Codebook for Dallas elites and upper-class dataset: "Kinship Interlocks" replication edition

Shay O'Brien

This codebook explains each variable in each sheet in the main data file, `blue2.xslx`. The dataset is all hand-constructed, and many of the variables are incomplete or ad hoc. You can think of it as somewhere between a traditional database and really structured note-taking. If you would like to use it in any way beyond a simple replication, please don't hesitate to reach out to me! I will be thrilled you're using it, and I would love to give more context on the variables you're interested in and point you to further resources to flesh them out as-needed.

Individuals sheet

The Individuals sheet contains the individual-level data. It is difficult to precisely describe the denominator of this population, because I add people for many different reasons. Usually, they were listed in a Dallas social register and/or they were a Dallas elite by some measure. Sometimes, they never lived in Dallas, and/or they weren't really elite or upper class by any standard measure, but they were intimately entangled with the Dallas elite in a way I felt was substantively important to something I'm interested in understanding. If you'd like to use the data for your own original analyses, I encourage you to subset this full population in a more precise way to help with interpretability.

Please note that although the dataset is constructed from publicly available sources, in this version, I have deleted anyone who could possibly still be alive (meaning that they were born after 1915 and I do not have a death year for them). This means that replications will generate slightly different denominators for elite/upper class categories than the ones in the paper. The main empirical patterns should be unaffected.

Variable	Type / Values	Description / Notes
id	Integer	Unique, random person identifier.
last	String	Surname. For married women, this is their last known surname.
first	String	First name.
middle	String	Middle name(s).
nickname	String	Nickname.
suffix	String	Name suffix (e.g., Jr., III).
maiden	String	Natal surname, for married women.
last_penult	String	Penultimate surname for multiply married women.

Continued on next page

Variable	Type / Values	Description / Notes
last_antipen	String	Antepenultimate surname for multiply married women.
last_antianti	String	Ante-ante-penultimate surname for multiply married women.
last_anti3	String	Third prior surname for multiply married women. I haven't needed another column yet (so far).
black	Binary (0/1)	Person socially categorized as Black based on state categories in archival records. I am slowly adding within-racial-category elites; there will be more (and more race variables) to come. If NA, you can assume they were white. One person in the dataset (id 19525) has a 1 in this column, but she lived sometimes as white and sometimes as black; see my paper "Unnatural Wills" for details.
woman	Binary (0/1)	1 = woman, 0 = man. I haven't yet found anyone socially categorized in any other way, or who I am certain crossed gender categories in their lifetimes.
byear	Integer	Birth year.
bmonth	Integer (1–12)	Birth month.
bdate	Integer (1–31)	Birth date.
birthtown	String	Birth town/city.
birthcounty	String	Birth county.
birthstate	String	Birth state.
bcountry	String	Birth country. If blank and the state is filled in, it's the U.S.
dyear	Integer	Death year.
dmonth	Integer (1–12)	Death month.
ddate	Integer (1–31)	Death day of month.
deathtown	String	Death town/city.
deathcounty	String	Death county.
deathstate	String	Death state.
deathcountry	String	Death country. If blank and the state is filled in, assume U.S.

Continued on next page

Variable	Type / Values	Description / Notes
death_cause	String	Cause of death. Mostly from Texas Death Certificates, but sometimes from other archival sources. Incomplete but not bad. I think someone could do something really interesting with this.
arrive_dallas	Integer / String	Year arrived in Dallas. If I know they never lived there, it says "never."
left_dallas	Integer	Year left Dallas (if known). If I know they never lived there, it says "never."
years_ed	Numeric	Years of education. Super incomplete. Could get much more complete data by linking to Census.
woman_everworkedpay	Binary (0/1)	Indicator that I know a woman ever worked for pay. This was unusual. Useful especially for women who briefly worked for pay but were mostly homemakers.
career	String	Primary occupation/career description. I wouldn't recommend using this as any kind of systematic categorical variable. Most are "Homemaker" or "Businessman." I mark women as "Homemaker" when I know they mostly didn't work for pay (important to note, though, that few did much physical housework themselves - they generally supervised the labor of of domestic workers). For men in these data, their jobs are more based on owning stuff than doing anything in particular, so I just write "Businessman."
industry_primary	String	Primary industry. Better as a categorical variable than `career`, but definitely could use some cleaning up. Note that if a woman mostly worked as a homemaker, I've left this blank.
estate	Numeric	Total estate size at death (if recorded in any source I have found). Not adjusted for inflation.

Continued on next page

Variable	Type / Values	Description / Notes
estate_real1860	Numeric	Real estate value in 1860 US Census. This and all Census variables are hand-coded. I should obviously do some kind of matching thing with the full count data but I just haven't. I like to look at every number for every person and think about what that means in terms of their lifecourse lol.
estate_pers1860	Numeric	Personal estate value in 1860 US Census.
estate_real1870	Numeric	Real estate value in 1870 US Census.
estate_pers1870	Numeric	Personal estate value in 1870 US Census.
home_30	Numeric	Home value in 1930 US Census.
rent_30	Numeric	Monthly rent in 1930 US Census.
home_40	Numeric	Home value in 1940 US Census.
rent_40	Numeric	Monthly rent in 1940 US Census.
income_40	Numeric	Income in 1940 US Census. Note that it is top-coded at 5000 (although some enumerators ignored this). Most elite men reported either 5000 or 0 (meaning that they didn't have wages- they were self-employed).
wages_50	Numeric	Wages in 1950 US Census. Same caveats as above, except top-coded at 10,000. Also, only some people were asked this question.
business_50	Numeric	Business income in 1950 US Census. Top-coded at 10,000 and only some people were asked this question.
dividends_50	Numeric	Dividends in 1950 US Census. Top-coded at 10,000 and only some people were asked this question
coloredamerica_whoswho	Binary (0/1)	Listed in *Colored America Who's Who*.
whoswho	Binary (0/1)	Listed in U.S. general Who's Who directory.
whoswho_firstyrfound	Integer	First year found in Who's Who.
deb_elsewhere	Binary (0/1)	Debutante outside of Dallas.
deb_where	String	Location(s) where they debuted, if it wasn't Dallas.

Continued on next page

Variable	Type / Values	Description / Notes
deb_jewish	Binary (0/1)	Jewish debutante indicator (for Jewish debutante ball that happened some years in Dallas).
deb	Binary (0/1)	Debutante or not.
deb_year	Integer	Debut year.
religion	String	Religion (Christian/Jewish).
denomination	String	Religious denomination. Not complete but not bad either.
father_birth	String	Father birthplace/state. Ad hoc.
mother_birth	String	Mother birthplace/state. Ad hoc.
cw_c	Binary (0/1)	Civil War Confederate veteran. Ad hoc but probably the most complete of the veteran variables bc they loved to brag about this.
cw_u	Binary (0/1)	Civil War Union veteran. Ad hoc.
ww1	Binary (0/1)	World War I veteran. Ad hoc and super not complete.
ww2	Binary (0/1)	World War II veteran. Ad hoc and super not complete.
y86_rb_reg	Binary (0/1)	Listed in 1895 social register. I accidentally named it 86 and have never bothered to change it.
y00_sr_reg	Binary (0/1)	Listed in 1900 Social Register.
y02_sr_reg	Binary (0/1)	Listed in 1902 Social Register.
y04_sr_reg	Binary (0/1)	Listed in 1904 Social Register.
y06_sr_reg	Binary (0/1)	Listed in 1906 Social Register.
y09_sr_reg	Binary (0/1)	Listed in 1909 Social Register.
y12_sbb_reg	Binary (0/1)	Listed in 1912 Social Register.
y14_sbb_reg	Binary (0/1)	Listed in 1914 Social Register.
y32_sr_reg	Binary (0/1)	Listed in 1932 Social Register- the other one.
1941_sr_reg	Binary (0/1)	Listed in 1941 Social Register.
1943_sr_reg	Binary (0/1)	Listed in 1943 Social Register.
1953_sr_reg	Binary (0/1)	Listed in 1953 Social Register.
1955_sr_reg	Binary (0/1)	Listed in 1955 Social Register.
y25_bb_reg	Binary (0/1)	Listed in 1925 Blue Book/Social Register.
y26_bb_reg	Binary (0/1)	Listed in 1926 Blue Book/Social Register.
y27_bb_reg	Binary (0/1)	Listed in 1927 Blue Book/Social Register.
y28_bb_reg	Binary (0/1)	Listed in 1928 Blue Book/Social Register.

Continued on next page

Variable	Type / Values	Description / Notes
y29_bb_reg	Binary (0/1)	Listed in 1929 Blue Book/Social Register.
y30_bb_reg	Binary (0/1)	Listed in 1930 Blue Book/Social Register.
y31_bb_reg	Binary (0/1)	Listed in 1931 Blue Book/Social Register.
y32_bb_reg	Binary (0/1)	Listed in 1932 Blue Book/Social Register.
y33_bb_reg	Binary (0/1)	Listed in 1933 Blue Book/Social Register.
y34_bb_reg	Binary (0/1)	Listed in 1934 Blue Book/Social Register.
y35_bb_reg	Binary (0/1)	Listed in 1935 Blue Book/Social Register.
y36_bb_reg	Binary (0/1)	Listed in 1936 Blue Book/Social Register.
y37_bb_reg	Binary (0/1)	Listed in 1937 Blue Book/Social Register.
y38_bb_reg	Binary (0/1)	Listed in 1938 Blue Book/Social Register.
y39_bb_reg	Binary (0/1)	Listed in 1939 Blue Book/Social Register.
y40_bb_reg	Binary (0/1)	Listed in 1940 Blue Book/Social Register.
y41_bb_reg	Binary (0/1)	Listed in 1941 Blue Book/Social Register.
y42_bb_reg	Binary (0/1)	Listed in 1942 Blue Book/Social Register.
sd7273	Binary (0/1)	Listed in Social Directory 1972-1973.
sd88	Binary (0/1)	Listed in Social Directory 1988.
sd94	Binary (0/1)	Listed in Social Directory 1994.
wikidataID	String	Wikidata identifier.
part_native	Binary (0/1)	Partial indigenous American ancestry indicator, based on multiple archival sources. Ad hoc. Does not include the many white people who tried to claim they were part indigenous to get land in Oklahoma but were denied.
jewish_ancestry	Binary (0/1)	Jewish ancestry indicator, based on multiple archival sources. I use this if someone or someone's very recent ancestor converted from Judaism to Christianity. Ad hoc.

Continued on next page

Variable	Type / Values	Description / Notes
prison	Binary (0/1)	They spent time in prison. Ad hoc.
institutionalized	Binary (0/1)	They were sent to live in an asylum. Ad hoc.
no_kids	Binary (0/1)	Indicator: no children. Useful because sometimes people did have children but the children aren't in the dataset for various reasons (never lived in Dallas, etc.) Not complete but pretty good.
never_married	Binary (0/1)	Indicator: never married. You can infer that women were never-married if the maiden name column is blank. This is useful for men. It's not complete but pretty good.
scandal	Binary (0/1)	I know they were part of a major scandal. Ad hoc.
dem_repub	Categorical	Political party affiliation. Ad hoc and very very incomplete. `D` is Democratic, `R` is Republican, `S` (rare...) is Socialist. A string of letters marks a sequence of shifts over their lifetime, such that `DR` means they switched from Democrat to Republican. Two really singular cases- one, `RSA`, means Republican to Socialist to Anarchist; the other, `CA`, means Communist to Anarchist.
gay	Numeric.	Graded indicator of strong archival evidence that they were queer. 4- confirmed by them; 3- confirmed by others; 2- probable; 1- suspected/possible. Very ad hoc!!!
riches2rags	0, .5, or 1.	Their class destination- 0 if I know they stayed rich/upper class until death, .5 if they became middle class, 1 if they became poor/working class. Ad hoc.

Continued on next page

Variable	Type / Values	Description / Notes
rags2riches	0, .5, or 1.	Their class origin- 0 if they were born rich/upper class, .5 if they were born middle class, 1 if they were born poor/working class. Ad hoc but should be complete for upper-class men who worked in cotton or oil and for their wives. Working on completeness for bigger subsets.

Relationships

These are all mutual kin ties between people in the Individuals sheet. Conceptually, you can frame this as a slight underestimate of all mutual kin ties between people in the data, because I still sometimes find new ties. But it's really very complete at this point - probably the most complete part of the whole dataset. Note that the replication code calculates more ties using these data and adds them to the edgelist. For example, I calculate and add sibling ties when the siblings are tied to the same parents in this edgelist (rather than adding the sibling ties to this edgelist by hand).

Variable	**Type / Values**	**Description / Notes**
p1	Integer	ID for person 1 in the kin tie. Corresponds to ID in individuals dataset. If this is a cross-generational tie (parent-child, auntuncle-niecenephew, etc.), this is the elder person.
first1	String	First name of person 1 (automatically grabbed from the Individuals sheet for readability).
p2	Integer	ID for person 1 in the kin tie. Corresponds to ID in individuals dataset. If this is a cross-generational tie (parent-child, auntuncle-niecenephew, etc.), this is the younger person.
first2	String	First name of person 2 (automatically grabbed from the Individuals sheet for readability).
edgetype	Categorical	Primary relationship type between `p1` and `p2` (e.g., spouse, parent-child, sibling).

Continued on next page

Variable	Type / Values	Description / Notes
et_2	Categorical	This coding provides details about the gender of the intermediary ties, for more complex kin ties direclty listed in the edgelist. The string of characters after the underscore tells you the gender of the ties connecting them, with m for man and w for woman. So an_mwm indicates an uncle/nephew tie connected through a woman. `p1` would be the uncle, `p2` would be the nephew, and they would be connected by their sister/mom, respectively. Useful if you want to get much more specific about kin ties in your analysis.
start	Integer	Start year of marriage (if known). For non-spousal ties, I don't fill anything in because you can infer it from birth/death years.
end	Integer	End year of the marriage. For non-spousal ties, I don't fill anything in because you can infer it from birth/death years.
notes	String	For marriages, indicates whether they ended in divorce or death.

Organization Variables

This dataset contains information on organizations referenced in the membership data. Each row corresponds to a distinct organization. I add them when I come across them.

Variable	Type / Values	Description / Notes
org_id	Integer	Unique, random identifier for the organization. Note that these are four digits and the person-ids are five digits.
name	String	Organization name.
location	String	Location of the organization. Often not filled in because it's Dallas or because I just.. knew it?
type	Categorical	Organization category/type (e.g., club, school, business, social group). Used this at the start but increasingly I neglect to fill this in.

Membership Edgelist

This dataset links individuals to organizations (clubs, schools, workplaces, etc). Each row corresponds to an individual–organization affiliation. I draw on it a bit for the kinship interlocks paper. The organizations I use for the kinship interlocks paper are mostly completely coded - see Appendix A for details- but overall these data are incomplete and ad hoc. If you would like to use them yourself to make, for example, a bipartite organizational network, please reach out to me and I can point you to complete membership data for relevant clubs, etc. that you can digitize and code (if I know of them). Some kinds of data - like company affiliations, or school or university attendance - would require more significant archival work pulling together a ton of different sources, and I'd be happy to help you find the right information for that, too.

Variable	Type / Values	Description / Notes
id	Integer	Unique identifier for the person.
last	String	Individual's last name (for reference/readability).
first	String	Individual's first name (for reference/readability).
org_id	Integer	Unique identifier for the organization.
org_name	String	Organization name (for reference/readability).
from_def	Integer	Membership start year/date ("definite" start year, if available).
from_incl	Integer	Earliest possible membership start year (inclusive bound).
to_incl	Integer	Latest possible membership end year (inclusive bound).
to_def	Integer	Membership end year/date ("definite" end year, if available).
title	String	Membership title. Super incomplete
founder	Binary (0/1)	Indicator that they helped found the organization.
pres	Binary (0/1)	Indicator that they were president.
pres_year_start	Integer	Start year of presidency term (if known).
pres_year_end	Integer	End year of presidency term (if known).
boardchair	Binary (0/1)	Indicator that they served as board chair.
board	Binary (0/1)	Indicator that they served on the board of the organization.
board_year_start	Integer	Start year of board service (if known).
board_year_end	Integer	End year of board service (if known).

Continued on next page

Variable	Type / Values	Description / Notes
vicepres	Binary (0/1)	Indicator that they served as vice president.

Social Years Lookup Table

This lookup table maps Social Register / directory indicator column names in the person-level dataset to the calendar year they represent. It is used to translate binary listing indicators into first/last listing years. There are certainly better ways to do this but I've never bothered to change it.

Variable	Type / Values	Description / Notes
column	String	Name of the indicator column in the person-level data (e.g., `y41_sr_reg`, `y30_bb_reg`).
year	Integer	Calendar year corresponding to the indicator column.