Econometric Analysis of Games with Multiple Equilibria

Áureo de Paula

Department of Economics, University College London, London WC1E 6BT, United Kingdom; Center for Microdata Methods and Practice (CeMMAP) and Institute for Fiscal Studies, London WC1E 7AE, United Kingdom; email: a.paula@ucl.ac.uk

Annu. Rev. Econ. 2013. 5:107–31

First published online as a Review in Advance on August 29, 2013

The Annual Review of Economics is online at economics.annualreviews.org

This article's doi: 10.1146/annurev-economics-081612-185944

JEL codes: C1, C3, C7

identification, multiplicity, social interactions

This article reviews the recent literature on the econometric analysis of games in which multiple solutions are possible. Multiplicity does not necessarily preclude the estimation of a particular model (and, in certain cases, even improves its identification), but ignoring it can lead to misspecifications. The review starts with a general characterization of structural models that highlights how multiplicity affects the classical paradigm. Because the information structure is an important guide to identification and estimation strategies, I discuss games of complete and incomplete information separately. Although many of the techniques discussed here can be transported across different information environments, some are specific to particular models. Models of social interactions are also surveyed. I close with a brief discussion of post-estimation issues and research prospects.

1. I D I

In this article I review the recent literature on the econometric analysis of games in which multiple solutions are possible. Equilibrium models are a defining ingredient of economics. Game theoretic models, in particular, have played a prominent role in various subfields of the discipline for many decades. When taking these models to data, one endows a sample of games represented by markets, neighborhoods, or economies with an interdependent payoff structure that depends on observable and unobservable variables (to the econometrician and potentially to the players), and participants choose actions. One pervasive feature in many of these models is the existence of multiple solutions for various payoff configurations, and this aspect carries over to estimable versions of such systems.

Although the existence of more than one solution for a given realization of the payoff structure does not preclude the estimation of a particular model (and, in certain cases, even improves their identification), ignoring its possible occurrence can potentially cause severe misspecifications and nonrobustness in the analysis of substantive questions. Fortunately, much has been learned in the recent past about the econometric properties of such models. The tools available benefit from advances in identification analysis, estimation techniques, and computational capabilities, and I discuss some of these below.

Owing to space limitations, the survey is by no means exhaustive. I nevertheless cover some of the main developments thus far. Because most of the literature has concentrated on parametric models, this is also my focus here. As in many other contexts, the parametric and functional restrictions that are imposed deserve careful deliberation, and some of the parametric and functional restrictions in the models I present can be relaxed (e.g., the linearity of the parametric payoff function and the distributional assumptions in theorem 2 of Tamer 2003 and the analysis of social interactions models in Brock & Durlauf 2007). Once point or partial identification has been established, estimation typically proceeds by applying well-understood methods such as maximum likelihood and method of moments (many times with the assistance of simulations) in the case of point-identified models or by carrying out recently developed methods for partially identified models. A thorough discussion of partially identified models would require much more space, and I leave that for other surveys covering those methods in more detail (see, e.g., Tamer 2010). I nevertheless do discuss estimation and computation aspects that are somewhat peculiar to the environments described below.

In the games analyzed here, given a set of payoffs for the economic agents involved, a solution concept defines the (possibly multiple) outcomes that are consistent with the economic environment. The solution concepts I use below essentially consist of mutual best responses (plus consistent beliefs when information is asymmetrically available), and I refer to those as equilibria or solutions indiscriminately (hopefully without much confusion to the reader). In the following sections, the solution concepts are Nash equilibrium for complete-information games, Bayes-Nash or Markov perfect equilibrium for incomplete-information games, and rational expectations equilibrium as defined in the social interactions literature for those types of models. Although these are commonly assumed solution concepts, others exist. Aradillas-Lopez & Tamer (2008), for example, consider rationalizable strategies, and network-formation games rely on pairwise stability or similar concepts. Multiplicity is often an issue for these alternative definitions, and many of the ideas discussed below (e.g., bounds) can be used when those concepts are adopted instead.

One important ingredient guiding identification and estimation strategies in these models is the information environment of a game. Whether a game is one of complete or incomplete (i.e., private) information may affect the econometric analysis in a substantive manner. Many techniques discussed below can be transported across these different information environments, but some are specific to particular models. In the next sections, I discuss identification and esti-

possible are intermediary cases in which, with some probability [potentially dependent on the realizations of and and the parameters in the model $u \equiv (b_1, b_2, \Delta_1, \Delta_2)$], say $l(\ ,\ , u)$ 2 [0, 1], one of the two equilibria is selected whenever payoffs fall in the multiplicity region. Different selection probabilities [i.e., $l(\ ,\ , u)$ in the example] will induce different distributions over the observable outcomes y_i.

One could (and in many examples below does) include the equilibrium selection mechanism into the structure. Nevertheless, one must bear in mind that modeling the equilibrium selection process requires extra assumptions. This opens up an additional avenue for misspecification. Moreover, an estimated equilibrium selection mechanism is more likely to be policy sensitive. This is because

¼ ðu_1, u_2þ 2 [$_1^u$ b_1 Δ$_1$, $_1^u$ b_1] × [$_2^u$ b_2 Δ$_2$, $_2^u$ b_2]. In this case, both ¼ (0, 0) and ¼ (1, 1) are possible equilibria. [¼ (0, 0) is a unique equilibrium when u_i < $_i^u$ b_i Δ$_i$, i ¼ 1, 2, and ¼ (1, 1) is a unique equilibrium when u_i > $_i^u$ b_i, i ¼ 1, 2.] The number of players choosing 1 is no longer the same across equilibria, but one could nonetheless mimic the previous strategy and consider the probability of events {(0, 1)}, {(1, 0)}, and {(1, 1), (0, 0)}, where one pools together any two outcomes that are both equilibria for some given and . Once this is done, singleton events correspond to outcomes that oc-240.7(oat)1470.

the previous subsection would be inconclusive, as in Jovanovic

$$\mathbb{P}(\ \mid\ ;(u, F\ )) = \int l(\ \mid E(\ ,\ ,u),\ ,\ ,u)\,\mathbf{1}_{\ \in E(\ ,\ ,u)} f(\ )\,d\ , \qquad (2)$$

where $f(\cdot)$ is the PDF for , which is assumed to be independent of .

Given the specification summarized in Equation 2, it is unclear whether the model identifies the equilibrium selection mechanism $l(\cdot)$. In fact, unless further restrictions are imposed, it does not. To take an extreme, but simple illustration, consider again the example in Jovanovic (1989) with no covariates. In this case, let l denote the probability that (1, 1) is selected whenever there are multiple equilibria (i.e., $(u_1, u_2) \in [-\Delta, 0]^2$). Then

$$\mathbb{P}(\ = (1,1)) = l\Delta^2 \quad \text{and} \quad \mathbb{P}(\ = (0,0)) = 1 - \mathbb{P}(\ = (1,1)).$$

It is not possible to pin down l and Δ from the distribution of outcomes. One of the issues here is that there are more unknowns than there are equations. To reduce the degrees of freedom in the problem, Bajari et al. impose additional structure. For example, the number of equations can be increased if the support of covariates is relatively large, generating additional conditional moments. To keep the number of parameters under control, Bajari et al. assume that selection

$$\sum_{g=1}^{G}\left(1_{\ g=a} - \widehat{\mathbb{P}}(\ \ = a |\ \ ; u, g, F\ )\right) h(\ \ ),$$

where $\widehat{\mathbb{P}}$ is a computer-simulated estimate of P.

expect, given the distribution of types for the other individual involved in the game (i.e., the private-information components of their payoffs, u_i). As before, I consider only pure-strategy equilibria for simplicity. Given i's opponent's strategy, the best response dictates that

$$y_i = 1 \quad \text{if} \quad {}_i^u b_i + \mathbb{P}(y_j = 1 | \; , u_i)\Delta_i + u_i \geq 0, \quad j \neq i, \qquad (3)$$

As in the previous section, I now discuss different approaches to inference in games of incomplete information like the one just outlined above.

4.1. D h

The previous discussion underscores the benefits of further restrictions on the equilibrium selection mechanism for the econometric analysis of incomplete-information games with possibly many equilibria. One common strategy is to assume that

$$I(k|E(\ , u, F), \ , u, F) = 1_{k=K}$$

for some $K \in E(\ , u, F)$. In words, whenever primitives and covariates coincide for two games, thus inducing an identical equilibrium set, the same equilibrium is played in these two games. One can nevertheless be agnostic about which equilibrium is selected [i.e., which element of $E(\ , u, F)$ is selected].

When is it realistic to assume that the same equilibrium is played across games? As Mailath (1998) points out, “

The assumption of a unique equilibrium in the data is crucial to travel from the conditional

Notice that with a unique equilibrium in the data, Equation 5 corresponds to the distribution of a binomial random variable (with parameters 2 and p). Using simulations, Sweeting (2009, p. 723) notes that when $\Delta > 0$, "a mixture generates greater variance in the number of stations choosing a particular outcome than can be generated by a single binomial component." As he points out, this suggests that multiplicity provides additional information about the payoff structure of the game under analysis.

De Paula & Tang (2012) formalize and generalize this idea in many directions. For the basic insight, take the expression in Equation 4 and compare two equilibria where $p_j^h(\ ) > p_j^l(\ )$. If $\Delta_i > 0$, it has to be the case that $p_i^h(\ ) > p_i^l(\ )$. (This is because $F_{u_{ij}}$ is increasing as it is a CDF.) Conversely, if $\Delta_i < 0$, one must have $p_i^h(\ ) < p_i^l(\ )$

Finally, I note that the essential assumption of the conditional independence of the latent variables is also commonly found in dynamic games of incomplete information. Optimal decision rules in those settings involve not only equilibrium beliefs but also continuation value functions that may change across equilibria. Nevertheless, the characterization of optimal policy rules in that context suggests that the existence of a unique equilibrium in the data can still be detected by the lack of correlation in actions across players of a given game, as presented here in the case of the static game. [The identification of sign(Δ_i) would require additional restrictions, however.] Because most of the known methods for the semiparametric estimation of incomplete information (static or dynamic) rely on the existence of a single equilibrium in the data (see above), a formal test for the assumption of a unique equilibrium in the data-generating process can be quite useful.

4.3. G -

To establish proposition 1 in de Paula & Tang (2012), it is paramount that the latent variables be conditionally independent. Any association between u_1 and u_2 will lead to correlation in actions even under a unique equilibrium, but it will also change the nature of equilibrium decision rules in important ways [i.e., $\mathbb{P}(y_j$ ¼ 1j , $u_i)$ in Equation 3 is now a nontrivial function of u_i

different editions of the game within a particular market, different equilibria are allowed across different markets. Similar ideas also appear in other papers in the literature, such as Bajari et al. (2007) and Pesendorfer & Schmidt-Dengler (2008). Even in the absence of long panels, Bajari et al. (2010a) suggest a few interesting strategies, such as the use of conditional likelihood methods when the n errors are logistic or the use of Manski's (1987) panel data rank estimator. These

$$u_i\left(b + \Delta \mathbb{E}\left[\frac{\sum_{j \neq i} y_j}{N - 1}\right]\right) = \mathbb{E}\left[u_i\left(b + \Delta \frac{\sum_{j \neq i} y_j}{N - 1}\right)\right],$$

the (equilibrium) expected utility agrees with the utility at the expected (equilibrium) profile of actions. This is, in particular, the best response predicament in my example with two players and two actions under incomplete information when $b_1 = b_2$ and $\Delta_1 = \Delta_2$

(2011). Bisin et al. present Monte Carlo evidence in a model with possibly many equilibria for certain parameter configurations that highlights the computational costs and statistical properties of the two estimators. Because the asymptotic approximations rely on $N \to 1$, I must point out that the econometric estimators in such large population games might present some delicate issues given the dependence in equilibrium outcomes within a game as the number of players grows. This is a topic of ongoing research (see, e.g., Menzel 2010, Song 2012).

5.2. A

As is the case in the previous section, group-level unobserved heterogeneity is potentially important in many applications. Ignoring it essentially rules out an important channel of unobserved contextual effects (or correlated effects) in the terminology coined by Manski (1993). Brock & Durlauf (2007, section 4) also discuss a series of potential scenarios that would allow the model to identify (at least partially) the parameters of interest. These include the use of panels, restrictions on the distribution of unobserved group shocks (i.e., large support, stochastic come to be played in the data. Even when the game is estimated under the assumption that a unique equilibrium is played in the data, the possession of estimated parameters allows one to go back and calculate all the potential equilibria for a particular parametric configuration. In their study of stock analysts' recommendations, for instance, Bajari et al. (2010a) notice the existence of multiple equilibria before New York Attorney General Eliot Spitzer launched a series of investigations on conflict of interest, with one equilibrium yielding much more optimistic ratings than those granted in the equilibrium post-Spitzer.

6.2. A h

Above I try to present many tools used in the econometric analysis of games with multiple equilibria. There is nevertheless still much to be understood in these settings. One interesting avenue that appears in some papers cited here is the connection with panel data methods. Just as the distribution of outcomes in game theoretic models is a mixture over equilibrium-specific outcome distributions under multiplicity, the observable distribution of outcomes in panel data models is a mixture over the distribution of individuals effects. Important idiosyncrasies such as the (typical) finiteness of the equilibrium set (which would correspond to a finite support for the individual effects) may help bring in interesting technical results in the panel literature to shed light on some properties of econometric game theoretic models. Examples of such studies include Hahn & Moon (2010) and Bajari et al. (2011). Here a important caveat, mentioned above, is that the cardinality of the equilibrium set, jE(u, F ,)j, will depend on the covariates and parametric configurations, whereas the support of the individual effects in the usual panel data model suffers no such restrictions. This might introduce important complications.

In a similar fashion, Grieco (2012) and Chen et al. (2011) suggest treating the equilibrium selection mechanisms as a (possibly infinite dimensional) nuisance parameter that is concentrated out in a profile sieve–maximum likelihood estimator procedure aimed at estimating semi-parametric partially identified models. Again, in this case, the dependence of the cardinality of the equilibrium set, jE(u, F ,)j, on the covariates and parametric configurations might introduce subtle complications, as the class of functions that contain the equilibrium set might have to vary with u, F , and to accommodate this dependence.

6.3. A

Recent applications in other areas of economics are available (e.g., Card & Giuliano 2011, Todd & Wolpin 2012) and are likely to become more common.

D I A

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

A K DG

I thank Tim Halliday, Chris Julliard, Elie Tamer, and Matthew Weinberg for discussions and suggestions. I am grateful to the Reviewing Editor, Chuck Manski, who provided detailed comments that markedly improved this review. Financial support from the Economic and Social Research Council through the ESRC Centre for Microdata Methods and Practice grant RES-589-28-0001 and the National Science Foundation through grant award number SES 1123990 is gratefully acknowledged.

I A I D

Ackerberg D, Gowrinsankaran G. 2006. Quantifying equilibrium network externalities in the ACH banking industry. Rand J. Econ. 37:738–61

Aguirregabiria V, Mira P. 2007. Sequential estimation of dynamic discrete games. Econometrica 75:1–53

Aguirregabiria V, Mira P. 2012. Structural estimation of games with multiple equilibria in the data. Work. Pap., Univ. Toronto

Andrews D, Berry S, Jia P. 2004. Confidence regions for parameters in discrete games with multiple equilibria, with an application to discount chain store location. Work. Pap., Yale Univ., New Haven, CT

Aradillas-Lopez A. 2010. Semi-parametric estimation of simultaneous games with incomplete information. J. Econom. 157:409–31

Aradillas-Lopez A, Tamer E. 2008. The identification power of equilibrium in simple games. J. Bus. Econ. Stat. 26:261–310

Brock W, Durlauf S. 2001. Discrete choice with social interactions. Rev. Econ. Stud. 62:235–60
Brock W, Durlauf S. 2007. Identification of bina

Manski C. 1988. Identification of binary response models. J. Am. Stat. Assoc. 83:729–38
Manski C. 1993. Identification of endogenous social effects: the reflection problem. Rev. Econ. Stud. 60:531–42
Menzel K. 2010. Inference for large games with exchangeable players. Work. Pap., New York Univ.

Annual Review of Economics

Volume 5, 2013

Contents

Indexes

Errata

An online log of corrections to articles may be found at http://econ.annualreviews.org