INDETERMINACY, NONPARAMETRIC CALIBRATION AND COUNTERFACTUAL EQUILIBRIA

By

Donald J. Brown and Ravi Kannan

June 2003

COWLES FOUNDATION DISCUSSION PAPER NO. 1426

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

YALE UNIVERSITY

Box 208281

New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Indeterminacy, Nonparametric Calibration and Counterfactual Equilibria

Donald J. Brown* and Ravi Kannan†

June 15, 2003

Abstract

We propose a nonparametric approach to multiple calibration of numerical general equilibrium models, where counterfactual equilibria are solutions to the Walrasian inequalities. We present efficient approximation schemes for deciding the solvability of Walrasian inequalities.

Keywords: Applied general equilibrium analysis, Walrasian inequalities, O-minimal structures, Monte Carlo algorithms

JEL Classification: C62, C63, D51, D58

1 Introduction

Numerical specifications of applied general equilibrium models are inherently indeterminate. Simply put, there are more unknowns (parameters) than equations (general equilibrium restrictions). Calibration of parameterized numerical general equilibrium models resolves this indeterminacy using market data from a "benchmark year"; parameter values gleaned from the empirical literature on production functions and demand functions; and the general equilibrium restrictions. The calibrated model allows the simulation and evaluation of alternative policy prescriptions, such as changes in the tax structure, by using Scarf's algorithm or one of its variants to compute counterfactual equilibria. Not surprisingly, the legitimacy of calibration as a methodology for specifying numerical general equilibrium models is the subject of an ongoing debate within the profession, ably surveyed by Dawkins et al. (2002). In their survey, they briefly discuss multiple calibration. That is, choosing parameter values for numerical general equilibrium models consistent with market data for two or more years. It is the implications of this notion that we explore in this paper.

*Cowles Foundation, Yale University

†Department of Computer Science, Yale University

Our approach derives from Varian's unique insight that nonparametric analysis of demand or production data admits extrapolation, i.e., "given observed behavior in some economic environments, we can forecast behavior in other environments," Varian (1982, 1984). The forecast behavior in applied general equilibrium analysis is the set of counterfactual equilibria. We extend the analyses of Brown and Matzkin (1996) and Brown and Shannon (2000), where the Walrasian and dual Walrasian inequalities are derived, to encompass the computation and evaluation of counterfactual equilibria.

Here is an example inspired by the discussion of extrapolation in Varian (1982), illustrating the nonparametric formulation of decideable counterfactual propositions in demand analysis. Suppose we observe a consumer chosing a finite number of consumption bundles x_i at market prices p_i, i.e., $(p_1, x_1), (p_2, x_2), ..., (p_n, x_n)$. If the demand data is consistent with utility maximization subject to a budget constraint, i.e., satisfies GARP, the generalized axiom of revealed preference, then there exists a solution of the Afriat inequalities, U, that rationalizes the data, i.e., if $p_i \cdot x \leq p_i \cdot x_i$ then $U(x_i) \geq U(x)$ for $i = 1, 2, ..., n$, where U is concave, continuous, monotone and nonsatiated (Afriat, 1967; Varian, 1983). Hence we may pose the following question for any two unobserved consumption bundles $\bar{x}$ and $\hat{x}$: Will $\bar{x}$ be revealed preferred to $\hat{x}$ for **every** solution of the Afriat inequalities? An equivalent formulation is the counterfactual proposition: $\bar{x}$ is not revealed preferred to $\hat{x}$ for some price vector p and some utility function U, a solution of the Afriat inequalities.

This proposition can be expressed in terms of the solution set for the following family of polynomial inequalities: The Afriat inequalities for the augmented data set $(p_1, x_1), (p_2, x_2), ..., (p_n, x_n)$, $(p, \hat{x})$ and the inequality $p \cdot \bar{x} > p \cdot \hat{x}$, where p is unobserved. If these inequalities are solvable, then the stated counterfactual proposition is true. If not, then the answer to our original question is yes. Notice that n of the Afriat inequalities are quadratic in the unobservables, i.e., the product of the marginal utility of income at $\hat{x}$ and the price vector p.

Brown and Matzkin (1996) characterized the Walrasian model of competitive market economies for data sets consisting of a finite number of observations on market prices, income distributions and aggregate demand. The Walrasian inequalities, as they are called here, are defined by the Afriat inequalities for individual demand and budget constraints for each consumer; the Afriat inequalities for profit maximization over a convex aggregate technology; and the aggregation conditions that observed aggregate demand is the sum of unobserved individual demands. The Brown–Matzkin theorem states that market data is consistent with the Walrasian model if and only if the Walrasian inequalities are solvable for the unobserved utility levels, marginal utilities of income and individual demands. Since individual demands are unobservable, the Afriat inequalities for each consumer are quadratic in the unobservables, i.e., the product of the marginal utilities of income and individual demands.

A decision method for this system of Walrasian inequalities constitutes a specification test for multiple calibration of numerical general equilibrium models, i.e., the market data is consistent with the Walrasian model if and only if the Walrasian inequalities are solvable. In our section on algorithms, we give an effective deterministic algorithm for this decision problem. If the system of Walrasian inequalities are solvable, then for every $\varepsilon > 0$, the algorithm computes a finite ε-net of solutions. The algorithm is based on the following observation: There is a finite set of candidate marginal utilities of income (one per agent per observation) such that every set of consumption bundles admitting a solution of the Afriat inequalities with strictly quadratically concave utilities, actually admits a solution with strictly concave utilities with one of our candidate marginal utilities of income. Moreover, this solution is the solution of a linear program.

An important point is that this set of candidates has cardinality $(1/\varepsilon)^{NT}$ where $\varepsilon > 0$ is a parameter, N is the number of observations and T the number of agents. Hence the algorithm will run in time bounded by a function which is polynomial in the **number of commodities** and exponential only in N and T. In situations involving a large number of commodities and a small N, T, this is very efficient. Note that trade between countries observed over a small number of periods is an example.

A more challenging problem is the computation of counterfactual equilibria. Fortunately, a common restriction in applied general equilibrium analysis is the assumption that consumers are maximizing homothetic utility functions subject to their budget constraints. Afriat (1981) and subsequently Varian (1983) derived a family of inequalities in terms of utility levels, market prices and incomes that characterize consumer's demands under this assumption. We shall refer to these inequalities as the homothetic Afriat inequalities. Brown and Lovász in Brown (1995) observed that the homothetic Afriat inequalities can be expressed as a finite family of quadratic polynomial inequalities over the exponential reals, i.e., the real number field with exponentiation. Wilke in 1991, published in Wilke (1996) proved that the exponential reals in an 0-minimal structure on the field of real numbers — see Van den Dries (1996) for a brief overview of 0-minimal structures, including the exponential reals. The importance of 0-minimal structures for this paper is the deep and surprising theorem of Laskowski (1992) that 0-minimal structures admit a uniform law of large numbers due to Vapnik and Čerovenkis (1971). We use this law of large numbers to construct an efficient randomized approximation scheme for deciding the solvability of a definable system of Walrasian inequalities. Here definable simply means that the Walrasian inequalities can be expressed as a finite family of polynomial inequalities over the exponential reals (or more generally a 0-minimal structure on the field of real numbers).

The intuition underlying our random decision method is simple. First, we rewrite the Walrasian inequalities as a finite family of strict polynomial inequalities over the exponential reals by substituting the budget constraints and aggregation conditions into the Afriat inequalities. If the solution set is not empty then it is open and therefore a set of positive Lebesgue measure. If we can compute a lower bound v on the measure of a nonempty solution set, then by randomly "guessing" possible solutions to the Walrasian inequalities, we estimate μ, the measure of the solution set. If with high probability $\mu < v$ then we decide that the given system of Walrasian inequalities are not solvable. It is here that we invoke the Vapnik–Čerovenkis uniform law of large numbers.

As an application of our approach, we revisit the Harberger tax-model as it is exposited in Shoven and Walley (1992). We discuss the simulation and evaluation of a change in the taxation of capital in a two-sector model, assuming the market data available in multiple calibration of a numerical general equilibrium model.

Finally a much discussed criticism of calibration in Dawkins et al. (2002) is the absence of statistical analysis in calibrated general equilibrium models, e.g., the effect of random shocks to tastes or technology. Recently Brown and Calsamiglia (2003) considered a class of random quasilinear utility functions. They showed if the shocks to tastes have compact support, then utility maximization of random quasilinear utility functions subject to a budget constraint has refutable implications on finite data sets. In effect, they extended Afriat's analysis to random quasilinear utility functions. Here the quasilinear Afriat inequalities are quadratic in the unobservables, i.e., the product of the random shocks and individual consumptions. In Brown and Kannan (2003) we rebut the econometric claim given above by extending our specification test to the case where the quasilinear Afriat inequalities are random.

The remainder of the paper is organized as follows: the second section contains the economic models. Then we have a section where we discuss our decision methods. The Harberger tax-model discussion is the final section of the paper.

2 Economic Models

We consider an economy with L commodities and T consumers. Each agent has $\mathbb{R}^L_+$ as her consumption set. We restrict attention to strictly positive market prices $S = \{p \in \mathbb{R}^L_{++} : \sum_{i=1}^L p_i = 1\}$. The Walrasian model assumes that consumers have utility functions $u_t : \mathbb{R}^L_+ \to \mathbb{R}$, income I_t and that aggregate demand $\bar{x} = \sum_{t=1}^T x_t$, where

$$u_t(x_t) = \max_{\substack{\text{s.t.}\, p\cdot x \leq I_t \\ x \geq 0}} u_t(x).$$

Suppose we observe a finite number N of profiles of income distributions $\{I_t^r\}_{t=1}^T$, market prices p^r and aggregate demand $\bar{x}^r$, where $r = 1, 2, ..., N$, but we do not observe the utility functions or demands of individual consumers. When are these data consistent with the Walrasian model of aggregate demand? The answer to this question is given by the following two theorems of Brown and Matzkin (1996).

Theorem 1 (Brown and Matzkin) *There exist nonsatiated, continuous, strictly concave, monotone utility functions* $\{u_t\}_{t=1}^T$ *and* $\{x_t^r\}_{t=1}^T$*, such that* $u_t(x_t^r) = \max_{p^r \cdot x \leq I_t^r} u_t(x)$

and $\sum_{t=1}^{T} x_t^r = \bar{x}^r$*, where* $r = 1, 2, ..., N$*, if and only if* $\exists$ $\{\hat{u}_t^r\}$, $\{\lambda_t^r\}$ *and* $\{x_t^r\}$ *for* $r = 1, ..., N$*;* $t = 1, ..., T$ *such that*

$$\hat{u}_t^r < \hat{u}_t^s + \lambda_t^s p^s \cdot (x_t^r - x_t^s) \ (r \neq s = 1, ..., N;\ t = 1, ..., T) \tag{1}$$

$$\lambda_t^r > 0,\ \hat{u}_t^r > 0 \textit{ and } x_t^r \geq 0 \ (r = 1, ..., N;\ t = 1, ..., T) \tag{2}$$

$$p^r \cdot x_t^r = I_t^r \ (r = 1, ..., N;\ t = 1, ..., T) \tag{3}$$

$$\sum_{t=1}^{T} x_t^r = \bar{x}^r \ (r = 1, ..., N) \tag{4}$$

(1) and (2) constitute the strict Afriat inequalities; (3) defines the budget constraints for each consumer; and (4) is the aggregation condition that observed aggregate demand is the sum of unobserved individual consumer demand. This family of condition is called here the (strict) Walrasian inequalities.[1] The observable variables in this system of inequalities are the I_t^r, p^r and $\bar{x}^r$, hence this is a nonlinear family of polynomial inequalities in unobservable utility levels $\hat{u}_t^r$, marginal utilities of income λ_t^r and individual consumer demands x_t^r.

Theorem 2 (Brown and Matzkin) *There exist nonsatiated, continuous, strictly concave homothetic monotone utility functions* $\{u_t\}_{t=1}^T$ *and* $\{x_t\}_{t=1}^T$ *such that* $u_t(x_t^r) = \max_{p^r \cdot x \leq I_t^r} u_t(x)$ *and* $\sum_{t=1}^{T} x_t^r = \bar{x}_t^r$*, where* $r = 1, 2, ..., N$ *if and only if* $\exists \{\hat{u}_t^r\}$ *and* $\{x_t^r\}$ *for* $r = 1, ..., N$*;* $t = 1, ..., T$ *such that*

$$\hat{u}_t^r < \hat{u}_t^s \frac{p^s \cdot x_t^r}{p^s \cdot x_t^s} \ \ (r \neq s = 1, ..., N;\ t = 1, ..., T) \tag{5}$$

$$\hat{u}_t^r > 0 \textit{ and } x^r \geq 0 \ \ (r = 1, ..., N;\ t = 1, ..., T) \tag{6}$$

$$p^r \cdot x_t^r = I_t^r \ \ (r = 1, ..., N;\ t = 1, ..., T) \tag{7}$$

$$\sum_{t=1}^{T} x_t^r = \bar{x}_t^r \ \ (r = 1, ..., N)$$

(5) *and* (6) *constitute the strict Afriat inequalities for homothetic utility functions.*

Following Brown and Lóvasz in Brown (1995), we make the change of variables $\hat{u}_t^r = e^{\hat{z}_t^r}$ and rewrite (5) and (6):

$$e^{\hat{z}_t^r - \hat{z}_t^s} < \frac{p^s \cdot x_t^r}{I_t^s} \ \ (r \neq s = 1, ..., N;\ t = 1, ..., T) \tag{8}$$

$$\hat{z}_t^r \in \mathbb{R} \ \ (r = 1, ..., N;\ t = 1, ..., T). \tag{9}$$

[1] Brown and Matzkin call them the equilibrium inequalities, but there are other plausible notions of equilibrium in market economies — see Brown and Calsamiglia (2003) for a discussion.

The observable variables in this system of inequalities are the I_t^s, p^r and $\bar{x}^r$, hence this is a family of polynomial inequalities over the exponential reals in unobservable utlity levels $\hat{z}_t^r$ and individual demands x_t^r. In addition, this is a family of smooth convex inequalities, where feasibility can be decided in polynomial time.

We now turn to the case of random quasilinear utility functions of the form considered by Brown and Calsamiglia (2003), i.e., $v(x, x_0, \varepsilon) = u(x) + \varepsilon \cdot x + x_0$, where $\bar{\varepsilon}_\ell \leq \varepsilon \leq \bar{\varepsilon}_u$ and prices are normalized such that x_0 is the numeraire good.

Theorem 3 (Brown and Calsamiglia) *There exist nonsatiated continuous strictly concave utility functions* $\{u_t\}_{t=1}^T$, $\{x_t^r\}_{t=1}^T$ *and* $\{\varepsilon_t^r\}_{t=1}^T$ *such that* $u_t(\hat{x}_t^r) + \varepsilon_t^r \cdot \hat{x}_t^r + \hat{x}_{0,t}^r = \max_{x_t^r}\{u_t(x_t^r) + \varepsilon_t^r \cdot x_t^r - p^r \cdot x_t^r + I_t^r\}$ *and* $\sum_{t=1}^T x_t^r = \bar{x}^r$, *where* $r = 1, 2, ..., N$, *if and only if* $\exists\{\hat{u}_t^r\}$, $\{\varepsilon_t^r\}$ *and* $\{x_t^r\}$ *for* $r = 1, ..., N$; $t = 1, ..., T$ *such that*

$$\hat{u}_t^r < \hat{u}_t^s + (p^s - \varepsilon_t^s) \cdot (x_t^r - x_t^s) \quad (r \neq s = 1, ..., N;\ t = 1, ..., T) \tag{10}$$

$$\hat{u}_t^r > 0,\ \bar{\varepsilon}_\ell^r \leq \varepsilon_t^r \leq \bar{\varepsilon}_u^r \text{and } x_t^r \geq 0 \quad (r = 1, ..., N;\ t = 1, ..., T) \tag{11}$$

$$p^r \cdot x_t^r = I_t^r \quad (r = 1, ..., N;\ t = 1, ..., T)$$

$$\sum_{t=1}^T x_t^r = \bar{x}^r \quad (r = 1, ..., N)$$

Brown and Shannon (2000) proposed a family of polynomial inequalities in terms of the dual strict Afriat inequalities which we find more useful for our analysis.

The dual strict Afriat inequalities for each consumer t can be expressed as follows:

$$\hat{v}_t^r > \hat{v}_t^s - \lambda_t^s I_t^s x_t^s \cdot \left(\frac{p^r}{I_t^r} - \frac{p^s}{I_t^s}\right) \ (r \neq s = 1, ..., N;\ t = 1, ..., T) \tag{12}$$

$$\lambda_t^r > 0 \text{ and } x_t^r \gg 0\ (r = 1, ..., N;\ t = 1, ..., T) \tag{13}$$

Theorem 4 (Brown and Shannon) *There exist numbers* $\hat{v}_t^r$, λ_t^r *and vectors* x_t^r *for* $(r = 1, ..., N;\ t = 1, ..., T)$ *satisfying the dual strict Afriat inequalities* (12) *and* (13) *if and only if there exist numbers* $\hat{u}_t^r$, λ_t^r *and vectors* x_t^r *for* $(r = 1, ..., N;\ t = 1, ..., T)$ *satisfying the strict Afriat inequalities* (1) *and* (2).

Hence we define the dual Walrasian inequalities as (3), (4), (12) and (13) where now the data is consistent with the Walrasian model of aggregate demand if and only if the dual Walrasian inequalities are solvable.

Brown and Shannon (2000) in their Lemma 1 also show that any solution of the dual strict Afriat inequalities gives rise to C^∞ functions $w_t : \mathbb{R}_+^{L+1} \to \mathbb{R}$ where $w_t(p, I)$ is convex in (p/I), strictly increasing in I and strictly decreasing in p such that $w_t(p_t^r, I_t^r) = \hat{v}_t^r$ and $D_{p/I} w_t(p_t^r, I_t^r) = -\lambda_t^r I_t^r x_t^r$ for $(r = 1, ..., N;\ t = 1, ..., T)$. Shannon and Zame (2002) define a smooth convex function $w_t : \mathbb{R}_+^J \to \mathbb{R}$ as (strictly) quadratically convex on a convex subset $Y \subset \mathbb{R}^J_+$ if there is a constant $k_t > 0$ such that for each $x, y \in Y$, where $x \neq y$:

$$w_t(y) > w_t(x) + Dw_t(x) \cdot (y - x) + k_t ||x - y||^2. \tag{14}$$

They point out that any smooth strictly convex function on a compact convex subset $Y \subset \mathbb{R}^J$ is (strictly) quadratically convex. Their observation follows from the second order Taylor expansion of $w_t(x)$ and the fact that $D^2 w_t(x)$ is positive definite on $\mathbb{R}^J$, for all $x \in Y$. Let $\lambda_{\min}(x)$ = minimum eigenvalue of $D^2 w_t(x)$, then $k_t = \min_{x \in Y} \lambda_{\min}(x)$.

We extend their result to families of smooth strictly convex functions on a compact convex subset $Y \subset \mathbb{R}^J$, by restricting attention to families that are compact in the C^2-topology. In this case, we can choose one constant $k > 0$ for all the functions in the family.

Both the Brown–Matzkin and Brown–Shannon analyses extend to production economies, if social endowments $\bar{e}^r$ are observed in each period and firms are price-taking profit maximizers. We simply add the restriction: $p^r \cdot \bar{y}^s \leq p^r \cdot \bar{y}^r$ for $r, s = 1, ..., N$, where $\bar{y}^r = \bar{x}^r - \bar{e}^r$.

3 Algorithms

A Monte Carlo algorithm is a randomized algorithm which produces the correct solution with high probability. Solvability of the Walrasian inequalities is a decision problem, i.e., in each instance the answer is yes or no. The Monte Carlo algorithm for the Walrasian decision problem has non-zero probability that it errs only when the output is no, i.e., it has a one-sided error. Our algorithm derives from the celebrated result of Vapnik–Čerovenkis (1971) on the uniform deviations of relative frequencies from probabilities over a VC class.

A collection $\mathcal{C}$ of subsets of some space $\mathcal{X}$ picks out a certain subset E of a finite set $\{x_1, ..., x_n\} \subset \mathcal{X}$ if $E = \{x_1, ..., x_n\} \cap A$ for some $A \in \mathcal{C}$. $\mathcal{C}$ is said to shatter $\{x_1, ..., x_n\}$ if $\mathcal{C}$ picks out each of its 2^n subsets. The VC dimension of $\mathcal{C}$, denoted $V(\mathcal{C})$, is the smallest n for which no set of size n is shattered by $\mathcal{C}$. A collection $\mathcal{C}$ of measurable sets is called a VC class if its dimension $V(\mathcal{C})$ is finite.

We construct a VC class of sets from the Walrasian inequalities as follows. Consider the Walrasian inequalities as a parameterized family of strict linear inequalities, where the parameter set is the family of possible assignments of utility levels and demands to consumers satisfying the budget constraints and the aggregation conditions in the Walrasian inequalities. This is a compact convex subset of $\mathbb{R}^L_{++}$, which we call the state space and denote as Ω. Let $\Lambda \subset \mathbb{R}^J_{++}$ be the set of possible assignments of marginal utilities of income to each consumer in each observation, then w.o.l.o.g. we can assume $\Lambda \subset [0,1]^J$ since the Walrasian inequalities are homogeneous of degree one in utility levels and marginal utilities of income, where $J = TN$. Let $\Gamma(\omega)$ be the set of $\delta \in \Lambda$ such that (ω, δ) satisfy the Walrasian inequalities, where $\omega = \{\hat{w}_t^r, x_t^r\}$. A deep and surprising result of Laskowski (1992) is that $\{\Gamma(\omega)\}_{\omega \in \Omega}$ has finite VC dimension. Since $\Gamma(\omega)$ is open for all $\omega \in \Omega$, $\{\Gamma(\omega)\}_{\omega \in \Omega}$ is a VC class.

Invoking an inequality of Shannon and Zame (2002) on C^2-compact families of smooth concave utility functions, we derive a uniform lower bound v on $\lambda(\Gamma(\omega))$, for $\Gamma(\omega) \neq \phi$, where λ is Lebesgue measure on $\mathbb{R}^J$. That is, $\inf_{\omega \in \Omega, \Gamma(\omega) \neq \phi} \lambda(\Gamma(\omega)) \geq v$.

Returning to the dual strict Afriat inequalities and assuming they were generated by utilities belonging to a C^2-compact family of smooth strictly convex utilities on a compact cube containing $\bar{x}^r$ in its interior for $r = 1, 2, ..., N$:

$$w_t^r > w_t^s - \lambda_t^s I_t^s x_t^s \cdot \left(\frac{p^r}{I_t^r} - \frac{p^s}{I_t^s} \right) + k \left\| \frac{p^r}{I_t^r} - \frac{p^s}{I_t^s} \right\|^2 \ (r, s = 1, ..., N; \ t = 1, ..., T) \quad (15)$$

$$\lambda_t^r > 0 \text{ and } x_t^r \gg 0 \ (r = 1, ..., N; \ t = 1, ..., T). \quad (16)$$

Suppose these inequalities have a solution $\hat{\lambda}_t^r$ for given w_t^r and x_t^r, recall that we observe the p^r and I_t^r. We wish to construct ε-balls around $\hat{\lambda}_t^r$ such that $\forall (\lambda_t^r) \in \mathbb{B}_\varepsilon(\hat{\lambda}_t^r)$, they solve the strict dual Afriat inequalities given in (12) and (13). To this end let

$$\begin{aligned}
\delta &= \min_{\substack{r \neq s \\ 1 \leq r,s \leq N \\ 1 \leq t \leq T}} k \left\| \frac{p^r}{I_t^r} - \frac{p^s}{I_t^s} \right\|^2 \\
\eta &= \max_{\substack{r \neq s \\ 1 \leq r,s \leq N \\ 1 \leq t \leq T}} \left| \bar{x}^r \cdot \left(\frac{p^r}{I_t^r} - \frac{p^s}{I_t^s} \right) \right| \times \max_{1 \leq r \leq N} \sum_{t=1}^{T} I_t^r \text{ and} \\
\varepsilon &= \min \left\{ \frac{\delta}{\eta}, \min_{\substack{1 \leq r \leq N \\ 1 \leq t \leq T}} (w_t^r, \lambda_t^r) \right\}.
\end{aligned}$$

Then

$$\begin{aligned}
\hat{w}_t^r &> \hat{w}_t^s - \hat{\lambda}_t^s I_t^s x_t^s \cdot \left(\frac{p^r}{I_t^r} - \frac{p^s}{I_t^s} \right) + k \left\| \frac{p^r}{I_t^r} - \frac{p^s}{I_t^s} \right\|^2 \\
&\geq \hat{w}_t^s - \hat{\lambda}_t^s I_t^s x_t^s \cdot \left(\frac{p^r}{I_t^r} - \frac{p^s}{I_t^s} \right) + \delta \\
&\geq \hat{w}_t^s - \hat{\lambda}_t^s I_t^s x_t^s \cdot \left(\frac{p^r}{I_t^r} - \frac{p^s}{I_t^s} \right) + \frac{\delta \left[I_t^s x_t^s \cdot \left(\frac{p^r}{I_t^r} - \frac{p^s}{I_t^s} \right) \right]}{\eta} \\
&\geq \hat{w}_t^s - \hat{\lambda}_t^s I_t^s x_t^s \cdot \left(\frac{p^r}{I_t^r} - \frac{p^s}{I_t^s} \right) + \varepsilon \left[I_t^s x_t^s \cdot \left(\frac{p^r}{I_t^r} - \frac{p^s}{I_t^s} \right) \right] \\
&= \hat{w}_t^s - (\hat{\lambda}_t^s - \varepsilon) I_t^s x_t^s \cdot \left(\frac{p^r}{I_t^r} - \frac{p^s}{I_t^s} \right).
\end{aligned}$$

That is,

$$\hat{w}_t^r > \hat{w}_t^s - (\hat{\lambda}_t^s - \varepsilon) I_t^s x_t^s \cdot \left(\frac{p^r}{I_t^r} - \frac{p^s}{I_t^s} \right).$$

Hence if we know k, then we can compute a uniform ball around any solution of the strict dual Afriat inequalities for given w_t^r and x_t^r and observed p^r and I_t^r for $r = 1, ..., N$ and $t = 1, ..., T$. The volume of this ball, call it v, defines a binomial distribution on the unit cube $[0,1]^J$, given the uniform distribution on $[0,1]^J$. That is, v gives a lower bound on the probability that a random draw from $[0,1]^J$ is a J-tuple of marginal utilities of income such that the given data w_t^r and x_t^r satisfy the dual Walrasian inequalities.

The classical inequality of Vapnik and Čerovenkis (1971) on uniform deviations of relative frequencies from probabilities over a VC class can now be used to estimate the *probability* $1-\delta$ that the solution set for the given Walrasian inequalities has measure less than v as a function of the the number of Monte Carlo iterations m, v and d, the VC dimension of $\{\Gamma(\omega)\}_{\omega\in\Omega}$. We randomly draw points from the unit cube, $[0,1]^J$, endowed with the uniform distribution. Given this random assignment of marginal utilities of income to consumers in each observation, we can decide in polynomial time if the family of linear inequalities derived from the Walrasian inequalities has a solution. If the answer is yes, then the algorithm terminates. If the answer is no, we continue drawing samples until we find a solution to the derived family of linear inequalities or the sample size is m.

If the sample size is m and we have not found a solution to the Walrasian inequalities then by the VC inequality we know with probability $1-\delta$ that the solution set has measure less than v. In this case we decide that the given Walrasian inequalities are not solvable. The running time of our randomized approximation scheme is polynomial in $1/v$, $\log 1/\delta$ and m, hence efficient in the family of randomized approximation schemes. Motwani and Raghavan (1995) discuss efficient randomized algorithms for decision problems in Chapter 11. Now, we give the details.

Let $\mathbb{P}$ be an arbitrary probability measure on the Borel field $\mathcal{F}$ on $\mathbb{R}^S$ and suppose $u_1, ..., u_m$ are drawn independently from $\mathbb{P}$. Then $\bar{u} = (u_1, ..., u_m)$ has probability distribution $\mathbb{P}^m$, the m-fold product measure defined by $\mathbb{P}$. $\mathbb{P}_{m;\bar{u}}$ is the empirical probability measure defined by $\bar{u}$, i.e., $\mathbb{P}_{m;\bar{u}}(E) =$ fraction of u_i which lie in E, for each $\mathcal{F}$-measurable, given triplet $(\mathbb{R}^J, \mathcal{F}, \mathbb{P})$, i.e., $E \in \mathcal{F}$, subset $E \subset \mathbb{R}^J$. We assume that $\mathcal{C}$ is a VC class of subsets, $C \subseteq \mathcal{F}$ measurable with respect to $\mathbb{P}$, with VC-dimension d. Define $\rho(\bar{u}) = \sup_{S\in\mathcal{C}} ||\mathbb{P}_{m;\bar{u}}(S) - \mathbb{P}(S)||$.

Theorem (Blumer, et al., 1989) *Assuming that $\rho(\bar{u})$ is $\mathcal{F}$-measurable and for fixed $\delta, v > 0$ if*

$$m \geq \max\left\{ \frac{4}{v} \log_2 \frac{2}{\delta},\ \frac{8d}{v} \log_2 \frac{13}{v} \right\} \text{ then}$$

$$\mathbb{P}^m(\{(u_1, ..., u_m) = \bar{u} : \rho(u_1, ..., u_m) > v\}) < \delta.$$

Since $\mathbb{P}^m\{(u_1,...,u_m) = \bar{u} : \exists S \in \mathcal{C}$ s.t. $\mathbb{P}(S) > v$ and $\mathbb{P}_{m;\bar{u}}(S) = 0\} \leq \mathbb{P}^m\{(u_1,...,u_m) : \rho(u_1,...,u_m) > v\} < \delta$ for sufficiently large m, we see that $\mathbb{P}^m(\{(u_1,...,u_m) : \exists S \in \mathcal{C}$ s.t. $\mathbb{P}(S) > v$ and $\mathbb{P}_{m;\bar{u}}(S) = 0\}) < \delta$.

Let

$$\begin{aligned} A &= \{(u_1,...,u_m) = \bar{u} : \exists S \in \mathcal{C} \text{ s.t. } \mathbb{P}(S) > v \text{ and } \mathbb{P}_{m;\bar{u}}(S) = 0\} \\ B &= \{(u_1,...,u_m) = \bar{u} : \rho(\bar{u}) > v\} \text{ and} \\ C &= \{(u_1,...,u_m) = \bar{u} : (\forall S \in \mathcal{C})\mathbb{P}_{m;\bar{u}}(S) = 0\} \end{aligned}$$

If $D = B \cap C$, then $D \subseteq A$. Hence $\mathbb{P}^m(D) \leq \mathbb{P}^m(A)$ and

$$\mathbb{P}^m(\{(u_1,...,u_m) = \bar{u} : \exists S \in \mathcal{C} \text{ s.t. } \mathbb{P}(S) > v \text{ and } (\forall S \in \mathcal{C})\mathbb{P}_{m;\bar{u}}(S) = 0\}) < \delta \quad (17)$$

To bound the confidence level that the Walrasian inequalities are not solvable, given that a sample of size m does not produce a solution, we compute a lower bound on $\mathbb{P}^m(C)$. For any sample point $\alpha \in [0,1]^J$ with the property that for some r, s and t, where $r \neq s$; $1 \leq r, s \leq N$ and $1 \leq t \leq T$:

$$\bar{x}_t I_t^s \cdot \left(\frac{p^r}{I_t^r}\right) - I_t^s < \frac{\hat{v}_t^s - \hat{v}_t^r}{\lambda_t^s}$$

violates the corresponding dual strict Afriat inequality:

$$\hat{v}_t^r > \hat{v}_t^s - \lambda_t^s I_t^s x_t^s \cdot \left(\frac{p^r}{I_t^r} - \frac{p^s}{I_t^s}\right)$$

for all $0 \ll x_t^s \leq \bar{x}_t$. This is an open condition, hence the set of such α is a set of positive Lebesgue measure, say ξ. Hence the confidence level $\mathbb{P}^m(B|C) \leq \delta/\xi$. That is,

$$\mathbb{P}^m\{(u_1,...,u_m) : \exists S \in \mathcal{C} \text{ s.t. } \mathbb{P}(S) > v | (\forall S \in \mathcal{C})\mathbb{P}_{m;\bar{u}}(S) = 0\} < \delta/\xi \quad (18)$$

Our algorithm is defined by (18), where $\mathcal{C} = \{\Gamma(\omega)\}_{\omega\in\Omega}$ and $\mathbb{P} = \lambda$, Lebesgue measure on $\mathbb{R}^J$. Given the data and k, we can compute v. Then $\lambda(\Gamma(\omega)) \geq v$ for all $\omega \in \Omega$. The VC-dimension of $\{\Gamma(\omega)\}_{\omega\in\Omega}$, d, is $s(J+2)$, i.e., for fixed ω the dual Walrasian inequalities are linear in the unknown marginal utilities of income, where s is the number of dual Walrasian inequalities after substituting the budget constraints and aggregation conditions into the (strict) dual Afriat inequalities. Here we use the well-known result on the VC dimension of the family of half spaces on $\mathbb{R}^J$ and the VC dimension of intersections of VC classes — see exercise 14 in Section 2.6 in Van der Vaart and Wellner (1996) for discussion. We choose $\delta \in (0,1)$ and draws from $[0,1]^J$ endowed with the uniform distribution, where $m \geq \max\{\frac{4}{v}\log_2\frac{2}{\delta}, \frac{8d}{v}\log_2\frac{13}{v}\}$. For each draw we use an interior-point polynomial time linear programming algorithm to see if there is a solution to the dual Walrasian inequalities for the chosen marginal utilities of income. We continue drawing samples until we find a solution or we have drawn m samples and found no solution. In this case $\mathbb{P}_{m;\bar{u}}(\Gamma(\omega)) = 0$ for all $\omega \in \Omega$, where $\bar{u} = (u_1, ..., u_m)$ is the realized sample, and we decide with confidence level $1 - \delta/\xi$ that the Walrasian inequalities are not solvable for the given data set. This algorithm is efficient, i.e., polynomial in $1/v$ and $\log 1/\delta$.

In this relatively simple case, we can do better. In fact, we now give a deterministic decision method suggested by our random decision scheme. The random scheme is necessary for deciding the existence of counterfactual equilibria as discussed below in Section 3.2.

3.1 A Deterministic Decision Method

In this section, we describe a deterministic algorithm to find solutions of the dual strict inequalities. We do so by simply showing that there is a finite set of "candidate" λ's - marginal utilities of income (one for each agent for each observation)- such that for every solution of x 's (individual consumption bundles), one of the candidate λ 's works with the x. First we assume here that the observables $p^r, \bar{x}^r, I_t^r,\ r = 1, 2, \ldots N, t = 1, 2, \ldots T$ are given and further, we assume that

$$\left|\left|\frac{p^r}{I_t^r} - \frac{p^s}{I_t^s}\right|\right| \geq \varepsilon_1, \forall r \neq s \quad \text{where } \varepsilon_1 > 0 \tag{19}$$

This assumption is reasonable; indeed if p^r, p^s as vectors make a small angle, then p^r is nearly a scalar multiple of p^s and the observation r may be discarded.We also let

$$\eta = \max_{r,t} I_t^r.$$

If the dual strict Afriat inequalities have a solution, we may scale the w, λ so that after scaling, we may assume that all λ are between 0 and 1. Similarly, for the strictly quadratically convex Afriat inequalities (15), (16), we may scale the w, λ as well as the k so that after scaling, we may assume that the solution satisfies $\lambda_t^r \in [0\ ,\ 1]$. We will make this assumption.

Lemma 1 *Suppose* $\hat{w}_t^r, \hat{\lambda}_t^r \in [0, 1], \hat{x}_t^r$ *solve the strictly quadratically convex dual Afriat inequalities* (15), (16), (3), (4). *Then for any* λ_t^r *with*

$$\hat{\lambda}_t^r \leq \lambda_t^r \leq \hat{\lambda}_t^r + \frac{k\varepsilon_1^2}{\eta} \quad \forall t, r,$$

we have that $\hat{w}_t^r$, λ_t^r, $\hat{x}_t^r$ *satisfy the strict dual Afriat inequalities* — (12), (13), (3), (4).

Proof

$$\begin{aligned}
&\hat{w}_t^s - \lambda_t^s I_t^s \hat{x}_t^s \left(\frac{p^r}{I_t^r} - \frac{p^s}{I_t^s}\right) \\
&\leq \hat{w}_t^s - \hat{\lambda}_t^s I_t^s \hat{x}_t^s \cdot \frac{p^r}{I_t^r} + \lambda_t^s I_t^s \frac{\hat{x}_t^s \cdot p^s}{I_t^s} \text{ since } \hat{\lambda}_t^s \leq \lambda_t^s \text{ and } I_t^s \hat{x}_t^s \cdot p^r / I_t^r \geq 0 \\
&\leq \hat{w}_t^s - \hat{\lambda}_t^s I_t^s \hat{x}_t^s \frac{p^r}{I_t^r} + \hat{\lambda}_t^s I_t^s \frac{\hat{x}_t^s \cdot p^s}{I_t^s} + \frac{k\varepsilon_1^2}{\eta} I_t^s \text{ using (3)} \\
&\leq \hat{w}_t^s - \hat{\lambda}_t^s I_t^s \hat{x}_t^s \left(\frac{p^r}{I_t^r} - \frac{p^s}{I_t^s}\right) + k\varepsilon_1^2 < \hat{w}_t^r \text{ by hypothesis.}
\end{aligned}$$

Now define

$$S = \left\{ \lambda : \lambda_t^r \text{ is an integer multiple of } \frac{k\varepsilon_1^2}{\eta} \quad 0 \leq \lambda_t^r \leq 1 + \left(\frac{k\varepsilon_1^2}{\eta}\right) \right\}.$$

Then, by the lemma it follows that

Lemma 2 *For any* $\hat{w}_t^r, \hat{\lambda}_t^r \in [0,1], \hat{x}_t^r$ *solving the strictly quadratically convex dual Afriat inequalities* (15), (16), (3), (4), *there exists a* $\lambda \in S$ *such that* $\hat{w}_t^r$, λ_t^r, $\hat{x}_t^r$ *satisfy the strict dual Afriat inequalities* — (12), (13), (3), (4).

This immediately yields an algorithm: we enumerate the set S and then for each candidate $\lambda \in S$, we now solve a **linear program** to see if the inequalities (12), (13), (3) and (4) have a solution. If there is no solution for any λ, then we conclude that the system (15), (16), (3), (4) has no solution with this k. Otherwise, we would have found a solution to (12), (13), (3), (4). Note that

$$|S| \leq \frac{(1 + (k\varepsilon_1^2/\eta))^{NT}}{(k\varepsilon_1^2/\eta)^{NT}} \approx \frac{1}{(k\varepsilon_1^2/\eta)^{NT}},$$

which is only exponential in NT and independent of L, the number of commodities. For each candidate λ, we solve a linear program where the computational time is bounded above by a polynomial in N, T, L. Thus, when the number of observations and number of agents are small compared to the number of commodities, this is a very efficient algorithm.

3.2 Applications

In multiple calibration, two or more years of market data together with empirical studies on demand and production functions and the general equilibrium restrictions are used to specify numerical general equilibrium models. The maintained assumption is that the market data in each year is consistent with the Walrasian model of market economies. This assumption which is crucial to the calibration approach is never tested, as noted in Dawkins et al. (2002).

The assumption of Walrasian equilibrium in the observed markets is testable, under a variety of assumptions on consumer's tastes, using the necessary and sufficient conditions stated in Theorems 1, 2, and 3 and the market data available in multiple calibration. In particular, Theorem 2 can be used as a specification test for the numerical general equilibrium models discussed in Shoven and Whalley (1992), where it is typically assumed that utility functions are homothetic.

Following Varian, we can extrapolate from the observed market data available in multiple calibration to unobserved market configurations. We simply augment the equilibrium inequalities defined by the observed data with additional polynomial inequalities characterizing possible but unobserved market configurations of utility levels, marginal utilities of income, individual demands, aggregate demands, income distributions and equilibrium prices. Counterfactual equilibria are defined as solutions to this augmented family of equilibrium inequalities.

In general, the Afriat inequalities in this system will be cubic in the product of unobserved marginal utilities of income, the unobserved equilibrium prices and unobserved individual demands. This is to be contrasted with observations that include the market prices where the Afriat inequalities are only quadratic in the product of the unobserved marginal utility of income and individual demand. It is here that we need the random decision methods.

4 The Harberger Tax-Model

We begin by recalling the two-sector model of the US economy. In this model there are two types of households or consumers; two types of firms or producers; two goods; and two factors of production, labor and capital. We assume that consumers have homothetic utility functions and are endowed with the factors of production. Firms have production functions that are homogeneous of degree one, hence make zero profits in equilibrium. Following Harberger, we assume that factors are inelastically supplied.

Assuming there are two years of data available, as is the case in a typical multiple calibration exercise, the Walrasian inequalities for the two-sector model constitute a specification test for the Harberger tax-model. In both years we observe: aggregate demand for the two goods; social endowments of capital and labor; the income distributions of households; and market prices of goods and factors, where labor is the numeraire good. We can now state the Walrasian inequalities for the two-sector model.

Households:

$$\hat{u}_t^r < \hat{u}_t^s \frac{p^s \cdot x_t^r}{p^s \cdot x_t^s} \quad (r \neq s = 1, 2; \; t = 1, 2) \tag{20}$$

$$\hat{u}_t^r > 0 \text{ and } x_t^r \geq 0 \quad (r = 1,2;\ t = 1,2) \tag{21}$$

$$p^r \cdot x_t^r = I_t^r \quad (r = 1,2;\ t = 1,2) \tag{22}$$

$$\sum_{t=1}^{T} x_t^r = \bar{x}^r \quad (r = 1,2) \tag{23}$$

Firms:

$$\hat{f}_t^r < \hat{f}_t^s \frac{q_t^s \cdot y_t^r}{q_t^s \cdot y_t^s} \quad (r \neq s = 1,2;\ t = 1,2) \tag{24}$$

$$\hat{f}_t^r > 0 \text{ and } y_t^r \geq 0 \quad (r = 1,2;\ t = 1,2) \tag{25}$$

$$q_t^r \cdot y_t^r = J_t^r \quad (r = 1,2;\ t = 1,2) \tag{26}$$

$$\sum_{t=1}^{T} y_t^r = \bar{y}^r \quad (r = 1,2) \tag{27}$$

where

x_t^r is consumer t's demand for goods in period r
$\hat{u}_t^r$ is consumer t's utility level in period r
I_t^r is consumer t's income in period r
p^r is the price vector for goods in period r
$\bar{x}^r$ is the aggregate demand for goods in period r

and

y_t^r is firm t's demand for factors in period r
$\hat{f}_t^r$ is firm t's output level in period r
J_t^r is firm t's revenue in period r
q_t^r is the price vector for factor demands by firm t in period r
$\bar{y}^r$ is the aggregate supply of factors in period r

If we now impose a per unit tax T on the use of capital by firm 1 then counterfactual equilibria are solutions to the equilibrium inequalities for the two-sector model, augmented in the following manner: The range of r in the equilibrium inequalities is now $\{1, 2, 3\}$ where $q_1^3 = (1, \bar{r} + T)$, $q_2^3 = (1, \bar{r})$ and $\bar{r} > 0$. In addition, we require $p^3 \gg 0$ and consumer's factor endowments to be the same in periods two and three. In each counterfactual equilibrium we compute the social loss due to the tax: $\frac{1}{2}T\Delta K_1$, where ΔK_1 is the change in demand for capital by firm 1. A decideable family of counterfactuals in this model are of the form: The augmented equilibrium inequalities and the inequality $\frac{1}{2}T\Delta K_1 > \alpha$, where α is known and fixed.

Acknowledgments The authors are pleased to acknowledge the comments and remarks of Rudiger Bachman, Brendan Beare and Michael Mahoney. The generous assistance of Glena Ames is also appreciated.

5 References

Afriat, S. (1967). "The Construction of a Utility Function from Expenditure Data," *International Economic Review*, 66–77.

——— (1981). "On the Constructability of Consistent Price Indices between Several Periods Simultaneously." In A. Deaton (ed.), *Essays in Applied Demand Analysis*. Cambridge: Cambridge University Presss.

Blumer, A., A. Ehrenfeucht, D. Haussler and M. Warmuth (1989). "Learnability and the Vapnik–Čerovenkis Dimension," *Journal of the ACM*, 36, 929–965.

Brown, D. (1995). "The Toronto Lectures," mimeographed.

Brown, D. and C. Calsamiglia (2003). "The Strong Law of Demand," Cowles Foundation Discussion Paper 1399, Yale University.

Brown, D. and R. Kannan (2003). "A Monte Carlo Algorithm for the Walrasian Decision Problem," mimeographed.

Brown, D. and R. Matzkin (1996). "Testable Restrictions on the Equilibrium Manifold," *Econometrica*, 64, 1249–1262.

Brown, D. and C. Shannon (2000). "Uniqueness, Stability, and Comparative Statics in Rationalizable Walrasian Markets," *Econometrica*, 68, 1529–1539.

Dawkins, C., T.N. Srinivasan and J. Whalley (2000). "Calibration." In J.J. Heckman and E. Leamer (eds.), *Handbook of Econometrics*, Vol. 5. New York: Elsevier.

Laskowsi, C. (1992). "Vapnik–Čerovenkis Classes of Definable Sets," *J. London Math. Soc.*, 245, 231–251.

Motwani, R. and P. Raghavan (1995). *Randomized Algorithms*. Cambridge: Cambridge University Press.

Shannon, C. and W. Zame (2002). "Quadratic Concavity and Determinacy of Equilibrium," *Econometrica*, 70, 631–662.

Shoven, J. and J. Walley (1992). *Applying General Equilibrium*. Cambridge: Cambridge University Press.

Van den Dries, L. (1996). "*O*-minimal Structures on the Field of Real Numbers," *Jber. d. Df. Math.-Verein*, 165–171.

Vapnik, V. and A. Čerovenkis (1971). "On the Uniform Convergence of Relative Frequencies of Events to Their Probabilities," *Theory Probab. Appl.*, 16, 264–280.

apnik, V. and A. Čerovenkis (1996). *Weak Convergence and Empirical Processes*, New York: Springer.

Varian, H. (1982). "The Nonparametric Approach to Demand Analysis," *Econometrica*, 50, 945–973.

——— (1983). "Non-parametric Tests of Consumer Behavior," *Review of Economic Studies*, 99–110.

——— (1984). "Nonparametric Approach to Production Analysis," *Econometrica*, 52, 579–597.

Wilke, A. (1996). "Model Completeness for Expansions of the Ordered Field of Real Numbers by Restricted Pfaffian Functions and the Exponential Function," *J.AMS*, 9, 1051–1094.