Problem Set 1

MACSS 30200

Tyler Amos

17-Apr-2018

Data Description

This dataset contains records of live births in the United States for the year 2016. A live birth is defined as "every product of conception that gives a sign of life after birth" (CDC 2016). Along with various demographic data (e.g., mother's race, education) the datset includes details of the infant's health and any medical conditions.

Access: This dataset can be downloaded from the Centers for Disease Control and Prevention (CDC) website (www.cdc.gov) as a compressed (e.g., .zip, .dat.Z) file. A derivation of this dataset is also available from the National Bureau of Economic Research website (www.nber.org) in a variety of formats (e.g., Stata, SAS).

Curation: The data is curated in its original source by the CDC through the National Center for Health Statistics' National Vital Statistics System. The National Bureau of Economic Research curates the derived datasets.

Previous Studies Using This Data

This data has been used to study both medical and demographic/sociological phenomena. Medical researchers have used this dataset for studies in perinatology (Shih et al. 2016) and pharmacy (Sahabi et al. 2018). In a demographic study, Walton (2009) examined the impact of racial segregation on children's birth weights. Lopoo et al. (2006) and DeLeire et al. (2011) assessed the impact of social policies like welfare and Medicaid on fertility.

Collection

The data were collected by National Institute for Health Statistics from a 100% sample taken from birth records in US states (NBER 2018). The National Bureau of Economic Research then consolidated the datasets into a variety of formats. This paper uses the downloadable CSV of the data available via NBER (see Data Access Links below).

Descriptive Statistics

Note: Outliers and missing data have been removed. Visualizations are based on a simple random sample of 10,000 cases.

Numeric Variables

Code	Description	Max	Min	Std	Mean
dob_tt	Time of Birth in the format *hhmm*	2400	0	640	1231
mager	Mother's age in years	50	12	5.8	28.7
fagecomb	Father's combined age in years	98	11	22.7	39.5
priorlive	Number of children by the same mother who are still living	21	0	4.6	1.3

Categorical Variables

Code	Description	Mode	Appended Visualization
dob_mm	Month of Birth	August	Distribution of Births by Month
dob_wk	Day of the Week of Birth	Tuesday	Distribution of Births by Day of the Week
meduc	Mother's Educational Attainment	High School Graduate/GED	Mother's Educational Attainment
feduc	Father's Educational Attainment	High School Graduate/GED	Father's Educational Attainment

One especially notable trend in the data is the differential influence of sex on the age of the parent at the time of birth. If we disaggregate these data points on the basis of education we see that although higher levels of educaiton appear to be associated with delayed fatherhood, this is not strictly true in the upper age ranges. By contrast, mother's data displays the same trend in the lower age ranges but has an additional upper limit. This is, of course, attributable to well-known biological factors. However the interaction with education, given that higher education appears to be associated with a lower number of children (see Figure 3), suggests highly-educated women have children later and in fewer numbers than other women.
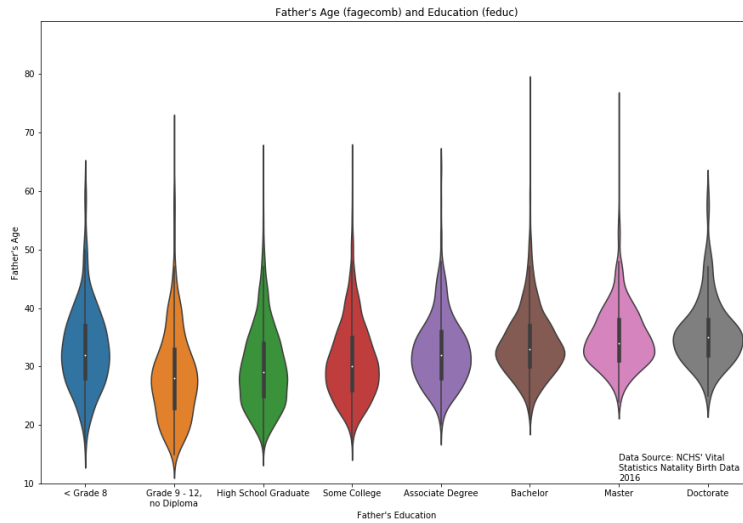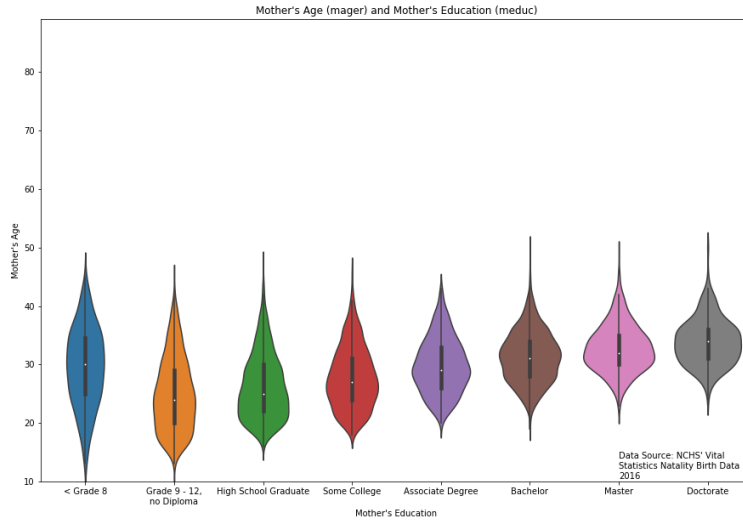

Figure 1: Father's Age and Education

Figure 2: Mother's Age and Education

Sliced Data by Mother's Educational Attainment

When slicing the data by Mother's Educational Attainment (see Appendix for visualizations) we observe a number of interesting trends. First, we see that education appears to be negatively associated with the number of prior living children. Second, the mother's education appears to have no association with the timing of the birth. Third, the mother's educaton has some association with the father's age, consistent with what we have seen when comparing father's education and mother's education to their respective ages. That is to say, more educated individuals tend to have children later in life.

References

CDC. "User Guide to the 2016 Natality Public Use File" Accessed via National Bureau of Economic Research: http://nber.org/natality/2016/natl2016.pdf Accessed: 16 April 2018

DeLeire, Thomas, Leonard M. Lopoo, and Kosali I. Simon. "Medicaid expansions and fertility in the United States." *Demography* 48, no. 2 (2011): 725-747.

National Bureau of Economic Research (NBER). "NCHS' Vital Statistics Natality Birth Data" http://nber.org/data/vital-statistics-natality-data.html Accessed: 16 April 2018

Lopoo, Leonard M., and Thomas DeLeire. "Did welfare reform influence the fertility of young teens?." *Journal of Policy Analysis and Management* 25, no. 2 (2006): 275-298.

Shahabi, Ahva, Desi Peneva, Devin Incerti, Kimmie McLaurin, and Warren Stevens. "Assessing variation in the cost of palivizumab for respiratory syncytial virus prevention in preterm infants." *PharmacoEconomics-open* 2, no. 1 (2018): 53-61.

Shih, Tiffany, Desi Peneva, Xiao Xu, Amelia Sutton, Elizabeth Triche, Richard A. Ehrenkranz, Michael Paidas, and Warren Stevens. "The rising burden of preeclampsia in the United States impacts both maternal and child health." *American journal of perinatology* 33, no. 04 (2016): 329-338.

Walton, Emily. "Residential segregation and birth weight among racial and ethnic minorities in the United States." *Journal of health and social behavior* 50, no. 4 (2009): 427-442.

Data Access Links

Natality Data 2016 Source: National Bureau of Economic Research Link: http://nber.org/natality/2016/natl2016.csv.zip

Appendix: Visualizations

All Replication code for these visualizations is available at: https://github.com/tamos/MACS30200proj/tree/mast
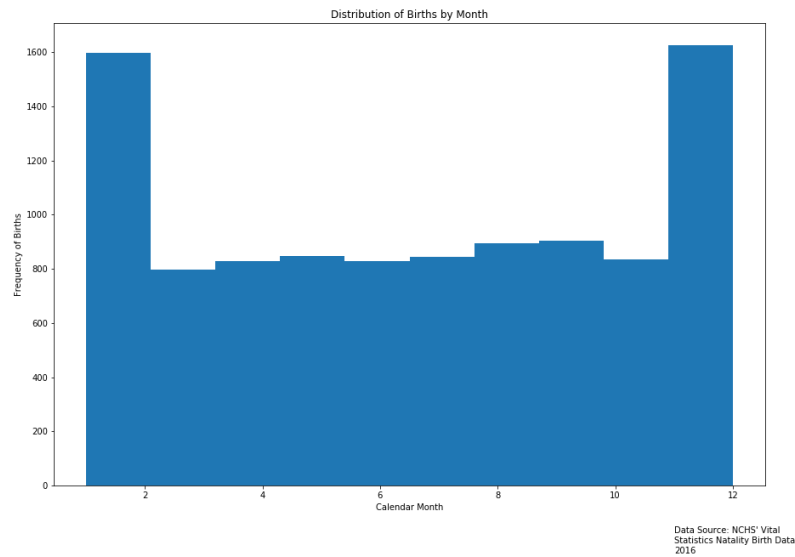

Figure 3: Note the majority of births occur at the end and beginning of the year.
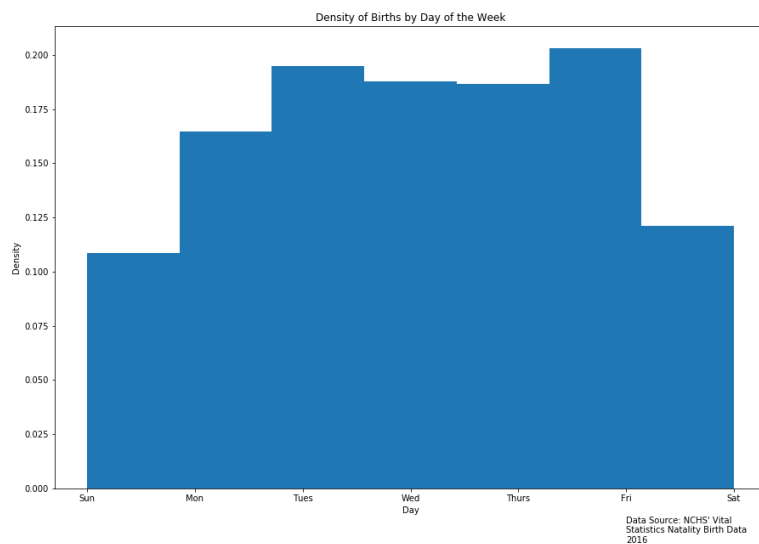
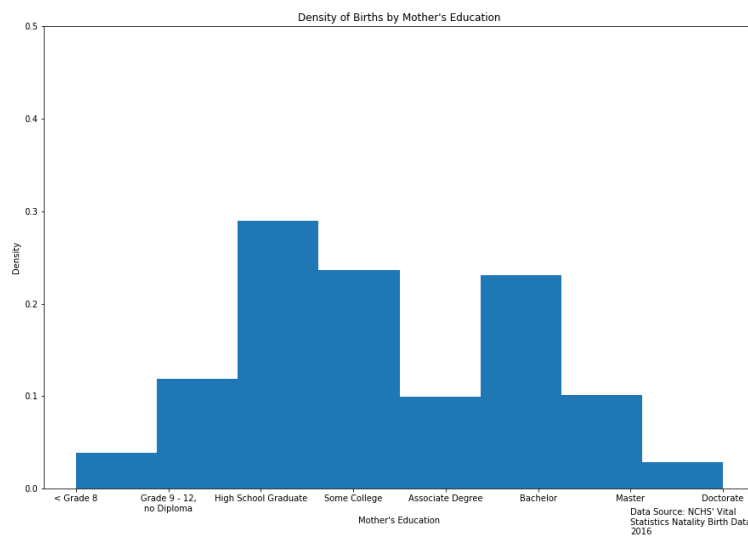Figure 4: Note the majority of births occur at the end and beginning of the year.

Figure 5: Note the near bimodality, with most births to High School Graduates and holders of Bachelor's Degrees
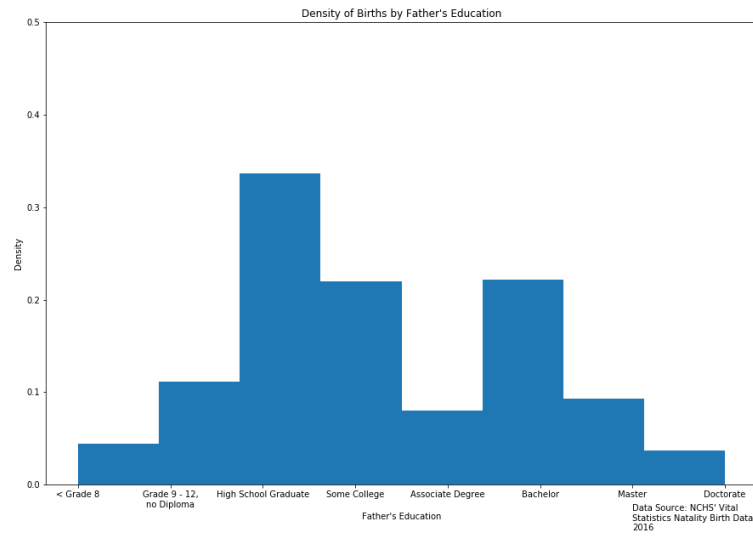
Figure 6: Once again, note the high number of births to High School Graduates, but unlike the Mothers' distribution, there is no near bimodality.
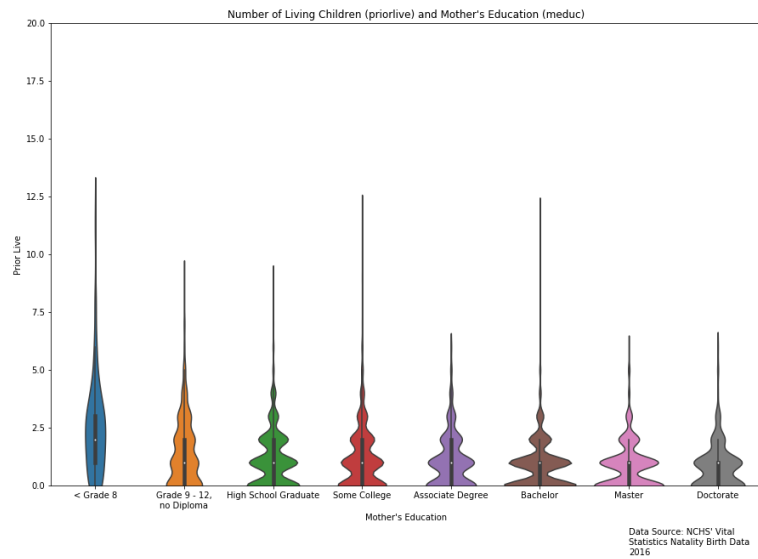

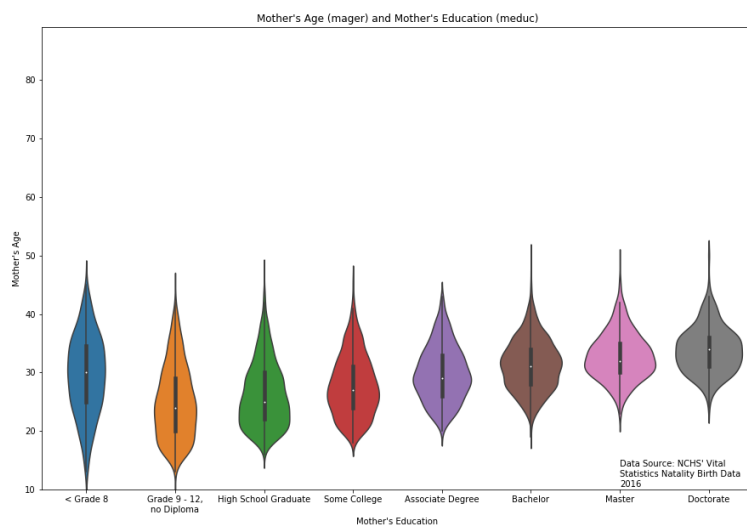Figure 7: Number of Living Children and Mother's Education
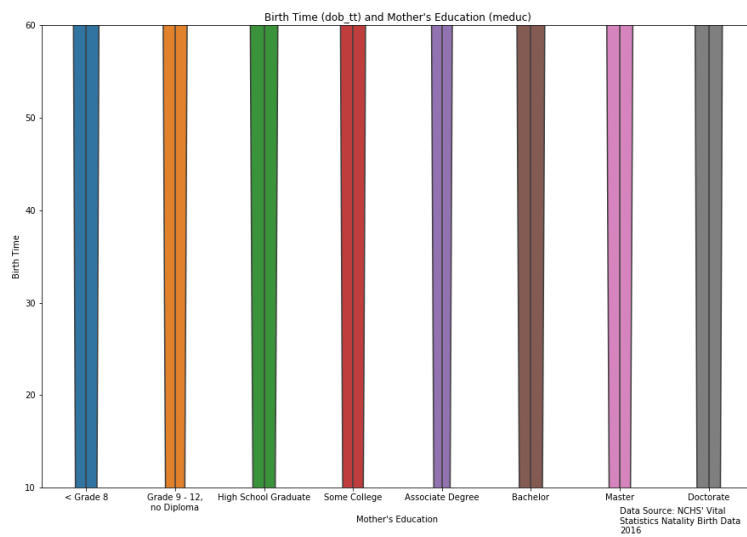
Figure 8: Mother's Age and Education
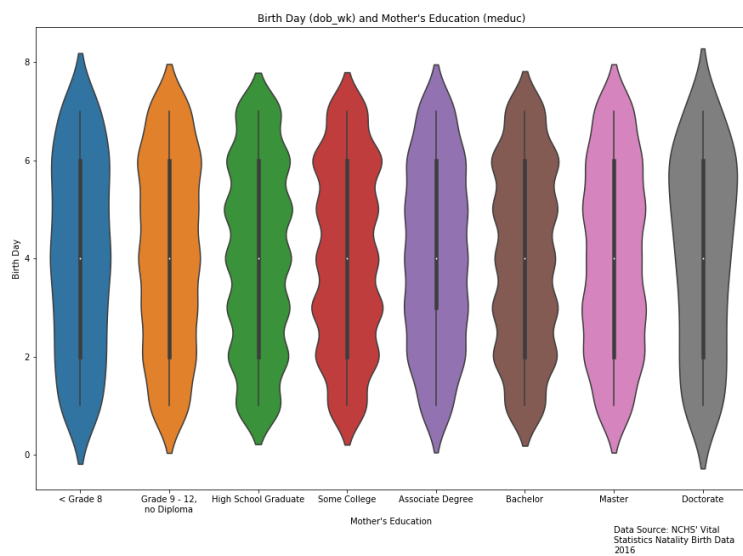

Figure 9: Birth Time and Mother's Education
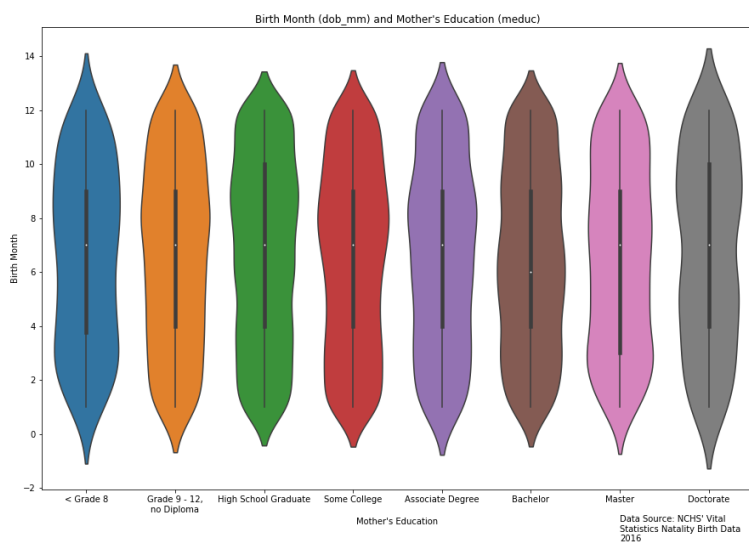
Figure 10: Birth Day and Mother's Education

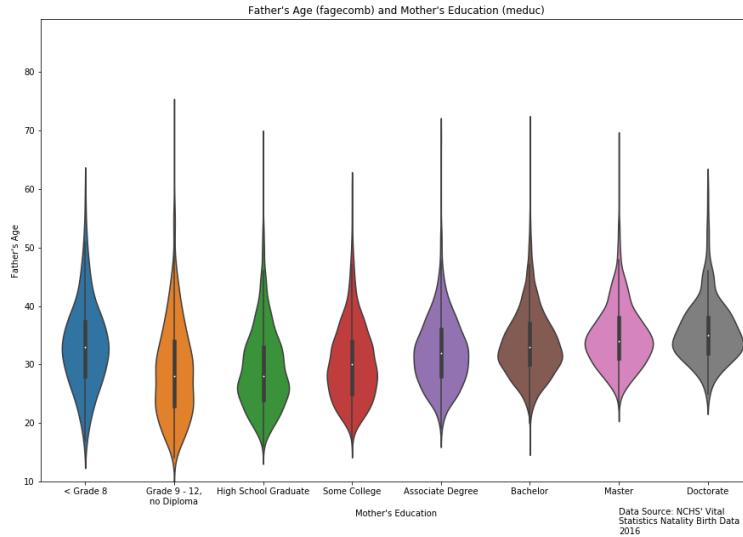Figure 11: Birth Month and Mother's Education

10

Figure 12: Father's Age and Mother's Education

Research Paper Critique

Paper: Roach Anleu, S., & Mack, K. (2015). Performing authority: Communicating judicial decisions in lower criminal courts. Journal of Sociology, 51(4), 1052-1069.

1. State the research question of your assigned paper.

This paper investigates the question: Are sentencing decisions communicated in distinct ways? The researchers' investigate two core hypotheses: (i) sentencing decisions will be communicated in a way that engages in a way that prioritizes greater performance of legitimacy/authority through engagement with the sentenced party; (ii) non-sentencing decisions will be communicated in more impersonal, generic ways. (Roach Anleu and Mack 2015, 1053) The researchers assert differences between these two phenomena can be explained by a differential need for legitimacy, legitimacy being more important in sentencing situations.

2. What data did the paper use?

The paper uses data hand-coded by (mostly) the authors via in-person observation of court proceedings in Australian lower criminal courts. The sample was drawn

from across Australia (all states and capitals) as well as suburbs and large cities. Magistrates of male and female gender, with a variety of experience were observed. Criminal proceedings were chosen as they are present across all jurisdictions. Decisions were chosen to be observed rather than trials because trials are actually scarce occurrences relative to decisions. (Roach Anleu and Mack 2015, 1057) The final dataset consists of 1,287 observations, with each legal "matter" a unit of observation. For each matter, the researchers coded the magistrates behaviour independently and reconciled differences afterwards to produce a single coding. (Ibid) The behaviour observed was: (i) if the magistrate looked directly at the defendant; (ii) the magistrate's ordering of the decision and rationale; (iii) the effect of legal representatives on these behaviours. (Ibid 1053) The researchers used both their observation in court and legal records to extract relevant information about the case such as the specific allegation. Additional research in court records was used to fill missing data.

3. What theory did the paper reference in order to interpret the data? (Note: it is possible that the paper has no reference to theory.)

The primary theory upon which the paper is based is Weber's theory of legitimacy and authority. From the Weberian perspective, an individual's authority to arbitrate (e.g., a judge) is based in part on their legitimacy in the eyes of those over whom they hold power. The author's explore this concept of legitimacy, specifically the means through which it is cultivated by performance of authority - the words and actions those in authority use to reinforce their legitimacy.

In relation to the research question discussed above, this theory is used to explain the expected difference in magistrate's behaviour between sentencing and non-sentencing decisions. Magistrates are expected to perform certain behaviours in order to cultivate legitimacy when it is most needed - at sentencing - and expected to not perform to the same extent during more routine decisions.

4. Was your assigned paper a descriptive study, an identification exercise, a numerical solution to system of equations study, or some combination of the three? (These are the three classifications we discussed in class.)

This paper is a mixture of a descriptive study and an identification study. The authors introduce a new data set which they coded themselves, which is an element of descriptive studies. Apart from this, the study is mostly focused on identification. The authors seek to understand how the type of decision a magistrate makes (exogenous) and the presence of legal council (exogenous) shapes the extent to which they perform their authority (endogenous),

as measured by their behaviour towards defendants. They also examine these same exogenous variables with respect to the order of the decision (endogenous).

This research design would allow the researchers to make some explanatory claims, asserting there is a causal mechanism at play in line with those identified by Weber. Specifically, the interpersonal communication strategies (eye contact, speaking directly to the defendant), as well as the structure of the decision are hypothesized to be the means through which the magistrate asserts their legitimacy and performs their authority for the court. Thus the need for legitimacy explains the variation in magistrate behaviour.

5. What computational methods did this paper use to answer the research question? What was their result or answer to the question?

The researchers hand-coded their dataset but used computational analyses to compare their results and establish statistical significance. Specifically, they used Chi-square tests to examine the difference in observed counts across their dependent variables. They find (Roach Anleu and Mack, 1058):

- The type of decision is significant and positively correlated with magistrates' speaking to/looking at defendants;
- The presence of council does not appear to impact the magistrates' behaviour in sentencing decisions - they continue to engage with the defendant directly;
- The presence of council does appear to impact the magistrates' behaviour on non-sentencing, more routine decisions;
- The type of decision has a significant effect on the structure of a decision, as it is communicated to the defendant.

In sum, the paper finds routine decisions which do not require great legitimacy (deciding a recess, for example) are delivered in more impersonal ways and may be directed to legal council rather than the defendant. These decisions are often given without much explanation or justification. Sentencing decisions, which require greater legitimacy, are more often given with personal engagement between the defendant and the magistrate. These are also more often prefaced with structured legal reasoning, providing a framework for defendants to understand the rationale for the decision.

6. Think of yourself as an academic referee. Give two suggestions to the author(s) of your assigned paper of things the authors might do to improve their results or strengthen their evidence for the answer to the question.

1. Adapt methodology to avoid biased hand-coding

The risks inherent in hand-coding, even when done by two trained researchers, are substantial. These include the potential of one researcher's perspective to dominate, in addition to the risks of variations in human emotion, mood, and attentiveness across time. Furthermore, using only two principal coders, likely of similar background, training, and beliefs, risks the introduction of systematic bias.

A more robust procedure would be to assess magistrate behaviour through:

- Audio analysis of recordings (where available) for characteristics such as valence and arousal;
- Analysis of court transcripts for personal pronoun usage, which would indicate the magistrate speaking directly to the defendant.

However, certain restrictions specific to courts may make these two approaches unfeasible. In that case, the expansion of the sample to include more cases coded by a more diverse pool of coders would help to mitigate this risk. The researchers should include explicit mentions of the coders demographic characteristics, as the differential impact of race on judicial experience is well-known.

2. Communicate causal mechanisms and specific research question more clearly

The researchers present their research question at various points in several ways. For example:

- "This paper investigates how magistrates perform their authority in the delivery of decisions in open court." (pp 1052)

- "The key research question is whether sentencing decisions are communicated in distinct ways." (pp 1053)

- "It examines the manner in which magistrates communicate their decisions within the parameters of the socio- legal setting of the courtroom and considers the implications of those behaviours for performing authority and maintaining legitimacy. ... This paper investigates how magistrates perform their authority in the face-to-face delivery of decisions, particularly to the defendant." (pp 1053)

While the question does not change in a meaningful way across these statements, it hurts the paper's clarity. I would suggest stating the question clearly at the

14

beginning of the paper, and providing a table of endogenous and exogenous variables with clear description of the hypothesized causal mechanisms at play in each of these formulations of the model.

References

Roach Anleu, S., & Mack, K. (2015). Performing authority: Communicating judicial decisions in lower criminal courts. Journal of Sociology, 51(4), 1052-1069.