Takehome 3- DID

Jordi Torres

November 15, 2025

Exercise 1

Let $(g,t) \in \{1,\dots,G\} \times \{1,\dots,T\} := N$ In our case, g stands for municipalities and t for term. Treatment in this case is defined as $d_{g,t} \in \{0,1\}$, where 1 indicates that the village was given the Seguro Popular at that time period.

The design was staggered because treatment was switched in only once $d_{g,t} = 1$ and then it remains 1 (there is one exception of a switcher that is in treatment, switches off and then switcher back in, probably a data cleaning mistake.). Intensity of treatment does not vary over time or change, no?
In table 1 we can see the proportion of villages that switch into or out of treatment and those that are stayers. For those that are stayers, I distinguish between the always treated and the always control. We can observe that most of the sample is always treated, while only 10% of villages switch.

group	n	rel_freq
Always Treated	151	0.77
Never Treated	26	0.13
Switcher	18	0.09

Table 1: Classification of municipalities

We can write the following model:

$$Y_{g,t} = \sum_{g'=1}^{G} \gamma_{g'} \mathbf{1}\{g = g'\} + \sum_{t'=1}^{T} \gamma_{t'} \mathbf{1}\{t = t'\} + \beta_{fe}^{X} \mathbf{1}\{d_{g,t} = 1\} + \mathbf{X}'\beta + \epsilon_{j,t}$$

Where we cluster $\epsilon_{j,t}$ at the village level.
We need to collapse the dataset at the level of analysis. I include the general case where we want to condition on a vector of characteristics at the g,t level to make sure assumptions below hold (as stated in Theorem 1 of Chaisemartin d'Haultefe DID)

Theorem 1. *In design CLA, if PT and NA hold, then:* $\mathbf{E}(\beta_{fe}^{X}) = ATT$

Assumptions needed

1. **SUTVA (randomness of treatment)**: this implies that g's potential outcome depends only on g's treatment and not on other villages. This implies that there are no spillovers across villages of the Seguro Popular treatment (which seems reasonable).–¿ is this needed 100%? Is this implied by any of the other two assumptions?

2. **No anticipation (NA)**: current outcome does not depend on future realizations of the treatment. Here implies that agents don't expect the change in the subsidy and adapt their behavior? Initial condition also? what if villages were exposed to treatment before we start measuring them?

3. **Parallel trend assumptions (PT)**: it implies that, in the absence of treatment, the trends that groups would have followed was the same. That is: for all $t \leq 2$: $\mathbf{E}(Y_{g,t}(0) - Y_{g,t-1}(0))$, which means for untreated potential outcomes would have stayed the same.

Limitations of these assumptions? **ADD** p.67 of their book, basically.

	2WFE	2WFE	FD	FD
treat	-0.04	-0.04	-0.07	-0.07
	(0.07)	(0.07)	(0.09)	(0.09)
mean_age		-0.02		-0.06
		(0.06)		(0.05)
share_formal		0.02		0.04
		(0.06)		(0.07)
mean_sec_occup		0.12		
		(0.09)		
(Intercept)			-0.00	0.01
			(0.02)	(0.02)
Num. obs.	772	772	577	577
Num. groups: cvemun	195	195		
Num. groups: quarter	4	4		
R^2 (full model)	0.72	0.72		
R^2 (proj model)	0.00	0.00		
Adj. R^2 (full model)	0.63	0.63		
Adj. R^2 (proj model)	-0.00	-0.00		
R^2			0.00	0.00
Adj. R^2			-0.00	-0.00

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 2: Fixed Effects and First Differences Models

Exercise 2

In table 2 I show the results of the TWFE estimator and the First difference estimator. Both coefficients are not significant; likely because of the low amount of switchers that we have in the dataset. Without forbidden comparisons. –¿solution to this? with and without forbidden comparisons?
textbfNote: revise this ugly ass table Solved a way to put latex tables directly from R to Latex but with the problem of 1. ugly as fuck, 2. unable to put it where I want to.
Here to revise a little bit the implementation of the first difference estimator etcetera in R and modify a little bit so the table is correct.

Exercise 5

$$\mathbf{E}(w_j) = U, \forall j \in \{f, i\}$$

$$\frac{w_f^* + B_f + \eta_f U}{\rho + \eta_f} = U$$

$$w_f^* = U\rho - Bj$$

Thus, although the total compensation will the be the same across informal and formal jobs (i.e $U\rho$), if $B_f > B_i \iff w_f > w_i$. Higher amenities in terms of social security conditions in formal jobs will decrease the reservation wage. (no differential in skill-type of contract/informality + assumes workers with different skills have the same bargaining power.)

Exercise 6

I will first try to simplify the expression of w_j^*, given that this is an essential equilibrium element in the likelihood. This follows the procedure we did in class.
We use the results of the previous exercise and the definition given in the exercise to write:

$$w_j^* = b - B_j + \sum_j \lambda_j \left(\int_{w_j < w_j^*} U \, dF_j(w_j) + \int_{w_j \geq w_j^*} E_j(w_j) \, dF_j(w_j) - U \right)$$

Collecting terms and noting that $\int_{w_j < w_j^*} dF_j(w_j) = F_j(w_j^*)$, we obtain:

$$w_j^* = b - B_j + \sum_j \lambda_j \left(U F_j(w_j^*) - U + \int_{w_j \geq w_j^*} E_j(w_j) \, dF_j(w_j) \right)$$

This simplifies to:

$$w_j^* = b - B_j + \sum_j \lambda_j \left(-U\big(1 - F_j(w_j^*)\big) + \int_{w_j \geq w_j^*} E_j(w_j) \, dF_j(w_j) \right).$$

Adding and subtracting U inside the integral and simplifying the $U\big(1 - F_j(w_j^*)\big)$ terms gives:

$$w_j^* = b - B_j + \sum_j \lambda_j \int_{w_j \geq w_j^*} \big(E_j(w_j) - U\big) \, dF_j(w_j)$$

We then substitute for the definition of $E_j(w_j)$:

$$E_j(w_j) = \frac{w_j + B_j + \eta_j U}{\rho + \eta_j}$$

Using $E_j(w_j^*) = U$ implies $E_j(w_j) - U = \dfrac{w_j - w_j^*}{\rho + \eta_j}$Replacing this into the expression above, we obtain:

$$w_j^* = b - B_j + \sum_j \frac{\lambda_j}{\rho + \eta_j} \int_{w_j \geq w_j^*} \big(w_j - w_j^*\big) \, dF_j(w_j),$$

which holds for each $j \in \{i, f\}$

Then we construct the likelihood contributions of employed and unemployed people in the following way.

Unemployed. Let $\tilde{F}_j(w_j^*) \equiv 1 - F_j(w_j^*)$ and define the sector-specific job-finding hazards $a_j \equiv \lambda_j \tilde{F}_j(w_j^*)$. The total exit hazard from unemployment is

$$h_u \;=\; a_i + a_f \;=\; \lambda_i \tilde{F}_i(w_i^*) + \lambda_f \tilde{F}_f(w_f^*)$$

The likelihood of an ongoing unemployment spell of duration t_u is

$$L(t_u, u) \;=\; \underbrace{h_u \exp(-h_u t_u)}_{\text{duration density}} \times \underbrace{p(u)}_{\text{steady-state mass at } U},$$

where, for the three-state system (U, E_i, E_f),

$$p(u) \;=\; \left(1 + \tfrac{a_i}{\eta_i} + \tfrac{a_f}{\eta_f}\right)^{-1}$$

Employed. For employed individuals with observed sector j and wage w,

$$L(w, e_j) \;=\; \underbrace{\frac{f_j(w)}{\tilde{F}_j(w_j^*)}}_{\text{accepted-wage density on } [w_j^*, \infty)} \times \underbrace{p(e_j)}_{\text{steady-state mass at } E_j}, \qquad p(e_j) \;=\; \frac{a_j/\eta_j}{1 + \frac{a_i}{\eta_i} + \frac{a_f}{\eta_f}}$$

If the sector is not observed, use the mixture

$$L(w, e) \;=\; \sum_{j \in \{i, f\}} \frac{f_j(w)}{\tilde{F}_j(w_j^*)} \, p(e_j)$$

Thus, the overall likelihood function can be written as:

$$L(\theta) = \prod_{i \in U} L(t_{u,i}, u) \times \prod_{i \in E} L(w_i, e).$$

And if we take logs the previous expression simply becomes:

$$LL(\theta) = \sum_{i \in U} (\log(h_u) - h_u t_{u,i} + \log p(u)) \, \mathbf{1}\{i \in U\} + \sum_{i \in e_j} \Big(\log(f_j(w)) - \log(\tilde{F}_j(w^*)) + \log p(e_j)\Big) \, \mathbf{1}\{i \in e_j\}$$

Derivatives
We want to argue that some parameters can't be identified. So we take derivatives with respect to the four primitives of the model. Focusing on how the hazard h_u and the truncation term $\tilde{F}_j(w_j^*)$ enter the likelihood, we can write (ignoring mixture issues and keeping N_u and N_e as the numbers of unemployed and employed observations, respectively):

$$\frac{\partial LL(\theta)}{\partial \lambda_j} = N_u \left[\left(\frac{1}{h_u} - \bar{t}_u \right) \frac{\partial h_u}{\partial \lambda_j} \right] - N_e \left[\frac{1}{\tilde{F}_j(w_j^*)} \frac{\partial \tilde{F}_j(w_j^*)}{\partial \lambda_j} \right],$$

where $\bar{t}_u$ is the average unemployment duration.
If we expand the expressions (ignoring cross–sector effects on the other reservation wage), then:

$$\frac{\partial h_u}{\partial \lambda_j} = \tilde{F}_j(w_j^*) + \lambda_j \frac{\partial \tilde{F}_j(w_j^*)}{\partial \lambda_j} = \tilde{F}_j(w_j^*) - \lambda_j f_j(w_j^*) \frac{\partial w_j^*}{\partial \lambda_j},$$

$$\frac{\partial \tilde{F}_j(w_j^*)}{\partial \lambda_j} = \frac{\partial \tilde{F}_j(w_j^*)}{\partial w_j^*} \frac{\partial w_j^*}{\partial \lambda_j} = -f_j(w_j^*) \frac{\partial w_j^*}{\partial \lambda_j},$$

where $f_j(\cdot)$ is the density of $F_j(\cdot)$.
Thus:

$$\begin{aligned} \frac{\partial LL(\theta)}{\partial \lambda_j} &= N_u \left[\left(\frac{1}{h_u} - \bar{t}_u \right) \left(\tilde{F}_j(w_j^*) - \lambda_j f_j(w_j^*) \frac{\partial w_j^*}{\partial \lambda_j} \right) \right] - N_e \left[\frac{1}{\tilde{F}_j(w_j^*)} \left(-f_j(w_j^*) \frac{\partial w_j^*}{\partial \lambda_j} \right) \right] \\ &= N_u \left[\left(\frac{1}{h_u} - \bar{t}_u \right) \left(\tilde{F}_j(w_j^*) - \lambda_j f_j(w_j^*) \frac{\partial w_j^*}{\partial \lambda_j} \right) \right] + N_e \left[\frac{f_j(w_j^*)}{\tilde{F}_j(w_j^*)} \frac{\partial w_j^*}{\partial \lambda_j} \right]. \end{aligned}$$

Hence, λ_j enters both directly (through $\tilde{F}_j(w_j^*)$ in h_u) and indirectly through the equilibrium object w_j^*.
Then, if we take derivatives with respect to η_j we can write:

$$\frac{\partial LL(\theta)}{\partial \eta_j} = N_u \left[\left(\frac{1}{h_u} - \bar{t}_u \right) \frac{\partial h_u}{\partial \eta_j} \; + \; \frac{1}{p_u} \frac{\partial p_u}{\partial \eta_j} \right] - N_e \left[\frac{1}{\tilde{F}_j(w_j^*)} \frac{\partial \tilde{F}_j(w_j^*)}{\partial \eta_j} \; + \; \frac{1}{p_e} \frac{\partial p_e}{\partial \eta_j} \right].$$

Doing the same type of expansion as before:

$$\begin{aligned} \frac{\partial h_u}{\partial \eta_j} &= \lambda_j \frac{\partial \tilde{F}_j(w_j^*)}{\partial \eta_j} + \lambda_i \frac{\partial \tilde{F}_i(w_i^*)}{\partial \eta_j} \\ &= -\lambda_j f_j(w_j^*) \frac{\partial w_j^*}{\partial \eta_j} - \lambda_i f_i(w_i^*) \frac{\partial w_i^*}{\partial \eta_j}, \\ \frac{\partial \tilde{F}_j(w_j^*)}{\partial \eta_j} &= -f_j(w_j^*) \frac{\partial w_j^*}{\partial \eta_j}, \end{aligned}$$

so that

$$\frac{\partial LL(\theta)}{\partial \eta_j} = N_u \left[\left(\frac{1}{h_u} - \bar{t}_u \right) \left(-\lambda_j f_j(w_j^*) \frac{\partial w_j^*}{\partial \eta_j} - \lambda_i f_i(w_i^*) \frac{\partial w_i^*}{\partial \eta_j} \right) + \frac{1}{p_u} \frac{\partial p_u}{\partial \eta_j} \right]$$

$$- N_e \left[\frac{1}{\tilde{F}_j(w_j^*)} \left(-f_j(w_j^*) \frac{\partial w_j^*}{\partial \eta_j} \right) + \frac{1}{p_e} \frac{\partial p_e}{\partial \eta_j} \right]$$

$$= N_u \left[\left(\frac{1}{h_u} - \bar{t}_u \right) \left(-\lambda_j f_j(w_j^*) \frac{\partial w_j^*}{\partial \eta_j} - \lambda_i f_i(w_i^*) \frac{\partial w_i^*}{\partial \eta_j} \right) + \frac{1}{p_u} \frac{\partial p_u}{\partial \eta_j} \right]$$

$$+ N_e \left[\frac{f_j(w_j^*)}{\tilde{F}_j(w_j^*)} \frac{\partial w_j^*}{\partial \eta_j} - \frac{1}{p_e} \frac{\partial p_e}{\partial \eta_j} \right].$$

Same thing happens here: η_j affects the likelihood only through equilibrium objects such as w_j^* and the steady-state probabilities p_u and p_e.
Finally, for the parameters of the wage offer distribution we can schematically write (for sector j):

$$\frac{\partial LL(\theta)}{\partial \mu_j} = \kappa \frac{(\ln w_j - \mu_j)}{\sigma_j^2} + \frac{\partial \Phi(\cdot)}{\partial \mu_j},$$

and for the variance parameter:

$$\frac{\partial LL(\theta)}{\partial \sigma_j^2} = \kappa \frac{(\ln w_j - \mu_j)^2}{2\sigma_j^4} + \frac{\partial \Phi(\cdot)}{\partial \sigma_j^2},$$

where κ collects constants and sample size terms, and $\Phi(\cdot)$ denotes the contribution of the truncation and duration components of the likelihood. The data only allow us to identify certain combinations of the structural parameters. For instance, $a_j = \lambda_j \tilde{F}_j(w_j^*)$ can be identified, but it is a composite of λ_j and the reservation wage w_j^* (and thus of μ_j, σ_j). The same argument applies to η_j.
intuition: in the likelihood λ_j and η_j never enter separately. From the data, we can identify only the reduced form objects $hu = ai + af$ the ratios $\frac{a_j}{\eta_j}$, and the shape of the accepted wage distributions.
FOCs and (non-)identification; collinearity.
Define the sector-specific job–finding hazard

$$a_j \equiv \lambda_j \tilde{F}_j(w_j^*).$$

All the elements that enter the likelihood — the total hazard h_u, the stationary probabilities $p(u)$ and $p(e_j)$, and the accepted–wage density — can be written in terms of a_j, η_j and the wage-distribution parameters (μ_j, σ_j), and, crucially, the reservation wage w_j^* itself is an equilibrium object that also depends on $(\lambda_j, \eta_j, \mu_j, \sigma_j)$ only through a_j.
Thus the log–likelihood can be rewritten as

$$LL(\theta) = \tilde{L}\big(a_i, a_f, \eta_i, \eta_f, \mu_i, \mu_f, \sigma_i, \sigma_f\big),$$

with λ_j appearing only via $a_j = \lambda_j \tilde{F}_j(w_j^*)$.
Using the chain rule,

$$\frac{\partial LL}{\partial \lambda_j} = \frac{\partial \tilde{L}}{\partial a_j} \frac{\partial a_j}{\partial \lambda_j}, \qquad \frac{\partial LL}{\partial \eta_j} = \frac{\partial \tilde{L}}{\partial a_j} \frac{\partial a_j}{\partial \eta_j} + \frac{\partial \tilde{L}}{\partial \eta_j}.$$

If, as in this model, $\tilde{L}$ depends on η_j only through the ratio a_j/η_j (via the stationary probabilities), then $\partial \tilde{L}/\partial \eta_j$ is itself a linear combination of $\partial \tilde{L}/\partial a_j$, and the two first–order conditions

$$\frac{\partial LL}{\partial \lambda_j} = 0, \qquad \frac{\partial LL}{\partial \eta_j} = 0$$

are *collinear*: they collapse to the same restriction in terms of a_j and a_j/η_j. In other words, the data pin down only the composite objects

$$a_j = \lambda_j \tilde{F}_j(w_j^*), \qquad \frac{a_j}{\eta_j},$$

but not λ_j and η_j separately. This is precisely the sense in which some of the structural parameters are not identified: different combinations of (λ_j, η_j) that yield the same $(a_j, a_j/\eta_j)$ produce the same likelihood and hence satisfy the same FOCs.