# Exam 4137 – Spring 2026 - Postponed

---

IMPORTANT: Always motivate your answers. Show knowledge and understanding of the concepts taught in the course. Make sure not only to repeat material from lectures and curriculum, but engage with the empirical examples and show that you understand what it means in the particular setting in addition to using math. Your answers should be as short as possible and as long as necessary. Each lettered subtask carries 10 points, for a total of 150.

---

1. (20 points) Short questions.

   (a) Interpret the following estimands

      i. $\mathbb{E}(Y_i^1 - Y_i^0 \mid D_i^1 > D_i^0)$

      This is the local average treatment effect. It is the average treatment effect for compliers: individuals who take treatment when the instrument is switched on, but wouldn't otherwise.

      ii. $\mathbb{E}(Y_i^1 - Y_i^0 \mid D_i = 0)$

      This is the average treatment effect for the untreated (ATUT).

   (b) TRUE or FALSE - No points without an explanation!

      iii. We do not need to specify the entire model correctly to consistently estimate it's parameters with maximum likelihood.

      TRUE. For a mean parameter, it is suffficient that the model for the conditional means are correctly specified. In that case, it is often called quasi-MLE. Regular variance estimates will however be incorrect..

      iv. Using person fixed effects in nonlinear models is a good way to get rid of all time-constant omitted variables.

      FALSE: In general not because of the incidental parameters problem - you cannot partition nonlinear models, and so we cannot get rid of the person FE's without consistently estimating them. In panel data where the time dimension is fixed, we cannot.

2. (80 points) You work in the Ministry of Finance, and have been tasked with evaluating a policy proposal to reduce the value added tax (VAT) on repairs of mobile phones and other small electronic devices. Reading up on the topic, you come across *"More and cheaper haircuts after VAT cut? On the efficiency and incidence of service sector consumption taxes"*, published by Tuomas Kosonen in the Journal of Public Economics in 2015. In that paper, he exploits the reduction in the VAT from 22 to 8% for hairdressers in Finland in 2007 in a Difference-in-differences setup, and uses a control group of beauty salons, day spas, massage services, physiotherapists, dry cleaning and repair services whose VAT rate was kept constant at 22%.

(a) Using potential outcomes notation, define and interpret the potential and realized log consumer prices for firm $i$ at time $t$.

$Y_{it}^0$ is the log consumer price of firm $i$'s output at time $t$ in the case where the firm faces the regular VAT rate of 22%.
$Y_{it}^1$ is the log consumer price of firm $i$'s output at time $t$ in the case where the firm faces the low VAT rate of 8%.
Treatment indicator $D_i = 1$ for hairdressers and $D_i = 0$ for firms in the control group.
Observed and potential outcomes are linked through SUTVA: $Y_{it} = D_i \times Y_{it}^1 + (1 - D_i) \times Y_{it}^0$. Students may also mention no pre-treatment effect either here or on the next subquestion: $Y_{it}^1 = Y_{it}^0$ if $t < T$, where $T$ is the reform period.

(b) Write down and interpret the regression equation you would run to obtain the difference in difference estimate.

$Y_{it} = \alpha_i + \gamma_t + \beta D_i \times 1(t \geq T) + e_{it}$
Here, $\alpha_i$ are firm fixed effects, $\gamma_t$ are quarter or time fixed effects, $\beta$ is the parameter of interest. It is OK to use $D_i$ in place of $\alpha_i$ and $1(t \geq T)$ in place of $\gamma_t$, but this presentation typically produce more precise estimates because it lowers residual variation. Time-varying treatment effects $\beta_g D_i \times 1(t = g)$ are also OK.

(c) What assumption do you need to make on the potential outcomes for the difference in difference estimate to have a causal interpretation? What interpretation is that?

We need the common trend assumption: $E(Y_{it}^0 - Y_{it'}^0 \mid D_i = 1) = E(Y_{it}^0 - Y_{it'}^0 \mid D_i = 0))$. The change in mean potential outcomes in the untreated state between any two years $t$ and $t'$ are the same for the treatment and control group. In this setting: The log of the consumer prices in the various types of firms would have increased by the same amount if they had all been subject to the 22% flat VAT. Because the outcome is in logs, thus means that prices would have changed with the same rate in the absence of treatment. Under that assumption, the diff in diff estimator recovers the ATT: The average effect of the VAT cut on the relative consumer prices for haircuts. Very good students could mention that it is indeed a variance-weighted average of the post-treatment quarter-specific ATT's.

(d) Figure 2 illustrates the main results from the paper. Does this figure support the identifying assumption? Why or why not?

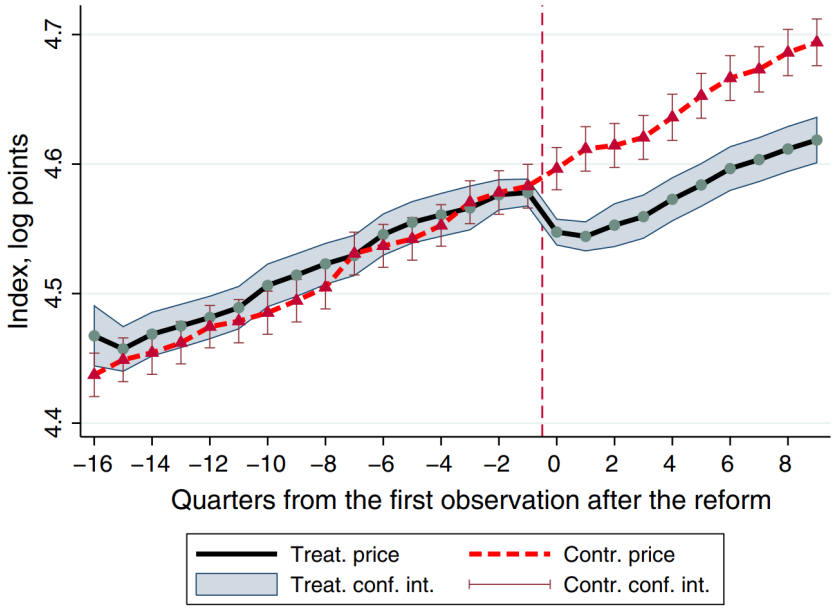



Fig. 2. from Kosonen (2015): Treatment and control prices as an index over time

*Note:* The figure presents the development of prices over time for a sample of firms. The price variable is a log of composite prices for hairdressing and other beauty services, an arithmetic average of different service prices in each firm and quarter. The mean value of the price variable is normalized at 1 at the end of 2006. The horizontal axis represents the quarters from the reform that took place in the beginning of 2007

We see that the trends in the log consumer price index is parallel for the hairdressers and other related industries in the control group prior to the VAT reform. This suggests that the trends between the treatment and control groups prior to the reform were parallel, which supports the common trends assumption - it is more plausible that the counterfactual trends in the absence of the reform would also have been parallel when we see that they were parallel in the pre-period, but the assumption is fundamentally untestable.

(e) Table 3 presents the main results in regression format. Interpret the DD coefficient in column (2).

The prices decrease on average by (approximately) 4.9%, because the outcome is in logs. This is an average treatment effect of the 22 to 8 percent VAT drop for the treated group - hairdressers.

(f) Is the magnitude of the price drop large or small? What price drop would you expect if the entire VAT reduction was passed on to consumer prices?

Full pass-through would imply a price drop of 1.08/1.22-1 = 0.115, or 11.5% price drop. The estimated coefficient is about half of that, suggesting that only half of the reduced VAT actually accrue to consumers.

(g) Explain why it is important to cluster standard errors on the firm level in this setting, as explained in the note to Figure 2 from Kosonen (2015).

**Table 3 from Kosonen (2015):** DD results on log prices.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | DD | DD | DD | Placebo |
| DD | $-0.1^{***}$ | $-0.049^{***}$ | $-0.058^{***}$ | 0.009 |
| | (0.008) | (0.007) | (0.013) | (0.012) |
| After | $0.124^{***}$ | $-0.004$ | 0.006 | $-0.011$ |
| | (0.006) | (0.006) | (0.012) | (0.011) |
| Log input | | | $0.008^{***}$ | 0.003 |
| | | | (0.003) | (0.003) |
| Net asset | | | $9.26\text{e}{-08}^{*}$ | $3.30\text{e}{-08}$ |
| | | | $(5.03\text{e}{-08})$ | $(5.47\text{e}{-08})$ |
| Income | | | $-5.80\text{e}{-09}$ | $-1.70\text{e}{-08}$ |
| | | | $(2.18\text{e}{-08})$ | $(3.77\text{e}{-08})$ |
| Time controls | No | Yes | Yes | Yes |
| N | 9499 | 9499 | 5535 | 2995 |
| $R^2$ | 0.247 | 0.436 | 0.294 | 0.324 |
| N firms | 702 | 702 | 531 | 502 |

*Note:* Fixed effects OLS regression results on log prices. The explanatory variables in column (1) are an indicator for the after period and a DD indicator. Column (2) adds to this a linear time trend for all, a separate time trend for treatment group and indicators for quarter of year. Column (3) adds firm accounting variables, shown. Column (4) is a placebo estimate where the after period starts from the beginning of 2005. Robust standard errors are clustered at firm level: ***p < 0.01, **p < 0.05, *p < 0.1

Regular, homoskedastic standard errors assume that the error term is identically and independently distributed for all rows in your data. This implies that the error term is independent across any two rows. This is questionable in a Diff-in-Diff designs like this. It could for instance be that the error term of a particular firm-quarter is correlated with the error term of the same firm in a different quarter. These concerns are particularly pressing when the treatment is determined at a level higher than the rows in your data, such as at the firm level in this setting. Clustered standard errors are robust to this problem in addition to heteroskedasticity. Students might also mention the Moulton factor.

(h) Explain the logic behind column (4) in Table 3 from Kosonen (2015). Why does this column support the identifying assumption?

This is an alternative way of supporting the common trend assumption. The authors pretend the reform happened in 2005 instead of 2007, and then run their diff in diff model on this placebo reform., The fact that they find nothing suggest that the prices of hairdressers and other similar businesses did not develop differently from before to after 2005.

3. (50 points) In the paper *"Incidence and distributional effects of Value Added Taxes"* (Economic Journal, 2019), Ingvil Gaarder instead exploits the reduction in Value Added Taxes for food in Norway in July 2001. The VAT rate was reduced from 24% to 12% for food, but not other commodities. Her main estimating equation is the following regression model:

$$y_{it} = \alpha + 1(t \geq c)[g_l(t-c) + \lambda] + 1(t < c)[g_r(t-c)] + e_{it}$$

, where $y_{it}$ is log consumer prices for good $i$ in month $t$, $c$ is the reform date (July 1st, 2001), $g_l$ and $g_r$ are unknown functions and $\lambda$ is the parameter of interest.

(a) Using potential outcomes, what assumptions do you need for the parameter $\lambda$ to recover causal effects of the VAT cut?

Continuity of potential outcomes $lim_{t\downarrow c}E(y_{it}^j) = lim_{t\uparrow c}E(y_{it}^j) \quad \forall j \in 0,1\backslash$. The mean of the potential consumer prices with the high $(j=0)$or the low $(j=1)$ VAT rate does not change discontinuously as we cross the reform threshold July 2001.$lim_{t\downarrow c}E(y_{it}^j \mid t)$ notation also fine.

(b) Interpret the estimate in column 2) of Table 3. What causal effect is this? Compare this estimate to the estimate from Kosonen (2015).

**Table 3 from Gaarder (2019):**

*Reform Effects on Consumer Prices of Food*

| | FD | RD | | |
|---|---|---|---|---|
| Dependent variable: log consumer price | (1) | (2) | (3) | (4) |
| Food and non-alcoholic beverages | $-0.105^{***}$ | $-0.106^{***}$ | $-0.103^{***}$ | $-0.106^{***}$ |
| | (0.008) | (0.014) | (0.005) | (0.006) |
| Item FE | No | No | Yes | Yes |
| Month effects | No | No | No | Yes |
| Hypothesis tests: | | p-value | | |
| $H_0 : \lambda = -0.097$ versus $H_1 : \lambda \neq -0.097$ | 0.315 | 0.533 | 0.226 | 0.142 |
| $H_0 : \lambda \geq -0.097$ versus $H_1 : \lambda < -0.097$ | 0.158 | 0.267 | 0.113 | 0.071 |

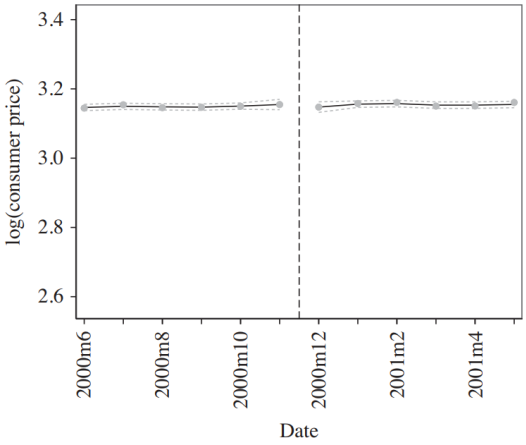
*Notes.* $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. The coefficient in column (1) is estimated using the FD model with log consumer prices as dependent variable. The coefficients in columns (2)–(4) are estimated using an RD model with log consumer prices as dependent variable, triangular weights and two months bandwidth. Column (2) reports the results with no controls, column (3) includes item fixed effects, and column (4) also controls for possible month effects using a DiD strategy. The standard errors are clustered at the firm level and robust to heteroscedasticity. We report p-values for the two-sided test that the VAT for food items is fully shifted and from one-sided tests of the null hypothesis that the VAT for food items is undershifted to consumer prices.

The coefficient estimate suggest that consumer prices on food dropped by (approximately) 10.6 percent. This is the average treatment effect of the VAT tax cut in July 2001 on food items, and is remarkably close to full pass-through: 1.012/1.024-1=0.097. The treatment effect (or pass-through) could be different from Kosonen (2015) because these are very different goods, and perhaps the elasticity of demand or supply is different (differences in market competition?), or in principle because VAT cuts have different effects in July than the rest of the year - the treatment effect is local to the cutoff.

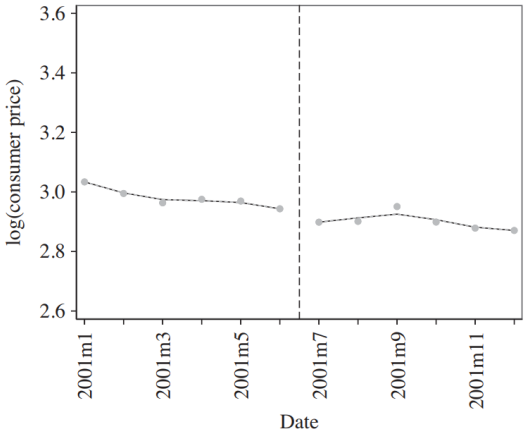
(c) The reform was announced in December 2000, six months before it was enacted. The author reports in Figure 4 the consumer price on food for months around the announcement. Explain the logic behind this robustness check and why the results in the figure supports the identification in the paper.

An important assumption is the continuity of mean potential outcomes. While time is non-manipulable, we could worry that the prices of food items had already responded before the introduction, in anticipation of the change. Figure 4 shows that absolutely nothing happened around the announcement, and therefore constitutes a particularly interesting placebo check.

(d) In Figure 5, the author reports the average predicted food prices based on oil prices and exchange rates around the time of the reform. Explain the logic behind this robustness test and why the results support the identifying assumptions of the paper.




(a) Fig. 4 from Gaarder (2019): Evolution of Consumer Price on Food Over Time




(b) Fig. 5 from Gaarder (2019): Evolution of Predicted Consumer Price on Food over Time

*Left Panel:* Each observation is the average consumer price for food reported on the 15th each month. The dashed vertical line denote the date the reform was announced. The solid lines are from a local linear regression with triangular weights on monthly consumer price data. The dashed lines represent the 95% confidence intervals. The y-axes are scaled to ±0.5 SD of the mean consumer price.

*Right Panel:* Predicted food consumer prices is given by a regression of food consumer prices on oil prices (Brent), euro/NOK, SEK/NOK, GDP/NOK and an import weighted exchange rate. The covariates are lagged four weeks and each observation is the average predicted consumer price for food reported on the 15th each month. The dashed vertical line denote the reform date. The solid lines are from a local linear regression with triangular weights on the predicted food consumer prices. The y-axes are scaled to ±0.5 SD of the mean predicted consumer price.

Again, we are worried about violation of mean potential outcomes. One way to strengthen our belief in that assumption is by looking at discontinuities in other covariates that we believe should be unaffected by the reform. Of particular interest are covariates that strongly predict consumer prices. Here, the author uses oil prices and exchange rates, which strongly influence consumer prices, and show that there are no effect on the predicted consumer prices based on these variables. This strengthens the identifying assumptions because then it is more plausible that potential consumer prices would also have been continuous at the threshold.

(e) Gaarder uses local linear regressions with triangular weights for $g_r$ and $g_l$ and limits attention to a short window round the reform date. Why is this?

A regression discontinuity design is fundamentally local to the cutoff: It identifies the shift at the threshold. Thus, we care only about estimating the two means correct at the threshold, and not about estimating the means correctly further away from the threshold. Using triangular weights essentially does this by assigning higher weights to observations closer to the threshold, while limiting the sample to one or two months around the threshold does this explicitly.