}{(1-\eta)+r\eta}\hat{q}(y|x, s_c)$ while for $(x, s_c) \in XS_c^{ns}(\sigma^*)$ we have $p_\pi(y|x, s_c; \sigma^*) = q(y|x, s_c)$. We can choose r close to one so that $\hat{q}(y|x, s_c)$ is arbitrarily close to $p(y|x, s_c)$. Therefore by Lemma A.5, we have that we can find r close enough to one such that σ^* is a Proxy Equilibrium due to the strict conditions (21).

Finally consider the case with only weak inequality and without full support. We can construct a proxy mapping that is η-close to perfect in the same manner. Take any system of beliefs $q \in Q$ satisfying the two conditions in the proposition statement. We can write π_c as

$$\pi_c(y^\bullet, x^\bullet, s_c^\bullet | y, x, s_c) = q(y^\bullet|x^\bullet, s_c^\bullet)\hat{p}(x^\bullet, s_c^\bullet)$$

with $\hat{p}$ as any joint distribution such that $\hat{p}(x^\bullet, s_c^\bullet) > 0$. The result then follows from the same construction as before with the changed π_c. □

Data availability

No data was used for the research described in the article.

References

Battigalli, P., 1987. Comportamento razionale ed equilibrio nei giochi e nelle situazioni sociali. Mimeo.
Battigalli, P., Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M., 2015. Self-confirming equilibrium and model uncertainty. Am. Econ. Rev. 105 (2), 646–677.
Battigalli, P., Francetich, A., Lanzani, G., Marinacci, M., 2019. Learning and self-confirming long-run biases. J. Econ. Theory 183, 740–785.
Blackwell, D., 1953. Equivalent comparisons of experiments. Ann. Math. Stat. 24 (2), 265–272.
Bohren, J.A., Hauser, D.N., 2021. Learning with heterogeneous misspecified models: characterization and robustness. Econometrica 89 (6), 3025–3077.
Bohren, J.A., Hauser, D.N., 2024a. The behavioral foundations of model misspecification: a decomposition. Mimeo.
Bohren, J.A., Hauser, D.N., 2024b. Misspecified models in learning and games. Annu. Rev. Econ. submitted for publication, https://doi.org/10.1146/annureveconomics-091324-032415.
Carroll, R.J., Ruppert, D., Stefanski, L.A., Crainiceanu, C.M., 2006. Measurement Error in Nonlinear Models: A Modern Perspective. Chapman and Hall/CRC.
Chalfin, A., McCrary, J., 2018. Are us cities underpoliced? Theory and evidence. Rev. Econ. Stat. 100 (1), 167–186.
Cooper, A.C., Woo, C.Y., Dunkelberg, W.C., 1988. Entrepreneurs' perceived chances for success. J. Bus. Venturing 3 (2), 97–108.
Coyle, D., 2015. GDP: A Brief but Affectionate History-Revised and Expanded Edition. Princeton University Press.
Dekel, E., Fudenberg, D., Levine, D.K., 2004. Learning to play Bayesian games. Games Econ. Behav. 46 (2), 282–303.
Eliaz, K., Galperti, S., Spiegler, R., 2024. False narratives and political mobilization. J. Eur. Econ. Assoc.
Eliaz, K., Spiegler, R., 2020. A model of competing narratives. Am. Econ. Rev. 110 (12), 3786–3816.
Eliaz, K., Spiegler, R., Thysen, H.C., 2021. Strategic interpretations. J. Econ. Theory 192, 105192.
Esponda, I., 2008. Behavioral equilibrium in economies with adverse selection. Am. Econ. Rev. 98 (4), 1269–1291.
Esponda, I., Pouzo, D., 2016. Berk–Nash equilibrium: a framework for modeling agents with misspecified models. Econometrica 84 (3), 1093–1130.
Esponda, I., Pouzo, D., 2017. Conditional retrospective voting in large elections. Am. Econ. J. Microecon. 9 (2), 54–75.
Eyster, E., Rabin, M., 2005. Cursed equilibrium. Econometrica 73 (5), 1623–1672.
Frick, M., Iijima, R., Ishii, Y., 2020. Misinterpreting others and the fragility of social learning. Econometrica 88 (6), 2281–2328.
Fudenberg, D., He, K., 2018. Learning and type compatibility in signaling games. Econometrica 86 (4), 1215–1255.
Fudenberg, D., Kamada, Y., 2015. Rationalizable partition-confirmed equilibrium. Theor. Econ. 10 (3), 775–806.
Fudenberg, D., Kamada, Y., 2018. Rationalizable partition-confirmed equilibrium with heterogeneous beliefs. Games Econ. Behav. 109, 364–381.
Fudenberg, D., Lanzani, G., Strack, P., 2021. Limit points of endogenous misspecified learning. Econometrica 89 (3), 1065–1098.
Fudenberg, D., Lanzani, G., Strack, P., 2024. Selective memory equilibrium. J. Polit. Econ. 132 (12), 3978–4020.

Fudenberg, D., Levine, D.K., 1993. Self-confirming equilibrium. Econometrica, 523–545.
Fudenberg, D., Levine, D.K., 2006. Superstition and rational learning. Am. Econ. Rev. 96 (3), 630–651.
Gonçalves, D., 2023. Sequential sampling and equilibrium. Mimeo.
Haltiwanger, J., 2015. Job creation, job destruction, and productivity growth: the role of young businesses. Annu. Rev. Econ. 7, 341–358.
Hayward, M.L., Shepherd, D.A., Griffin, D., 2006. A hubris theory of entrepreneurship. Manag. Sci. 52 (2), 160–172.
Heckman, J.J., 1979. Sample selection bias as a specification error. Econometrica, 153–161.
Heidhues, P., Kőszegi, B., Strack, P., 2018. Unrealistic expectations and misguided learning. Econometrica 86 (4), 1159–1214.
Jehiel, P., 2005. Analogy-based expectation equilibrium. J. Econ. Theory 123 (2), 81–104.
Jehiel, P., 2018. Investment strategy and selection bias: an equilibrium perspective on overoptimism. Am. Econ. Rev. 108 (6), 1582–1597.
Jehiel, P., Koessler, F., 2008. Revisiting games of incomplete information with analogy-based expectations. Games Econ. Behav. 62 (2), 533–557.
Lipnowski, E., Sadler, E., 2019. Peer-confirming equilibrium. Econometrica 87 (2), 567–591.
Malmendier, U., Tate, G., 2005. Ceo overconfidence and corporate investment. J. Finance 60 (6), 2661–2700.
Nakamura, E., Steinsson, J., 2013. Price rigidity: microeconomic evidence and macroeconomic implications. Annu. Rev. Econ. 5, 133–163.
Office for National Statistics, 2023. Business demography, UK (released 17 November 2022). https://www.ons.gov.uk/businessindustryandtrade/business/activitysizeandlocation/datasets/businessdemographyreferencetable. (Accessed 5 February 2023).
Osborne, M.J., Rubinstein, A., 1998. Games with procedurally rational players. Am. Econ. Rev. 88 (4), 834–847.
Osborne, M.J., Rubinstein, A., 2003. Sampling equilibrium, with an application to strategic voting. Games Econ. Behav. 45 (2), 434–441.
Patil, S., Salant, Y., 2024. Optimal sample sizes and statistical decision rules. Theor. Econ. 19 (2), 583–604.
Salant, Y., Cherry, J., 2020. Statistical inference in games. Econometrica 88 (4), 1725–1752.
Schennach, S., 2022. Measurement systems. J. Econ. Lit. 60 (4), 1223–1263.
Schumacher, H., Thysen, H.C., 2022. Equilibrium contracts and boundedly rational expectations. Theor. Econ. 17 (1), 371–414.
Spiegler, R., 2016. Bayesian networks and boundedly rational expectations. Q. J. Econ. 131 (3), 1243–1290.
Spiegler, R., 2017. 'Data monkeys': a procedural model of extrapolation from partial statistics. Rev. Econ. Stud. 84 (4), 1818–1841.
Spiegler, R., 2020. Can agents with causal misperceptions be systematically fooled? J. Eur. Econ. Assoc. 18 (2), 583–617.
Spiegler, R., 2021. Modeling players with random "data access". J. Econ. Theory 198, 105374.
Steiner, J., Stewart, C., 2008. Contagion through learning. Theor. Econ. 3 (4), 431–458.
Tetlow, R.J., Von zur Muehlen, P., 2001. Robust monetary policy with misspecified models: does model uncertainty always call for attenuated policy? J. Econ. Dyn. Control 25 (6–7), 911–949.
Wooldridge, J.M., 2020. Introductory Econometrics: A Modern Approach, 7th ed. Cengage Learning.