Report 3: Gaps in Marriage Dates

A casual report

Ricky Heinrich for Mélanie Méthot

2023-08-18

Contents

Intro

Wondering about the gap in years between marriages? That's what we are here to investigate. You will find 2 [ACTION NEEDED] bolded in this report.

Problem: more cleaning needed ...

A quick overview: the first column has the difference in years between the first and the second marriage, and the second column the count. It is worth mentioning that the way this was computed was a straight substraction of the year: if somebody's first marriage was in December 1912, and they commited bigamy in January 1913, effectively just one month after, the computation results in '1 year' difference. If the first marriage was November 1912 and the second December 1912, then it is '0 year' difference. This limitation is a result of unstandardized data collection: sometimes, only the year was/could be collected. So it had to be used as a common denominator for all. Future work may be to take the time to include the month and day information of those that have it, in the case that the information is available for both their first and second marriages, and to take a more comprehensive difference. But that require much more time and effort to work through so for now we move forward with our limitations.

We immediately see some things 'wrong': there are negative 'years' (ie second marriage comes before the first marriage).

Var1	Freq
-20	1
0	57
1	116
2	156
3	176
4	192
5	201
6	159
7	133
8	136
9	118
10	98
11	99
12	76
13	83
14	73
15	52
16	41
17	49
18	46
19	34
20	28
21	22
22	19
23	16
24	14
25	16
26	12
27	8
28	14
29	13
30	16
31	3
32	3
33	8
34	2
35	1
36	5
37	2
38	2
39	1
40	3
42	1
47	1
48	1

There are 10 cases where the date of the first marriage appear to be after that of the second. These cases are shown in the following table. I suppose it may be a mixup from the transcriber, where they meant the 'first' marriage is the one dealt with 'now' and the second one is the previous one. Or they may be typos, like in the case of the second row here, where it may be either 1918, 1928, or possibly even 1912 (I quickly looked on the original spreadsheet to see if I confirm myself with primary sources but for this example there were no links...). I also quickly looked at the primary sources for the first item, to see if I could decide on

something, but didn't find any mention of years in the first three links I looked through and decided to use my time better. **[ACTION NEEDED: fix this typo]**

	state	FIRST_	SECOND_	THIRD_
1	Queensland	Jan 17, 1925	April 30, 1905	NA

There are four cases where the third marriage appears at an earlier date than the second marriage. Does this suggest that the 'third marriage' in this case is in fact the 'first' chronologically? [post jul 6 meeting: yes]

	state	FIRST_	SECOND_	THIRD_
1	South_Australia	1893, May 17 (5)	1913, June 14	1883, May 23

There a two cases where the third marriage appears before the first marriage, which were included as well in the previous table. I checked just in case there would be some that are before the second but after the second, for whatever reason.

	state	FIRST_	SECOND_	THIRD_
1	South_Australia	1893, May 17 (5)	1913, June 14	1883, May 23

Given the complication of some cases having a 'third' marriage that seems to be the first chronologically, I won't code up a solution for these cases under the logical assumption of a transcriber mixup, but rather wait on your thoughts, and then we can 'fix' it in the original spreadsheets. [Post Jul6 meeting: I will go and fix these in the original] For now, I will just remove these observations from the analysis.

I had fixed a typo using my best judgement (2938 -> 1938), but there is a typo I am not sure how to fix: what could a year of 8274 for a third marriage be? It seems like you have a whole document on this guy (which I don't have access to) so I'm sure you'll be able to find faster than me the correct year. **[ACTION NEEDED: fix typo]**

state	NAME	FIRST_	SECOND_	THIRD_
Victoria	Edward Seymour	Sep. 17, 1872	May 14, 1877	did not marry

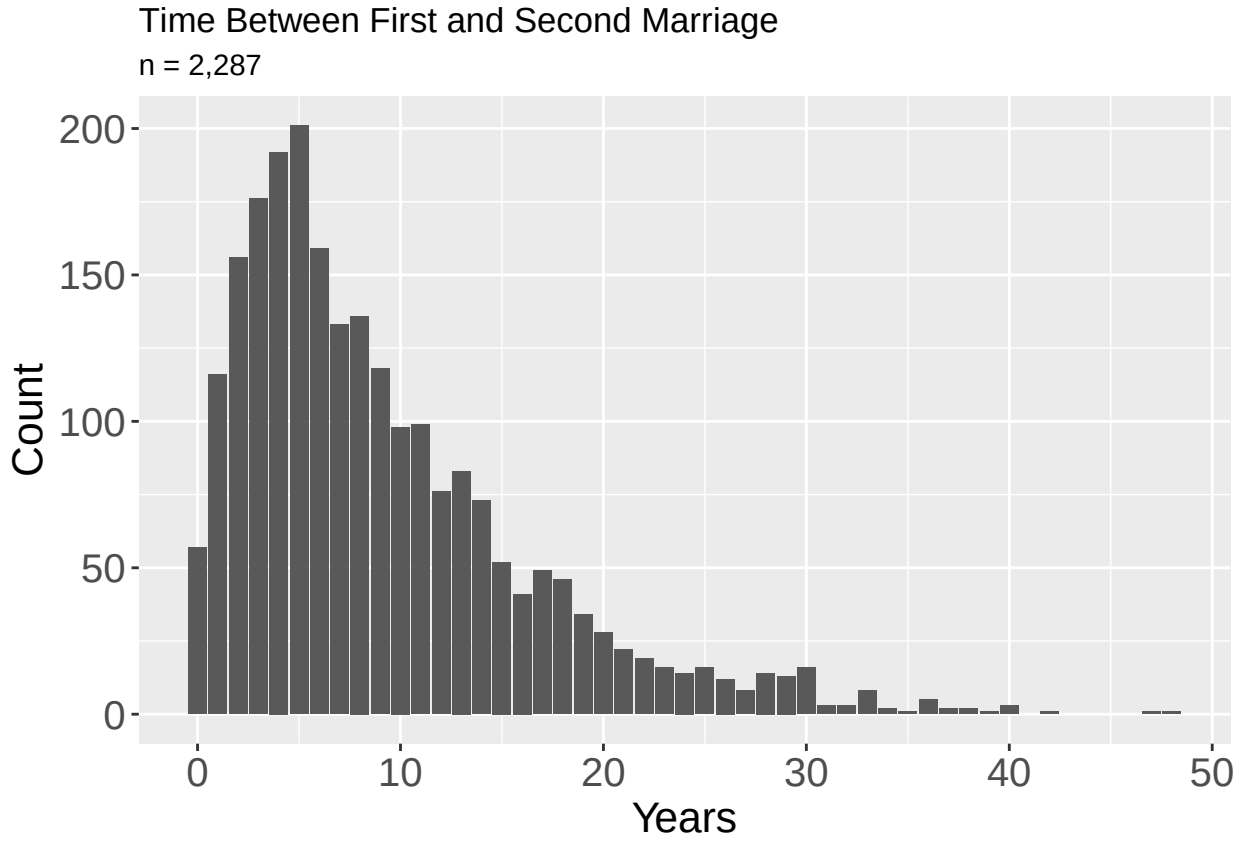
Analysis

First and Second

There are 2,287 cases where a year for a first marriage and a second marriage is recorded, out of 3070 cases available. There are 155 cases where there is a date for the 'second' marriage, but nothing written for the 'first'. 17 of these have a known date for a third 'marriage'.

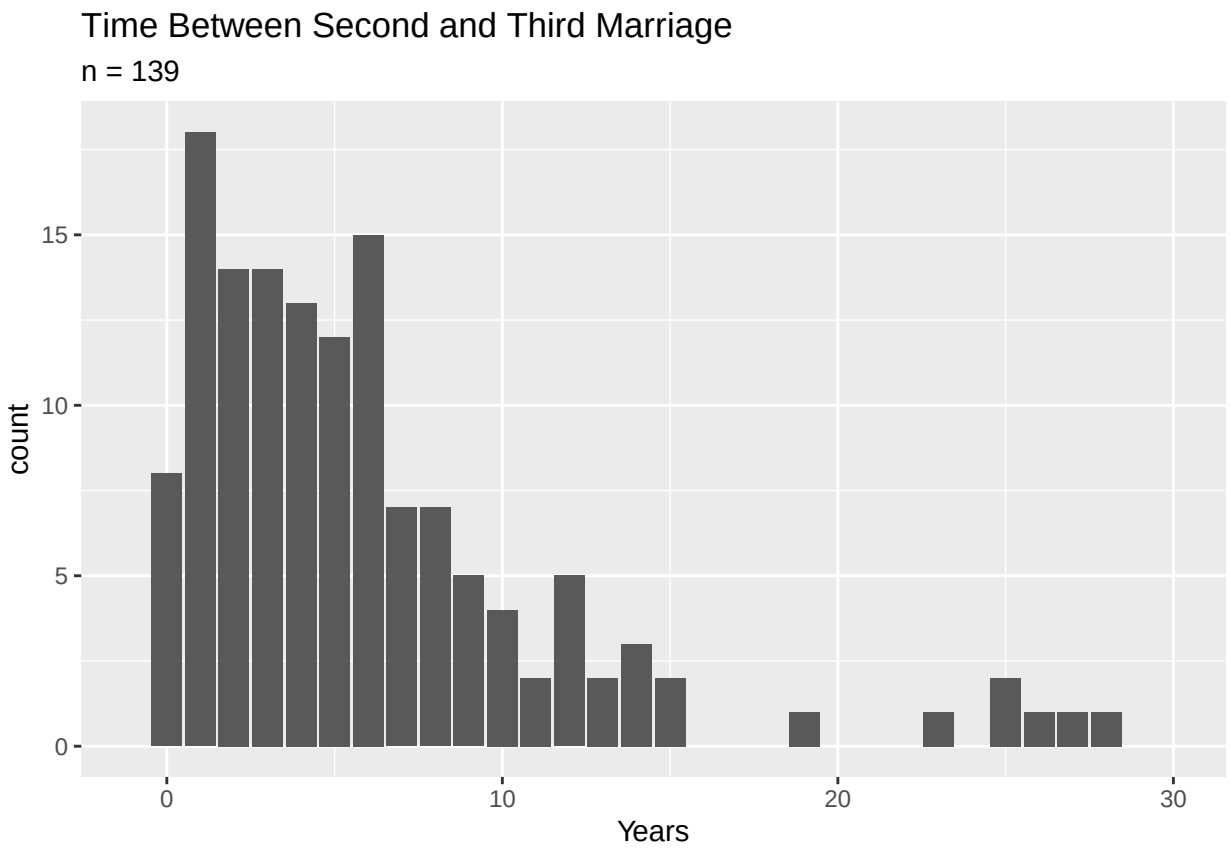
In the following plot, we see a graph of the time between first and second marriages, with the count associated. This plot only includes the cases where the first and second years of marriages were known in the data. The differences range from 0 (happened in the same year) to 48. 99% of differences are under 33 years, 95% are under 24 years, 90% are under 19 years, 75% are under 13 years. Another way to phrase this is that only 10% of cases are over 19 years, 5% over 24 years, and 1% over 33 years. The mean (average) difference is 9 years, and the median is 7 years. The median is the value at which half of

the values (50%) occur below and half of them above. The difference occurring the most is 6 years, followed by 5, then 4. The distribution is left-skewed, ie have a tendency of occuring at 'smaller' differences.



Second and Third

There is a much smaller count of cases where there are both years for a second marriage and a third marriage recorded: only 139. The mean difference in years is 6, and the median is 5. 75% occur under 8 years, 90% under 13, 95% under 17, and 99% under 27.

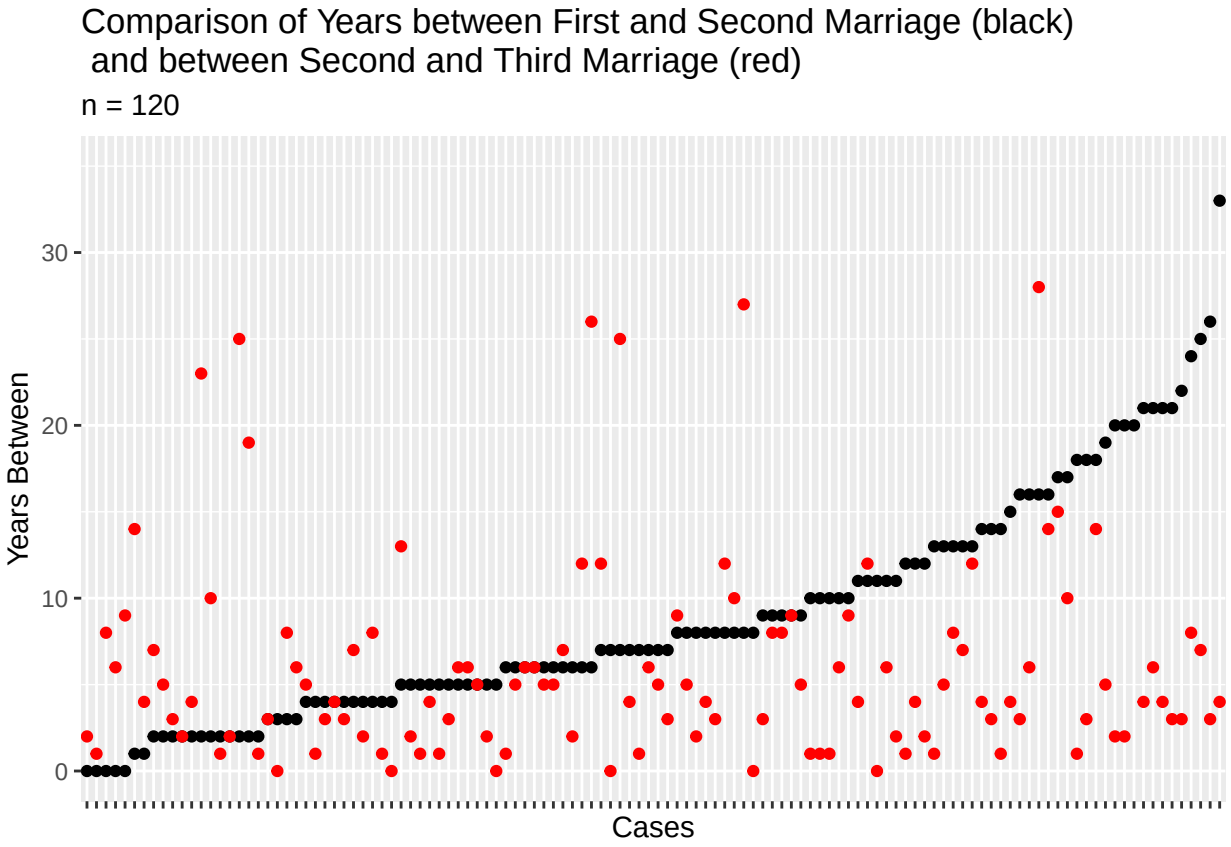


First, Second, AND Third

This next plot needs some work to be more understandable, but provides an initial idea of the years between first and second marriages versus the years between second and third. There are only 120 cases where there is a year recorded for all three events. The black points represent the difference in the first set, and the red points the second. Any red point that is above a black point means that more years elapsed between the second and third compared to the first and the second. Any red point below is the opposite: there was less time elapsed between the second and the third than the first and the second.

It looks like there are more points below the black 'curve', indicating that it seems like in general, less time elapsed between the second and the third marriage. In fact, that was the case 78 times, whereas more time elapsed before getting married a third time 34 times. There was only 8 times when the number of years was equal (when the a red point overlaps directly a black point).

Of interest might be the first few points on the left side of the plot, where the first and second marriage happened in the same calendar year, and the third marriage followed 1 and 2 years behind.


Comparison of Years between First and Second Marriage (black)
and between Second and Third Marriage (red)
n = 120
Years Between
30
20
10
0
Cases