Structural Stability Implies Robustness to Bounded Rationality[1]

Luca Anderlini

Department of Economics, University of Southampton, Highfield, Southampton SO17 1BJ, United Kingdom; and Department of Economics, Georgetown University, Washington DC 20057, U.S.A.
E-mail: Luca.Anderlini@soton.ac.uk

and

David Canning

School of Management and Economics, Queen's University Belfast, University Road, Belfast, BT7 1NN, United Kingdom
E-mail: d.canning@qub.ac.uk

Received August 20, 2000; revised December 5, 2000

The introduction of a small amount of bounded rationality into a model sometimes has little effect, and sometimes has a dramatic impact on predicted behavior. We call a model robust to bounded rationality if small deviations from rationality result only in small changes in the equilibrium set. We also say that a model is structurally stable if the equilibrium set (given fully rational agents) varies continuously with the parameter values of the model.

It is easy to see that when the equilibrium set is discontinuous, bounded rationality can have a very large impact on behavior in the neighborhood of the discontinuity. We go further and show that it is only at such discontinuities that bounded rationality can have large effects. It follows that a model is robust to bounded rationality if and only if it is structurally stable. Thus, we can characterize which models will be robust to bounded rationality and which will not, independently of the exact form that the bounded rationality takes.
Journal of Economic Literature Classification Numbers: C69, C79, D51, E19.

1. INTRODUCTION

1.1. Motivation

The assumption of perfect rationality that underlies most economic models is far too strict. Ideally, we would like a model of bounded rational-

[1]We wish to thank Paolo Siconolfi and David Webb for insightful discussions and comments. During the preparation of this paper David Canning visited the Department of Economics of Harvard University. He is grateful for their generous hospitality.

ity that fits observed behavior and could be used as a basis for economic analysis. This is both an empirical and a theoretical question which, for the moment, lacks a completely satisfactory answer.[2] For instance, Ido and Roth [19] investigate how well simple learning models fit experimental data; while these simple models out-perform theories based on perfect rationality they can still not be considered as complete descriptions of actual behavior.

Given the difficulty of constructing a fully specified procedural model of bounded rationality it is interesting to ask how robust economic models are to relaxing the assumption of perfect rationality. This allows us to investigate which models give predictions that are unlikely to change much when we introduce a small amount of bounded rationality, and which may be giving misleading predictions that rely crucially on the assumption of perfect rationality.

We measure the degree of rationality by means of an arbitrary continuous function that is normalized so that zero corresponds to full rationality. The advantage of this somewhat abstract framework is that it captures many common approaches to modeling bounded rationality. This allows us to study the impact of introducing bounded rationality without having to specify the precise rationality measure under consideration.

A problem with a more concrete approach is that many different measures of rationality have been suggested. For example, in Radner [25] the notion of ϵ-Equilibrium is based on how far each agent's payoff is from that obtained by a fully rational agent. In macroeconomics and finance it is common to assume that most agents are fully rational, but some small proportion follow ad hoc or arbitrary rules; in this case we can use the proportion of irrational agents in the model as our rationality measure. In a rational expectations macro-model we could assume that each agent maximizes given his beliefs, and measure rationality by how far an agent's beliefs are from those of a fully rational agent. This way of proceeding may be particularly relevant in learning models, since even if learning rules lead beliefs to converge, the beliefs they generate are close to, but not perfectly, correct. However, as we shall see, each of these suggested measures satisfies our assumption of continuity and in each case a zero corresponds to full rationality. It follows that our results are applicable in each case.

The effect of bounded rationality has been investigated in many specific models (Conlisk [9], Mankiw [21], Russel and Thaler [28], Evans and Ramey [12], Sethi and Franke [30], among others). We are particularly interested in the impact of very small amounts of bounded rationality. Following

[2]The literature on bounded rationality is vast and growing very rapidly. It would be fool-hardy to even attempt a survey here. From a theoretical point of view, we refer to the two comprehensive treatments of Weibull [32] and Rubinstein [27]. An authoritative introduction to the vast experimental literature is Roth [26].

Radner [25], one strand of the literature (Haltiwanger and Waldman [14], [16], [15], Bonfim and Diebold [7], among others) defines robustness to bounded rationality in terms of whether or not a small amount of bounded rationality only has a small impact on behavior. This is the approach we will use here. An alternative approach, which we do not follow, is to define robustness in terms of the ratio of the effect on behavior to the amount of bounded rationality, when there is only a small amount of bounded rationality. This definition of robustness, which has been put forward in the literature (Akerlof and Yellen [2], [1], Jones and Stock [20], Wang [31], among others) is harder to satisfy since, even if the effect on behavior is very small, the ratio of the effect to the amount of bounded rationality may be large.

Using our definition of robustness, it is clear that introducing a small amount of bounded rationality makes little difference in some situations but can have a large impact in others. The aim of this paper is to address the issue in a general setting and develop a characterization of which models are robust to a small amount of bounded rationality and which ones are not.

We begin by defining an abstract framework within which we define a model that depends on some exogenously given parameters, and an equilibrium notion. In addition, we assume that for all parameter values of the model an equilibrium (with full rationality) exists. Many economic models fit the abstract framework that we define below: we give as explicit examples N-player normal form games, a general equilibrium (pure exchange) model, a macroeconomic model with strategic complementarities, and a rational expectations macroeconomic model.

We define a model as robust to bounded rationality if small deviations from rationality result in only small changes in the equilibrium set. We then introduce the notion of structural stability: a model is structurally stable if the equilibrium set (under full rationality) varies continuously with changes in the parameter values. Structural stability fails at critical parameter values where the equilibrium set changes in a discontinuous way. Note that this is also a form a robustness; predictions in structurally stable models will be robust to small changes in the parameter values while predictions in models that are not structurally stable will not be robust near critical parameter values.

We show that there is a strong connection between the notions of robustness to bounded rationality and structural stability for the class of models we investigate. In particular, a model is robust to bounded rationality if and only if it is structurally stable. As a corollary of this result we have that introducing a small amount of bounded rationality has, and can only have, a large effect on behavior in a model near critical parameter values, where the equilibrium set is changing in a discontinuous way.

Intuitively, when there is a discontinuity in the equilibrium correspondence, equilibria that suddenly disappear live on in the neighborhood of the discontinuity as boundedly rational equilibria. This is fairly easy to demonstrate and gives boundedly rational equilibria that are far from rational behavior for parameter values near the discontinuity. This intuition is similar to that employed by Fudenberg and Levine [13] who use ϵ-Equilibria in finite approximations to games with infinite strategy sets to characterize the (fully rational) equilibria in the limit, even when the equilibrium set is not continuous at the limit.

What is less obvious is that a discontinuity in the equilibrium set is the *only* way in which a small amount of bounded rationality can have a large impact on behavior. Indeed, this part of our result depends crucially on the compactness of the parameter space that we assume for our model; when the parameter space is unbounded it is easy to construct examples where bounded rationality can make a large difference, despite the model being structurally stable.

Our general result, characterizing robustness to bounded rationality via the structural stability of the equilibrium mapping provides a simple method for analyzing the effects of bounded rationality in many specific models, independently of the exact nature of the bounded rationality. All that needs to be done is to study the structural stability of the model with fully rational agents and identify the critical parameter values.

1.2. Overview

The rest of the paper is organized as follows. In the next section we describe our abstract framework in full detail. In Section 3 we first set up the apparatus that allows us to define robustness to bounded rationality and structural stability for a model, and then we proceed to state the main result of the paper. In Section 4 we briefly discuss the relationship between our results and the literature on equilibrium refinements and bounded rationality.

Section 5 contains four leading examples that fit the class of models described in Section 2. These are, finite strategic-form games (Subsection 5.1), pure exchange general equilibrium (Subsection 5.2), a macroeconomic model with strategic complementarities (Subsection 5.3), and a rational expectations macroeconomic model (Subsection 5.4).

Section 6 contains some brief concluding remarks. For ease of exposition, all proofs are relegated to the Appendix. The Appendix also contains some standard definitions concerning compact-valued correspondences that we reproduce there purely for the sake of completeness. In the numbering of equations, Definitions, Lemmas and so on, a prefix of "A" means that the corresponding item is to be found in the Appendix.

2. THE MODEL

We start by defining the primitives of our analysis. For reasons of generality, we work with an abstract formulation of what constitutes a model and the associated *rationality function*.

The abstract framework which we set up now can be thought of intuitively as a parameterized class of 'generalized games' together with an associated abstract rationality function. We choose this way to proceed because it guarantees that our framework is sufficiently general to encompass many interesting economic models as is shown by the examples that we analyze in Section 5 below.

Definition 2.1 (Model). *A model $\mathcal{M}$ consists of a quadruple* $(\Lambda, \mathcal{A}, \mathcal{F}, \mathcal{R})$ *with the following interpretation.*

The set Λ*, with typical element* λ*, is the* parameter space*, while* $\mathcal{A}$*, with typical element* a*, is the* action space.

The correspondence $\mathcal{F} : \Lambda \times \mathcal{A} \rightrightarrows \mathcal{A}$ *is the* feasibility correspondence[3] *that associates a subset of* $\mathcal{A}$ *to each element of* $\Lambda \times \mathcal{A}$*.*

The feasibility correspondence $\mathcal{F}$ *clearly induces a further correspondence from* Λ *into* $\mathcal{A}$*, which we denote by* $\mathcal{C}$ *and we refer to as the* consistent actions correspondence. *Formally, given* $\mathcal{F}$*, we define* $\mathcal{C} : \Lambda \rightrightarrows \mathcal{A}$ *by setting* $\mathcal{C}(\lambda) = \{a \in \mathcal{A} \text{ such that } a \in \mathcal{F}(\lambda, a)\}$*.*

Finally, we let $\mathcal{G} \subseteq \Lambda \times \mathcal{A}$ *be the graph of the consistent actions correspondence* $\mathcal{C}$ *so that* $\mathcal{G} = \{(\lambda, a) \in \Lambda \times \mathcal{A} \text{ such that } a \in \mathcal{C}(\lambda)\}$*. This enables us to define* $\mathcal{R} : \mathcal{G} \to \mathbb{R}_+$ *as the* rationality function *for the model.*

We can think of the rationality function as measuring how far the actions chosen in the model are from a fully rational choices. We shall take a value of zero of the rationality function to denote full rationality and assume that the degree of rationality decreases as the value of the rationality function increases.

Within the abstract framework that we have just defined, we focus on the class of models that satisfy the following three properties.

Assumption 2.1 (Compactness). Λ *and* $\mathcal{A}$ *are both compact subsets of some complete separable metric spaces. Throughout the paper,* d_Λ *and* $\mathrm{d}_\mathcal{A}$ *denote the metrics used on* Λ *and* $\mathcal{A}$ *respectively.*

---

[3]Throughout the paper, the symbol $\rightrightarrows$ between two sets, say $\mathcal{A}$ and $\mathcal{B}$, denotes a correspondence that associates a subset of $\mathcal{B}$ to each element of $\mathcal{A}$, while, as is standard, $\to$ denotes a function associating an element of $\mathcal{B}$ to each element of $\mathcal{A}$.

ASSUMPTION 2.2 (Continuity of Consistent Actions). *The consistent actions correspondence* $\mathcal{C} : \Lambda \rightrightarrows \mathcal{A}$ *is non-empty, continuous and compact-valued.*[4]

ASSUMPTION 2.3 (Continuous Rationality). *The rationality function* $\mathcal{R} : \mathcal{G} \to \mathbb{R}_+$ *is continuous.*

Given any λ, the set of ϵ-Equilibria consists of those consistent actions which are within ϵ of full rationality.

DEFINITION 2.2 (ϵ-Equilibrium). *Given* $\mathcal{M}$*, any* $\lambda \in \Lambda$*, and any* $\epsilon \geq 0$*, the set of* ϵ*-Equilibria at* λ *is defined as*

$$\mathcal{E}(\lambda, \epsilon) = \{a \in \mathcal{A} \quad \text{such that} \quad a \in \mathcal{C}(\lambda) \quad \text{and} \quad \mathcal{R}(\lambda, a) \leq \epsilon\} \tag{1}$$

Note that the term ϵ-Equilibrium is usually taken to mean that agents are within ϵ of their maximal payoff in a game. Here, we use this term in the expanded sense of being within ϵ of full rationality according to an arbitrary rationality function associated with a model.

A special case of ϵ-Equilibria occurs when $\epsilon = 0$, and we obtain the equilibrium set of the model under full rationality.

DEFINITION 2.3 (Equilibrium). *Given* $\mathcal{M}$ *and any* $\lambda \in \Lambda$*, the equilibrium set at* λ *is defined as* $\mathcal{E}(\lambda, 0)$*. Sometimes the equilibrium set at* λ *will be denoted simply by* $\mathcal{E}(\lambda)$*.*

Note that we use the mapping $\mathcal{E}$ in several senses. In general, it denotes the set of ϵ-Equilibria at a parameter value λ. Sometimes, however, it is useful to think of how the equilibrium set varies with ϵ for a fixed λ or of how it varies with λ for a fixed ϵ. Sometimes we will write $\mathcal{E}(\lambda, \epsilon)$ as $\mathcal{E}_\lambda(\epsilon)$, depending on what is convenient. In what follows no confusion should arise in the multiple uses of the notation $\mathcal{E}$.

ASSUMPTION 2.4 (Existence). *For every* $\lambda \in \Lambda$ *the equilibrium set* $\mathcal{E}(\lambda)$ *is not empty.*

Clearly, Assumption 2.4 ensures that the set $\mathcal{E}(\lambda, \epsilon)$ is also non-empty for every $\lambda \in \Lambda$ and for every $\epsilon \geq 0$. A great deal of effort has gone into

[4]The definition of a continuous correspondence that we use is the standard one. A correspondence is continuous if and only if it is both upper hemicontinuous and lower hemicontinuous. See Definitions A.1, A.2 and A.3 in the Appendix.

proving the existence of an equilibrium for certain models. Here we are concerned only with the robustness of the equilibrium set once it has been established that an equilibrium exists.

We conclude this section with a preliminary result that characterizes a property of our set up which will be useful below.

LEMMA 2.1 (ϵ-Equilibrium Correspondence). *The correspondence $\mathcal{E}(\lambda, \epsilon)$ is compact-valued and upper hemicontinuous.*

3. STRUCTURAL STABILITY AND ROBUSTNESS TO ϵ-EQUILIBRIA

The central question with which we are concerned is how the equilibrium set changes when we allow agents to be boundedly rational. To study this question we clearly need to be able to measure the distance between two sets.

DEFINITION 3.4 (Haussdorf Metric). *Let (X, d_X) be a metric space, and let $\mathcal{Q}(X)$ be the collection of all closed, nonempty subsets of X. Throughout the rest of the paper, the Hausdorff distance induced by d_X between two sets $A, B \in \mathcal{Q}(X)$ is denoted by h_X and is defined as*

$$\mathrm{h}_X(A, B) \;=\; \sup\left\{ \sup_{a\in A}\left[\inf_{b\in B} \mathrm{d}_X(a,b)\right],\; \sup_{b\in B}\left[\inf_{a\in A} \mathrm{d}_X(a,b)\right]\right\} \tag{2}$$

Since our ϵ-Equilibrium Correspondence is compact-valued, the sets of actions that it generates are certainly closed. It follows that we can apply the Hausdorff metric $\mathrm{h}_{\mathcal{A}}$ induced by $\mathrm{d}_{\mathcal{A}}$ to measure the distance between them.

We now state the definition of robustness to bounded rationality which we will work with for the rest of the paper.

DEFINITION 3.5 (Robustness). *The model $\mathcal{M} = (\Lambda, \mathcal{A}, \mathcal{F}, \mathcal{R})$ is robust to ϵ-Equilibria if and only if*

$$\forall\, \gamma > 0,\;\; \exists \overline{\epsilon} > 0 \quad \text{such that} \quad \epsilon < \overline{\epsilon} \;\Rightarrow\; \mathrm{h}_{\mathcal{A}}[\mathcal{E}(\lambda,\epsilon), \mathcal{E}(\lambda)] < \gamma,\; \forall\, \lambda \in \Lambda \tag{3}$$

In other words, a model is robust to ϵ-Equilibria if and only if for ϵ sufficiently small, the ϵ-Equilibrium set is close to the equilibrium set.

Another way to look at Definition 3.5 is as follows. Let $\mathcal{Q}(\mathcal{A})$ be the set of all closed nonempty subsets of $\mathcal{A}$. We can then view the ϵ-Equilibrium Correspondence $\mathcal{E}_\lambda : \mathbb{R}_+ \rightrightarrows \mathcal{A}$ as a *function* $\mathcal{E}^*_\lambda : \mathbb{R}_+ \to \mathcal{Q}(\mathcal{A})$.

Studying the properties of $\mathcal{E}_\lambda$ and studying the properties of $\mathcal{E}^*_\lambda$ are largely equivalent exercises. Our next lemma shows that continuity of the correspondence $\mathcal{E}_\lambda$ is equivalent to continuity of the function $\mathcal{E}^*_\lambda$ in the Hausdorff metric.[5]

Lemma 3.1 (Hausdorff Continuity). *Let (Y, d_Y) be a complete separable metric space, $\mathcal{Q}(Y)$ be the set of all closed nonempty subsets of Y, and h_Y be the Hausdorff metric induced by d_X on $\mathcal{Q}(Y)$ as in Definition 3.4. Let X be a compact subset of a complete separable metric space, and $f : X \rightrightarrows Y$ be a compact-valued correspondence. Next consider the function $g : X \to \mathcal{Q}(Y)$ which is equivalent to f in the sense that for every $x \in X$ we have that $g(x) = f(x)$. Then f is a continuous correspondence (see Definition A.3) if and only if g is a continuous function in the Hausdorff metric h_Y.*

Notice that we require the value of $\overline{\epsilon}$ in (3) to be independent of λ. Intuitively, to call a model robust to ϵ-Equilibria we require that, as ϵ becomes small, the ϵ-Equilibrium set should be close to the actual equilibrium set in a manner that can be 'predicted' purely on the basis of the degree bounded rationality present in the model. From the definition of equicontinuity we then immediately have the following, which is stated without proof (see Definition A.4 in the Appendix).

Remark 3.1 (Equicontinuity). *The model $\mathcal{M} = (\Lambda, \mathcal{A}, \mathcal{F}, \mathcal{R})$ is robust to ϵ-Equilibria if and only if the family of functions $\mathcal{E}^*_\lambda(\epsilon)$ is equicontinuous at $\epsilon = 0$.*

A weaker notion of robustness to bounded rationality would be to insist only that the ϵ-Equilibrium correspondence is continuous at zero for each parameter value. However, such a definition has no bite; every model satisfying our assumptions would be robust to bounded rationality in this sense of the term. This is the content of our next preliminary result.

Lemma 3.2 (Continuity at $\epsilon = 0$). *For every $\lambda \in \Lambda$, the correspondence $\mathcal{E}_\lambda(\epsilon)$ is continuous at $\epsilon = 0$.*

Note that $\mathcal{E}_\lambda$ need not be continuous everywhere, only at zero. Lemma 3.2 implies that, for given λ, small deviations from rationality will have only a small effect on the equilibrium set.

For fixed λ, as ϵ decreases, the ϵ-Equilibrium set converges to the equilibrium set. However, the equilibrium set need not be continuous in λ. If

[5] Lemma 3.1 below is stated without proof since it is a standard result. See for instance Hildenbrand [18, Problem I.B.4]

this is the case, for fixed $\epsilon > 0$ we can always find a λ such that the set of ϵ-Equilibria is 'far' from the actual equilibrium set. This is the problem captured in our definition of robustness to bounded rationality above.

We define a model to be structurally stable if a small change in its parameters leads to a correspondingly small change in the equilibrium set. In our definition of structural stability below we focus on the continuity of the equilibrium correspondence as the parameters vary.

Definition 3.6 (Structural Stability). *The model $\mathcal{M} = (\Lambda, \mathcal{A}, \mathcal{F}, \mathcal{R})$ is structurally stable if and only if the equilibrium correspondence $\mathcal{E} : \Lambda \rightrightarrows \mathcal{A}$ is continuous.*

We can define the parameters of the model as being at a critical value if the equilibrium mapping is not continuous at that value.

In a model that lacks structural stability, as λ approaches a critical value, the equilibrium set suddenly jumps in size to include new points. Note that, by Lemma 2.1, the equilibrium mapping is upper hemicontinuous; the problem is that it looses lower hemicontinuity at the critical points.

We are now ready to state our main result.

Theorem 3.1 (Robustness and Stability). *The model $\mathcal{M} = (\Lambda, \mathcal{A}, \mathcal{F}, \mathcal{R})$ is robust to ϵ-Equilibria if and only if it is structurally stable.*

The intuition behind the proof of Theorem 3.1 is that as we move away from critical values of the parameters the equilibrium set suddenly gets smaller. However, the equilibria that disappear live on in the form of ϵ-Equilibria at nearby parameter values. For parameters near the critical values we can always find ϵ-Equilibria that are far from the equilibrium set for that parameter. No matter how small we make ϵ, the problem persists for parameters sufficiently close to the critical values.

Therefore, for parameters sufficiently close to the critical values, the set of ϵ-Equilibria must be far from the equilibrium set under full rationality. Conversely, for parameters that are not near the critical values, a small amount of bounded rationality cannot have a disproportionately large impact on the equilibria of the model.

As we mentioned above, the more surprising part of Theorem 3.1 is that the *only* way in which a small amount of bounded rationality can have a large impact on behavior is a discontinuity in the equilibrium mapping. In other words, what is less obvious about our main result is that the structural stability of the model implies that it is robust to bounded rationality. The intuition we have given above for the proof of Theorem 3.1 is not quite sufficient to bring out some crucial parts of the argument. In particular, the proof of Theorem 3.1 hinges on the compactness of the parameter and action spaces in a way that may not be apparent at first sight.

To bring out this point more clearly, we conclude this section with a simple example of a model that violates compactness and is structurally stable, but is not robust to bounded rationality.

Consider a model $\mathcal{M}$ as follows. The action and parameter spaces are both equal to the non-negative real numbers, so that $\Lambda = \mathcal{A} = \mathbb{R}_+$.[6] The feasibility correspondence is trivial in the sense that $\mathcal{F}(\lambda) = \mathcal{A}$ for every $\lambda \in \Lambda$, so that the consistent actions correspondence is given by $\mathcal{C}(\lambda) = \mathcal{A}$ for every $\lambda \in \Lambda$. The rationality function is given by $\mathcal{R}(\lambda, a) = |a^{1/2} - \lambda|$.

Given this specification of $\mathcal{M}$, given any $\lambda \in \Lambda$ we obviously have that the equilibrium set is $\mathcal{E}(\lambda) = \lambda^2$, so that the model is obviously structurally stable according to Definition 3.6.

The set of ϵ-Equilibria for the model is easily computed as

$$\mathcal{E}(\lambda, \epsilon) \;=\; \{\, a \in \mathcal{A} \;\text{ such that }\; \lambda - \epsilon \,\le\, a^{\frac{1}{2}} \,\le\, \lambda + \epsilon \,\} \tag{4}$$

Using (4), straightforward computations give us that the Hausdorff distance between the equilibrium set and the set of ϵ-Equilibria is given by

$$\mathrm{h}_{\mathcal{A}}[\mathcal{E}(\lambda, \epsilon), \mathcal{E}(\lambda)] \;=\; \epsilon^2 + 2\,\epsilon\,\lambda \tag{5}$$

Equation (5) clearly tells us that for any given ϵ we can always make the distance $\mathrm{h}_{\mathcal{A}}[\mathcal{E}(\lambda, \epsilon), \mathcal{E}(\lambda)]$ arbitrarily large by choosing a value of λ that is large enough. Therefore, the model is not robust to bounded rationality according to Definition 3.5.

Finally, notice that if we modify the model $\mathcal{M}$ we have just described to make Λ and $\mathcal{A}$ some compact subsets of $\mathbb{R}_+$, then robustness to bounded rationality would immediately be restored (as Theorem 3.1 implies). Indeed, from (5) it is immediate that if Λ is bounded above, then (3) must hold so that the model is structurally stable according to Definition 3.5.

4. EQUILIBRIUM SELECTION AND BOUNDED RATIONALITY

It is immediate from Definition 2.2 above that the approach we have taken in this paper leads us to view bounded rationality as a weaker notion than full rationality; a fully rational equilibrium is always an ϵ-Equilibrium under bounded rationality. This is a little at odds with the literature on using bounded rationality as an equilibrium selection device. In that literature (Selten [29], Myerson [23], among very many others) a fully rational

[6]Notice that this choice of Λ and $\mathcal{A}$ violates Assumption 2.1.

equilibrium is not always boundedly rational, and this can be used as a refinement of equilibrium.

Since our basic approach has been to argue that bounded rationality always increases the size of the equilibrium set, our notion of bounded rationality can never act as a selection device. To use bounded rationality as an equilibrium selection device the fully rational model must not automatically satisfy the conditions of bounded rationality.

However, equilibrium selection can be reconciled with our framework by replacing our definition of an ϵ-Equilibrium having a rationality level *at most* ϵ with the definition of an 'Exact' ϵ-Equilibrium with a rationality level *exactly equal* to ϵ. In this way, a fully rational equilibrium need not be an equilibrium under bounded rationality. We can then define a 'Perfect' equilibrium as the limit of any sequence of such Exact ϵ-Equilibria as ϵ approaches zero. Our next definition formalizes the notion we have just outlined.

DEFINITION 4.7 (Exact ϵ-Equilibrium). *Given $\mathcal{M}$, any $\lambda \in \Lambda$, and any $\epsilon > 0$, the set of Exact ϵ-Equilibria at λ is defined as*

$$\mathcal{E}^X(\lambda, \epsilon) = \{a \in \mathcal{A} \quad \text{such that} \quad a \in \mathcal{C}(\lambda) \quad \text{and} \quad \mathcal{R}(\lambda, a) \; = \; \epsilon\} \tag{6}$$

Taking the limit of Exact ϵ-Equilibria as ϵ shrinks to zero, gives us the set of Perfect Equilibria.

DEFINITION 4.8 (Perfect Equilibrium). *Given $\mathcal{M}$ and any $\lambda \in \Lambda$, the set of Perfect Equilibria at λ is defined as the limit of any sequence of Exact ϵ-Equilibria as $\epsilon \to 0$.*

Our next remark shows that, if we restrict ourselves to Exact ϵ-Equilibria, then our set up is consistent with the use of bounded rationality as an equilibrium selection device.

REMARK 4.2 (Rational Equilibria and Perfect Equilibria). *Given $\mathcal{M}$ and any $\lambda \in \Lambda$, the set of Perfect Equilibria at λ is (weakly) contained in the set $\mathcal{E}(\lambda)$ of fully rational equilibria.*

5. EXAMPLES

5.1. Finite Strategic-Form Games

Our first example of a class of economic models that fit Definition 2.1 above, and at the same time satisfy Assumptions 2.1, 2.2, 2.3 and 2.4 is that of Finite Strategic-Form Games.

Consider the following set of finite-action strategic-form N-player games. For each player n, fix a finite action set α_n, and let m_n denote the cardinality of α_n, while α denotes the joint action set. Let also $Q = N \prod_{n=1}^{M} m_n$. Next, let $\lambda \in \mathbb{R}^Q$ denote the array of all possible payoff vectors corresponding to all possible action profiles α. We take the metric d_Λ on the space of all possible λ to be the standard Euclidean metric, and we assume that λ is restricted to lie in some compact subset Λ of $\mathbb{R}^Q$.

For each player n let $\mathcal{A}_n$ be the set of mixed strategies available to player n in the game above. Let also $\mathcal{A}$ be the set of all possible mixed strategy profiles. Clearly, $\mathcal{A}$ is a finite-dimensional Euclidean space. We take $\mathrm{d}_\mathcal{A}$ to be the standard Euclidean metric on $\mathcal{A}$.

The feasibility correspondence in this case is straightforward since all mixed strategies in $\mathcal{A}_n$ are always available to n, regardless of the strategies of other players and of the value of λ. Hence we set $\mathcal{F}(\lambda, a) = \mathcal{A}$ for every $\lambda \in \Lambda$ and for every $a \in \mathcal{A}$.[7]

The payoffs to each player n in the mixed extension of game that we now consider can be written as $\pi_n(\lambda, a_n, a_{-n})$. A natural rationality function for player n is given by

$$\mathcal{R}_n(\lambda, a) \;=\; \max_{\tilde{a}_n \in \mathcal{A}_n} \pi_n(\lambda, \tilde{a}_n, a_{-n}) \;-\; \pi_n(\lambda, a) \tag{7}$$

We take the overall rationality function to be $\mathcal{R} = \max\{\mathcal{R}_1, \ldots, \mathcal{R}_N\}$, with each component defined as in (7).

The model $\mathcal{M} = (\Lambda, \mathcal{A}, \mathcal{F}, \mathcal{R})$ that we have just defined clearly satisfies Assumption 2.1 by construction, and it trivially satisfies Assumption 2.2. By the maximum theorem of Berge [5], each component of the rationality function defined in (7) is continuous, and therefore $\mathcal{R}$ is continuous as required. Hence $\mathcal{M}$ satisfies Assumption 2.3. Finally, every finite-action strategic-form N-player game has a Nash equilibrium in mixed strategies as shown by Nash [24]. Therefore $\mathcal{M}$ satisfies Assumption 2.4.

Examples of critical parameter values at which the equilibrium correspondence of $\mathcal{M}$ looses lower hemicontinuity are not hard to find. For instance, if we let λ tend to the vector of zero payoffs for all players and for all possible outcomes of the game, we approach a critical point. There is an interesting interpretation of this case if we think of the limit described above as a process of scaling down payoffs, say by multiplying them all by a scalar that shrinks to zero. In this case, we can interpret Theorem 3.1 as telling us that the scale of payoffs (in absolute value), relative to the scale of bounded rationality (ϵ) matters in determining whether the introduction of bounded rationality has a small or a large effect on the equilibrium set.

[7] Notice that this implies that the consistent actions correspondence is also straightforward in this case since $\mathcal{F}$ as above implies that $\mathcal{C}(\lambda) = \mathcal{A}$ for every $\lambda \in \Lambda$.

This may be a matter of concern for a variety of experimental settings since the scale of payoffs there is often limited by budget constraints.

Less trivial examples can be constructed exploiting the topological properties of the Nash equilibrium correspondence for finite strategic-form games (Wu and Jia-He [34], Wilson [33], Harsanyi [17]). For instance, we know that every finite strategic form game that has a continuum of Nash equilibria, must be one at which the Nash equilibrium correspondence loses lower hemicontinuity. In this case we can be assured that robustness to bounded rationality also fails.

5.2. Pure Exchange General Equilibrium

Our second example of a class of models that fits our framework it that of a set of Pure Exchange General Equilibrium models.

Consider a pure exchange economy with L goods and $N-1$ traders, indexed by $n = 1, \dots, N-1$, defined as follows.

$U_n : \mathrm{I\!R}^L \to \mathrm{I\!R}$ is the utility function of trader n, assumed to be continuous, without local maxima and strictly quasi-concave over the consumption set $\mathrm{I\!R}^L_+$. Let $\lambda_n = (\lambda_n^1, \dots, \lambda_n^L) \in \mathrm{I\!R}^L_{++}$ be the endowment vector of trader n, while $\lambda \in \mathrm{I\!R}^{L(N-1)}_{++}$ represents the array of endowments in the economy. Assume that λ is constrained to lie in some compact subset Λ of $\mathrm{I\!R}^{L(N-1)}_{++}$. For every $\lambda \in \Lambda$, let $r(\lambda) = [r^1(\lambda), \dots, r^L(\lambda)]$ be the vector of aggregate endowments of each good in the economy, and let r^* be the L-dimensional vector with components $r^{*\ell} = \max_{\lambda \in \Lambda} r^\ell(\lambda)$ for $\ell = 1, \dots, L$. Notice that since Λ is compact, the vector r^* is well defined. Let $\mathcal{B}$ be the set of vectors $a_n \in \mathrm{I\!R}^L_+$ which satisfy $a_n^\ell \leq r^{*\ell}$ for every $\ell = 1, \dots, L$. Finally, for every $n = 1, \dots, N-1$ let $\mathcal{A}_n = \mathrm{I\!R}^L_+ \cap \mathcal{B}$.

The N-th agent in the economy is the auctioneer. Let $\mathcal{A}_N$ be the $L-1$-dimensional simplex Δ^{L-1}, and finally let $\mathcal{A} = \mathop{\times}\limits_{n=1}^{N} \mathcal{A}_n$. We take both Λ and $\mathcal{A}$ to be equipped with their respective standard Euclidean metrics.

The feasibility correspondence for each trader is given by the budget sets. For every $n = 1, \dots N-1$, let

$$\mathcal{F}_n(\lambda, a) \; = \; \left\{ a_n \in \mathcal{A}_n \quad \text{such that} \sum_{\ell=1}^{L} a_N^\ell a_n^\ell \leq \sum_{\ell=1}^{L} a_N^\ell \lambda_n^\ell \right\} \tag{8}$$

The feasibility correspondence for the auctioneer is trivial since he can call any prices in $\mathcal{A}_N$, regardless of λ and of the actions of traders $n = 1, \dots, N-1$. In other words, we take it to be the case that

$$\mathcal{F}_N(\lambda, a) \; = \; \mathcal{A}_N \qquad \forall \, \lambda \in \Lambda, \;\; \forall \, a \in \mathcal{A} \tag{9}$$

We take the overall feasibility correspondence to be given by $\mathcal{F} = (\mathcal{F}_1, \ldots, \mathcal{F}_N)$, with each component defined as in (8) and in (9).

A natural rationality function for every trader $n = 1, \ldots, N-1$ is given by

$$\begin{aligned} \mathcal{R}_n(\lambda, a) \;&=\; \max_{\tilde{a}_n \in \mathcal{A}_n} U_n(\tilde{a}_n) \,-\, U_n(a_n) \\ &\text{s.t.}\;\; \tilde{a}_n \in \mathcal{F}_n(\lambda, a) \end{aligned} \tag{10}$$

To obtain market clearing, the best that the auctioneer can do is to set prices so as to maximize the value of aggregate excess demand. Therefore, a natural rationality function for the auctioneer is given by

$$\begin{aligned} \mathcal{R}_N(\lambda, a) \;=\; \max_{\tilde{a}_N \in \mathcal{A}_N} \sum_{\ell=1}^{L} \tilde{a}_N^\ell \left[\sum_{n=1}^{N-1} a_n^\ell \,-\, \sum_{n=1}^{N-1} \lambda_n^\ell \right] - \\ \sum_{\ell=1}^{L} a_N^\ell \left[\sum_{n=1}^{N-1} a_n^\ell \,-\, \sum_{n=1}^{N-1} \lambda_n^\ell \right] \end{aligned} \tag{11}$$

We take the overall rationality function to be $\mathcal{R} = \max\{\mathcal{R}_1,\, \ldots,\, \mathcal{R}_N\}$, with each component defined as in (10) and in (11).

The model $\mathcal{M} = (\Lambda, \mathcal{A}, \mathcal{F}, \mathcal{R})$ that we have just defined clearly satisfies Assumption 2.1 by construction.

Since we have bounded the action of traders $n = 1, \ldots, N-1$ to be in $A_n = \mathbb{R}^L_+ + \cap \mathcal{B}$, every component of the feasibility correspondence defined in (8) is in fact continuous. Trivially, the N-th component of the feasibility correspondence is also continuous. It is then a matter of routine to check that the consistent actions correspondence $\mathcal{C}$ induced by $\mathcal{F}$ is continuous. Therefore $\mathcal{M}$ satisfies Assumption 2.2.

By the maximum theorem [5], each component of the rationality function defined in (10) and in (11) is continuous, and therefore $\mathcal{R}$ is continuous as required. Hence $\mathcal{M}$ satisfies Assumption 2.3.

Finally, it is well known that every pure exchange economy satisfying our assumptions has a competitive equilibrium (Debreu [10], Arrow and Debreu [3], Debreu [11]). Therefore $\mathcal{M}$ satisfies Assumption 2.4.

Clearly, the model that we have described involves ϵ-rationality from the part of the traders and of the auctioneer. However, we can adapt the set up to encompass one other possible interesting case; rational traders with an ϵ-rational auctioneer so that while traders are rational we only obtain approximate market-clearing.[8]

[8]Since ϵ-rational traders can choose any consumption bundle in a set with a non-empty interior, the symmetric case of exact market-clearing with ϵ-rational traders is hard to accommodate in our set up. Whatever the prices picked by the auctioneer, market-clearing cannot be guaranteed.

To look at this case, we need to modify both the feasibility correspondence and the rationality function of $\mathcal{M}$. The feasibility correspondence for the auctioneer remains as in (9), while the feasibility correspondence for each trader is constructed so that they are forced to pick the utility-maximizing actions given their budget sets. In other words, for every $n = 1, \ldots, N-1$, we set

$$\begin{aligned} \mathcal{F}_n(\lambda, a) &= \arg \max_{a_n \in \mathcal{A}_n} U_n(a_n) \\ & \quad \text{s.t.} \sum_{\ell=1}^{L} a_N^\ell a_n^\ell \leq \sum_{\ell=1}^{L} a_N^\ell \lambda_n^\ell \\ & \quad a_n \in \mathcal{A}_n \end{aligned} \tag{12}$$

Notice, that since the preferences of all traders are assumed to be strictly quasi-concave, the new feasibility correspondence that we have defined is continuous as before, and hence it is again a matter of routine to check that our new model satisfies Assumption 2.2.

The rationality function of the new model is simply set to be equal to the auctioneer component of the rationality function of the previous case, as in (11). Therefore, the new model satisfies Assumption 2.3. Finally, the new model always has an equilibrium for exactly the same reasons as before. Hence it satisfies Assumption 2.4.

Classes of pure exchange economies which are not structurally stable are not difficult to find. All we need to find is a set of pure exchange economies that includes one or more 'critical' endowment vectors. When utility functions are smooth, such examples can be constructed in well known ways.[9] For example, let λ' be an array of endowments for which a unique equilibrium price vector exists (for instance because the array of endowments is itself a Pareto-efficient allocation). Let also λ'' be an array of endowments for which there exists multiple, say three, equilibrium price vectors. Then as we move the parameters of the economy from λ' to λ'' (for instance taking the linear combinations of these two extreme points) we must 'cross' a critical economy λ^* where the equilibrium correspondence looses lower hemicontinuity.[10] Hence, the model will be neither structurally stable nor robust to bounded rationality. For parameter values near λ^*, the introduction of a small amount of bounded rationality will have a large effect on the equilibrium set.

5.3. A Macroeconomic Model with Strategic Complementarities

[9] See for instance Mas-Colell [22] or Balasko [4].

[10] Our claim here is a direct consequence of known results. See for instance Balasko [4, Proposition 4.2.5].

The model in this section is a simplified version of the model in Haltiwanger and Waldman [15]. In this framework there is a positive feed back from the aggregate level of economic activity to each individual's optimal activity level (strategic complementarity). Most agents optimize given their beliefs which are based on rational expectations but a small number of people in the model follow an ad hoc expectation formation mechanism. For simplicity, we assume they have fixed expectations. The measure of rationality in this model is the proportion of people who follow the simple (non-rational) expectations rule.

There is a continuum of agents of size one indexed by n. Each agent must decide whether to produce one unit of output this period or not. The earnings from producing output are $r = \alpha + \beta Y$ where α and β are fixed parameters and Y is the aggregate level of output. Set against this is the cost of production c. For each agent, the cost of production is a random draw with uniform distribution on the interval $[0, 1]$. This cost is known to the agent.

A proportion σ of the agents are of bounded rationality and simply set their expected earnings to $r^e = s$. The others, a fraction $1-\sigma$, have rational expectations and set $r^e = E(r)$.

Agents with production costs less than their expected earning will decide to produce output, while those with costs above their expected earnings will remain idle. Therefore, total output under these expectation formation rules is given by

$$Y(r^e, \sigma, s) = (1-\sigma)\int_0^{r^e} dn + \sigma \int_0^{s} dn = (1-\sigma)\, r^e + \sigma\, s \quad (13)$$

together with

$$r^e = \alpha + \beta Y(r^e, \sigma, s) \quad (14)$$

Notice that once expected earnings are one or more, all the agents with rational expectations produce output.

The model embodied in (13) and (14) can be fitted in our framework in a fairly straightforward way. We take the parameter space of the model to have typical element $\lambda = (\alpha, \beta, s)$. We assume that α, β and s are all constrained to lie in the closed interval $[0, 1]$.

It is convenient to include the proportion of agents σ who are boundedly rational as an element of the 'action' component of the model.[11] We take

[11] Of course, this should not be interpreted literally. As will become clear in a moment, it is simply a device to fit the model we have just outlined into the class of models to which Theorem 3.1 applies.

the action space of the model to have typical element $a = (\sigma, r^e, Y)$; the rational agents must decide on their expectations and we must set aggregate output. We assume that each of σ, r^e and Y lie in the unit interval $[0,1]$. We take the metric on Λ and $\mathcal{A}$ to be the standard Euclidean one. Therefore it is clear that the model we have described satisfies Assumption 2.1.

We define the rationality function of the model to be simply given by

$$\mathcal{R}(\lambda, a) \;=\; \sigma \tag{15}$$

so that it is trivial that the rationality function is continuous as required by Assumption 2.3.

The feasibility correspondence is constructed so that aggregate output must satisfy (13) and expectations are indeed rational. In other words, for every λ and a, we define $\mathcal{F}(\lambda, a)$ to be the following set[12]

$$\begin{aligned}\{\, a \in \mathcal{A} \;\mid\; \sigma \in [0,1] \;\text{ and }\; r^e \;=\; \min\{1, \alpha + \beta Y\} \;\text{ and}&\\ Y \;=\; (1-\sigma) r^e + \sigma s \,\}&\end{aligned} \tag{16}$$

It is easy to show that the consistent actions correspondence $\mathcal{C}$ induced by $\mathcal{F}$ defined in (16) is as follows. When $\lambda = (\alpha, \beta, s) \neq (0, 1, s)$, we have that $\mathcal{C}(\lambda)$ is the set of triplets (σ, r^e, Y) that satisfy

$$\begin{aligned}\sigma \in [0,1], \quad r^e = \min\left\{1, \alpha + \beta \frac{(1-\sigma)\alpha + \sigma s}{1-(1-\sigma)\beta}\right\} \quad \text{and}&\\ Y = \min\left\{1, \frac{(1-\sigma)\alpha + \sigma s}{1-(1-\sigma)\beta}\right\}&\end{aligned} \tag{17}$$

while when $\lambda = (\alpha, \beta, s) = (0, 1, s)$, we have that $\mathcal{C}(\lambda)$ is the set of triplets (σ, r^e, Y) that satisfy either

$$\sigma \in (0,1], \quad r^e = \min\{1, \alpha + \beta s\}, \quad Y = s \tag{18}$$

or

$$\sigma = 0, \quad r^e \in [0,1], \quad Y = r^e \tag{19}$$

By inspection, it is straightforward to show that the consistent actions correspondence defined in (17), (18) and (19) is continuous over Λ. Therefore the model satisfies Assumption 2.2 as required.

[12]Notice in (16) we have bounded r^e above by 1. This is without loss of generality since, as we noted above, once expected returns are 1 or more, all rational agents in the model produce one unit of output.

The (fully rational) equilibria of the model are easy to compute given (13) and (14), and using the fact that we have bounded above r^e above by 1. When $\beta < 1$ there is a unique equilibrium given by

$$r^e \;=\; \min\left\{1, \alpha + \beta\frac{\alpha}{1-\beta}\right\} \qquad \text{and} \qquad Y \;=\; \min\left\{1, \frac{\alpha}{1-\beta}\right\} \tag{20}$$

When $\beta = 1$ and $\alpha > 0$, there is also a unique equilibrium given by

$$r^e \;=\; \min\{1, \alpha + \beta\} \;=\; 1 \qquad \text{and} \qquad Y \;=\; 1 \tag{21}$$

Finally, when $\beta = 1$ and $\alpha = 0$ there is a continuum of equilibria given by

$$r^e \;= \in [0, 1] \qquad \text{and} \qquad Y \;=\; r^e \tag{22}$$

In this case, any level of output can be sustained as an equilibrium by the expectation that exactly that level of output will be produced.

Clearly, it follows from (20), (21) and (22) that the model's equilibrium set is non-empty for any parameter values. Therefore, the model satisfies Assumption 2.4 as required.

It is also evident from (20), (21) and (22) that the equilibrium correspondence is not continuous at ($\beta = 1$, $\alpha = 0$) since at this point it changes from a unique value to a continuum. The model is structurally unstable. It follows from Theorem 3.1 above that the equilibrium is not robust to introducing a small about of bounded rationality (that is, a small value of $\sigma > 0$) near this critical point.

The non-robustness of the model to bounded rationality near this critical point can be seen more directly in the following way. Given a *fixed* proportion σ of boundedly rational agents, for $\beta < 1$ the model we have described has a unique equilibrium given by

$$\begin{aligned} r^e \;&=\; \min\left\{1, \alpha + \beta\frac{(1-\sigma)\alpha + \sigma s}{1-(1-\sigma)\beta}\right\} \quad \text{and} \\ &\qquad\qquad Y \;=\; \min\left\{1, \frac{(1-\sigma)\alpha + \sigma s}{1-(1-\sigma)\beta}\right\} \end{aligned} \tag{23}$$

Notice now that for *any* given positive value of σ, as $\beta \to 1$ and $\alpha \to 0$, the formulae in (23) tell us that $r^e \to s$ and $Y \to s$. In other words, no matter how small we choose to make σ, for parameter values sufficiently close to $\beta = 1$ and $\alpha = 0$, we can find a boundedly rational equilibrium which is arbitrarily close to $r^e = s$ and $Y = s$. These values can be quite far (depending on the 'direction' from which $\beta = 1$ and $\alpha = 0$ is being approached) from the fully rational equilibrium given by (20).

Haltiwanger and Waldman [15] work a more complex model that has a similar structure to the one we have used here. They limit their investigation to an open set of parameter values for which there is a unique equilibrium in the fully rational model. However, they note the existence of a continuum of equilibria for boundary parameter values. Our results show that it is precisely this structural instability in the model at the boundary that drives the fact that their model is not robust to bounded rationality near the boundary of the parameter space.

5.4. A Rational Expectations Macroeconomic Model

Our last example focuses on the effect of relaxing the constraint that agents have rational expectations in a macroeconomic model. We take as a measure of bounded rationality the difference between the agents' subjective expectation and the actual expected value of the variable given their information. Our goal is to investigate if a small amount of bounded rationality in this sense can have large repercussions on behavior and outcomes.

The rational expectation macroeconomic model that we work with in this section is taken from Blanchard and Fischer [6, Ch. 10.3]

Aggregate demand is given by

$$y^d = m - p + v \tag{24}$$

where m is the money stock, p is the price level and v is a non-degenerate random shock with an expected value of zero.

The equation for aggregate supply is

$$y^s = \beta(p - w + u) \tag{25}$$

where w is the wage rate, u is a non-degenerate random shock with an expected value of zero, and β is a non-negative constant. Labor demand is given by

$$n^d = \gamma(p - w + \alpha u) \tag{26}$$

with γ and α non-negative constants. Labor supply is written as

$$n^s = \delta(w - p) \tag{27}$$

where δ is a non-negative constant.

To these behavioral relationships they add the market clearing condition

$$y^d = y^s \tag{28}$$

In other words, market clearing in the goods market occurs automatically through flexible prices.

On the other hand, the labor market need not clear. Instead, wages are set so that expected labor demand equals expected labor supply. Under full rationality (rational expectations) the wage is set so that

$$E[n^d] = E[n^s] \tag{29}$$

where $E[\cdot]$ denotes the 'rational' expectation of a given variable, namely the expected value of that variable conditional on all available information.[13]

The microeconomic foundations of this canonical model (with rational expectations) are discussed at length in Blanchard and Fischer [6] and we will not repeat their discussion here. Instead, we begin by identifying the solution to the model under full rationality. This a useful benchmark case from which to start. Using (26) and (27), the fact that the wage must clear the labor market in expected terms, as given by (29), yields that

$$(\gamma + \delta)E[p] \; = \; (\gamma + \delta)w \tag{30}$$

Using market clearing in the goods market we can then determine the price level as follows

$$p \; = \; \frac{m + v + \beta(w - u)}{1 + \beta} \tag{31}$$

Using (30) and (31), it is immediate to see that if $(\gamma + \delta) > 0$ then the model has a unique equilibrium under full rationality given by

$$w \; = \; m \qquad \text{and} \qquad p \; = \; \frac{m + v + \beta(w - u)}{1 + \beta} \; = \; m + \frac{v - \beta u}{1 + \beta} \tag{32}$$

When $\gamma + \delta = 0$, equation (30) does not pin down the equilibrium wage anymore. In this case we assume that the wage can only take values in a given interval $[0, \overline{w}]$. Therefore, when $\gamma + \delta = 0$, the model has a continuum of fully rational equilibria given by

$$w \; \in \; [0, \overline{w}] \qquad \text{and} \qquad p \; = \; \frac{m + v + \beta(w - u)}{1 + \beta} \tag{33}$$

In terms of the abstract framework that we defined in Section 2, we define a typical element of the parameter space of the model as

$$\lambda = (\alpha, \beta, \gamma, \delta, m, u, v) \tag{34}$$

[13] Note that, later on, when we allow bounded rationality into the model, equation (29) will no longer necessarily hold.

Note that in this specification we treat the actual realizations of the random variables as parameters.

A typical element of the action space is

$$a = (w, p) \tag{35}$$

Note that the other outcomes, output and employment, can be written as simple functions of wages and prices using the behavioral equations of the model above.

To make the model fit our abstract framework developed in Section 2 we assume that the parameter and action spaces are compact subsets of $\mathbb{R}$. In other words, we assume that the parameters and action variables of the model satisfy[14,15]

$$\begin{array}{c} \alpha \in [0,1] \quad \beta \in [0,\overline{\beta}] \quad \gamma \in [0,\overline{\gamma}] \quad \delta \in [0,\overline{\delta}] \\ m \in [0,\overline{m}] \quad w \in [0,\overline{w} \quad u \in [\underline{u},\overline{u}] \quad v \in [\underline{v},\overline{v}]] \\ p \in [\underline{v} - \overline{\beta}\overline{u}, \overline{m} + \overline{v} + \overline{\beta}\overline{w} - \overline{\beta}\underline{u}] \end{array} \tag{36}$$

Clearly, given (36), the model satisfies Assumption 2.1.

We assume that wages are set on the basis that the labor market is expected to clear, but these expectations need not be correct. The measure of bounded rationality in the model is given by how far from clearing (because of non-rational expectations) the labor market is. We define the rationality function for the model as

$$R(\lambda, a) \;=\; \left| E(n^d - n^s) \right| \tag{37}$$

In other words, under bounded rationality, in an ϵ-Equilibrium, the wage is set so that the absolute value of the expected excess demand for labor is ϵ or less.

Given wages, we can determine prices using the equilibrium condition in the goods market. The feasibility correspondence constrains the price level to be the one that clears the goods market. In other words, we set

$$F(\lambda, a) = \left\{ w \in [0, \overline{w}],\ p = \frac{m + v + \beta(w-u)}{1+\beta} \right\} \tag{38}$$

Notice next that, given (38), the consistent actions correspondence is simply given by $\mathcal{C}(\lambda) = \mathcal{F}(\lambda, a)$. Therefore, by inspection of (38), the model satisfies Assumption 2.2 as required.

[14]Note that since we are assuming that $E[v] = E[u] = 0$, and since the random shocks are non-degenerate, it follows that in (36) above we have that $\underline{v} < 0$, $\underline{u} < 0$, $\overline{v} > 0$ and $\overline{u} > 0$. We also assume that $\overline{w} \geq \overline{m}$, and that $\overline{\beta}$, $\overline{\gamma}$, $\overline{\delta}$, $\overline{m}$ and $\overline{w}$ are all strictly positive.

[15]Note that the bounds for p in (36) are simply the result of plugging the bounds for all other variables and parameters of the model into the expression for p in (32) above.

Taking expectations, the rationality function defined in (37) can be rewritten as

$$R(\lambda, a) = (\gamma + \delta)\ |E[p] - w| = (\gamma + \delta) \left| \frac{m - w}{1 + \beta} \right| \tag{39}$$

Clearly, (39) shows that the rationality function is continuous as required, and therefore the model satisfies Assumption 2.3. It is also immediate that $R(\lambda, a) = 0$ corresponds to the equilibrium set of the original model with full rationality. Since we have already explicitly computed the equilibria of the model under full rationality[16] this is also enough to show that Assumption 2.4 is satisfied by the model. Hence, the model satisfies all the assumptions required by Theorem 3.1.

However, the model is not structurally stable since it is immediate from (32) and (33) that the equilibrium correspondence is discontinuous at $(\gamma + \delta) = 0$. Therefore, by Theorem 3.1, the model is not robust to bounded rationality in the neighborhood of this point.

Intuitively, what is happening is that γ and δ measure the elasticity of the demand and supply of labor with respect to the wage at the equilibrium point. Near the critical parameter values the demand and supply curves are very inelastic; small mistakes in predicting labor demand or supply can have arbitrarily large implications for nominal wage setting. These mistakes in nominal wage setting then have large impacts on prices and output (provided that $\beta > 0$).

To see this more directly, note that for parameters near the critical point, we can set the wage arbitrarily and still be epsilon-rational since $R(\lambda, a) = (\gamma + \delta)|(m - w)/(1 + \beta)|$ remains small. In other words, for any $w \in [0, \overline{w}]$ and for any $\epsilon > 0$ there exist a $z > 0$ such that whenever $0 < \gamma + \delta < z$, we can be sure that w is an ϵ-Equilibrium of the model. Clearly, this can be quite far from the unique fully rational equilibrium of the model, which, given (32), has $w = m$.

Blanchard and Fischer [6] analyze only the case where there is a unique rational expectations equilibrium by assuming $\gamma > 0$. We have extended the equilibrium analysis to include $\gamma = 0$. However, simply assuming $\gamma > 0$ does not get rid of the problem. If bounded rationality takes the form of expectations that are close to, but not quite rational, then when $\gamma + \delta$ is positive, but small, the model is not robust to bounded rationality.

6. CONCLUSIONS

The study of the effects of small amounts of bounded rationality can be carried out in many cases simply by studying the structural stability

[16] See (32) and (33) above.

of the equilibrium correspondence. This should greatly simplify the theoretical study of the impact of bounded rationality in particular models. In addition, since our results hold for any continuous rationality function, once we have studied robustness to one type of bounded rationality we can generalize the results to many other types.

Our results also have implications for experimental economics. Even in cases where an experimental situation has a unique equilibrium prediction, our results suggest that small deviations from full rationality will have a large impact on observed behavior if and only if the experimental setup is close to a critical value, where equilibrium behavior becomes structurally unstable. Failure of equilibrium predictions in such cases is not surprising, and might not lead us to abandon the assumption of rational (or close to rational) behavior. However, large scale failure of equilibrium predictions in experimental settings that are far from critical values of the equilibrium correspondence, does undermine the assumption of rational (or close to rational) behavior.

APPENDIX

A.1. DEFINITIONS

For the sake of completeness, we begin with some standard definitions[1] which are used throughout the paper.

Definition A.1 (Upper Hemicontinuity). *Let X and Y be complete separable metric spaces. The compact-valued correspondence $\mathcal{H} : X \rightrightarrows Y$ is upper hemicontinuous at x if and only if given any sequence $\{x^n, y^n\}$ such that $x^n \to x$, $y^n \to y$, and $y^n \in \mathcal{H}(x^n)$ for every n, we have that $y \in \mathcal{H}(x)$.*

The compact-valued correspondence $\mathcal{H} : X \rightrightarrows Y$ is upper hemicontinuous if an only if it is upper hemicontinuous at every $x \in X$.

Definition A.2 (Lower Hemicontinuity). *Let X and Y be complete separable metric spaces. The compact-valued correspondence $\mathcal{H} : X \rightrightarrows Y$ is lower hemicontinuous at x if and only if $x^n \to x$ and $y \in \mathcal{H}(x)$ imply that there exists a sequence $\{y^n\}$ such that $y^n \in \mathcal{H}(x^n)$ for every n and $y^n \to y$.*

The compact-valued correspondence $\mathcal{H} : X \rightrightarrows Y$ is lower hemicontinuous if an only if it is lower hemicontinuous at every $x \in X$.

[1]There are several different ways of characterizing upper and lower hemicontinuity. However, Border [8] shows that the sequential characterization used in Definitions A.1 and A.2 below is equivalent to the more general definition, using upper and lower inverses, when the mapping is compact-valued.

Definition A.3 (Continuity). *Let X and Y be complete separable metric spaces. The compact-valued correspondence $\mathcal{H} : X \rightrightarrows Y$ is continuous at x if and only if it is both upper and lower hemicontinuous at x.*

The compact-valued correspondence $\mathcal{H} : X \rightrightarrows Y$ is continuous if and only if it is continuous at every $x \in X$.

Definition A.4 (Equicontinuity). *Let (X, d_X) and (Y, d_Y) be complete separable metric spaces. Let $f_z : X \to Y$ describe a family of functions from X to Y as z varies in Z. In other words, for each given $z \in Z$, f_z is a function taking each point in X to a point in Y.*

The family of functions $f_z : X \to Y$ is said to be equicontinuous at x^* if and only if for every $\gamma > 0$ there exists an $\epsilon > 0$ such that

$$\mathrm{d}_X(x, x^*) \leq \epsilon \quad \Rightarrow \quad \mathrm{d}_Y\left[f_z(x), f_z(x^*)\right] \leq \gamma \qquad \forall\, z \in Z \tag{A.1}$$

A.2. PROOFS

Proof of Lemma 2.1 (ϵ-Equilibrium Correspondence). The fact that $\mathcal{E}$ is compact-valued is immediate from (1), using the fact that $\mathcal{C}$ is compact-valued by Assumption 2.2 and $\mathcal{R}$ is continuous by Assumption 2.3.

To see that $\mathcal{E}$ is upper hemicontinuous, consider a sequence $\{\lambda^n, \epsilon^n\}$ converging to (λ^*, ϵ^*), and any corresponding sequence a^n with $a^n \in \mathcal{E}(\lambda^n, \epsilon^n)$ for every n, and let a^n converge to a^*. We need to show that, necessarily, $a^* \in \mathcal{E}(\lambda^*, \epsilon^*)$.

The fact that $a^* \in \mathcal{C}(\lambda^*)$ is immediate from the fact that $\mathcal{C}$ is continuous by Assumption 2.2. It remains to show that $\mathcal{R}(\lambda^*, a^*) \leq \epsilon^*$. But since $a^n \in \mathcal{E}(\lambda^n, \epsilon^n)$ for every n, we must have that $\mathcal{R}(\lambda^n, a^n) \leq \epsilon^n$ for every n. Since $\mathcal{R}$ is continuous by Assumption 2.3 this directly implies $\mathcal{R}(\lambda^*, a^*) \leq \epsilon^*$, as required. ▌

Proof of Lemma 3.2 (Continuity at $\epsilon = 0$). By Lemma 2.1 we know that $\mathcal{E}_\lambda(\epsilon)$ is upper hemicontinuous. Using (1) it is immediate that if $a \in \mathcal{E}_\lambda(0)$, then $a \in \mathcal{E}_\lambda(\epsilon)$ for every $\epsilon > 0$. Therefore, $\mathcal{E}_\lambda(\epsilon)$ is lower hemicontinuous at $\epsilon = 0$. Hence it is continuous at $\epsilon = 0$. ▌

Proof of Theorem 3.1 (Robustness and Stability). We start by showing that robustness to ϵ-Equilibria implies structural stability.

We proceed by contradiction. Assume that we can find a model $\mathcal{M}$ that is robust to ϵ-Equilibria but not structurally stable.

Since $\mathcal{M}$ is not structurally stable, we can find some $\lambda^* \in \Lambda$ such that $\mathcal{E}(\lambda)$ is not lower hemicontinuous at λ^*. It follows that we can find a λ^*, an $a^* \in \mathcal{E}(\lambda^*)$, a sequence $\lambda^n \to \lambda^*$, and a $\delta > 0$ such that

$$\inf_{a \in \mathcal{E}(\lambda^n)} \mathrm{d}_{\mathcal{A}}(a^*, a) > \delta \quad \forall n \tag{A.2}$$

By (1), $a^* \in \mathcal{E}(\lambda^*)$ implies $a^* \in \mathcal{C}(\lambda^*)$. Therefore, since $\mathcal{C}$ is continuous and non-empty by Assumption 2.2, we can also find a sequence $a^n \to a^*$ with $a^n \in \mathcal{C}(\lambda^n)$ for every n. Therefore $(\lambda^n, a^n) \in \mathcal{G}$ for every n. Since $\{\lambda^n, a^n\} \to (\lambda^*, a^*)$ by construction, and since $\mathcal{R}$ is continuous by Assumption 2.3, using the fact that $a^* \in \mathcal{E}(\lambda^*)$, we can now conclude that as $n \to \infty$ it must be the case that $\mathcal{R}(\lambda^n, a^n) \to 0$. Therefore

$$\forall \epsilon > 0 \;\; \exists \overline{n} \quad \text{such that} \quad n > \overline{n} \;\Rightarrow\; a^n \in \mathcal{E}(\lambda^n, \epsilon) \tag{A.3}$$

Using (A.3) we now know that given any sequence $\epsilon^k \to 0$, we can find subsequences $\lambda^{n^k} \to \lambda^*$ and $a^{n^k} \to a^*$ such that

$$a^{n^k} \in \mathcal{E}(\lambda^{n^k}, \epsilon^k) \quad \forall k \tag{A.4}$$

Using (2) to write out in full $\mathrm{h}_{\mathcal{A}}[\mathcal{E}(\lambda^{n^k}, \epsilon^k),\, \mathcal{E}(\lambda^{n^k})]$, it is immediate to verify that (A.4) implies that

$$\mathrm{h}_{\mathcal{A}}\left[\mathcal{E}(\lambda^{n^k}, \epsilon^k), \mathcal{E}(\lambda^{n^k})\right] \geq \inf_{a \in \mathcal{E}(\lambda^{n^k})} \mathrm{d}_{\mathcal{A}}(a^{n^k}, a) \quad \forall k \tag{A.5}$$

Using the triangular inequality on the right-hand side of (A.5) now yields

$$\mathrm{h}_{\mathcal{A}}\left[\mathcal{E}(\lambda^{n^k}, \epsilon^k), \mathcal{E}(\lambda^{n^k})\right] \geq \inf_{a \in \mathcal{E}(\lambda^{n^k})} \left[\mathrm{d}_{\mathcal{A}}(a^*, a) - \mathrm{d}_{\mathcal{A}}(a^*, a^{n^k})\right] \;\; \forall k \tag{A.6}$$

Using (A.2), from (A.6) we can then directly conclude that

$$\mathrm{h}_{\mathcal{A}}\left[\mathcal{E}(\lambda^{n^k}, \epsilon^k), \mathcal{E}(\lambda^{n^k})\right] \geq \delta - \mathrm{d}_{\mathcal{A}}(a^{n^k}, a^*) \quad \forall k \tag{A.7}$$

Since, by construction, $a^{n^k} \to a^*$, we obviously have that $\mathrm{d}_{\mathcal{A}}(a^{n^k}, a^*) \to 0$. Using (A.7), it is therefore clear that we have showed the following. Given any sequence $\epsilon_n \to 0$, we can find some $\gamma > 0$ and a sequence λ^n such that

$$\mathrm{h}_{\mathcal{A}}[\,\mathcal{E}(\lambda^n, \epsilon^n), \mathcal{E}(\lambda^n)\,] > \gamma \quad \forall n \tag{A.8}$$

Finally, notice that (A.8) directly contradicts (3) and hence contradicts the hypothesis that $\mathcal{M}$ is robust to ϵ-Equilibria. Therefore, this is enough to conclude the proof that robustness to ϵ-Equilibria implies structural stability.

To show that structural stability implies robustness to ϵ-Equilibria we proceed by contradiction again. Let $\mathcal{M}$ be a model that is structurally stable but not robust to ϵ-Equilibria.

Since $\mathcal{M}$ is not robust to ϵ-Equilibria we must be able to find some $\delta > 0$, a sequence λ^n and a sequence $\epsilon^n \to 0$ with $\epsilon^n > 0$ for every n such that

$$\mathrm{h}_{\mathcal{A}}[\,\mathcal{E}(\lambda^n, \epsilon^n), \mathcal{E}(\lambda^n)\,] > \delta \qquad \forall\, n \tag{A.9}$$

Since it must always be the case that $\mathcal{E}(\lambda^n) \subseteq \mathcal{E}(\lambda^n, \epsilon^n)$, (A.9) implies that we can now select some sequence a^n with $a^n \in \mathcal{E}(\lambda^n, \epsilon^n)$ and $a^n \notin \mathcal{E}(\lambda^n)$ for every n and find some $\gamma > 0$ such that

$$\inf_{a \in \mathcal{E}(\lambda^n)} \mathrm{d}_{\mathcal{A}}(a^n, a) \,>\, \gamma \qquad \forall\, n \tag{A.10}$$

Notice next that since $a^n \in \mathcal{E}(\lambda^n, \epsilon^n)$, we have that $a^n \in \mathcal{C}(\lambda^n)$ and $\mathcal{R}(\lambda^n, a^n) \leq \epsilon^n$ for every n. Since Λ and $\mathcal{A}$ are compact and $\epsilon^n \to 0$, using the continuity of $\mathcal{C}$ and $\mathcal{R}$, we can now find convergent subsequences $\lambda^{n^k} \to \lambda^*$ and $a^{n^k} \to a^*$ such that $a^* \in \mathcal{C}(\lambda^*)$ and $\mathcal{R}(\lambda^*, a^*) = 0$. Notice also that the latter implies that $a^* \in \mathcal{E}(\lambda^*)$

Since $a^{n^k} \to a^*$, (A.10) implies that

$$\inf_{a \in \mathcal{E}(\lambda^{n^k})} \mathrm{d}_{\mathcal{A}}(a^*, a) \,\geq\, \gamma \qquad \forall\, k \tag{A.11}$$

Since $a^* \in \mathcal{E}(\lambda^*)$ this implies that

$$\mathrm{h}_{\mathcal{A}}[\,\mathcal{E}(\lambda^*), \mathcal{E}(\lambda^{n^k})] \geq \gamma \qquad \forall\, k \tag{A.12}$$

But since $\lambda^{n^k} \to \lambda^*$, (A.12) clearly contradicts the fact that, since $\mathcal{M}$ is structurally stable, $\mathcal{E}$ must be continuous at λ^*. This is enough to prove that every model that is structurally stable must also be robust to ϵ-Equilibria. Hence the proof of the theorem is complete. ▮

Proof of Remark 4.2 (Rational Equilibria and Perfect Equilibria). By (1) and (6) it is obvious that any Exact ϵ-Equilibrium is in fact an ϵ-Equilibrium.

The claim is now an immediate consequence of the fact that, by Lemma 2.1, for any given $\lambda \in \Lambda$, the ϵ-Equilibrium Correspondence is upper hemi-continuous at $\epsilon = 0$. ▮

REFERENCES

1. George A. Akerlof and Janet L. Yellen. A Near-Rational Model of the Business Cycle with Wage and Price Inertia. *Quarterly Journal of Economics*, 100:823–838, 1985.
2. George A. Akerlof and Janet L. Yellen. Can Small Deviations from Rationality Make Significant Differences to Economic Equilibria? *American Economic Review*, 75:708–720, 1985.
3. Kenneth J. Arrow and Gerard Debreu. Existence of Equilibrium for a Competitive Economy. *Econometrica*, 22:265–290, 1954.
4. Yves Balasko. *Foundations of the Theory of General Equilibrium*. Academic Press, New York, 1988.
5. C. Berge. *Topological Spaces*. MacMillan, New York, 1963.
6. Olivier Jean Blanchard and Stanley Fischer. *Lectures on Macroeconomics*. The MIT Press, Cambridge, Massachusetts, 1989.
7. Antulio N. Bonfim and Francis X. Diebold. Bounded Rationality and Strategic Complementarity in a Macroeconomic Model: Policy Effects, Persistence and Multipliers. *Economic Journal*, 107:1358–1374, 1997.
8. K. C. Border. *Fixed Point Theorems with Applications to Economics and Game Theory*. Cambridge University Press, Cambridge, 1985.
9. John Conlisk. Costly Optimizers Versus Cheap Imitators. *Journal of Economic Behavior and Organization*, 1:275–293, 1980.
10. Gerard Debreu. A Social Equilibrium Existence Theorem. *Proceedings of the National Academy of Sciences*, 38:886–893, 1952.
11. Gerard Debreu. *Theory of Value*. Yale University Press, New Haven, 1959. Cowles Foundation Monograph No. 17.
12. George Evans and Garey Ramey. Expectations Calculations and Macroeconomic Dynamics. *American Economic Review*, 82:207–224, 1992.
13. Drew Fudenberg and David Levine. Limit games and limit equilibria. *Journal of Economic Theory*, 38:261–279, 1986.
14. J. Haltiwanger and M. Waldman. Rational Expectations and the Limits of Rationality: An Analysis of Heterogeneity. *American Economic Review*, 75:326–340, 1985.
15. John C. Haltiwanger and Michael Waldman. Limited Rationality and Strategic Complements: The Implications for Macroeconomics. *Quarterly Journal of Economics*, 104:463–483, 1989.
16. John C. Haltiwanger and Michael Waldman. Rational Expectations in the Aggregate. *Economic Inquiry*, 27:619–636, 1989.
17. John Harsanyi. Oddness of the Number of Equilibrium Points. *Internaional Journal of Game Theory*, 2:235–250, 1973.
18. Werner Hildenbrand. *Core and Equilibria of a Large Economy*. Princeton University Press, Princeton, 1974.
19. Erev Ido and Alvin E. Roth. Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique Mixed Strategy Equilibria. *American Economic Review*, 88:848–881, 1998.

20. Stephen R. G. Jones and James H. Stock. Demand Disturbances and Aggregate Fluctuations: The Implications of Near Rationality. *Economic Journal*, 97:49–64, 1987.

21. N. Gregory Mankiw. Small Menu Costs and Large Business Cycles: A Macroeconomic Model. *Quarterly Journal of Economics*, 100:529–538, 1985.

22. Andreu Mas-Colell. *The Theory of General Economic Equilibrium: A Differentiable Approach*. Cambridge University Press, Cambridge, 1985.

23. Roger Myerson. Refinements of the Nash Equilibrium Concept. *International Journal of Game Theory*, 7:73–80, 1978.

24. John Nash. Equilibrium Points in N-Person Games. *Proceedings of the National Academy of Sciences*, 36:38–39, 1950.

25. Roy Radner. Collusive Behaviour in Oligopolies with Long but Finite Lives. *Journal of Economic Theory*, 32:136–156, 1980.

26. Alvin E. Roth. Introduction to Experimental Economics. In John H. Kagel and Alvin E. Roth, editors, *The Handbook of Experimental Economics*, pages 3–109. Princeton University Press, Princeton, 1995.

27. Ariel Rubinstein. *Modeling Bounded Rationality.* M.I.T. Press, Cambridge, Massachusetts, 1998.

28. Thomas Russell and Richard Thaler. The Relevance of Quasi Rationality in Competitive Markets. *American Economic Review*, 75:1071–1082, 1985.

29. Reinhard Selten. A Re-examination of the Perfectness Concept for Equilibrium Points in Extensive Games. *International Journal of Game Theory*, 4:25–55, 1975.

30. Rajiv Sethi and Reiner Franke. Behavioural Heterogeneity under Evolutionary Pressure: Macroeconomic Implications of Costly Optimization. *Economic Journal*, 105:583–600, 1995.

31. Y. Wang. Non-Rational Behaviour in Financial Market Fluctuations. *Economic Journal*, 103:1462–1478, 1993.

32. Jorgen W. Weibull. *Evolutionary Game Theory.* M.I.T. Press, Cambridge, Massachusetts, 1995.

33. Robert Wilson. Computing Equilibria of N-Person Games. *SIAM Journal of Applied Mathematics*, 21:80–87, 1971.

34. Wen-Tsün Wu and Jiang Jia-He. Essential Equilibrium Points of N-Person Non-Cooperative Games. *Scientia Sinica*, 11:1307–1322, 1962.