

The plot of our course

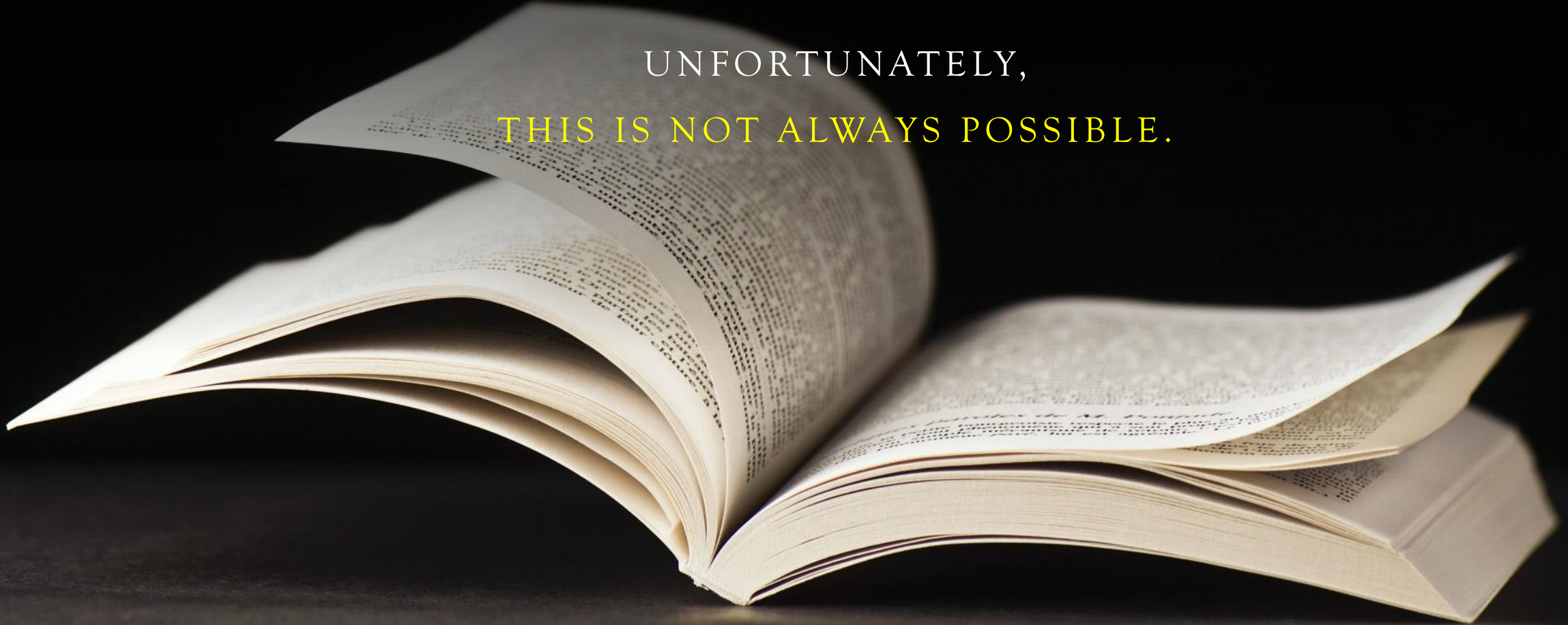
FELIPE BUCHBINDER



WHEN WE BUILD A REGRESSION MODEL,
WE WOULD LIKE IT TO CONTAIN ALL X'S
THAT ARE IMPORTANT TO EXPLAIN OUR Y...

UNFORTUNATELY,

THIS IS NOT ALWAYS POSSIBLE.



Why not?





We might not have all the variables



Some variables may not be observable

Car insurance

$$\text{Expected cost} = \beta_0 + \beta_1 \cdot X + \epsilon$$

Observable variables:

- Age
- Gender
- Lives in a metropolitan area?
- (any other ideas?)

Car insurance

$$\text{Expected cost} = \beta_0 + \beta_1 \cdot X + \epsilon$$

Unobservable variables:

- Driving recklessly
- Goes to party often?
- Health issues (e.g. likelihood of syncope due to heart arrhythmia)
- (any other ideas?)

Another example:

Impact of democracy on child mortality

Ross, M. (2006). Is democracy good for the poor?. *American Journal of Political Science*, 50(4), 860-874.

$$\text{Child Mortality} = \beta_0 + \beta_1 \cdot \text{Democracy} + \beta_2 \cdot X + \epsilon$$

Observable variables

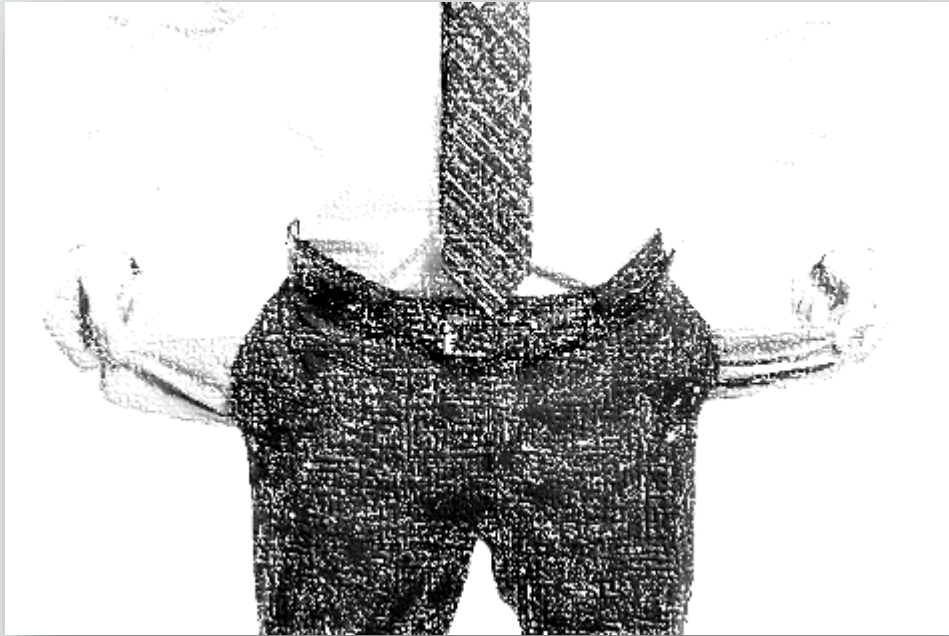
- GDP per capita
- Investment in health per capita
- Number of hospitals per capita
- (any other ideas?)

$$\text{Child Mortality} = \beta_0 + \beta_1 \cdot \text{Democracy} + \beta_2 \cdot X + \epsilon$$

Unobservable variables

- Health habits
- Dietary habits
- Cultural aspects regarding the caring and nourishing of children
- Existence of conflicts zones
- Etc.

Variables that differ between entities but are not observable are called **unobserved heterogeneities**.



We might not have all the variables

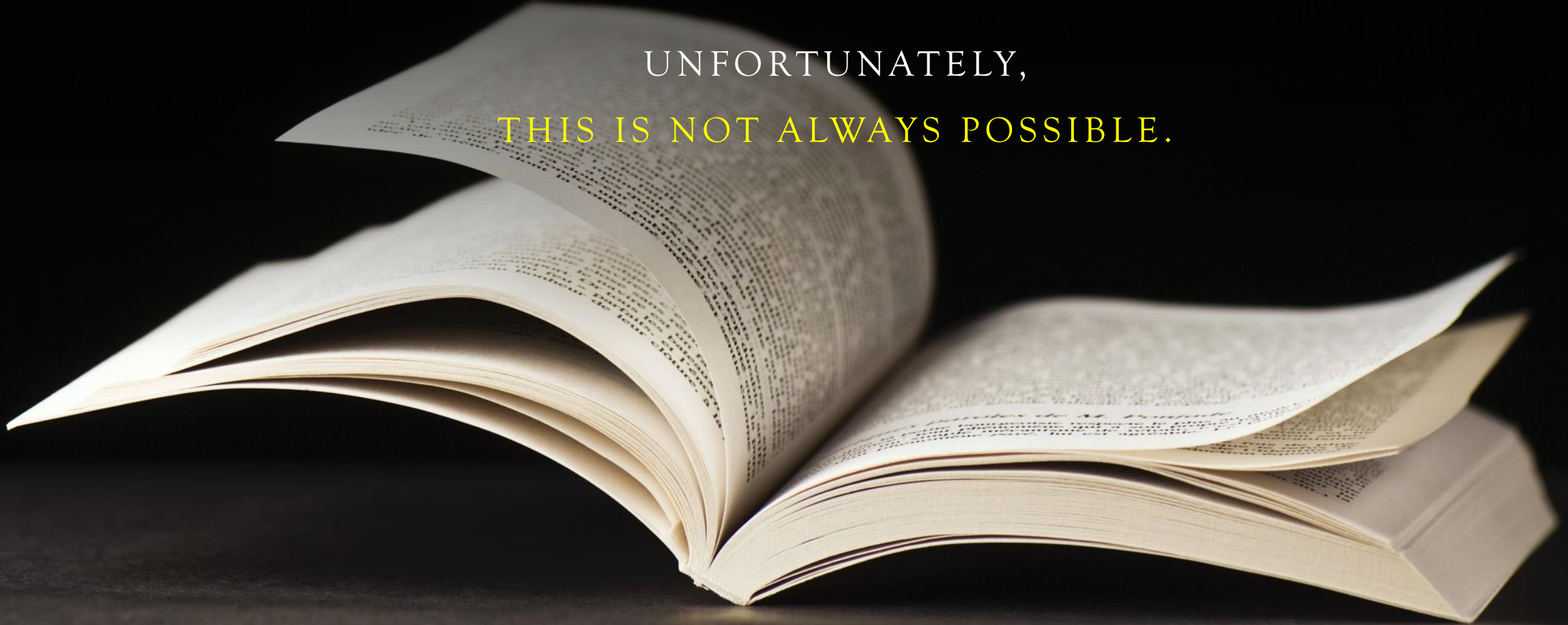


Some variables may not be observable

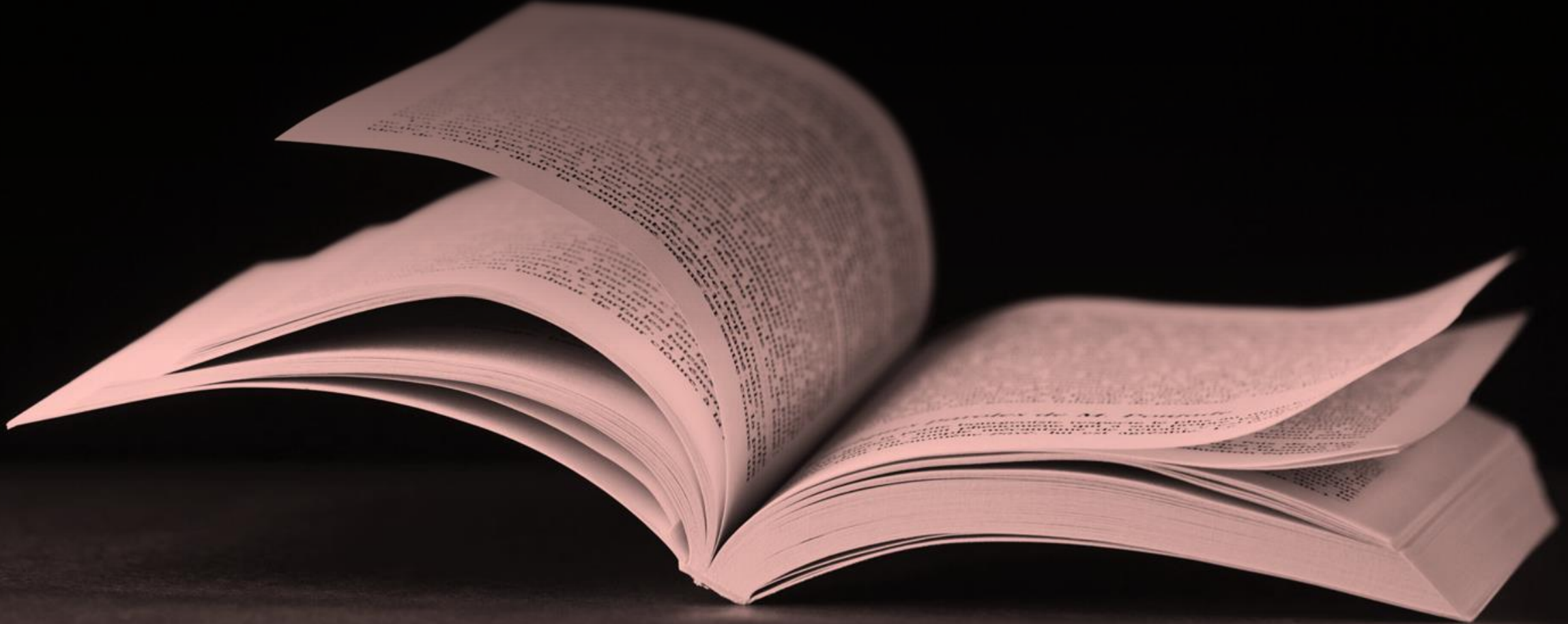
WHEN WE BUILD A REGRESSION MODEL,
WE WOULD LIKE IT TO CONTAIN ALL X'S
THAT ARE IMPORTANT TO EXPLAIN OUR Y...

UNFORTUNATELY,

THIS IS NOT ALWAYS POSSIBLE.



WHEN AN IMPORTANT X IS MISSING IN OUR
MODEL, **BAD THINGS CAN HAPPEN.**



For example...?



Biased
regression
coefficients



Linear Regression refresher

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$
$$\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}; \sigma^2 \mathbf{I})$$

$$\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$$

\mathbf{b} is unbiased, meaning $\mathbb{E}(\mathbf{b}) = \boldsymbol{\beta}$

Now suppose there's a variable, \mathbf{U} , that affects \mathbf{Y} but we fail to put it in our model. We simply calculate \mathbf{b} without it!

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{U} + \boldsymbol{\epsilon}$$
$$\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}; \sigma^2 \mathbf{I})$$

$$\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$$

Now suppose there's a variable, \mathbf{U} , that affects \mathbf{Y} but we fail to put it in our model. We simply calculate \mathbf{b} without it!

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{U} + \boldsymbol{\epsilon}$$
$$\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}; \sigma^2 \mathbf{I})$$

$$\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$$

In this case, \mathbf{b} is no longer an unbiased estimate of $\boldsymbol{\beta}$!

$$\begin{aligned}
\mathbf{b} &= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \\
&= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T (\mathbf{X} \boldsymbol{\beta} + \mathbf{U} + \boldsymbol{\epsilon}) \\
&= \underbrace{(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X}}_{\mathbf{I}} \boldsymbol{\beta} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{U} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \boldsymbol{\epsilon} \\
&= \boldsymbol{\beta} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{U} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \boldsymbol{\epsilon}
\end{aligned}$$

Taking the expected value...

$$\begin{aligned}
\mathbb{E}(\mathbf{b}) &= \boldsymbol{\beta} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{U} + (\mathbf{X}^T \mathbf{X})^{-1} \underbrace{\mathbb{E}(\mathbf{X}^T \boldsymbol{\epsilon})}_0 \\
&\quad \therefore \\
\mathbb{E}(\mathbf{b}) &= \boldsymbol{\beta} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{U}
\end{aligned}$$

Thus, in general, $\mathbb{E}(\mathbf{b}) \neq \boldsymbol{\beta}$:

b is a biased estimate of **β**

I propose a name for this Theorem...

$$\begin{aligned}\mathbf{b} &= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \\ &= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T (\mathbf{X} \boldsymbol{\beta} + \mathbf{U} + \boldsymbol{\epsilon}) \\ &= \underbrace{(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X}}_{\mathbf{I}} \boldsymbol{\beta} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{U} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \boldsymbol{\epsilon} \\ &= \boldsymbol{\beta} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{U} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \boldsymbol{\epsilon}\end{aligned}$$

Taking the expected value...

$$\mathbb{E}(\mathbf{b}) = \boldsymbol{\beta} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{U} + (\mathbf{X}^T \mathbf{X})^{-1} \underbrace{\mathbb{E}(\mathbf{X}^T \boldsymbol{\epsilon})}_0$$

$$\begin{aligned}&\therefore \\ \mathbb{E}(\mathbf{b}) &= \boldsymbol{\beta} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{U}\end{aligned}$$

Thus, in general, $\mathbb{E}(\mathbf{b}) \neq \boldsymbol{\beta}$:

b is a biased estimate of **β**

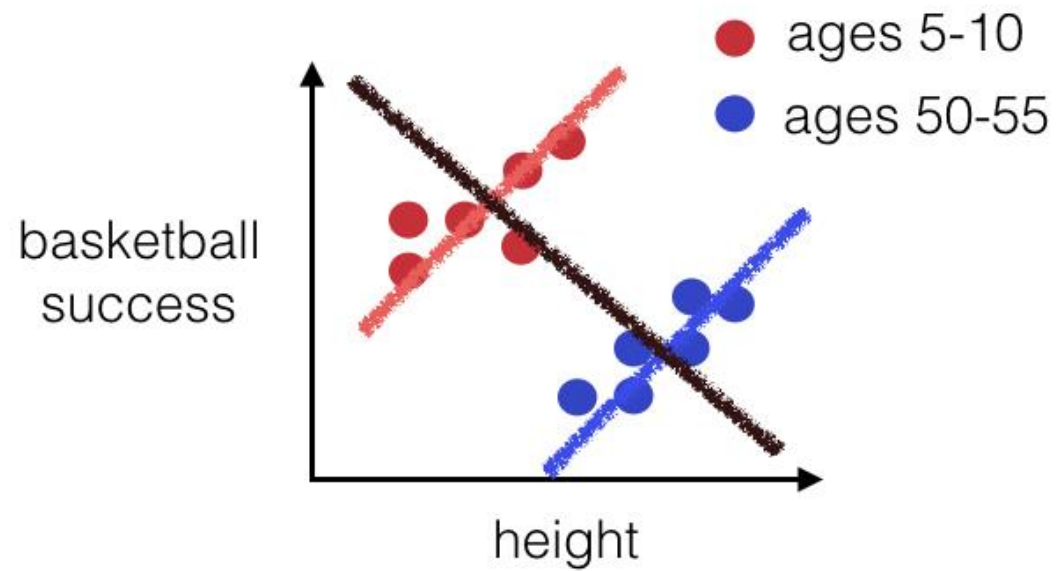
A close-up portrait of Maleficent, the character from Disney's 2014 film "Maleficent". She is shown from the chest up, wearing her iconic black horned headdress and a black cape with a blue lining. She has pale skin, green eyes, and red lips. The background is a misty, mountainous landscape with a soft, golden light. A light gray, shield-shaped callout box with a dotted border is positioned on the right side of the image, containing the title and text.

Maleficent's Theorem

A variable that wasn't invited
to a regression will curse the
coefficients of those that
were, making them biased

An extreme scenario:
Simpson's Paradox

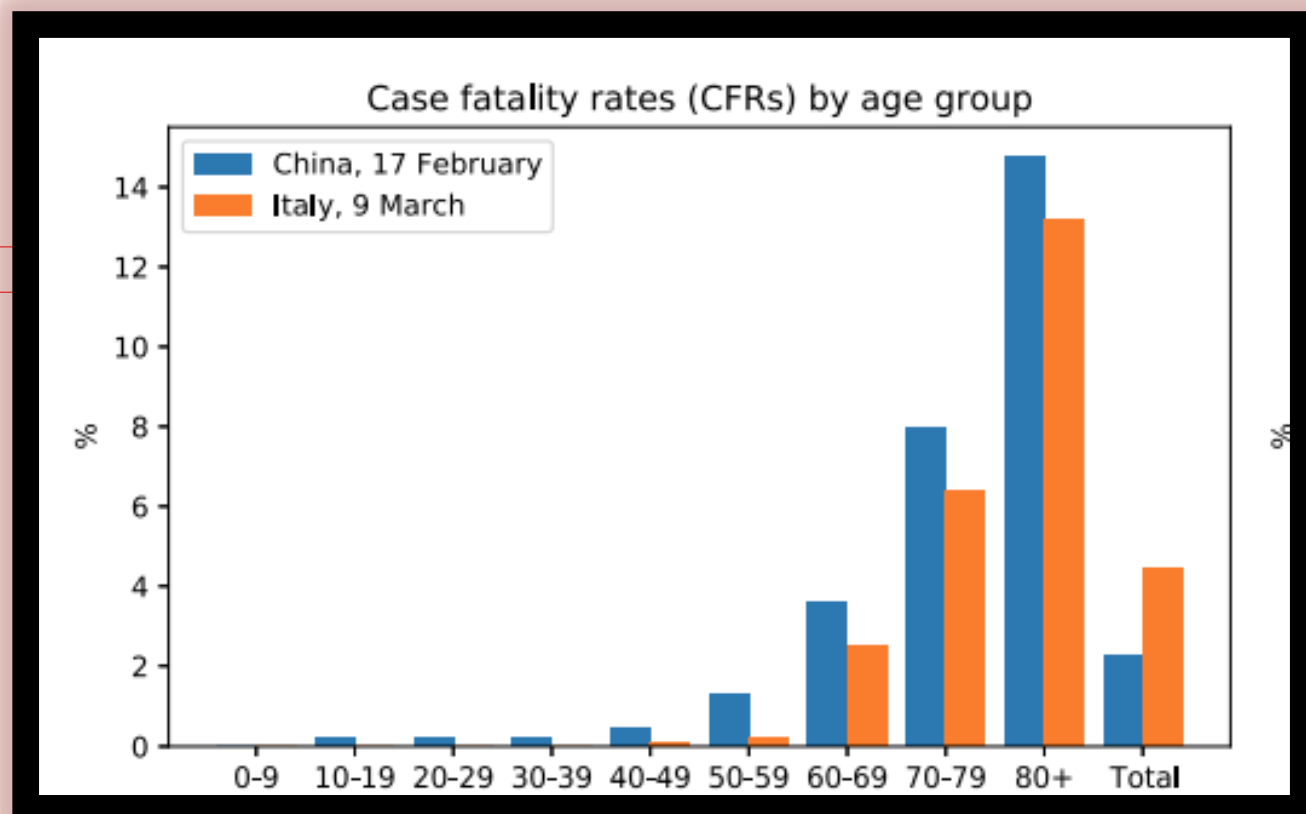
Wanna be good at Basketball? Be short!



Simpson's Paradox in COVID-19 Case Fatality Rates: A Mediation Analysis of Age-Related Causal Effects

Julius von Kügelgen^{1b}, Luigi Gresele^{1b}, and Bernhard Schölkopf^{1b}

Abstract—We point out an instantiation of Simpson's paradox in COVID-19 case fatality rates (CFRs): comparing a large-scale study from China (February 17) with early reports from Italy (March 9), we find that CFRs are lower in Italy for every age group, but higher overall. This phenomenon is explained by a stark difference in case demographic between the two countries. Using this as a motivating example, we introduce basic concepts from mediation analysis and show how these can be used to quantify different direct and indirect effects when assuming a coarse-grained causal graph involving country, age, and case fatality. We curate an age-stratified CFR dataset with >750 k cases and conduct a case study, investigating total, direct, and indirect (age-mediated) causal effects between different countries and at different points in time. This allows us to separate age-related effects from others unrelated to age and facilitates a more transparent comparison of CFRs across countries at different stages of the COVID-19 pandemic. Using longitudinal data from Italy, we discover a sign reversal of the direct causal effect in mid-March, which temporally aligns with the reported collapse of the healthcare system in parts of the country. Moreover, we find that direct and indirect effects across 132 pairs of countries are only weakly correlated, suggesting that a country's policy and case

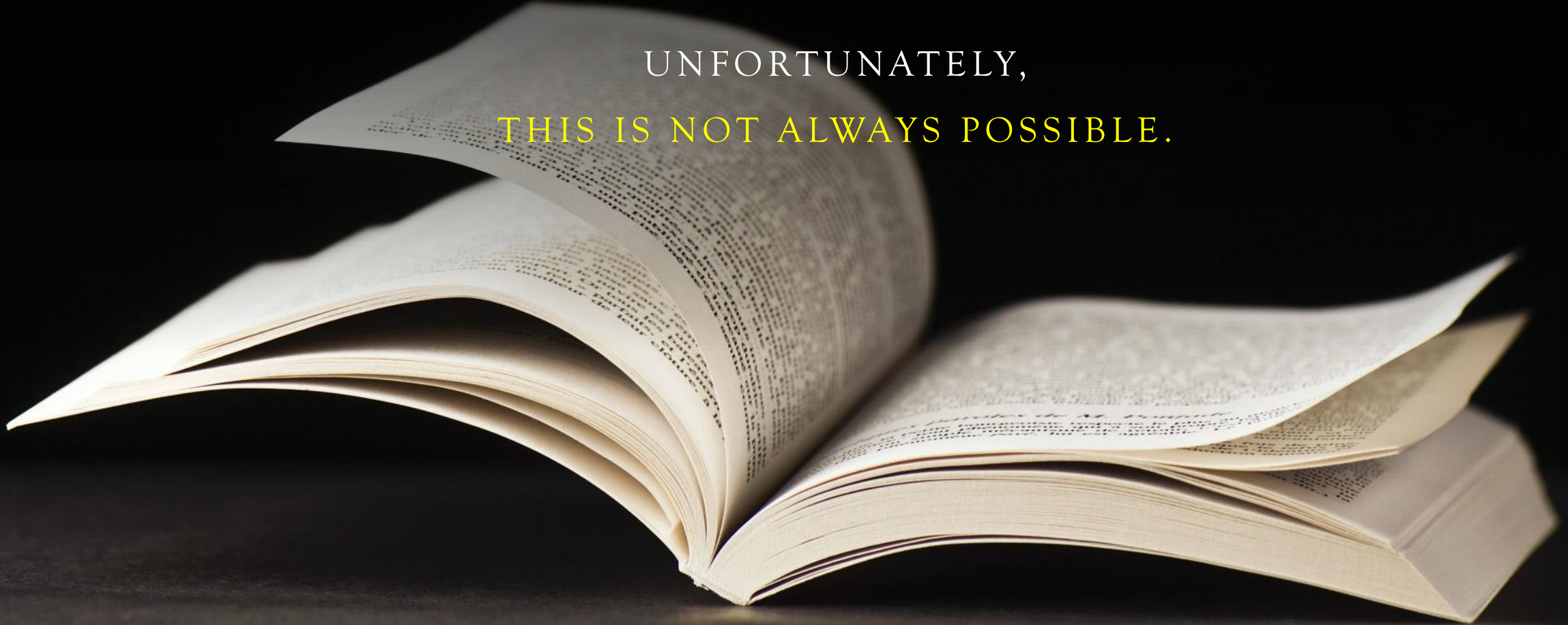


being reported across multiple countries all over the world, ultimately leading to the World Health Organization declaring it a pandemic on March 11, 2020 [1]. As of September 28, 2020, the pandemic led to more than 33 million confirmed cases and

WHEN WE BUILD A REGRESSION MODEL,
WE WOULD LIKE IT TO CONTAIN ALL X'S
THAT ARE IMPORTANT TO EXPLAIN OUR Y...

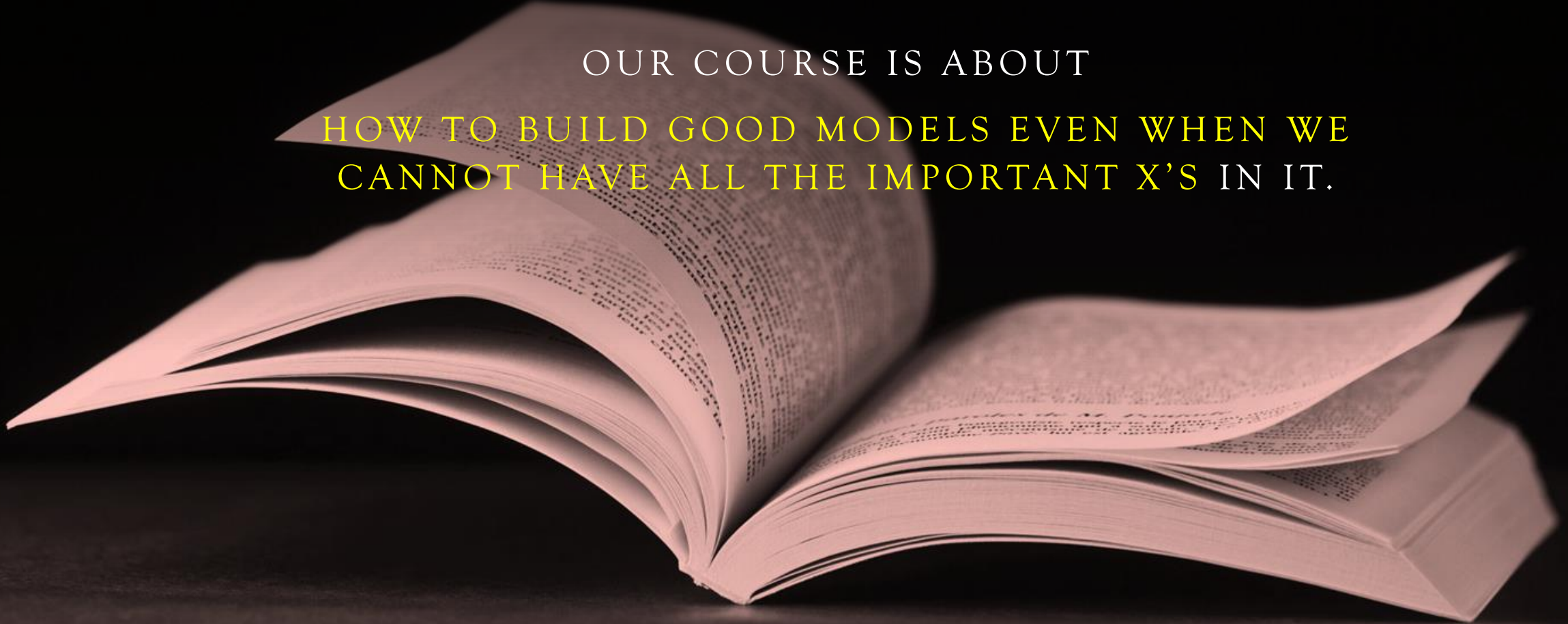
UNFORTUNATELY,

THIS IS NOT ALWAYS POSSIBLE.



WHEN AN IMPORTANT X IS MISSING IN OUR
MODEL, **BAD THINGS CAN HAPPEN.**

OUR COURSE IS ABOUT
**HOW TO BUILD GOOD MODELS EVEN WHEN WE
CANNOT HAVE ALL THE IMPORTANT X'S IN IT.**



Our strategy will be to
analyze things over time,
so we can have a feeling of how things
usually are.

If an X matters, it's effect should be
made visible by observing something
over time and comparing it with
others.



Why do we date before we get engaged?



$$\begin{pmatrix} \text{Will} \\ \text{I be happy?} \end{pmatrix} = \beta_0 + \beta_1 \begin{pmatrix} \text{Is s/he} \\ \text{fun?} \end{pmatrix} + \beta_2 \begin{pmatrix} \text{Do we} \\ \text{have chemistry?} \end{pmatrix} + \beta_3 (\text{Does s/he love me?}) + \epsilon$$

$$\begin{pmatrix} \text{Will} \\ \text{I be happy?} \end{pmatrix} = \beta_0 + \beta_1 \begin{pmatrix} \text{Is s/he} \\ \text{fun?} \end{pmatrix} + \beta_2 \begin{pmatrix} \text{Do we} \\ \text{have chemistry?} \end{pmatrix} + \beta_3 \begin{pmatrix} \text{Does s/he love me?} \end{pmatrix} + \epsilon$$

Unobservable

$$\begin{pmatrix} \text{Will} \\ \text{I be happy?} \end{pmatrix} = \beta_0 + \beta_1 \begin{pmatrix} \text{Is s/he} \\ \text{fun?} \end{pmatrix} + \beta_2 \begin{pmatrix} \text{Do we} \\ \text{have chemistry?} \end{pmatrix} + \beta_3 \begin{pmatrix} \text{Does s/he love me?} \end{pmatrix} + \epsilon$$

Unobservable



Over time, you'll see how
s/he usually acts towards you.

$$\begin{pmatrix} \text{Will} \\ \text{I be happy?} \end{pmatrix} = \beta_0 + \beta_1 \begin{pmatrix} \text{Is s/he} \\ \text{fun?} \end{pmatrix} + \beta_2 \begin{pmatrix} \text{Do we} \\ \text{have chemistry?} \end{pmatrix} + \beta_3 \begin{pmatrix} \text{Does s/he love me?} \end{pmatrix} + \epsilon$$

Unobservable

By dating other people
(not at once! ☺)

you'll be able to see how other people usually treat you
and assess if s/he is special

Over time, you'll see how
s/he usually acts towards you.

$$\begin{pmatrix} \text{Will} \\ \text{I be happy?} \end{pmatrix} = \beta_0 + \beta_1 \begin{pmatrix} \text{Is s/he} \\ \text{fun?} \end{pmatrix} + \beta_2 \begin{pmatrix} \text{Do we} \\ \text{have chemistry?} \end{pmatrix} + \beta_3 \begin{pmatrix} \text{Does s/he love me?} \end{pmatrix} + \epsilon$$

Unobservable

By dating other people
(not at once! ☺)

you'll be able to see how other people usually treat you
and assess if s/he is special

Over time, you'll see how
s/he usually acts towards you.

A less silly example:

Impact of democracy on child mortality

Ross, M. (2006). Is democracy good for the poor?. *American Journal of Political Science*, 50(4), 860-874.

What control variables would you use in this model?

$$\text{Child Mortality} = \beta_0 + \beta_1 \cdot \text{Democracy} + \beta_2 \cdot X + \epsilon$$

What could X be?

- GDP per capita
- Investment in health per capita
- Number of hospitals per capita
- (any other ideas?)

Some things that affect child mortality are
unobserved

$$\text{Child Mortality} = \beta_0 + \beta_1 \cdot \text{Democracy} + \beta_2 X + U + \epsilon$$

What could U be?

- Health habits
- Dietary habits
- Cultural aspects regarding the caring and nourishing of children
- Existence of conflicts zones
- Etc.

Some things that affect child mortality are
unobserved

$$\text{Child Mortality} = \beta_0 + \beta_1 \cdot \text{Democracy} + \beta_2 X + U + \epsilon$$

What could U be?

- Health habits
- Dietary habits
- Cultural aspects regarding the caring and nourishing of children
- Existence of conflicts zones
- Etc.

We can get a sense of these
by observing many countries over time



*LOTS OF THINGS ARE INVISIBLE, BUT WE DON'T
KNOW HOW MANY BECAUSE WE CAN'T **SEE** THEM.*

★ *L'essentiel est invisible
pour les yeux*



To deal with unobserved
heterogeneities, we'll observe
many entities over time



4 kinds of data



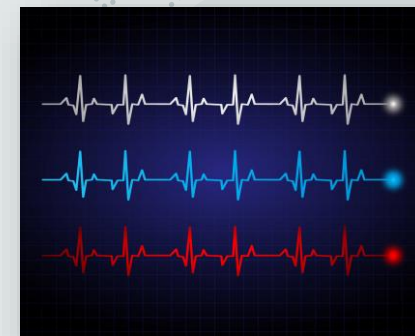
Cross-Sectional



Time Series



Pooled



Panel

The general panel data model

a.k.a. the most important equation in our course



The general panel data model

a.k.a. the most important equation in our course

$$y_{it} = \underbrace{\beta_0}_{\text{Intercept}} + \underbrace{\beta_1 X_{1it} + \dots + \beta_p X_{pit}}_{\substack{X\text{'s are called covariates} \\ \text{and are observed}}} + \underbrace{U_i}_{\text{Unobserved}} + \underbrace{\epsilon_{it}}_{\substack{\text{Idyossincratic} \\ \text{error}}}$$

or, in matrix notation,

$$\mathbf{Y}_{it} = \mathbf{X}_{it}\boldsymbol{\beta} + \mathbf{U}_i + \boldsymbol{\epsilon}_{it}$$

The big question is how to estimate $\boldsymbol{\beta}$ well even though we do not know \mathbf{U} , since it is unobserved.

The general panel data model

a.k.a. the most important equation in our course

$$y_{it} = \underbrace{\beta_0}_{\text{Intercept}} + \underbrace{\beta_1 X_{1it} + \dots + \beta_p X_{pit}}_{\substack{\text{X's are called covariates} \\ \text{and are observed}}} + \underbrace{U_i}_{\text{Unobserved}} + \underbrace{\epsilon_{it}}_{\substack{\text{Idyossincratic} \\ \text{error}}}$$

or, in matrix notation,

$$\mathbf{Y}_{it} = \mathbf{X}_{it}\boldsymbol{\beta} + \mathbf{U}_i + \boldsymbol{\epsilon}_{it}$$

The big question is how to estimate $\boldsymbol{\beta}$ well even though we do not know \mathbf{U} , since it is unobserved.

Data is for each entity
at each time → Panel Data

The general panel data model

a.k.a. the most important equation in our course

$$y_{it} = \underbrace{\beta_0}_{\text{Intercept}} + \underbrace{\beta_1 X_{1it} + \dots + \beta_p X_{pit}}_{\substack{\text{X's are called covariates} \\ \text{and are observed}}} + \underbrace{U_i}_{\text{Unobserved}} + \underbrace{\epsilon_{it}}_{\substack{\text{Idyossincronic} \\ \text{error}}}$$

or, in matrix notation,

$$\mathbf{Y}_{it} = \mathbf{X}_{it}\boldsymbol{\beta} + \mathbf{U}_i + \boldsymbol{\epsilon}_{it}$$

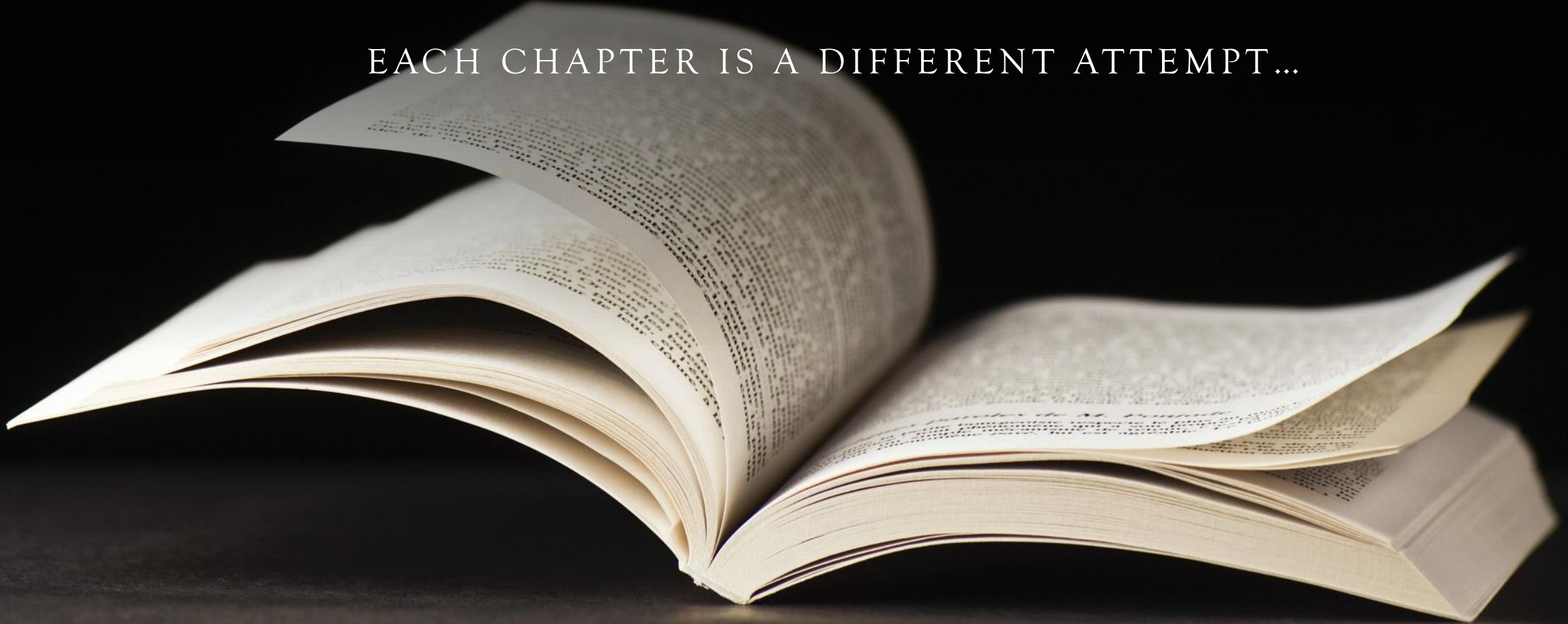
The big question is how to estimate $\boldsymbol{\beta}$ well even though we do not know \mathbf{U} , since it is unobserved.

Data is for each entity
at each time → Panel Data

Unobserved heterogeneity
is assumed to be a characteristic of the entity
that does not vary through time
(entity-fixed effect)

OUR COURSE IS THE STORY OF HOW
HUMANITY TACKLED THE CHALLENGE OF
MAKING A REGRESSION MODEL WITHOUT
KNOWING ALL THE X'S...

EACH CHAPTER IS A DIFFERENT ATTEMPT...



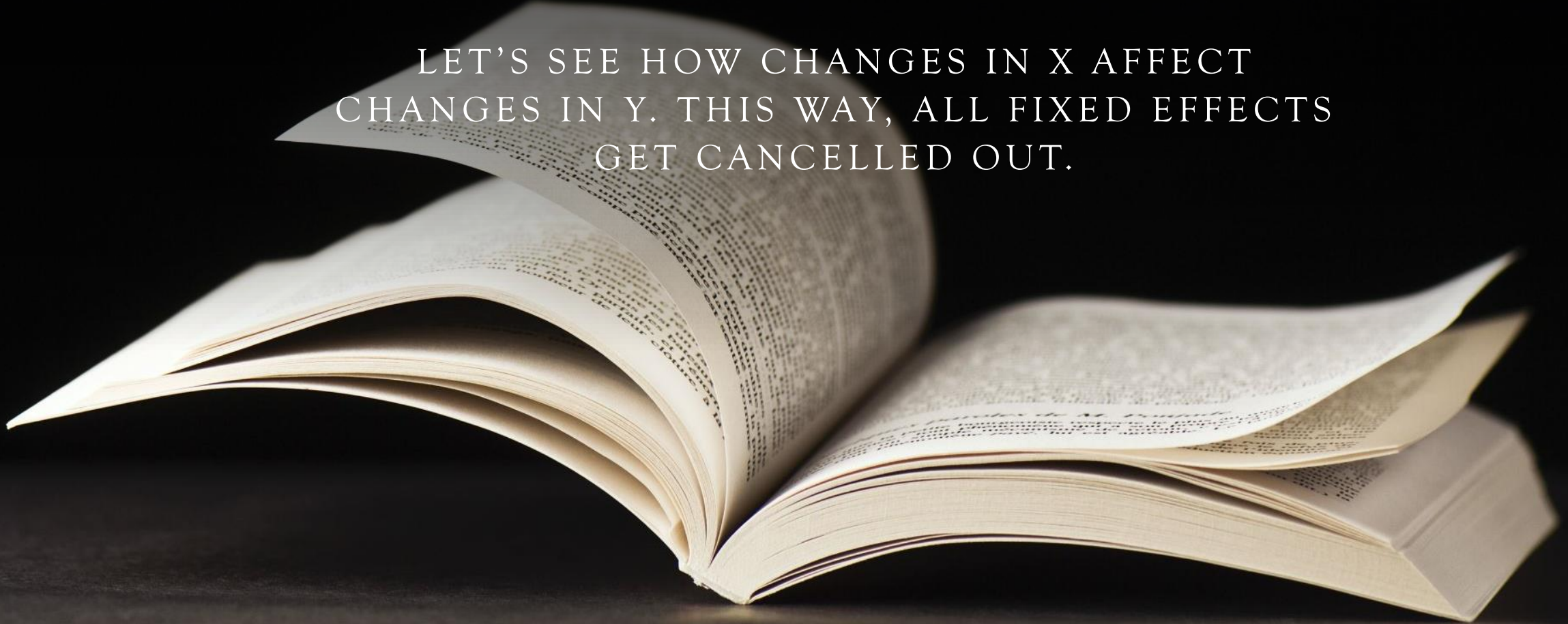
CHAPTER 1: POOLED REGRESSION

WHEN I FORGOT I HAD PANEL DATA AND
SIMPLY RAN A TRADITIONAL REGRESSION ON
EVERYTHING!



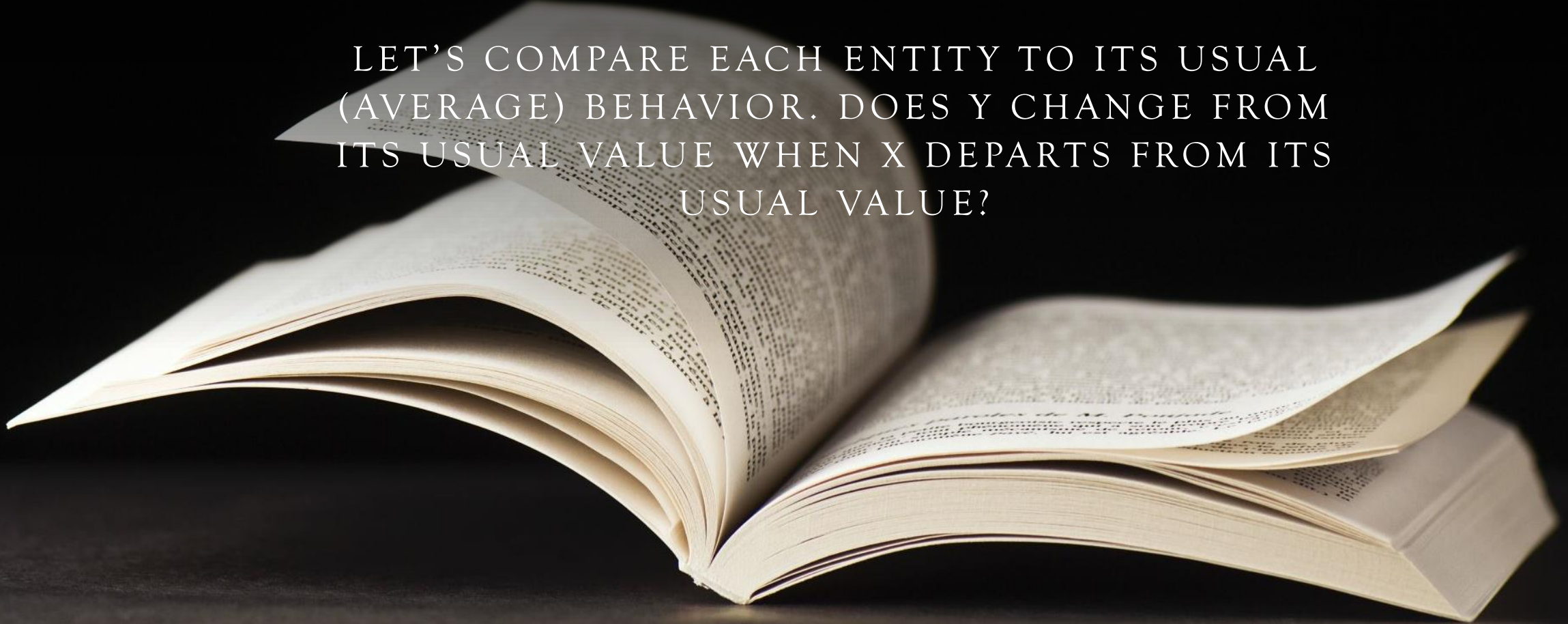
CHAPTER 2: FIRST DIFFERENCES

LET'S SEE HOW CHANGES IN X AFFECT
CHANGES IN Y. THIS WAY, ALL FIXED EFFECTS
GET CANCELLED OUT.



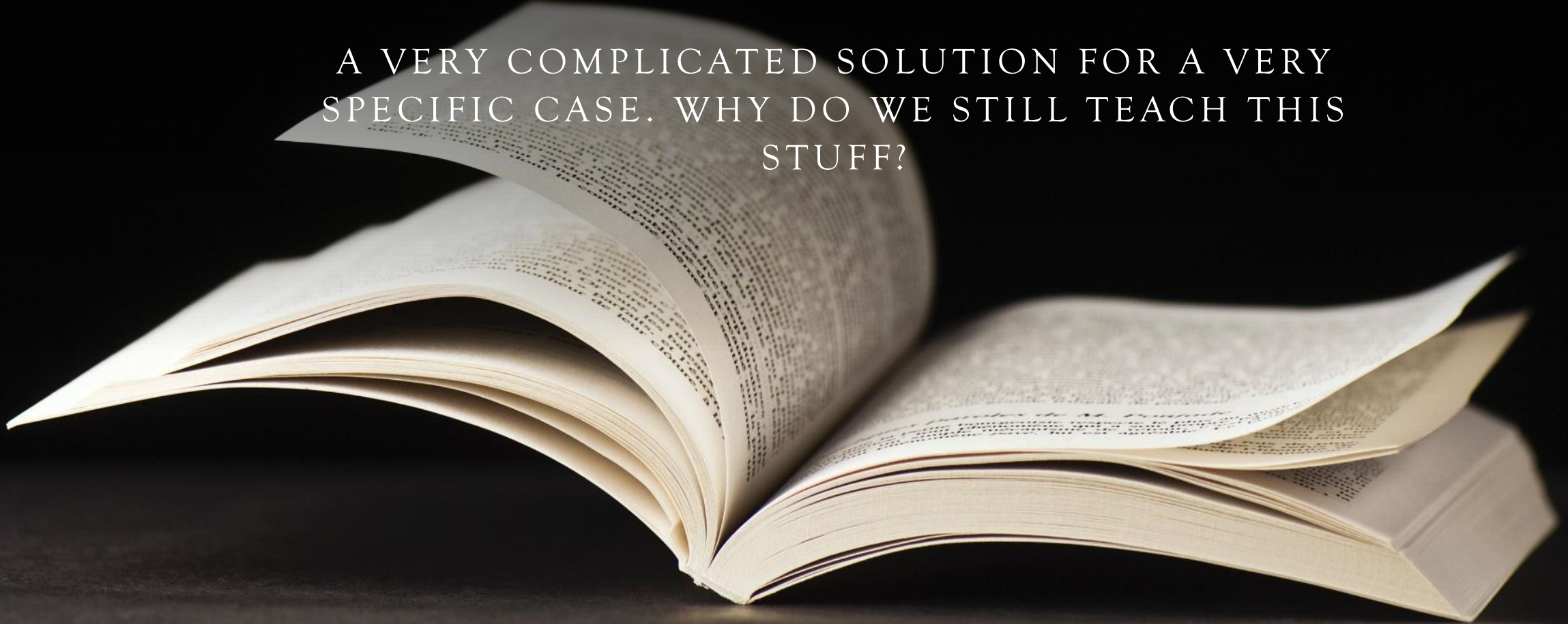
CHAPTER 3: FIXED EFFECTS

LET'S COMPARE EACH ENTITY TO ITS USUAL
(AVERAGE) BEHAVIOR. DOES Y CHANGE FROM
ITS USUAL VALUE WHEN X DEPARTS FROM ITS
USUAL VALUE?



CHAPTER 4: RANDOM EFFECTS

A VERY COMPLICATED SOLUTION FOR A VERY
SPECIFIC CASE. WHY DO WE STILL TEACH THIS
STUFF?



CHAPTER 5: MODELS FOR RESIDUALS CORRELATED AS AN AR(1)

WHEN THINGS THAT HAPPEN IN VEGAS DON'T
JUST STAY IN VEGAS...



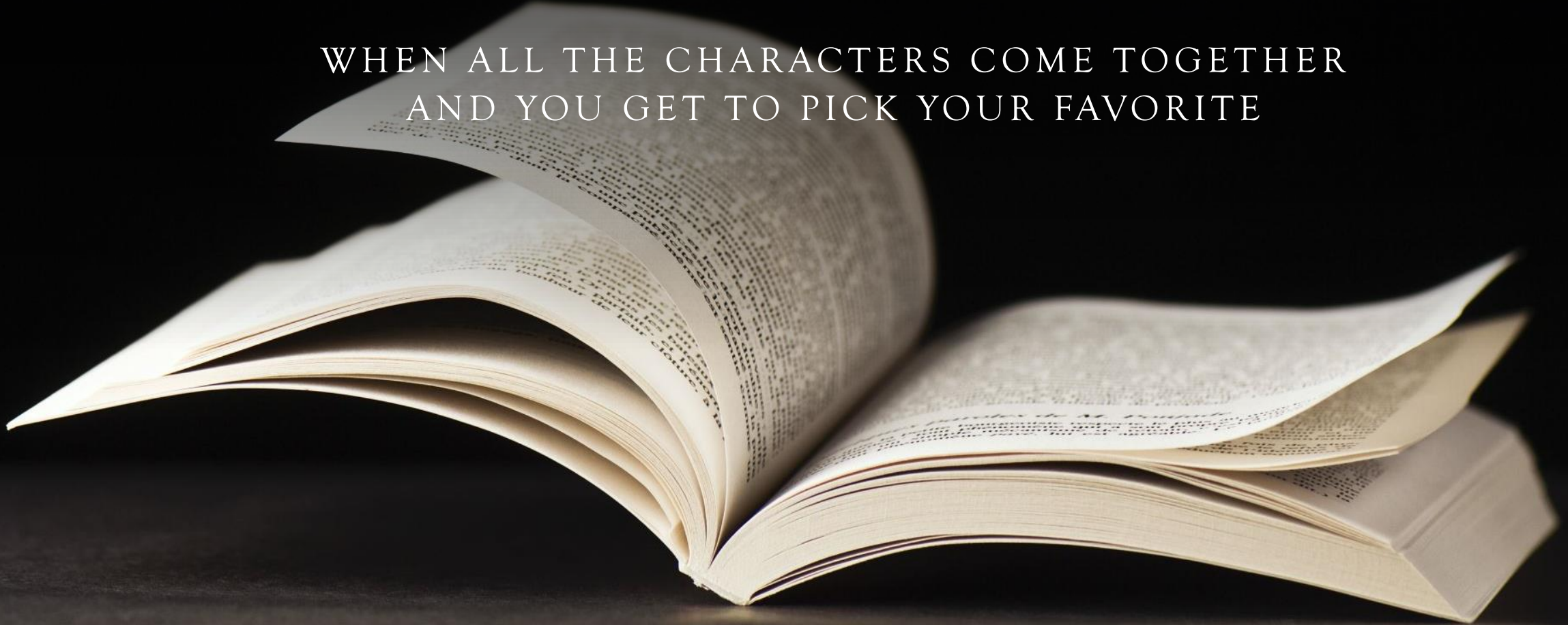
CHAPTER 5: MODELS FOR RESIDUALS CORRELATED AS AN AR(1)

WHEN THINGS THAT HAPPEN IN VEGAS DON'T
JUST STAY IN VEGAS...



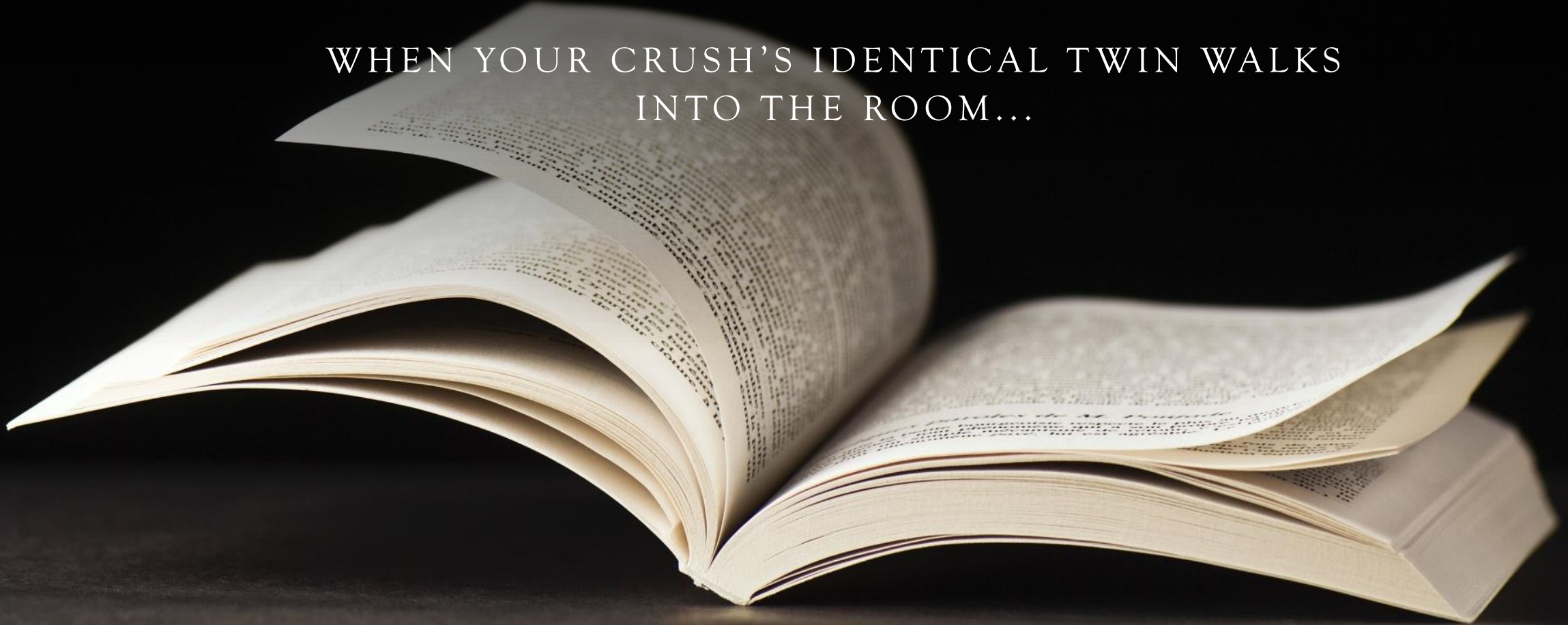
CHAPTER 6: CHOOSING THE BEST MODEL

WHEN ALL THE CHARACTERS COME TOGETHER
AND YOU GET TO PICK YOUR FAVORITE



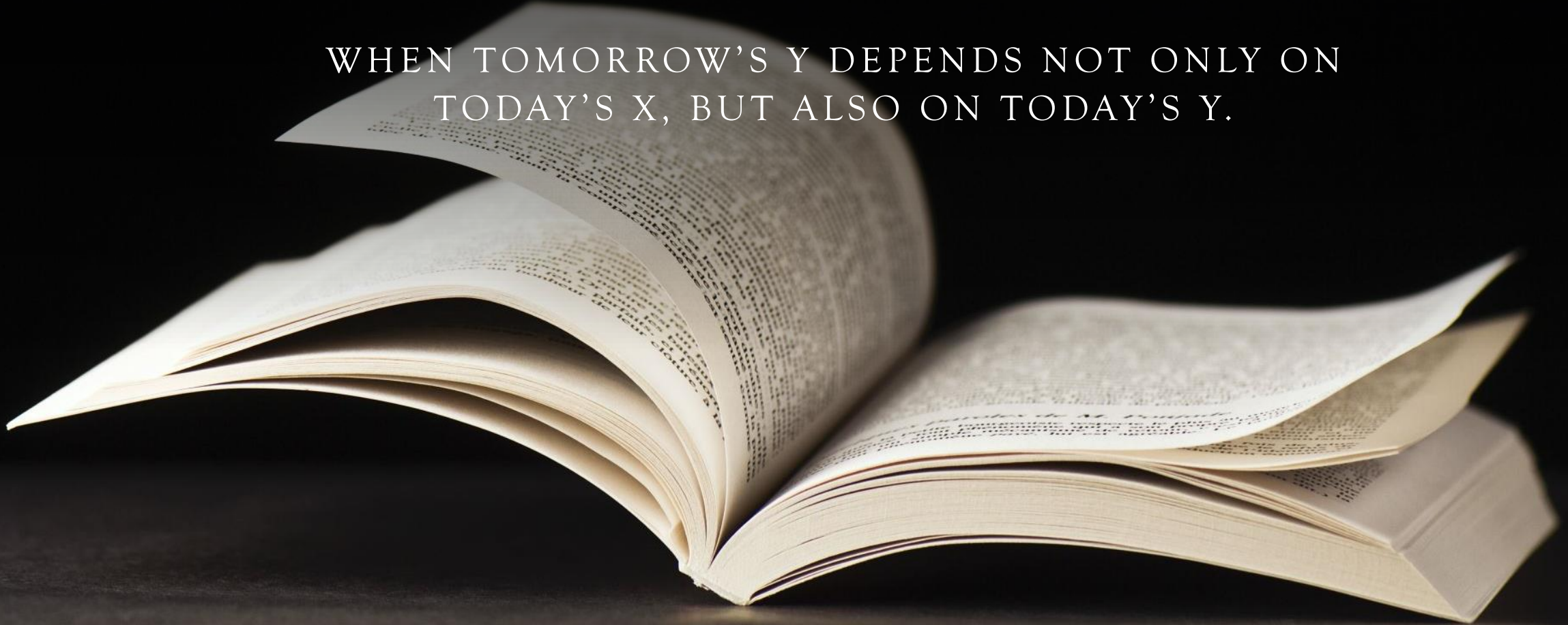
CHAPTER 7: INSTRUMENTAL VARIABLES

WHEN YOUR CRUSH'S IDENTICAL TWIN WALKS
INTO THE ROOM...



CHAPTER 8: DYNAMIC PANELS

WHEN TOMORROW'S Y DEPENDS NOT ONLY ON
TODAY'S X , BUT ALSO ON TODAY'S Y .

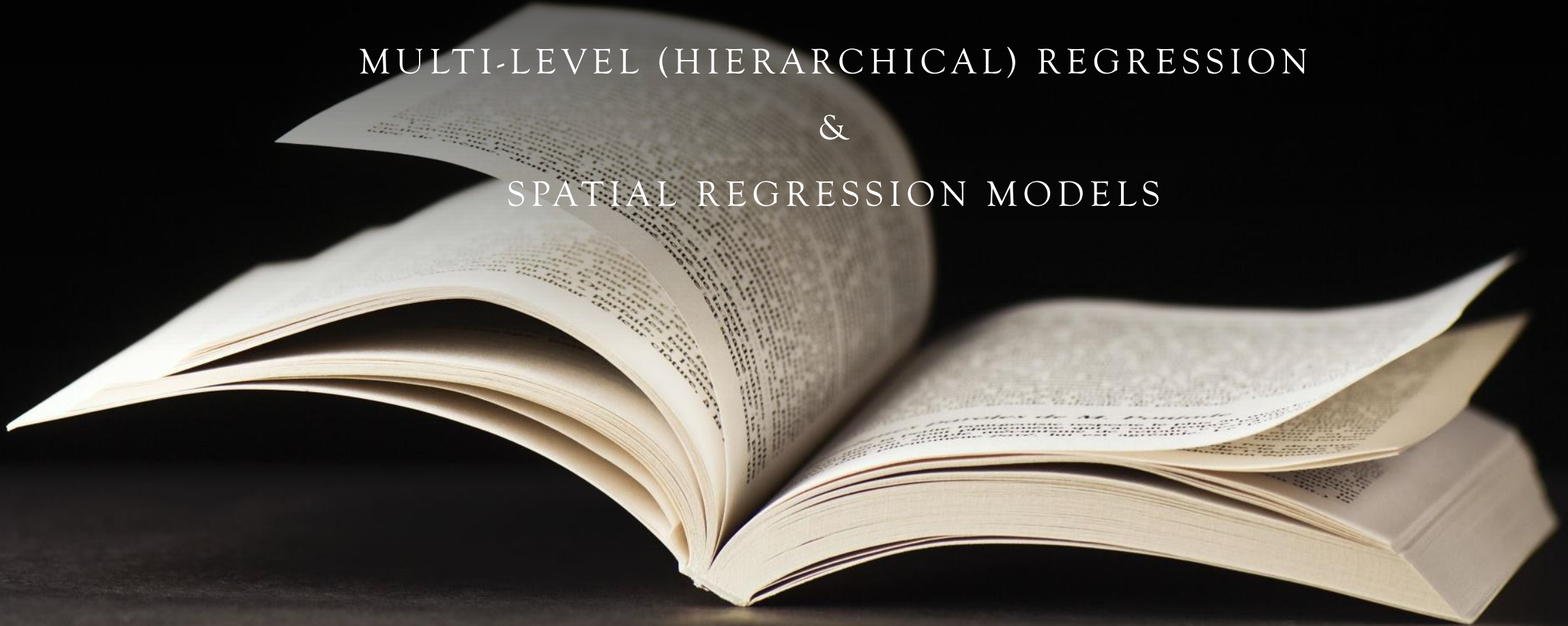


CHAPTER 9: GENERALIZED TIME MODELS

MULTI-LEVEL (HIERARCHICAL) REGRESSION

&

SPATIAL REGRESSION MODELS



Some final comments
before you go...



There's another reason to use Panel Data...

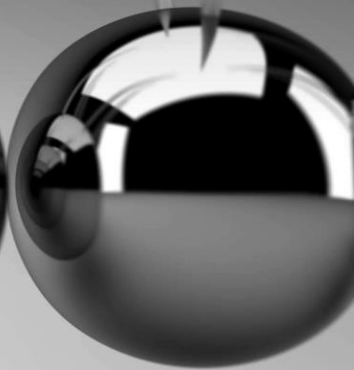
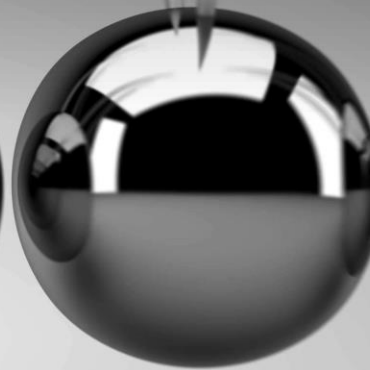
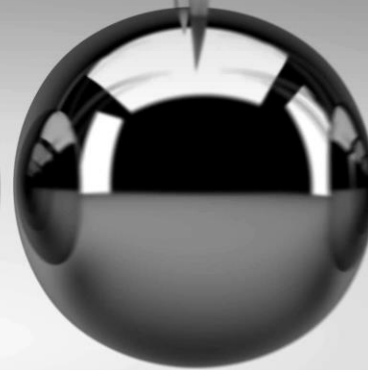
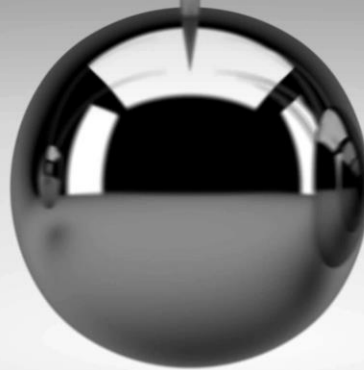
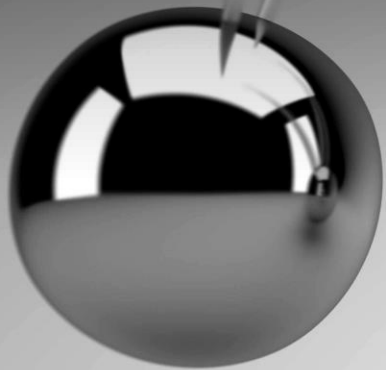


Stock Market Report

330.40	341.94	354.71	367.80	380.50
395.23	411.58	428.00	445.00	462.00
478.00	495.00	512.00	529.00	546.00
563.00	580.00	597.00	614.00	631.00
646.00	663.00	680.00	697.00	714.00
727.00	744.00	761.00	778.00	795.00
812.00	829.00	846.00	863.00	880.00
897.00	914.00	931.00	948.00	965.00
982.00	1000.00	1017.00	1034.00	1051.00
1068.00	1085.00	1102.00	1119.00	1136.00
1153.00	1170.00	1187.00	1204.00	1221.00
1238.00	1255.00	1272.00	1289.00	1306.00
1321.00	1338.00	1355.00	1372.00	1389.00
1406.00	1423.00	1440.00	1457.00	1474.00
1491.00	1508.00	1525.00	1542.00	1559.00
1576.00	1593.00	1610.00	1627.00	1644.00
1661.00	1678.00	1695.00	1712.00	1729.00
1746.00	1763.00	1780.00	1797.00	1814.00
1831.00	1848.00	1865.00	1882.00	1899.00
1916.00	1933.00	1950.00	1967.00	1984.00
2001.00	2018.00	2035.00	2052.00	2069.00
2086.00	2103.00	2120.00	2137.00	2154.00
2171.00	2188.00	2205.00	2222.00	2239.00
2256.00	2273.00	2290.00	2307.00	2324.00
2341.00	2358.00	2375.00	2392.00	2409.00
2426.00	2443.00	2460.00	2477.00	2494.00
2511.00	2528.00	2545.00	2562.00	2579.00
2596.00	2613.00	2630.00	2647.00	2664.00
2681.00	2698.00	2715.00	2732.00	2749.00
2766.00	2783.00	2800.00	2817.00	2834.00
2851.00	2868.00	2885.00	2902.00	2919.00
2936.00	2953.00	2970.00	2987.00	3004.00
3021.00	3038.00	3055.00	3072.00	3089.00
3106.00	3123.00	3140.00	3157.00	3174.00
3191.00	3208.00	3225.00	3242.00	3259.00
3276.00	3293.00	3310.00	3327.00	3344.00
3361.00	3378.00	3395.00	3412.00	3429.00
3446.00	3463.00	3480.00	3497.00	3514.00
3531.00	3548.00	3565.00	3582.00	3599.00
3616.00	3633.00	3650.00	3667.00	3684.00
3701.00	3718.00	3735.00	3752.00	3769.00
3786.00	3803.00	3820.00	3837.00	3854.00
3871.00	3888.00	3905.00	3922.00	3939.00
3956.00	3973.00	3990.00	4007.00	4024.00
4041.00	4058.00	4075.00	4092.00	4109.00
4126.00	4143.00	4160.00	4177.00	4194.00
4211.00	4228.00	4245.00	4262.00	4279.00
4296.00	4313.00	4330.00	4347.00	4364.00
4381.00	4398.00	4415.00	4432.00	4449.00
4466.00	4483.00	4500.00	4517.00	4534.00
4551.00	4568.00	4585.00	4602.00	4619.00
4636.00	4653.00	4670.00	4687.00	4704.00
4721.00	4738.00	4755.00	4772.00	4789.00
4806.00	4823.00	4840.00	4857.00	4874.00
4891.00	4908.00	4925.00	4942.00	4959.00
4976.00	4993.00	5010.00	5027.00	5044.00
5061.00	5078.00	5095.00	5112.00	5129.00
5146.00	5163.00	5180.00	5197.00	5214.00
5231.00	5248.00	5265.00	5282.00	5299.00
5316.00	5333.00	5350.00	5367.00	5384.00
5401.00	5418.00	5435.00	5452.00	5469.00
5486.00	5503.00	5520.00	5537.00	5554.00
5571.00	5588.00	5605.00	5622.00	5639.00
5656.00	5673.00	5690.00	5707.00	5724.00
5741.00	5758.00	5775.00	5792.00	5809.00
5826.00	5843.00	5860.00	5877.00	5894.00
5911.00	5928.00	5945.00	5962.00	5979.00
5996.00	6013.00	6030.00	6047.00	6064.00
6081.00	6098.00	6115.00	6132.00	6149.00
6166.00	6183.00	6200.00	6217.00	6234.00
6251.00	6268.00	6285.00	6302.00	6319.00
6336.00	6353.00	6370.00	6387.00	6404.00
6421.00	6438.00	6455.00	6472.00	6489.00
6506.00	6523.00	6540.00	6557.00	6574.00
6591.00	6608.00	6625.00	6642.00	6659.00
6676.00	6693.00	6710.00	6727.00	6744.00
6761.00	6778.00	6795.00	6812.00	6829.00
6846.00	6863.00	6880.00	6897.00	6914.00
6931.00	6948.00	6965.00	6982.00	6999.00
7016.00	7033.00	7050.00	7067.00	7084.00
7101.00	7118.00	7135.00	7152.00	7169.00
7186.00	7203.00	7220.00	7237.00	7254.00
7271.00	7288.00	7305.00	7322.00	7339.00
7356.00	7373.00	7390.00	7407.00	7424.00
7441.00	7458.00	7475.00	7492.00	7509.00
7526.00	7543.00	7560.00	7577.00	7594.00
7611.00	7628.00	7645.00	7662.00	7679.00
7696.00	7713.00	7730.00	7747.00	7764.00
7781.00	7798.00	7815.00	7832.00	7849.00
7866.00	7883.00	7900.00	7917.00	7934.00
7951.00	7968.00	7985.00	8002.00	8019.00
8036.00	8053.00	8070.00	8087.00	8104.00
8121.00	8138.00	8155.00	8172.00	8189.00
8206.00	8223.00	8240.00	8257.00	8274.00
8291.00	8308.00	8325.00	8342.00	8359.00
8376.00	8393.00	8410.00	8427.00	8444.00
8461.00	8478.00	8495.00	8512.00	8529.00
8546.00	8563.00	8580.00	8597.00	8614.00
8631.00	8648.00	8665.00	8682.00	8699.00
8716.00	8733.00	8750.00	8767.00	8784.00
8801.00	8818.00	8835.00	8852.00	8869.00
8886.00	8903.00	8920.00	8937.00	8954.00
8971.00	8988.00	9005.00	9022.00	9039.00
9056.00	9073.00	9090.00	9107.00	9124.00
9141.00	9158.00	9175.00	9192.00	9209.00
9226.00	9243.00	9260.00	9277.00	9294.00
9311.00	9328.00	9345.00	9362.00	9379.00
9396.00	9413.00	9430.00	9447.00	9464.00
9481.00	9498.00	9515.00	9532.00	9549.00
9566.00	9583.00	9600.00	9617.00	9634.00
9651.00	9668.00	9685.00	9702.00	9719.00
9736.00	9753.00	9770.00	9787.00	9804.00
9821.00	9838.00	9855.00	9872.00	9889.00
9906.00	9923.00	9940.00	9957.00	9974.00
9991.00	10008.00	10025.00	10042.00	10059.00
10076.00	10093.00	10110.00	10127.00	10144.00
10161.00	10178.00	10195.00	10212.00	10229.00
10246.00	10263.00	10280.00	10297.00	10314.00
10331.00	10348.00	10365.00	10382.00	10399.00
10416.00	10433.00	10450.00	10467.00	10484.00
10501.00	10518.00	10535.00	10552.00	10569.00
10586.00	10603.00	10620.00	10637.00	10654.00
10671.00	10688.00	10705.00	10722.00	10739.00
10756.00	10773.00	10790.00	10807.00	10824.00
10841.00	10858.00	10875.00	10892.00	10909.00
10926.00	10943.00	10960.00	10977.00	10994.00
11011.00	11028.00	11045.00	11062.00	11079.00
11096.00	11113.00	11130.00	11147.00	11164.00
11181.00	11198.00	11215.00	11232.00	11249.00
11266.00	11283.00	11300.00	11317.00	11334.00
11351.00	11368.00	11385.00	11402.00	11419.00
11436.00	11453.00	11470.00	11487.00	11504.00
11521.00	11538.00	11555.00	11572.00	11589.00
11606.00	11623.00	11640.00	11657.00	11674.00
11691.00	11708.00	11725.00	11742.00	11759.00
11776.00	11793.00	11810.00	11827.00	11844.00
11861.00	11878.00	11895.00	11912.00	11929.00
11946.00	11963.00	11980.00	12000.00	12017.00
12034.00	12051.00	12068.00	12085.00	12102.00
12119.00	12136.00	12153.00	12170.00	12187.00
12204.00	12221.00	12238.00	12255.00	12272.00
12289.00	12306.00	12323.00	12340.00	12357.00
12374.00	12391.00	12408.00	12425.00	12442.00
12459.00	12476.00	12493.00	12510.00	12527.00
12544.00	12561.00	12578.00	12595.00	12612.00
12629.00	12646.00	12663.00	12680.00	12697.00
12714.00	12731.00	12748.00	12765.00	12782.00
12799.00	12816.00	12833.00	12850.00	12867.00
12884.00	12901.00	12918.00	12935.00	12952.00
12969.00	12986.00	13003.00	13020.00	13037.00
13054.00	13071.00	13088.00	13105.00	13122.00
13139.00	13156.00	13173.00	13190.00	13207.00
13224.00	13241.00	13258.00	13275.00	13292.00
13309.00	13326.00	13343.00	13360.00	13377.00
13394.00	13411.00	13428.00	13445.00	13462.00
13479.00	13496.00	13513.00	13530.00	13547.00
13564.00	13581.00	13598.00	13615.00	13632.00
13649.00	13666.00	13683.00	13700.00	13717.00
13734.00	13751.00	13768.00	13785.00	13802.00
13819.00	13836.00	13853.00	13870.00	13887.00
13904.00	13921.00	13938.00	13955.00	13972.00
13989.00	14006.00	14023.00	14040.00	14057.00
14074.00	14091.00	14108.00	14125.00	14142.00
14159.00	14176.00	14193.00	14210.00	14227.00
14244.00	14261.00	14278.00	14295.00	14312.00
14329.00	14346.00	14363.00	14380.00	14397.00
14414.00	14431.00	14448.00	14465.00	14482.00
14499.00	14516.00	14533.00	14550.00	14567.00
14584.00	14601.00	14618.00	14635.00	14652.00
14669.00	14686.00	14703.00	14720.00	14737.00
14754.00	14771.00	14788.00	14805.00	14822.00
14839.00	14856.00	14873.00	14890.00	14907.00
14924.00	14941.00	14958.00	14975.00	14992.00
15009.00	15026.00	15043.00	15060.00	15077.00
15094.00	15111.00	15128.00	15145.00	15162.00
15179.00	15196.00	15213.00	15230.00	15247.00
15264.00	15281.00	15298.00	15315.00	15332.00
15349.00	15366.00	15383.00	15400.00	15417.00
15434.00	15451.00	15468.00	15485.00	15502.00
15519.00	15536.00	15553.00	15570.00	15587.00
15604.00	15621.00	15638.00	15655.00	15672.00
15689.00	15706.00	15723.00	15740.00	15757.00
15774.00	15791.00	15808.00	15825.00	15842.00
15859.00	15876.00	15893.00	15910.00	15927.00
15944.00	15961.00	15978.00	15995.00	16012.

Causality

(Is this a strong argument?)



Scattered comments about panel data

N entities over T time periods

- Typically $N \gg T$ (why?)
- Balanced vs. Imbalanced panel data
- Entities have characteristics that do not change through time (entity-fixed effects)
- Time periods may have characteristics that affect all entities equally (time-fixed effects)

Could you give examples of entity–fixed and time-fixed effects?

How would this change the “general” Panel Data model?

- Entity-fixed effects are much more common, and time-fixed effects are often neglected. Indeed, we often speak of “fixed effects” (with no qualification) to refer to “entity-fixed effects”.