Econometrics 2

Regression with Panel Data

Lasha Chochua

2026

Panel Data – Introduction

Introduction

- So far, we have only focused on **cross-sectional data**. Now we introduce **time**, which leads us to **Panel Data**.
- A **panel dataset** contains observations on multiple entities, where each entity is observed at two or more points in time.
- If the dataset includes variables X and Y, the observations are denoted:

$$(X_{it}, Y_{it}), \quad i = 1, \ldots, n \text{ and } t = 1, \ldots, T$$

  - The subscript i refers to the **entity** being observed
  - The subscript t refers to the **time** at which the observation is recorded

- Depending on whether some time periods are missing for some entities:
  - A **balanced panel** has no missing observations
  - An **unbalanced panel** has missing observations for some (i, t) pairs

Introduction – Data Structure

- Each row = one (i, t) observation
- Entities stacked by time period (or vice versa)
- Panel structure allows **within-unit** comparisons over time

Observation Number	State	Year	Cigarette Sales (packs per capita)	Average Price per Pack (including taxes)	Total Taxes (cigarette excise tax + sales tax)
1	Alabama	1985	116.5	$1.022	$0.333
2	Arkansas	1985	128.5	1.015	0.370
3	Arizona	1985	104.5	1.086	0.362
⋮	⋮	⋮	⋮	⋮	⋮
47	West Virginia	1985	112.8	1.089	0.382
48	Wyoming	1985	129.4	0.935	0.240
49	Alabama	1986	117.2	1.080	0.334
⋮	⋮	⋮	⋮	⋮	⋮
96	Wyoming	1986	127.8	1.007	0.240
97	Alabama	1987	115.8	1.135	0.335
⋮	⋮	⋮	⋮	⋮	⋮

Example – Panel Data on Traffic Fatalities

- **Panel data** (also called **longitudinal data**) track multiple entities across time.
- In this lecture, we use a dataset on **state-level traffic fatalities in the U.S.**

- The data include:
  - $n = 48$ states
  - $T = 7$ years: 1982, ..., 1988
  - Total observations:

$$48 \times 7 = 336$$

- Each observation contains traffic and policy variables for a specific state and year.
- We use this dataset to illustrate panel data methods throughout.

Context – Traffic Deaths and Alcohol Policies

- About **40,000 highway fatalities** occur annually in the U.S.
- Around **25%** of fatal crashes involve a drinking driver.
- Risk of fatal crash is much higher between **1 a.m. and 3 a.m.**

- The dataset includes:
  - Annual **traffic fatality rates** per 10,000 people
  - **Alcohol taxes** (real tax on a case of beer, inflation-adjusted to 1988 dollars)
  - **Drunk driving laws** and their type/severity

- **Goal:** study how government policies (like beer taxes) affect fatality rates

OLS with Cross-Sectional Data – 1982 and 1988

- We regress **fatality rate** on **beer tax** using data from a single year.

- **1982:**

$$\widehat{FatalityRate} = \underset{(0.15)}{2.01} + \underset{(0.13)}{0.15} \cdot BeerTax$$

- **1988:**

$$\widehat{FatalityRate} = \underset{(0.11)}{1.86} + \underset{(0.13)}{0.44} \cdot BeerTax$$

- **Interpretation:**
  - The 1988 coefficient is **statistically significant**, but both years show a **positive** effect
  - This is counterintuitive: higher beer taxes appear to **increase** traffic deaths
  - Likely explanation: **omitted variable bias**

Why Panel Data Helps

- Many factors affect traffic fatality rates:
  - Road quality, car density, driving culture, etc.
  - These are **hard to observe** and may correlate with beer tax

- With **cross-sectional data**, we cannot control for these unobserved factors.
- With **panel data**, we observe each state over time.

Key Insight

If unobserved factors are **constant within a state over time**, we can control for them using **state fixed effects** – even without directly observing them.

Before-and-After Comparisons

Before-and-After Comparisons

- When each state is observed for $T = 2$ periods, we can compare the second period to the first.
- This **differences method** holds constant unobserved factors that vary across states but are constant over time.

- Let Z_i be a variable that affects fatalities in state i but does **not change over time** (e.g., cultural norms).
- Then the regression model is:

$$FatalityRate_{it} = \beta_0 + \beta_1 BeerTax_{it} + \beta_2 Z_i + u_{it}$$

Before-and-After Comparisons (Cont.)

- For 1982:

$$FatalityRate_{i,1982} = \beta_0 + \beta_1 BeerTax_{i,1982} + \beta_2 Z_i + u_{i,1982}$$

- For 1988:

$$FatalityRate_{i,1988} = \beta_0 + \beta_1 BeerTax_{i,1988} + \beta_2 Z_i + u_{i,1988}$$

- Subtracting the 1982 equation from the 1988 equation:

$$FatalityRate_{i,1988} - FatalityRate_{i,1982} = \beta_1 (BeerTax_{i,1988} - BeerTax_{i,1982}) + (u_{i,1988} - u_{i,1982})$$

Interpreting the Differenced Equation

- Intuition: if cultural norms Z_i don't change between 1982 and 1988, they won't affect **changes** in fatality rates.

- Changes in fatalities must then be driven by:
  - **Changes in beer tax**
  - **Changes in the error term**

Important

Key Result

By focusing on changes, we eliminate unobserved time-invariant variables Z_i, reducing **omitted variable bias** in the causal estimate.

OLS on First Differences

- The regression based on changes:

$$\Delta \widehat{FatalityRate}_i = \underset{(0.065)}{-0.072} - \underset{(0.36)}{1.04} \cdot \Delta BeerTax_i$$

- **Interpretation:**
  - Slope is **statistically significant** at the 5% level
  - A \$1 increase in beer tax reduces traffic deaths by **1.04 per 10,000 people**
  - This is a **large effect**: the average fatality rate is around 2 in these data
  - Implication: traffic deaths could be **cut in half** with a \$1 tax increase

> **Note**
>
> The differenced regression controls for **fixed factors** like cultural norms, but not for time-varying confounders.

Caution and Limitations

- The differenced regression controls for factors **fixed over time**.
- But other variables may also **change over time** and correlate with beer taxes.
- Omitting them still leads to **omitted variable bias**.

- Also, before-and-after methods are designed for $T = 2$.
- Our dataset has **7 years** – we need a method that uses all the data.

Solution

Fixed effects regression allows us to fully exploit the panel structure for any $T \geq 2$.

Fixed Effects Regression

Fixed Effects Regression

- Fixed effects regression controls for **omitted variables** in panel data that vary across entities but do **not change over time**.
- Unlike before-and-after methods, fixed effects work with **more than two time periods**.

- The model allows for a **different intercept** for each entity:
  - These intercepts absorb all time-invariant differences across entities
  - They can be represented by a set of **binary (indicator) variables**

Key Requirement

Fixed effects are appropriate when omitted factors differ across units but are **constant within units over time**.

The Fixed Effects Regression Model

- Start from the population model:

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 Z_i + u_{it}$$

  - Y_{it}: fatality rate in state i at time t
  - X_{it}: beer tax
  - Z_i: unobserved factor (e.g., culture) that does **not** vary over time

- Since Z_i is time-invariant, define $\alpha_i \equiv \beta_0 + \beta_2 Z_i$. Then:

Fixed Effects Model

$$Y_{it} = \beta_1 X_{it} + \alpha_i + u_{it}$$

α_i is a **unit-specific intercept** (entity fixed effect).

Interpretation of Fixed Effects

- Each α_i represents the average effect of **being in entity** i.
- These are called **entity fixed effects**.

- They account for all unobserved factors that:
  - Vary **across** entities (states)
  - Remain **constant** over time

- In practice, we do **not** estimate α_i directly as a standalone parameter.
- Instead, we use **binary dummy variables** for each entity.

Dummy Variable Formulation

- Define binary dummies:
  - $D_{2i} = 1$ if $i = 2$, 0 otherwise
  - $D_{3i} = 1$ if $i = 3$, 0 otherwise
  - ... up to D_{ni}
- Omit the first dummy (D_1) to avoid perfect multicollinearity.
- The model becomes:

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \gamma_2 D_{2i} + \gamma_3 D_{3i} + \cdots + \gamma_n D_{ni} + u_{it}$$

- This model includes:
  - One common intercept β_0
  - $n - 1$ entity dummies
  - Same slope β_1 for all entities

Equivalence of the Two Formulations

Fixed Effects Form	Dummy Variable Form
$Y_{it} = \beta_1 X_{it} + \alpha_i + u_{it}$	$Y_{it} = \beta_0 + \beta_1 X_{it} + \sum_{j=2}^{n} \gamma_j D_{ji} + u_{it}$
For $i = 1$: $\alpha_1 = \beta_0$	Baseline intercept for omitted group
For $i \geq 2$: $\alpha_i = \beta_0 + \gamma_i$	Shift relative to baseline

- Both models:
  - Control for Z_i (unobserved, fixed over time)
  - Impose **same slope** β_1 for all entities
  - Differ only in how intercepts are parameterized

Extension to Multiple Regressors

- If other variables besides X affect Y and vary over time, they must be included.
- Otherwise: **omitted variable bias** in estimating β_1.

Fixed Effects Model – Multiple Regressors

$$Y_{it} = \beta_1 X_{1,it} + \cdots + \beta_k X_{k,it} + \alpha_i + u_{it}$$

Equivalently, using binary indicators (omitting the first entity):

$$\begin{aligned} Y_{it} = \beta_0 + \beta_1 X_{1,it} + \cdots + \beta_k X_{k,it} \\ + \gamma_2 D_{2i} + \gamma_3 D_{3i} + \cdots + \gamma_n D_{ni} + u_{it} \end{aligned}$$

Estimation and the Within Estimator

Estimation and Inference

- The fixed effects model can in principle be estimated by OLS.
- The model includes $k+n$ regressors: k covariates plus $n-1$ binary dummies plus an intercept.

- In practice, this direct OLS approach can be:
  - **Tedious** to implement
  - **Infeasible** in software when n is large

Practical Solution

Econometric software uses **special routines** mathematically equivalent to OLS with full dummies, but computationally much faster – exploiting the algebraic structure of fixed effects.

The Entity-Demeaned OLS Algorithm

- The fixed effects estimator is computed in two steps:

  1. **Demeaning step:** subtract entity-specific time averages from each variable
  2. **OLS step:** run OLS on the demeaned variables

- Consider the fixed effects model with a single regressor:

$$Y_{it} = \beta_1 X_{it} + \alpha_i + u_{it}$$

- Take time averages for each entity i:

$$\bar{Y}_i = \frac{1}{T}\sum_{t=1}^{T} Y_{it}, \quad \bar{X}_i = \frac{1}{T}\sum_{t=1}^{T} X_{it}, \quad \bar{u}_i = \frac{1}{T}\sum_{t=1}^{T} u_{it}$$

Demeaning Step – Remove Fixed Effects

- Averaging the model over time gives:

$$\bar{Y}_i = \beta_1 \bar{X}_i + \alpha_i + \bar{u}_i$$

- Subtracting from the original equation:

$$Y_{it} - \bar{Y}_i = \beta_1 (X_{it} - \bar{X}_i) + (u_{it} - \bar{u}_i)$$

- Define **demeaned variables**:

$$\tilde{Y}_{it} \equiv Y_{it} - \bar{Y}_i, \quad \tilde{X}_{it} \equiv X_{it} - \bar{X}_i$$

- The fixed effect α_i **drops out** – this is exactly the point.

OLS on Demeaned Variables

- The demeaned regression model:

$$\tilde{Y}_{it} = \beta_1 \tilde{X}_{it} + \tilde{u}_{it}$$

- Estimate β_1 by running OLS on $\tilde{Y}_{it}$ and $\tilde{X}_{it}$.

Equivalence Result

This method yields the **same estimator** as OLS with $n - 1$ entity dummies, but is more computationally efficient. Software uses this approach automatically.

The Within Estimator

- The estimator obtained from demeaning is called the **within estimator**.
- It focuses exclusively on **within-unit variation**:
  - Does **not** use differences **between** entities
  - Ignores whether a unit has high or low average levels of X and Y

Interpretation

The within estimator captures how **deviations from entity means** in X explain **deviations from entity means** in Y. Only the **fluctuations around each unit's own average** matter for estimation.

Fixed Effect Estimator – Derivation (1/4)

- The minimization problem:

$$(\hat{\beta}, \hat{\alpha}_1, \ldots, \hat{\alpha}_n) = \arg \min_{b,\, a_1, \ldots, a_n} \sum_{i=1}^{n} \sum_{t=1}^{T} (Y_{it} - bX_{it} - a_i)^2$$

- Note on double summation:

$$\sum_{i=1}^{n} \sum_{t=1}^{T} X_{it} = \sum_{i=1}^{n} (X_{i1} + X_{i2} + \cdots + X_{iT})$$

- That is: sum over all time periods for each entity, then sum across entities.

Fixed Effect Estimator – Derivation (2/4)

- First-order conditions (FOC) for b:

$$\sum_{i=1}^{n}\sum_{t=1}^{T}\left(Y_{it}-\hat{\beta}X_{it}-\hat{\alpha}_i\right)X_{it}=0$$

- First-order conditions for each a_i:

$$\sum_{t=1}^{T}\left(Y_{it}-\hat{\beta}X_{it}-\hat{\alpha}_i\right)=0$$

Fixed Effect Estimator – Derivation (3/4)

- From the FOC for a_i, solving for $\hat{\alpha}_i$:

$$\hat{\alpha}_i = \frac{1}{T}\sum_{t=1}^{T}\left(Y_{it} - \hat{\beta}X_{it}\right) = \bar{Y}_i - \bar{X}_i\hat{\beta}$$

- Plug $\hat{\alpha}_i$ into the FOC for b:

$$\sum_{i=1}^{n}\sum_{t=1}^{T}(Y_{it} - X_{it}\hat{\beta} - \bar{Y}_i + \bar{X}_i\hat{\beta})X_{it} = 0$$

$$\left(\sum_{i=1}^{n}\sum_{t=1}^{T}(Y_{it} - \bar{Y}_i)X_{it}\right) - \hat{\beta}\left(\sum_{i=1}^{n}\sum_{t=1}^{T}(X_{it} - \bar{X}_i)X_{it}\right) = 0$$

Fixed Effect Estimator – Derivation (4/4)

- Solving for $\hat{\beta}$:

$$\hat{\beta}_{FE} = \frac{\sum_{i=1}^{n}\sum_{t=1}^{T}(X_{it} - \bar{X}_i)(Y_{it} - \bar{Y}_i)}{\sum_{i=1}^{n}\sum_{t=1}^{T}(X_{it} - \bar{X}_i)^2} = \frac{\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{X}_{it}\tilde{Y}_{it}}{\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{X}_{it}^2}$$

- This is the **same estimator** as the one obtained in the demeaned method – confirming equivalence.

Assumptions, Properties, and Inference

Fixed Effects Regression Assumptions

- The simple fixed effects model: $Y_{it} = \beta_1 X_{it} + \alpha_i + u_{it}$
  - **Assumption 1 – Strict exogeneity:**

$$E(u_{it} \mid X_{i1}, X_{i2}, \dots, X_{iT}, \alpha_i) = 0$$

  - **Assumption 2 – i.i.d. across entities:**

$$(X_{i1}, \dots, X_{iT}, u_{i1}, \dots, u_{iT}) \text{ are i.i.d. across } i$$

  - **Assumption 3:** Large outliers are unlikely (finite fourth moments).
  - **Assumption 4:** No perfect multicollinearity.
- For multiple regressors, X_{it} represents $X_{1,it}, \dots, X_{k,it}$.

Statistical Properties – Unbiasedness and Consistency

- Recall:

$$\hat{\beta}_{FE} = \frac{\sum_{i,t} \tilde{X}_{it}\tilde{Y}_{it}}{\sum_{i,t} \tilde{X}_{it}^2}$$

- Substituting $\tilde{Y}_{it} = \beta_1 \tilde{X}_{it} + \tilde{u}_{it}$:

$$\hat{\beta}_{FE} = \beta_1 + \frac{\sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{X}_{it}\tilde{u}_{it}}{\sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{X}_{it}^2}$$

Result

Under Assumptions 1–4, $\hat{\beta}_{FE}$ is **unbiased** and **consistent**. This parallels the derivation of the OLS estimator.

Asymptotic Distribution of $\hat{\beta}_{FE}$

- From the decomposition above:

$$\hat{\beta}_{FE} = \beta_1 + \frac{\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{X}_{it}\tilde{u}_{it}}{\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{X}_{it}^2}$$

- Multiply both sides by $\sqrt{nT}$:

$$\sqrt{nT}(\hat{\beta}_{FE} - \beta_1) = \frac{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_i}{\hat{Q}_{\tilde{X}}}, \quad \text{where} \quad \eta_i = \sqrt{\frac{1}{T}}\sum_{t=1}^{T}\tilde{X}_{it}\tilde{u}_{it}, \quad \hat{Q}_{\tilde{X}} = \frac{1}{nT}\sum_{i,t}\tilde{X}_{it}^2$$

- In panel data, typically $n \gg T$, so we use asymptotics with $n \to \infty$, T fixed.

Asymptotic Distribution (Cont.)

- Under fixed effects assumptions:

$$\hat{Q}_{\tilde{X}} \xrightarrow{p} Q_{\tilde{X}} = E\left[T^{-1}\sum_{t=1}^{T}\tilde{X}_{it}^2\right]$$

- The η_i are i.i.d. with mean 0 and variance σ_η^2.

- By the Central Limit Theorem:

$$\sqrt{nT}(\hat{\beta}_{FE} - \beta_1) \xrightarrow{d} N\left(0, \frac{\sigma_\eta^2}{Q_{\tilde{X}}^2}\right)$$

- Asymptotic variance of $\hat{\beta}_{FE}$:

$$\text{var}(\hat{\beta}_{FE}) = \frac{1}{nT} \cdot \frac{\sigma_\eta^2}{Q_{\tilde{X}}^2}$$

Clustered Standard Errors (1/2)

- Use sample counterparts to estimate the asymptotic variance:

$$SE(\hat{\beta}_{FE}) = \sqrt{\frac{1}{nT} \cdot \frac{s_\eta^2}{\hat{Q}_{\tilde{X}}^2}}$$

- Sample variance of $\hat{\eta}_i$:

$$s_\eta^2 = \frac{1}{n-1} \sum_{i=1}^{n} \hat{\eta}_i^2, \quad \text{where} \quad \hat{\eta}_i = \sqrt{\frac{1}{T}} \sum_{t=1}^{T} \tilde{X}_{it} \hat{u}_{it}$$

- Note: $\bar{\hat{\eta}} = 0$ because regressors and residuals are orthogonal by construction.

Clustered Standard Errors (2/2)

- The $\hat{\eta}_i$ are formed by **aggregating over time** for each entity – this is clustering at the entity level.

Important

Why Clustering Matters

Clustered standard errors are robust to:

- **Autocorrelation** within each entity i over time
- **Heteroskedasticity** across entities i

Without clustering, we incorrectly treat nT observations as independent.

- t-tests using clustered SEs are valid as $n \to \infty$
- For joint tests: F-statistic follows $F_{q,\infty}$ asymptotically

Why Clustering Matters – An Example

Example – Worker Productivity

Let Y_{it} = worker i's monthly productivity in month t.
If u_{it} includes unobserved factors like motivation or stress, a stressful January (u_{i1}) may carry over into February (u_{i2}). So u_{i1} and u_{i2} are **positively correlated within the same worker**.

- This is **autocorrelation within the same individual over time**.
- Ignoring it leads to:
  - **Underestimated** standard errors
  - **Overconfident** inference – we reject too often

- **Clustered standard errors** allow u_{it} to be correlated **within** i but **not across** i.

Time Fixed Effects and Two-Way Fixed Effects

Regression with Time Fixed Effects

- Time fixed effects control for factors that:
  - **Vary over time**, but are **constant across entities**
- **Example:** national automobile safety standards improve each year – they affect all states equally

- Omitted variable bias may arise from S_t in:

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 Z_i + \beta_3 S_t + u_{it}$$

- S_t = unobserved variable that changes **over time** but not across entities
- If S_t correlates with X_{it}, omitting it biases $\hat{\beta}_1$

Time-Only Fixed Effects Model

- Suppose Z_i is absent:

$$Y_{it} = \beta_1 X_{it} + \lambda_t + u_{it}$$

- λ_t is a **time fixed effect** for period t:
  - Captures the effect of omitted variables (like S_t) that **change over time**
  - Analogous to entity fixed effects, but varies by **time**, not entity
  - Absorbs time-specific shocks **common to all units**

Note

In this model, λ_t varies across t but is constant across all entities in period t.

Both Entity and Time Fixed Effects – Two-Way FE

- Some omitted variables vary across entities (e.g., culture) while others vary over time (e.g., national policies).

Two-Way Fixed Effects Model

$$Y_{it} = \beta_1 X_{it} + \alpha_i + \lambda_t + u_{it}$$

where α_i = entity fixed effect, λ_t = time fixed effect.

- Equivalently, using dummy variables:

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \gamma_2 D2_i + \cdots + \gamma_n Dn_i + \delta_2 B2_t + \cdots + \delta_T BT_t + u_{it}$$

- Controls for both time-invariant and entity-invariant omitted variables simultaneously.

Estimating Two-Way Fixed Effects

- Estimate via OLS with:
  - $n-1$ entity dummies
  - $T-1$ time dummies
- Alternative (efficient) method in balanced panels – **double demeaning**:
  - Demean Y_{it} and X_{it} by entity mean
  - Then demean by time-period mean
  - Regress double-demeaned Y on double-demeaned X

Note

Special case: if $T=2$, two-way fixed effects reduces to the first-difference model:

$$\Delta Y_i = \beta_1 \Delta X_i + \Delta u_i$$

Common in difference-in-differences setups.

Application – Drunk Driving Laws

Application – Drunk Driving Laws and Traffic Deaths

- States differ in drunk driving punishments, beer tax levels, and other economic/legal policies.
- Omitting these differences introduces **omitted variable bias** in OLS regressions.

- Example: OLS without fixed effects:
  - Coefficient on beer tax $= +0.36 \rightarrow$ suggests tax **increases** fatalities
  - Likely biased due to omitted factors

Key Finding

Including **state fixed effects** reverses the sign: the beer tax coefficient becomes -0.66, indicating that higher taxes are associated with **fewer** deaths.

Estimation Results

- Column (1): OLS, no FE – positive biased coefficient
- Column (2): State FE only – sign reversal
- Column (3): State + time FE – sign preserved, less precise
- Column (4): Full controls added

Dependent variable: traffic fatality rate (deaths per 10,000).

Regressor	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Beer tax	0.36 (0.05) [0.26, 0.46]	−0.66 (0.29) [−1.23, −0.09]	−0.64 (0.36) [−1.35, 0.07]	−0.45 (0.30) [−1.04, 0.14]	−0.69 (0.35) [−1.38, 0.00]	−0.46 (0.31) [−1.07, 0.15]	−0.93 (0.34) [−1.60, −0.26]
Drinking age 18		0.10		0.03 (0.07) [−0.11, 0.17]	−0.01 (0.08) [−0.17, 0.15]		0.04 (0.10) [−0.16, 0.24]
Drinking age 19				−0.02 (0.05) [−0.12, 0.08]	−0.08 (0.07) [−0.21, 0.06]		−0.07 (0.10) [−0.26, 0.13]
Drinking age 20				0.03 (0.05) [−0.07, 0.13]	−0.10 (0.06) [−0.21, 0.01]		−0.11 (0.13) [−0.36, 0.14]
Drinking age						0.00 (0.02) [−0.05, 0.04]	
Mandatory jail or community service?				0.04 (0.10) [−0.17, 0.25]	0.09 (0.11) [−0.14, 0.31]	0.04 (0.10) [−0.17, 0.25]	0.09 (0.16) [−0.24, 0.42]
Average vehicle miles per driver				0.008 (0.007)	0.017 (0.011)	0.009 (0.007)	0.124 (0.049)
Unemployment rate				−0.063 (0.013)		−0.063 (0.013)	−0.091 (0.021)
Real income per capita (logarithm)				1.82 (0.64)		1.79 (0.64)	1.00 (0.68)
Years	1982–88	1982–88	1982–88	1982–88	1982–88	1982–88	1982 & 1988 only
State effects?	no	yes	yes	yes	yes	yes	yes
Time effects?	no	no	yes	yes	yes	yes	yes
Clustered standard errors?	no	yes	yes	yes	yes	yes	yes

Adding Time Effects and Controls

- Time fixed effects: small impact on beer tax coefficient, now less precisely estimated.
- Controls added (column 4):
  - Drunk driving laws (minimum drinking age, punishment type)
  - Driving/economic controls: vehicle miles, unemployment, log income
- Estimated beer tax coefficient:

$$-0.45 \quad \text{(vs. } -0.66 \text{ in the state-FE-only model)}$$

- Interpretation: a \$0.50 tax increase → approximately 0.23 fewer deaths per 10,000

Note

Wide confidence intervals – interpret the magnitude with caution.

Results for Policy Variables

1. **Minimum drinking age:**
   - Coefficients are small
   - Not jointly significant ($p = 0.786$)
2. **Punishment severity:**
   - Also small and statistically insignificant
3. **Economic variables:**
   - Higher unemployment $\rightarrow$ fewer fatalities:

$$1\% \text{ increase} \Rightarrow -0.063 \text{ deaths per 10,000}$$

   - Higher income $\rightarrow$ more fatalities:

$$1\% \text{ increase} \Rightarrow +0.0182 \text{ deaths per 10,000}$$

Note

Interpretation: better economic conditions $\rightarrow$ more driving $\rightarrow$ more fatalities.

Sensitivity Checks and Interpretation

- Columns (5)–(7) check robustness:
  - Dropping economic variables → beer tax becomes significant again
  - Results stable across specifications
- Main conclusions:
  - **State and time fixed effects are crucial –** without them, results are misleading
  - **Beer taxes appear to reduce fatalities**, likely via reduced alcohol consumption
  - **Other policy variables** show limited direct effect

Takeaway

Beer taxes may be effective, but the results must be interpreted carefully given wide confidence intervals and potential remaining confounders.

Summary

Summary – Regression with Panel Data (1/2)

1. Panel data track multiple entities (n) over two or more time periods (T).
2. **Entity fixed effects** control for unobserved variables that differ across entities but are **constant over time**.
3. With $T = 2$, fixed effects regression reduces to a **first-difference** regression:

$$\Delta Y_i = \beta_1 \Delta X_i + \Delta u_i$$

4. Entity fixed effects are estimated via OLS by including $n - 1$ entity dummies plus the X variables and an intercept.

Summary – Regression with Panel Data (2/2)

5. **Time fixed effects** control for unobserved variables that are constant **across entities** but **vary over time**.

6. **Two-way fixed effects** include:
   - $n - 1$ entity dummies
   - $T - 1$ time dummies
   - The X's and an intercept

7. Panel data often exhibit **autocorrelation within entities** over time:
   - Standard errors must allow for within-entity autocorrelation and heteroskedasticity
   - **Clustered standard errors** (clustered at the entity level) are the standard solution

Required Reading

- **Stock and Watson (2020)** *Introduction to Econometrics*, Chapter 10
  - Sections 10.1 – 10.7
  - Appendices 10.2