Nonparametric inference on the number of equilibria

Maximilian Kasy

Department of Economics, UC Berkeley

Three goals

1. Inference on the number of roots of functions which are nonparametrically identified
2. Relating different notions of equilibrium to the roots of identifiable functions
3. Testing whether there are multiple equilibria in the dynamics of neighborhood composition in US cities

This paper proposes

- a superconsistent estimator of the number of equilibria of economic systems,
- an inference procedure based on non-standard asymptotics.

More precisely, suppose:

- the equilibria of a system are solutions of $g(x) = 0$.
- g is nonparametrically identified by a conditional moment restriction.

This paper provides confidence sets for the number $Z(g)$ of solutions to the equation $g(x) = 0$.

Examples of multiple equilibria in economics

- **Urban segregation**: Becker and Murphy (2000), Card, Mas, and Rothstein (2008)
- **Household level poverty traps**: Dasgupta and Ray (1986)
- **Social norms**: Young (2008)
- **Agglomerations in economic geography**: Krugman (1991)
- **Market entry of firms**: Bresnahan and Reiss (1991), Berry (1992)
- **Poverty traps in macro models of economic growth:** Quah (1996), Azariadis and Stachurski (2005), Bowles, Durlauf, and Hoff (2006)
- **Financial market bubbles**: Stiglitz (1990), Lux (1995)
- ...

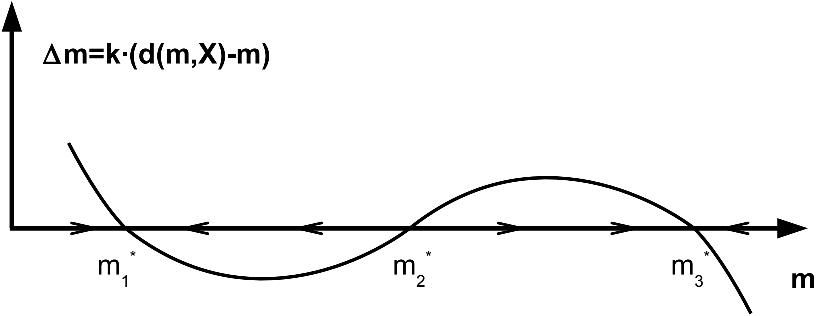
Example: Dynamics of neighborhood composition

In the search-model of the housing market proposed in my paper on "Identification in models of sorting with social externalities"

$$\Delta m = \kappa \cdot (d(m, X) - m), \tag{1}$$

where

- m is the minority share among households living in the neighborhood,
- d is the minority share among households *wanting* to live in the neighborhood,
- Δm is the change of m over a given time period,
- X are exogenous factors influencing relative demand,
- κ is a parameter reflecting search frictions.


Δm=k·(d(m,X)-m)
m₁*
m₂*
m₃*
m

Why should we care about multiple equilibria?

- They explain persistent inequality.
- They imply history dependence.
- They imply that “Big Push” interventions have a lasting effect.

Why should we care about inference on the number of equilibria?

- Because we should care about multiple equilibria.
- The statistical theory is mathematically interesting.

Two general setups, relating equilibria to roots

1) Static games of incomplete information (See Bajari, Hong, Krainer, and Nekipelov (2006)):

- Two players i, actions $a_i \in \{0,1\}$, public information s observed by the econometrician.
- Under exclusion restrictions, we can estimate the average response function of a player to the expected action σ_{-i} of the other player, $g_i\left(\sigma_{-i}, s\right)$.
- Bayesian Nash equilibria are given by solutions to $g(\sigma_1, s) = g_1(g_2(\sigma_1, s), s) - \sigma_1 = 0$.

2) Stochastic difference equations

- $\Delta X_{i,t+1} = X_{i,t+1} - X_{i,t} = g(X_{i,t}, \epsilon_{i,t})$
- number of roots of g in $X \approx$ number of "equilibrium regions"
- number of roots of nonparametric quantile regressions of $\Delta X_{i,t+1}$ on $X_{i,t} \geq$ number of roots of g in X

Roadmap

1. Inference procedure and its asymptotic justification, baseline case
2. Monte Carlo evidence
3. Generalizations: control variables, higher dimensional systems, stable and unstable equilibria
4. Identification and inference for games and for difference equations
5. Application to data on neighborhood composition (from Card, Mas, and Rothstein (2008))
6. Conclusion

Baseline case - Assumptions

Parameter of interest: number of roots

$$Z(g) := |\{x \in \mathscr{X} : g(x) = 0\}| \tag{2}$$

Assume:

- g has one-dimensional and compact domain and range (generalized later)
- observable data are i.i.d. draws of (Y_i, X_i)
- the density of X is bounded away from 0 on $\mathscr{X}$
- g is identified by a conditional moment restriction

$$g(x) = \text{argmin}_y E_{Y|X}[m(Y - y)|X = x] \tag{3}$$

Examples of conditional moment restrictions:

- $m(\delta) = \delta^2$ for conditional mean regression
- $m_q(\delta) = \delta(q - \mathbf{1}(\delta < 0))$ for conditional q^{th} quantile regression

Assumptions - continued

Assume, furthermore, that g is continuously differentiable and generic:

Definition (Genericity)

A continuously differentiable function g is called generic if

$$\{x : g(x) = 0 \ and \ g'(x) = 0\} = \varnothing$$

and if all roots of g are in the interior of $\mathscr{X}$.

Genericity of g implies that g has only a finite number of roots.

The estimator

Let K_τ and L_ρ be kernel functions.

1. Estimate $g(.)$ and $g'(.)$ using local linear m-regression:

$$\left(\widehat{g}(x), \widehat{g'}(x)\right) = \mathrm{argmin}_{a,b} \sum_i K_\tau(X_i - x) m(Y_i - a - b(X_i - x)).$$

2. Estimate $Z(g)$ by $\widehat{Z} = Z_\rho\left(\widehat{g}(.), \widehat{g'}(.)\right)$, where Z_ρ is defined as

$$Z_\rho\left(g(.), g'(.)\right) := \int_{\mathscr{X}} L_\rho(g(x))|g'(x)|dx.$$

K_τ and L_ρ are assumed to be Lipschitz continuous, positive symmetric kernel functions integrating to 1 with bandwidth τ and ρ and support $[-\tau, \tau]$ and $[-\rho, \rho]$.

Inference

3. Estimate the variance and bias of $\widehat{Z}$ relative to Z using bootstrap.
4. Construct integer valued confidence sets for Z using t-statistics based on $\widehat{Z}$ and the bootstrapped variance and bias.

Justification and asymptotic properties

Proposition

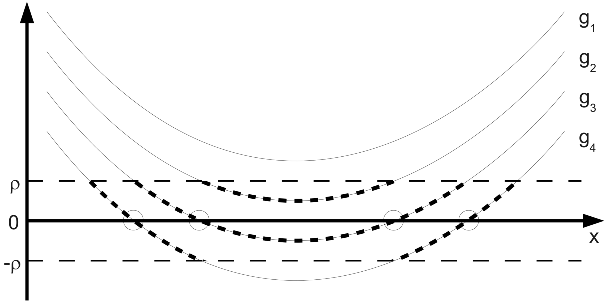
For g continuously differentiable and generic, if $\rho > 0$ is small enough, then

$$Z_\rho(g(.), g'(.)) = Z(g(.)).$$

Idea of proof:

- Consider the subset of $\mathscr{X}$ where $L_\rho(g) \neq 0$, i.e., $g(x) < \rho$.
- If ρ is small enough, this subset is partitioned into disjoint neighborhoods of the roots of g, and g is monotonic in each of these neighborhoods.
- A change of variables, setting $y = g(x)$, shows that the integral over each of these neighborhoods equals one.

Figure: Z AND Z_ρ



Notes: $Z(g_1) = Z_\rho(g_1) = 0$,
$Z(g_2) = 0 < Z_\rho(g_2) < 1$
$Z(g_3) = 2 > Z_\rho(g_3) > 1$,
and $Z(g_4) = Z_\rho(g_4) = 2$.

Local constancy of Z and Z_ρ

Definition ($\mathscr{C}^1$ norm)

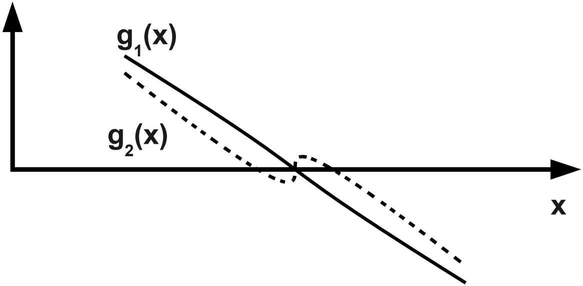
Let $\mathscr{C}^1(\mathscr{X})$ denote the space of continuously differentiable functions on the compact domain $\mathscr{X}$. The norm $||.||$ on $\mathscr{C}^1(\mathscr{X})$ is defined by

$$||g|| := \sup_{x\in\mathscr{X}} |g(x)| + \sup_{x\in\mathscr{X}} |g'(x)|.$$

Proposition (Local constancy)

$Z(.)$ is constant in a neighborhood, with respect to the norm $||.||$, of any generic function $g \in C^1$, and so is Z_ρ if ρ is small enough.

Figure: On the importance of wiggles



Corollary (Superconsistency)

If $\left(\widehat{g}, \widehat{g'}\right)$ *converges uniformly in probability to* (g, g')*, if* g *is generic and if* $\alpha_n \to \infty$ *is some arbitrary diverging sequence, then*

$$\alpha_n(Z(\widehat{g}) - Z(g)) \to^p 0.$$

Furthermore, if ρ *is small enough so that* $Z_\rho(g, g') = Z(g)$ *holds, then*

$$\alpha_n(Z_\rho\left(\widehat{g}, \widehat{g'}\right) - Z(g)) \to^p 0.$$

Asymptotics for the first stage

Assumption (Bahadur expansion)

$$\left(\widehat{g}(x), \widehat{g'}(x)\right) - (g(x), g'(x)) = R -$$

$$-f_x^{-1}(x)s^{-1}(x)I_n(x)\frac{1}{n}\sum_i K_\tau(X_i - x)\phi(Y_i - g(x) - g'(x)(X_i - x))\left(\frac{1}{\tau}, \frac{X_i - x}{\nu_2\tau^3}\right) \quad (4)$$

where

- f_x *is the density of* x, $\nu_2 := \int K(x)x^2 dx$
- $\phi := m'$ *(in a piecewise derivative sense),* $s(x) = \frac{\partial}{\partial g(x)}E[\phi(Y - g(x))|X = x]$
- $I_n(x)$ *is a non-random matrix converging uniformly to the identity matrix*
- $R = o_p\left(\left(\widehat{g}(x), \widehat{g'}(x)\right) - (g(x), g'(x))\right)$ *uniformly in* x

Kong, Linton, and Xia (2010) provide regularity conditions under which $R = \left(1, \frac{1}{\tau}\right) O_p\left(\frac{log(n)}{n\tau}\right)^\lambda$ uniformly in x, for some $\lambda \in (0,1)$ as $n \to \infty$.

Non-standard sequence of experiments

- In order to get nondegenerate asymptotics of $\widehat{Z}$
- we need a nondegenerate limit of $\widehat{g'}$
- which requires sequences of experiments with increasing amounts of "noise" relative to "signal."

Assume:

- for the n^{th} experiment we observe $(Y_{i,n}, X_{i,n})$ for $i = 1, \ldots, n$
- $$\begin{aligned} X_{i,n} &\sim^{iid} f_x(.) && (5) \\ \gamma_{i,n}|X_{i,n} &\sim f_{\gamma|X} && (6) \\ Y_{i,n} &= g(X_{i,n}) + r_n \gamma_{i,n}, && (7) \end{aligned}$$
- where $\{r_n\}$ is a real-valued sequence
- and $0 = \mathrm{argmin}_a E[m(\gamma - a)|X] = \mathrm{argmin}_a E[m(r_n \gamma - a)|X]$.

The central result

Theorem (Asymptotic normality)

Under the above model assumptions, if $r_n = (n\tau^5)^{1/2}$, $n\tau \to \infty$, $\rho \to 0$ *and* $\tau/\rho^2 \to 0$, *then there exist* $\mu > 0$ *and* V *such that*

$$\sqrt{\frac{\rho}{\tau}}\left(\widehat{Z} - \mu - Z\right) \to N(0, V)$$

for $\widehat{Z} = Z_\rho\left(\widehat{g}, \widehat{g'}\right)$. *Both* μ *and* V *depend on the data generating process only via the asymptotic mean and variance of* $\widehat{g'}$ *at the roots of* g, *which in turn depend upon* f_X, g', s *and* $Var(\phi|X)$ *evaluated at the roots of* g.

Outline of proof:

- Ignore variation in $\widehat{g}$, since variation in $\widehat{g'}$ asymptotically dominates.
- Apply various Taylor approximations.
- Partition the range of integration into intervals of length 2τ.
- Using a Poisson approximation, show that the number of X_i falling into each range is approximately independent.
- Deduce that the integrals over each range of length 2τ are approximately independent except for immediate neighbors.
- Using distributional convergence of $\widehat{g'}$, show distributional convergence to nondegenerate limit of the integral over a range 2τ.
- Apply a CLT for m-dependent variables, writing the integral as sum of sub-integrals.

Bootstrap inference

- $\widehat{Z} - Z$ converges to a normal distribution, after bias correction and rescaling.
- We could perform a t-test on hypotheses about Z, if we had
  1. a consistent estimator of V
  2. an estimator of μ converging at a rate faster than $\sqrt{\rho/\tau}$.
- Bootstrap provides such estimators -
- if the sample size grows fast enough relative to $\sqrt{\rho/\tau}$ and τ.

Relative efficiency of Z_ρ

- Increasing ρ reduces the variance without affecting the bias in the limit.
- The reason: Asymptotically the difficulty in estimating Z driven entirely by fluctuations in $\widehat{g'}$. Larger ρ averages out these fluctuations over a larger range of X.
- $\Rightarrow Z_{\rho_1}$ is asymptotically inefficient relative to Z_{ρ_2} for $\rho_1 < \rho_2$.
- $Z(g) = \lim_{\rho \to 0} Z_\rho(g)$
- This suggests that the simple plug-in estimator $Z(\widehat{g})$ is asymptotically inefficient relative to $\widehat{Z}$.
- This is only a heuristic argument, however: We can not exchange the limits with respect to ρ and with respect to n to obtain the limit distribution of $Z(\widehat{g})$.

Monte Carlo evidence

- Data generated by

$$\begin{aligned} X_i &\sim^{iid} \mathrm{Uni}[0,1] \\ \gamma_i | X_i &\sim f_{\gamma|X} \\ Y_i &= g^j(X_i) + \gamma_i \end{aligned} \tag{8}$$

- $f_{\gamma|X}$ is an appropriately centered and scaled uniform or normal distribution.
- Two functions g^j:
  1. with one root: $g^1(x) = 0.5 - x$
  2. with three roots: $g^2(x) = 0.5 - 5x + 12x^2 - 8x^3$
- g is estimated by
  1. median regression
  2. mean regression
  3. 0.9 quantile regression

Figure: Density of $\widehat{Z}$ in Monte Carlo experiments

uniform errors, median regression
$g^1(x) = 0.5 - x$, $Z(g^1) = 1$

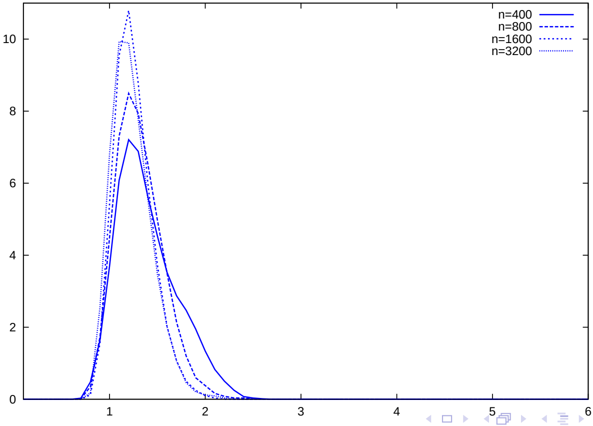


Figure: Density of $\widehat{Z}$ in Monte Carlo experiments

uniform errors, median regression

$$g^2(x) = 0.5 - 5x + 12x^2 - 8x^3,\ Z(g^2) = 3$$

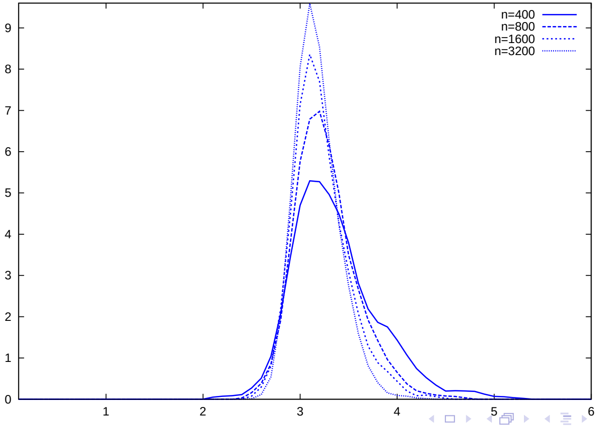


Table: MONTECARLO REJECTION PROBABILITIES UNDER A TEST OF ASYMPTOTIC LEVEL α =5%.

n	τ	**r**	$\widehat{P}(\zeta > z_\alpha)$	$\widehat{P}(\zeta < -z_\alpha)$
400	0.065	0.179	0.05	0.01
800	0.059	0.194	0.03	0.02
1600	0.055	0.231	0.02	0.01
3200	0.052	0.290	0.02	0.01
400	0.065	0.268	0.03	0.02
800	0.059	0.292	0.01	0.02
1600	0.055	0.347	0.01	0.01
3200	0.052	0.434	0.01	0.02

Notes: The columns show in turn

- sample size, regression bandwidth, error standard deviation,
- and the rejection probabilities of one-sided tests.

The g are estimated by mean regression, the errors are uniformly distributed, and the first 4 experiments are generated using g^1, the next 4 using g^2.

Generalizations

So far: no controls, one-dimensional X, inference on the total number of roots.

We will now extend this to:

1. estimation controlling for covariates
2. higher dimensional systems
3. inference on the number of stable and unstable equilibria

Conditioning on covariates

Consider functions g identified by

$$g(x, w_1) = \operatorname{argmin}_y E_{W_2}\left[E_{Y|X,W}[m(Y-y)|X=x, W_1=w_1, W_2]\right]. \quad (9)$$

The parameter of interest is $Z(g(., w_1))$.

Condition 9 can be rationalized by a structural model of the form:

- $Y = h(X, W_1, \epsilon)$
- $\epsilon \perp (X, W_1)|W_2$ ("selection on observables")
- $g(x, w_1) := \operatorname{argmin}_y E_\epsilon [m(h(x, w_1, \epsilon) - y)]]$

W^2 serves as vector of controls variables.
If $m(\delta) = \delta^2$, equation 9 identifies the average structural function.
This will be important in the discussion of games of incomplete information.

Estimate g by

$$\left(\widehat{g}(x, w_1), \widehat{g'}(x, w_1)\right) = \text{argmin}_{a,b} M(a, b, x, w_1),$$

where $M(a, b, x, w_1) =$

$$\frac{1}{n} \sum_j \frac{\sum_i K_\tau(X_i - x, W_{1i} - w_1, W_{2i} - W_{2j}) m(Y_i - a - b(X_i - x))}{\sum_i K_\tau(X_i - x, W_{1i} - w_1, W_{2i} - W_{2j})}. \quad (10)$$

The paper states an asymptotic normality result for this context, generalizing the previous one. Crucial steps of the proof:

1. To obtain a sequence of experiments, such that $\widehat{g}$ converges uniformly to g while $\widehat{g'}$ has a non degenerate limiting distribution.
2. To obtain an approximation of $\widehat{g'}$ equivalent to equation 4.

This can be done using results on partial means from Newey (1994).

Higher dimensional systems

Suppose g is a function from $\mathbb{R}^d$ to $\mathbb{R}^d$.

We can define $\widehat{Z}$ as

$$\widehat{Z} := \int L_\rho(\widehat{g}) |\det \widehat{g'}|. \tag{11}$$

The paper states an asymptotic normality result for this context, generalizing the previous one, with slower rates of convergence.

Stable and unstable roots

Consider the number of “stable” and “unstable” roots, Z^s and Z^u:

Definition

$$Z^s(g) := |\{x \in \mathscr{X} : g(x) = 0 \textit{ and } g'(x) < 0\}|$$

and

$$Z^u(g) := |\{x \in \mathscr{X} : g(x) = 0 \textit{ and } g'(x) > 0\}|.$$

We can define smooth approximations of these parameters as follows:

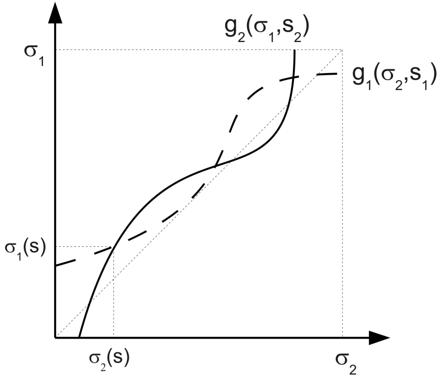
$$\begin{aligned} Z^s_\rho(g(.), g'(.)) &:= \int_{\mathscr{X}} L_\rho(g(x))|g'(x)|\mathbf{1}(g'(x) < 0)dx \\ Z^u_\rho(g(.), g'(.)) &:= \int_{\mathscr{X}} L_\rho(g(x))|g'(x)|\mathbf{1}(g'(x) > 0)du \end{aligned}$$

Static games of incomplete information

Assume:

- Two players i, two actions a, observations indexed by j.
- We observe an i.i.d. sample of $(a_{1,j}, a_{2,j}, s_j)$, the players' realized action and the public information of the game.
- $a_{i,j} \in \{0,1\}$ for $i = 1,2$ and $s \in \mathbb{R}^k$.
- $\sigma_i(s) := E[a_i|s]$: rational expectation beliefs of player $-i$ about the expected action of player i.
- $g_i(\sigma_{-i}, s) := E[a_i|\sigma_{-i}, s]$: average response of player i (averaging over private information of i).
- Bayesian Nash Equilibrium: $\sigma_i(s) = g_i(\sigma_{-i}(s), s)$ for $i = 1,2$.

Figure: Response functions and Bayesian Nash Equilibria



If one component of s is excluded from g_i, we can identify $g_i(\sigma_{-i}, s_i) = E[a_i|\sigma_{-i}, s_i]$, since there is independent variation of σ_{-i} and s_i.

Bayesian Nash Equilibria are solutions to

$$g(\sigma_1, s) = g_1(g_2(\sigma_1, s), s) - \sigma_1.$$

The following procedure is a nonparametric variant of the procedure proposed by Bajari, Hong, Krainer, and Nekipelov (2006).

1. Estimate beliefs by local linear mean regression:

$$(\widehat{\sigma}(s), \widehat{\sigma}'(s)) = \mathrm{argmin}_{b,c} \sum_j K_\tau(s_j - s)\left(a_{i,j} - b - c(s_j - s)\right)^2$$

2. Estimate average response functions by local linear mean regression, again:

$$\left(\widehat{g}_i(\bar{\sigma}_{-i}, s_i), \widehat{g'}_i(\bar{\sigma}_{-i}, s_i)\right) =$$

$$\operatorname{argmin}_{b,c} \sum_j K_\tau(\widehat{\sigma}_{-i,j} - \bar{\sigma}_{-i}, s_{i,j} - s_i)\left(a_{i,j} - b - c\left(\widehat{\sigma}_{-i,j} - \bar{\sigma}_{-i}, s_{i,j} - s_i\right)\right)^2$$

3. Plugging $\widehat{g}_2$ into $\widehat{g}_1$, both estimated by 2, yields the following estimator of g,

$$\widehat{g}(\bar{\sigma}_1, s) = \widehat{E}\left[a_1 \middle| \widehat{\sigma}_2 = \widehat{E}\left[a_2 | \widehat{\sigma}_1 = \bar{\sigma}_1, s_2\right], s_1\right] - \bar{\sigma}_1.$$

4. Perform inference on the number of Bayesian Nash Equilibria given s, $Z(g(.,s))$, using

$$\widehat{Z} = Z_\rho\left((\widehat{g}(.,s), \widehat{g'^1}(.,s)\right).$$

We can again show asymptotic normality for an appropriate sequence of experiments. The following sequence works. It shrinks response functions to the diagonal.

$$
\begin{aligned}
s_{j,n} &\sim^{iid} f_s(.) && (12)\\
a_{i,j,n}|s_{j,n} &\sim \mathrm{Bin}(\sigma_{i,n}(s_{j,n})) && (13)\\
\sigma_{i,n}(s) &= g_{i,n}(\sigma_{-i,n}(s), s_i) && (14)\\
g_{1,n}(\sigma_2, s_1) &= \frac{1}{r_n} g_{1,0}(\sigma_2, s_1) + \left(1 - \frac{1}{r_n}\right)\sigma_2 && (15)\\
g_{2,n}^{-1}(\sigma_2, s_2) &= \frac{1}{r_n} g_{2,0}^{-1}(\sigma_2, s_2) + \left(1 - \frac{1}{r_n}\right)\sigma_2 && (16)
\end{aligned}
$$

This setup implies

$$
r_n g_n(\sigma_1, s) \to g_{1,0}(\sigma_1, s_1) - g_{2,0}^{-1}(\sigma_1, s_2).
$$

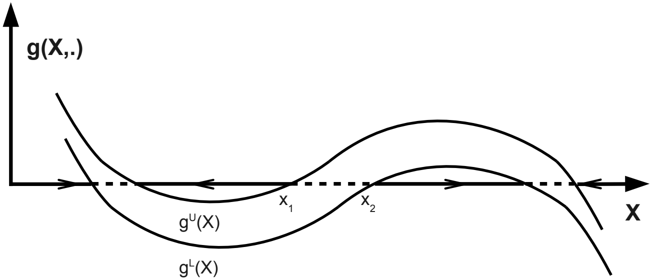
Stochastic difference equations

$$\Delta X_{i,t+1} = X_{i,t+1} - X_{i,t} = g(X_{i,t}, \epsilon_{i,t}) \tag{17}$$

We can show:

1. The number of roots of g allows to characterize the qualitative dynamics of the stochastic difference equation in terms of equilibrium regions.
2. If we find only one root in cross-sectional quantile regressions of ΔX on X, this implies that there is only one stable root for a family of conditional average structural functions.

Figure: Qualitative dynamics of stochastic difference equations



If there are unstable equilibria structurally, then quantile regressions should exhibit multiple roots.

Assume:

- $\Delta X = g(X, \epsilon)$.
- First order stochastic dominance: $\mathbb{P}(g(x', \epsilon) \leq Q|X)$ is non-increasing as a function of X, holding x' constant.
- Global stability: $g(\inf \mathscr{X}, \epsilon) > 0$, $g(\sup \mathscr{X}, \epsilon) < 0$ for all ϵ.

Then:

Proposition (Unstable equilibria in dynamics and quantile regressions)

If $Q_{\Delta X|X}(q|X)$ has only one root X for all q, then

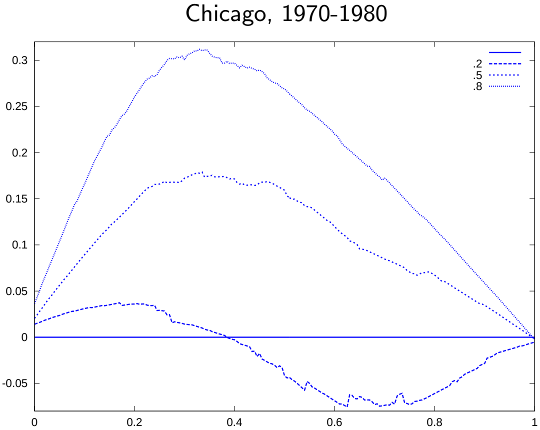
$$E\left[\frac{\partial}{\partial X} g(X, \epsilon) \middle| \Delta X = 0, X\right] \leq 0$$

for all X, where $(0, X)$ is in the support of $(\Delta X, X)$.

Application to data on neighborhood composition

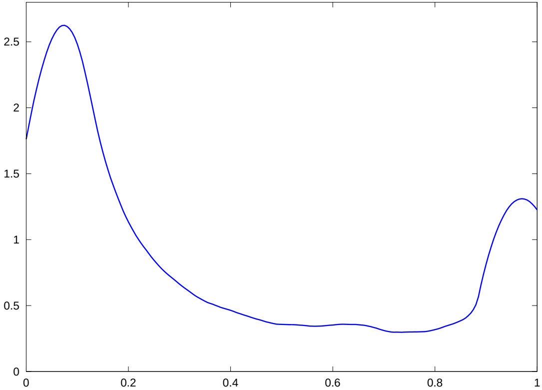
- Data: Those used for analysis of neighborhood composition dynamics by Card, Mas, and Rothstein (2008)
- Neighborhood Change Database (NCDB): aggregates US census variables to the level of census tracts, matching observations from the same geographic area over time.
- We will study the dynamics of minority share in a neighborhood.
- This paper
  1. runs local linear quantile regressions
  2. of the change in minority share on initial minority share in a neighborhood
  3. separately for each MSA and decade
  4. and performs inference on the number of roots.

This figure shows local linear quantile regressions
of the change in neighborhood minority share 1970-1980
on initial minority share 1970
for the .2, .5 and .8th conditional quantile.

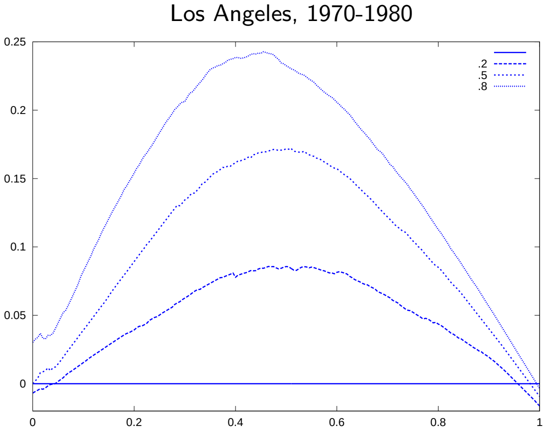


This figure shows the density
of the distribution of minority share across neighborhoods
(not weighted by neighborhood size)

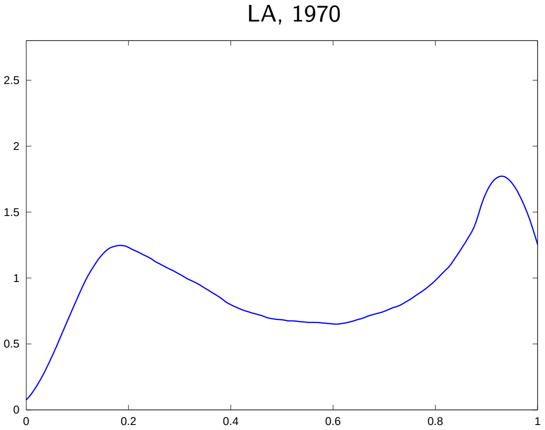
Chicago, 1970



This figure shows, again, local linear quantile regressions of the change in neighborhood minority share 1970-1980 on initial minority share 1970
for the .2, .5 and $.8^{th}$ conditional quantile.



This figure shows the density
of the distribution of minority share across neighborhoods
(not weighted by neighborhood size)



Selected results

Table: .95 CONFIDENCE SETS FOR $Z(g)$ FOR SELECTED MSAS BY QUANTILE, CHANGE IN MINORITY SHARE

MSA	70s		
	$q = .2$	$q = .5$	$q = .8$
Atlanta, GA MSA	[1,1]	[1,1]	[0,0]
Boston, MA-NH PMSA	[0,1]	[0,1]	[0,1]
Chicago, IL PMSA	[0,1]	[0,1]	[0,1]
Detroit, MI PMSA	[1,2]	[0,1]	[0,1]
Los Angeles-Long Beach, CA PMSA	[1,1]	[1,1]	[0,1]
New York, NY PMSA	[0,1]	[0,1]	[0,0]
San Francisco, CA PMSA	[1,1]	[0,1]	[0,1]

Selected results

Table: .95 confidence sets for $Z(g)$ for selected MSAs by quantile, change in minority share

MSA	80s		
	$q = .2$	$q = .5$	$q = .8$
Atlanta, GA MSA	[2,3]	[0,0]	[0,0]
Boston, MA-NH PMSA	[0,1]	[0,1]	[0,0]
Chicago, IL PMSA	[2,2]	[0,1]	[0,1]
Detroit, MI PMSA	[0,1]	[0,1]	[0,1]
Los Angeles-Long Beach, CA PMSA	[0,1]	[0,1]	[0,1]
New York, NY PMSA	[0,0]	[0,0]	[0,0]
San Francisco, CA PMSA	[0,0]	[0,1]	[0,0]

Discussion

- There is not much evidence of Z exceeding 1.
- The data indicate a general rise in minority shares that is largest for neighborhoods with intermediate initial share, but not to the extent of leading to tipping behavior.
- Proposition 3: No multiple roots in quantile regressions $\Rightarrow$ no multiple equilibria in the underlying structural relationship.
- I would conclude that tipping is not a widespread phenomenon in US ethnic neighborhood composition (in contrast to Card, Mas, and Rothstein (2008)).

Cautionary remarks

Potential causes of bias in the estimated number of equilibria:

- Range of integration:
  - If a root lies right on the boundary of the chosen range of integration, it enters Z_ρ as $1/2$ only.
  - Extending the range of integration beyond the unit interval might lead to an upward bias, if extrapolated regression functions intersect with the horizontal axis.
  - There might be roots of g in the unit interval but beyond the support of the data (not an issue here).
- If the bandwidth parameter ρ is chosen too large, this might bias the estimated number of equilibria downwards.

Summary and conclusion

- The paper proposes an inference procedure for the number of roots of functions nonparametrically identified using conditional moment restrictions
- based on a smoothed plug-in estimator of the number of roots which is super-consistent under i.i.d. asymptotics,
- but asymptotically normal under non-standard asymptotics,
- and asymptotically efficient relative to the "naive" plug-in estimator.

- The results are extended to cover:
  - covariates as controls,
  - higher dimensional domain and range,
  - and inference on the number of equilibria with various stability properties.
- Two theoretical applications were discussed:
  - static games of incomplete information
  - stochastic difference equations
- In an empirical application to neighborhood composition dynamics in the United States, no evidence of multiplicity of equilibria is found.

Potential further applications

- household level poverty traps
- intergenerational mobility
- efficiency wages
- macro models of economic growth
- financial market bubbles (herding)
- market entry
- social norms

The Matlab/Octave code written for this paper is available upon request.

Thanks for your time!