

# INCREASING TRANSPARENCY THROUGH A MULTIVERSE ANALYSIS

---

Sara Steegen, Francis Tuerlinckx, Andrew Gelman, and Wolf Vanpaemel

**KU LEUVEN**

# Multiverse analysis

- Data are not just “collected” or “observed”
- Data are **constructed**: translating observed raw data to processed data set ready for analysis
  - discretization
  - transformation
  - data exclusion
  - ...
- In data construction there are often many researcher degrees of freedom (Simmons, Nelson & Simonsohn, 2011; Gelman & Loken, 2014): researcher has to choose among several options
- Different choices can lead to different data sets: **multiverse of data sets**
- Different data sets can lead to different statistical results: **multiverse of results**

# Multiverse analysis

- Researchers typically consider a single processed data set: **single data set analysis**
- Problematic since choosing among data processing options is often arbitrary
  - arbitrary? when there are different reasonable options + no clear justification
- Arbitrariness is inherited by result
- Is the result robust?
  - which other reasonable results could have been obtained?

# Multiverse analysis

- Researchers typically consider a single processed data set: **single data set analysis**
- Problematic since choosing among data processing options is often arbitrary
  - arbitrary? when there are different reasonable options + no clear justification
- Arbitrariness is inherited by result
- Is the result robust?
  - which other reasonable results could have been obtained?
- Solution: **multiverse analysis**
  - if data processing choices are arbitrary, without a clear justification...
  - ... consider **all** reasonable options
  - perform the analysis of interest across multiverse of data sets
  - summarize results across multiverse

# Demonstration

*Research Article*



## The Fluctuating Female Vote: Politics, Religion, and the Ovulatory Cycle

**Kristina M. Durante<sup>1</sup>, Ashley Rae<sup>1</sup>, and Vldas Griskevicius<sup>2</sup>**

<sup>1</sup>College of Business, University of Texas, San Antonio, and <sup>2</sup>Carlson School of Management, University of Minnesota

Psychological Science  
XX(X) 1–10  
© The Author(s) 2013  
Reprints and permissions:  
sagepub.com/journalsPermissions.nav  
DOI: 10.1177/0956797612466416  
pss.sagepub.com



### **Abstract**

Each month, many women experience an ovulatory cycle that regulates fertility. Although research has found that this cycle influences women's mating preferences, we proposed that it might also change women's political and religious views. Building on theory suggesting that political and religious orientation are linked to reproductive goals, **we tested how fertility influenced women's politics, religiosity, and voting in the 2012 U.S. presidential election.** In two studies with large and diverse samples, **ovulation had drastically different effects on single women and women in committed relationships.** Ovulation led single women to become more liberal, less religious, and more likely to vote for Barack Obama. In contrast, ovulation led women in committed relationships to become more conservative, more religious, and more likely to vote for Mitt Romney. In addition, ovulation-induced changes in political orientation mediated women's voting behavior. Overall, the ovulatory cycle not only influences women's politics but also appears to do so differently for single women than for women in relationships.

# Raw data

- answers to three statements on religiosity
- answers to five statements on social political attitudes
- answer to “what is your current romantic relationship status?” (choose 1 of 4 answer options)
- answers to fertility related questions
  - the start date of the last period
  - the typical cycle length
  - how sure are you about the start date of the last period
  - ...

# Data processing: Durante et al. (2013)

Translating the observed, raw data to the processed data ready for analysis involved several data processing steps

1. cycle day assessment: transforming and combining variables
2. fertility assessment: dichotomization (high vs low fertility)
  - high in fertility when cycle day is between 7 and 14
  - low in fertility when cycle day is between 17 and 25
3. relationship status assessment: dichotomization (single vs committed relationship)
4. exclusion criteria based on cycle length
5. exclusion criteria based on variable sure

# Data processing: Durante et al. (2013)

Translating the observed, raw data to the processed data ready for analysis involved several data processing steps

→ 1 processed data set

→ single data set analysis: effect of fertility x relationship status



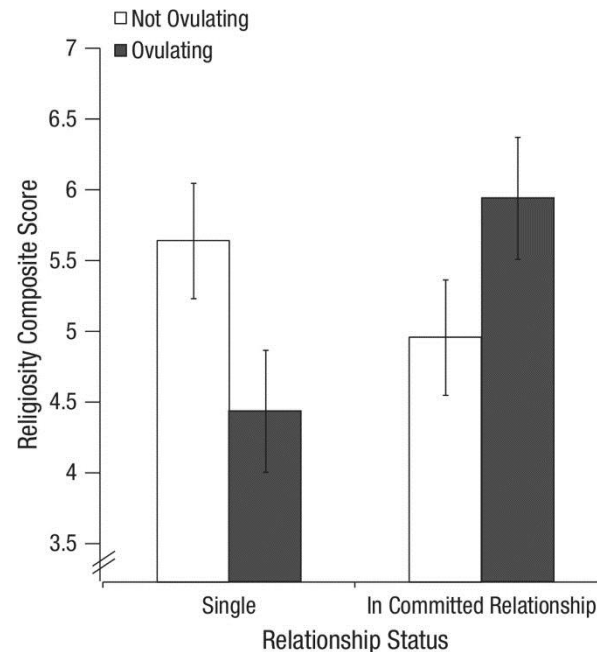
# Data processing: Durante et al. (2013)

Translating the observed, raw data to the processed data ready for analysis involved several data processing steps

→ 1 processed data set

→ single data set analysis: effect of fertility x relationship status

- ... on religiosity



$F(1,159) = 6.46, p = .012$

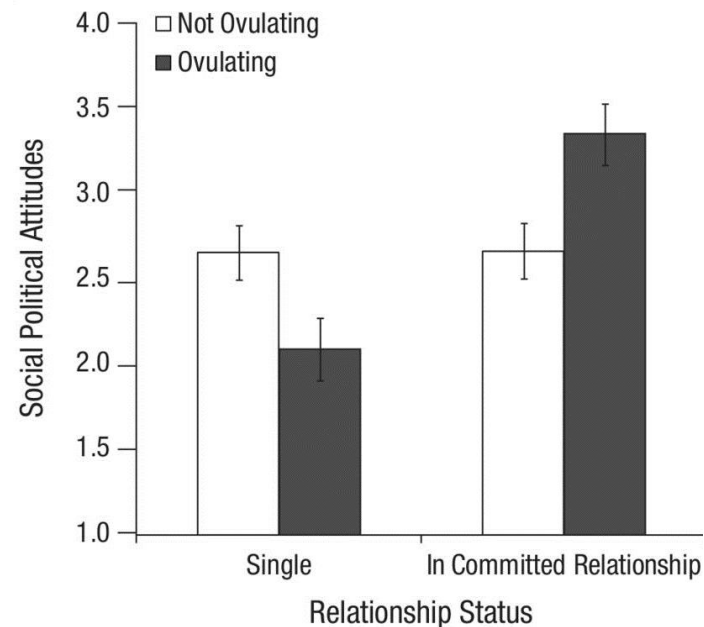
# Data processing: Durante et al. (2013)

Translating the observed, raw data to the processed data ready for analysis involved several data processing steps

→ 1 processed data set

→ single data set analysis: effect of fertility x relationship status

- ... on social political attitudes



$F(1,299) = 12.26, p = .001$

# Data processing: Durante et al. (2013)

Translating the observed, raw data to the processed data ready for analysis involved several data processing steps

→ 1 processed data set

→ single data set analysis: effect of fertility x relationship status

- ... on religiosity:  $F(1,159) = 6.46, p = .012$
- ... on social political attitudes:  $F(1,299) = 12.26, p = .001$

# Data processing: Durante et al. (2013)

Translating the observed, raw data to the processed data ready for analysis involved several data processing steps

→ 1 processed data set

→ single data set analysis: effect of fertility x relationship status

- ... on religiosity:  $F(1,159) = 6.46, p = .012$
- ... on social political attitudes:  $F(1,299) = 12.26, p = .001$

→ are these results robust? which results do we get with other processing choices?

# Data processing: Durante et al. (2013)

Translating the observed, raw data to the processed data ready for analysis involved several data processing steps

→ 1 processed data set

→ single data set analysis: effect of fertility x relationship status

- ... on religiosity:  $F(1,159) = 6.46, p = .012$
- ... on social political attitudes:  $F(1,299) = 12.26, p = .001$

→ are these results robust? which results do we get with other processing choices?

→ multiverse analysis

# Data processing: other reasonable options

1. cycle day assessment: transforming and combining variables
2. fertility assessment: dichotomization (high vs low fertility)
  - high in fertility when cycle day is between 7 and 14
  - low in fertility when cycle day is between 17 and 25
3. relationship status assessment: dichotomization (single vs committed relationship)
4. exclusion criteria based on cycle length
5. exclusion criteria based on variable sure

# Data processing: other reasonable options

1. cycle day assessment: transforming and combining variables
2. **fertility assessment: dichotomization (high vs low fertility)**
  - **high** in fertility when cycle day is between **7 and 14**
  - **low** in fertility when cycle day is between **17 and 25**
3. relationship status assessment: dichotomization (single vs committed relationship)
4. exclusion criteria based on cycle length
5. exclusion criteria based on variable sure

# Data processing: other reasonable options

1. cycle day assessment: transforming and combining variables
2. fertility assessment: dichotomization (high vs low fertility)
  - high in fertility when cycle day is between 6 and 14
  - low in fertility when cycle day is between 17 and 27(Durante et al., 2011)
3. relationship status assessment: dichotomization (single vs committed relationship)
4. exclusion criteria based on cycle length
5. exclusion criteria based on variable sure



# Data processing: other reasonable options

1. cycle day assessment: transforming and combining variables
2. **fertility assessment: dichotomization (high vs low fertility)**
  - **high** in fertility when cycle day is between **9 and 17**
  - **low** in fertility when cycle day is between **18 and 25**(Durante et al., 2012)
3. relationship status assessment: dichotomization (single vs committed relationship)
4. exclusion criteria based on cycle length
5. exclusion criteria based on variable sure

# Data processing: other reasonable options

1. cycle day assessment: transforming and combining variables
2. fertility assessment: dichotomization (high vs low fertility)
  - **high** in fertility when cycle day is between **8 and 14**
  - **low** in fertility when cycle day is between **1 and 7 or 15 and 28**(Durante et al., 2014)
3. relationship status assessment: dichotomization (single vs committed relationship)
4. exclusion criteria based on cycle length
5. exclusion criteria based on variable sure

# Data processing: other reasonable options

1. cycle day assessment: transforming and combining variables
2. fertility assessment: dichotomization (high vs low fertility)
  - **high** in fertility when cycle day is between **9 and 17**
  - **low** in fertility when cycle day is between **1 and 8 or 18 and 28**(Durante et al., 2015)
3. relationship status assessment: dichotomization (single vs committed relationship)
4. exclusion criteria based on cycle length
5. exclusion criteria based on variable sure

# Data processing: other reasonable options

1. cycle day assessment (3 choice options)
2. fertility assessment (5 choice options)
3. relationship status assessment (3 choice options)
4. exclusion criteria based on cycle length (3 choice options)
5. exclusion criteria based on variable sure (2 choice options)

# Data processing: other reasonable options

1. cycle day assessment (3 choice options)
2. fertility assessment (5 choice options)
3. relationship status assessment (3 choice options)
4. exclusion criteria based on cycle length (3 choice options)
5. exclusion criteria based on variable sure (2 choice options)

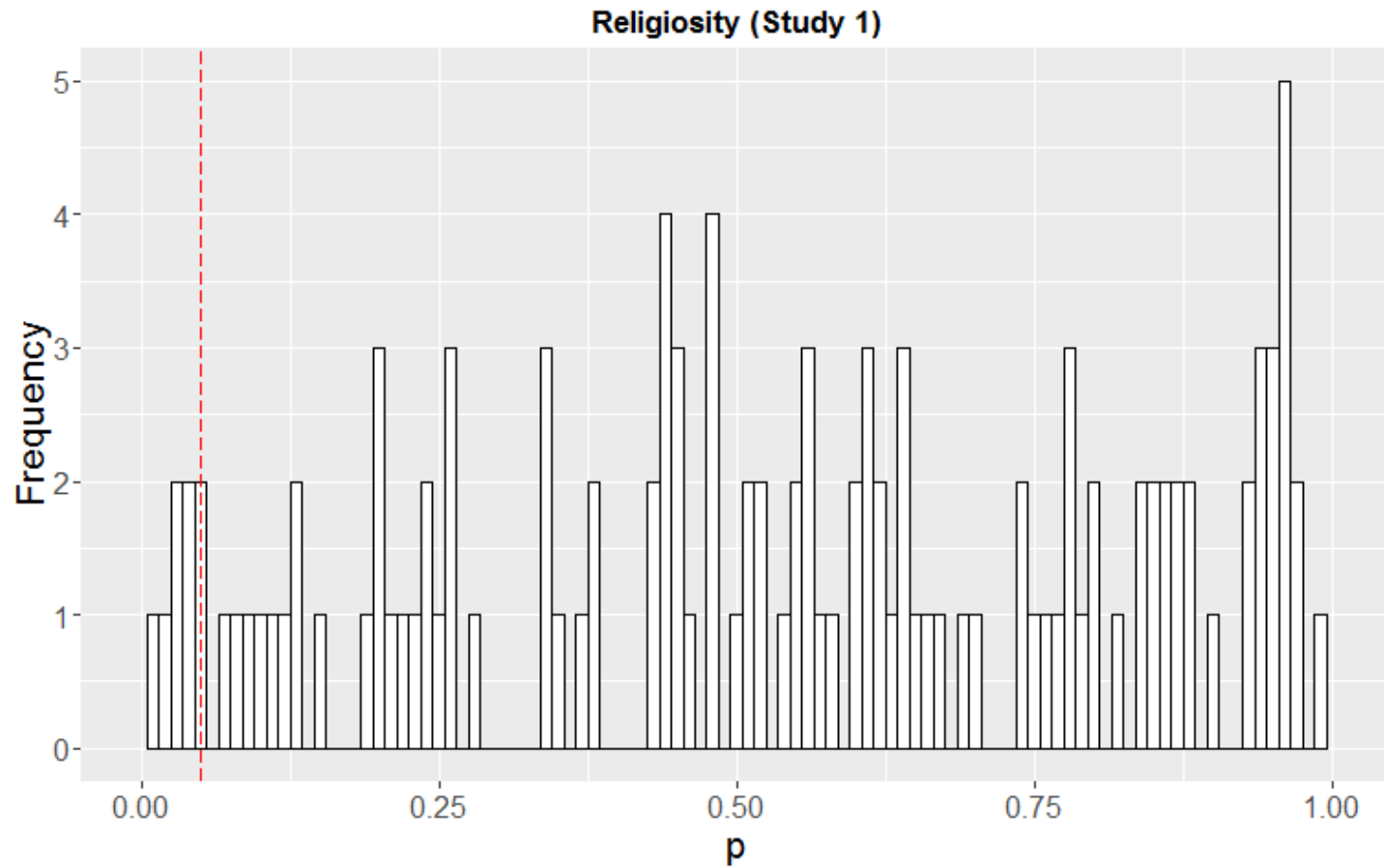
rather than arbitrarily choosing 1 option and creating a single processed data set,  
we consider all combinations of options, which give rise to a multiverse of data sets

# Multiverse analysis: results

perform analysis on each data set in multiverse → multiverse of results

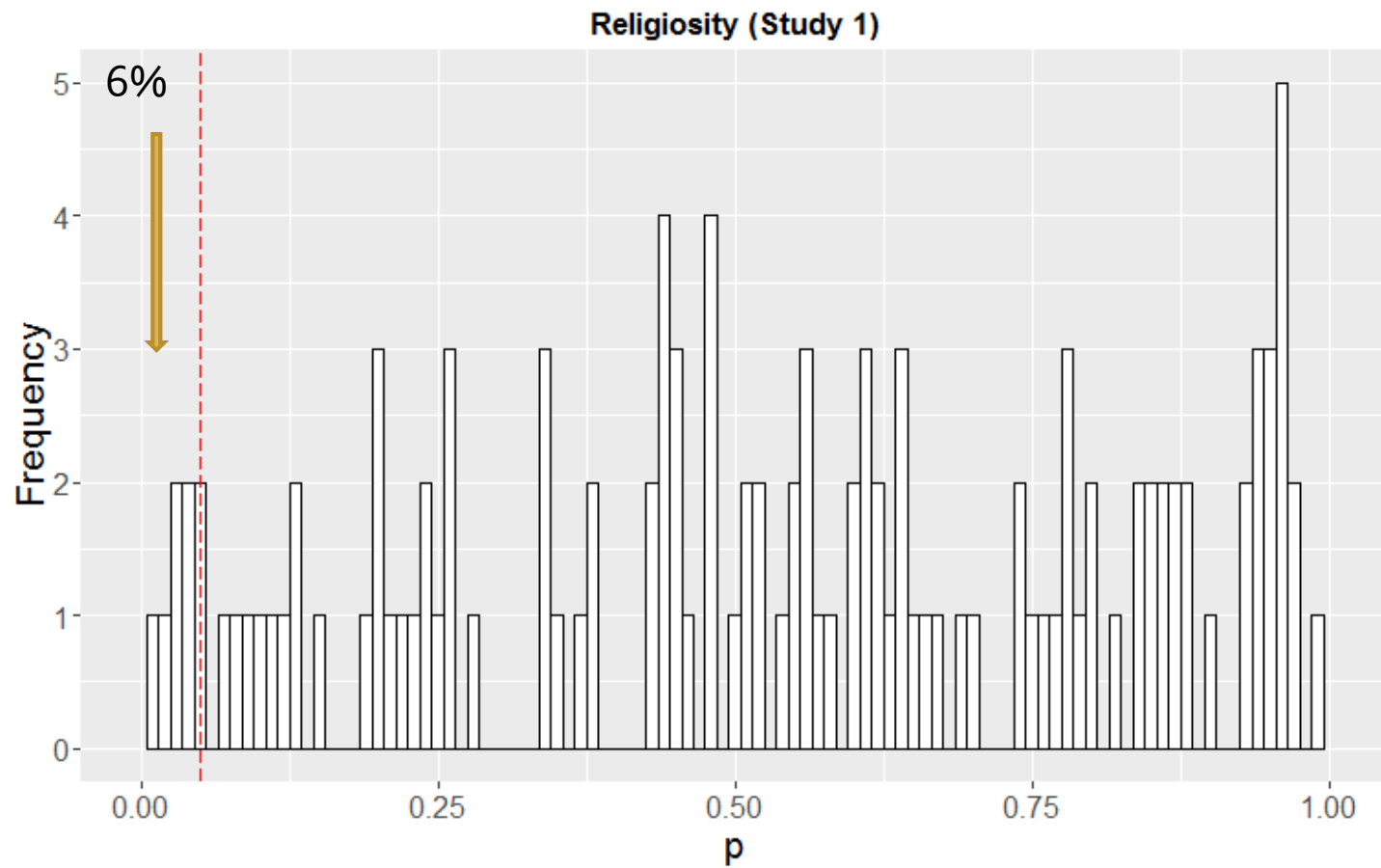
# Multiverse analysis: results

religiosity - fertility x relationship status



# Multiverse analysis: results

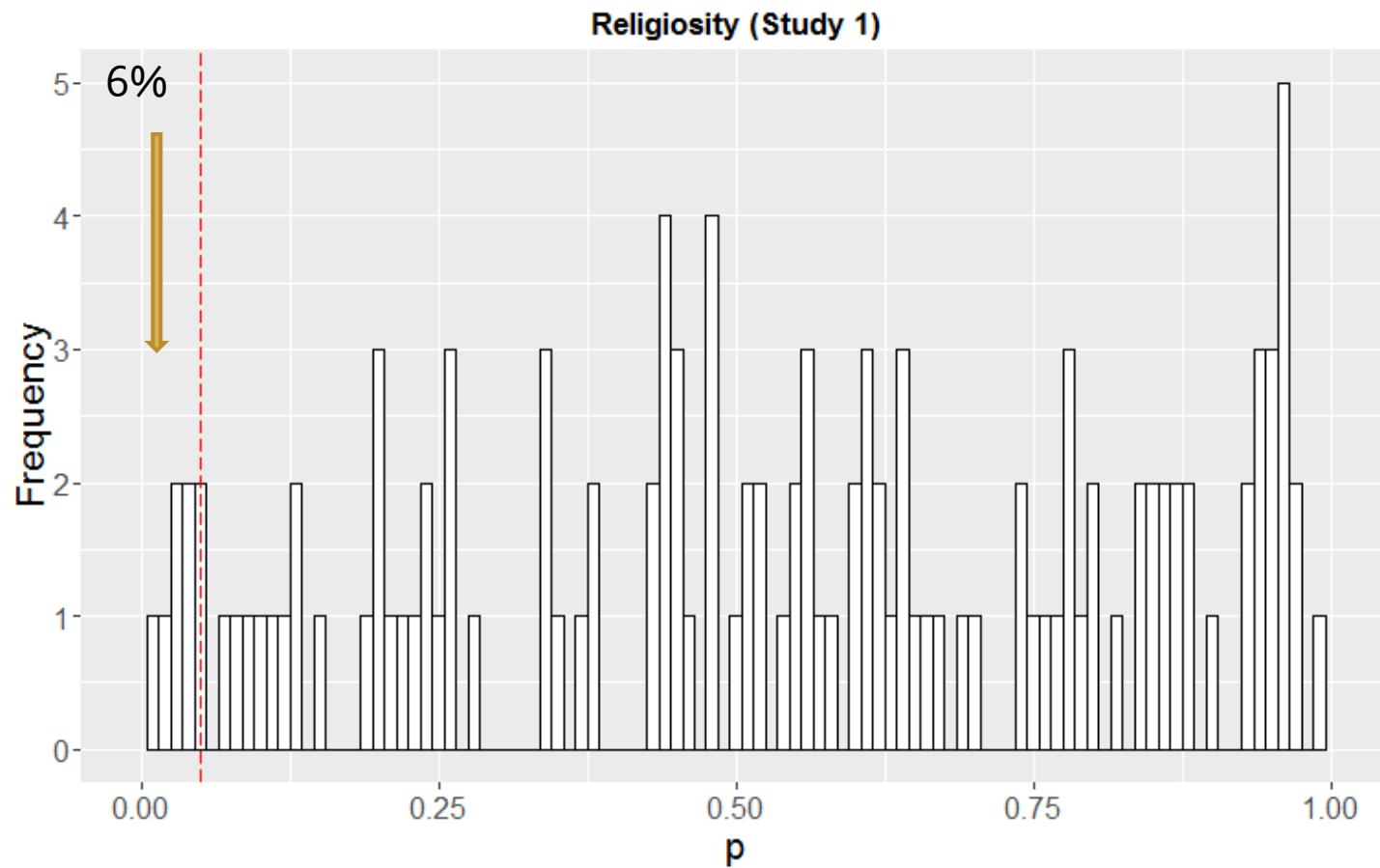
religiosity - fertility x relationship status





# Multiverse analysis: results

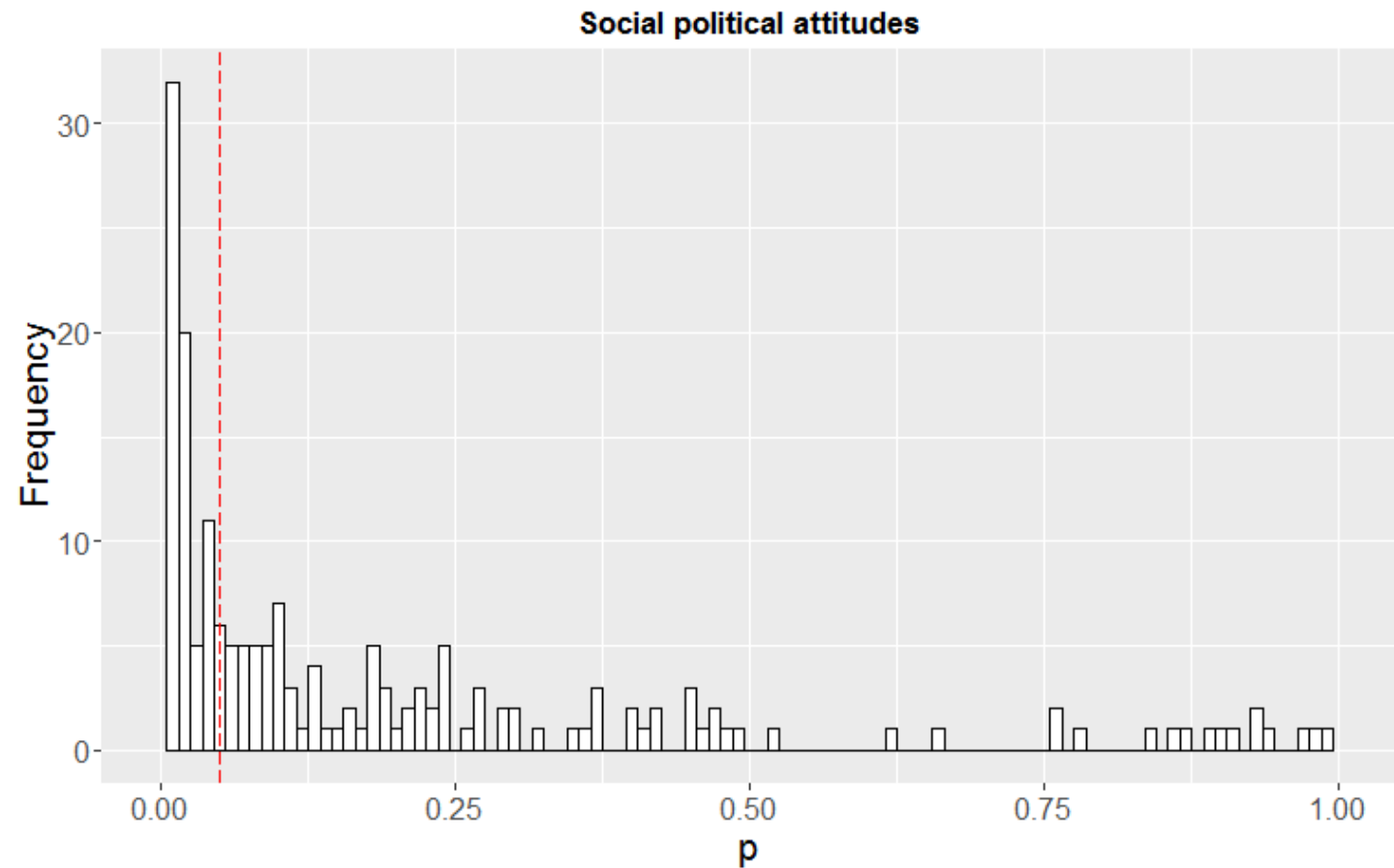
religiosity - fertility x relationship status



effect is too fragile to  
be taken seriously

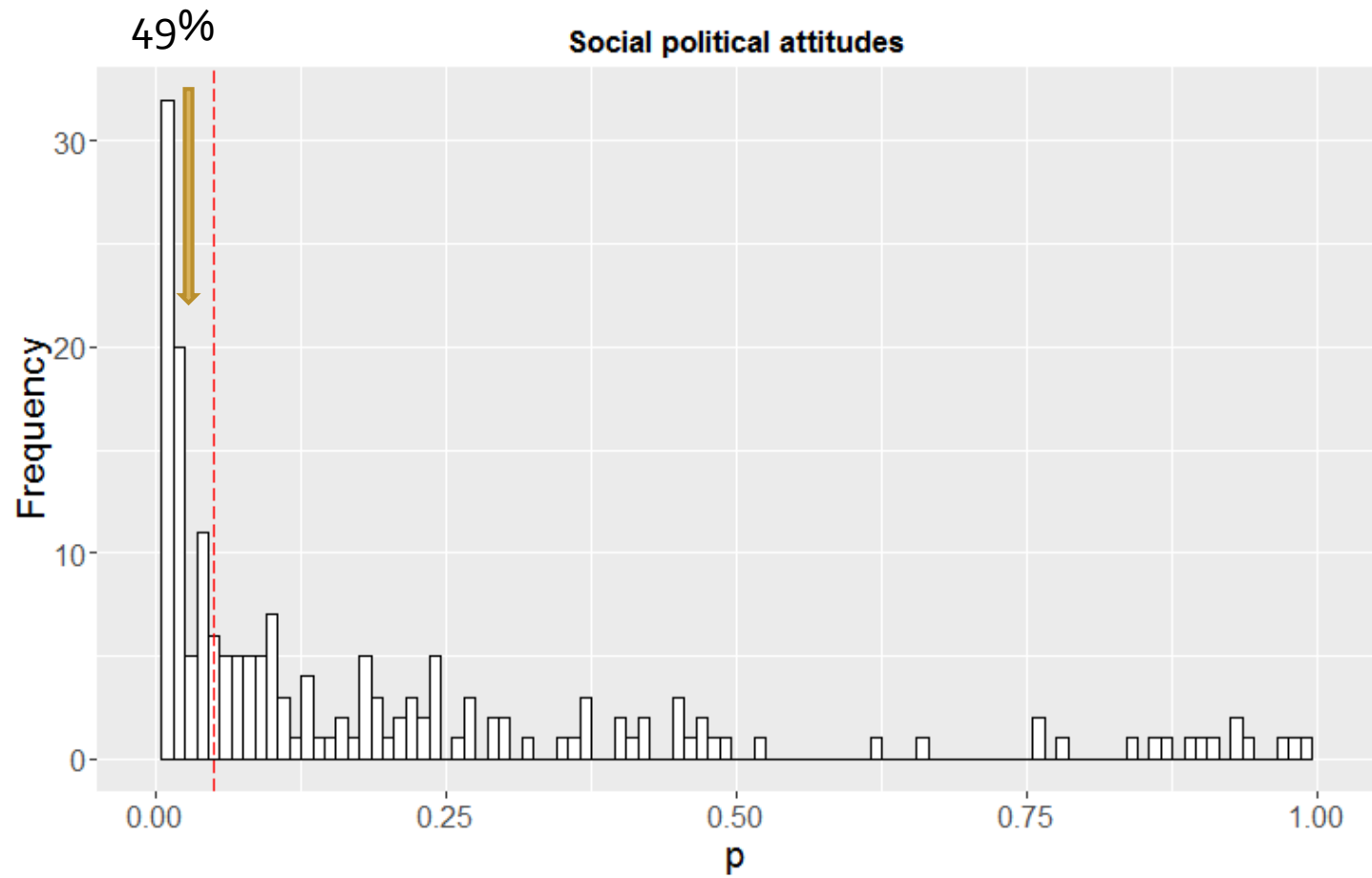
# Multiverse analysis: results

social political attitudes - fertility x relationship status



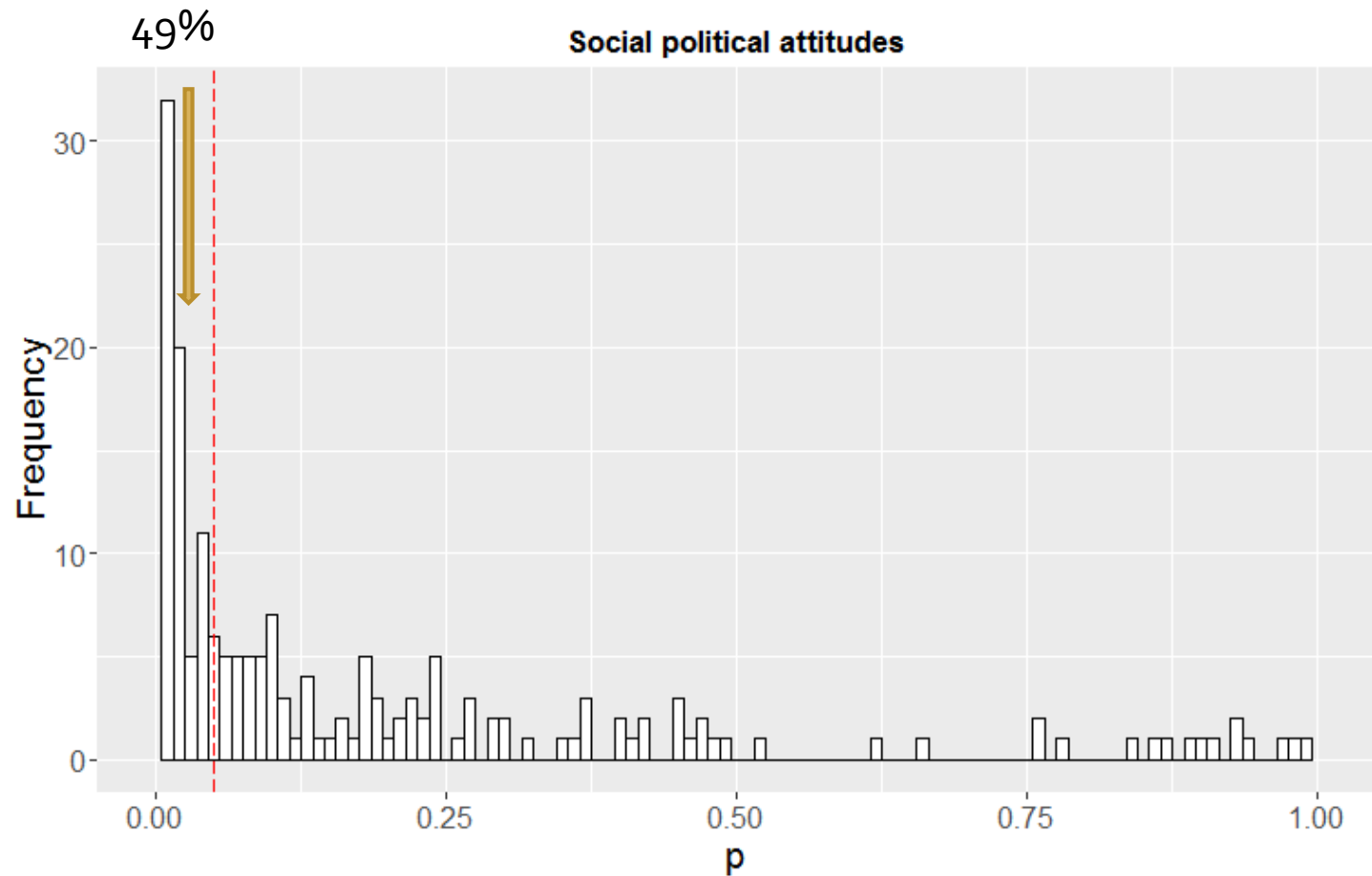
# Multiverse analysis: results

social political attitudes - fertility x relationship status



# Multiverse analysis: results

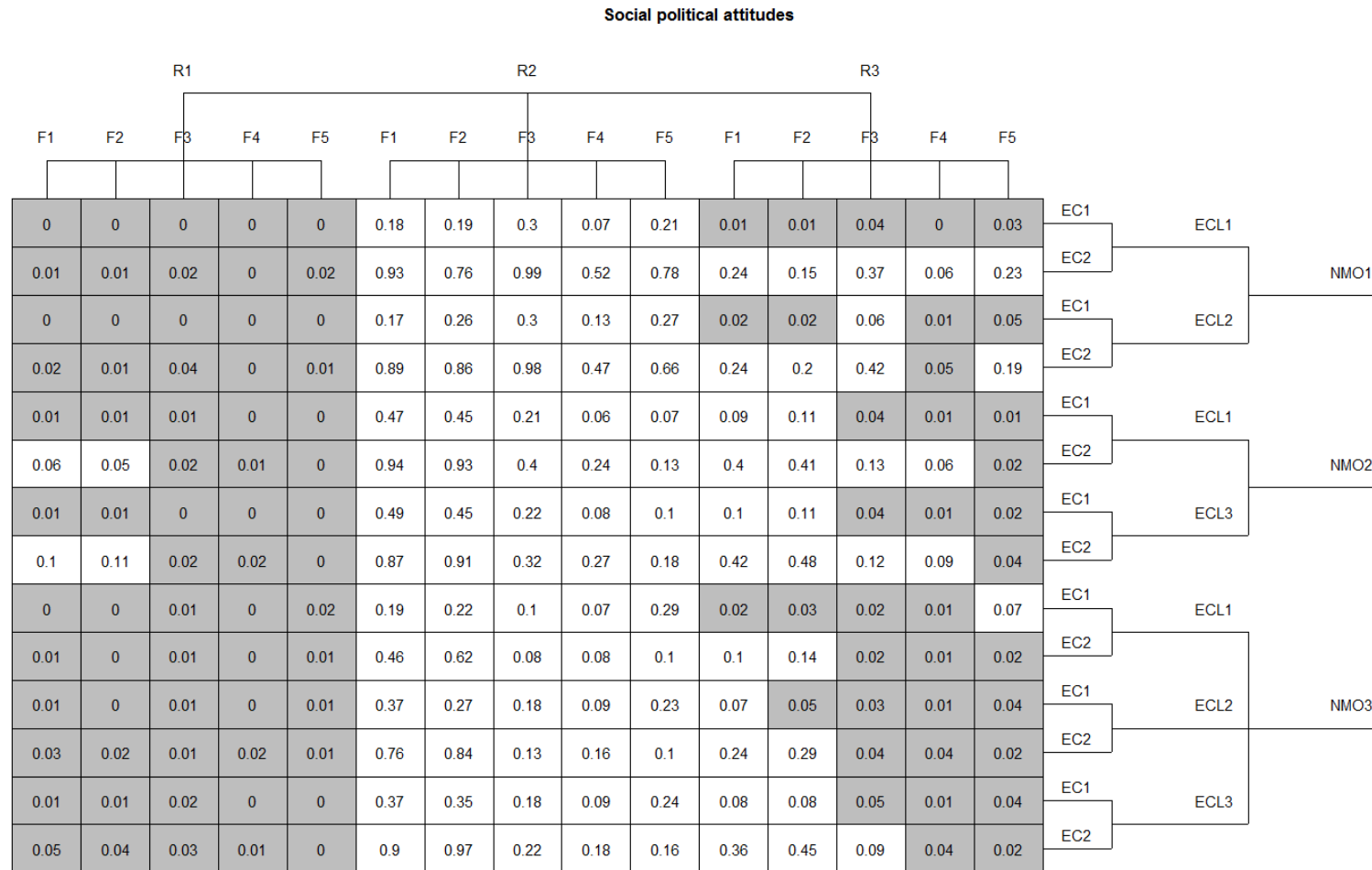
social political attitudes - fertility x relationship status



the belief in the effect heavily depends on the belief in the different processing choices  
→ closer inspection

# Multiverse analysis: results

social political attitudes - fertility x relationship status



the belief in the effect heavily depends on the belief in the different processing choices  
 → closer inspection

# Multiverse analysis: results

social political attitudes - fertility x relationship status

**Social political attitudes**

R1					R2					R3							
F1	F2	F3	F4	F5	F1	F2	F3	F4	F5	F1	F2	F3	F4	F5			
0	0	0	0	0	0.18	0.19	0.3	0.07	0.21	0.01	0.01	0.04	0	0.03	EC1		ECL1
0.01	0.01	0.02	0	0.02	0.93	0.76	0.99	0.52	0.78	0.24	0.15	0.37	0.06	0.23	EC2		
0	0	0	0	0	0.17	0.26	0.3	0.13	0.27	0.02	0.02	0.06	0.01	0.05	EC1		ECL2
0.02	0.01	0.04	0	0.01	0.89	0.86	0.98	0.47	0.66	0.24	0.2	0.42	0.05	0.19	EC2		
0.01	0.01	0.01	0	0	0.47	0.45	0.21	0.06	0.07	0.09	0.11	0.04	0.01	0.01	EC1		ECL1
0.06	0.05	0.02	0.01	0	0.94	0.93	0.4	0.24	0.13	0.4	0.41	0.13	0.06	0.02	EC2		
0.01	0.01	0	0	0	0.49	0.45	0.22	0.08	0.1	0.1	0.11	0.04	0.01	0.02	EC1		ECL3
0.1	0.11	0.02	0.02	0	0.87	0.91	0.32	0.27	0.18	0.42	0.48	0.12	0.09	0.04	EC2		
0	0	0.01	0	0.02	0.19	0.22	0.1	0.07	0.29	0.02	0.03	0.02	0.01	0.07	EC1		ECL1
0.01	0	0.01	0	0.01	0.46	0.62	0.08	0.08	0.1	0.1	0.14	0.02	0.01	0.02	EC2		
0.01	0	0.01	0	0.01	0.37	0.27	0.18	0.09	0.23	0.07	0.05	0.03	0.01	0.04	EC1		ECL2
0.03	0.02	0.01	0.02	0.01	0.76	0.84	0.13	0.16	0.1	0.24	0.29	0.04	0.04	0.02	EC2		
0.01	0.01	0.02	0	0	0.37	0.35	0.18	0.09	0.24	0.08	0.08	0.05	0.01	0.04	EC1		ECL3
0.05	0.04	0.03	0.01	0	0.9	0.97	0.22	0.18	0.16	0.36	0.45	0.09	0.04	0.02	EC2		

if you trust R1, there is a significant interaction effect

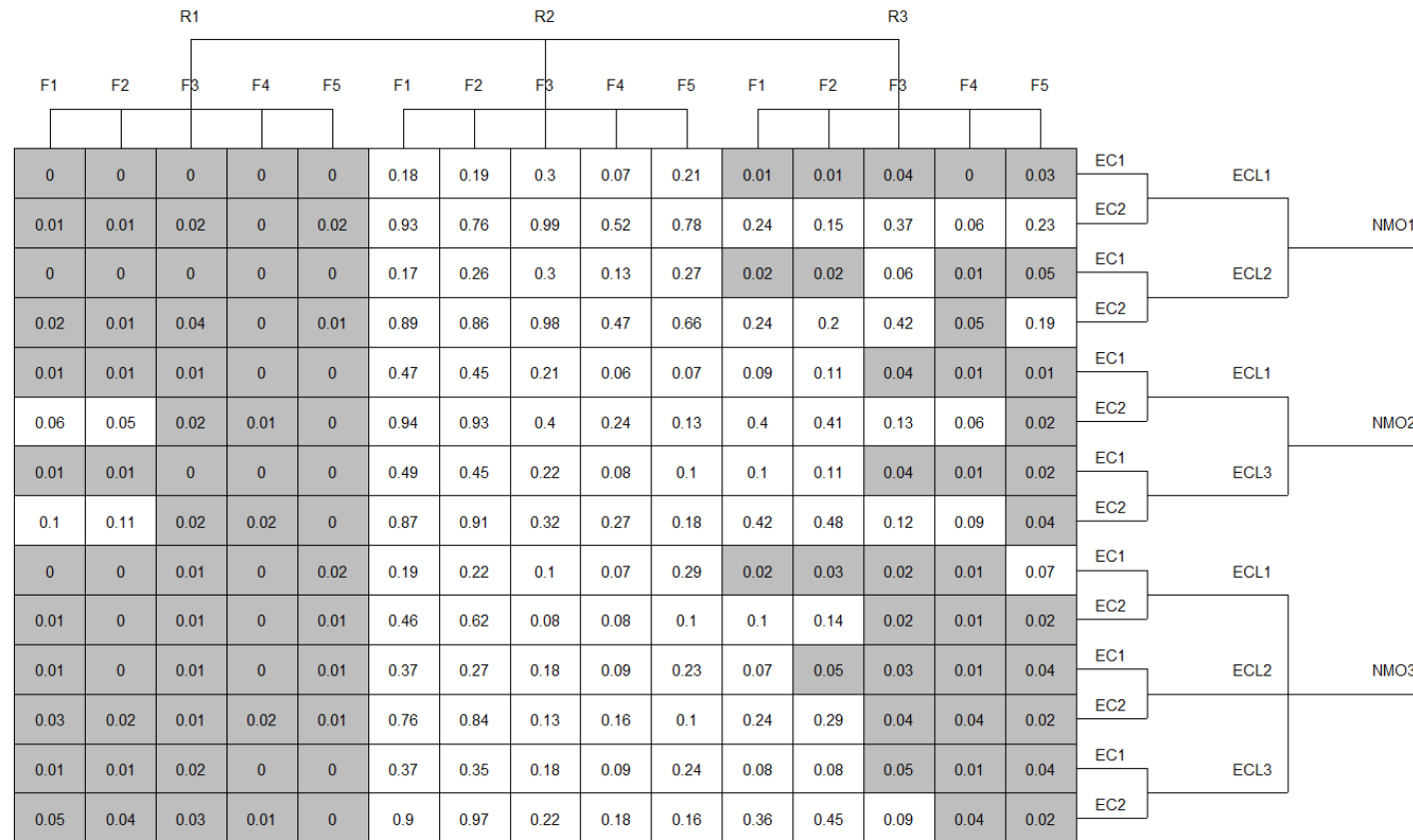
if you trust R2, there is no significant interaction effect

if you trust R3 or don't know what to trust

- you are uncertain about the interaction effect
- your main conclusion should be a methodological one, i.e. we need to **deflate** the multiverse

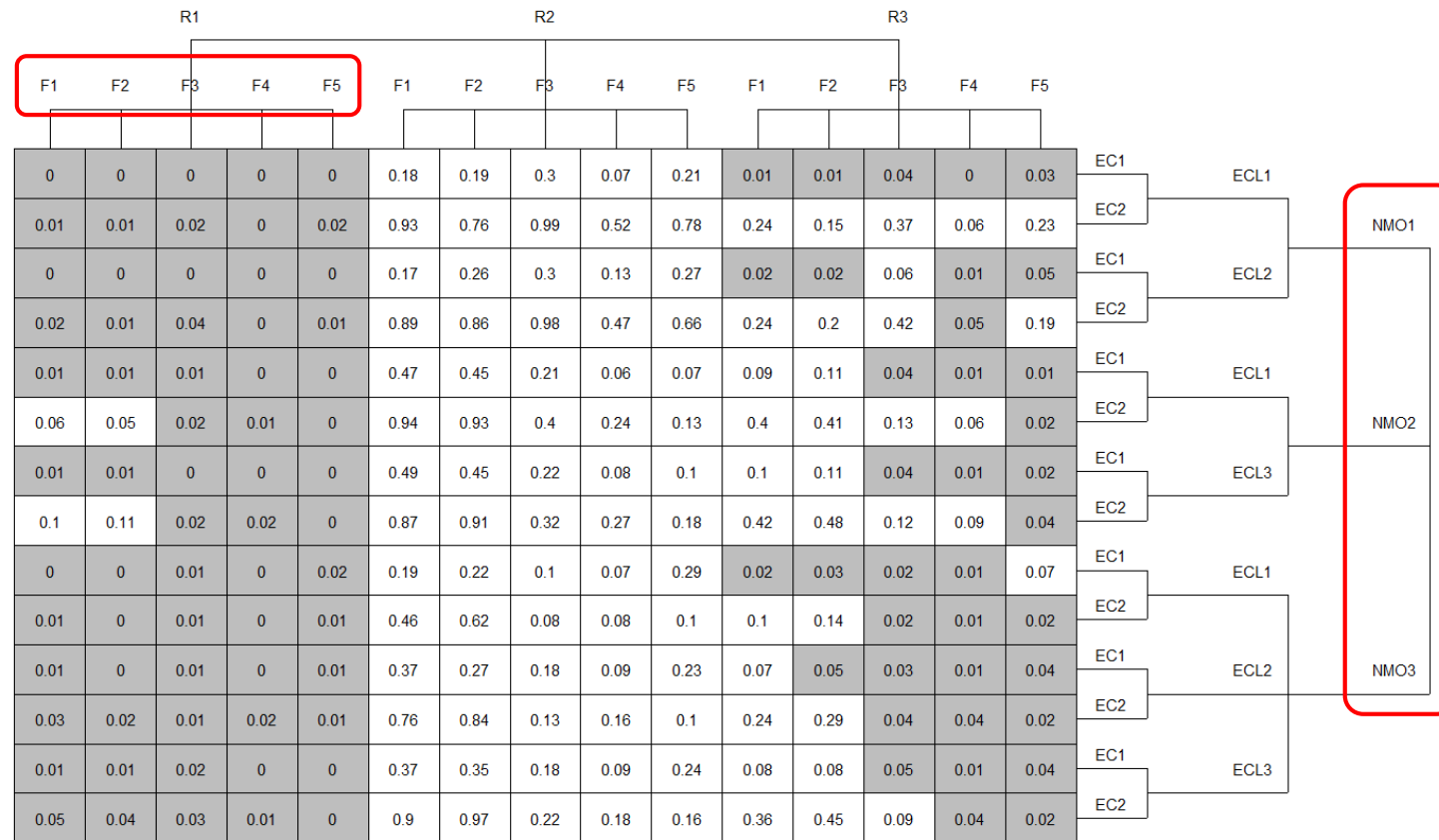
# Deflating the multiverse

- Better theory (assessment of fertility)



# Deflating the multiverse

- Better theory (assessment of fertility)





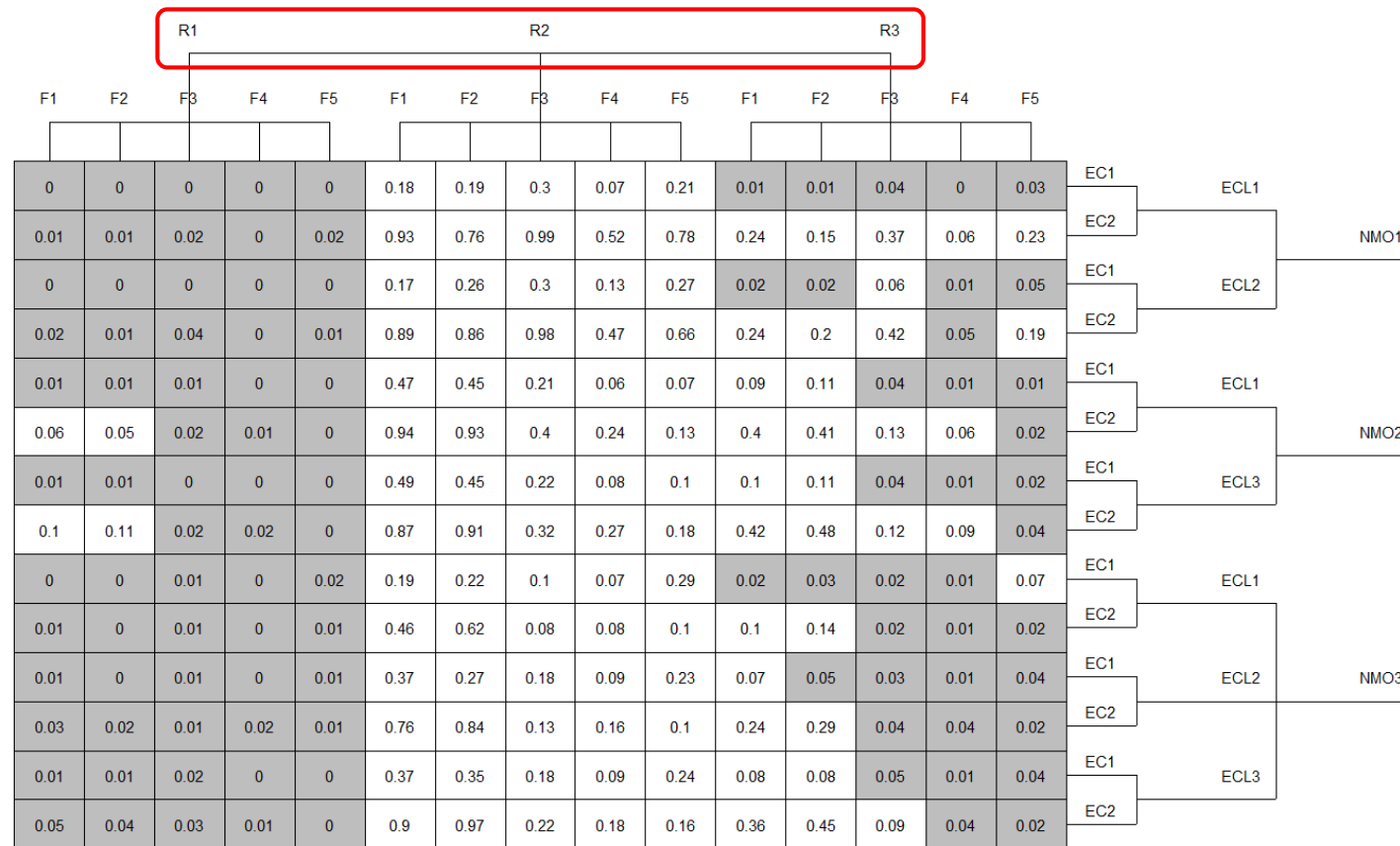
# Deflating the multiverse

- Better theory (assessment of fertility)
- Cleaner design (assessment of relationship status)

R1					R2					R3								
F1	F2	F3	F4	F5	F1	F2	F3	F4	F5	F1	F2	F3	F4	F5				
0	0	0	0	0	0.18	0.19	0.3	0.07	0.21	0.01	0.01	0.04	0	0.03	EC1		ECL1	
0.01	0.01	0.02	0	0.02	0.93	0.76	0.99	0.52	0.78	0.24	0.15	0.37	0.06	0.23	EC2			NMO1
0	0	0	0	0	0.17	0.26	0.3	0.13	0.27	0.02	0.02	0.06	0.01	0.05	EC1		ECL2	
0.02	0.01	0.04	0	0.01	0.89	0.86	0.98	0.47	0.66	0.24	0.2	0.42	0.05	0.19	EC2			
0.01	0.01	0.01	0	0	0.47	0.45	0.21	0.06	0.07	0.09	0.11	0.04	0.01	0.01	EC1		ECL1	
0.06	0.05	0.02	0.01	0	0.94	0.93	0.4	0.24	0.13	0.4	0.41	0.13	0.06	0.02	EC2			NMO2
0.01	0.01	0	0	0	0.49	0.45	0.22	0.08	0.1	0.1	0.11	0.04	0.01	0.02	EC1		ECL3	
0.1	0.11	0.02	0.02	0	0.87	0.91	0.32	0.27	0.18	0.42	0.48	0.12	0.09	0.04	EC2			
0	0	0.01	0	0.02	0.19	0.22	0.1	0.07	0.29	0.02	0.03	0.02	0.01	0.07	EC1		ECL1	
0.01	0	0.01	0	0.01	0.46	0.62	0.08	0.08	0.1	0.1	0.14	0.02	0.01	0.02	EC2			
0.01	0	0.01	0	0.01	0.37	0.27	0.18	0.09	0.23	0.07	0.05	0.03	0.01	0.04	EC1		ECL2	NMO3
0.03	0.02	0.01	0.02	0.01	0.76	0.84	0.13	0.16	0.1	0.24	0.29	0.04	0.04	0.02	EC2			
0.01	0.01	0.02	0	0	0.37	0.35	0.18	0.09	0.24	0.08	0.08	0.05	0.01	0.04	EC1		ECL3	
0.05	0.04	0.03	0.01	0	0.9	0.97	0.22	0.18	0.16	0.36	0.45	0.09	0.04	0.02	EC2			

# Deflating the multiverse

- Better theory (assessment of fertility)
- Cleaner design (assessment of relationship status)



# Multiverse analysis: discussion

- shows the robustness/fragility of a finding
- helps identifying the sources of the fragility
- is not a formal test to assess strength of evidence
- is primarily meant to be used on *own* data
- is subjective, but less so than a single data set analysis
- is not restricted to p-values
  - e.g., Bayes factors are equally sensitive to choices in data construction
- in case of pre-registration: pre-register a multiverse analysis
- shows a small part only of a larger multiverse of statistical results
  - there is not only a data multiverse but also a model multiverse (e.g., Gelman & Loken, 2014; Patel, Burford, & Ioannidis, 2015)

# Thank you for listening

Interested to do a PhD on responsible research practices? Contact [wolf.vanpaemel@kuleuven.be](mailto:wolf.vanpaemel@kuleuven.be)

- Durante, K. M., & Arsena, A. R. (2015). Playing the field: The effect of fertility on women's desire for variety. *Journal of Consumer Research*, 41, 1372–1391.
- Durante, K. M., Griskevicius, V., Cantú, S. M., & Simpson, J. A. (2014). Money, status, and the ovulatory cycle. *Journal of Marketing Research*, 51, 27–39.
- Durante, K. M., Griskevicius, V., Hill, S. E., Perilloux, C., & Li, N. P. (2011). Ovulation, female competition, and product choice: Hormonal influences on consumer behavior. *Journal of Consumer Research*, 37, 921–934.
- Durante, K. M., Griskevicius, V., Simpson, J. A., Cantú, S. M., & Li, N. P. (2012). Ovulation leads women to perceive sexy cads as good dads. *Journal of Personality and Social Psychology*, 103, 292–305.
- Durante, K. M., Rae, A., & Griskevicius, V. (2013). The fluctuating female vote: Politics, religion, and the ovulatory cycle. *Psychological Science*, 24, 1007–1016.
- Gelman, A., & Loken, E. (2014). The statistical crisis in science. *American Scientist*, 102, 460–465.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22, 1359–1366.
- Steegen, Tuerlinckx, Gelman and Vanpaemel (2016). Increasing Transparency Through a Multiverse Analysis. *Perspectives on Psychological Science*, 11, 702-712.

# Data processing: other reasonable options

## Exclusion criteria

answers to fertility related questions

- the start date of the last period
- the start date of the period before the last period
- the typical cycle length
- the expected start date of the next period
- how sure are you about the start of the last period
- how sure are you about the start date of the period before the last period

} exclude women who are not sure  
about their start dates  
(Durante et al., 2014)

# Data processing: other reasonable options

## Exclusion criteria

answers to fertility related questions

- the start date of the last period
- the start date of the period before the last period
- the typical cycle length
- the expected start date of the next period
- how sure are you about the start of the last period
- how sure are you about the start date of the period before the last period

} exclude women who have irregular cycle lengths  
(calculated cycle length)

# Data processing: other reasonable options

## Fertility status

answers to fertility related questions

- the start date of the last period
- the start date of the period before the last period
- the typical cycle length
- the expected start date of the next period
- how sure are you about the start of the last period
- how sure are you about the start date of the period before the last period

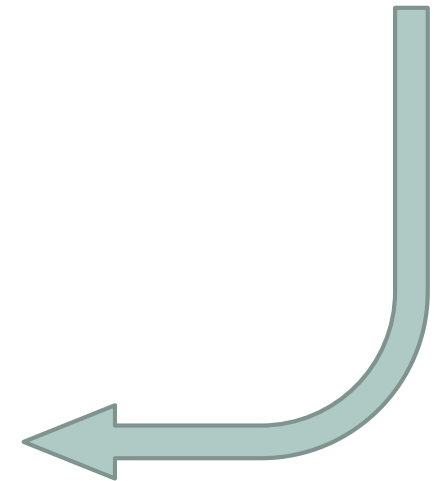
} cycle length → next menstrual onset → cycle day

→ dichotomization: high vs low fertility

high in fertility when cycle day is between 9 and 17

low in fertility when cycle day is between 1 and 8 or 18 and 28

(Durante et al., 2015)



# Data processing: other reasonable options

## Fertility status

answers to fertility related questions

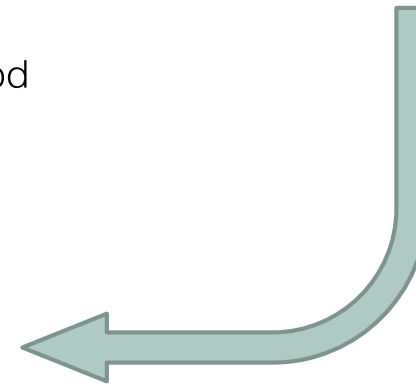
- the start date of the last period
- the start date of the period before the last period
- the typical cycle length
- the expected start date of the next period } next menstrual onset → cycle day
- how sure are you about the start of the last period
- how sure are you about the start date of the period before the last period

→ dichotomization: high vs low fertility

high in fertility when cycle day is between 9 and 17

low in fertility when cycle day is between 1 and 8 or 18 and 28

(Durante et al., 2015)





# Deflating the multiverse

assessment of relationship status

- answer to “what is your current romantic relationship status?”

(1) *not dating/romantically involved with anyone*

(2) *dating or involved with only one partner*

(3) *engaged or living with my partner*

(4) *married*

# Data processing: other reasonable options

- |    |  |            |              |
|----|--|------------|--------------|
| 1. | cycle day assessment (3 choice options)                      | → option 1 | } data set 1 |
| 2. | fertility assessment (5 choice options)                      | → option 1 |              |
| 3. | relationship status assessment (3 choice options)            | → option 1 |              |
| 4. | exclusion criteria based on cycle length (3 choice options)  | → option 1 |              |
| 5. | exclusion criteria based on variable sure (2 choice options) | → option 1 |              |

rather than arbitrarily choosing 1 option and creating a single processed data set,  
we consider all combinations of options, which give rise to a multiverse of data sets

# Data processing: other reasonable options

- |    |  |            |                            |
|----|--|------------|----------------------------|
| 1. | cycle day assessment (3 choice options)                      | → option 2 | } data set 2<br>data set 1 |
| 2. | fertility assessment (5 choice options)                      | → option 1 |                            |
| 3. | relationship status assessment (3 choice options)            | → option 1 |                            |
| 4. | exclusion criteria based on cycle length (3 choice options)  | → option 1 |                            |
| 5. | exclusion criteria based on variable sure (2 choice options) | → option 1 |                            |

rather than arbitrarily choosing 1 option and creating a single processed data set,  
we consider all combinations of options, which give rise to a multiverse of data sets

# Data processing: other reasonable options

1.	cycle day assessment (3 choice options)	→ option 3	}	
2.	fertility assessment (5 choice options)	→ option 1		
3.	relationship status assessment (3 choice options)	→ option 1		data set 3
4.	exclusion criteria based on cycle length (3 choice options)	→ option 1		data set 2
5.	exclusion criteria based on variable sure (2 choice options)	→ option 1		data set 1

rather than arbitrarily choosing 1 option and creating a single processed data set,  
we consider all combinations of options, which give rise to a multiverse of data sets

# Data processing: other reasonable options

1.	cycle day assessment (3 choice options)	→ option 1	}	
2.	fertility assessment (5 choice options)	→ option 2		
3.	relationship status assessment (3 choice options)	→ option 1		data set 4
4.	exclusion criteria based on cycle length (3 choice options)	→ option 1		data set 3
5.	exclusion criteria based on variable sure (2 choice options)	→ option 1		data set 2
				data set 1

rather than arbitrarily choosing 1 option and creating a single processed data set,  
we consider all combinations of options, which give rise to a multiverse of data sets

# Data processing: other reasonable options

- |    |  |            |   |            |
|----|--|------------|---|------------|
| 1. | cycle day assessment (3 choice options)                      | → option 1 | } |            |
| 2. | fertility assessment (5 choice options)                      | → option 3 |   |            |
| 3. | relationship status assessment (3 choice options)            | → option 1 |   | data set 5 |
| 4. | exclusion criteria based on cycle length (3 choice options)  | → option 1 |   | data set 4 |
| 5. | exclusion criteria based on variable sure (2 choice options) | → option 1 |   | data set 3 |

data set 2

data set 1

rather than arbitrarily choosing 1 option and creating a single processed data set,  
we consider all combinations of options, which give rise to a multiverse of data sets

# Data processing: other reasonable options

- |   |            |   |            |
|---|------------|---|------------|
| 1. cycle day assessment (3 choice options)                      | → option 1 | } |            |
| 2. fertility assessment (5 choice options)                      | → option 1 |   |            |
| 3. relationship status assessment (3 choice options)            | → option 2 |   | data set 6 |
| 4. exclusion criteria based on cycle length (3 choice options)  | → option 1 |   | data set 5 |
| 5. exclusion criteria based on variable sure (2 choice options) | → option 1 |   | data set 4 |

rather than arbitrarily choosing 1 option and creating a single processed data set,  
we consider all combinations of options, which give rise to a multiverse of data sets

data set 3  
data set 2  
data set 1

# Data processing: other reasonable options

- |   |            |   |            |
|---|------------|---|------------|
| 1. cycle day assessment (3 choice options)                      | → option 1 | } |            |
| 2. fertility assessment (5 choice options)                      | → option 1 |   |            |
| 3. relationship status assessment (3 choice options)            | → option 3 |   | data set 7 |
| 4. exclusion criteria based on cycle length (3 choice options)  | → option 1 |   | data set 6 |
| 5. exclusion criteria based on variable sure (2 choice options) | → option 1 |   | data set 5 |
- data set 4  
data set 3  
data set 2  
data set 1

rather than arbitrarily choosing 1 option and creating a single processed data set,  
we consider all combinations of options, which give rise to a multiverse of data sets



# Data processing: other reasonable options

- |   |            |   |
|---|------------|---|
| 1. cycle day assessment (3 choice options)                      | → option 3 | } data set 210<br>...<br>data set 7<br>data set 6<br>data set 5<br>data set 4<br>data set 3<br>data set 2<br>data set 1 |
| 2. fertility assessment (5 choice options)                      | → option 3 |   |
| 3. relationship status assessment (3 choice options)            | → option 5 |   |
| 4. exclusion criteria based on cycle length (3 choice options)  | → option 3 |   |
| 5. exclusion criteria based on variable sure (2 choice options) | → option 2 |   |
- rather than arbitrarily choosing 1 option and creating a single processed data set,  
we consider all combinations of options, which give rise to a multiverse of data sets

# Data processing: other reasonable options

- |   |            |
|---|------------|
| 1. cycle day assessment (3 choice options)                      | → option 3 |
| 2. fertility assessment (5 choice options)                      | → option 3 |
| 3. relationship status assessment (3 choice options)            | → option 5 |
| 4. exclusion criteria based on cycle length (3 choice options)  | → option 3 |
| 5. exclusion criteria based on variable sure (2 choice options) | → option 2 |

rather than arbitrarily choosing 1 option and creating a single processed data set,  
we consider all combinations of options, which give rise to a multiverse of data sets



data multiverse