Approximate generalizations and applied equilibrium analysis*

Felix Kubler
Department of Economics
University of Mannheim
fkubler@rumms.uni-mannheim.de

First Version: April 18, 2004
This Version: April 29, 2005

Abstract

In this paper I derive conditions on the fundamentals of general equilibrium models that allow for a generalization of finitely many examples to statements about (infinite) classes of economies and I show how these approximate generalizations can be applied in computational experiments.

If there exist upper bounds on the number of connected components of one-dimensional linear subsets of the set of parameters for which a conjecture is true, one can conclude that it is correct for all parameter values in the class considered, except for a small residual set, once one has verified it for a predetermined finite set of points. I spell out assumptions on economic fundamentals which ensure that these bounds on the number of connected components exist, and that the residual set can be bounded from above.

I argue that utility- and production functions used in applied equilibrium analysis satisfy these conditions. Using the theoretical results, I show how computational experiments can be used to explore qualitative and quantitative implications of economic models. I give examples for actual upper bounds in realistically calibrated economies and discuss both deterministic and random algorithms for generalizing examples in these economies.

*I thank Don Brown, Dave Cass, Ken Judd, Narayana Kocherlakota, Mordecai Kurz, George Mailath, Marcel Richter, Klaus Ritzberger as well as seminar participants at various universities and conferences for helpful comments and discussions.

1 Introduction

Computational methods are widely used as a tool to study quantitative features of non-linear general equilibrium models. For a given example-economy, i.e. a given specification of preferences, technologies and market-arrangements these methods compute an (approximate) equilibrium and allow for quantitative statements about one equilibrium of the example-economy. In general, it is not feasible to use computational methods to prove that for all parameter values in a given (infinite) set these quantitative features of the model hold. Even a billion examples cannot prove a proposition which describes a continuum of different cases. However, under weak assumptions on the economic model, sufficiently many examples can tell us something about the size of the set of exogenous parameters for which the proposition is true. It is possible to use computational experiments to formulate general conjectures and to infer from the experiments that these conjectures are correct except perhaps for a small set of parameters. This method is applicable to competitive equilibrium models as well as to models with strategic interaction.

I consider economic questions which can be formulated as a statement (which I call formula[1]) which takes the form $\phi(e, a)$ where e is a vector of exogenous variables and a is a vector summarizing relevant quantitative properties of the endogenous and/or exogenous variables. For example, for a model with asset markets a typical statement might be *there exist prices and allocations such that markets clear and agents' first order conditions hold and excess returns on stocks are above a percent when the coefficient of risk aversion and individual endowments are given by* e. Of course, purely qualitative statements of the form $\phi(e)$ are a special case of the general framework.

I assume that the exogenous variables lie in some compact set $E \subset \mathbb{R}^l$. Suppose a series of computational experiments can find some $\bar{a}$ such that the statement is true for all e in a given finite set of parameter values i.e. for all $e \in F$. I introduce assumptions on ϕ which ensure that for any $\epsilon > 0$, the set F can be constructed so that one can infer that ϕ is true for all elements in E except for a subset with Lebesgue measure smaller than ϵ.

This is quite simple if the set of interest, $\Phi = \{e \in E : \phi(e) \text{ holds}\}$, can be shown to be convex. Unfortunately, in interesting economic problems this is almost never the case. However, depending on utility and production functions (for example if they are all Cobb-Douglas) it might be the case that the set Φ can be described as the Boolean combination of sets of the form

$$\{x \in \mathbb{R}^l : f(x) > 0\} \text{ or } \{x \in \mathbb{R}^l : g(x) = 0\}$$

for polynomials with real coefficients f and g, i.e. they are semi-algebraic sets. For semi-algebraic sets one can explicitly bound the number of connected components.

Koiran (1995) shows that if there exist upper bounds on the number of connected

[1] As will be made precise below I consider sentences in a first order language underlying an o-minimal expansion of the real field. For now the reader should think of this simply as a well defined statement which is either true or false.

components of the intersection of axes-parallel lines and the set Φ one can construct lower bounds on the size of the set Φ by knowing that ϕ holds at sufficiently many points in a grid. More specifically, Koiran shows that if the bound is given by λ and if $\phi(\bar{e})$ holds for all $\bar{e}$ on a grid, $F = [1/N, ..., 1]^l$, then the l-dimensional Lebesgue measure of $\Phi_{\bar{a}} \cap [0,1]^l$ is at least $1 - l\lambda/N$. The problem of proving that conjectures hold approximately thus reduces to finding bounds on the number of connected components of the set defined by the economic statement.

While it turns out that for many economic problems the sets of interest will not be semi-algebraic, I will argue that they are *definable in an 'o-minimal' structure*, as in Richter and Wong (1999) and Blume and Zame (1993) (what this means will be explained in Section 2). This ensures that there exist bounds on the relevant number of connected components. In particular, utility functions and production functions used in applied work are almost always 'Pfaffian functions' (see e.g. Khovanskii (1991) for definitions and motivations) and I show how the available bounds from Gabrielov and Vorobjov (2004) on the number of connected components of the set of solutions to Pfaffian equations can be used to derive upper bounds on the number of connected components of sets which are relevant for computational experiments.

In applied general equilibrium analysis two practical problems occur which make a direct application of the methods difficult.

First, due to rounding and truncation errors it is often not possible to compute exact equilibria and one has to resort to ϵ-equilibria instead. In this case, given a $\bar{e} \in E$ one often cannot determine whether $\phi(\bar{e})$ holds. Instead, all one can verify is that there exists some $\tilde{e}$ with $\|\tilde{e} - \bar{e}\| < \epsilon$ such that $\phi(\tilde{e})$ holds (see Kubler and Schmedders (2005) for a discussion of approximate and exact equilibria). I show that by using bounds on the number of connected components of the intersection of Φ with axes-parallel cylinders, the basic argument can be extended to this case. Verifying a conjecture about exact equilibria at sufficiently many ϵ-equilibria allows statements about the size of the set of parameters for which the conjecture holds. This gives a justification for the commonly used practice to consider ϵ-equilibria in applications.

Secondly, interesting formulas ϕ contain existential quantifiers (such as 'there exist allocations and prices such that agents optimize and markets clear and...') and even for semi-algebraic economies (i.e. for economies where all fundamentals can be described by polynomial inequalities and equalities) known bounds on the number of connected components grow exponentially as the dimension of the sets grows. Moreover, even with a fixed number of connected components, the number of examples one has to compute grows exponentially with the dimension of E. It turns out that, even for medium-sized problems, the methods are therefore often not directly applicable. An alternative is to use a random algorithm and to make statements about the size of the set of interest which are correct with high probability (see Judd (1997) or Blum et al. (1998)). Using a pseudo-random number generator, one can randomly draw values for the exogenous parameters and after

sufficiently many draws, for any given δ, my results then imply bounds on the probability that the true residual set is less than δ. These random algorithms are applicable even for problems for which known bounds on the number of connected components are relatively large, as long as they are orders of magnitude smaller than the errors in the computations.

The paper is organized as follows. In Section 2, I provide the necessary mathematical results. Section 3 explains how these results can be used in computational general equilibrium analysis and gives examples of computational experiments. Section 4 concludes.

2 Approximate generalizations

The key idea of the paper is to consider models where one can find bounds on the number of connected components of the sets of interest. In order to state the theorem which gives sufficient conditions to ensure that the number of connected components can be bounded, I use some notation from model theory. The following section gives a very brief overview and introduces some notation. See Marker (2002) for an introduction to model theory.

2.1 A little model theory

I consider a structure on the real numbers $\mathbb{R}$ which I denote by $\mathcal{M}$. The underlying first order language, $\mathcal{L}$, includes as primitives at least $+$, $-$, $\cdot$ and $<$. Given the primitive functions, one can construct an expanded class of functions by repeated composition of these basic functions. These are called terms. Basic formulas are $t_1 = t_2$, $t_1 < t_2$ for terms t_1, t_2. Recursively, from the basic formulas, one constructs formulas by Boolean operations ('and', 'or' and 'not') and by existential quantification over $\mathbb{R}$. I write $\phi(e_1, ..., e_l, a_1, ..., a_m)$ for a $\mathcal{L}$-formula with free variables $(e, a) \in \mathbb{R}^l \times \mathbb{R}^m$ and write $\mathcal{M} \models \phi(\bar{e}, \bar{a})$ to express that $\phi(\bar{e}, \bar{a})$ is true in $\mathcal{M}$ (one might be tempted to simply write $\phi(\bar{e}, \bar{a})$ to indicate that the statement is true - however, this notation allows to differentiate between the statement and saying that the statement is true in the structure $\mathcal{M}$).

It will be assumed throughout of the paper[2] that the economic statement of interest can be expressed as an $\mathcal{L}$-formula which I write as $\phi(e, a)$, where $e \in \mathbb{R}^l$ can be interpreted as the exogenous variables and $a \in \mathbb{R}^m$ describes quantitative features of the equilibrium; $\mathcal{M} \models \phi(\bar{e}, \bar{a})$ will simply mean that at $\bar{e}$ the property $\phi(\bar{e}, \bar{a})$ holds.

A set $\Phi \in \mathbb{R}^l$ is called definable if there is an $\mathcal{L}$ formula $\phi(x)$ such that

$$\Phi = \{\bar{x} : \mathcal{M} \models \phi(\bar{x})\}.$$

Throughout this paper, definable means *definable with parameters*, i.e. I will consider sets which are definable using constants from $\mathbb{R}$ (see van den Dries (1998, page 22)). A function is called definable if its graph is a definable set.

[2] In Section 3, I give concrete examples of economies and computational experiments which fit this framework.

2.1.1 Computational experiments

Ideally computational experiments take as given $\bar{e} \in \mathbb{R}^l$ and find some $\bar{a} \in \mathbb{R}^m$ such that

$$\mathcal{M} \models \phi(\bar{e}, \bar{a}).$$

Most existing algorithm used in practice only find one of possibly several solutions and often the algorithm only finds an approximate solution. The fact that they only find one solution limits the economic statements one can consider (for example, one can generally not make statement about all solutions), but is irrelevant for the approximate generalization suggested in this paper – I will give a concrete description of computational experiments in general equilibrium analysis in Section 3. The fact that only approximate solutions can be obtained often means that the computational experiment can only determine an $\epsilon > 0$ such that there exists a $\tilde{e}$ with $\|\tilde{e} - \bar{e}\| \leq \epsilon$ and $\mathcal{M} \models \phi(\tilde{e}, \bar{a})$. I show below that for sufficiently small ϵ this is sufficient for approximate generalizations.

Even if the experiment can be conducted at 'low' computational cost, it can only be be repeated for finitely many different values of e. While one can often assume that it is known that all interesting values of e lie in some compact set $E \subset \mathbb{R}^l$, it is unfortunately computationally infeasible (or, depending on the underlying structure even impossible) to find some $\bar{a}$ such that for all $\bar{e} \in E$ the statement is true. i.e. such that

$$\mathcal{M} \models \forall \bar{e}[\phi(\bar{e}, \bar{a}) \vee \neg[\bar{e} \in E]].$$

As explained in the introduction I instead want to derive a statement about how 'likely' it is that there exists $\tilde{e}$ for which the statement cannot possibly be true after one has verified that $\mathcal{M} \models \phi(\bar{e}, \bar{a})$ for sufficiently many appropriately chosen $\bar{e} \in E$. I assume that E is a compact, Lebesgue-measurable and definable subset of $\mathbb{R}^l$, $E = \{x : \mathcal{M} \models \eta(e)\}$, for some $\mathcal{L}$ formula η, and denote the l-dimensional Lebesgue measure by ν. This leads to the definition of an approximately correct statement

DEFINITION 1 *Assume that for $\bar{a} \in \mathbb{R}^m$, $\{\bar{e} : \mathcal{M} \models \phi(\bar{e}, \bar{a}) \wedge \eta(\bar{e})\}$ is ν-measurable. Then the statement*

$$\forall \bar{e}[\phi(\bar{e}, \bar{a}) \vee \neg\eta(\bar{e})]$$

is said to be δ-approximately correct if

$$\frac{\nu(\{\bar{e} : \mathcal{M} \models \phi(\bar{e}, \bar{a}) \wedge \eta(\bar{e})\})}{\nu(\{\bar{e} : \mathcal{M} \models \eta(\bar{e})\})} > 1 - \delta.$$

While in general one cannot rule out that there exist some $e \in E$ for which the statement is false it might often be useful to find bounds on the size of sets of variables for which the conjecture might be wrong. I call this set the potentially exceptional set and show in Corollary 1 below that a bound on the number of connected components of the intersection of Φ and axes-parallel cylinders can be used to make a statement on δ, i.e. bound the size

of the potentially exceptional set. The method will say nothing about where this set might be located, in particular the exceptional set might be empty. This statement is, of course, much weaker than showing that the statement is true for all elements of E but the point is that in many applications this is just not possible.

The philosophy is somewhat related with the idea underlying genericity analysis for smooth economies. There one is concerned with showing that the exceptional set has measure zero. It might very well be possible that *all* of the economically relevant specifications fall into the residual set of measure zero, but it is simply not true (or cannot be shown) that the residual set is empty. Of course, there is a huge quantitative difference between showing that the residual set has small positive measure and showing that it has zero measure - in this respect generic results are much stronger than the ones aimed for in this paper.

2.1.2 O-minimality

In this paper, I argue that virtually all economies that are examined in applied work are definable in an o-minimal expansion of the real field $\mathbb{R}$. A structure $\mathcal{M}$ is said to be o-minimal if for any definable set $X \subset \mathbb{R}$ there are finitely many intervals $I_1, ..., I_n$ with endpoints in $\mathbb{R} \cup \{-\infty, +\infty\}$ and a finite set X_0 such that

$$X = X_0 \cup I_1 \cup ... \cup I_n.$$

It is beyond the scope of this paper to discuss the assumption of o-minimality in detail. It will be made clear below that this assumption is very useful for the analysis. For an thorough reference on o-minimal structures see van den Dries (1998). A well known example of an o-minimal structure is the ordered field of real numbers, terms in the underlying language are simply polynomials and definable sets are the semi-algebraic sets. It will become clear below that for some economic applications this structure is actually rich enough. However, in many cases one wants to include exponentiation in the underlying first order language, this does not destroy o-minimality (see Wilkie (1994)). In Section 2.1.3 I give an example of an o-minimal structure which is particularly interesting for economic applications.

The following 2 theorems are important for our analysis. The first theorem is a standard result for o-minimal structures (see e.g. van den Dries (1998)).

THEOREM 1 *Let $\Phi \subset \mathbb{R}^l$ be a definable set in an o-minimal expansion of $\mathbb{R}$. There is a uniform bound B such that for any affine set $L \subset \mathbb{R}^l$, the set $\Phi \cap L$ has at most B connected components.*

A set is affine if it can be defined by a system of linear equations. Berarducci and Servi (2004) show that there are algorithms to compute upper bounds on B from the defining formula ϕ. In practice, however, this result is rather useless for obtaining actual bounds. In Section 2.3 below, I give specific examples in which bounds can be easily derived from the defining formulas.

The second result that is important for the analysis of this paper follows from the cell-decomposition theorem (see van den Dries (1998) for a statement and proof of the cell-decomposition theorem).

THEOREM 2 *If $\mathcal{M}$ is an o-minimal expansion of $\mathbb{R}$, all definable sets are Lebesgue-measurable.*

2.1.3 O-minimal economic models

Given an o-minimal structure, preferences over consumption bundles in some definable set X, $\succeq$, are called definable if all better sets are definable, i.e. for all $x \in X$, $\{y : y \succeq x\}$ is definable in $\mathcal{M}$. Richter and Wong (2000) prove that definable preferences can be represented by definable utility functions. It is easy to see that definable utility functions give rise to definable best response correspondences and that in pure exchange economies the equilibrium manifold is definable if preferences are definable. Blume and Zame (1992) apply o-minimality to consumer theory and general equilibrium analysis and prove a definable analogue of Debreu's theorem on generic local uniqueness.

Both Blume and Zame and Richter and Wong argue that the assumption that preferences and technologies are definable in an o-minimal structure is very natural and satisfied in almost all (finite) applied general equilibrium models. I want to apply Theorems 1-2 to examine sets that are defined by quantitative statements which can be expressed as first order formulas in an o-minimal structure.

To give concrete examples of utility functions which are frequently used in applications and which can be defined in an o-minimal structure it is useful to define Pfaffian functions. The definition is from Khovanskii (1991) who shows that these functions maintain many of the finiteness properties of polynomials.

DEFINITION 2 *A Pfaffian chain of order $r \geq 0$ and degree $\alpha \geq 1$ in an open domain $G \subset \mathbb{R}^n$ is a sequence of analytic functions $f_1, ..., f_r$ in G satisfying differential equations*

$$df_j(x) = \sum_{1 \leq i \leq n} g_{ij}(x, f_1(x), ..., f_j(x))dx_i$$

for $1 \leq j \leq r$. The g_{ij} are polynomial in $x = (x_1, ..., x_n)$, $y_1, ..., y_j$ of degree not exceeding α. A function $f(x) = p(x, f_1(x), ..., f_r(x))$, with p being a polynomial of degree β is called a Pfaffian function of order r and degree (α, β).

The following simple facts about Pfaffian functions are easy to verify.

- $\exp(x)$ is a Pfaffian function of order 1 and degree $(1, 1)$ in $\mathbb{R}$; $f(x) = \log(x)$ is a Pfaffian function of order 2 and degree (2,1) on $\mathbb{R}_{++}$ since $f'(x) = 1/x$ and $f''(x) = -(f'(x))^2$.
- Given two Pfaffian functions of order r with the same underlying chain and degrees (α_1, β_2) and (α_2, β_2) respectively the sum is a Pfaffian function of order r and degree $(\max(\alpha_1, \alpha_2), \max(\beta_1, \beta_2))$.

The product of the two functions is Pfaffian of order r and degree $(\max(\alpha_1, \alpha_2), \beta_1 + \beta_2)$.

- A partial derivative of a Pfaffian function of order r and degree (α, β) is a Pfaffian function with the same Pfaffian chain of order r and degree $(\alpha, \alpha + \beta - 1)$

This shows that all commonly used utility function (e.g. CES, LRT) as well as first order conditions for agents' optimality can be written as Pfaffian functions. Wilkie (1996) proved that the structure generated by Pfaffian functions is o-minimal. Hence almost all finite economies considered in applied work are definable in an o-minimal structure.

2.2 The main result

I now show how bounds on the number of connected components can be used to obtain approximately correct conjectures. The main theorem in this section is a generalization of Koiran (1995). The proofs are fairly straightforward generalizations of Koiran's proofs but they are included for completeness.

For simplicity, it is assumed that the set of unknown parameters, E, is the unit cube in $\mathbb{R}^l$, i.e. $E = [0,1]^l$ (as long as the true set of interest can be mapped into a cube by a continuous, measure-preserving and o-minimal function, the argument can be modified relatively easily - this will become clearer in Section 3). For a positive integer N define F to be the set of evenly spaced grid-points with distance $1/N$, i.e. $F = \{1/N, 2/N, ..., 1\}^l$. Throughout, fix $\|.\|$ to denote the 2-norm, it is trivial to alter the argument for other norms but the 2-norm will prove useful in deriving actual bounds below.

Suppose that the computational experiment finds an $\bar{a}$ and ϵ, $1/N > \epsilon \geq 0$, such that for all $e \in F$ there is $\tilde{e}$ with $\|\tilde{e} - e\| \leq \epsilon$ and with

$$\mathcal{M} \models \phi(\bar{e}, \bar{a}).$$

Given $\bar{a}$, define the set

$$\Phi = \{\bar{e} : \mathcal{M} \models \phi(\bar{e}, \bar{a})\} \cap [0,1]^l$$

(to simplify notation I drop the subscript a and define Φ to be a subset of $[0,1]^l$ right away). The goal is to give good lower bounds on the size (Lebesgue measure) of Φ. Define a generalized indicator function $\Im^\epsilon(x)$ to be 1 if here is a $y \in \Phi$ with $\|y - x\| \leq \epsilon$ and zero otherwise. For $\epsilon = 0$ this is the simple indicator function and the Lebesgue measure of Φ is given by $\int_{[0,1]^l} \Im^0(x)dx$.

For $x \in F$, define a cylinder of radius ϵ centered around $(x_1, ..., x_{\bar{i}-1}, x_{\bar{i}+1}, ..., x_l)$ by

$$C^\epsilon_{-\bar{i}}(x) = \{y \in \mathbb{R}^l : \|y_i - x_i\| \leq \epsilon \text{ for } i \neq \bar{i}\}.$$

For a set A denote by $\kappa(A)$ the number of connected components. The following lemma generalizes Lemma 2 in Koiran (1995).

LEMMA 1 *Given $\bar{x} \in F$, define $Q = C^{\epsilon}_{-1}(\bar{x}) \cap \Phi$. Denote by $\kappa(Q)$ the number of its connected components. Then*

$$\left| \int_0^1 \Im^{\epsilon}(y, \bar{x}_2, ..., \bar{x}_l) dy - \frac{1}{N} \sum_{i=1}^{N} \Im^{\epsilon}(i/N, \bar{x}_2, ..., \bar{x}_l) \right| \leq \kappa(Q)/N.$$

Proof. The set of $x \in \mathcal{C}^{\epsilon}_{-1}$ for which $\Im^{\epsilon}(x) = 1$ can be written as the union of K disjoint connected pieces with $K \leq \kappa$, i.e. there exist $a_1 < b_1 < ... < a_K < b_K$ such that

$$\{x \in C^{\epsilon}_{-1} : \Im^{\epsilon}(x) = 1\} = \cup_{k=1}^{K} \{x \in C^{\epsilon}_{-1}, x_1 \in [a_k, b_k]\}.$$

Then

$$\left| \frac{1}{N} \sum_{i=1}^{N} \Im^{\epsilon}(i/N, \bar{x}_2, ..., \bar{x}_l) - \int_0^1 \Im^{\epsilon}(y, \bar{x}_2, ..., \bar{x}_l) dy \right| \leq$$
$$\sum_{k=1}^{K} \left| \int_{a_k}^{b_k} \Im^{\epsilon}(y, \bar{x}_2, ..., \bar{x}_l) dy - \frac{1}{N} \sum_{i: a_k \leq i/N \leq b_k} \Im^{\epsilon}(i/N, \bar{x}_2, ..., \bar{x}_l) \right|$$

The definition of $\Im^{\epsilon}$ implies that for all k

$$\left| \int_{a_k}^{b_k} \Im^{\epsilon}(y, \bar{x}_2, ..., \bar{x}_l) dy - \frac{1}{N} \sum_{i: a_k \leq i/N \leq b_K} \Im^{\epsilon}(i/N, \bar{x}_2, ..., \bar{x}_l) \right| \leq 1/N.$$

The result follows directly from this. □

Define λ to be the maximal number of connected components across all intersections of Φ with all possible cylinders C^{ϵ}, i.e.

$$\lambda = \max_{i=1,...,l; x \in F} \kappa(C^{\epsilon}_{-i}(x) \cap \Phi).$$

The following theorem is the main tool for the analysis in this paper.

THEOREM 3 *Given a bound on connected components λ, one can estimate the size of Φ from verifying ϕ on a grid as follows.*

$$\left| \frac{1}{N^l} \sum_{i_1,...,i_l}^{N} \Im^{\epsilon}(i_1/N, ..., i_l/N) - \int_{[0,1]^l} \Im^0(x) dx \right| \leq \frac{l + 2\epsilon}{N} \lambda \tag{1}$$

Proof.

The theorem is proved by induction. For $l = 1$, one only needs to modify the last step of the proof of Lemma 1 and obtains

$$\left| \frac{1}{N} \sum_{i=1}^{N} \Im^{\epsilon}(i/N) - \int_{[0,1]} \Im^0(x) dx \right| \leq \kappa(\Phi)(1 + 2\epsilon)/N$$

For $l > 1$, the induction goes as follows follows.

$$\left|\frac{1}{N^l}\sum_{i_1,\ldots,i_l}^{N}\Im^{\epsilon}(i_1/N,\ldots,i_l/N)-\int_{[0,1]^l}\Im^{0}(x)dx\right| \leq$$
$$\left|\int_{[0,1]}\frac{1}{N^{l-1}}\sum_{i_2,\ldots,i_l}^{N}\Im^{\epsilon}(x_1,i_2/N,\ldots,i_l/N)dx_1-\int_{[0,1]^l}\Im^{0}(x)dx\right| +$$
$$\left|\frac{1}{N^l}\sum_{i_1,\ldots,i_l}^{N}\Im^{\epsilon}(i_1/N,\ldots,i_l/N)-\int_{[0,1]}\frac{1}{N^{l-1}}\sum_{i_2,\ldots,i_l}^{N}\Im^{\epsilon}(x_1,i_2/N,\ldots,i_l/N)dx_1\right|.$$

Assuming that (1) holds for $l-1$, one obtains that for all $x_1 \in [0,1]$,

$$\left|\frac{1}{N^{l-1}}\sum_{i_2,\ldots,i_l}^{N}\Im^{\epsilon}(x_1,i_2/N,\ldots,i_l/N)-\int_{[0,1]^{l-1}}\Im^{0}(x_1,\tilde{x})d\tilde{x}\right| \leq \lambda\frac{l-1+2\epsilon}{N}.$$

By Lemma 1,

$$\left|\frac{1}{N^l}\sum_{i_1,\ldots,i_l}^{N}\Im^{\epsilon}(i_1/N,\ldots,i_l/N)-\int_{[0,1]}\frac{1}{N^{l-1}}\sum_{i_2,\ldots,i_l}^{N}\Im^{\epsilon}(x_1,i_2/N,\ldots,i_l/N)dx_1\right| \leq \frac{\lambda}{N}.$$

The result then follows by integrating the first term over $[0,1]$ and adding the result to the second expression. □

Koiran (1995) considers the (important) special case $\epsilon = 0$. With bounds on the number of connected components of the intersection of Φ with axes parallel lines this provides a method for bounding the measure of Φ. In practice, these bounds are often orders of magnitude better than bounds on connected components of the intersection with general cylinders C^{ϵ}.

It is unclear, if the bounds in the theorem are tight and if the choice of the grid-points is optimal. In particular, the question whether one can find points in higher dimension which do not require the number of points to grow at the exponential rate of Equation 1 is subject to further research.

The theorem immediately implies the following corollary which is the central theoretical result of the paper.

COROLLARY 1 *Given a class of o-minimal economies parameterized by $E = [0,1]^l$ and a sentence $\phi(e)$, for any $\delta > 0$ one can compute a finite N such that if $\phi(\bar{e})$ holds for all points $\bar{e} \in \{1/N,\ldots,1\}^N$, then ϕ is δ-approximately correct according to Definition 1.*

The methods introduced in this section are obviously not the only ones, one can use to show that a given formula holds for a rich class of parameters. Since the real closed

field is decidable, one can apply algorithmic quantifier elimination and use an algorithm to verify if a given semi-algebraic statement of interest is true *for all* parameters in a given (semi-algebraic) set (see e.g. Basu et al (2003)). However, for more complicated structures decidability is an open problem and there are certainly no algorithm available for quantifier elimination at this time. Moreover, even for the semi-algebraic case, the methods in this paper seem to be much more tractable than quantifier elimination.

In applying these method, the 'only' challenge is to find reasonable bounds on λ. For general formulas good upper bounds are fairly difficult to obtain, however, computational experiments usually consider very specific formulas, for which it is easy to obtain bounds.

2.3 Bounds on connected components of solutions to equations

In order to bound the number of connected components of relevant sets arising in computational experiments, I will use two results that bound the number of connected components of solutions to definable systems of equations. Given definable functions $f : G \subset \mathbb{R}^l \to \mathbb{R}^n$, what is the maximal number of connected components of

$$S = \{x : f(x) = 0\} \subset \mathbb{R}^l.$$

I discuss two cases, in the first case, f consists only of polynomial functions. The second case considers semi-Pfaffian sets, i.e. $f_1, ..., f_n$ are Pfaffian functions. The following bounds are, to the best of my knowledge, the best currently known. However, it is clear that these are upper bounds which are often not obtained in economic applications. In particular, there is a large gap between known upper and lower bounds in the Pfaffian case.

Semi-algebraic sets

The following bound is from Rojas (2000). Suppose $f_1, ..., f_n$ are polynomial and $G = \mathbb{R}^l$. Consider the convex hull of the union of the l unit vectors in $\mathbb{R}^l$ together with the origin and the exponents of all monomials in the equalities which define S (i.e. for the monomial $x_1^{\alpha_1}...x_l^{\alpha_l}$ one would take the vector $\alpha \in \mathbb{R}^l$). For a set $Q \subset \mathbb{R}^l$ denote by $\text{vol}(Q)$ the l-dimensional volume which is standardized to obtain volume 1 for the l-dimensional simplex.

Then the number of connected components of S, $\kappa(S)$ can be bounded as follows.

$$\kappa(S) \leq 2^{l-1}\text{vol}(Q) \tag{2}$$

Semi-Pfaffian sets

The following result is from Gabrielov and Vorobjov (2004). Suppose all f_i, $1 \leq i \leq n$ are Pfaffian functions on a domain $G \subset \mathbb{R}^l$, with either $G = \mathbb{R}^l$ or $G = \mathbb{R}^l_{++}$, having common Pfaffian chain of order r and degrees (α, β_i) respectively. Let $\beta = \max_i \beta_i$. Then the number of connected components of $\{x : f_1(x) = ... = f_n(x) = 0\}$ does not exceed

$$2^{\frac{r(r-1)}{2}+1}\beta(\alpha + 2\beta - 1)^{l-1}((2l-1)(\alpha+\beta) - 2l + 2)^r \tag{3}$$

3 Computational experiments in general equilibrium

I now give an abstract description of computational experiments in general equilibrium analysis, provide two concrete examples to illustrate the analysis and show how the mathematical results from the previous section can be used to derive general conjectures from these experiments.

Throughout this section, I fix an o-minimal structure $\mathcal{M}$ and assume that the economic statement of interest, for a given specification of exogenous variables, e, can be written as follows.

$$\phi(e,a) \leftrightarrow \exists (x_1,...,x_k) \in \mathbb{R}^n : \begin{array}{rcl} h_1(x_1,e) & = & 0 \\ & \vdots & \\ h_k(x_k,e) & = & 0 \\ \psi(x_1,...,x_k,e,a) & > & 0 \end{array} \tag{4}$$

where $h_1,...,h_k$ and ψ are definable functions. For each $i = 1,...,k$ the (possibly multivariate) function h_i should be understood to summarize the equilibrium conditions for a given economy, i.e. they consist of necessary and sufficient first order conditions together with market clearing or in the simplest case simply of the aggregate excess demand functions. The vector x_i is supposed to contain all endogenous variables (e.g. allocations and prices) for this economy. Different h_i correspond to different specifications of the economy, for example h_1 could summarize the equilibrium conditions for an economy without taxes while in h_2 some taxes are introduced to the economy. The function ψ makes the comparative statics comparisons which are of interest in the particular application.

Before explaining how bounds on the number of connected components can be computed, I give, as an illustration, two concrete examples and show how they fit into the general framework of Equation (4).

Example 1

The first example considers a simple qualitative comparative statics exercise. A motivation and a theoretical examination can be found in Geanakoplos and Kubler (2003). Consider a model of an exchange economy with two periods and a single perishable good. There are S possible states $s = 1,...,S$ in the second period, and the first period state is denoted by $s = 0$. There are two agents, A and B, with initial endowments $e^h \equiv (e_0^h, \tilde{e}^h) \in \mathbb{R}^{S+1}_{++}$ and time-separable, von Neumann-Morgenstern utility functions

$$U^h(x) = u^h(x_0) + \sum_{s=1}^{S} \pi_s u^h(x_s)$$

with strictly increasing, strictly concave and C^2 functions u^h for $h = A, B$. I assume the state probabilities π_s are identical across agents. Suppose that agents can only trade via a single safe asset paying 1 unit of the good in states $s = 1,...,S$. Denote agent h's holding of the asset by $\theta^h \in \mathbb{R}$, and the price of the asset by q.

It is well known that even though the incomplete market equilibrium is generally not Pareto-optimal, the Arrow-Debreu equilibrium might not Pareto-dominate it. Some agents might lose through the price effect of financial innovations. In fact, the price effect will always hurt some agent in that he is not able to afford his incomplete markets consumption at the new Arrow Debreu prices. It is then an interesting question to examine under which conditions the lender is always *observably better off* in the complete markets equilibrium, i.e. at the new prices he could have afforded his old consumption bundle.

In order to tackle this question, one can use a computational experiment. One has to fix the number of states to some S and assume that utility functions belong to some parametric class of Pfaffian functions. An example is the class of utility functions exhibiting linear risk tolerance, i.e. $u'(x)/u''(x)$ is a linear function in x. The exogenous variables are individual endowments e_s^h as well as parameters of the utility function and probabilities π_s, $h = A, B$, $s = 0, ..., S$.

Assuming that first order conditions are necessary and sufficient, the statement ϕ then is as follows.
$\exists(q, \theta, c_0, c_1, c_2)$:

$$\begin{aligned}
-qu'_A(e_0^A - q\theta) + \sum_{s=1}^{S} \pi_s u'_A(e_s^A + \theta) &= 0 \\
-qu'_B(e_0^B + q\theta) + \sum_{s=1}^{S} \pi_s u'_B(e_s^B - \theta) &= 0 \\
\frac{u'_A(c_1)}{u'_A(c_0)} - \frac{u'_B(e_1 - c_1)}{u'_B(e_0 - c_0)} &= 0 \\
\frac{u'_A(c_2)}{u'_A(c_0)} - \frac{u'_B(e_2 - c_2)}{u'_B(e_0 - c_0)} &= 0 \\
u'_A(c_0)(c_0 - e_0^A) + \sum_{s=1}^{S} \pi_s u'_A(c_s)(c_s - e_s^A) &= 0 \\
\theta \left(u'_A(c_0)(c_0 - (e_0^A - q\theta)) + \sum_{s=1}^{S} \pi_s u'_A(c_s)(c_s - (e_s^A + \theta)) \right) &> 0
\end{aligned}$$

Note that the last inequality describes the qualitative comparative statics statement because if A is the borrower, he must be observably worse off for the statement to be true. Since all functions are Pfaffian, there exist bounds on the number of connected components and one can bound the size of the set of parameters for which the lender observably gains.

Example 2

The second example re-examines a famous result in applied general equilibrium analysis. Following Shoven (1976), I ask about the welfare effects of capital taxation in the two-sector model of the US economy. While Shoven provided results only for fixed preferences and

production functions, one can use the methods of this paper to assess how robust these results are if one considers a whole class of functions.

One important aspect of Shoven's analysis is that unknown exogenous parameters are 'calibrated' so that in the equilibrium with taxes, the model matches key aspects of aggregate data. In this case, it is not straightforward to map the set of exogenous parameters E into the unit cube. I explain how some of the exogenous variables can be chosen to ensure that the benchmark model matches observed quantities. The statement of interest then quantifies over these, given values for the 'free' parameters.

In a static economy, 2 consumption goods are produced by 2 sectors, $j = 1, 2$ using as input capital and labor. Production functions are CES and therefore specified up to unknown parameters $\xi = (\beta_1, \gamma_1, \sigma_1, \beta_2, \gamma_2, \sigma_2)$,

$$f_j(y_l, y_k; \xi) = \gamma_j(\beta_j y_l^{\sigma_j} + (1 - \beta_j) y_k^{\sigma_j})^{1/\sigma_j}.$$

Two individuals, $i = 1, 2$, are endowed with capital and labor (k_i, l_i). Prices are (p_1, p_2, p_k, p_l), throughout I normalize $p_l = 1$. Utilities are CES, hence individual demand functions are specified up to unknown parameters $\zeta = (\alpha_1, \nu_1, \alpha_2, \nu_2)$ and endowments,

$$x_i(p; l_i, k_i, \zeta) = \begin{pmatrix} \frac{p_k k_i + l_i}{\alpha_i p_1^{1-\nu_i} + (1-\alpha_i) p_2^{1-\nu_i}} \frac{\alpha_i}{p_1^{\nu_i}} \\ \frac{p_k k_i + l_i}{\alpha_i p_1^{1-\nu_i} + (1-\alpha_i) p_2^{1-\nu_i}} \frac{1-\alpha_i}{p_2^{\nu_i}} \end{pmatrix}.$$

In the benchmark equilibrium there is a tax on capital for one sector, τ, the revenue is distributed among the two agents in equal shares. Since only nominal quantities are known, Shoven sets all prices to be equal to one in the benchmark equilibrium (this is without loss of generality).

Given individual endowments in capital and labor as well as exogenous parameters in demand- and production functions, k, l, ξ, ζ, which ensure that the equilibrium with the tax matches aggregate quantities (i.e. total labor input, capital input and total output per sector), the computational experiment then measures the effects on GDP (measured at old prices) of abolishing the tax on capital in sector 1. Of course, for many values of exogenous variables, the benchmark equilibrium will not match observed quantities. In order to examine for which range of exogenous parameters the statement is true one therefore needs to decide which parameters one can choose freely and which are pinned down by the benchmark equilibrium. For example it is reasonable to pin down endowments of the second agent by aggregate endowments of capital and labor. One could then use aggregate output to pin down γ_1 and γ_2 in the production function and use aggregate consumption to pin down α_i for $i = 1, 2$. The free variables are then l_1, k_1, σ_j, β_j, $j = 1, 2$ and ν_1, ν_2. The crucial inside is to quantify not only over the endogenous variables but also over those exogenous variables that are determined by the benchmark equilibrium.

Given values for the free exogenous variables a classical computational experiment thus 'measures' a such that $\exists (p_1, p_2, p_k), (\bar{y}_{jk}, \bar{y}_{jl})_{j=1,2}, (y_{jk}, y_{jl})_{j=1,2}$ as well as $l_2, k_2, \gamma_1, \gamma_2, \alpha_1, \alpha_2$

with

$$\begin{aligned}
x^1(\mathbf{1}, l^1, k^1, \zeta) + x^2(\mathbf{1}, l^2, k^2, \zeta) &= X \\
f_j(\bar{y}_{jk}, \bar{y}_{jl}; \xi) &= X_j, \ j = 1, 2 \\
\frac{\partial f_j(\bar{y}_j; \xi)}{\partial l} &= 1, \quad j = 1, 2 \\
\frac{\partial f_1(\bar{y}_1; \xi)}{\partial k} &= (1 - \tau) \\
\frac{\partial f_2(\bar{y}_2; \xi)}{\partial k} &= 1 \\
\bar{y}_{1k} + \bar{y}_{2k} &= K \\
\bar{y}_{1l} + \bar{y}_{2l} &= L \\
k^1 + k^2 &= K \\
l^1 + l^2 &= L
\end{aligned}$$

as well as

$$\begin{aligned}
x^1_j(p, l^1, k^1, \zeta) + x^2_j(p, l^2, k^2, \zeta) &= f_j(y_{jk}, y_{jl}; \xi), \ j = 1, 2 \\
\frac{\partial f_j(y_j; \xi)}{\partial l} &= p_l, \quad j = 1, 2 \\
\frac{\partial f_j(y_j; \xi)}{\partial k} &= p_k, \quad j = 1, 2 \\
y_{1k} + y_{2k} &= K \\
y_{1l} + y_{2l} &= L \\
f_1(y_{1k}, y_{1l}; \xi) + f_2(y_{2k}, y_{2l}; \xi) &> a(X_1 + X_2)
\end{aligned}$$

After suitable rescaling this experiment can be conducted on a grid of free exogenous variables. Since (again) all functions are Pfaffian, the number of connected components can be bounded and one can make statements about the relative size of the set of preference and technology parameters for which the abolishment of taxes leads to a GDP growth of at least a percent.

3.1 Bounding the number of connected components

Suppose that equilibria can be computed with error[3] $\epsilon \geq 0$ and that for some large N the computational experiment finds $\bar{a}$ such that for all $e \in \{1/N, ..., 1\}^l$, there is a $\bar{e} \in [0, 1]^l$ with $\|\bar{e} - e\| \leq \epsilon$ and with

$$\mathcal{M} \models \phi(\bar{e}, \bar{a}).$$

In order to bound from below the size (Lebesgue measure) of the set $\{e \in E : \mathcal{M} \models \phi(e, \bar{a})\}$ and to apply the results from Section 2, one needs to bound the number of connected

[3]In Appendix A, I provide a justification for the special case $\epsilon = 0$. Considering this case separately is worthwhile because the available bounds on the number of connected components are much sharper.

components of $\Phi \cap C^\epsilon$ where C^ϵ is a generic axis-parallel cylinder in $\mathbb{R}^l$. I.e. for a given $\bar{i}$, consider the system of inequalities (4) with the additional equation

$$\|(e_1, ...e_{\bar{i}-1}, e_{\bar{i}+1}, ..., e_l) - (\bar{e}_1, ..., \bar{e}_{\bar{i}-1}, \bar{e}_{\bar{i}+1}, ..., \bar{e}_l)\| \leq \epsilon \tag{5}$$

The following simple result on the number of connected components will be key for the analysis to follow.

THEOREM 4 *Let $x = (x^1, x^2)$. Given a definable set $\Phi = \{x \in \mathbb{R}^l : \mathcal{M} \models \phi(x)\}$ with $\kappa(\Phi)$ connected components, the projection*

$$P_{x^1}(\Phi) = \{x^1 : \mathcal{M} \models \exists x^2 \phi(x^1, x^2)\}$$

is definable and has at most $\kappa(\Phi)$ connected components.

To prove Theorem 4 simply observe that if there are two points $y_1, y_2 \in P_{x^1}(\Phi)$ there must exist z_1, z_2 such that $(y_1, z_1), (y_2, z_2) \in \Phi$. If there is no continuous path between y_1 and y_2, since projection is continuous, there can obviously be none between (y_1, z_1) and (y_2, z_2).

The theorem implies that it suffices to consider the number of connected components of

$$\begin{aligned} Q^{\bar{i}} = \{(x_1, ..., x_k, e) \in \mathbb{R}^m \times \mathbb{R}^l : \quad & h_i(x_i, e) = 0, \quad i = 1, ..., k \\ & \psi(x, e, a) < 0 \\ & \textstyle\sum_{i \neq \bar{i}}^{l} (e_i - \bar{e}_i)^2 \leq \epsilon^2\} \end{aligned} \tag{6}$$

for some generic $\bar{e} \in F$.

Note that by examining the set in the space $\mathbb{R}^m \times E$, the dimension increases substantially and the bounds become large fairly fast.

3.1.1 Strict and weak inequalities

The bounds (2) and (3) from Section 2 apply to solutions to systems of equations, while (6) also consists of inequalities.

Weak inequalities can turned into strict inequalities by adding a slack variable. The number of connected components of $\{x \in \mathbb{R}^n : f(x) = 0 \wedge g(x) \leq 0\}$ is no larger then the number of connected components of $\{(x, y) \in \mathbb{R}^{n+1} : f(x) = 0 \wedge g(x) + y^2 = 0\}$.

Strict inequalities can be turned into inequalities in the following way. Given J inequalities $g_1(x) > 0$, ..., $g_J(x) > 0$ and a system of equations $f(x) = 0$ the number of connected components of

$$\{x : f(x) = 0 \wedge g_1(x) > 0 \wedge ... \wedge g_J(x) > 0\}$$

is bounded by the number of connected components of

$$\{x : f(x) = 0 \wedge g_1(x) \neq 0 \wedge ... \wedge g_J(x) \neq 0\}$$

which is bounded by the number of connected components of

$$\{(x,\gamma) : 1 - \gamma\Pi_{j=1}^{J} g_j(x) = 0\}.$$

It follows from Rojas (2000) that generally the number of connected components grows slower in the number of inequalities if these are strict inequalities rather than weak ones. This is the reason why I defined computational experiments to make statements about strict inequality.

If the function ψ is multivariate, i.e. $\psi(x,e,a) = (\psi_1(x,e,a),...,\psi_n(x,e,a))$, bounds on the number of connected components of $Q^{\bar{i}}$ might grow exponentially with the number of inequalities. However, defining N separate sets

$$\begin{aligned} Q_n^{\bar{i}} = \{(x_1,...,x_k,e) \in \mathbb{R}^m \times \mathbb{R}^l : \quad & h_i(x_i,e) = 0, \quad i = 1,...,k \\ & \psi_n(x,e,a) < 0 \\ & \textstyle\sum_{i\neq\bar{i}}^{l} (e_i - \bar{e}_i)^2 \leq \epsilon^2\} \end{aligned}$$

for $n = 1,...,N$, one can apply Equation (1) for each set separately and by additivity of the Lebesgue measure the total residual set cannot be larger than the sum of the residual sets for each n.

Note that for $\epsilon = 0$ the last inequality can only hold with equality and $e_i = \bar{e}_i$ for all $i \neq \bar{i}$. Therefore the number of unknowns and the dimension of the set whose number of connected components one is interested in shrinks by l (the slack variable for the last inequality as well as e_i, $i \neq \bar{i}$). Bounds often become order of magnitudes smaller.

3.2 How many examples are necessary ?

It is a well known problem that even for the semi-algebraic case, known upper bounds on the number of zeros or connected components are often astronomical and considering Pfaffian functions makes matters much worse. I illustrate the problem with example 1.

To simplify matters, assume that there are only two states $s = 1, 2$ and that they have known probabilities of 0.5 each. I also assume that both agents have log-utility. In this case one can explicitly solve for the Arrow-Debreu equilibrium and the first order conditions which characterize the incomplete market equilibrium can be written as polynomials. The following system then characterizes the computational experiment.

$$\begin{aligned} \exists q,\theta,c : 4q\theta^2 + 3q\theta e_1^A + 3q\theta e_2^A + 2qe_1^A e_2^A - e_0^A e_1^A - e_0^A e_2^A - 2e_0^A\theta &= 0 \\ 4q\theta^2 - 3q\theta e_1^A - 3q\theta e_2^B + 2qe_1^B e_2^B - e_0^B e_1^B - e_0^B e_2^B - 2e_0^B\theta &= 0 \\ c_0 e_0 e_1 e_2 - \left(\frac{1}{2} e_0 (e_0^A e_1 e_2 + 0.5(e_1^A e_0 e_2 + e_2^A e_0 e_1))\right) &= 0 \\ c_1 e_0 e_1 e_2 - \left(\frac{1}{2} e_1 (e_0^A e_1 e_2 + 0.5(e_1^A e_0 e_2 + e_2^A e_0 e_1))\right) &= 0 \\ c_2 e_0 e_1 e_2 - \left(\frac{1}{2} e_2 (e_0^A e_1 e_2 + 0.5(e_1^A e_0 e_2 + e_2^A e_0 e_1))\right) &= 0 \\ \theta\left(2c_0c_1c_2 - \left(c_1c_2(e_0^A - q\theta) + c_0c_2(e_1^A + \theta) + c_0c_1(e_2^A + \theta)\right)\right) &> 0 \end{aligned}$$

Substituting for c_0, c_1, c_2 and hence eliminating the last 3 equations, the inequality simplifies considerably and reads as follows.

$$\theta\left(e_0^A e_1 e_2 + 0.5(e_1^A e_0 e_2 + e_2^A e_0 e_1) - q\theta e_1 e_2 - 0.5(\theta e_0 e_2 + \theta e_0 e_1)\right) > 0.$$

Since the system is polynomial it is reasonable to assume that equilibria can be computed exactly (see the appendix for details). Given any $\bar{s} = 0, 1, 2$ and $\bar{h} = 1, 2$, one therefore simply needs to bound the number of connected components of the set

$$\begin{array}{lcl}
\{(q, \theta, e_{\bar{s}}^{\bar{h}}): & & \\
4q\theta^2 + 3q\theta e_1^A + 3q\theta e_2^A + 2qe_1^A e_2^A - e_0^A e_1^A - e_0^A e_2^A - 2e_0^A\theta & = & 0 \\
4q\theta^2 - 3q\theta e_1^A - 3q\theta e_2^B + 2qe_1^B e_2^B - e_0^B e_1^B - e_0^B e_2^B - 2e_0^B\theta & = & 0 \\
1 - \theta\left(e_0^A e_1 e_2 + 0.5(e_1^A e_0 e_2 + e_2^A e_0 e_1) - q\theta e_1 e_2 - 0.5(\theta e_0 e_2 + \theta e_0 e_1)\right) & = & 0\}
\end{array}.$$

Using the bounds from Equation (2), I obtain that the number of connected components is bounded by $2^2 \times 7 = 28$. Although this bound is fairly small, Equation (1) implies that in order to bound the size of the residual set of individual endowments in $\mathbb{R}_+^6$ for which the lender does not observably gain by 0.1 one needs to perform the experiment on a grid with $10 \times 6 \times 28$ points in each direction. Even for this trivial case, one therefore has to compute a fairly large number of examples. The main problem is that the number of points one has to consider goes up exponentially with the dimension of unknown parameters in E. This 'curse of dimensionality' can be partly overcome by using random methods.

3.3 A random algorithm

Suppose one has access to a random number generator and can draw uniformly and independently random $\tilde{e} \in \{1/N, ..., 1\}^l$. Despite the fact that in practice one can often only use pseudo-random numbers this is a reasonable assumption (see e.g. L'Ecuyer (2004) for a discussion on generating random and quasi-random numbers). The random numbers are naturally integer valued (see Blum et al. (1998) for a more elaborate discussion on *probabilistic machines*) and therefore lie on a grid.

For a given $\bar{a}$, if after M draws, $\tilde{e}^1, ..., \tilde{e}^M$ it turns out that for each i there is an $\bar{e}$ with $\|\bar{e} - \tilde{e}^i\| \leq \epsilon$ and with

$$\mathcal{M} \models \phi(\bar{e}, \bar{a}),$$

by the Binomial formula one obtains

$$\text{Prob}\left[\frac{1}{N^l}\sum_{i_1,...,i_l} \Im^\epsilon(i_1/N, ..., i_l/N) < 1 - \delta\right] \leq (1-\delta)^M.$$

Using Equation (1) one can now infer probabilistic statements about the size of Φ from probabilistic statements about the number of points in the finite grid for which the statement

is true. If $\epsilon < 1/N$, with a bound on the maximal number of connected components of $\Phi \cap C_i^\epsilon$, λ, one obtains

$$\text{Prob}\left[\int_{[0,1]^l} \Im^0(x)dx < 1 - \delta - \frac{l+2\epsilon}{N}\lambda\right] \leq (1-\delta)^M \tag{7}$$

See Koiran (1995) or Blum et al. (1998, Chapter 17.4) for the case $\epsilon = 0$ and a discussion of the result.

Note that the number of connected components is not relevant if N is sufficiently large. However, in practice a bound for N arises naturally from the precision with which equilibria can be computed, i.e. with ϵ. Therefore these methods are applicable if and only if the number of connected components is several orders of magnitude smaller than $1/\epsilon$.

While these random methods are much more efficient (the number of points needed is independent of the dimension which only enters through bounds on the number of connected components), it is not clear how to interpret a statement like '29 is a prime number with probability 0.99'. While it is well known in theoretical computer science that random algorithms often reduce the complexity of the problem considerably, these algorithms usually solve a specific problem and it can often be checked, that the candidate solution produced by the algorithm is an actual solution (without probabilities attached). One possible interpretation of Equation (7) is the following. Suppose *nature* draws randomly a vector of parameters e uniformly from $[0,1]^l$. Equation (7) implies that the overall probability that this parameter will lie in Φ is at least $(1-\delta-\frac{l+2\epsilon}{N}\lambda)(1-(1-\delta)^N)$. This therefore allows for statistical statements about how likely it is that the conjecture is true for randomly selected parameters.

4 Conclusion

Computational experiments that make statements about one specific example economy can be generalized to infinite classes of economies when the economic fundamentals are definable in an o-minimal structure. Corollary 1, the main theoretical result, makes precise under which conditions finitely many examples suffice to make statements about sets of parameters with positive Lebesgue measure.

I argue that this theoretical insight can be fruitfully put to work in applied general equilibrium analysis. For all commonly used specifications of utility and production functions one can easily compute how many examples are needed to make statements about *large* sets of parameters. These statements are possible, even if equilibria cannot be computed exactly.

However, it turns out that in large problems the number of examples needed is astronomically high and it is therefore not feasible to make general statements using a deterministic algorithm. Fixing the exact statement of interest ex ante, a random algorithm can be used to make statements about the probability that a given conjecture holds for a set of relative size $1-\delta$.

Computing numerous random examples and then using statistical inference to summarize the findings is not a new idea (see e.g. Judd (1997)) but has not previously been formalized taking into account finite precision arithmetics of actual computations. The for this method most important practical insight of this paper is about the interplay of errors in computation, ϵ, and the number of connected components. One can estimate the Lebesgue measure of the set Φ by randomly drawing examples if the number of connected components is orders of magnitude smaller than $1/\epsilon$. Otherwise, it is not possible to even make probabilistic statements about the size of Φ.

The most important shortcoming of the paper is its focus on finite economies. It is subject to further research to generalize the results to recursive equilibria of dynamic models.

Appendix A: Computation of exact equilibria

For finite economies, it is well known since Scarf (1967) how to compute an ϵ-equilibrium for any given economy and any $\epsilon > 0$. This implies that it is possible, for any l, $\epsilon > 0$ and e to find one x_l such that $\|f_l(x_l, e)\| < \epsilon$. I refer to these x_l as ϵ-equilibria.

In general it is difficult to relate these ϵ-equilibria to exact equilibria. However, if the economic fundamentals are analytic, Smale's so called 'alpha-method' can be used to determine bounds on how far the computed equilibrium allocation and prices are from an exact equilibrium. Since it is not very well known in economics, Smale's method is summarized for completeness. The following results are from Blum et al (1998), Chapter 8.

Let $D \subset \mathbb{R}^n$ be open and let $f : D \to \mathbb{R}^n$ be analytic. For $z \in D$, define $f^{(k)}(z)$ to be the k'th derivative of f at z. This is a multi-linear operator which maps k-tuples of vectors in D into $\mathbb{R}^n$. Define the norm of an operator A to be

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}.$$

Suppose that the Jacobian of f at z, $f^{(1)}(z)$ is invertible and define

$$\gamma(z) = \sup_{k \geq 2} \left\| \frac{(f^{(1)}(z))^{-1} f^{(k)}(z)}{k!} \right\|^{\frac{1}{(k-1)}}$$

and

$$\beta(z) = \|(f^{(1)}(z))^{-1} f(z)\|.$$

THEOREM 5 *Given a $\bar{z} \in D$, suppose the ball of radius $(1 - \frac{\sqrt{2}}{2})/\gamma(\bar{z})$ around $\bar{z}$ is contained in D and that*

$$\beta(\bar{z})\gamma(\bar{z}) < 0.157.$$

Then there exists a $\tilde{z} \in D$ with

$$f(\tilde{z}) = 0 \text{ and } \|\bar{z} - \tilde{z}\| \leq 2\beta(\bar{z}).$$

In the economic example in Section 3.2 I use Newton's method to compute approximate zeros. The value of the equilibrium functions at the solutions are around 10^{-10} and it can be verified that β is very small and γ sufficiently large to ensure interiority of the ball around $\bar{z}$.

References

[1] Basu, S., R. Pollack and M.-F. Roy, (2003), *Algorithms in Real Algebraic Geometry*, Springer Verlag, New York.

[2] Berarducci, A. and T. Servi, (2004), An effective version of Wilkie's theorem of the complement and some effective o-minimality results, *Annals of Pure and Applied Logic*, 125, 43–74.

[3] Blum, L., F. Cucker, M. Shub and S. Smale, (1998), *Complexity and Real Computation*, Springer Verlag, New York.

[4] Blume, L. and W.R. Zame, (1993), The algebraic geometry of competitive equilibrium, in W. Neuefeind, ed., *Essays in General Equilibrium and International Trade: In Memoriam Trout Rader*, Springer-Verlag, New York.

[5] van den Dries, L., (1999), *Tame Topology and O-minimal Structures*, Cambridge University Press.

[6] Geanakoplos, J. and F. Kubler, (2003), Dollar denominated debt and optimal security design, Cowles Foundation DP 1449.

[7] Gabrielov, A. and N. Vorobjov, (2004), Complexity of Computation with Pfaffian and Noetherian Functions, in: *Normal Forms, Bifurcations and Finiteness Problems in Differential Equations*, Kluwer.

[8] Judd, K. (1997), Computational economics and economic theory: Complements or substitutes?, *Journal of Economic Dynamics and Control*, 21, 907–942.

[9] Khovanskii, (1991), *Fewnomials*, Translations of Mathematical Monographs, Vol. 88, AMS, Providence.

[10] Koiran, P., (1995), Approximating the volume of definable sets, *Proc. 36th IEEE Symposium on Foundations of Computer Science*, 134-141.

[11] Kubler, F. and K. Schmedders, (2005), Approximate versus exact equilibria in dynamic economies, *Econometrica*, forthcoming.

[12] L'Ecuyer, P., (2004), Random number generation, in: J. E. Gentle, W. Haerdle, and Y. Mori, eds.: *Handbook of Computational Statistics*, Springer-Verlag.

[13] Marker, D., (2002), *An Introduction to Model Theory*, Springer-Verlag.

[14] Richter, M.K. and K.-C. Wong, (2000), Definable utility in o-minimal structures *Journal of Mathematical Economics*, 34, 159–172.

[15] Rojas, J.M., (2000), Some speed-ups and speed limits in real algebraic geometry, *Journal of Complexity*, 16, 552–571.

[16] Scarf, H., (1967), On the computation of equilibrium prices, In: *Ten Studies in the Tradition of Irving Fisher* John Wiley and Sons: New York.

[17] Shoven, J.B., (1976), The incidence and efficiency effects of taxes on income from capital, *Journal of Political Economy*, 84, 1261–1283.

[18] Wilkie, A.J., (1996), Model completeness results of restricted Pfaffian functions and the exponential function, *Journal of the AMS*, 9, 1051–1094.