Title
A note on the triple difference in economic models

Journal
Applied Economics Letters, 23(4)

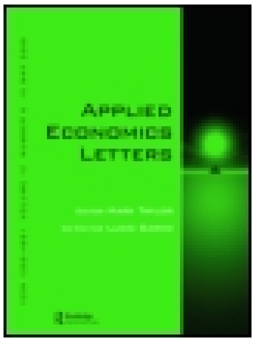
ISSN
1350-4851 1466-4291

Authors
Berck, Peter
Villas-Boas, Sofia B

Publication Date
2015-08-03

DOI
10.1080/13504851.2015.1068912

Peer reviewed



Routledge
Taylor & Francis Group

Applied Economics Letters

ISSN: 1350-4851 (Print) 1466-4291 (Online) Journal homepage: http://www.tandfonline.com/loi/rael20

A note on the triple difference in economic models

Peter Berck & Sofia B. Villas-Boas

To cite this article: Peter Berck & Sofia B. Villas-Boas (2015): A note on the triple difference in economic models, Applied Economics Letters, DOI: 10.1080/13504851.2015.1068912

To link to this article: http://dx.doi.org/10.1080/13504851.2015.1068912

APPLIED ECONOMICS LETTERS, 2015
http://dx.doi.org/10.1080/13504851.2015.1068912

Routledge
Taylor & Francis Group

A note on the triple difference in economic models

Peter Berck and Sofia B. Villas-Boas

Department of Agricultural and Resource Economics, University of California, Berkeley, CA, USA

ABSTRACT
This article shows when a triple difference strategy using an imperfect control category improves on the double difference strategy for estimating an average treatment effect. For example, a product is treated in one place and not another leading to a double difference strategy. When does comparison with an untreated product in triple difference strategy improve accuracy?

KEYWORDS
Difference in difference; triple difference; policy evaluation; average treatment effect

JEL CLASSIFICATION
C21; C25; C51

I. Introduction

Difference-in-differences estimation is one of the most important identification strategies in applied economics (Meyer 1995; Angrist and Krueger 2005; Bertrand, Duflo, and Mullainathan 2004; Athey and Imbens 2006). Exploiting a quasi-experimental panel design identification of the treatment effect uses a difference-in-differences econometric estimation strategy to measure the effect of a policy change on an outcome variable of interest. The difference-in-difference model measures the effect of policy by removing the effects of time and place. When the outcome variable is determined by policy, time, place and yet another variable, a triple difference strategy may reduce the bias in the estimate of the effect of the policy change. Published examples of a triple difference strategy include (1) comparing men's hours of labour before and after California extended its requirement for overtime after 8 hours worked in a day to men. The comparison to other states provided the double difference and the further comparison to women, the triple difference (Hamermesh and Trejo, 2000) and (2) the effect of daylight savings time on energy use was found by comparing energy use in Sydney when it expanded its use of daylight savings time for the Olympics with other states in Australia. The double difference is given by comparing time and place and the triple difference comes from comparing to mid-day where energy use is not affected by daylight savings time (Kellogg and Wolf 2008).

II. Model

We start by presenting the commonly known linear difference-in-differences model for a continuous and uncensored outcome. A policy change is observed in place $c = 1$, from time $t = 1$ forwards. The policy does not change in place $c = 0$. The policy in place 1 is $p = p_0$ and then from time 1 is p_1. For example, a tax is imposed on water bottles in Washington in July 2010 and it is not imposed in Oregon. Furthermore, the tax is imposed on good $i = 1$ and not on good $i = 0$, and in the example, the tax is imposed on water but not on juice. The outcome of interest is the number of water bottles sold, because the water bottles are a major ingredient in litter. For a second example, Sweden reduces its tax on restaurants while Norway does not. Moreover, the tax is changed on restaurant meals but not on hotels or on other retail. The outcome of interest is restaurant employment because the tax change was meant to increase employment.

The intuition behind the difference in differences is to measure the effect of the policy p on the outcome x_i in the treated group relative to changes in the outcome x_i in the control group. If the panel data have an untreated outcome $x_{i=0}$ that shares some possible common determinants with the outcome

CONTACT: Sofia B. Villas-Boas sberto@berkeley.edu

of interest $x_{i=1}$, another control structure can be used in a triple difference strategy to estimate the treatment effect of the policy change by comparing the double differences in $x_{i=1}$ with the double difference in $x_{i=0}$ and in so doing, controlling for more factors that could bias the average treatment effect.

These facts give the raw material for a triple difference: time $t<>1$, place $c = 0$, 1 and good $i = 0$, 1. But when is the triple difference, using variation from time, place and good, better than the double difference, just using time and place variation in the data to estimate the average treatment effect of the policy change?

The economic relationship at issue is the change in outcome x due to the policy p. As usual, the estimation is set up so that x is a function of time t, place c and the policy p. We add a possibly confounding variable y. For instance, in the examples, y would be income, and it varies by time and place and affects the demand for x (restaurant meals and water bottles). In this article, we maintain the hypothesis that the policy does not affect either the untreated place, time or good.

Let the true equation to be estimated be given by

$$x_{ict} = \alpha_c + \beta_i t + \gamma_i \ln p_{ict} + \delta_i \ln y_{ict}, \tag{1}$$

where the Greek letters are parameters to be estimated and in particular γ denotes the average effect of the policy on the outcome x. The model is written in logs because it is natural in the examples to think of growth rates, but the algebra will be much the same in the linear case. The variable δ_i is the semi-income elasticity of good i and it is not assumed that it is the same as the semi-elasticity of good j. Nor is it assumed that income growth was the same in both regions. Let Δ be the difference operator. In the case of time, the difference is taken from the initial value, where $t = 0$. Differencing out in time yields

$$\Delta_t x_{ict} = \beta_i t + \gamma_i \ln(p_{ict}/p_{ict=0}) + \delta_i \ln(y_{ict}/y_{ict=0}) \tag{2}$$

And additionally differencing out in space yields

$$\Delta_{tc}\mathrm{x}_{\mathrm{ict}} = \gamma_i \ln(p_{i,c=1,t}/p_{i,c=1,t=0}) + \delta_i\left[\ln(y_{i,c=1,t}/y_{i,c=1,t=0}) - \ln(y_{i,c=0,t}/y_{i,c=0,t=0})\right] \tag{3}$$

Letting

$$\Delta_{tc}\ln p = \ln(p_{i,c=1,t}/p_{i,c=1,t=0}) \text{ and } \Delta_{tc}\ln y = \left[\ln(y_{i,c=1,t}/y_{i,c=1,t=0}) - \ln(y_{i,c=0,t}/y_{i,c=0,t=0})\right] \tag{4}$$

we can write this more compactly as

$$\Delta_{tc}x_{ict} = \gamma_i \Delta_{tc}\ln p + \delta_i \Delta_{tc}\ln y \tag{5}$$

Equation 5 is the double difference specification. It depends upon the policy in region 0 being the same before and after $t = 1$, so that $\ln(p_{i,c=0,t}/p_{i,c=0,t=0}) = 0$. As set up, the double difference removes the effect of y whenever both regions have the same growth in y, which is in the example the same growth in income. If the data on y are exogenous, then the data can directly be included in a regression and easily controlled for. However if there are no available data on y in one of the periods or if y is an endogenous variable, then it would be desirable to eliminate y with a further differencing strategy. In our example of water bottles, the observations could be on purchases in stores where the income of the customers is not known. For the restaurants, the policy change was meant to affect income by increasing employment and income is may be endogenous to this system. With either missing data or endogeneity, a triple difference may be desirable. The triple difference specification as given by taking an additional difference in good i:

$$\Delta_{tci}x_{ict} = \gamma_i \ln(p_{i=1,c=1,t}/p_{i=1,c=1,t=0}) + (\delta_{i=1} - \delta_{i=0})\left[\ln(y_{i,c=1,t}/y_{i,c=1,t=0}) - \ln(y_{i,c=0,t=1}/y_{i,c=0,t=0})\right]. \tag{6}$$

or

$$\Delta_{tci}x_{\mathrm{ict}} = \gamma_{\mathrm{i}}\Delta_{tc}\ln p + (\delta_1 - \delta_0)\Delta_{tc}\ln y \tag{7}$$

Equation 6 is dependent upon there being no change in p for good 0 and growth in y being independent of the good. If y was the income term from a demand equation, that would be the case.

III. Specification error and estimation

The estimated models omit the terms in y (because of endogeneity or data problems). That is, one estimates the equations: $\Delta_{tc}x_{ict} = \gamma_i \Delta_{tc}\ln p$ or $\Delta_{tci}x_{ict} = \gamma_{\mathrm{i}}\Delta_{tc}\ln p$. The bias in the estimators for these equations can be derived using Thiel's specification error theorem (Thiel, 1971, 549).

$$
\begin{aligned}
\mathbf{E}\gamma_1^{tc} &= (\Delta_{tc}\ln p'\Delta_{tc}\ln p)^{-1}\Delta_{tc}\ln p' x_{tc} \\
&= (\Delta_{tc}\ln p'\Delta_{tc}\ln p)^{-1}\Delta_{tc}\ln p'[\gamma_1\Delta_{tc}\ln p + \delta_1\Delta_{tc}\ln y] \\
&= \gamma_1 + \{(\Delta_{tc}\ln p'\Delta_{tc}\ln p)^{-1}\Delta_{tc}\ln p'\Delta_{tc}\ln y\}\delta_1
\end{aligned} \tag{8}
$$

The first line is the formula for ordinary least squares (OLS) and **E** is the expectation operator. The second line substitutes the formula for the true model for x. The third line shows the difference between the OLS estimator and the true γ is the coefficient from regressing the income term on price times the true coefficient on the income term. The double difference estimator is unbiased if $\delta_1 = 0$ or if the regression of twice differenced $\ln y$ on twice differenced $\ln p$ has a zero coefficient. Another way to look at the bias is that whenever the product $(\Delta_{tc}\ln y)\delta_1$ is near zero, the double difference estimator will have little bias.

The same algebraic steps give the bias for the triple difference

$$
\begin{aligned}
\mathbf{E}\gamma_1^{tci} &= (\Delta_{tc}\ln p'\Delta_{tc}\ln p)^{-1}\Delta_{tc}\ln p' x_{tc} \\
&= (\Delta_{tc}\ln p'\Delta_{tc}\ln p)^{-1}\Delta_{tc}\ln p'[\gamma_1\Delta_{tc}\ln p \\
&\quad + (\delta_1 - \delta_0)\Delta_{tc}\ln y] \\
&= \gamma_1 + \{(\Delta_{tc}\ln p'\Delta_{tc}\ln p)^{-1}\Delta_{tc}\ln p'\Delta_{tc}\ln y\}(\delta_1 - \delta_0)
\end{aligned} \tag{9}
$$

The bottom line is that the triple difference has lower bias than the double difference whenever y is important and the two goods have a similar response to the omitted variable $\ln y$. Specifically whenever

$$|\delta_1 - \delta_0| < |\delta_1| \tag{10}$$

IV. Discussion

When should $|\delta_1 - \delta_0| < |\delta_1|$? When δ_1 and δ_0 are income elasticities for goods, one would expect them to be different by less than their magnitude. For instance, in Park et al. (1996), the income elasticities for foods eaten at home and foods away from home for poverty and nonpoverty groups are given. For each group for foods eaten at home, the income elasticities are within 0.2 and larger than 0.2, so a triple difference will lessen the bias. Blanciforti and Green (1983) look at income elasticities between groups of items. The income elasticities vary from 4.4 to .24, so a triple difference strategy absent prior knowledge of the income elasticities may not reduce bias.

In the context of a markup type model, let x be the price of a good, p be the policy and y is the cost of distribution. So long as the markup on the cost of distribution is close in the two regions, the triple difference is more accurate than the double difference.

What drives these examples is that the goods $i = 0$ and $i = 1$ share a common 'shift' variable y, either income or cost of distribution. So even though they may not respond exactly the same to that variable, and they may not have exactly the same changes in that variable, the product of the differences is all that matters and small times small is very small.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

We thank the Giannini Foundation for support.

References

Angrist, J. D., and A. B. Krueger. 2005. "Empirical Strategies in Labor Economics." *Handbook of Labor Economics* 3 (Part A): 1277–1366.

Athey, S., and G. Imbens. 2006. "Identification and Inference in Nonlinear Difference-in-Differences Models." *Econometrica* 74 (2, March): 431–497. doi:10.1111/ecta.2006.74.issue-2.

Bertrand, M., E. Duflo, and S. Mullainathan. 2004. "How Much Should We Trust Differences-In-Differences Estimates?" *The Quarterly Journal of Economics* 119 (1): 249–275. doi:10.1162/003355304772839588.

Blanciforti, L., and R. Green. 1983. "An Almost Ideal Demand System Incorporating Habits: An Analysis of Expenditures on Food and Aggregate Commodity Groups." *The Review of Economics and Statistics* 65 (3): 511–515.

Hamermesh, D., and S. Trejo. 2000. "The Demand for Hours of Labor: Direct Evidence from California." *The Review of Economics and Statistics* 82 (1): 38–47. doi:10.1162/003465300558614.

Kellogg, R., and H. Wolf. 2008. "Daylight Time and Energy: Evidence from an Australian Experiment." *Journal of*

Environmental Economics and Management 56: 207–220. doi:10.1016/j.jeem.2008.02.003.

Meyer, B. D. 1995. “Natural and Quasi-Experiments in Economics.” *Journal of Business & Economic Statistics* 13 (2) JBES Symposium on Program and Policy Evaluation (April): 151–161.

Park, J. L., R. B. Holcom, K. C. Raper, and O. Capps Jr. 1996. “A Demand Systems Analysis of Food Commodities by U.S. Households Segmented by Income.” *The Review of Economics and Statistics* 65 (3, August): 511–515.

Thiel, H. 1971. *Principles of Econometrics*. New York: John Wiley.