

DISCUSSION:
PAYMENT CRISES AND CONSEQUENCES
BY: KOCH, CHEN, RICHARDSON, SHARMA

Gabriel Chodorow-Reich
Harvard University and NBER

NBER Monetary Economics
November 13, 2020

PLAN

- ① What kind of crisis is this?
- ② Synthetic controls and inference.
- ③ Anything else special about Rhode Island?
- ④ Conclusion.

RHODE ISLAND TIMELINE

Date	Frozen deposit share (%)	
	Of RISDIC	Of total
January 1, 1991	100	18
January 5, 1991	81	15
May 1991	74	13
June 1991	51	9
May 1992	31	6
June 1992	5	1
1993	0	0

- Maryland, Nebraska, Ohio: deposits available (possibly with limit) within a few weeks.

SIMILAR TO...

- Different from consequences of letting major banks fail in 2008:
 - ▶ Limited chain reaction through counterparty risk.
 - ▶ No concentrated wealth losses, only liquidity.
 - ▶ Bank *lending* channel offset by new entry, although firm-bank relationships could still matter.
 - ▶ Corporations not directly impacted (or did they have frozen deposits?).
- More similar to 2016 demonetization in India:
 - ▶ Households temporarily lost access to liquid wealth (cash).
 - ▶ Short-run effects on output (Chodorow-Reich, Gopinath, Mishra, Narayanan, 2020).
- Or to hypothetical cyber attack that freezes banking system.

SYNTHETIC CONTROLS OVERVIEW

- Assume data generating process (eq. (2) in the paper):

$$y_{1,t}^N = \mathbf{w}' \mathbf{y}_{\mathbf{0},t}^N + \varepsilon_{1,t},$$

$$y_{1,t}^I = y_{1,t}^N + \beta_t.$$

- ▶ $y_{1,t}^N$: potential outcome of unit 1 if no treatment.
- ▶ $\mathbf{w}$: vector of non-negative weights summing to 1.
- ▶ $\mathbf{y}_{\mathbf{0},t}^N$: vector of outcomes of donor units.
- ▶ $\varepsilon_{1,t} \sim N(0, \sigma^2)$: idiosyncratic shock to y_1 .
- ▶ $y_{1,t}^I$: potential outcome of unit 1 if intervention.
- ▶ β_t : treatment effect.
- ▶ Observe $\mathbf{y}_{\mathbf{0},t}^N \forall t$, $y_{1,t}^N$ for $t = 0, 1 \dots, T_0 - 1$, $y_{1,t}^I$ for $t = T_0, \dots$

SYNTHETIC CONTROLS OVERVIEW

- Assume data generating process (eq. (2) in the paper):

$$y_{1,t}^N = \mathbf{w}' \mathbf{y}_{0,t}^N + \varepsilon_{1,t},$$

$$y_{1,t}^I = y_{1,t}^N + \beta_t.$$

- ▶ $y_{1,t}^N$: potential outcome of unit 1 if no treatment.
 - ▶ $\mathbf{w}$: vector of non-negative weights summing to 1.
 - ▶ $\mathbf{y}_{0,t}^N$: vector of outcomes of donor units.
 - ▶ $\varepsilon_{1,t} \sim N(0, \sigma^2)$: idiosyncratic shock to y_1 .
 - ▶ $y_{1,t}^I$: potential outcome of unit 1 if intervention.
 - ▶ β_t : treatment effect.
 - ▶ Observe $\mathbf{y}_{0,t}^N \forall t$, $y_{1,t}^N$ for $t = 0, 1, \dots, T_0 - 1$, $y_{1,t}^I$ for $t = T_0, \dots$
- Ordinary approach: $\hat{\mathbf{w}} = \arg \min \sum_{t=1}^{T_0-1} \left(y_{1,t} - \hat{\mathbf{w}}' \mathbf{y}_{0,t}^N \right)^2$. Under DGP, consistent estimator of $\mathbf{w}$ as length of pre-intervention period $\rightarrow \infty$.

SYNTHETIC CONTROLS OVERVIEW

- Assume data generating process (eq. (2) in the paper):

$$y_{1,t}^N = \mathbf{w}' \mathbf{y}_{0,t}^N + \varepsilon_{1,t},$$

$$y_{1,t}^I = y_{1,t}^N + \beta_t.$$

- ▶ $y_{1,t}^N$: potential outcome of unit 1 if no treatment.
 - ▶ $\mathbf{w}$: vector of non-negative weights summing to 1.
 - ▶ $\mathbf{y}_{0,t}^N$: vector of outcomes of donor units.
 - ▶ $\varepsilon_{1,t} \sim N(0, \sigma^2)$: idiosyncratic shock to y_1 .
 - ▶ $y_{1,t}^I$: potential outcome of unit 1 if intervention.
 - ▶ β_t : treatment effect.
 - ▶ Observe $\mathbf{y}_{0,t}^N \forall t$, $y_{1,t}^N$ for $t = 0, 1, \dots, T_0 - 1$, $y_{1,t}^I$ for $t = T_0, \dots$
- Ordinary approach: $\hat{\mathbf{w}} = \arg \min \sum_{t=1}^{T_0-1} \left(y_{1,t} - \hat{\mathbf{w}}' \mathbf{y}_{0,t}^N \right)^2$. Under DGP, consistent estimator of $\mathbf{w}$ as length of pre-intervention period $\rightarrow \infty$.
 - If $\sigma^2 \rightarrow 0$, $y_{1,t}^N$ in convex hull of $\mathbf{y}_{0,t}^N$ and point identify β_t .

SYNTHETIC CONTROLS OVERVIEW

- Assume data generating process (eq. (2) in the paper):

$$y_{1,t}^N = \mathbf{w}' \mathbf{y}_{0,t}^N + \varepsilon_{1,t},$$

$$y_{1,t}^I = y_{1,t}^N + \beta_t.$$

- $y_{1,t}^N$: potential outcome of unit 1 if no treatment.
 - $\mathbf{w}$: vector of non-negative weights summing to 1.
 - $\mathbf{y}_{0,t}^N$: vector of outcomes of donor units.
 - $\varepsilon_{1,t} \sim N(0, \sigma^2)$: idiosyncratic shock to y_1 .
 - $y_{1,t}^I$: potential outcome of unit 1 if intervention.
 - β_t : treatment effect.
 - Observe $\mathbf{y}_{0,t}^N \forall t$, $y_{1,t}^N$ for $t = 0, 1, \dots, T_0 - 1$, $y_{1,t}^I$ for $t = T_0, \dots$
- Ordinary approach: $\hat{\mathbf{w}} = \arg \min \sum_{t=1}^{T_0-1} \left(y_{1,t} - \hat{\mathbf{w}}' \mathbf{y}_{0,t}^N \right)^2$. Under DGP, consistent estimator of $\mathbf{w}$ as length of pre-intervention period $\rightarrow \infty$.
 - If $\sigma^2 \rightarrow 0$, $y_{1,t}^N$ in convex hull of $\mathbf{y}_{0,t}^N$ and point identify β_t .
 - Otherwise, only identify β_t up to unknown $\varepsilon_{1,t}$.

INFERENCE

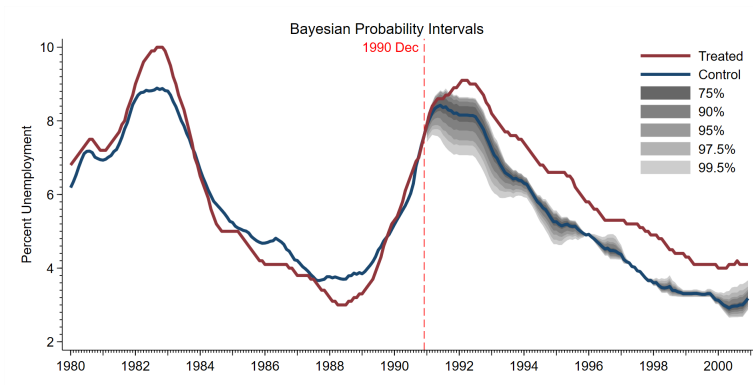
- Inference non-standard because β not point-identified if $\sigma^2 \neq 0$.
- Abadie and Gardeazabal (2003), Abadie, Diamond and Hainmueller (2010): permutation test of placebo treatment of each control unit.
 - ▶ Apply synthetic control to each potential control unit.
 - ▶ Intuition: if σ^2 small, post-treatment evolution of control units is similar and placebo treatment effects will all be close to zero.
 - ▶ Exact test of null of no treatment effect (or any null), but not confidence set.
- KCRS Bayesian approach: draw repeatedly from posterior using MCMC and report coverage ratio quantiles.
 - ▶ KCRS intuition: if post treatment evolution of control units is similar, doesn't matter which units are used to form synthetic control.

BAYESIAN APPROACH (ROUGHLY)

- ① Draw candidate $\hat{\mathbf{w}}^i, \hat{\sigma}^i$ in i th draw.
- ② Compute $\hat{\boldsymbol{\varepsilon}}_{1,t < T_0}^i = \mathbf{y}_{1,t < T_0} - \mathbf{Y}_{0,t < T_0} \hat{\mathbf{w}}^i$.
- ③ Evaluate likelihood $\hat{\mathcal{L}}^i$ using $\hat{\boldsymbol{\varepsilon}}_{1,t}^i \sim N(0, \hat{\sigma}^i)$.
- ④ Accept jump to $\hat{\mathbf{w}}^i, \hat{\sigma}^i$ with probability $\alpha = \min\{1, \hat{\mathcal{L}}^i / \mathcal{L}^{i-1}\}$.
- ⑤ Draw $\mathbf{w}^{i+1}, \sigma^{i+1}$ from t distribution centered around $\mathbf{w}^i, \sigma^i$.
- ⑥ Obtain mean, quantiles of $\mathbf{w}' \mathbf{y}_{0,t}^N$ from MCMC chain.

EXAMPLE: UNEMPLOYMENT RATE

(a) Counterfactuals and Probability Intervals



INFERENCE REVISITED

- What happens if likelihood of a few units dominates the rest?
- Common for $\mathbf{w}$ to be sparse.
- Limiting case: there exists $j : y_{1,t}^N = y_{j,t}^N + \varepsilon_{1,t}$.
- How? Suppose units 1 and j share common shock or factor structure:

$$y_{1,g,t}^N = \alpha_{g_t} + \varepsilon_{1,t},$$

$$y_{j,g,t}^N = \alpha_{g_t} + \varepsilon_{j,t}.$$

- If $\text{Var}(\alpha_{g_t}) \gg \sigma^2$, posterior should converge to $w_j = 1, w_i = 0 \ \forall i \neq j$.

INFERENCE REVISITED

- What happens if likelihood of a few units dominates the rest?
- Common for $\mathbf{w}$ to be sparse.
- Limiting case: there exists $j : y_{1,t}^N = y_{j,t}^N + \varepsilon_{1,t}$.
- How? Suppose units 1 and j share common shock or factor structure:

$$y_{1,g,t}^N = \alpha_{g_t} + \varepsilon_{1,t},$$

$$y_{j,g,t}^N = \alpha_{g_t} + \varepsilon_{j,t}.$$

- If $\text{Var}(\alpha_{g_t}) \gg \sigma^2$, posterior should converge to $w_j = 1, w_i = 0 \ \forall i \neq j$.
- Coverage of $y_{1,t}^N$ degenerate despite arbitrary uncertainty about β (from unknown realizations of $\varepsilon_{1,t}$).

INFERENCE REVISITED

- What happens if likelihood of a few units dominates the rest?
- Common for $\mathbf{w}$ to be sparse.
- Limiting case: there exists $j : y_{1,t}^N = y_{j,t}^N + \varepsilon_{1,t}$.
- How? Suppose units 1 and j share common shock or factor structure:

$$y_{1,g,t}^N = \alpha_{g_t} + \varepsilon_{1,t},$$

$$y_{j,g,t}^N = \alpha_{g_t} + \varepsilon_{j,t}.$$

- If $\text{Var}(\alpha_{g_t}) \gg \sigma^2$, posterior should converge to $w_j = 1, w_i = 0 \ \forall i \neq j$.
- Coverage of $y_{1,t}^N$ degenerate despite arbitrary uncertainty about β (from unknown realizations of $\varepsilon_{1,t}$).
- Inference should augment the coverage plots with information from the posterior of σ^2 (I think).

INFERENCE REVISITED

- What happens if likelihood of a few units dominates the rest?
- Common for $\mathbf{w}$ to be sparse.
- Limiting case: there exists $j : y_{1,t}^N = y_{j,t}^N + \varepsilon_{1,t}$.
- How? Suppose units 1 and j share common shock or factor structure:

$$y_{1,g,t}^N = \alpha_{g_t} + \varepsilon_{1,t},$$

$$y_{j,g,t}^N = \alpha_{g_t} + \varepsilon_{j,t}.$$

- If $\text{Var}(\alpha_{g_t}) \gg \sigma^2$, posterior should converge to $w_j = 1, w_i = 0 \ \forall i \neq j$.
- Coverage of $y_{1,t}^N$ degenerate despite arbitrary uncertainty about β (from unknown realizations of $\varepsilon_{1,t}$).
- Inference should augment the coverage plots with information from the posterior of σ^2 (I think).
- In practice $\varepsilon_{1,t}$ probably serially correlated. Does this matter?

MECHANISM

- Were economic effects due to suspension of deposits?
- Or bank run itself?
- Or the factors that made the banking system vulnerable?
- Synthetic controls does not tell us.
- Narrative evidence: RI crisis ramped up in fall 1990 when regulators discovered that 52% of reported assets of Heritage Loan and Investment Company were ghost loans that instead went to bank's president, who also sat on board of directors of RISDIC.

KCRS: In each of the four states, mismanagement or defalcation at a large insured institution resulted in losses that threatened the insurance fund's stability. The revelation of the problem surprised regulators and the public. Unexpected bad news triggered runs at healthy members of the state insurance fund.

WHAT ELSE: GENERAL CORRUPTION

- Heritage Loan collapsed because of corrupt bank president.
- Deposits frozen by new governor hours into his administration. Previous governor later plead guilty to bribery, extortion and racketeering and served a year in prison.
- Mayor of providence resigned in 1984 after pleading *no lo contendre* to assault charge, received suspended 5 year sentence, won new term in 1991, then arrested in 2001 on federal criminal charges of racketeering, conspiracy, extortion, witness tampering, and mail fraud and served 5 years in prison.

WHAT ELSE: BANK UNDERCAPITALIZATION

English (1993): *NCUA examinations in November 1990 showed that Marquette Credit Union had a negative net worth of \$30 million, Davisville Credit Union had a negative net worth of \$18 million, and Rhode Island Central Credit Union (RICCU) had a negative net worth of \$19 million...*

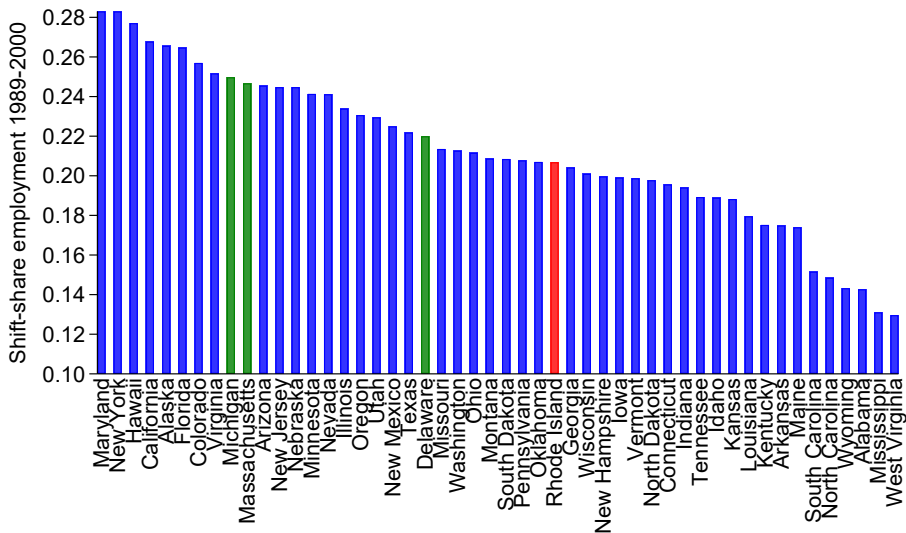
A federal examination of Marquette Credit Union, RISDIC's largest member, in November of 1990 revealed that 97% of Marquette's loans were classified as substandard, doubtful, or loss...

After Heritage's failure in November, runs began on other RISDIC-insured institutions. Having put much of its money into Heritage, RISDIC assessed members to replenish its reserves. In December 1990 a major run developed on RICCU. The credit union exhausted its credit line with the Rhode Island Credit Union League's Corporate Credit Union on December 31, 1990. On that day the RISDIC board decided to ask the state to appoint a conservator for RISDIC...

The state ordered all of the institutions to apply for federal insurance. The smaller members were able to obtain it, but 14 institutions remained closed...

The problems with RISDIC and several of its members were well-known in 1985 because the state attorney general commissioned a report on RISDIC. The report was extremely clear on the problems facing RISDIC and its members, especially three of its largest members. All three had made substantial loans to insiders without proper approval or documentation. The report claimed that all three were likely insolvent at that time...

WHAT ELSE: INDUSTRIAL COMPOSITION



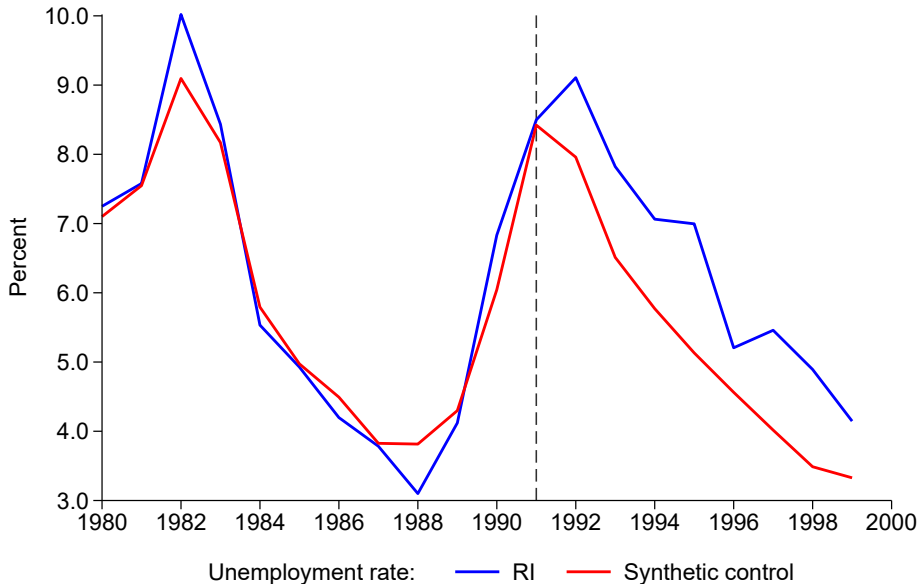
$$y_{i,1989,2000} = \sum_j \left(\frac{\text{Emp}_{i,j,1989}}{\text{Emp}_{i,1989}} \right) \times \left(\frac{\text{Emp}_{-i,j,2000}}{\text{Emp}_{-i,j,1989}} - 1 \right) \text{ using 394 SIC industries.}$$

INDUSTRIAL DIFFERENCES BETWEEN RI AND MA

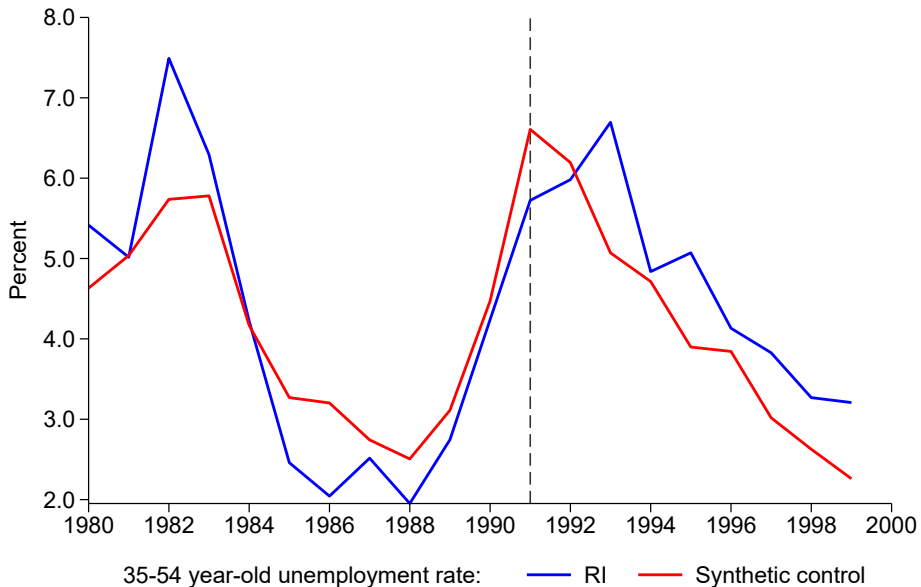
Industry title	Weight (%)		Leave-out growth (%)	Contribution (%)		
	MA	RI		MA	RI	Δ
SIC 396 Costume Jewelry	0.02	3.65	-33.2	-0.0	-1.2	1.2
SIC 737 Data Processing	1.44	0.86	193.5	2.8	1.7	1.1
Total						4.0

- 3.7% of RI employment in 1989 in costume jewelry.
- MA already had head start in computer and data processing.

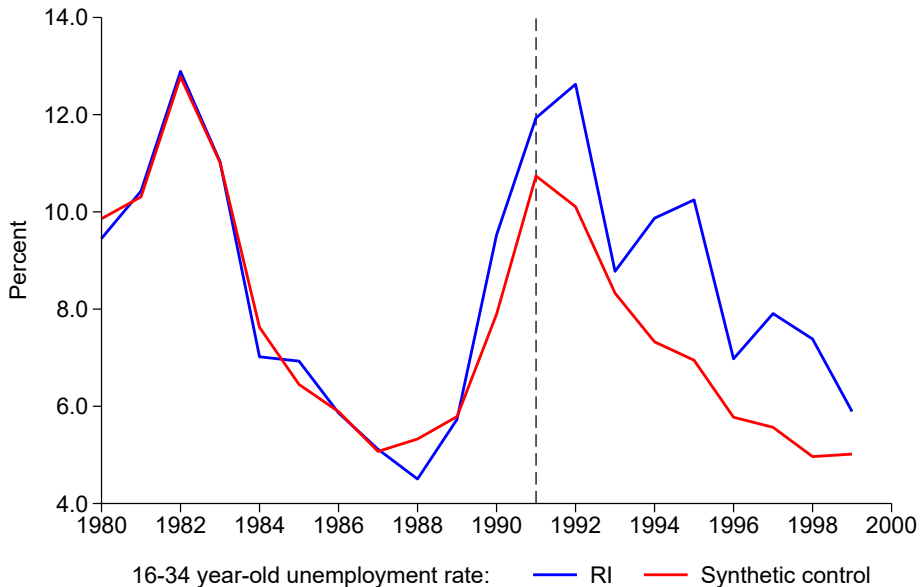
WHAT ELSE: AGE



WHAT ELSE: AGE



WHAT ELSE: AGE



OTHER SUGGESTIONS

- Specification checks carry over from classical implementation:
 - ▶ Select weights in training sample and report pre-treatment coverage.
 - ▶ Exclude neighbors due to SUTVA concerns: what happens in RI affects MA (a little) and what happens in MA affects RI (a lot).
 - ▶ Weights seem to differ a lot across outcomes. Why?
- Report diagnostics from the MCMC:
 - ▶ Algorithm draws 50 parameters (49 state weights + idiosyncratic variance σ^2 of RI) in one block.
 - ▶ Likelihood possibly bimodal: low σ^2 and large weights on closest comparators or high σ^2 and dispersed weights.
 - ▶ Show chain convergence for states with positive weight.
 - ▶ Evaluate coverage sets using Monte Carlo exercise.

CONCLUSION

- Why I like the paper:
 - ▶ Important topic that had been somewhat lost to history.
 - ▶ Econometric contribution.
 - ▶ Clearly written.
 - ▶ Sharp results.
- Main suggestions:
 - ▶ More formal justification of Bayesian procedure (or a companion paper...).
 - ▶ More on mechanism.
 - ▶ More diagnostic tests of synthetic control validity in this setting.
 - ▶ More diagnostic tests of MCMC.