Does Sociosexuality Affect Political Beliefs?

The Development, Validation, and Application of a Novel Causal Framework to solve the Non-Manipulable Nature of Sociosexuality using Synthetic Respondents

Keystrokes: 10,387 + (800 * 2) = 11,987

Introduction

Sociosexuality refers to individual differences in preferences and strategies for attracting and retaining mates. Conceptualized as a continuum ranging from short-term to long-term mating strategies (Buss & Schmitt, 1993), *short-term mating strategies* encapsulate uncommitted sexual activity (e.g., hook-ups), while *long-term mating strategies* cover committed relationships (Simpson & Gangestad, 1991). Existing studies have found correlations between a preference for short-term mating strategies and liberal political beliefs, while long-term mating strategies correlate with conservative beliefs (Petersen, 2018), contributing with a surprisingly sociosexual perspective on potential causes of cultural war and affective polarization in contemporary democracies (Goren & Chapp, 2017). However, the question of *causation* remains unanswered due to researchers being unable to manipulate sociosexuality, leaving results possibly confounded and unreliable.

Recent advances in *large language models* (LLMs), such as OpenAI's GPT-models, have improved the simulation of human language and intelligence (e.g., Yenduri *et al.*, 2024). This has been exploited by researchers across disciplines to generate *synthetic data* (Sufi, 2024), for example, to realistically simulate patients to train medical doctors (e.g., Holderried *et al.*, 2024; Pang *et al.*, 2024). Its advantages include reduced costs, though at the risk of reducing generalizability of results. In line with this trend, this project aims to produce a novel causal framework that exploits LLMs to generate data, where sociosexuality can be manipulated to infer causation, thus addressing the question:

RQ: To which degree does sociosexuality affect political beliefs? Can a reliable causal framework be developed to solve the non-manipulable nature of sociosexuality?

This research question is answered by (1) formalizing such a causal framework using potential outcomes; (2) validating the framework by comparing results between synthetic and human data; and (3) applying the framework to test whether theorized relationships between sociosexuality and

political beliefs are causal. Contributions thus include (*i*) novel causal insights for the sociosexuality literature; and (*ii*) an original, broadly applicable causal framework for political methodology.

Theory

Sociosexuality & Political Beliefs

The literature defines *sociosexuality* as 'individual differences in willingness to engage in uncommitted sexual relations' (Simpson & Gangestad, 1991: 870), with a low (high) preference reflected in long-term (short-term) mating strategies. Sociosexuality has rather surprisingly been correlated to political differences on societally relevant issues such as stances on abortion (Weeden & Kurzban, 2014), recreational drugs (Kurzban *et al.*, 2010; Quintelier *et al.*, 2013; Zhang, 2017), and rights of sexual minorities (Pinsof & Haselton, 2016; Kanazawa, 2022). Similar to *predispositions* (Hibbing *et al.*, 2019), a causal relationship is theoretically motivated by reproductive self-interest causing individuals to favor policies that help facilitate their sexual lifestyle (Weeden & Kurzban, 2014; 2017). However, the difficulty in credibly manipulating individual sociosexuality due to its biologically rigid nature (Lippa, 2009), the failure of researchers to identify *as-if-random* (Dunning, 2008) changes in sociosexuality, and questionable internal validity of alternative frameworks (e.g., *matching*; Wan *et al.*, 2024), has hindered causal inference. This leaves the literature with only correlational evidence, where relationships may be false positives or negatives (i.e., *confounding* and *masking*; McElreath, 2019), and basing policy decisions on these insights risk inefficiency if not counterproductivity. However, if found to be causal, these insights could be used in future research on depolarization (e.g., Severson, 2020).

Synthetic Causal Framework

The proposed synthetic causal framework draws on the *Neyman-Rubin Potential Outcomes Framework* (Imbens & Rubin, 2015) to offer a *synthetic solution* to the *Fundamental Problem of Causal Inference* (Holland, 1986). The term *synthetic* is used to denote the synthesis of simulated and human data (e.g., El Emam *et al.*, 2020), and while it has seen previous success in causal inference (e.g., Abadie *et al.*, 2010; Abadie *et al.*, 2015; Gilchrist *et al.*, 2023), its use within this framework can be considered novel. The framework exploits that LLMs are trained to infer human behavior (Ornstein *et al.*, 2024) and can be prompted to closely imitate human respondents (Jansen *et al.*, 2023; Coletta *et al.*, 2024). This involves specifying a character for the LLM to imitate, whose

characteristics are determined proportionally to a population of interest, and prompting it to generate a *life-history* (Rosenthal, 2018) for its character, which serves to make the character more realistic. Then, prior to treatment (D), each synthetic respondent can be expected to possess some potential outcome, $(Y_i(D = d)|\ t = 0)$, and to draw causal inference, each synthetic respondent is randomized to a treatment condition (e.g., sociosexual preference), producing, for example, the outcome: $(Y_i(D = 0)|\ t = 1)$. Following this, using the same characteristics and life-history, each respondent is restored to the time before receiving this treatment, thus erasing its memory of treatment, and is then assigned to the other treatment condition to have their counterfactual outcome measured: $(Y_i(D = 1)|\ t = 1)$. With both potential outcomes observed for each respondent (Westreich *et al.*, 2015), the *individual treatment effect*, $\tau_i = (Y_i(D = 1)|\ t = 1) - (Y_i(D = 0)|\ t = 1)$, and the *average treatment effect*, $\bar{\tau} = \frac{1}{n}\sum_{i=1}^{n}(Y_i(D = 1)|\ t = 1) - (Y_i(D = 0)|\ t = 1)$, can be estimated.

While the synthetic causal framework would provide a solution for drawing causal inference between sociosexuality and political beliefs, the generalizability of results to human populations are a concern. Similar to validating other frameworks (e.g., Hainmueller *et al.*, 2015; Sheagley & Clifford 2023), this can be done by comparing results to an established (causal) finding from another literature. This will involve replicating the comparable relationship between social deservingness and welfare attitudes (Aarøe & Petersen, 2014), previously proving robust across widely different cultures (i.e., Denmark and the United States) and multiple replications (e.g., Mummelo & Peterson, 2019; Kane *et al.*, 2023). This relationship can be expected to replicate in synthetic data if the framework produces reliable results:

H1: The effect of social deservingness on welfare attitudes is the same whether derived from a sample of synthetic respondents or a sample of human respondents.

Assuming a successful replication, causal hypotheses of the aforementioned correlational relationships can be more reliably tested and inferred to human populations:

H2a: A low level of sociosexuality increases conservatism.
H2b: A low level of sociosexuality decreases support for decriminalization of recreational drugs.
H2c: A low level of sociosexuality decreases support for abortion rights.
H2d: A low level of sociosexuality decreases support for homosexual rights.
H2e: A low level of sociosexuality decreases support for transgender rights.

Methodology

Paper 1

The first paper is purely theoretical, serving to further develop the synthetic causal framework. Similar to other contributions to political methodology (e.g., King, 1998), it will formalize assumptions and limitations, as well as recommendations on how to engineer reliable and replicable prompts to make the LLMs behave as realistically as possible. This will expectedly improve the validation of the framework in paper 2 and application in paper 3.

Paper 2

The second paper serves to validate the reliability of the framework by comparing the effect of social deservingness on welfare attitudes across synthetic and human data. The specific result to be replicated was derived from a meta-analysis of Aarøe and Petersen (2014) and the replicated US-based results (Mummelo & Peterson, 2019; Kane *et al.*, 2023) using a Bayesian mixed-effects model (Bürkner, 2017) with *weakly informative priors* (Gelman *et al.*, 2014). This produced a weighted average effect on welfare support of seeing an 'unlucky' compared to a 'lazy' welfare recipient of .32 percentage points (*pp*; ± .07; 95% CI[.16; .48]). A *power analysis* (Cohen, 1988) found using the same sample size as the original study (*n* = 1,009) 99.7% likely to detect this effect.

TABLE 1: Example of Deriving Causal Inference from a Synthetic Respondent

	$Y(D = 0) \mid t = 1$	$Y(D = 1) \mid t = 1$
Background Characteristics	Citizen of the United States of America, resident of Mississippi, white, male, between 20 and 24 years old, educated to a level less than high school, affiliated with the Republican party.	
	Lazy Recipient	**Unlucky Recipient**
Treatment Condition	"Imagine a man who is currently on social welfare. He has never had a regular job, but he is fit and healthy. He is not motivated to get a job"	"Imagine a man who is currently on social welfare. He has always had a regular job, but has now been the victim of a work-related injury. He is very motivated to get back to work again"
Outcome	"To what extent do you disagree or agree that the eligibility requirements for social welfare should be tightened for persons like him?"	
Response[*]	4 Agree	2 Disagree

NOTE: Phrasing of treatment conditions and outcome taken from Aarøe and Petersen (2014). [*]7-point Likert scale with response generated using the *GPT-4o* model on 24/8/2024.

To provide a proof-of-concept, a pilot study was conducted by generating 60 US-based synthetic respondents (for an example, see table 1). Assigning these to the lazy and unlucky welfare recipient conditions and measuring their welfare attitudes, this produced a comparable effect of .48 *pp* ($\pm$ .05; 95% CI[.38; .58]). While the small sample size likely caused the slight overestimation (Gelman & Carlin, 2014), the result is consistent with a larger sample replicating human results. While one replication is unlikely to provide definitive proof of the reliability of the framework, it does provide credibility for its application in paper 3.

Paper 3

Finally, the third paper applies the synthetic causal framework to investigate the theorized causal relationship between sociosexual strategies and political beliefs. Similar to Petersen (2018), this uses cross-national samples for robustness, generating synthetic US and Danish respondents, their characteristics generated using census data to be proportional to each national population. Given that correlational and causation estimates may vary, the 80%-powered sample size of 3,845 is determined by specifying the *smallest effect size of interest* (Lakens *et al.*, 2018) to correspond to a Pearson (1895) *correlation coefficient* of .05, a 'tiny' effect in psychology (Funder & Ozer, 2019). Using GPT-4, sociosexual strategies are manipulated using items from the *sociosexual orientation inventory* (Penke & Asendorpf, 2008) to correspond to a low/high sociosexuality in human samples. Measures of political beliefs mimic those used in the literature to ensure consistency, thus enabling the identification of causal effects to answer to the research question.

Preliminary Time Plan

Spring 2025	- Literature review of Prompt Engineering for Paper 1. - Further Development of the Synthetic Causal Framework. - Teach course *Methods II*.
Summer 2025	- Write Paper 1. - Literature review for Paper 2.
Fall 2025	- Finish and submit Paper 1. - Prepare design for Paper 2. - Teach course *Methods I*.
Winter 2025	- Submit Preregistration of the Replication on OSF of Paper 2. - Generate synthetic respondents for Paper 2 using OpenAI's API for GPT-4.
Spring 2026	- Analyze results and Write Paper 2. - Literature review for Paper 3.
Summer 2026	- Continue writing Paper 2. - Literature review for Paper 3.
Fall 2026	- Take PhD-Course Advanced Quantitative Empirical Methods (5 ECTS) - Take PhD-Course Social Research as a Craft (10 ECTS) - Finish and submit Paper 2. - Prepare design for Paper 3.
Winter 2026	- Submit Preregistration on OSF of Paper 3. - Generate Synthetic Respondents for Paper 3 using OpenAI's API for GPT-4.
Spring 2027	- Analyze and write Paper 3. - Prepare candidate seminar on Sociosexual Strategies. - Visit foreign university (e.g., UC Santa Barbara).
Summer 2027	- Continue writing Paper 3. - Finish preparing candidate seminar.
Fall 2027	- Teach candidate seminar. - Finish and submit Paper 3.
Winter 2027	- Begin writing project summary.
Spring 2027	- Finish writing project summary and complete PhD project.

6. Literature

Aarøe, L. & M. B. Petersen (2014): 'Crowding Out Culture: Scandinavians and Americans Agree on Social Welfare in the Face of Deservingness Cues', *The Journal of Politics*, 76(3): 684-697. DOI: 10.1017/S002238161400019X

Abadie, A., A. Diamond & J. Hainmueller (2010): 'Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program', *Journal of the American Statistical Association*, 105(490): 493-505. DOI: 10.1198/jasa.2009.ap08746

Abadie, A., A. Diamond & J. Hainmueller (2015): 'Comparative Politics and the Synthetic Control Method', *American Journal of Political Science*, 59(2): 495-510. DOI: 10.1111/ajps.12116

Bürkner, P.-C. (2017): 'brms: An R package for Bayesian multilevel models using Stan', *Journal of Statistical Software*, 80(1): 1-28. DOI: 10.18637/jss.v080.i01

Buss, D. M. & D. P. Schmitt (1993): 'Sexual Strategies Theory: An evolutionary perspective on human mating', *Psychological Review*, 100(2): 204–232. DOI: 10.1037/0033-295X.100.2.204

Cohen, J. (1988): *Statistical Power Analysis for the Behavioral Sciences*, 2nd edition. Lawrence Erlbaum Associates.

Colette, A., K. Dwarakanath, P. Liu, S. Vyetrenko & T. Balch (2024): 'LLM-driven Imitation of Subrational Behavior: Illusion or Reality?', *arXiv*: 2402.08755v1.

Dunning, T. (2008): 'Improving Causal Inference: Strengths and Limitations of Natural Experiments', *Political Research Quarterly*, 61(2): 282-293. DOI: 10.1177/1065912907306470

El Emam, K., L. Mosquera & R. Hoptroff (2020): *Practical Synthetic Data Generation: Balancing Privacy and the Broad Availability of Data*. O'Reilly Media.

Funder, D. & D. Ozer (2019): 'Evaluating Effect Size in Psychological Research: Sense and Nonsense', *Advances in Methods and Practices in Psychological Science*, 2(2): 155-168. DOI: 10.1177/2515245919847202

Gelman, A. & J. B. Carlin (2014): 'Beyond power calculations: Assessing type S (sign) and type M (magnitude) errors', *Perspectives on Psychological Science*, 9(6): 641-651. DOI: 10.1177/1745691614551642

Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari & D. B. Rubin (2014): *Bayesian Data Analysis*, 3rd edition. Chapman & Hall/CRC Press.

Gilchrist, D., T. Emery, N. Garoupa & R. Spruk (2024): 'Synthetic Control Method: A tool for comparative case studies in economic history', *Journal of Economic Surveys*, 37(2): 409-445. DOI: 10.1111/joes.12493

Goren, P. & C. Chapp (2017): 'Moral Power: How Public Opinion on Culture War Issues Shapes Partisan Predispositions and Religious Orientations', *American Political Science Review*, 111(1): 110–128. DOI: 10.1017/S0003055416000435

Hainmueller, J., D. Hangartner & T. Yamamoto (2015): 'Validating vignette and conjoint survey experiments against real-world behavior', *Proceedings of the National Academy of Sciences*, 112(8): 2395-400. DOI: 10.1073/pnas.1416587112

Hibbing, J. R., K. B. Smith & J. R. Alford (2019): *Predisposed. Liberals, Conservatives, and the Biology of Political Differences*. Routledge

Holderried, F., C. Stegemann-Philips, L. Herschbach, J.-A. Moldt, A. Nevins, J. Griewatz, M. Holderried, A. Herrmann-Werner, T. Festl-Wietek & M. Mahling (2024): 'A Generative Pretrained Transformer (GPT)—Powered Chatbot as a Simulated Patient to Practice History Taking: Prospective, Mixed Methods Study', *JMIR Medical Education*, 10: e53961. DOI: 10.2196/53961

Holland, P. W. (1986): 'Statistics and Causal Inference', *Journal of the American Statistical Association*, 81(396): 945-960. DOI: 10.1080/01621459.1986.10478354

Imbens, G. W. & D. B. Rubin (2015): *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press.

Jansen, B. J., S. Jung & J. Salminen (2023): 'Employing large language models in survey research', *Natural Language Processing Journal*, 4: 100020. DOI: 10.1016/j.nlp.2023.100020

Kanazawa, S. (2022): 'Unrestricted Sociosexuality Decreases Women's (but not Men's) Homophobia', *Sexuality & Culture*, 26(4): 1422–1431. DOI: 10.1007/s12119-022-09951-z

Kane, J. V., Y. R. Velez & J. Barabas (2023): 'Analyze the Attentive and Bypass Bias: Mock Vignette Checks in Survey Experiments', *Political Science Research and Methods*, 11(2): 293- 310. DOI: 10.1017/psrm.2023.3

King, G. (1998): *Unifying Political Methodology: The Likelihood Theory of Statistical Inference*. The University of Michigan Press.

Kurzban, R., A. Dukes & J. Weeden (2010): 'Sex, drugs and moral goals: reproductive strategies and views about recreational drugs', *Proceedings of the Royal Society B: Biological Sciences*, 277(1699): 3501–3508. DOI: 10.1098/rspb.2010.0608

Lakens, D., A. M: Scheel & P. M. Isager (2018): 'Equivalence Testing for Psychological Research: A Tutorial', *Advances in Methods and Practices in Psychological Science*, 1(2): 259-269. DOI: 10.1177/2515245918770963

Lippa, R. A. (2009): 'Sex Differences in Sex Drive, Sociosexuality, and Height across 53 Nations: Testing Evolutionary and Social Structural Theories', *Archives of Sexual Behavior*, 38(5): 631–651. DOI: 10.1007/s10508-007-9242-8

McElreath, R. (2019): *Statistical Rethinking: A Bayesian Course with Examples in R*, 2nd edition. Chapman & Hall/CRC Press. DOI: 10.1201/9780429029608

Mummolo, J. & E. Peterson (2019): 'Demand Effects in Survey Experiments: An Empirical Assessment', *American Political Science Review*, 113(2): 517-529. DOI: 10.1017/S0003055418000837

Ornstein, J. T., E. N. Baslingame & J. T. Truscott (2024): 'How to Train Your Stochastic Parrot: Large Language Models for Political Texts', Unpublished Manuscript on *GitHub*: (https://joeornstein.github.io/publications/ornstein-blasingame-truscott.pdf).

Pang, C., X. Jiang, N. P. Pavinkurve, K. S. Kalluri, E. L. Minto, J. Patterson, L. Zhang, G. Hripcsak, G. Gürsoy, N. Elhadad & K. Natarajan (2024): 'CEHR-GPT: Generating Electronic Health Records with Chronological Patient Timelines', *arXiv*: 2402.04400v2. DOI: 10.48550/arxiv.2402.04400

Pearson, K. (1895): 'Notes on regression and inheritance in the case of two parents', *Proceedings of the Royal Society of London*, 58: 240-242.

Penke, L. & J. B. Asendorpf (2008): 'Beyond global sociosexual orientations: A more differentiated look at sociosexuality and its effects on courtship and romantic relationships', *Journal of Personality and Social Psychology*, 95(5): 1113–1135. DOI: 10.1037/0022-3514.95.5.1113

Petersen, M. B. (2018): 'Reproductive interests and dimensions of political ideology', *Evolution and Human Behavior*, 39(2): 203–211. DOI: 10.1016/j.evolhumbehav.2017.12.002

Pinsof, D. & M. Haselton (2016): 'The Political Divide Over Same-Sex Marriage: Mating Strategies in Conflict?', *Psychological Science*, 27(4): 435–442. DOI: 10.1177/0956797615621719

Quintelier, K. J. P., K. Ishii, J. Weeden, R. Kurzban & J. Braeckman (2013): 'Individual differences in reproductive strategy are related to views about recreational drug use in Belgium, The Netherlands, and Japan', *Human Nature*, 24(2): 196–217. DOI: 10.1007/s12110-013-9165-0

Rosenthal, G. (2018): *Interpretive Social Research: An Introduction*. Universitätsverlag Göttingen. DOI: 10.17875/gup2018-1103.

Severson, A. (2020): 'Homo Politicus Was Born this Way: How Understanding the Biology of Political Belief Promotes Depolarization', *Open Science Framework*. DOI: 10.31235/osf.io/7qgcr

Sheagley, G. & S. Clifford (2023): 'No Evidence that Measuring Moderators Alters Treatment Effects', *American Journal of Political Science*, 0(0): 1-15. DOI: 10.1111/ajps.12814

Simpson, J. A. & S. W. Gangestad (1991): 'Individual differences in sociosexuality: Evidence for convergent and discriminant validity', *Journal of Personality and Social Psychology*, 60(6): 870–883. DOI: 10.1037/0022-3514.60.6.870

Wan, F., S. Sutcliffe, J. Zhang & D. Small (2024): 'Does matching introduce confounding or selection bias into the matched case-control design?', *Observational Studies*, 10(1): 1-9. DOI: 10.1353/obs.2024.a929114

Weeden, J. & R. Kurzban (2014): *The Hidden Agenda of the Political Mind: How Self-Interest Shapes Our Opinions and Why We Won't Admit It*. Princeton University Press. DOI: 10.1515/9781400851966

Weeden, J. & R. Kurzban (2017): 'Self-Interest Is Often a Major Determinant of Issue Attitudes', *Political Psychology*, 38(S1): 67–90. DOI: 10.1111/pops.12392

Westreich, D., J. K. Edwards, S. R. Cole, R. W. Platt, S. L. Mumford & E. F. Schisterman (2015): 'Imputation approaches for potential outcomes in causal inference', *International Journal of Epidemiology*: 1731-1737. DOI: 10.1093/ije/dyv135

Yenduri, G., M. Ramalingam, G. C. Selvi, Y. Supriya, G. Srivastava, P. K. R. Maddikunta, G. D. Raj, R. J. Jhaveri, B. Prabadevi, W. Wang, A. V. Vasilakos & T. R. Gadekallu (2024): 'GPT (Generative Pre-Trained Transformer)—A Comprehensive Review on Enabling Technologies,

Potential Applications, Emerging Challenges, and Future Direction', *IEEE Access*, 12: 54608-54649. DOI: 10.1109/ACCESS.2024.3389497

Zhang, J. (2017): 'Is Support of Censoring Controversial Media Content for the Good of Others? Sexual Strategies and Support of Censoring Pro-Alcohol Advertising', *Evolutionary Psychology*, 15(4): 1474704917742808. DOI: 10.1177/1474704917742808