Identifying Agglomeration Spillovers: New Evidence from Large Plant Openings

Carlianne Patrick, Georgia State University

Mark Partridge, The Ohio State University

Motivation

- Luring large industrial facilities is the primary local economic development strategy in the US and the practice is becoming more widespread throughout the developed and developing worlds (Combes et al. 2010; Bartik 2012; Patrick 2014).
- “Winning” the competition for a large industrial facility carries the promise of permanently changing an area’s economic well-being.
- The economic justification for local industrial programs relies critically on the size and nonlinearity of agglomeration externalities as well as multiple equilibria.

This paper . . .

- Uses confidential Census micro data to estimate the effects of incentivized plants openings on plants in the same geographical areas

- Compares the outcomes for plants in a county that “wins” a new plant (as reported by *Site Selection* and *Good Jobs First)* to plants in similar counties that did not to receive the new plant.

This paper . . .

- Uses quasi-experimental research designs to test three hypotheses:

  1) whether the plant opening generates positive externalities for incumbent firms (e.g., estimate the size of the agglomeration externality);

  2) whether residual TFP differs in a non-linear way depending on the density of incumbent firms in the area (e.g., test for nonlinearity of the agglomeration function); and

  3) whether the higher productivity due to a new plant pushes the location into a new and permanent equilibrium, or a transitory shock (e.g., test for multiple equilibria).

- Constitutes the most comprehensive analysis of the effects of large new plant openings to date

Background

- Agglomeration externalities are a form of localized increasing returns to scale arising from market and non-market interactions generating productivity and cost benefits due to proximity.
  - Some argue that economic development incentives are compensation for agglomeration spillovers generated by new firms.
- Most prominent study of agglomeration spillovers generated by highly subsidized, large, new industrial facilities is Greenstone, Hornbeck and Moretti (2010) (GHM)
  - Increase incumbent firm productivity by 12.5% over 5 years
  - Large compared to the typical range of productivity elasticities
  - Patrick (2016) finds smaller aggregate effects of GHM plantsWe start by trying to replicate the GHM findings

Background

- Glaeser and Gottlieb (2008) note that the effect on aggregate economic activity of reallocating economic activity across space depends upon the shape of the agglomeration function.
  - If the agglomeration function is substantially non-linear, then relocating economic activity across space can result in national output gains (or losses).
- Kline and Moretti (2014) formalize Glaeser and Gotlieb's (2008) proposition that aggregate gains rely on non-linear externalities
  - Test the effects of the Tennessee Valley Authority (TVA) program using aggregate county data over a 70 year period.
  - Cannot reject a constant elasticity of agglomeration.

Background

- A large class of theoretical models predicts that there are multiple steady-state distributions of economic activity, or multiple equilibria.
  - The selected steady-state depends upon initial conditions and the history of shocks or agents' expectations (Redding, Sturm, and Wolf 2011).
  - The combination of aggregate increasing returns to scale and multiple equilibria suggests (policy-induced) shocks may drastically change the spatial organization of economic activity.
  - Similarly, the theory underlying "Big Push" development strategies requires a large shock that will push the location beyond some threshold and out of a 'bad' equilibrium into a 'good' equilibrium.

Background

- A small body of literature testing solely for multiple equilibria yields mixed evidence in favor of the hypothesis.
  - The seminal paper by Davis and Weinstein (2002) finds Japanese cities returned to their pre-WWII equilibria as defined by population and manufacturing output shares.
  - Davis and Weinstein (2008) also rejects multiple equilibria in city-industry shares.
  - On the other hand, Bosker et al. (2007) determine that post-WWII, German city-shares are best described by two equilibria.
  - Kline and Moretti (2014) also find evidence that the TVA program investments caused permanent increases in manufacturing activity in Appalachia – albeit at the expense of other locations in the country.

Data: Large new plant openings

- Large new plant openings as the source of shocks (MDPs)

Case Set #	Description	Source	Years
1	Greenstone, Hornbeck, and Moretti (2010) MDPs	*Restricted-access replication programs*	1982-1993
2	All large, new plant opening appearing in *Site Selection* magazine from 1982-1993 (excluding GHM cases that do not appear in the magazine) and large, incentivized plants in the Good Jobs First data from 1988-1993	*Site Selection* magazine; Good Jobs First Subsidy Tracker Database	1982-1993
3	All large, new plant opening appearing in *Site Selection* magazine from 1982-1997 and large, incentivized plants in the Good Jobs First data from 1988-1997	*Site Selection* magazine; Good Jobs First Subsidy Tracker Database	1982-1997
4	Random sample of 500 "new" plants from the firm births with above the 95th percentile employment for new establishments	Census micro-data	1982-1997

Data: Large new plant openings

	MDP Case Set			
	1	2	3	4
MDP Shock as Share of Winner Output	1.23	1.662	0.7422	0.0994
	(2.736)	(5.996)	(3.024)	(0.5602)
MDP Shock as Share of Winner Value Added	1.247	1.657	0.8237	0.1079
	(2.641)	(5.989)	(3.038)	(0.4485)
MDP Shock Employment	2,645	1,110	1,459	288.3
	(5,532)	(2,033)	(2,783)	(1,013)
MDP Shock Payroll	142,800	100,800	213,800	265,700
	(283,700)	(247,300)	(1,684,000)	(4,112,000)
MDP Ratio of Other to Production Payroll	3.040	2.682	2.970	2.195
	(5.692)	(5.002)	(6.814)	(5.219)
MDP Cases	50	100	550	500

Data: Spillover Sample Plants

- "Treated" plants are (continuously-appearing) incumbent plants located in the same county as the large, new plant
- Counterfactual plants from counties that are:
  - Identified as "runners-up" by *Site Selection* magazine (the GHM revealed rankings identification strategy)
  - "Similar losers" determined by geographic proximity (100-250 miles) to the "winner" and matching on:
    - observables (propensity score), and
    - industry locational advantage

Data: Spillover Sample Plants

	Case Set 1			Case set 2		Case set 3		Case set 4	
	Winners	GHM Losers	PScore Losers	Winners	PScore Losers	Winners	Pscore Losers	Winners	PScore Losers
Plants									
(log) Output	10.65	10.61	10.75	10.52	10.62	10.55	10.556	10.51	10.45
	(1.262)	(1.308)	(1.063)	(1.133)	(1.08)	(1.118)	(1.101)	(1.103)	(1.067)
(log) Labor	6.546	6.374	6.475	6.426	6.417	6.332	6.324	6.295	6.254
	(1.156)	(1.138)	(0.9526)	(0.9827)	(0.9656)	(0.9923)	(0.9927)	(1.103)	(0.9926)
Counties									
Incumbent Plants	14.98	22.91	11.79	15.15	11.44	15.08	9.782	19.6	10.19
	(16.14)	(21.88)	(13.62)	(40.56)	(16.11)	(33.07)	(15.25)	(42.3)	(16.8)
Counties	50	80	80	70	100	300	450	300	450
Total Counties	100		100	200		650		600	

Spillover Estimates: Empirical Methodology

Estimate GHM spillover equations:

Model 1

$$\ln\left(Y_{pijt}\right) = \beta_1 \ln\left(L_{pijt}\right) + \beta_2 \ln\left(K^B_{pijt}\right) + \beta_3 \ln\left(K^E_{pijt}\right) + \beta_4 \ln\left(M_{pijt}\right) + \delta 1(Winner)_{pj} + \kappa 1(\tau \geq 0)_{jt} + \theta_1\left[1(Winner)_{pj} \times 1(\tau \geq 0)_{jt}\right] + \alpha_p + \mu_{it} + \lambda_j + \varepsilon_{pijt}$$

Model 2

$$\ln\left(Y_{pijt}\right) = \beta_1 \ln\left(L_{pijt}\right) + \beta_2 \ln\left(K^B_{pijt}\right) + \beta_3 \ln\left(K^E_{pijt}\right) + \beta_4 \ln\left(M_{pijt}\right) + \delta 1(Winner)_{pj} + \psi Trend_{jt} + \Omega\left[Trend_{jt} \times 1(Winner)_{pj}\right] + \kappa 1(\tau \geq 0)_{jt} + \gamma\left[Trend_{jt} \times 1(\tau \geq 0)_{jt}\right] + \theta_1\left[1(Winner)_{pj} \times 1(\tau \geq 0)_{jt}\right] + \theta_2\left[Trend_{jt} \times 1(Winner)_{pj} \times 1(\tau \geq 0)_{jt}\right] + \alpha_p + \mu_{it} + \lambda_j + \varepsilon_{pijt}$$

Spillover Estimates: Empirical Methodology

- With plant and case fixed effects,
  - The δ parameter for $1(Winner)_{pj}$ is identified by within-plant variation in winner status.
    - In other words, δ is identified by plants that are in a winning county for at least one case and in a losing county for at least one case.
    - If no county appears as both a winner and loser in a sample of cases, then δ cannot be identified from Equations (1) and (2).

Results: Weighted spillover estimates

	Case set 1		Case set 2	Case set 3	Case set 4
	GHM Losers	Pscore Losers			
	(1)	(2)	(3)	(4)	(5)
			Model 1		
Mean shift	0.01688	0.006965	-0.005743	3.714e-04	-7.353e-04
	(0.02087)	(0.02032)	(0.01602)	(0.002475)	(0.005489)
			Model 2		
Change after 5 years	0.08759**	0.1677***	0.03731	0.02020**	0.01481
	(0.04323)	(0.06209)	(0.03702)	(0.008461)	(0.01752)
Observations	27,000	17,500	30,500	103,000	123,000
R-squared	0.985	0.987	0.983	0.980	0.978
Plant FE	Y	Y	Y	Y	Y
Industry X Year FE	Y	Y	Y	Y	Y
Case FE	Y	Y	Y	Y	Y

NOTES: The table presents the results of estimating Equations 1 and 2 with five samples of continuously-appearing incumbent plants weighted by plants' total value of shipments in year $\tau = -8$. Columns 1 and 2 employ the continuously-appearing incumbent plant samples of GHM MDP winning counties with the plants in GHM losing counties and nearest propensity score losing counties, respectively. Columns (3) – (5) present estimates for continuously-appearing incumbent plants in the case set 2, 3, and 4, respectively, winning counties and their nearest propensity score neighbors. The reported mean shift if the equivalent of the $\theta_1$parameter from estimating equation 1 and the change after 5 years is calculated as $\theta_1 + 6\theta_2$ from estimating equation 2. Standard errors clustered at the county level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Results: Unweighted spillover estimates

	Case set 1 GHM Losers	Case set 1 Pscore Losers	Case set 2	Case set 3	Case set 4
	(1)	(2)	(3)	(4)	(5)
			Model 1		
Mean shift	0.006804	-0.003427	-0.003459	-7.063e-04	-0.002225
	(0.01076)	(0.01304)	(0.009956)	(0.002784)	(0.004283)
			Model 2		
Change after 5 years	0.03066	0.04259	-0.01047	-0.002356	0.003883
	(0.03244)	(0.04214)	(0.03131)	(0.009351)	(0.01320)
Observations	27,000	17,500	30,500	103,000	123,000
R-squared	0.967	0.968	0.963	0.957	0.955
Plant FE	Y	Y	Y	Y	Y
Industry X Year FE	Y	Y	Y	Y	Y
Case FE	Y	Y	Y	Y	Y

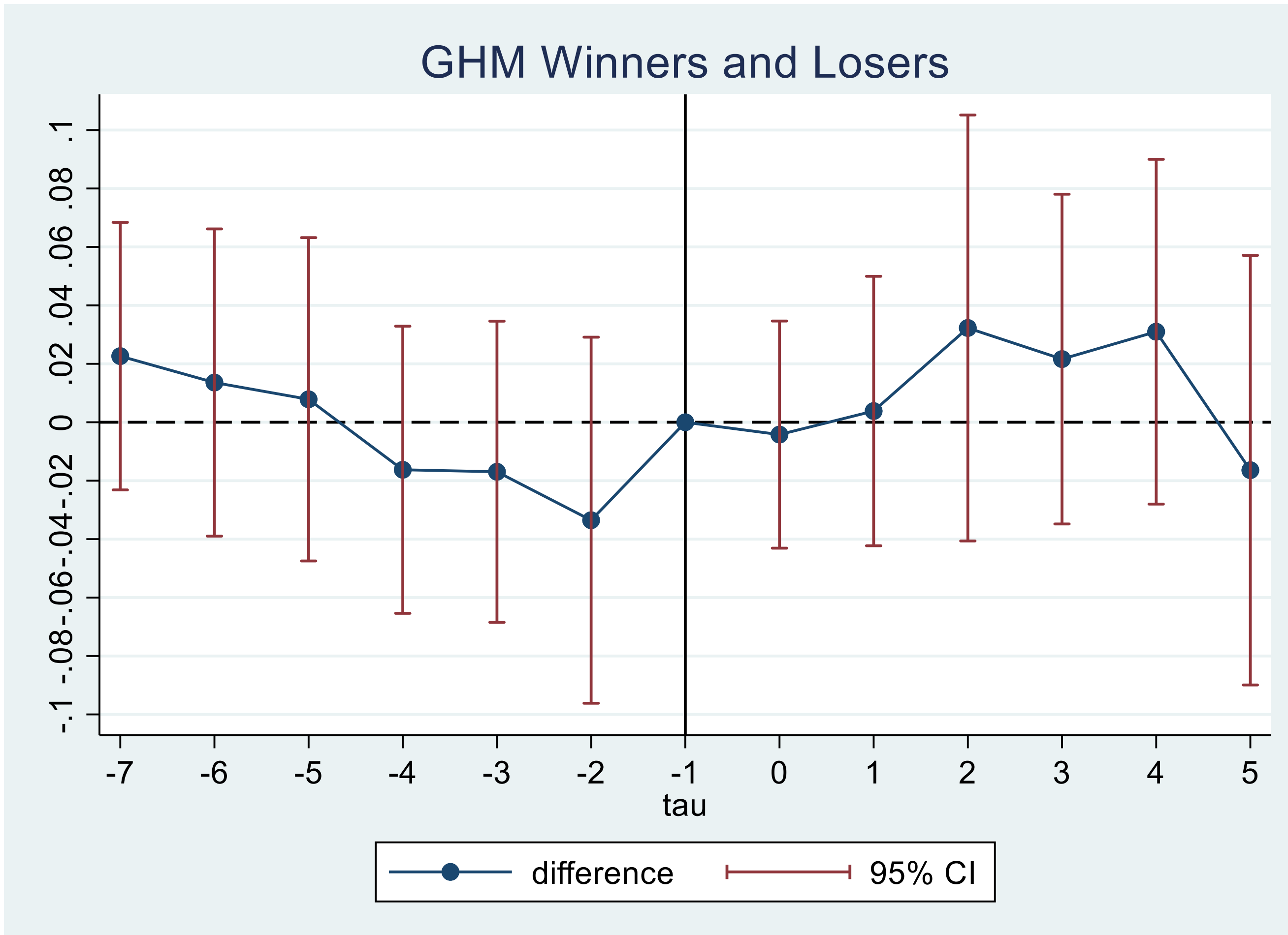
NOTES: The table presents the unweighted results of estimating equations 1 and 2 with five samples of continuously-appearing incumbent plants. Columns 1 and 2 employ the continuously-appearing incumbent plant samples of GHM MDP winning counties with the plants in GHM losing counties and nearest propensity score losing counties, respectively. Columns (3) – (5) present estimates for continuously-appearing incumbent plants in the case set 2, 3, and 4, respectively, winning counties and their nearest propensity score neighbors. The reported mean shift if the equivalent of the $\theta_1$parameter from estimating equation 1 and the change after 5 years is calculated as $\theta_1 + 6\theta_2$ from estimating equation 2. Standard errors clustered at the county level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Spillover Estimates: Test of Identifying Assumptions

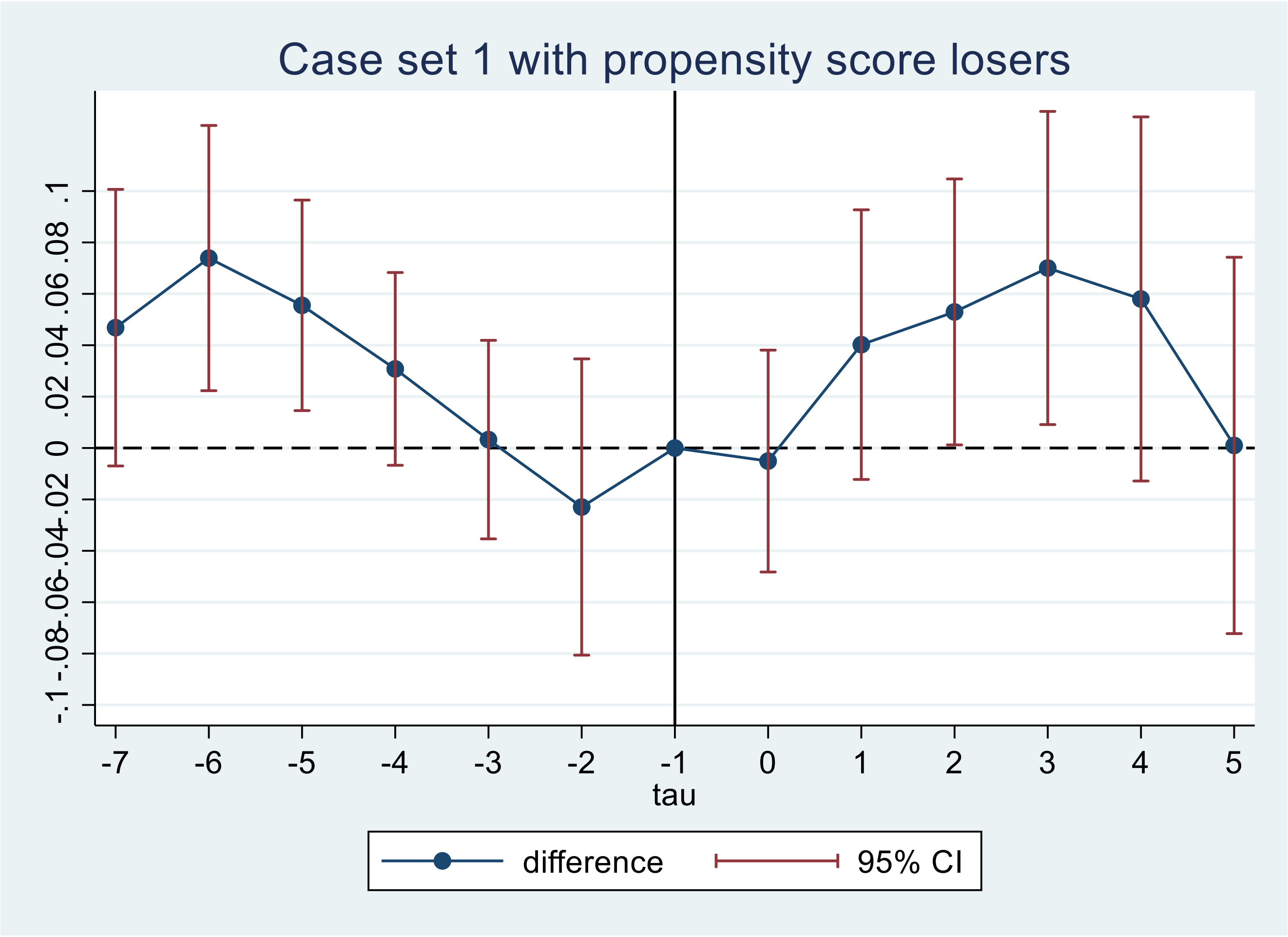
- Estimate

$$\ln\left(Y_{pijt}\right) = \beta_1 \ln\left(L_{pijt}\right) + \beta_2 \ln\left(K_{pijt}^{B}\right) + \beta_3 \ln\left(K_{pijt}^{E}\right) + \beta_4 \ln\left(M_{pijt}\right) + \theta_{w\tau} \sum_{\tau=-7}^{5}\left[1(Winner)_{pj} \times 1(t=\tau)_{jt}\right] + \theta_{l\tau} \sum_{\tau=-7}^{5}\left[1(Loser)_{pj} \times\right.$$

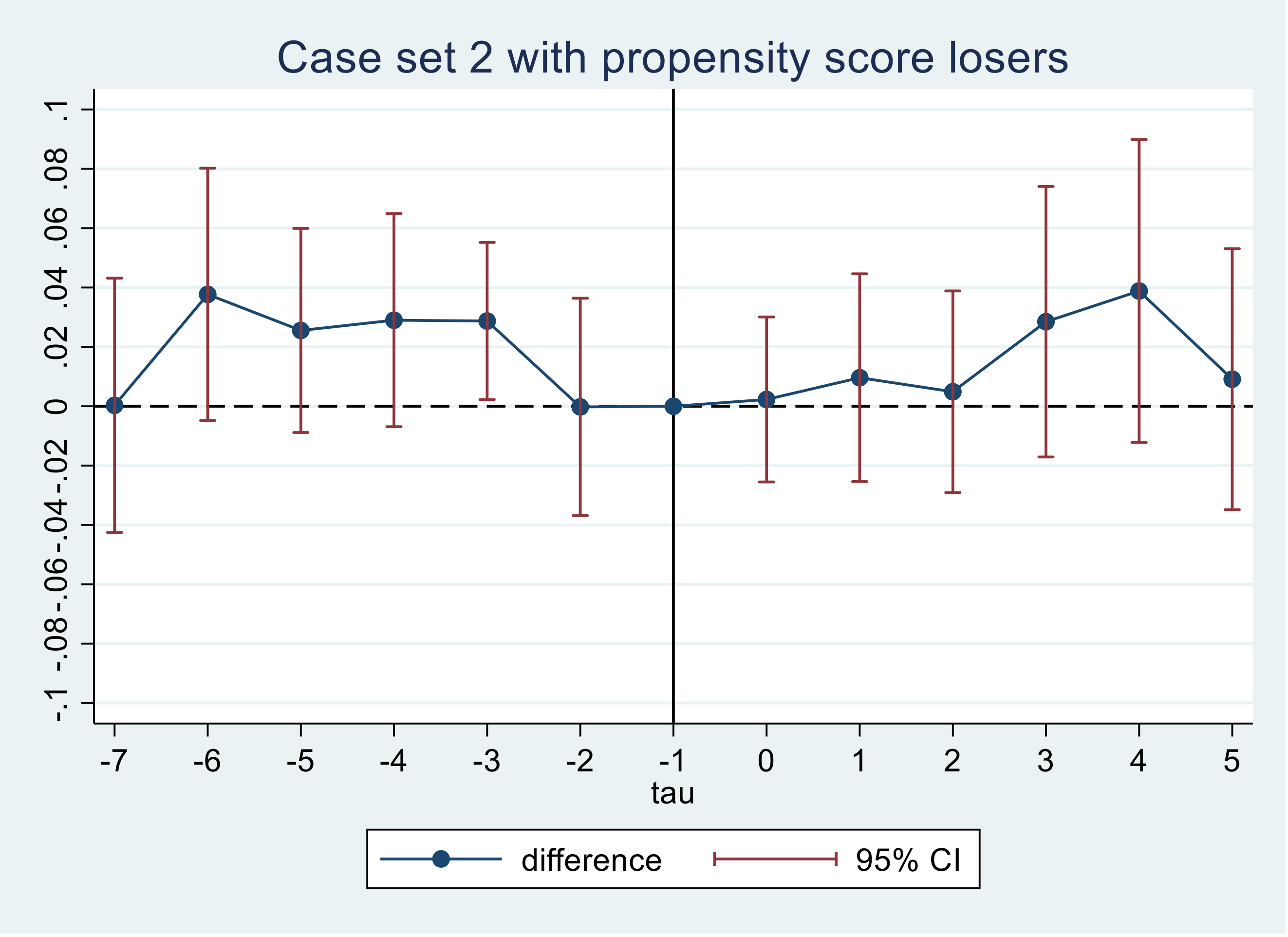
Spillover event study: GHM winners and losers



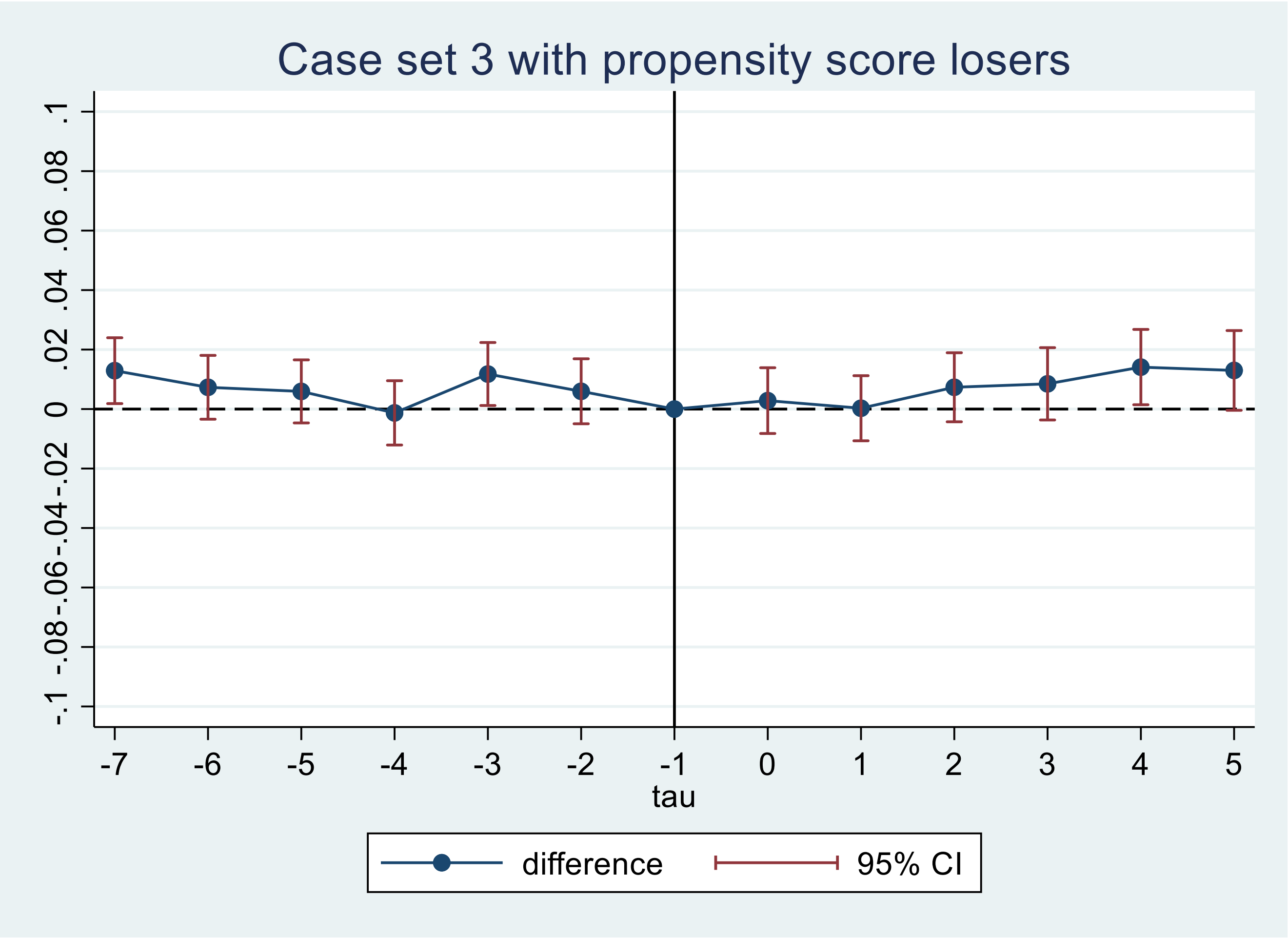
Spillover event study: GHM winners and propensity score losers



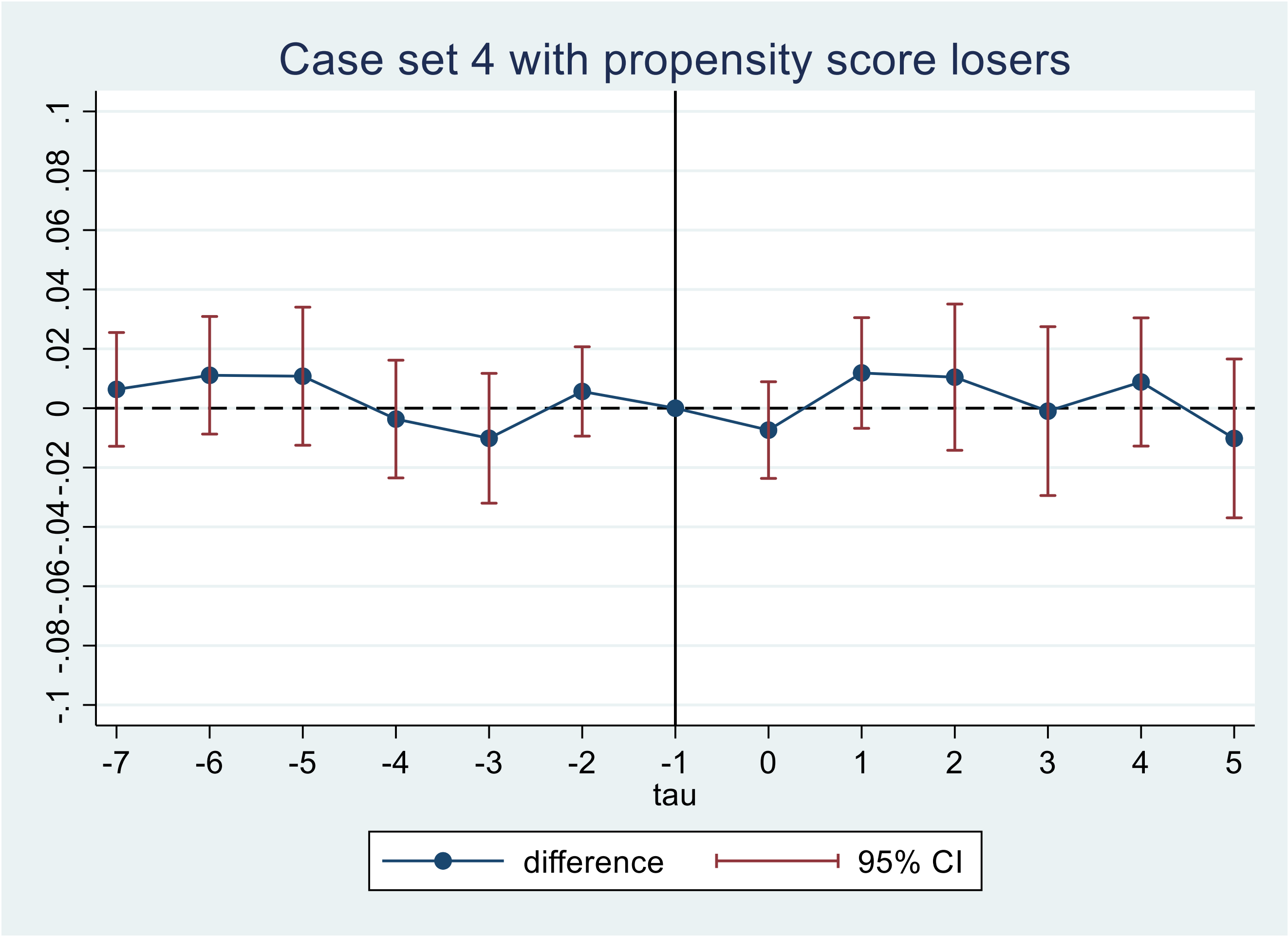
Spillover event study: Case set 2 winners and propensity score losers



Spillover event study: Case set 3 winners and propensity score losers



Spillover event study: Case set 4 winners and propensity score losers



Spillover Estimates: Robustness

- Incumbent plant output, capital expenditure, labor, and material inputs are simultaneously determined by the firm and these decisions may also be affected by time-varying unobservables that affect both selection and incumbent plant TFP.

- Estimates the spillover effects using a two-step procedure

  – A variant of the Combes et al. (2008, 2010) two-stage estimator adapted to a production function and our context.

  – Our preferred two-step method directly addresses the simultaneity of the output and inputs as well as firm heterogeneity using the Levinsohn-Petrin (2003) estimator in the first stage.

Spillover Estimates: Robustness

- Our preferred variant of the two-stage procedure reintroduces the plant fixed effect in the first-stage and estimates

$$\ln\left(Y_{p\tilde{\imath}jt}\right) = \beta_1 \ln\left(L_{p\tilde{\imath}jt}\right) + \beta_2 \ln\left(K^{B}_{p\tilde{\imath}jt}\right) + \beta_3 \ln\left(K^{E}_{p\tilde{\imath}jt}\right) + \beta_4 \ln\left(M_{p\tilde{\imath}jt}\right) + \alpha_p + \mu_{it} + \varepsilon_{p\tilde{\imath}jt}$$

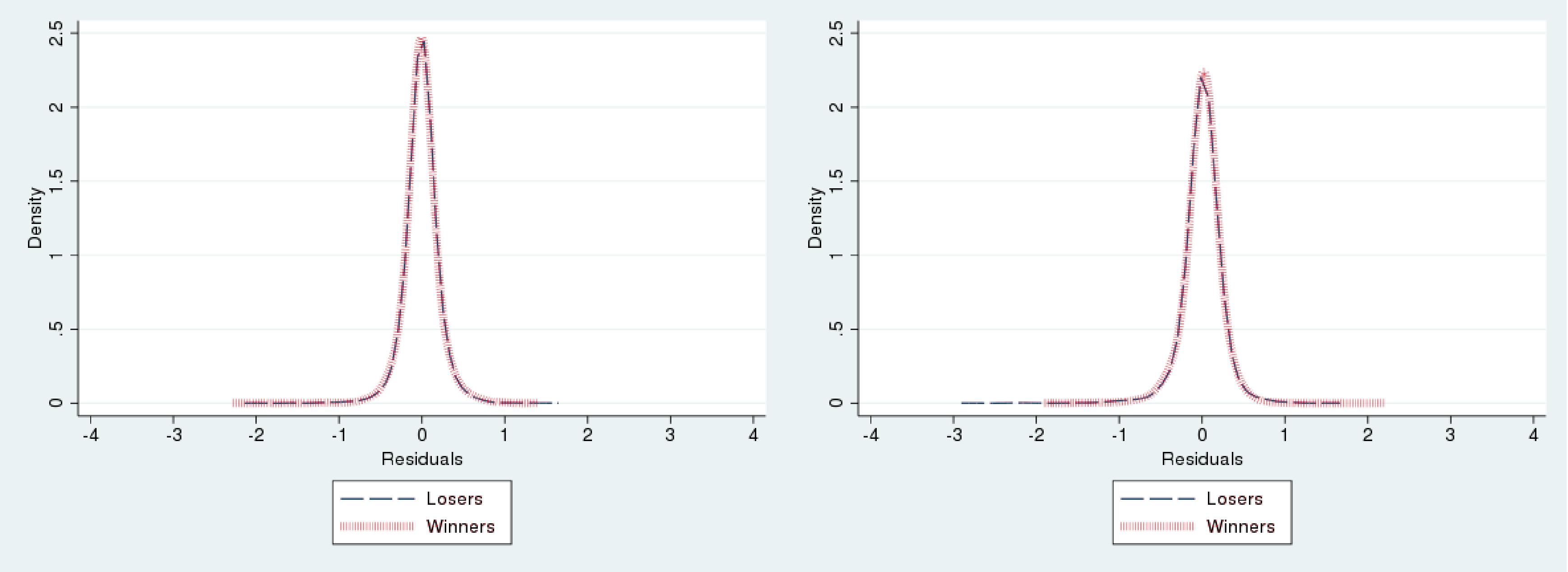
  using the Levinsohn-Petrin (2003) estimator.

- We then predict the residual for each plant and examine the distribution before estimating the spillover effect

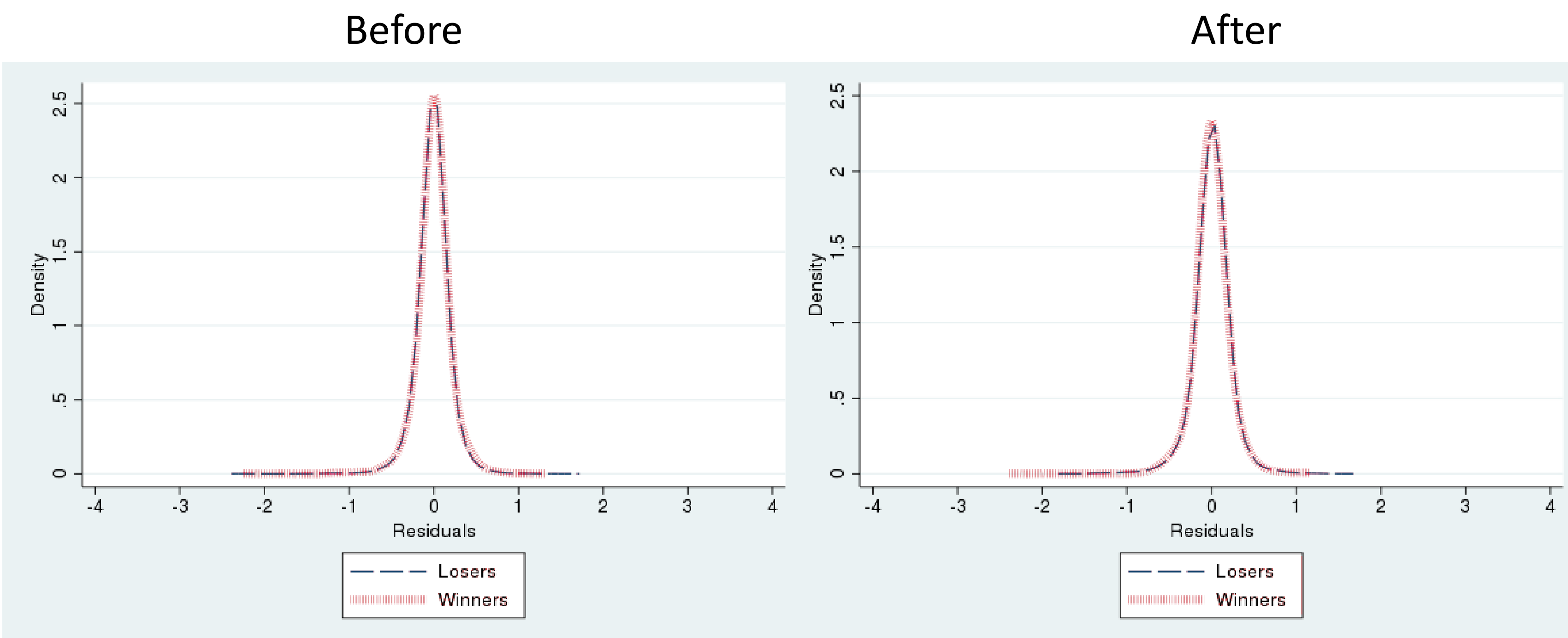
LP First-stage Residual TFP: Case Set 1 GHM Losers

Before

After

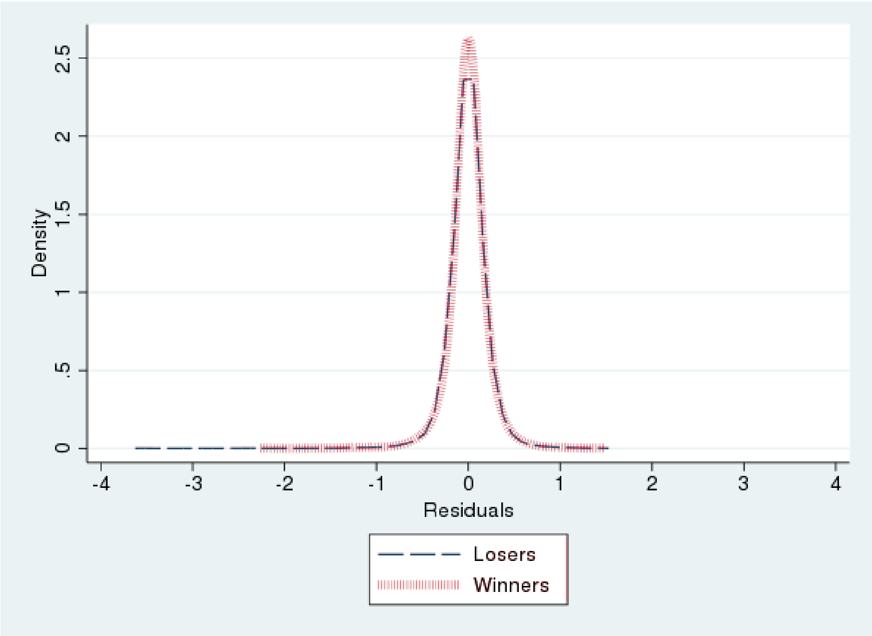


LP First-stage Residual TFP: Case Set 1 Propensity Score Losers

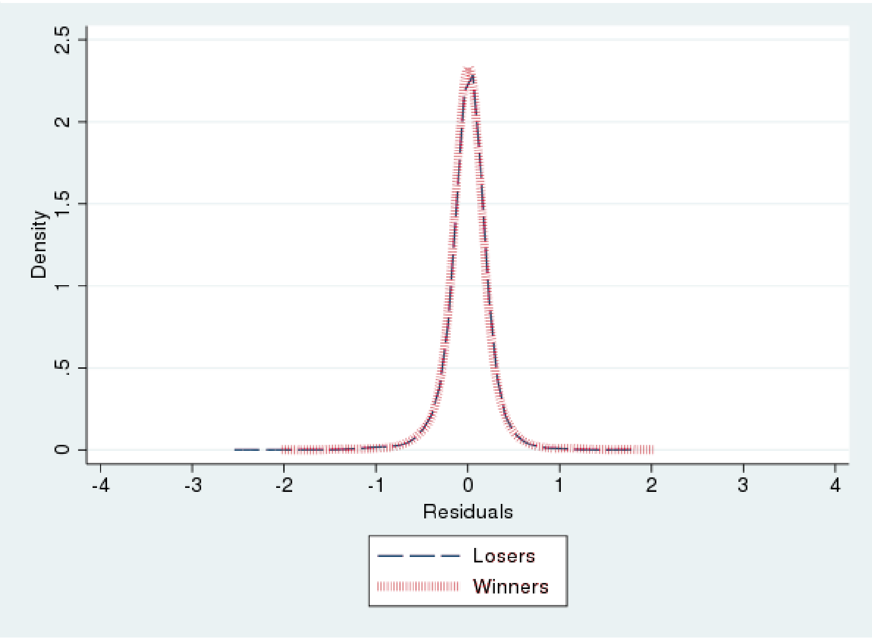


LP First-stage Residual TFP: Case Set 2

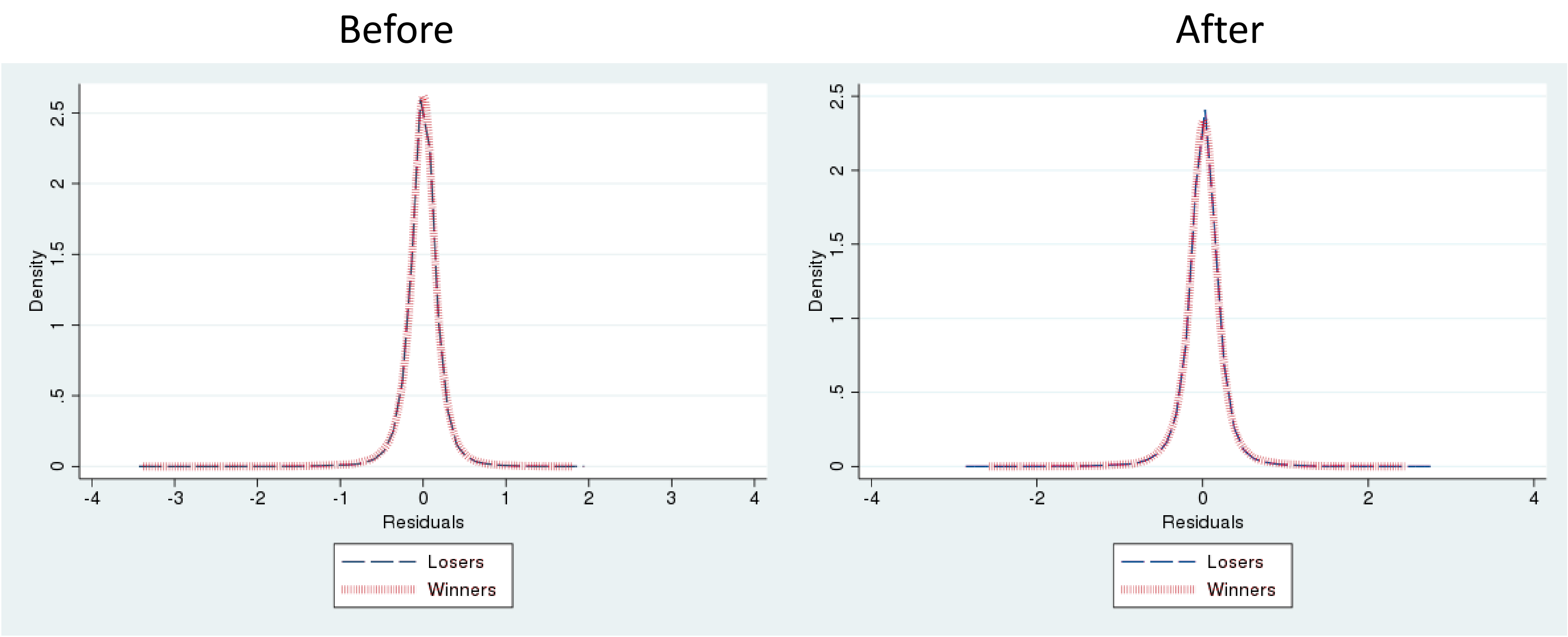
Before



After



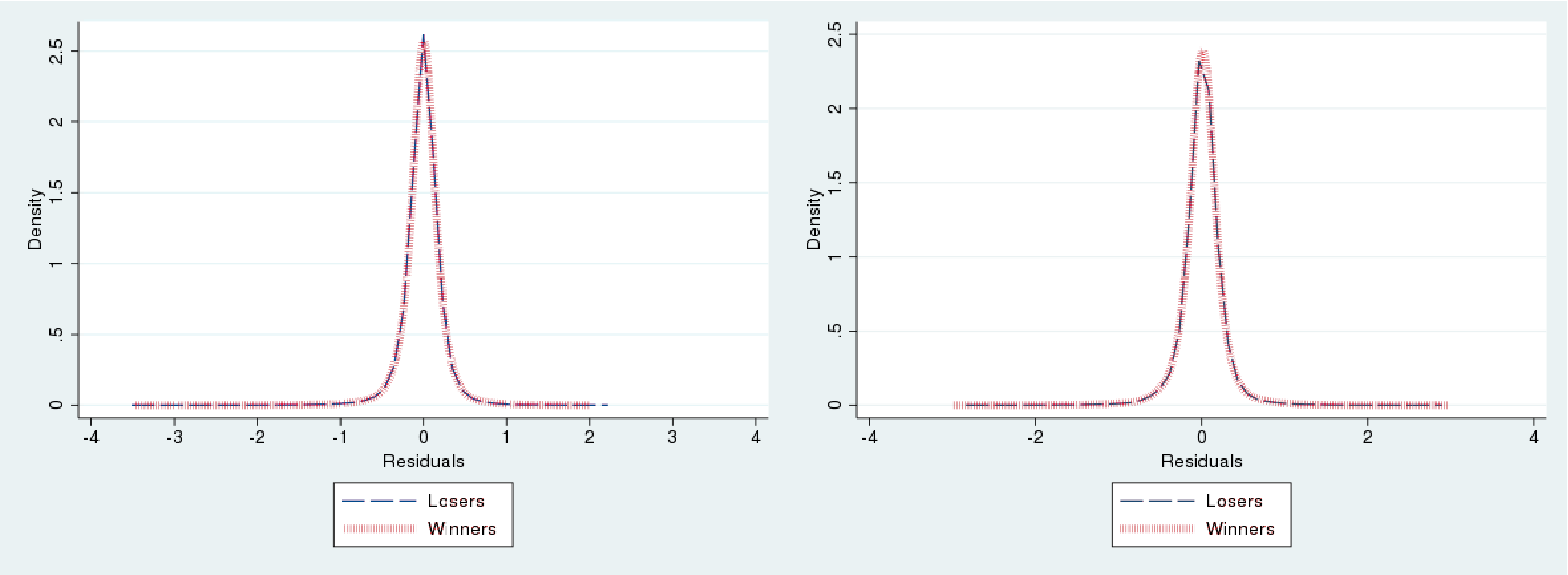
LP First-stage Residual TFP: Case Set 3



LP First-stage Residual TFP: Case Set 4

Before

After



Spillover Estimates: Robustness

- We then average the predicted residuals by county, 3 digit SIC code industry, and year to get $\hat{\hat{\varepsilon}}_{c(j)\tilde{\imath}t}$, the predicted average residual TFP in each county-industry-year .

- The second-stage estimates the spillover effects with

$$\hat{\hat{\varepsilon}}_{c(j)\tilde{\imath}t} = \delta 1(Winner)_{cj} + \kappa 1(\tau \geq 0)_{jt} + \theta_1 \big[1(Winner)_{cj} \times$$

Spillover Estimates: Second-stage LP estimates

	Case set 1		Case set 2	Case set 3	Case set 4
	GHM Losers	Pscore Losers			
	(1)	(2)	(3)	(4)	(5)
			Model 1		
Mean shift	-1.982e-04	-0.008209	0.002808	7.675e-04	-5.032e-04
	(0.01179)	(0.01079)	(0.009574)	(0.003655)	(0.003432)
Change after 5 years	0.02662	0.01944	-0.01535	-0.01894	-2.978e-04
	(0.03494)	(0.03677)	(0.03174)	(0.01256)	(0.01014)
Second Stage Obs.	18,500	13,500	20,000	74,000	85,000
Second Stage R-squared	0.114	0.051	0.041	0.031	0.028
First Stage Obs.	27000	17,500	30,500	103000	123000
First Stage R-squared	0.983	0.983	0.983	0.979	0.977
Plant fixed effects	Y	Y	Y	Y	Y
Industry X Year FE	Y	Y	Y	Y	Y
Case fixed effects	Y	Y	Y	Y	Y

Non-linearity: Empirical Methodology

To assess whether the spillovers are related to local plant density in a nonlinear way, we again use a two-step procedure and estimate

$$\ln\left(Y_{p\tilde{\imath}jt}\right) = \beta_1 \ln\left(L_{p\tilde{\imath}jt}\right) + \beta_2 \ln\left(K^B_{p\tilde{\imath}jt}\right) + \beta_3 \ln\left(K^E_{p\tilde{\imath}jt}\right) + \beta_4 \ln\left(M_{p\tilde{\imath}jt}\right) + \alpha_p + \varepsilon_{p\tilde{\imath}jt}$$

using the Levinsohn-Petrin (2003) estimator.

- We then predict the residual for each plant and average them by county, 3 digit SIC code industry, and year to get $\hat{\bar{\varepsilon}}_{c(j)\tilde{\imath}t}$

Non-linearity: Empirical Methodology

In the second stage, we semi-parametrically estimate

$$\hat{\hat{\varepsilon}}_{c(j)\tilde{\iota}t} = \delta 1(Winner)_{cj} + \kappa 1(\tau \geq 0)_{jt} + \theta_1 \left[1(Winner)_{cj} \times 1(\tau \geq 0)_{jt}\right] + g\left(ln\left(\frac{E_{i(p),j,c(j),t-s}}{R_{c(j)}}\right)\right) + \lambda_j + \epsilon_{cj\tilde{\iota}t},$$

where $g\left(ln\left(\frac{E_{i(p),j,c(j),t-s}}{R_{c(j)}}\right)\right)$ is an unknown function of the log (weighted) number of employees *E* per square mile in plant *p*'s county *c* for case *j* in year *t*-*s*.

Non-linearity: Empirical Methodology

- We consider four definitions of local plant density (corresponding to the weights in the density calculation):
  - establishments in the plants’ own-industry,
  - supplier industries,
  - customer industries, and
  - industries which share similar labor
- Each county-3 digit SIC industry-year combination therefore has its own density measure
- We instrument for local density using 1940 market potential (distance-weighted county income)

Non-linearity: Changes in plants densities

Weighting	All (1)	Winners Before (2)	Losers Before (3)	Winners After (4)	Losers After (5)
Panel A: GHM Sample					
Own Industry	2.645	2.243	2.888	2.476	2.74
	(11.03)	(15.73)	(5.608)	(18.48)	(5.298)
Proximity to mfg. input suppliers	1.626	0.9169	2.085	0.9392	1.904
	(2.786)	(1.168)	(3.337)	(1.399)	(3.091)
Proximity to mfg. output customers	2.34	1.938	2.619	2.16	2.389
	(10.78)	(15.7)	(4.885)	(18.48)	(4.404)
Labor pooling: CPS worker transitions	8.837	5.596	11.03	5.617	10.21
	(15.05)	(13.56)	(15.17)	(15.48)	(14.72)
Panel B: Case set 1 and propensity score losers					
Own Industry	2.03	1.536	2.293	1.522	2.19
	(3.525)	(2.02)	(4.268)	(1.878)	(3.539)
Proximity to mfg. input suppliers	1.172	0.9248	1.348	0.8824	1.205
	(1.917)	(1.167)	(2.354)	(1.038)	(1.824)
Proximity to mfg. output customers	1.518	1.257	1.696	1.184	1.502
	(2.567)	(1.821)	(3.054)	(1.652)	(2.374)
Labor pooling: CPS worker transitions	6.531	5.232	7.293	4.947	6.693
	(10.35)	(5.42)	(12.8)	(4.891)	(9.972)

Non-linearity: Changes in plants densities

Weighting	All (1)	Winners Before (2)	Losers Before (3)	Winners After (4)	Losers After (5)
Panel C: Case set 2 and propensity score losers					
Own Industry	3.104	3.072	3.381	2.829	3.025
	(6.485)	(5.912)	(7.574)	(4.414)	(5.961)
Proximity to mfg. input suppliers	2.185	2.255	2.335	2.007	1.936
	(4.25)	(3.646)	(5.07)	(2.813)	(3.673)
Proximity to mfg. output customers	2.782	3.262	2.816	2.701	2.445
	(5.342)	(6.363)	(5.293)	(4.271)	(4.64)
Labor pooling: CPS worker transitions	11.68	12.61	12.23	11.25	10.39
	(18.92)	(16.99)	(21.97)	(14.12)	(17.03)
Panel D: Case set 3 and propensity score losers					
Own Industry	2.894	2.936	3.104	2.69	2.871
	(6.357)	(6.178)	(6.749)	(5.654)	(5.812)
Proximity to mfg. input suppliers	1.861	1.969	1.975	1.787	1.74
	(3.585)	(3.412)	(3.968)	(2.972)	(3.214)
Proximity to mfg. output customers	2.399	2.566	2.56	2.25	2.181
	(5.001)	(4.889)	(5.376)	(4.202)	(4.333)
Labor pooling: CPS worker transitions	10.01	10.54	10.73	9.373	9.504
	(17.36)	(15.25)	(19.79)	(13.29)	(16.66)

Non-linearity: Changes in plants densities

Weighting	All (1)	Winners Before (2)	Losers Before (3)	Winners After (4)	Losers After (5)
Case set 4 and propensity score losers					
Own Industry	3.113	3.562	3.013	3.482	2.818
	(9.69)	(10.03)	(7.063)	(14.7)	(6.409)
Proximity to mfg. input suppliers	2.016	2.496	1.86	2.239	1.632
	(4.39)	(4.946)	(3.527)	(5.69)	(2.845)
Proximity to mfg. output customers	2.868	3.689	2.556	3.335	2.234
	(12.85)	(16.88)	(5.975)	(16.96)	(5.124)
Labor pooling: CPS worker transitions	11.41	14.11	10.35	13.43	9.194
	(35.65)	(37.69)	(18.87)	(58.58)	(15.82)

Non-linearity: Results

- Residual TFP increases linearly with interactions between economically-close plants for the range of densities most frequently observed in the data.

- Non-linearities observed only at density levels many standard deviations above the mean in the data.

  - Also ranges over which congestion externalities dominate agglomeration externalities

  - Semiparametric fits

- This suggests little gains in overall U.S. manufacturing output associated with moving plants from one location to another.

Multiple Equilibria: Empirical Methodology

- Compare county-manufacturing (and county-manufacturing-industry shares of national manufacturing (and manufacturing-industry) from before the MDP location with those 20 years after the large plant opening.

- Define county-manufacturing (log) share as

$$s_{c(j)\tau} \equiv \ln\left(S_{c(j)\tau}\right) = \ln\left(\frac{manufacturing\ output_{c(j)\tau}}{manufacturing\ output_{US\tau}}\right)$$

- Assuming a unique, stable equilibrium, county manufacturing shares at time τ can be modeled as

$$s_{c(j)\tau} = \Pi_{c(j)} + \varepsilon_{c(j)\tau},$$

where $\Pi_{\mathrm{c(j)}}$ is the initial equilibrium size in county *c* and $\varepsilon_{c(j)\tau}$ is a location-specific shock to manufacturing share.

Multiple Equilibria: Empirical Methodology

- Persistence of shocks takes the form

  $$\varepsilon_{c(j),\tau+1} = \rho\varepsilon_{c(j),\tau} + \nu_{c(j),\tau+1},$$

  where $\rho \in [0,1)$ is the persistence parameter.

- Let $v_{c(j),5}$ be the MDP shock to county output during the first five years after opening, and

- $v_{c(j),20}$ be the typical idiosyncratic location-specific shock to manufacturing share around the new post-MDP equilibrium $s_{c(j),20}$.

Multiple Equilibria: Empirical Methodology

- Then we can write the effect of the MDP shock to winning county *c(j)*'s share of manufacturing output as :

$$s_{c(j),20} - s_{c(j),5} = (\rho - 1)\nu_{c(j),5} + \left[\nu_{c(j),20} + \rho(1-\rho)\varepsilon_{c(j),-5}\right]$$

- If $\rho = 1$, then the shock is permanent and shares follow a random walk.
- If $\rho = 0$, then the shock has no effect.
- $0 \leq \rho < 1$ suggests a mean-reverting process, which may or may not be consistent with multiple equilibria

Multiple Equilibria: Empirical Methodology

- Possible that $\rho \neq 0$ because there is some correlation between the future changes in county manufacturing shares and past changes that we do not model.
  - Include pre-MDP opening growth in manufacturing share as a control
- It is also possible that the MDP shock is correlated with the error term
  - Instrument with for the shock:
    - Average national establishment output for firms in the MDP's 4-digit SIC industry and average national new entrant output in the MDP's 3-digit industry in time $\tau = -1$ expressed as a share of initial winning county manufacturing output.

Multiple Equilibria :Empirical Methodology

- Then, our estimating equation therefore becomes:
  $$s_{c(j),20} - s_{c(j),5} = \alpha\hat{v}_{c(j),5} + \beta_0 + \zeta PreMDP_{c(j)} + error_{c(j)}$$
- Estimate of $\alpha = (\rho - 1)$ gives evidence about whether the data support or reject the null of a unique equilibria, but doesn’t test against multiple equilibria

Multiple Equilibria: Empirical Methodology

- In the case of three equilibria, a county's share of manufacturing a county's share of manufacturing output at the new post-MDP equilibrium may be written:

$$s_{c(j),20} = \begin{cases} \Pi_{c(j)} + \Delta_1 + \varepsilon^1_{c(j),20} & if\ \nu_{c(j),5} < b_1 \\ \Pi_{c(j)} + \varepsilon^2_{c(j)20} & if\ b_1 < \nu_{c(j),5} < b_2 \\ \Pi_{c(j)} + \Delta_3 + \varepsilon^3_{c(j)20} & if\ \nu_{c(j),5} > b_2 \end{cases}$$

where $\Delta_1$and Δ_3 are the respective differences in log-shares from the initial equilibrium and the new equilibrium, $b_1$and $b_2$are the respective thresholds, and

$$\varepsilon^1_{c(j),20} = \rho\left(\varepsilon_{c(j),5} - \Delta_1\right) + \nu_{c(j)20} \quad if\ \nu_{c(j),5} < b_1$$
$$\varepsilon^2_{c(j),20} = \rho\varepsilon_{c(j),5} + \nu_{c(j),20} \quad if\ b_1 < \nu_{c(j),5} < b_2$$
$$\varepsilon^3_{c(j),20} = \rho\left(\varepsilon_{c(j),5} - \Delta_3\right) + \nu_{c(j),20} \quad if\ \nu_{c(j),5} > b_2$$

Multiple Equilibria: Empirical Methodology

- We assume that the period is long enough for the shock to have dissipated ($\rho = 0$) and estimate:

  $$s_{c(j),20} - s_{c(j),0} = (1-\rho)\Delta_1 I_1\left(b_1, \widehat{v_{c(j),5}}\right) + (1-\rho)\Delta_3 I_3\left(b_2, \widehat{v_{c(j),5}}\right) + \zeta PreMDP_{c(j)} + v_{c(j),20},$$

  where $I_1\left(b_1, v_{c(j),5}\right)$ is an indicator variable equal to one if $v_{c(j),5} < b_1$ and $I_3\left(b_2, v_{c(j),5}\right)$ is an indicator variable equal to one if$v_{c(j),5} > b_2$

- Use a maximum likelihood grids search method to determine thresholds

- Consider one, two, three, and four equilibria specifications and use the value of the likelihood functions to determine which bests describes the data.

Multiple Equilibria: Empirical Methodology

- Choose the equilibria specification that maximizes the Schwarz Criterion
- For multiple equilibria specification, we also require that a larger positive shock be associated with larger, new equilibrium share and that the thresholds lie between equilibrium shares:
  - Following Davis and Weinstein (2008), we therefore impose the following intercept ordering criterion:

$$\delta_1 + \delta_2 < b_1 < \delta_2 < b_2 < \delta_2 + \delta_3 < b_3 < \delta_2 + \delta_4.$$

Multiple equilibria mfg. share tests

	$s_{c(j),20} - s_{c(j),5}$	$s_{c(j),20} - s_{c(j),-1}$			
	IV Estimate	1 Equilibrium	2 Equilibria	3 Equilibria	4 Equilibria
	(1)	(2)	(3)	(4)	(5)
Case Set 1					
MDP Shock	-0.05057				
	(0.05477)				
Intercept ordering	N/A	N/A	Pass	Fail	Fail
Schwarz Criterion	N/A	-46.57	-52.18	-53.25	-53.56
Case set 2					
MDP Shock	0.09841*				
	(0.05061)				
Intercept ordering	N/A	N/A	Pass	Pass	Fail
Schwarz Criterion	N/A	-182.8	-214.4	-214	-217.8
Case set 3					
MDP Shock	0.06165**				
	(0.02390)				
Intercept ordering	N/A	N/A	Pass	Fail	Fail
Schwarz Criterion	N/A	-851	-1006	-993.4	-998.8
Case set 4					
MDP Shock	0.6852				
	(0.4926)				
Intercept ordering	N/A	N/A	Pass	Fail	Fail
Schwarz Criterion	N/A	-484.9	-559	-563.1	-567.5

Full results

Multiple equilibria mfg.-industry share tests

	$s_{c(j),20} - s_{c(j),5}$	$s_{c(j),20} - s_{c(j),-1}$			
	IV Estimate	1 Equilibrium	2 Equilibria	3 Equilibria	4 Equilibria
	(1)	(2)	(3)	(4)	(5)
Case Set 1					
MDP Shock	8.568e-05***				
	(1.784e-05)				
Intercept ordering	N/A	N/A	Pass	Fail	Fail
Schwarz Criterion	N/A	-141.7	-172.3	-175	-177.7
Case set 2					
MDP Shock	8.694e-05***				
	(1.468e-05)				
Intercept ordering	N/A	N/A	Fail	Fail	Fail
Schwarz Criterion	N/A	-270.7	-335.2	-338.6	-342
Case set 3					
MDP Shock	8.804e-05***				
	(2.746e-05)				
Intercept ordering	N/A	N/A	Pass	Fail	Fail
Schwarz Criterion	N/A	-1887	-2408	-2412	-2417
Case set 4					
MDP Shock	-0.03428				
	(0.1258)				
Intercept ordering	N/A	N/A	Pass	Fail	Fail
Schwarz Criterion	N/A	-839.9	-632.2	-636.1	-641.4

Full results

Multiple equilibria: Births

	Case set 1		Case set 2	Case set 3	Case set 4
	GHM Losers	Pscore Losers			
	(1)	(2)	(3)	(4)	(5)
Difference-in-differences	0.009528	-0.02972	-0.02337	-0.06017***	-0.09044***
	(0.04393)	(0.04268)	(0.03204)	(0.01477)	(0.01451)
Obs.	2,700	3,100	5,900	22,000	20,000
R-squared	0.927	0.91	0.9	0.896	0.901
Year fixed effects	Y	Y	Y	Y	Y
Case fixed effects	Y	Y	Y	Y	Y

Birth trends

Multiple equilibria: Deaths

	Case set 1		Case set 2	Case set 3	Case set 4
	GHM Losers	Pscore Losers			
	(1)	(2)	(3)	(4)	(5)
Difference-in-differences	0.05927	0.02695	-0.003901	-0.01664	-0.05173***
	(0.04504)	(0.04072)	(0.03499)	(0.01486)	(0.01493)
Obs.	2,700	3,100	5,900	22,000	20,000
R-squared	0.944	0.93	0.93	0.921	0.929
Year fixed effects	Y	Y	Y	Y	Y
Case fixed effects	Y	Y	Y	Y	Y

Death trends

Conclusion

- We find a significant cumulative increase in incumbent plant productivity after 5 years associated with the GHM MDP openings, albeit without the large mean shift estimated by GHM.
  - Econometrically identified by a unique subset of plants that continuously operate in counties that are both a winner and a loser for more than one case
  - Also appear to be unique to particular MDP openings in the GHM sample
- We find much weaker spillovers associated with other highly-incentivized MDP openings
  - Identified from a unique set of plants given the estimation strategy.

Conclusion

- Find that the agglomeration function is linear in all 4 economically-close employment density measures over the range of densities most observed in the data
- Cannot rule out the possibility of multiple equilibria
- However, the data most strongly support one unique equilibrium county share of manufacturing activity
  - Evidence of multiple equilibria in county-industry manufacturing shares for the set of shocks with the highest wages

Conclusion

- Suggests that even in the presence of significant spillovers for some incumbent plants, these large plant openings are not a sufficiently large positive shock to push locations into a new equilibrium
  - This may be due to countervailing congestion forces or weaker than anticipated spillovers or both.
  - Changes in long-term births and death suggest, at least some, congestion externalities

Next steps

- Relate the spillovers to the size of the incentives
- Incorporate comments and suggestions

Thank you!

Carlianne Patrick
Department of Economics
Andrew Young School of Policy Studies
Georgia State University
cpatrick@gsu.edu

Spillover Estimates: Robustness

In our variant of the Combes et al. (2008, 2010) two-stage procedure, we first estimate

$$\ln(Y_{pijt}) = \beta_1 \ln(L_{pijt}) + \beta_2 \ln(K^B_{pijt}) + \beta_3 \ln(K^E_{pijt}) + \beta_4 \ln(M_{pijt}) + B_{c(j)it} + \varepsilon_{pijt},$$

where $B_{c(j)it}$ is a vector county-2 digit SIC code industry-time fixed effects.

Note that equation (3) does not include a plant fixed effect as we cannot separately identify $B_{c(j)it}$with plant fixed effects included.

Spillover Estimates: Robustness

We estimate the spillover effect in the second-stage with

$$B_{c(j)it} = \delta 1(Winner)_{cj} + \kappa 1(\tau \geq 0)_{jt} + \theta_1 \left[1(Winner)_{cj} \times 1(\tau \geq 0)_{jt}\right] + \mu_{it} + \lambda_j + \epsilon_{cjit}$$

$$B_{c(j)it} = \delta 1(Winner)_{pj} + \psi Trend_{jt} + \Omega\left[Trend_{jt} \times 1(Winner)_{pj}\right] + \kappa 1(\tau \geq 0)_{jt} + \gamma\left[Trend_{jt} \times 1(\tau \geq 0)_{jt}\right] + \theta_1\left[1(Winner)_{pj} \times 1(\tau \geq 0)_{jt}\right] + \theta_2\left[Trend_{jt} \times 1(Winner)_{pj} \times 1(\tau \geq 0)_{jt}\right] + \mu_{it} + \lambda_j + \epsilon_{cjit}$$

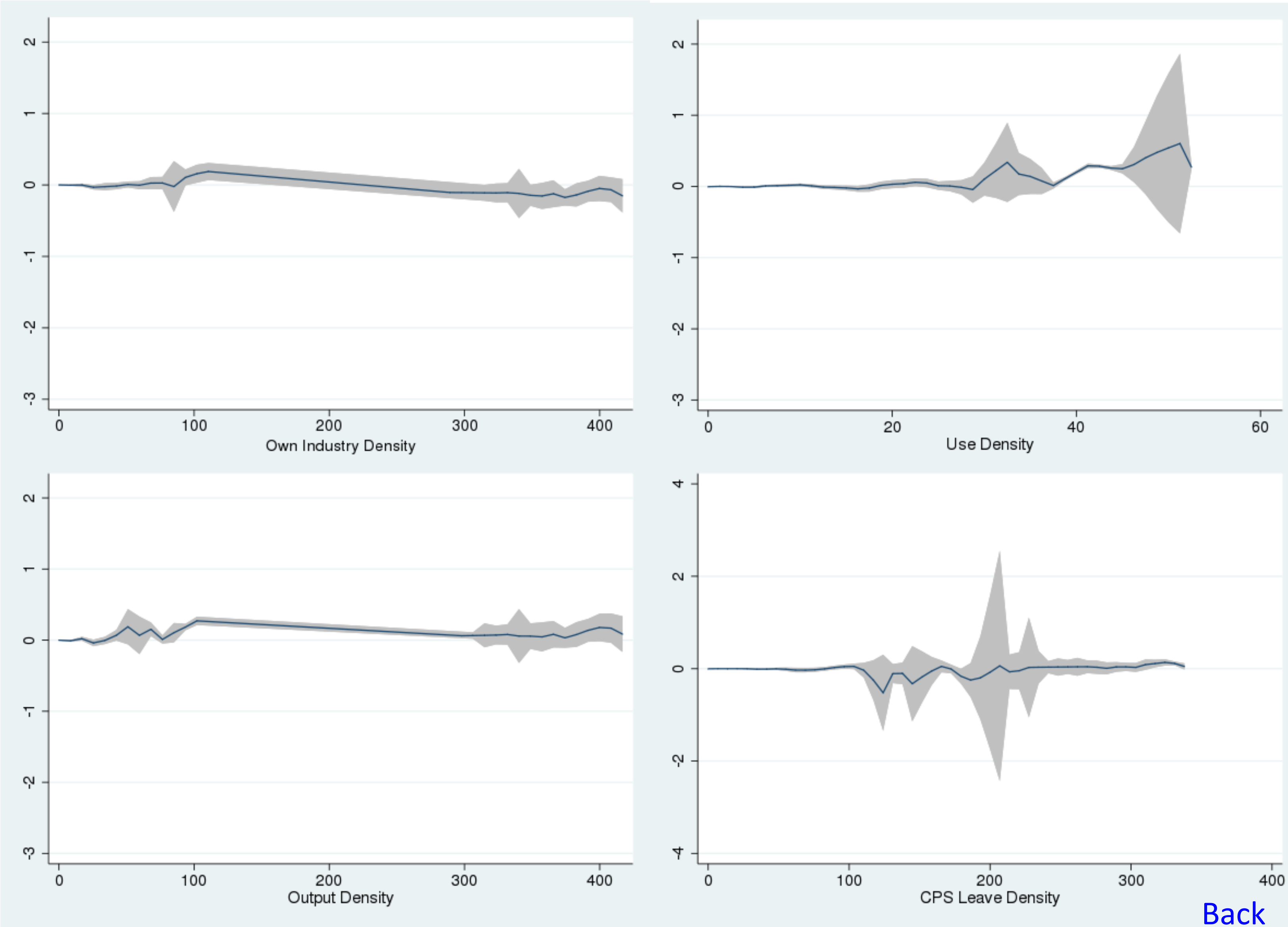
Spillover Estimates: Second-stage FE estimates

	Case set 1		Case set 2	Case set 3	Case set 4
	GHM Losers	Pscore Losers			
	(1)	(2)	(3)	(4)	(5)
			Model 1		
Mean shift	-0.008865	0.001391	0.008381	-0.001266	-0.003091
	(0.01431)	(0.01835)	(0.01440)	(0.005362)	(0.005280)
			Model 2		
Change after 5 years	0.08736**	0.08943**	0.03537	-0.02315	0.02270
	(0.03657)	(0.04311)	(0.04120)	(0.01591)	(0.01611)
Second Stage Obs.	10,000	8,500	12,000	49,000	53,500
Second Stage R2	0.285	0.330	0.333	0.304	0.306
First Stage Obs.	27,000	17,500	30500	103000	123000
First Stage R2	0.981	0.982	0.975	0.979	0.967
Plant fixed effects	Y	Y	Y	Y	Y
Industry X Year FE	Y	Y	Y	Y	Y
Case fixed effects	Y	Y	Y	Y	Y

Spillover Estimates: Location Quotient 2-stage LP estimates

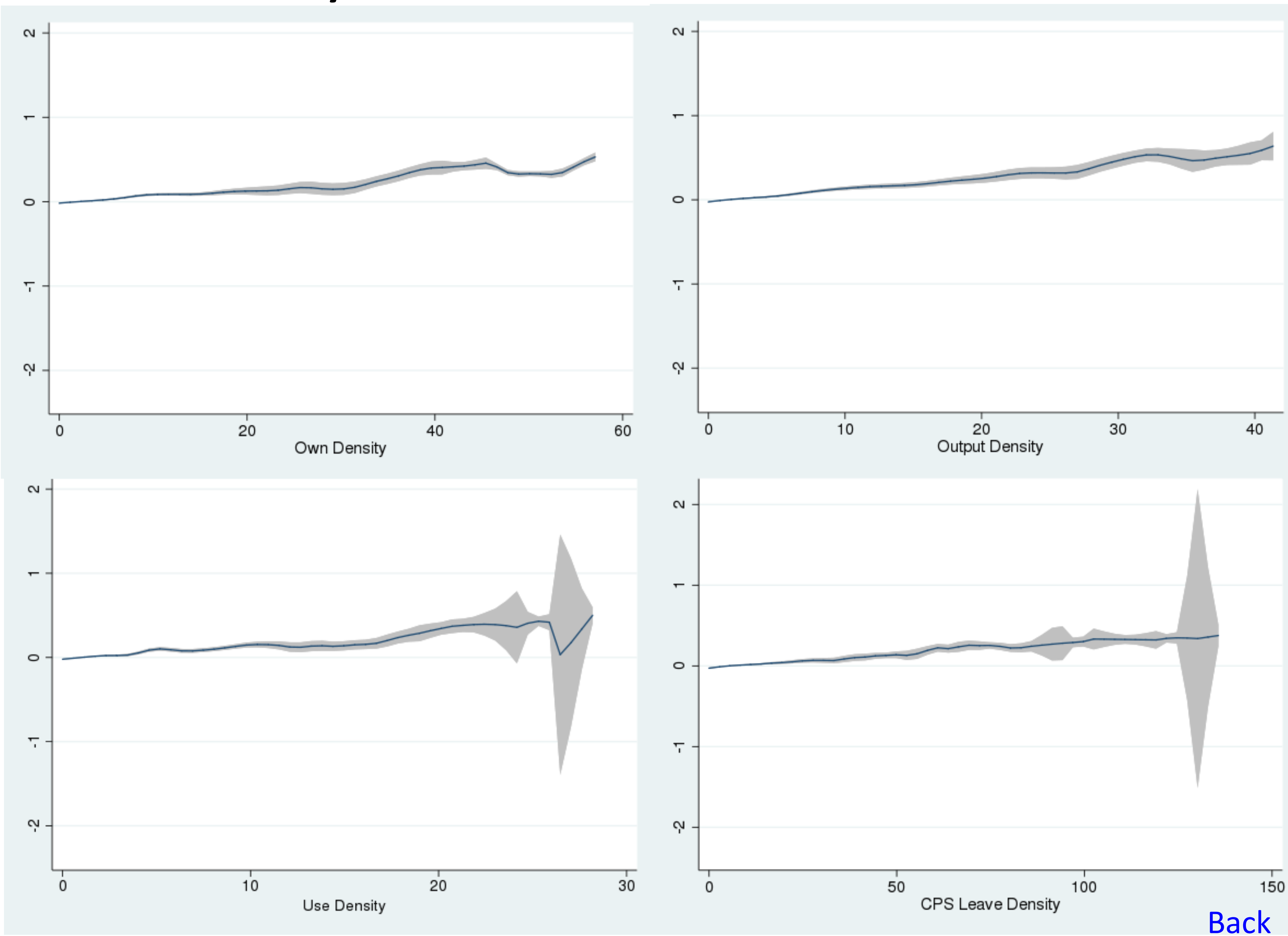
	(1)	(2)	(3)	(4)
	Model 1			
Mean shift	-0.008218	-8.978e-04	-7.885e-04	-0.003946
	(0.01079)	(0.01146)	(0.004880)	(0.003955)
	Model 2			
Change after 5 years	0.01944	0.003124	0.02278	0.01514
	(0.03676)	(0.03252)	(0.01443)	(0.01304)
Second Stage Obs.	13,500	13500	66,000	82,500
Second Stage R-squared	0.051	0.05	0.032	0.027
First Stage Obs.	17500	20000	92000	119000
First Stage R-squared	0.986	0.975	0.979	0.977
Plant fixed effects	Y	Y	Y	Y
Industry X Year FE	Y	Y	Y	Y
Case fixed effects	Y	Y	Y	Y

Non-linearity: Case set 1 and GHM loser results



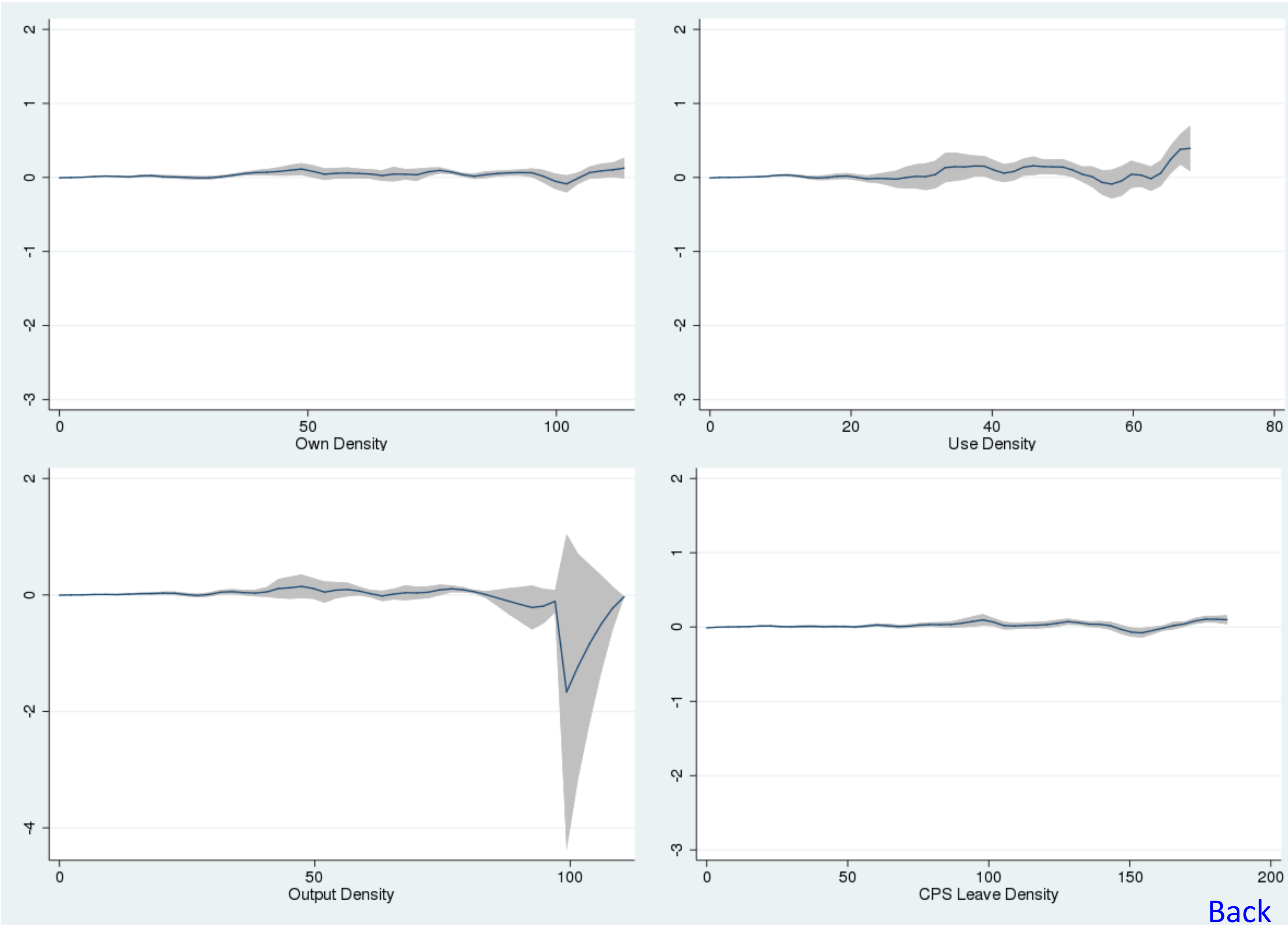
[Back](#)

Non-linearity: Case set 1 and PS loser results



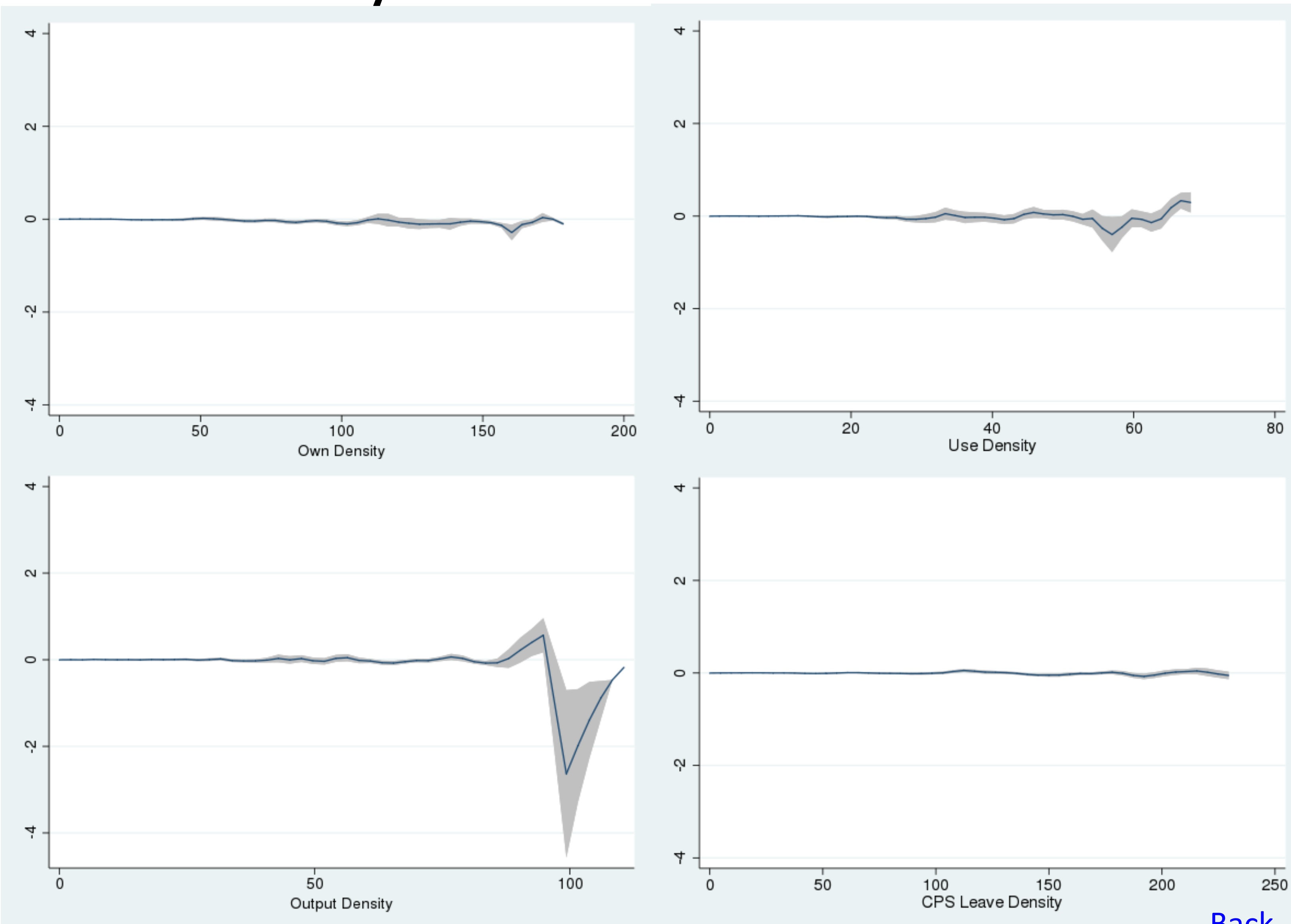
[Back]

Non-linearity: Case set 2 results



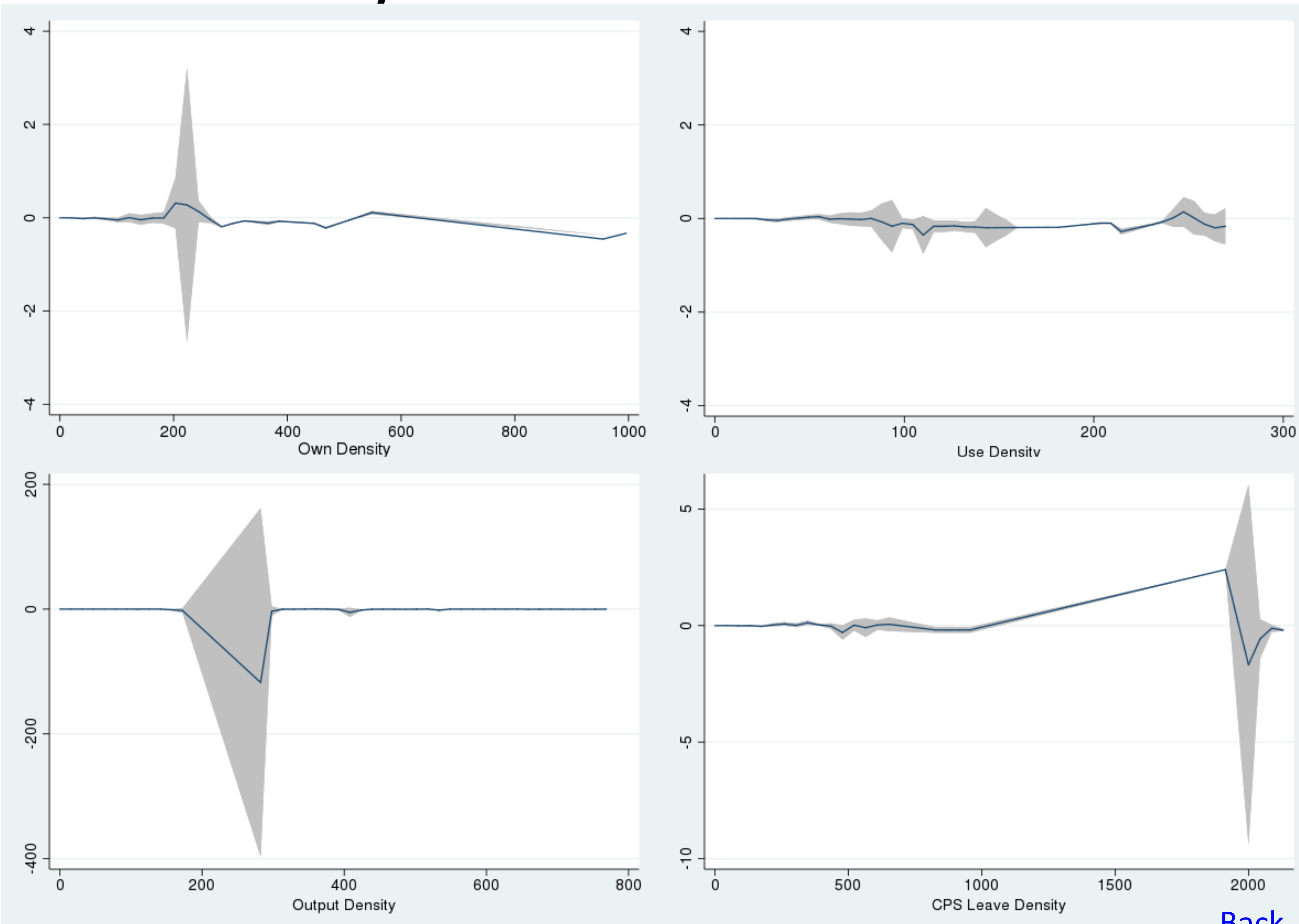
Back

Non-linearity: Case set 3 results



Back

Non-linearity: Case set 4 results



Back

Case set 1 multiple equilibria tests: mfg. shares

	$s_{c(j),20} - s_{c(j),5}$	$s_{c(j),20} - s_{c(j),-1}$			
	IV Estimate	1 Equilibrium	2 Equilibria	3 Equilibria	4 Equilibria
	(1)	(2)	(3)	(4)	(5)
Growth Rate $\tau = -5$	0.06642	1.325***	1.572***	1.29*	2.824***
to $\tau = 0$	(0.3079)	(0.4708)	(0.5147)	(0.5035)	(0.7544)
MDP Shock	-0.05057				
	(0.05477)				
δ_1			-1.771***	1.064**	-4.010**
			(0.5575)	(0.5036)	(1.628)
δ_3				1.952***	-1.323***
				(0.5328)	(0.4536)
δ_4					1.987***
					(0.4928)
Constant	0.2208*	0.4331**	2.004***	0.07808	1.224***
	(0.1215)	(0.1836)	(0.5218)	(0.2150)	(0.4043)
Thresholds					
b_1			0.3323	-0.7615	-4.253
b_2				0.4855	-0.3208
b_3					0.4855
Intercept ordering	N/A	N/A	Pass	Fail	Fail
Schwarz Criterion	N/A	-46.57	-52.18	-53.25	-53.56
Counties	30	30	30	30	30
R-squared		0.215	0.397	0.483	0.577

Back

Case set 2 multiple equilibria tests: mfg. shares

	$s_{c(j),20} - s_{c(j),5}$	$s_{c(j),20} - s_{c(j),-1}$			
	IV Estimate	1 Equilibrium	2 Equilibria	3 Equilibria	4 Equilibria
	(1)	(2)	(3)	(4)	(5)
Growth Rate $\tau = -5$	-3.193*	0.6045	1.515	1.462	2.185*
to $\tau = 0$	(1.723)	(0.8233)	(1.200)	(1.137)	(1.291)
MDP Shock	0.09841*				
	(0.05061)				
δ_1			-5.165***	-0.8436	2.360
			(1.267)	(1.26)	(2.518)
δ_3				6.488***	2.830
				(1.742)	(2.496)
δ_4					6.931***
					(1.428)
Constant	0.1606	0.9163***	5.351***	0.8958	-2.413
	(0.1764)	(0.3199)	(1.169)	(1.174)	(2.460)
Thresholds					
b_1			1.076	0.6336	0.04455
b_2				1.739	0.0716
b_3					1.739
Intercept ordering	N/A	N/A	Pass	Pass	Fail
Schwarz Criterion	N/A	-182.8	-214.4	-214	-217.8
Counties	70	70	70	70	70
R-squared		0.007	0.199	0.295	0.304

Back

Case set 3 multiple equilibria tests: mfg. shares

	$s_{c(j),20} - s_{c(j),5}$	$s_{c(j),20} - s_{c(j),-1}$			
	IV Estimate	1 Equilibrium	2 Equilibria	3 Equilibria	4 Equilibria
	(1)	(2)	(3)	(4)	(5)
Growth Rate $\tau = -5$	-0.08337	-0.05676	-0.2042	-0.2420	-0.2443
to $\tau = 0$	(0.06114)	(0.1150)	(0.1638)	(0.1570)	(0.1570)
MDP Shock	0.06165**				
	(0.02390)				
δ_1			-3.053***	-0.7947**	-0.2749
			(0.3346)	(0.3853)	(0.2548)
δ_3				3.553***	0.6676*
				(0.5121)	(0.4021)
δ_4					3.549***
					(0.5088)
Constant	-0.02123	0.4791***	3.212***	0.9009**	0.2414
	(0.04595)	(0.08505)	(0.3107)	(0.3643)	(0.1815)
Threshholds					
b_1			0.3316	0.23	0.0244
b_2				0.5663	0.23
b_3					0.5663
Intercept ordering	N/A	N/A	Pass	Fail	Fail
Schwarz Criterion	N/A	-851	-1006	-993.4	-998.8
Counties	450	450	450	450	450
R-squared		0.001	0.164	0.234	0.237

Back

Case set 4 multiple equilibria tests: mfg. shares

	$s_{c(j),20} - s_{c(j),5}$	$s_{c(j),20} - s_{c(j),-1}$			
	IV Estimate	1 Equilibrium	2 Equilibria	3 Equilibria	4 Equilibria
	(1)	(2)	(3)	(4)	(5)
Growth Rate $\tau = -5$	0.08275	0.3628***	0.3015***	0.3025***	0.3019***
to $\tau = 0$	(0.2011)	(0.02336)	(0.02776)	(0.02774)	(0.02766)
MDP Shock	0.6852				
	(0.4926)				
δ_1			-0.3799***	-0.8039***	-0.4935**
			(0.09322)	(0.2157)	(0.2211)
δ_3				-0.4937**	-0.2173*
				(0.2217)	(0.1202)
δ_4					0.7127***
					(0.2210)
Constant	-0.06347	-3.856e-04	0.1892***	0.6729***	-0.03981
	(0.04527)	(0.03918)	(0.06289)	(0.2074)	(0.07788)
Thresholds					
b_1			0.0003642	0.0007467	0.00009817
b_2				0.001002	0.0007538
b_3					0.001002
Intercept ordering	N/A	N/A	Pass	Fail	Fail
Schwarz Criterion	N/A	-484.9	-559	-563.1	-567.5
Counties	400	400	400	400	400
R-squared		0.374	0.262	0.269	0.275

Back

Case set 1 multiple equilibria tests: ind. shares

	$s_{c(j),20} - s_{c(j),5}$	$s_{c(j),20} - s_{c(j),-1}$			
	IV Estimate (1)	1 Equilibrium (2)	2 Equilibria (3)	3 Equilibria (4)	4 Equilibria (5)
Growth Rate $\tau = -5$ to $\tau = 0$	-0.003255 (0.005816)	-3.613 (9.779)	0.7158 (65.93)	0.7424 (69.19)	2.208 (125.1)
MDP Shock	8.568e-05*** (1.784e-05)				
δ_1			-41,040*** (12,870)	-72.48 (18,280)	470 (32,880)
δ_3				40,970* (21,420)	522.8 (36,060)
δ_4					41,000** (18,340)
Constant	0.2928 (0.4315)	729.7 (700.8)	41,020*** (11,840)	49.02 (17,490)	-505.4 (34,100)
Thresholds					
b_1			247.8	72.68	17.02
b_2				254.6	64.92
b_3					254.6
Intercept ordering	N/A	N/A	Pass	Fail	Fail
Schwarz Criterion	N/A	-141.7	-172.3	-175	-177.7
Counties	20	20	20	20	20
R-squared		0.010	0.464	0.464	0.464

Back

Case set 2 multiple equilibria tests: ind. shares

	$s_{c(j),20} - s_{c(j),5}$	$s_{c(j),20} - s_{c(j),-1}$			
	IV Estimate	1 Equilibrium	2 Equilibria	3 Equilibria	4 Equilibria
	(1)	(2)	(3)	(4)	(5)
Growth Rate $\tau = -5$	-0.002111				
to $\tau = 0$	(0.008214)				
MDP Shock	8.694e-05***	-1.553	-4.2830	-5.503	-8.199
	(1.468e-05)	(11.50)	(89.78)	(103.5)	(129.2)
δ_1			-20,510***	215.6	-631.7
			(7,345)	(8,566)	(16,390)
δ_3				20,540**	-634.5
				(7,576)	(14,520)
δ_4					20,550**
					(7,757)
Constant	0.1832	324.5	20,540***	1.140	636.1
	(0.2371)	(326.3)	(6,839)	(2,969)	(14,190)
Thresholds					
b_1			23.34	-0.5184	-8.618
b_2				60.57	0.2002
b_3					60.57
Intercept ordering	N/A	N/A	Fail	Fail	Fail
Schwarz Criterion	N/A	-270.7	-335.2	-338.6	-342
Counties	30	30	30	30	30
R-squared		0.001	0.225	0.225	0.225

Back

Case set 3 multiple equilibria tests: ind. shares

	$s_{c(j),20} - s_{c(j),5}$	$s_{c(j),20} - s_{c(j),-1}$			
	IV Estimate	1 Equilibrium	2 Equilibria	3 Equilibria	4 Equilibria
	(1)	(2)	(3)	(4)	(5)
Growth Rate $\tau = -5$	7.444e-04	-0.02986	0.4695	1.076	1.118
to $\tau = 0$	(0.009371)	(2.405)	(20.67)	(20.60)	(21.05)
MDP Shock	8.804e-05***				
	(2.746e-05)				
δ_1			-2,655***	-3.838	16.42
			(1,009)	(717.2)	(1,613)
δ_3				3,915***	18.83
				(1,263)	(1,639)
δ_4					3,915***
					(1,265)
Constant		41.55	2,654***	0.8327	-18.03
		(39.67)	(941.8)	(549.4)	(1,547)
Threshholds					
b_1			59.52	5.329	4.033
b_2				86.36	5.307
b_3					86.36
Intercept ordering	N/A	N/A	Pass	Fail	Fail
Schwarz Criterion	N/A	-1887	-2408	-2412	-2417
Counties		250	250	250	250
R-squared		0.000	0.028	0.044	0.044

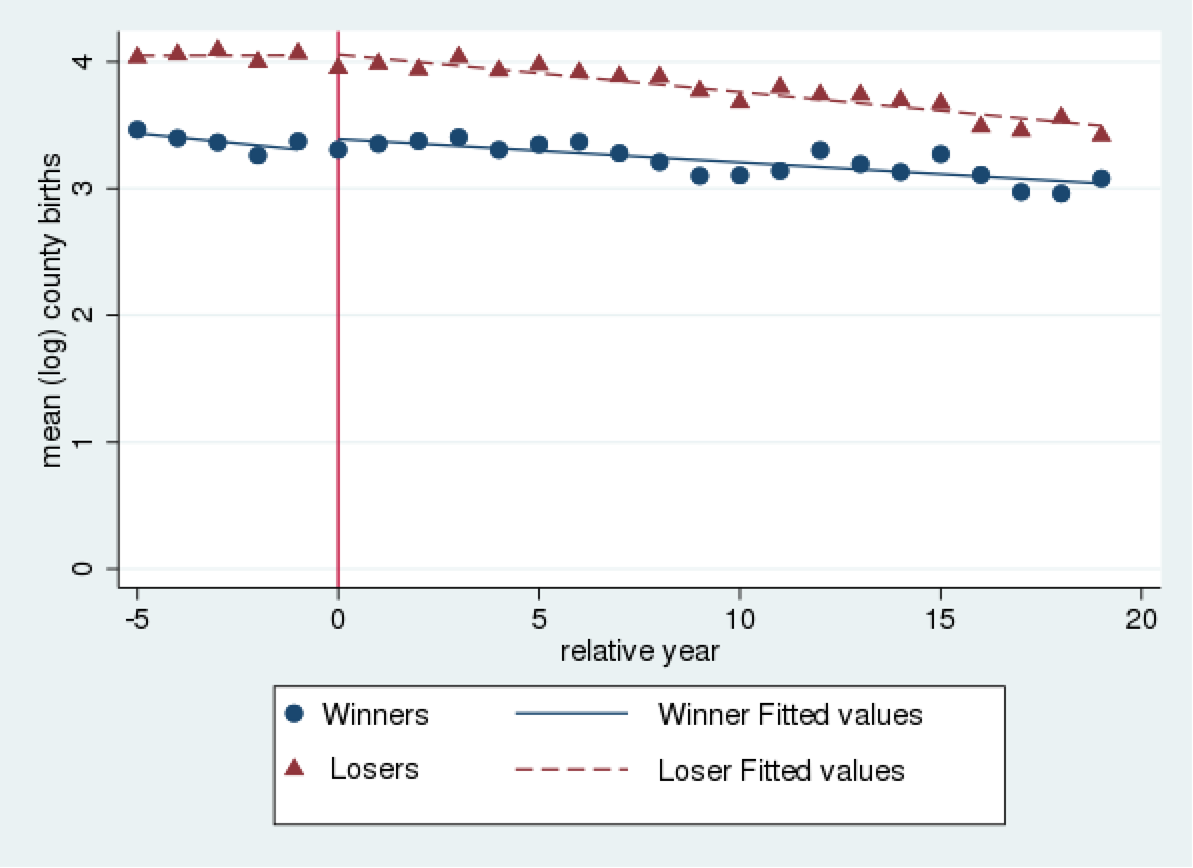
Back

Case set 4 multiple equilibria tests: mfg. shares

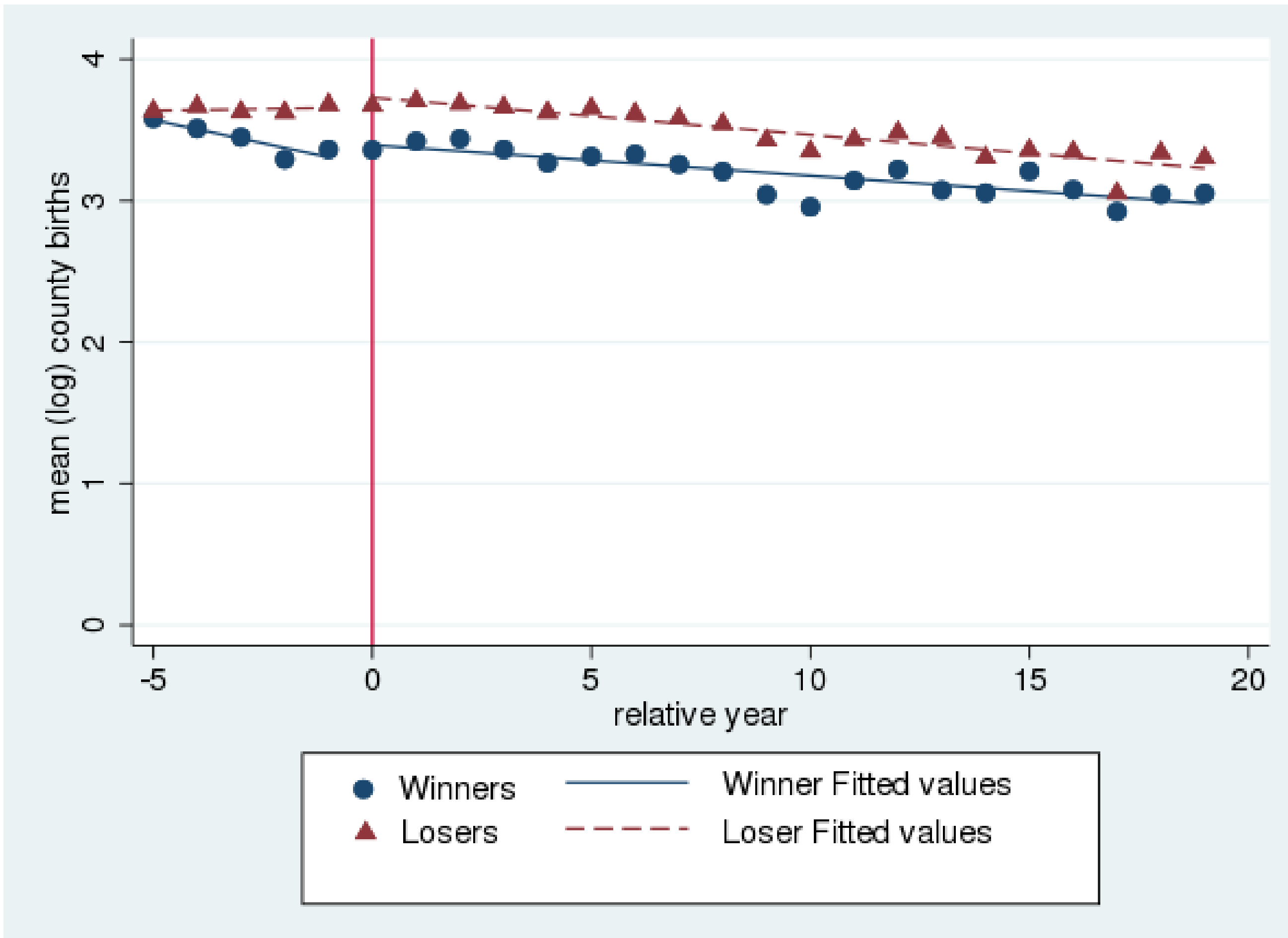
	$s_{c(j),20} - s_{c(j),5}$	$s_{c(j),20} - s_{c(j),-1}$			
	IV Estimate	1 Equilibrium	2 Equilibria	3 Equilibria	4 Equilibria
	(1)	(2)	(3)	(4)	(5)
Growth Rate $\tau = -5$	-0.002376	-0.001606	-0.002849	-0.003307	-0.003346
to $\tau = 0$	(0.009062)	(0.009231)	(0.003995)	(0.004011)	(0.004013)
MDP Shock	-0.03428				
	(0.1258)				
δ_1			-2.368***	-1.003*	0.4592
			(0.4885)	(0.6017)	(0.8238)
δ_3				1.910**	1.451
				(0.7861)	(0.9758)
δ_4					1.900**
					(0.7784)
Constant	0.5476	0.8182**	2.538***	1.137**	-0.3042
	(0.4252)	(0.4056)	(0.4513)	(0.5717)	(0.7990)
Thresholds					
b_1			2.051	1.512	1.153
b_2				3.017	1.512
b_3					3.017
Intercept ordering	N/A	N/A	Pass	Fail	Fail
Schwarz Criterion	N/A	-839.9	-632.2	-636.1	-641.4
Counties	250	250	250	250	250
R-squared		0.000	0.087	0.098	0.100

Back

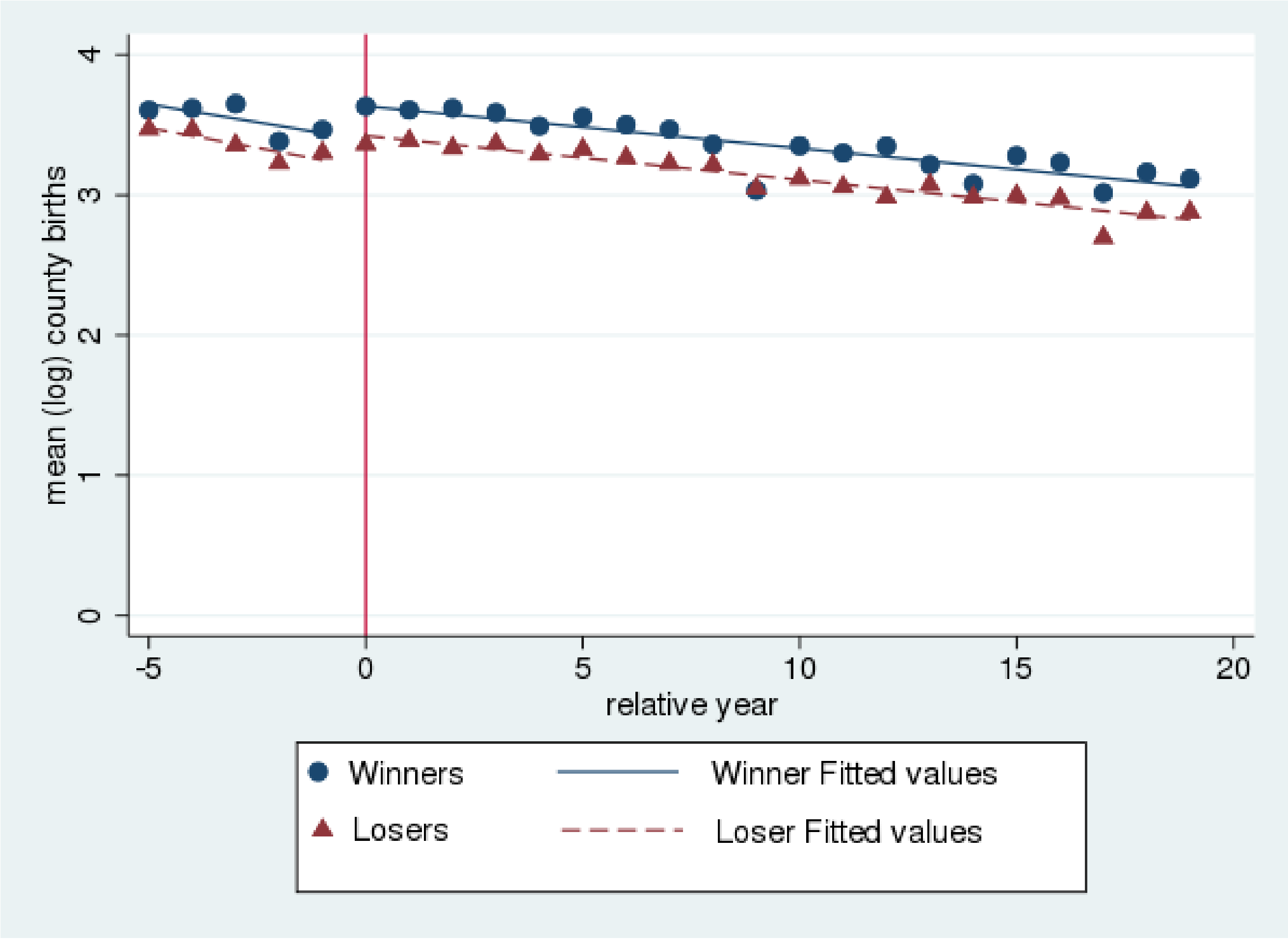
Birth Trends: GHM Winners and Losers



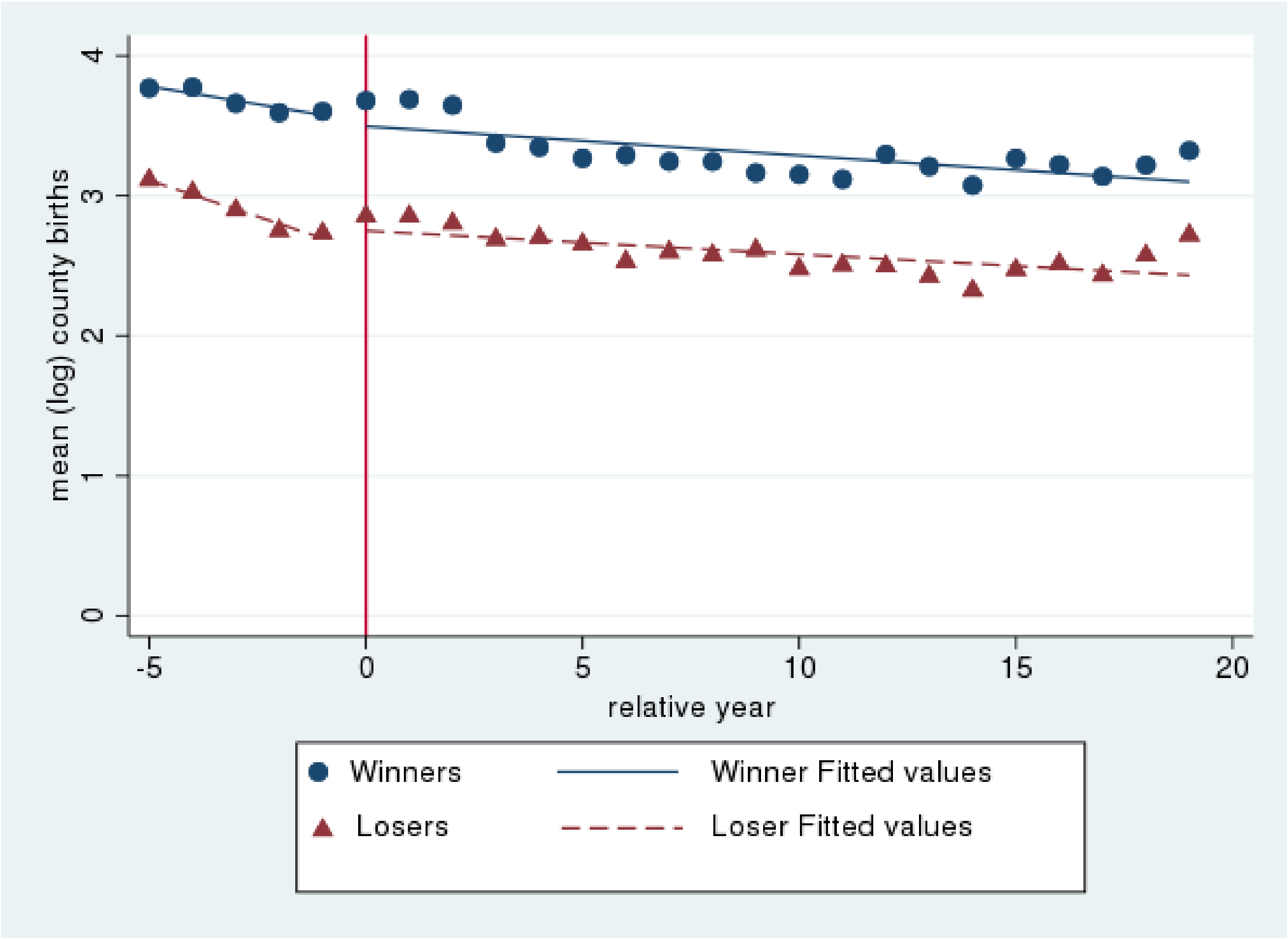
Birth Trends: Case set 1, GHM winners and propensity score losers



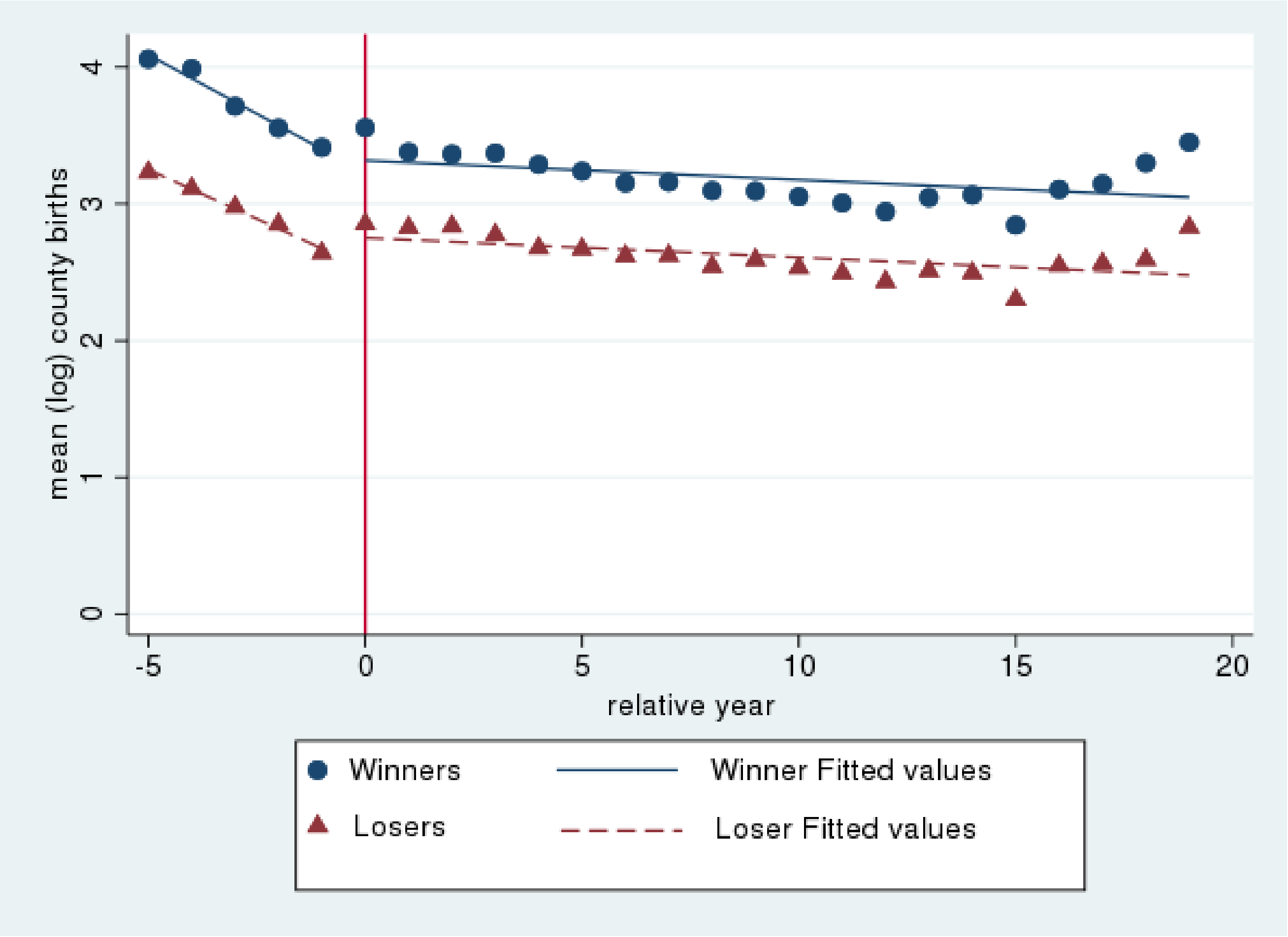
Birth Trends: Case set 2



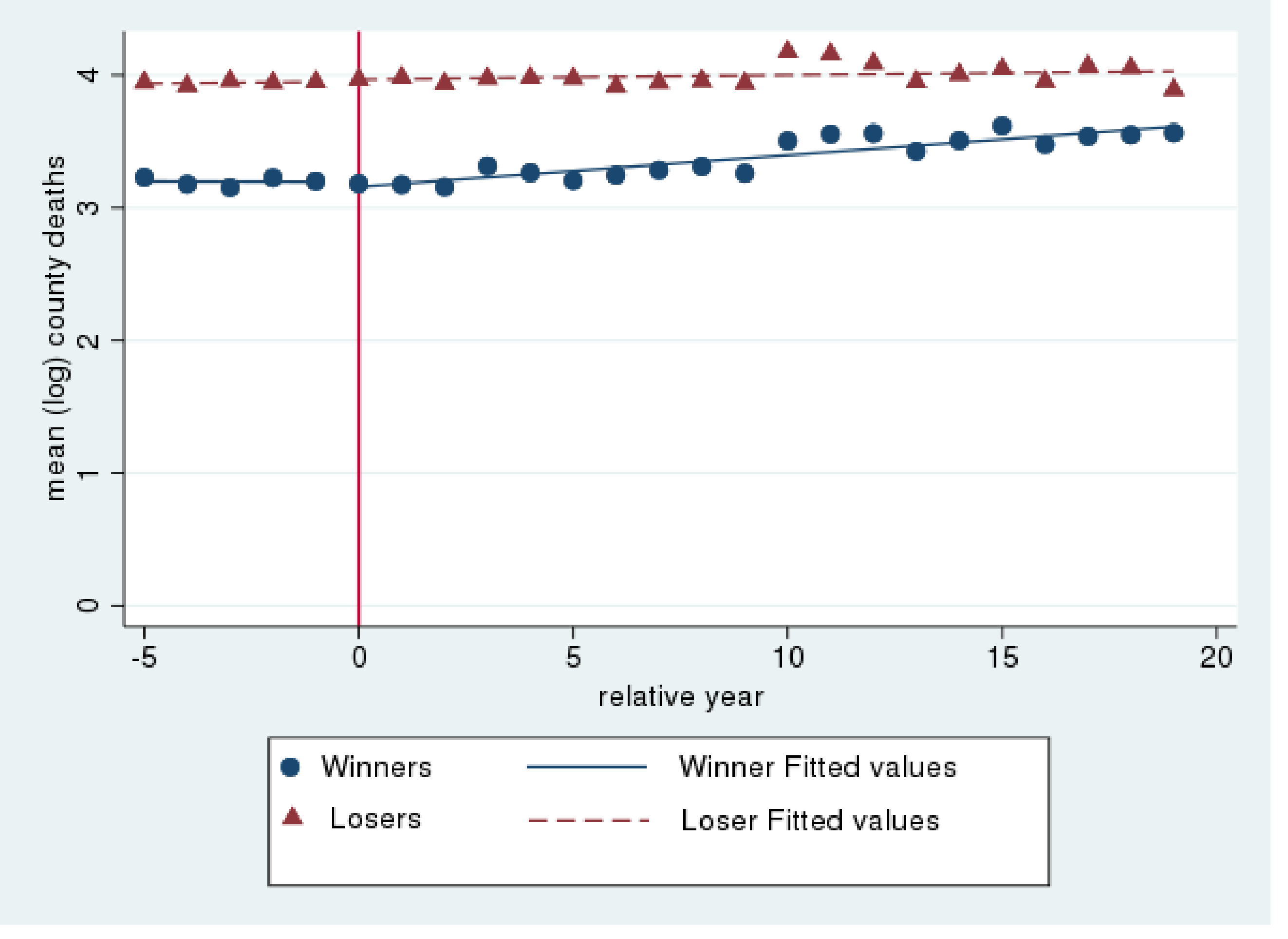
Birth Trends: Case set 3



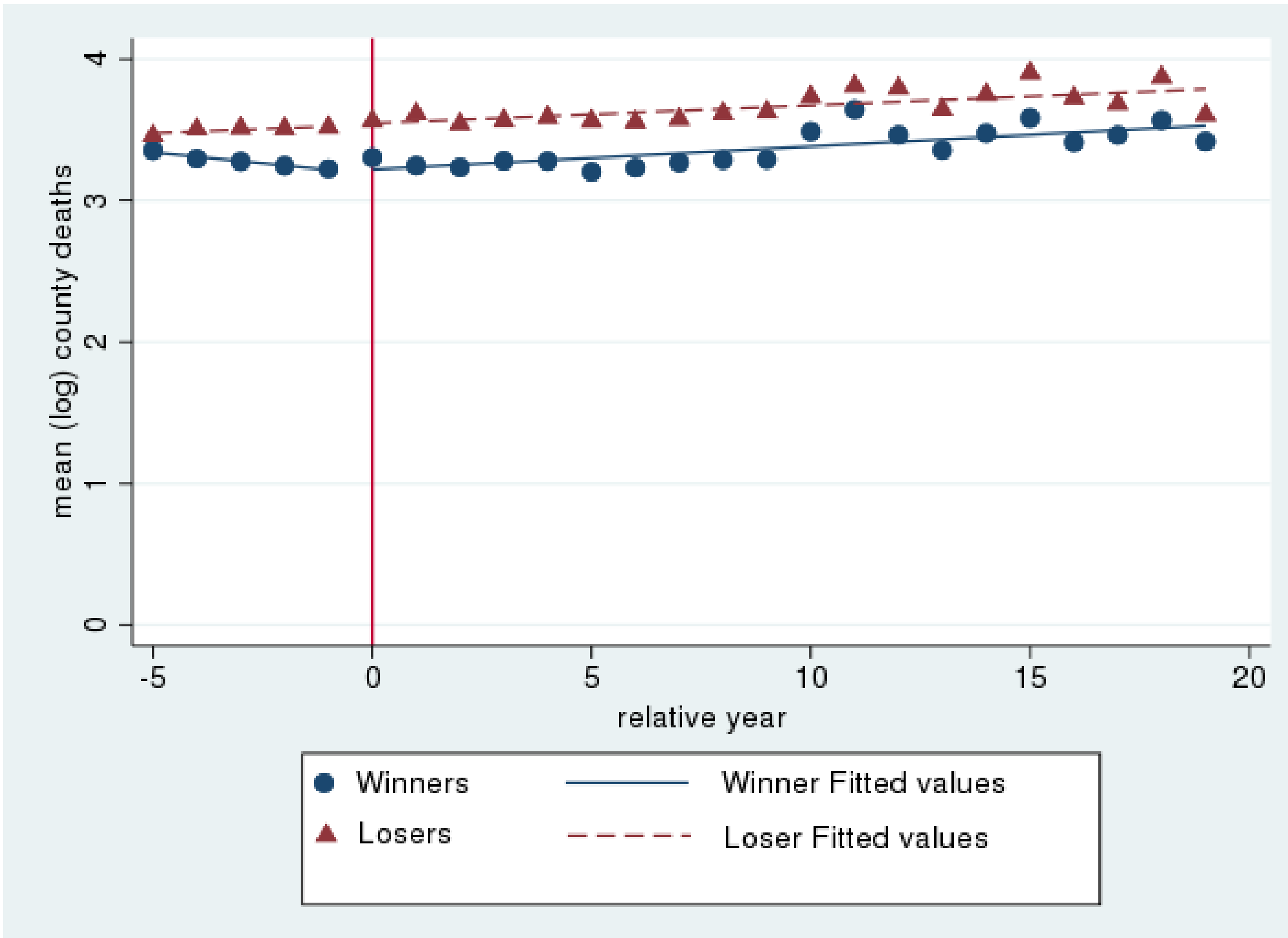
Birth Trends: Case set 4



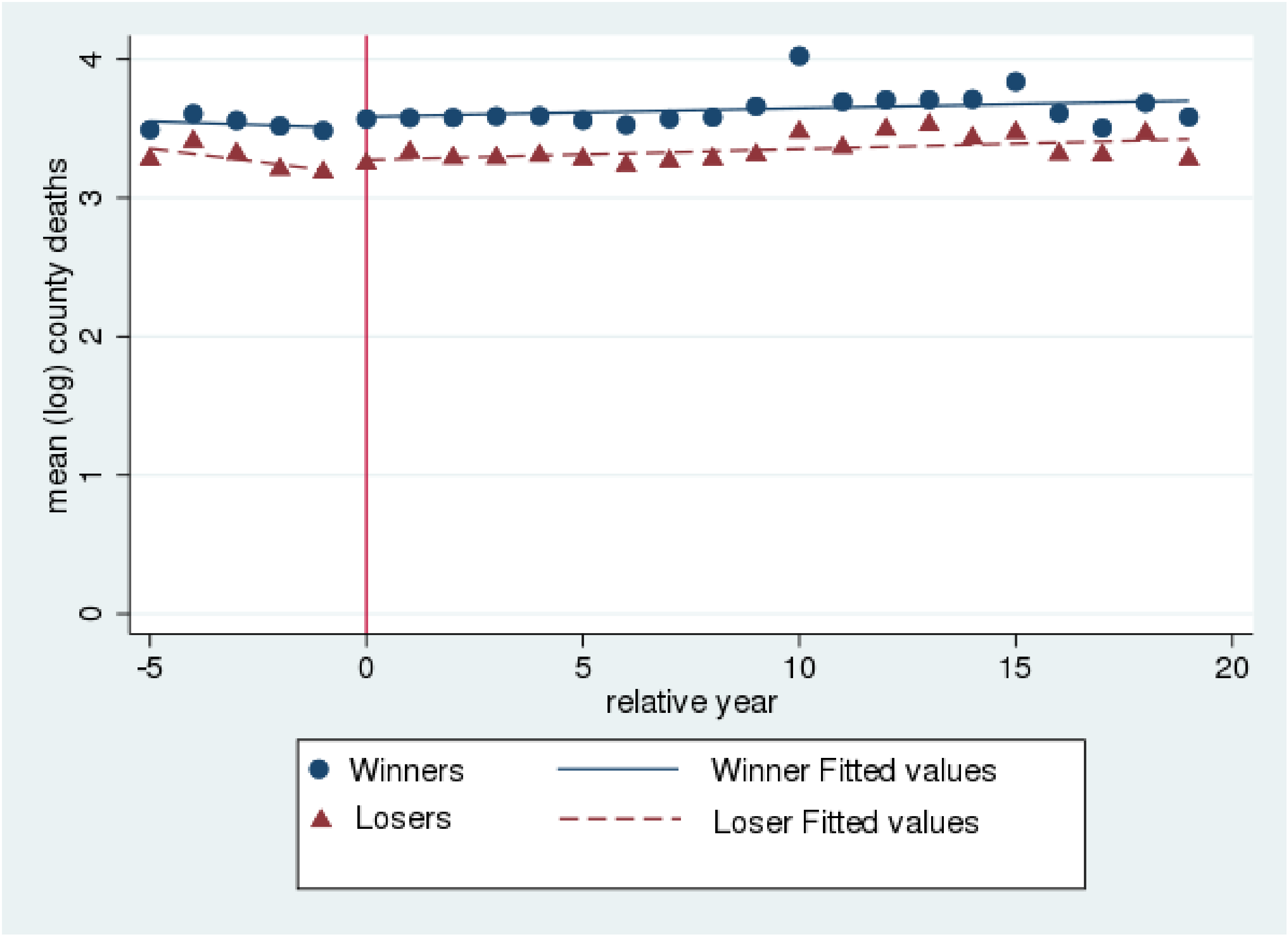
Death Trends: GHM Winners and Losers



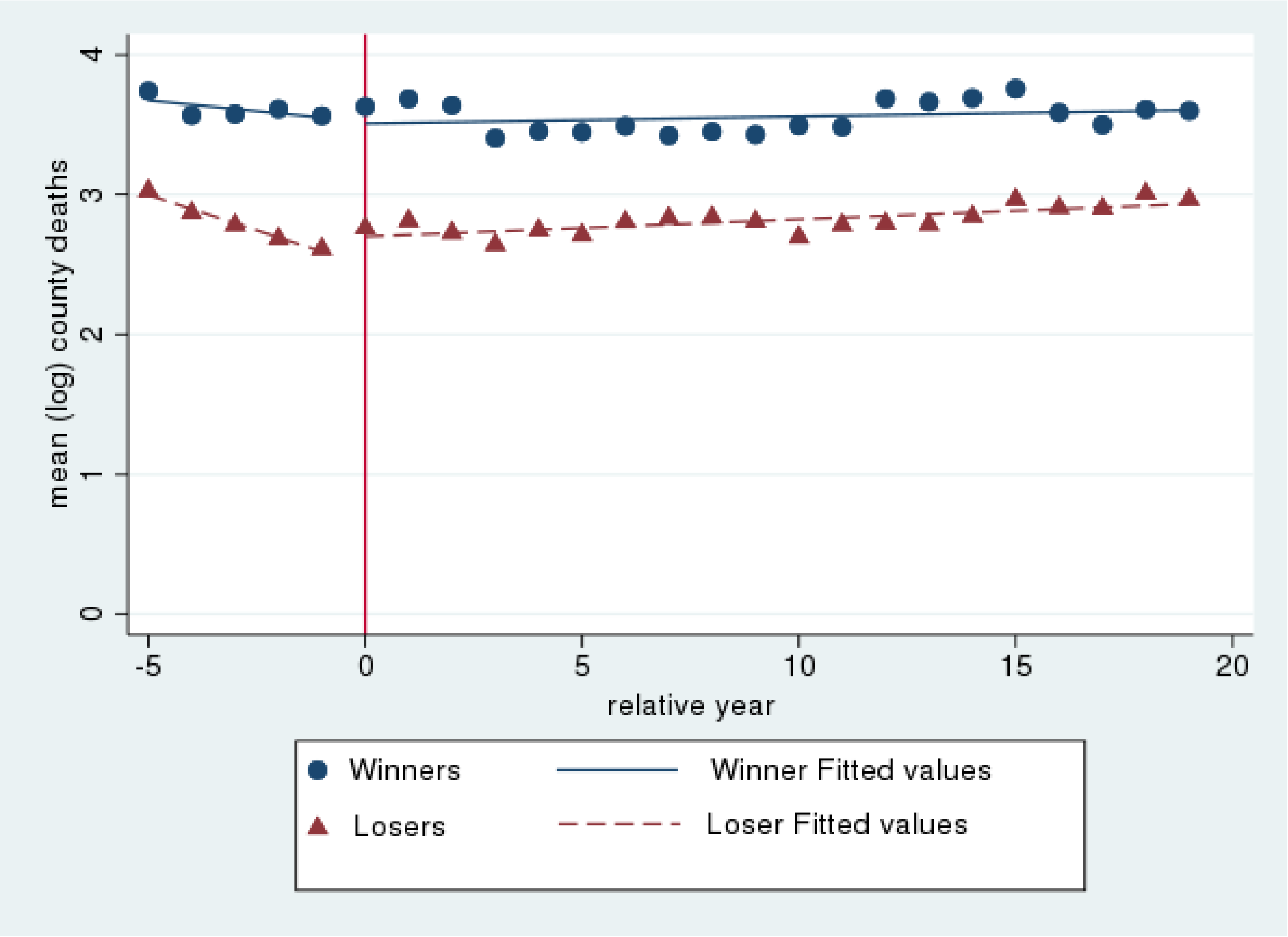
Death Trends: Case set 1, GHM winners and propensity score losers



Death Trends: Case set 2



Death Trends: Case set 3



Death Trends: Case set 4