

how to read quantitative papers when you don't know stats

some tips

Charlotte Lloyd, Harvard Sociology Department Writing Fellow, 2017-2018

start with the “big picture”

- even if you don't have a stats background, you can always understand the “big picture”
- identify the inputs and outcomes, don't worry about the math



- **WARNING:** be aware that papers may present MULTIPLE models, and that these models may have different inputs (and outcomes!)

input: what is it?

- models basically take input in the form of a data table
 - n number of observations (rows)
 - x number of variables (columns)

observation #	fruit	month	day	year	city	state
1	apple	June	8	1996	Denver	CO
2	banana	April	27	1988	Boulder	CO
...
100	kiwi	May	4	1980	Aspen	CO

input: be able to answer these questions

- observations
 - what is the unit of n ? (people, cities, countries, etc.)
 - how many are there? (i.e. what's the "n"?)
- variables
 - what is the independent variable? what are the dependent variables?
 - what is the source of data for each variable? (each variable may have a different source)
 - does the author mention serious missing data problems for any variables?

outcomes: some major types

- **regression models**

- outcome will be numerical “coefficients” that describe the relationship between variables
- positive coefficients mean there is a positive relationship and vice-versa

- **logit / probit models**

- primary outcome will be a “probability” or “risk” of an event happening (scale 0-1)
- will also generate numerical coefficients to show how variables contribute to outcomes

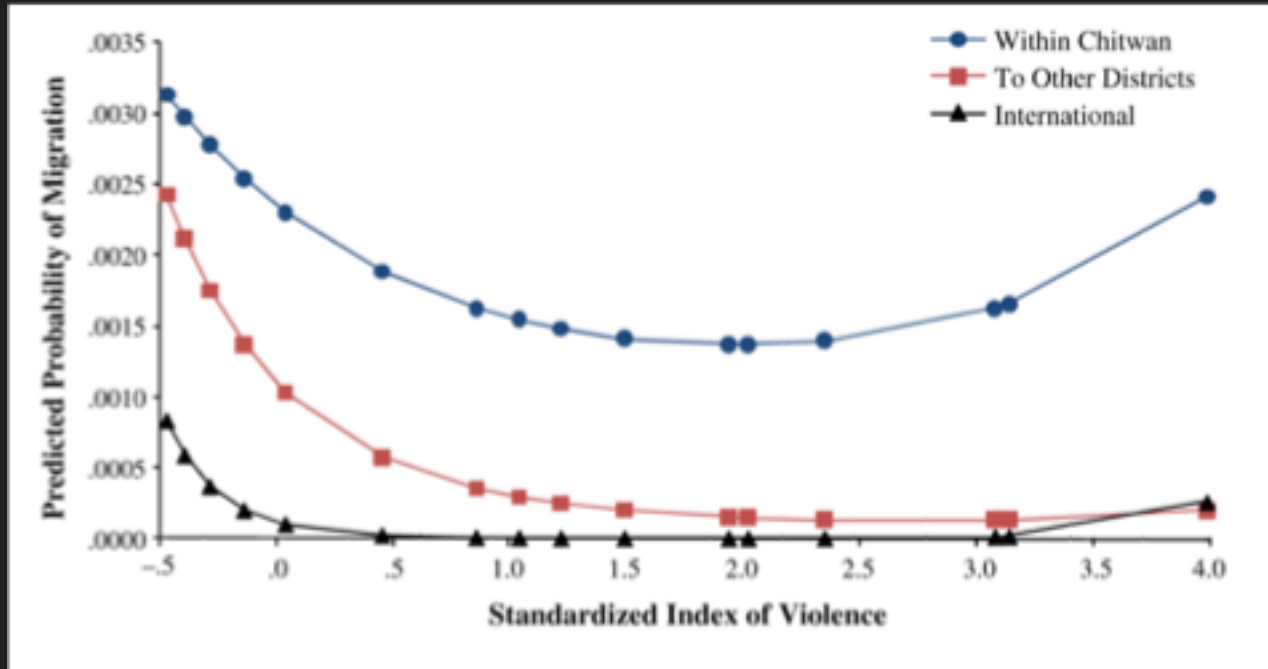
- **clustering, PCA, topic models**

- outcome will be assigning each observation to a group
- the point is to understand differences between the groups of observations

trust that they did the math correctly

- if it's a peer-reviewed paper, the math is probably fine... or close enough 😊
- if you don't know stats, it's fine to ignore mathy things like p-values, R-squared, model fit, heteroskedasticity, and variance. just focus on the inputs and outcomes!

tip: look for an actual “big picture”



- generally papers provide figures that you can use to jumpstart your understanding of the “big picture”
- figures most often show the relationship between the independent variable (one of the inputs) and the outcome

tip: look for tables summarizing variables

Bohra-Mishra and Massey Page 19

Table 1

Definition of variables

Variable	Definition
Outcome Variables	
Migrate to three competing locations	Whether respondent migrated in month $t+1$ to regions within Chitwan (1) migrate to other districts (2) migrate to other countries (3), or didn't migrate at all in month $t+1$ (4)
Migrate to two competing locations	Whether respondent migrated in month $t+1$ to regions within Chitwan (1) migrate to other districts in other countries (2) or didn't migrate at all in month $t+1$
Proximate Causes	
Standardized index of violence	Derived through factor analysis using six indicators of violence: number of bomb blasts, number of casualties from bomb blasts, and number of casualties from major incidents of violence. All variables are measured for Chitwan and the surrounding districts separately.
Root Causes	
Physical capital	
Oven furnished	1 if respondent's household owns furnished, 0 otherwise.
Standardized index of household amenities	Derived through factor analysis using data on the materials used to build the floor and roof of the respondent's house, and electric household.
Standardized index of goods owned	Derived through factor analysis using data on families owned by the household: a radio, TV, bicycle, motorcycle, car, tractor, pumpset, and irrigation.
Standardized index of livestock owned	Derived through factor analysis using data on number of chickens, goats, buffaloes, heifers, cows, sheep, goats, and pigs owned by the household.
Human capital	
Education	Number of years of schooling completed by the respondent.
Salary job	1 if respondent holds a salary job, 0 otherwise.
Age	Respondent's age, monthly event.
Age squared	Respondent's age squared, monthly event.
Increasing returns	
Social capital	
House member migrated within Chitwan	1 if any member from the respondent's household migrated within Chitwan in 1996, 0 otherwise.
House member migrated to other districts	1 if any member from the respondent's household migrated to other districts in 1996, 0 otherwise.
House member as international migrant	1 if any member from the respondent's household migrated to other countries in 1996, 0 otherwise.
Demographic Variables	
Female	1 if respondent is a female, 0 otherwise.
Married	1 if respondent was ever married, 0 otherwise, monthly event.
Number of household members	Number of people in the household as of 1996.
Ethnicity	
High upper caste	1 if Brahmin upper caste, 0 otherwise.
High lower caste	1 if Shukla lower caste, 0 otherwise.
BTI (Tharu/Burman)	1 if BTI (Tharu/Burman) caste, 0 otherwise.

Demography: Author manuscript, available in PMC 2012 June 01.

Bohra-Mishra and Massey Page 21

Table 2

Index weights for the composite index of violence and physical capital variables

	Index Weights
Violence Variables	
Chitwan and neighboring districts violence	
Number of bomb blast in Chitwan	0.331
Number of casualties from bomb blasts in Chitwan	0.265
Number of casualties from major attacks in Chitwan ^a	0.013
Number of bomb blasts in neighboring districts	0.318
Number of casualties from bomb blasts in neighboring districts	0.296
Number of casualties from major attacks in neighboring districts ^a	0.189
Violence explained by first factor	0.421
Physical Capital Variables	
Household amenities	
Roof of house is made of slate, tin, or concrete	0.248
Floor of house is made of concrete	0.262
Does not have own drinking water source	-0.214
No toilet	-0.217
Has electricity	0.242
Violence explained by first factor	0.687
Goods owned	
Household has a radio	0.190
Household has a TV	0.198
Household has a bicycle	0.188
Household has a motorcycle	0.200
Household has a car	0.165
Household has a tractor	0.188
Household has a pumpset for irrigation	0.110
Household has a biogas plant	0.210
Violence explained by first factor	0.489
Livestock owned	
Number of chickens and ducks	0.081
Number of goats	0.203
Number of buffaloes	0.300
Number of cows	0.271
Number of male buffaloes	0.231
Number of female buffaloes	0.328
Number of sheep and goats	0.422
Number of pigs	0.141
Violence explained by first factor	0.211

^aMajor attacks refer to clashes between the Maoists and the state, major attacks by other party, and major landmine explosions resulting in more than two casualties.

Demography: Author manuscript, available in PMC 2012 June 01.

Bohra-Mishra and Massey Page 22

Table 3

Descriptive statistics for the measures used to create composite index of violence and physical capital

	Count	Min.	Max.	SD	Mean
Proximate Causes					
Chitwan and neighboring districts violence					
Number of bomb blast in Chitwan	108	0	3	0.592	0.204
Number of casualties from bomb blasts in Chitwan	108	0	5	0.798	0.213
Number of casualties from major attacks in Chitwan ^a	108	0	38	3.740	0.528
Number of bomb blasts in neighboring districts	108	0	9	1.792	0.796
Number of casualties from bomb blasts in neighboring districts	108	0	12	2.259	0.750
Number of casualties from major attacks in neighboring districts ^a	108	0	86	10.325	3.000
Root Causes					
Physical capital					
Household amenities					
Roof of house is made of slate, tin, or concrete	1,391	0	1	0.500	0.491
Floor of house is made of concrete	1,391	0	1	0.430	0.245
Does not have own drinking water source	1,391	0	1	0.498	0.455
No toilet	1,391	0	1	0.480	0.361
Has electricity	1,391	0	1	0.475	0.342
Goods owned					
Household has a radio	1,391	0	1	0.500	0.523
Household has a TV	1,391	0	1	0.329	0.124
Household has a bicycle	1,391	0	1	0.485	0.624
Household has a motorcycle	1,391	0	1	0.175	0.032
Household has a car	1,391	0	1	0.260	0.073
Household has a tractor	1,391	0	1	0.089	0.008
Household has a pumpset for irrigation	1,391	0	1	0.179	0.033
Household has a biogas plant	1,391	0	1	0.208	0.045
Livestock owned					
Number of chickens and ducks	1,391	0	2,210	121.637	20.160
Number of pigeons	1,391	0	150	7.013	1.154

^aMajor attacks refer to clashes between the Maoists and the state, major attacks by other party, and major landmine explosions resulting in more than two casualties.

Demography: Author manuscript, available in PMC 2012 June 01.

tip: you can look at the outcomes table...

- ...but always read the findings section where authors *explain and interpret* the outcomes.