

Estimating Child Marriage Ban Impacts on Labour Market Outcomes in Mexico using Staggered DiD and IV-UQR Analysis

Erendira Leon Bravo (e.leonbravo@westminster.ac.uk)

University of Westminster

Lorraine Mackenzie (lorraine.mackenzie@stir.ac.uk)

University of Stirling

UK Stata Conference

3-4 September 2026

- Analysis based on the **impacts of legislative reforms that raised the minimum legal age of marriage to 18** on labour market outcomes
- A quasi-natural experiment based on a **staggered state-level** rollout of the child bans between 2014 and 2021

First stage: Staggered Difference-in-Differences (DiD) estimation

- Address the **endogeneity of marriage**
- Validity of **parallel trends**
- Using administrative data from the Marriage Statistics (EMAT) 2008-2024 to construct a cohort-state-year level instrument

Second stage: IV Unconditional Quantile Regression (IV-UQR)

- Analyse **wage heterogeneity** across the unconditional distribution
- Using earnings from the National Survey of Occupation and Employment (ENOE) 2008-2024

Staggered DiD and Validity of parallel trends

In spirits of Wang & Wang (2017) to causally identify the compliance of the legal system with the ban.

The equation for the staggered DiD by gender is as follows:

$$MMAL_{sc,j} = \gamma_{0,j} + \gamma_{1,j}DiD * T_{sc,j} + \mu_{sj} + \lambda_{c,j} + \rho_t + v_{sc,j} \quad (1)$$

reghdfe – Linear regression with multiple fixed effects. Also supports individual FEs with group-level outcomes.

```
reghdfe depvar [indepvars] [if] [in] [weight], [absorb(absvars)]  
[options]
```

absorb(absvars), categorical variables representing the fixed effects to be absorbed.

Staggered DiD and Validity of parallel trends

Example:

**For women*

```
reghdfe MMAL_cs did_term i.year if sex==0, absorb(state bth_year) cluster(state)
predict mmal_inst_men, xb
```

**For men*

```
reghdfe MMAL_cs did_term i.year if sex==1, absorb(state bth_year) cluster(state)
predict mmal_inst_women, xb
```

Parallel pre-trends and absence of anticipation effects using **csdid** following Callaway and Sant'Anna (2021)

```
csdid depvar [indepvars] [if] [in] [weight], [ivar(varname)]
[time(varname)] [gvar(varname)] [options]
```

ivar(varname), variable used as panel identifier. e.g., *state*.

time(varname), variable to identify time. e.g., *year*.

gvar(varname), variable identifying treatment groups or cohorts. Any positive value indicates when a group was initially treated.

Staggered DiD and Validity of parallel trends

Post estimation:

csdid simple, aggregates all post-treatment $ATT(g,t)$.

csdid event, for an event-study (dynamic treatment effects) table with leads and lags.

csdid pretrend, for a formal test on the parallel trends assumption.

csdid_plot, for creating plots of the results.

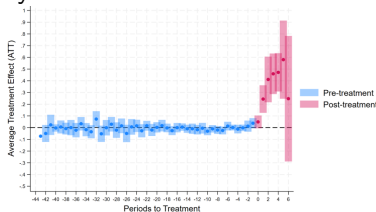
Example:

```
csdid age_marr, ivar(state) time(bth_year) gvar(first_affected_cohort) notyet
estat pretrend
estat simple
estat event

csdid_plot, title("{bf:Event Study:Dynamic effects of the ban on age of marriage}", size(medium)) ///
ytitle("Average Treatment Effect (ATT)") ///
xlabel(-44(2)6, labsize(vsmall)) ///
ylabel(-0.5(0.1)1, labsize(vsmall))
```

Staggered DiD and Validity of parallel trends

Event-Study: Dynamic effects of the ban on age of marriage



The staggered implementation of the child marriage bans successfully raised the assigned minimum legal marriage age on average by approximately 1.64 years for the affected cohorts for women and by 1.77 years for men.

	(1) Women	(2) Men
Minimum Legal Age of Marriage	1.64*** (0.324)	1.77*** (0.335)
No. of Obs.	8,253,752	7,939,713

Notes: *p<0.1, ** p<0.05, *** p<0.01
Standard errors in parentheses. Staggered Difference-in-Differences (DiD) for the instrumental variable of the age at marriage. The sample is constructed from the 2008-2024 Marriage Statistics, observations aged 15 to 50. The models is significant overall (Prob >F = 0.0000).

IV Unconditional Quantile Regression (IV-UQR)

Based on Firpo et al. (2009) in applying the **Re-centred Influence Function (RIF)** procedure

It allows the average effects to be interpreted at different quantiles of unconditional wage distribution as the dependent variables (e.g., the 5th, 95th percentiles or other intermediate quantiles)

The RIF quantile for the log hourly earnings can be expressed as:

$$RIF(w_{sc,j}; q_\tau) = q_\tau + [\tau - 1(w \leq q_\tau)]/f_w(q_\tau) \quad (2)$$

rifhdreg – Recentered influence function regression with high-dimensional fixed effects.

```
rifhdreg depvar [indepvars] [if] [in] [weight], rif(RIF_options)
```

IV Unconditional Quantile Regression (IV-UQR)

rif(*RIF_options*), specify the RIF statistic used as outcome.

The RIF statistic is first obtained using the **egen** function **rifvar()**
-then used as the dependent variable in an OLS model.

Then, the second stage equation for a given RIF quantile is implemented through the following specification for each gender at the individual level:

$$\widehat{RIF}(w_{sc,j}, \hat{q}_\tau) = \beta_{0\tau} + \beta_{1,\tau} \widehat{MMAL}_{sc,j\tau} + \beta_{2,\tau}(X_{sc,j\tau}) + \rho_\tau + \epsilon_{sc,j\tau} \quad (3)$$

xtreg – Linear models for panel data.

Fixed-effects (FE) model

```
xtreg depvar [indepvars] [if] [in] [weight], fe [FE_options]
```

IV Unconditional Quantile Regression (IV-UQR)

where *depvar* is the dependent variable, and *indepvars* are independent variables

weights are allowed for the fixed-effects model and for the population-averaged model

options, for the type of standard error reported. E.g., *robust*, *cluster*, etc.

A panel variable must be specified with the use of **xtset**

Example:

```
*****  
***Getting the RIFs  
*****  
foreach x of numlist 5/95 {  
  quietly rifhdreg lg_inc, rif(q(`x')) att retain(rifq`x')  
}  
  
*****  
**IV-UQR  
*****  
  
*For women  
foreach q of numlist 5(5)95 {  
  xtreg rifq`q'_m mmal_inst_women i.year i.bth_state if sex==0, fe cluster(state_cohort_id) robust  
  est store women_q`q'  
}  
  
*For men  
foreach q of numlist 5(5)95 {  
  xtreg rifq`q'_m mmal_inst_men i.year i.bth_state if sex==1, fe cluster(state_cohort_id) robust  
  est store men_q`q'  
}
```

IV-UQR Results

The empirical evidence suggests strong heterogeneity impacts. Delaying marriage boosts real earnings at the 10th and 25th percentiles by 23.0% and 16.2% for women, and 28.9% and 15.5% for men

	Women					
	Mean	RIF-10	RIF-25	RIF-50	RIF-75	RIF-90
<i>Minimum Legal Age of Marriage</i>	0.092*** (0.027)	0.230*** (0.061)	0.162*** (0.036)	0.074*** (0.025)	-0.004 (0.022)	-0.01 (0.025)
No. of Obs.	501,352	501,352	501,352	501,352	501,352	501,352

	Men					
	Mean	RIF-10	RIF-25	RIF-50	RIF-75	RIF-90
<i>Minimum Legal Age of Marriage</i>	0.082*** (0.013)	0.289*** (0.052)	0.155*** (0.020)	0.037*** (0.016)	-0.014 (0.014)	-0.028* (0.017)
No. of Obs.	749,330	749,330	749,330	749,330	749,330	749,330

Notes: Standard errors in parentheses

*p<0.1, ** p<0.05, *** p<0.01

IV-UQR regressions for the impacts of the age at marriage on logarythm of hourly earnings for women and men. The sample is constructed from the 2008-2024 Mexican National Occupations and Employment Survey (ENOE) and Marriage Statistics Data (EMAT). The models is significant overall (Prob >F = 0.0000).

Thank you!

References

- Callaway, Brantly and Sant'Anna, Pedro H. C. (2021). Difference-in-Differences with multiple time periods. *Journal of Econometrics*, 225(2): 200-230.
- Firpo, S. P., N. M. Fortin, and T. Lemieux (2009). Unconditional quantile regressions. *Econometrica* 77: 953-973.
<https://doi.org/10.3982/ECTA6822>
- Wang, C., and Wang L. (2017) Knot yet: minimum marriage age law, marriage delay, and earnings. *Journal of Population Economics*, 30:771–804. DOI 10.1007/s00148-017-0632-5

IV Unconditional Quantile Regression (IV-UQR)

rifvar() – is an egen extension that can be used to create RIFs for a large set of distributional statistics and in combination with other statistics

The flexibility and simplicity of this tool extends the analysis to the Gini inequality index, using linear regressions

```
egen newvar = rifvar(varname) [if] [in] [, options]
```

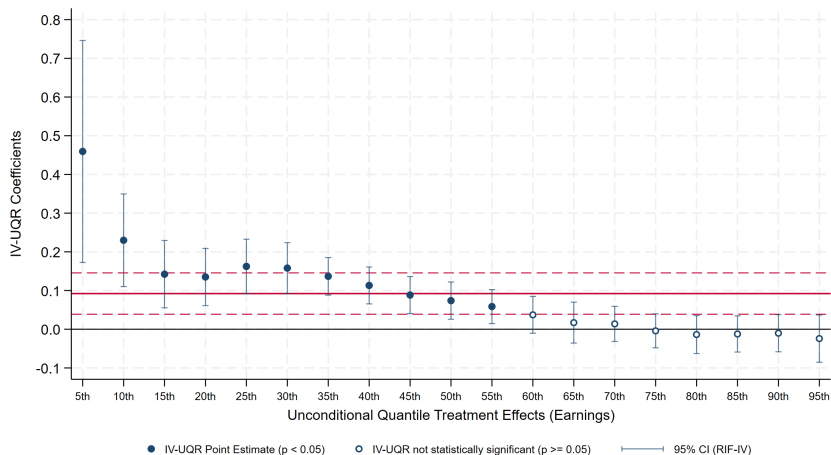
where *newvar* is the new variable created, and *varname* the variable of interest

Then:

```
xtreg depvar [indepvars] [if] [in] [weight], fe [FE_options]
```

IV-UQR Plot Results

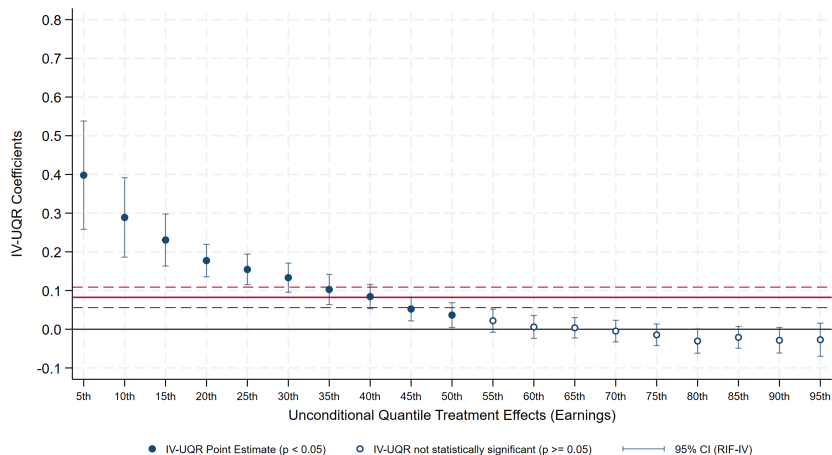
IV-UQR Treatment Effects Across the Female Earnings Distribution



Note: Solid red line represents OLS mean estimation with 95% confidence bounds (dashed red lines).

IV-UQR Plot Results

IV-UQR Treatment Effects Across the Male Earnings Distribution



Note: Solid red line represents OLS mean estimation with 95% confidence bounds (dashed red lines).