Causal data analysis

Chapter 22: Difference-in-Differences

Álmos Telegdy

Corvinus University of Budapest

Endogeneity

The **estimated** effect of x on y has two components

→ The causal effect of x on y

→ The selection effect of x on y (endogenity)

Endogenity is hard (impossible) to capture with control variables

- Only observed heterogeneity can be captured
- Unobserved heterogeneity may be substantial
  - Some confounders are not in the data
  - Some mechanisms causing endogeneity are not uncovered

Blood pressure and eating healthy food

Reverse causality: the doctor advises to high BP patients to eat healthy food → bias

Blood pressure and eating healthy food

Reverse causality: the doctor advises to high BP patients to eat healthy food $\rightarrow$ bias

Imagine we augment the data from **the period before** the doctoral consultation

- Two data points
  - BP before eating healthy food
  - BP after eating healthy food

Blood pressure and eating healthy food

Reverse causality: the doctor advises to high BP patients to eat healthy food $\rightarrow$ bias

Imagine we augment the data from **the period before** the doctoral consultation

- Two data points
  - BP before eating healthy food
  - BP after eating healthy food
- We can gauge the impact of beginning to eat healthy food
  - Can get rid of a lot of confounders: **for the same person** we measure BP under two regimes

Solution to reducing selection bias: use better data

- What about having information about subjects **before** treatment?
  - One can control for pre-treatment differences between subjects
- Need information on the same subject from multiple time periods ⇒ **panel data**
  - For now, periods: before/after treatment

Difference-in-differences estimation

- Compares the average of the outcome among units...
- ...in the treatment group and in the non-treatment group
- ...before and after the treatment

Difference-in-differences

	Untreated	Treated	Diff: Treated-Untreated
Before	$\bar{y}_{untreated,before}$	$\bar{y}_{treated,before}$	$\bar{y}_{treated,before} - \bar{y}_{untreated,before}$
After	$\bar{y}_{untreated,after}$	$\bar{y}_{treated,after}$	$\bar{y}_{treated,after} - \bar{y}_{untreated,after}$
Diff: After-Before	$\Delta\bar{y}_{untreated}$	$\Delta\bar{y}_{treated}$	$\Delta\bar{y}_{treated} - \Delta\bar{y}_{untreated}$

$$\hat{\beta}_{DiD} = (\bar{y}_{\text{ treated, after}} - \bar{y}_{\text{ treated, before}}) - (\bar{y}_{\text{ control, after}} - \bar{y}_{\text{ control, before}}) \tag{1}$$

Intuition

$\Delta_1 = (\bar{y}_{\text{treated, after}} - \bar{y}_{\text{treated, before}}) \longrightarrow$ removes heterogeneity within the treated group

$\Delta_0 = (\bar{y}_{\text{control, after}} - \bar{y}_{\text{control, before}}) \longrightarrow$ controls for changes in the environment

$\hat{\beta}_{DiD} = \Delta_1 - \Delta_0 \longrightarrow$ what is the average change in the outcome variable in the group of treated subjects relative to the group of control subjects

Example: privatization in the 90s

Question: does privatization decrease firm employment?

- $\Delta_1 = \bar{L}_{\text{priv, after}} - \bar{L}_{\text{priv, before}}$
  - $\Delta_1 < 0$; firms fire workers because efficiency reasons (as a result of privatization) and because the recession
- $\Delta_0 = \bar{L}_{\text{state, after}} - \bar{L}_{\text{state, before}}$
  - $\Delta_0 < 0$; firms fire workers because of the recession
- $\Delta_1 - \Delta_0 \approx 0$

Difference-in-differences with regression

$$\Delta y^E = \alpha + \beta x \tag{2}$$

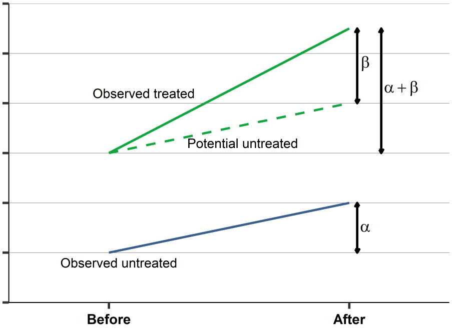
Difference-in-differences with regression

$$\Delta y^E = \alpha + \beta x \quad (2)$$

Coefficients

- α: average change of y in the control group ($x = 0$)
- $\alpha + \beta$: average change of y in the treated group ($x = 1$)
- β: average change of y in the treated group *relative to* the average change in the control group

The diff-in-diff graph



- $\alpha =$ average change in the control group (Δ_0)
- $\alpha + \beta =$ average change in the treated group (Δ_1)
- $\beta =$ the difference between the two changes ($\Delta_1 - \Delta_0$)

Identification assumption

The condition for a causal interpretation: **the change of y in the control group must be the a valid counterfactual of the change of y in the treated group**

- If y changed in the same way in the treated and control group if there were not treatment...
- ...then β is an unbiased estimator of ATE

This is a less stringent condition than random assignment: the condition refers only to Δy and not the level of y

Case study: airline merger's effect on ticket prices

- Two large US airlines merged
  - 2011: American Airlines filed for bankruptcy in November 2011
  - 2012: US Airways takeover bid
  - 2015: merger approved by the Competition Authority
  - 2015: booking web pages merged
- Expected effects of the merger

Case study: airline merger's effect on ticket prices

- Two large US airlines merged
  - 2011: American Airlines filed for bankruptcy in November 2011
  - 2012: US Airways takeover bid
  - 2015: merger approved by the Competition Authority
  - 2015: booking web pages merged
- Expected effects of the merger
  - Scale effect → larger company → more efficient company → **lower price**
  - Market power → **higher price**
- Analysis
  - Ex ante (authority) – price increase unlikely
  - Ex post (here): **what is the effect of the merger on ticket prices?**

Data

US Dept. for Transportation: a 10% sample of all ticket transactions ⇒ **Big Data** – 3 – 3.5 million observations quarterly (15GB, very complex)

- Observation: airline x route x year x quarter
- Variables: airport, price, number of passengers, airline
- Before= 2011; after = 2016

Definition of market

- Aggregate at the market level
  - Defined by their origin and destination airport, and whether they are one-way or return routes (e.g., New York – San Francisco)
  - Alternative def: take into account middle airports (e.g., New York – Atlanta, Atlanta – San Francisco)
  - Return ticket: if destination airport not known, drop from the data
- 460 airports (2011): 460x459=211,140 markets, in reality 141,712
  - median = 7 passengers; average = 140 passengers
- Small and large markets important for consumer surplus, but competition is different – separate analysis
  - Small market: median/mean = 7/100 (99% of markets, 60% of passengers)
  - Large market: median/mean = 8,000/10,000
- Unbalanced panel: 30 000 markets before **or** after treatment only (1% of passengers) → drop from the analysis

Treatment variable

- Treated market = American Airlines **and** U.S. Airways present in 2011
  - These are the markets where the merger can raise market power
- Control market= Neither AA, nor US present in 2011
  - Treatment depends on market characteristics before treatment
- If only AA or US present on the market in 2011 $\rightarrow$ drop from the sample
- 12 000 treated markets (14 million tickets), 72 000 control markets (4 million tickets)

Regression equation

Outcome variable: average ticket price

$$(\Delta \ln \bar{p})^E = \alpha + \beta AAUS_{before} \tag{3}$$

Diff-in-diff estimation

VARIABLES	(1) All markets	(2) Small markets	(3) Large markets
$AAUS_{before}$	-0.18**	-0.16**	-0.26**
	(0.01)	(0.01)	(0.03)
Constant	0.16**	0.14**	0.24**
	(0.01)	(0.01)	(0.02)
Observations	112,632	111,745	887
R-squared	0.05	0.04	0.09

Results

- In untreated markets, the average ticket price increased by 16% between 2011 and 2016
- Tickets increased more on large markets than on small markets (24% and 14%)
  - Prices not deflated, the increase is partly because inflation
- Price increase on treated markets 18% **lower** on average
- Treatment effect similar on small and large markets (reduces total price increase to approximately 0)

Diff-in-diff table

Regression results are identical to computing averages

	(1) Untreated	(2) Treated	(3) Difference: Treated - Untreated
Before	4.92	4.96	+0.04
After	5.08	4.94	-0.14
Difference: After - Before	+0.16	-0.02	-0.18

Note: Average log price, weighted by the number of passengers at baseline.

When is $\hat{\beta}_{DiD}$ causal?

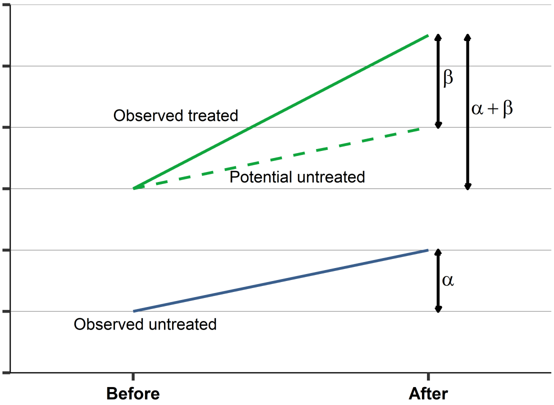
- Taking differences takes care of problems related to **levels** $\Rightarrow$ DiD estimation allows for selection and reverse causality related to **levels**
  - Whichever unit is treated may be related to initial levels of outcome
- As long as treatment is not related to changes, the measured effect is unbiased

Parallel trends assumption I

Identification assumption: in absence of the treatment, the treated group's average outcome change would be identical to the control group's average outcome change $\Rightarrow$ **parallel trends assumption**

$$\beta_{DiD} = ATET$$

PTA graph



Observed treated
Potential untreated
Observed untreated
β
α + β
α
Before
After

Parallel trends assumption II

Additional assumption: under treatment, the control group's average outcome change would be identical to the treated group's average outcome change

- If this holds, $\beta_{DiD} = ATE$
- This assumption usually not required
  - $ATET$ more interesting than ATE
  - Under some conditions, parallel trends assumption is verifiable to some extent (if we have data on the outcome variable before the treatment)

Trends before the treatment

Parallel trends assumption (PTA) is an **identification assumption** $\Rightarrow$ cannot be verified with data

- For this, observed outcomes should be compared to the counterfactual (which is unobserved)

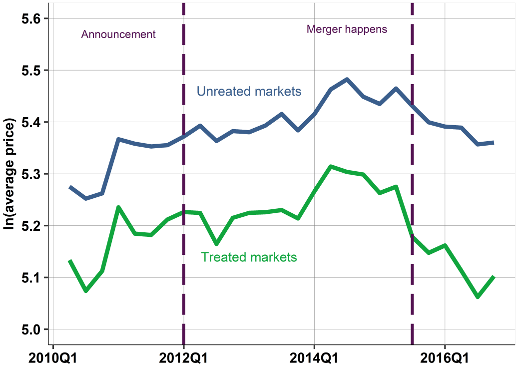
BUT: indirect evidence: look at **pre-treatment trends**

- What is the relation of the outcome variable between treated and control groups in the pre-treatment period?
    - For this, we need more than one pre-treatment data point

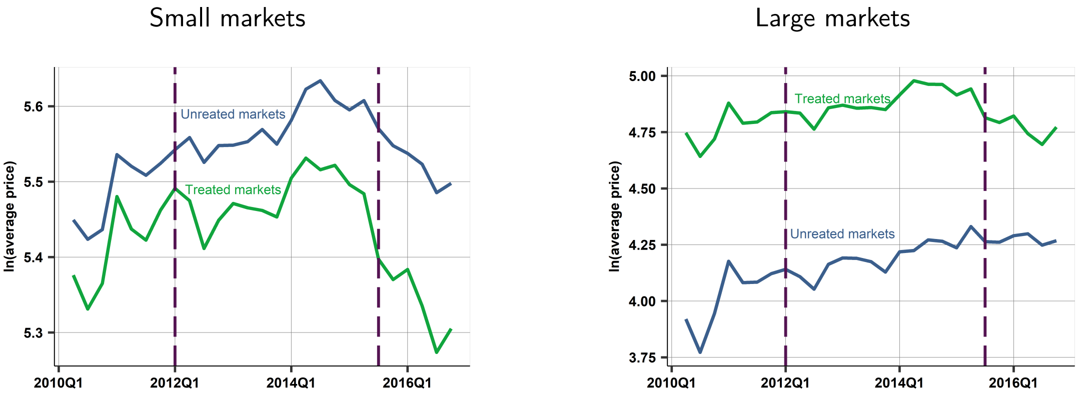
Case study

- PTE: without merger, prices would have changed in the same way on treated and control markets
- Indirect test of PTE: pre-trends
  - First data point: 2010 Q2
  - Merger announcement: 2012 Q4; merger: 2015 Q3

Pre-trends in the treated and control groups



Pre-trends in small and large markets



What does the figure suggest?

- The two trends move together but there are differences, especially in large markets
- The two lines fan out after 2015 – differences much smaller before treatment than after
- Are pre-trends parallel?
  - Not exactly
  - But they have very similar patterns pre-treatment
  - Pre-treatment differences much smaller, than the estimated effect
- Evidence not bullet proof, but okay
  - Possible that the magnitude of the true effect is different from the estimated effect
  - Strong indication that the merger decreases prices

Use of control variables

If PTA not valid, $\hat{\beta}_{DiD}$ will be biased

Attenuation of the bias: use control variables

$\Rightarrow$ PTA has to be true only *conditionally*

$\Rightarrow$ The pre-trend of the treated group has to be similar to the control group's but only after we partialled out the effect of control variables

Example: effect of privatization on employment

- Employment changes of firms depend on the industry ($IND \rightarrow y$)
- In some industries privatization was more intensive ($IND \rightarrow x$)
  - Without industry controls, PTA does not hold
- If we control for industry, PTA will be correct again
  - If there are no other confounders

Regression equation with controls

General form of diff-in-diff regression with $z_1,\ z_2, \ldots$ control variables

$$\Delta y^E = \alpha + \beta \cdot Treated + \gamma_1 z_1 + \gamma_2 z_2 + \ldots \tag{4}$$

$\Rightarrow \beta$ equals the average chance of y between the treated and control groups

Types of controls

Panel structure provides more possibilities than cross sectional data

- z may come from the pre-treatment period
- z may come from the post-treatment period
- z may be the difference between post- and pre-treatment

Which to choose? It depends on the type of endogeneity

- Post-treatment z may be part of the mechanism (bad control)
- Δz good control, if we believe that selection was along changes
  - E.g., the new owner knows that the firm sales are on a growth trajectory

Case study: the use of control variables

- What is behind the variability of the treatment variable?
  - Why was AA and US present in some markets, but not on others?

Case study: the use of control variables

- What is behind the variability of the treatment variable?
  - Why was AA and US present in some markets, but not on others?
- Siz of the market
  - More likely that on big markets both airlines are present
  - Passengers on large markets more likely to be wealthy (they link big cities)
- Competition on the market (the number of airlines)
  - Competition decreases prices
- Route features
  - Airport hub
  - How many cities are in between
  - Return market or not

Regression with control variables

Variables	(1) All markets	(2) Small markets	(3) Large markets
$AAUS_{before}$	-0.11**	-0.10**	-0.13**
	(0.01)	(0.01)	(0.03)
ln $no.passengers_{before}$	-0.00	0.00	0.06**
	(0.00)	(0.00)	(0.02)
Return route	0.19**	0.20**	0.17**
	(0.01)	(0.01)	(0.03)
Number of stops	-0.03**	0.00	-0.07**
	(0.01)	(0.01)	(0.02)
Share of largest carrier	0.26**	0.21**	0.43**
	(0.02)	(0.02)	(0.07)
Constant	-0.15**	-0.17**	-0.74**
	(0.03)	(0.02)	(0.21)
Observations	112,632	111,745	887
R-squared	0.14	0.11	0.23

Treatment variable continuous

Instead of binary x use quantitative $x \rightarrow$ the regression equation is the same, but the meaning of the coefficients is different

1. x measured before treatment

$$\Delta y^E = \alpha + \beta x_{\text{before}} \tag{5}$$

Treatment variable continuous

Instead of binary x use quantitative $x \rightarrow$ the regression equation is the same, but the meaning of the coefficients is different

1. x measured before treatment

$$\Delta y^E = \alpha + \beta x_{\text{before}} \tag{5}$$

α = average change of y when $x = 0$
β = average change of y between two subjects which differ in x by one unit

Treatment variable continuous

Instead of binary x use quantitative $x \rightarrow$ the regression equation is the same, but the meaning of the coefficients is different

1. x measured before treatment

$$\Delta y^E = \alpha + \beta x_{\text{before}} \tag{5}$$

α = average change of y when $x = 0$
β = average change of y between two subjects which differ in x by one unit

2. x is the difference between before and after treatment

$$\Delta y^E = \alpha + \beta \Delta x \tag{6}$$

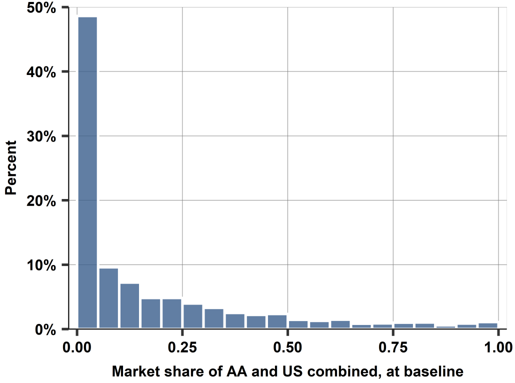
α = average change of y if x constant
β = average change of y between two subjects which differ in Δx by one unit

Case study: continuous treatment

- So far: AA and US present on the market before treatment
- Continuous treatment: $x =$ share(AA)+share(US) before treatment
  - Instead of two states of treatment ($x \in \{0,1\}$) we define many states of treatment ($x \in (0,1)$)
- Regression equation:

$$
\begin{aligned}
(\Delta \ln p)^E = \beta_0 + \beta_1 AAUSshare_{baseline} + \beta_2 \ln passengers_{baseline} \\
+\beta_3 return + \beta_4 stops + \beta_5 sharelargest_{baseline}
\end{aligned}
\tag{7}
$$

The distribution of the joint share of AA and US on markets



- Markets are weighted by the number of passengers

Regression results

Variables	(1) All markets	(2) Small markets	(3) Large markets
Market share before	-0.27**	-0.17**	-0.42**
	(0.02)	(0.02)	(0.05)
ln no. passengers before	-0.01**	-0.01**	0.05*
	(0.00)	(0.00)	(0.02)
Return route	0.21**	0.21**	0.19**
	(0.01)	(0.01)	(0.02)
Number of stops	-0.03**	0.00	-0.07**
	(0.01)	(0.01)	(0.02)
Share of largest carrier	0.31**	0.26**	0.47**
	(0.02)	(0.02)	(0.05)
Constant	-0.12**	-0.15**	-0.72**
	(0.03)	(0.02)	(0.20)
Observations	112,632	111,745	887
R-squared	0.15	0.11	0.30

Continuous treatment: interpretation

- Prices declined by 27% in markets where AA+US had a 100% share
- A difference of 10% in the joint share of AA and US generated a price decline of 2.7%

Continuous treatment: interpretation

- Prices declined by 27% in markets where AA+US had a 100% share
- A difference of 10% in the joint share of AA and US generated a price decline of 2.7%
  - The effect may not be linear in x

RECAP: Repeated observations of the same unit

So far

⇒ Same units observed before and after the treatment (individual, family, firm, market, country...)

- Simplest panel data structure: each unit observed twice
- Basic diff-in-diff applies to such data

Different units before and after treatment

What if there are no data on the treated subjects before treatment?

- If there are data from **different subjects before and after treatment**...
  - we can estimate a β_{DiD}...
  - ...but we need further assumptions
    - *Weaker data* $\Rightarrow$ *stronger assumptions*

- Compose 4 groups
  - Treated/untreated; before/after
  - Not the same units before and after, but similar ones

    $\Rightarrow$ BUT: prove they are similar

Example: building a factory close to a residential neighborhood

What is its effect on house prices?

- Not possible to observe twice the price of the same house (sales not frequent) $\Rightarrow$ cannot observe the same unit before and after treatment
- Four groups
  1. **Treated before**: p of houses close by, sold before factory building started
  2. **Treated after**: p of houses close by, sold after factory building started
  3. **Control before**: p of houses farther away, before factory building started
  4. **Control after**: p of houses farther away, after factory building started
- We can get β_{DiD} from these four groups

Estimation

- We estimate the effect from the four groups' average y

$$\beta_{DiD} = (\bar{y}_{\text{treatment, after}} - \bar{y}_{\text{treatment, before}}) - (\bar{y}_{\text{no treatment, after}} - \bar{y}_{\text{no treatment, before}}) \tag{8}$$

- Treated and control groups have different subjects before-after treatment

Pooled cross section: identification assumptions

- Same as so far: PTA
  - In absence of treatment, the average value of the outcome variable would have changed in the same way in the treated and control groups
- NEW: selection (different observations before and after)
  - **Treated and control subjects similar before and after treatment**

Regression form

In regression form the outcome variable is not Δy, because it cannot be computed

- Use each observation separately
  - As in cross sectional data
- *Treated* binary variable (time invariant)
  - Subject i belongs to the treatment group if $Treated = 1$
  - Subject i belongs to the control group if $Treated = 0$
- T binary variable (time variant)
  - Subject i observed after treatment if $T = 1$
  - Subject i observed before treatment if $T = 0$

Regression equation

$$y^E = \alpha + \beta\, Treated_i + \gamma\, T_t + \delta\, Treated_i \cdot T_t \tag{9}$$

Regression equation

$$y^E = \alpha + \beta\, Treated_i + \gamma\, T_t + \delta\, Treated_i \cdot T_t \tag{9}$$

(1) $E[y|Treated = 0, T = 0] = \alpha$

Regression equation

$$y^E = \alpha + \beta\, Treated_i + \gamma\, T_t + \delta\, Treated_i \cdot T_t \tag{9}$$

(1) $E[y|Treated = 0, T = 0] = \alpha$

(2) $E[y|Treated = 0, T = 1] = \alpha + \gamma$

Regression equation

$$y^E = \alpha + \beta \, Treated_i + \gamma \, T_t + \delta \, Treated_i \cdot T_t \tag{9}$$

(1) $E[y|Treated = 0, T = 0] = \alpha$

(2) $E[y|Treated = 0, T = 1] = \alpha + \gamma$

(3) $E[y|Treated = 1, T = 0] = \alpha + \beta$

Regression equation

$$y^E = \alpha + \beta \, Treated_i + \gamma \, T_t + \delta \, Treated_i \cdot T_t \tag{9}$$

(1) $E[y|Treated = 0, T = 0] = \alpha$
(2) $E[y|Treated = 0, T = 1] = \alpha + \gamma$
(3) $E[y|Treated = 1, T = 0] = \alpha + \beta$
(4) $E[y|Treated = 1, T = 1] = \alpha + \beta + \gamma + \delta$

Regression equation

$$y^E = \alpha + \beta\, Treated_i + \gamma\, T_t + \delta\, Treated_i \cdot T_t \tag{9}$$

(1) $E[y|Treated = 0, T = 0] = \alpha$
(2) $E[y|Treated = 0, T = 1] = \alpha + \gamma$
(3) $E[y|Treated = 1, T = 0] = \alpha + \beta$
(4) $E[y|Treated = 1, T = 1] = \alpha + \beta + \gamma + \delta$

- $\Delta_0 = (2) - (1) = \gamma$
- $\Delta_1 = (4) - (3) = \gamma + \delta$
- $\beta_{DiD} = \Delta_1 - \Delta_0 = \delta$

The two regression equations are interchangable

We can use this equation if the same units are observed before-after treatment.
Difference form:

$$\Delta y^E = \gamma^{'} + \delta^{'} \cdot x_{before} \tag{10}$$

Same estimated coefficients

- $\gamma \approx \gamma^{'}$
- $\delta \approx \delta^{'}$

Advantages – drawbacks

Advantages

- Can study a new set of problems:
  - No multiple period data for the same observations
    - Often impossible
- Can interact controls with time ($z \cdot T$)
- Larger sample

Disadvantages

- Need a new assumption
  - OLD: PTA
  - NEW: Groups similar before-after treatment
    - This is hard to prove (especially if not true...) $\Rightarrow$ selection

Case study: pooled cross section

Add to the balanced panel the unbalanced part

- Cannot add markets observed only after treatment (22 851 markets, 0.5 million passengers) $\Rightarrow$ treatment variable defined on the pre-treatment sample
- Add markets observed only before treatment (28 665 markets, 100,000 passengers)

Regression equation

$$(\ln p)^E = \alpha + \beta AAUS_{before} + \gamma \cdot After + \delta AAUS_{before} \cdot After \tag{11}$$

- Coefficients very similar to those when outcome variable in differences
  - Not surprising: weighted with no. passengers
- Add interactions between time and control variables

Regression results with controls interacted with time

Variables	All markets	Small markets	Large markets
$AAUS_{before} \times after$	-0.11**	-0.10**	-0.14**
	(0.01)	(0.01)	(0.03)
$AAUS_{before}$	0.43**	0.29**	0.69**
	(0.02)	(0.01)	(0.05)
$after$	-0.19**	-0.14**	-0.82**
	(0.04)	(0.03)	(0.22)
$\ln no.passengers_{before}$	-0.37**	-0.29**	-0.54**
	(0.01)	(0.00)	(0.05)
Return route	0.84**	0.84**	0.95**
	(0.02)	(0.01)	(0.05)
Number of stops	0.07**	0.14**	0.03
	(0.02)	(0.01)	(0.04)
Share of largest carrier	-1.45**	-1.40**	-1.69**
	(0.05)	(0.03)	(0.13)
$\ln passengers_{before} \times after$	-0.00	-0.00	0.06*
	(0.00)	(0.00)	(0.02)
Return route $\times after$	0.20**	0.20**	0.19**
	(0.01)	(0.01)	(0.03)
Number of stops $\times after$	0.01	-0.00	0.01
	(0.02)	(0.02)	(0.05)
Share of largest carrier $\times after$	0.30**	0.23**	0.44**
	(0.02)	(0.02)	(0.07)
Constant	7.85**	7.33**	9.23**
	(0.06)	(0.03)	(0.40)
Observations	254,178	252,404	1,774
R-squared	0.68	0.68	0.56