Response to "Reply to Fitzgerald: COVID-19 Vaccine Mandates and Voluntary Vaccination Behavior"

Jack Fitzgerald

October 2, 2024

1 Introduction

Rains & Richards (2024c) find that US states that mandated COVID-19 vaccination now have lower uptake of COVID-19 booster shots and both adult and child flu vaccines than states that banned local COVID-19 vaccination mandates. Rains & Richards (2024c) interpret these differences as causal effects of vaccine mandates. Fitzgerald (2024) replicates Rains & Richards (2024c) and shows that these findings are driven by the inclusion of a single bad control variable. Rains & Richards' (2024c) models on COVID-19 booster/flu vaccine uptake all control for contemporaneous COVID-19 vaccination rates, which are likely affected by both COVID-19 vaccination mandates and by unobserved factors that also impact COVID-19 booster/flu vaccine uptake, such as vaccine hesitancy (see Cohn et al. 2022; Karaivanov et al. 2022; Mills & Ruttenauer 2022). Controlling for contemporaneous COVID-19 vaccination rates thus likely induces collider bias in Rains & Richards' (2024c) estimates on COVID-19 vaccination mandates (Cinelli, Forney, & Pearl 2024). Fitzgerald (2024) shows that removing COVID-19 vaccination rates from Rains & Richards' (2024c) models sign-flips the paper's headline results on COVID-19 booster/flu vaccine uptake. I.e., though Rains & Richards (2024c) find that states that mandated COVID-19 vaccination have significantly *less* COVID-19 booster/flu vaccine uptake, Fitzgerald (2024c) shows that after removing the bad control from the model, states that mandated COVID-19 vaccination see significantly *more* COVID-19 booster/flu vaccine uptake.

Rains & Richards (2024a) respond to Fitzgerald (2024) by reporting new estimates from

models where they continue to control for contemporaneous COVID-19 vaccination rates, but add new variables controlling for baseline measures of mandate/vaccine acceptance. Per Rains & Richards (2024a), there are two new sets of control variables. The first variable is state-level attitudes towards requiring vaccinations for all individuals in January 2021, as measured by the COVID States Project (Lin et al. 2023). The second set of variables is state-level flu vaccination rates for healthy adults aged 18-64, as well as for children, from 2018-2019, per data from the US Centers for Disease Control and Prevention (CDC 2024). These new models appear to broadly support Rains & Richards' (2024c) original conclusions. The new estimates on the difference between mandate states and states that banned mandates are all negative, as in the original paper, with nearly all of these new estimates being statistically significantly different from zero.

This note demonstrates problems with the empirical analyses in Rains & Richards (2024a). I show that the new control strategy applied in Rains & Richards (2024a) still potentially exposes their estimates of interest to collider bias because these models still control for contemporaneous COVID-19 vaccination rates. Further, with the new baseline vaccination and attitudes data that Rains & Richards (2024a) collect, it is now unnecessary to incur risks of collider bias, as one can control for baseline vaccine attitudes without controlling for contemporaneous COVID-19 vaccination rates. I thus run a robustness check that removes COVID-19 vaccination rates from Rains & Richards' (2024a) new models. As in my original replication, this robustness check sign-flips every estimate of interest in Rains & Richards (2024a). I.e., even when controlling for baseline vaccine attitudes/uptake, states that mandated COVID-19 vaccination see higher, not lower, uptake of COVID-19 boosters and flu vaccines. The results in Rains & Richards (2024a) are thus likely driven by collider bias arising from unnecessarily controlling for contemporaneous COVID-19 vaccination rates. The replication repository for this analysis can be found at https://osf.io/9cn38/.

2 Risk of Collider Bias

Rains & Richards (2024a) argue that their new controls can eliminate the collider bias that I highlight in my original replication. Fitzgerald (2024) points to vaccine hesitancy as an un-

observed factor that impacts both COVID-19 vaccination rates and COVID-19 booster/flu vaccine uptake. The existence of such a factor implies that controlling for contemporaneous COVID-19 vaccination rates will induce collider bias because COVID-19 vaccination mandates impact COVID-19 vaccination rates (Cohn et al. 2022; Karaivanov et al. 2022; Mills & Ruttenauer 2022; Cinelli, Forney & Pearl 2024). Rains & Richards (2024a) argue that additionally controlling for baseline vaccination rates eliminates potential collider bias arising from vaccine hesitancy's joint impact on COVID-19 vaccination rates and COVID-19 booster/flu vaccine uptake. Appendix Figure A1 replicates Figure 1 in Rains & Richards (2024a) and shows that this claim is true: adding measures of baseline vaccine hesitancy does close collider paths arising from vaccine hesitancy.

However, it is an odd choice to continue controlling for contemporaneous COVID-19 vaccination rates when Rains & Richards (2024a) now have baseline measures of mandate/vaccine acceptance. Rains & Richards (2024a, 2024c) justify controlling for COVID-19 vaccination rates because they capture revealed attitudes about vaccination and mandates. From Rains & Richards (2024a):

> "We included COVID-19 vaccination rates for theoretical reasons. As previously discussed in our manuscript and supporting information, research on psychological reactance theory indicates that the efficacy of government mandates depends on people's preexisting feelings about the restricted behavior."

However, in Rains & Richards (2024a), this goal of controlling for "preexisting feelings about the restricted behavior" is now already achieved by controlling for baseline flu vaccine uptake and mandate attitudes. It is no longer necessary to control for COVID-19 vaccination rates to capture these attitudes. The models in Rains & Richards (2024a) could in principle achieve the same level of control over these baseline attitudes by omitting controls for COVID-19 vaccination rates, which would minimize the risk of collider bias.

Continuing to control for contemporaneous COVID-19 vaccination rates still likely induces collider bias in Rains & Richards' (2024a) new models. In order to rule out all collider bias induced by controlling for contemporaneous COVID-19 vaccination rates, collider paths from *all* unobserved factors that simultaneously affect COVID-19 vaccination rates and

COVID-19 booster/flu vaccine uptake must be closed. Appendix Figure A2 shows that even if controlling for baseline vaccine/mandate acceptance closes collider paths that run through vaccine hesitancy, if there is any other unobserved collider whose path is not closed by an observed control variable, then there will still be a collider path that biases estimates of the relationship between vaccine mandates and COVID-19 booster/flu vaccine uptake. For example, healthcare access simultaneously affects COVID-19 vaccination rates and COVID-19 booster/flu vaccine uptake, and its collider path is not blocked by any observed control variable. Many similar unobserved factors likely exist.

3 Replication

To assess the robustness of the new findings to potential collider bias, I replicate the results in Rains & Richards (2024a) and conduct a robustness check where I remove controls for contemporaneous COVID-19 vaccination rates. Unfortunately, Rains & Richards (2024a) do not provide a replication repository for the results in their reply. I thus combine the data from the replication repository for the original paper (Rains & Richards 2024b) with the new data sources cited in the reply (Lin et al. 2023; CDC 2024).

Examining the newer data sources reveals ambiguities in Rains & Richards' (2024a) variable descriptions that make reproducibility challenging without a replication repository. In particular, Rains & Richards (2024a) report obtaining "flu vaccination rates for healthy adults ages 18 to 64 during the 2018-2019 flu season from the US CDC, and child flu vaccination rates during 2018 to 2019 from the CDC." Several ambiguities become clear when comparing these variable descriptions to the CDC data.

1. **Timing**: The CDC (2024) data is stored monthly, not seasonally or annually, so the CDC (2024) data does not report just one '2018-2019' flu vaccination rate for any state. This implies either that a given state's monthly baseline flu vaccination rates are aggregated into one seasonal rate or that this state's individual monthly baseline vaccination rates take on different values that are somehow matched to that state's endline vaccination rates. The aggregation/matching process that Rains & Richards (2024a) employ is unclear.

2. **Missing values**: Several states are missing at least some vaccination rates in some or all months during the 2018-2019 flu season. It is unclear how these missing values are handled when aggregating or matching baseline vaccination rates to endline COVID-19 booster/flu vaccination rates.

3. **Vaccination types**: The CDC reports vaccination rates against the H1N1 swine flu, against the seasonal flu, or against either. It is unclear which of these latter two vaccine rates Rains & Richards (2024a) are collecting.

For my replication, I aggregate state i's vaccination rates by taking a simple mean of all available monthly vaccination rates against the seasonal flu for state i during the 2018-2019 flu season, ignoring missing values. I obtain these values both for non-high-risk adults aged 18-64 and for children aged six months to 17 years.

Appendix Table A1 shows the results of my best attempt to reproduce the findings in Rains & Richards (2024a). I cannot exactly reproduce the results in Rains & Richards (2024a). However, the conclusions arising from all coefficients on the 'mandate state' dummy in my reproduction attempt are nearly identical to those arising from the published estimates in Rains & Richards (2024a). This suggests that I have largely recovered the data collection and analytical strategies employed in Rains & Richards (2024a).

Table 1 compares the results of my reproduction to identical models that remove controls for contemporaneous COVID-19 vaccination rates; this robustness check sign-flips every estimate of interest in Rains & Richards (2024a). As in their original paper, when Rains & Richards (2024a) control for contemporaneous COVID-19 vaccination rates, states that mandated COVID-19 vaccination rates appear to have significantly *lower* uptake of COVID-19 boosters and flu vaccines than states which banned local COVID-19 vaccination mandates. However, after eliminating the risk of collider bias by removing contemporaneous COVID-19 vaccination rates from Rains & Richards' (2024a) models, mandate states have significantly *higher* uptake of COVID-19 boosters and flu vaccines, even after controlling for baseline mandate attitudes and baseline vaccination rates.

This check provides even stronger evidence that Rains & Richards' (2024a; 2024c) results are driven by collider bias than my original replication (Fitzgerald 2024). Removing

	COVID-19 booster uptake	Adult flu vaccine uptake	Child flu vaccine uptake	COVID-19 booster uptake	Adult flu vaccine uptake	Child flu vaccine uptake
Panel A						
Mandate state	-0.075	-0.121	-0.183	0.046	0.06	0.08
	(0.028)	(0.021)	(0.029)	(0.021)	(0.016)	(0.022)
Baseline mandate	-0.001	-0.001	-0.001	0	0.001	0.001
attitudes	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
COVID-19	3.686	1.522	2.194			
vaccination rate	(0.056)	(0.071)	(0.029)			
N	1025	205	1025	1025	205	1025
Panel B						
Mandate state	-0.07	-0.12	-0.182	0.045	0.058	0.063
	(0.029)	(0.022)	(0.03)	(0.022)	(0.017)	(0.021)
Baseline adult flu	0	0	0.003	0.001	0.001	0.008
vaccination rate	(0.004)	(0.003)	(0.004)	(0.003)	(0.002)	(0.003)
COVID-19	3.675	1.545	2.182			
vaccination rate	(0.057)	(0.071)	(0.03)			
N	1000	200	1000	1000	200	1000
Panel C						
Mandate state	-0.063	-0.117	-0.175	0.046	0.059	0.055
	(0.03)	(0.021)	(0.03)	(0.023)	(0.016)	(0.02)
Baseline child flu	-0.002	-0.002	-0.001	0	-0.001	0.007
vaccination rate	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)
COVID-19	3.686	1.52	2.194			
vaccination rate	(0.056)	(0.07)	(0.029)			
N	1025	205	1025	1025	205	1025

Note: Standard errors are reported in parentheses.

Table 1: Results With and Without Controls for COVID-19 Vaccination Rates

contemporaneous COVID-19 vaccination rates from Rains & Richards' (2024c) original models creates a bias tradeoff. On one hand, removing contemporaneous COVID-19 vaccination rates as a control minimizes the risk of collider bias. On the other hand, this check also risks introducing some bias by potentially re-opening backdoor paths from unobserved vaccine hesitancy and/or mandate attitudes. However, the robustness checks in this paper now control for measures of this baseline acceptance, which in principle should close those backdoor paths. This provides strong evidence that the differences between my estimates and those in Rains & Richards (2024a; 2024c) arise because their models are exposed to a unique collider bias. This implies that the results in both Rains & Richards (2024a) and Rains & Richards (2024c) are primarily driven by collider bias arising from adjustments for a bad control variable.

References

Centers for Disease Control and Prevention (2024). *Influenza vaccination coverage for all ages (6+ months)*. Dataset. URL: https://data.cdc.gov/Flu-Vaccinations/Influenza-Vaccination-Coverage-for-All-Ages-6-Mont/vh55-3he6/about_data (visited on 09/30/2024).

Cinelli, Carlos, Andrew Forney, and Judea Pearl (May 2022). "A crash course in good and bad controls". *Sociological Methods & Research* 53.3, pp. 1071–1104. DOI: 10.1177/00491241221099552.

Cohn, Ezra et al. (Sept. 2022). "The effect of a proof-of-vaccination requirement, incentive payments, and employer-based mandates on COVID-19 vaccination rates in New York City: A synthetic-control analysis". *The Lancet Public Health* 7.9, E754–E762. DOI: 10.1016/s2468-2667(22)00196-7.

Fitzgerald, Jack (2024). "US states that mandated COVID-19 vaccination see higher, not lower, take-up of COVID-19 boosters and flu vaccines". *Proceedings of the National Academy of Sciences* 121.41, e2403758121. DOI: 10.1073/pnas.2403758121.

Karaivanov, Alexander et al. (June 2022). "COVID-19 vaccination mandates and vaccine uptake". *Nature Human Behaviour* 6.12, pp. 1615–1624. DOI: 10.1038/s41562-022-01363-1.

Lin, Jennifer et al. (2023). *COVID states project*. Dataset. OSF, osf.io/j3tqd. URL: https://lazerlab.shinyapps.io/Behaviors_During_COVID/ (visited on 09/30/2024).

Mills, Melinda C and Tobias Ruttenauer (Jan. 2022). "The effect of mandatory COVID-19 certificates on vaccine uptake: Synthetic-control modelling of six countries". *The Lancet Public Health* 7.1, e15–e22. DOI: 10.1016/s2468-2667(21)00273-5.

Rains, Stephen A. and Adam S. Richards (2024a). "Reply to Fitzgerald: COVID-19 vaccine mandates and voluntary vaccination behavior". *Proceedings of the National Academy of Sciences* 121.41, e2409246121. DOI: 10.1073/pnas.2409246121.

— (2024b). *State COVID-19 Vaccination Mandates and Vaccine Adoption*. Deposited. OSF, osf.io/j3tqd.

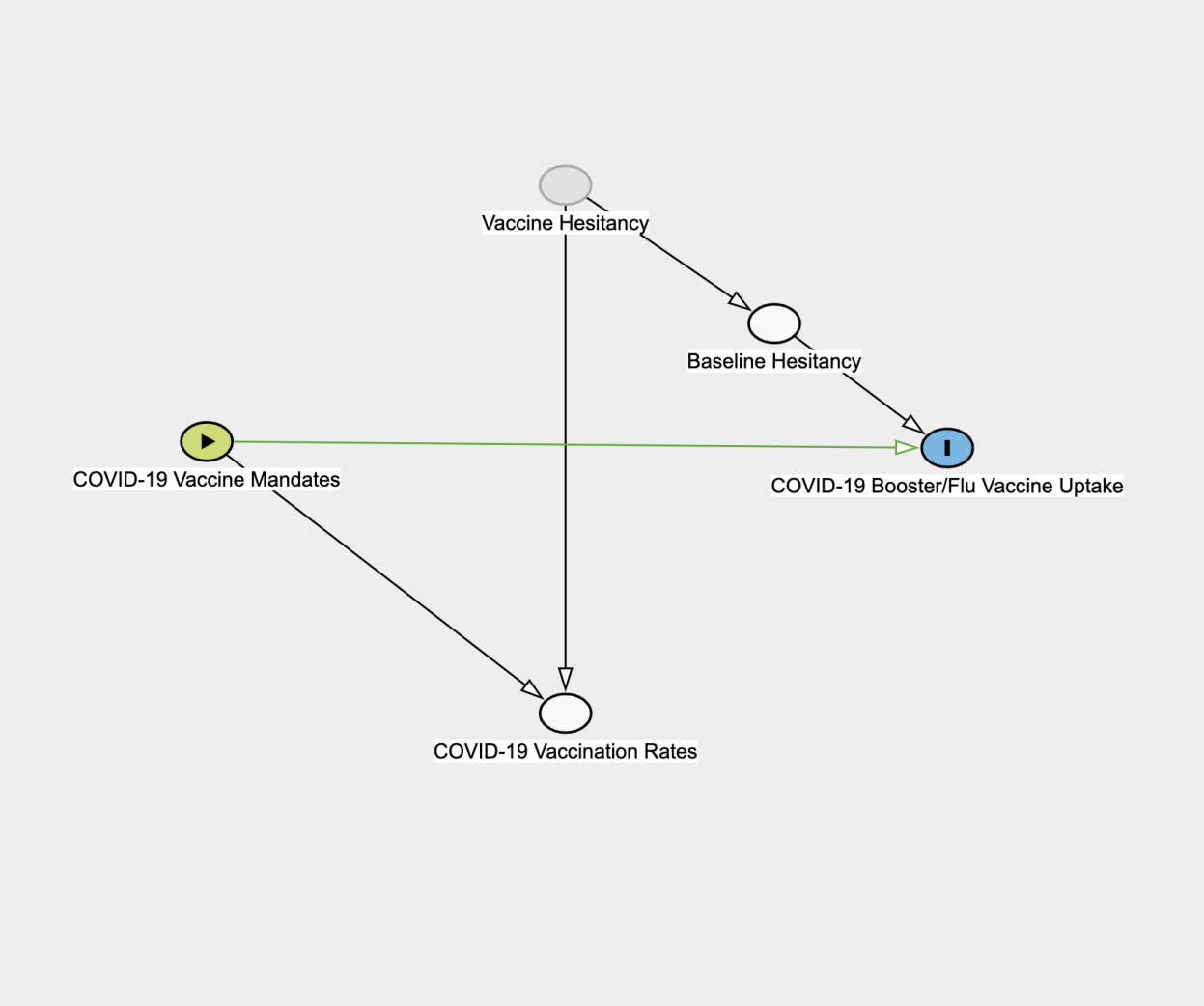
Rains, Stephen A. and Adam S. Richards (2024c). "US state vaccine mandates did not influence COVID-19 vaccination rates but reduced uptake of COVID-19 boosters and flu vaccines compared to bans on vaccine restrictions". *Proceedings of the National Academy of Sciences* 121.8, e2313610121. DOI: 10.1073/pnas.2313610121.

Appendix

	COVID-19 booster uptake	COVID-19 booster uptake	Adult flu vaccine uptake	Adult flu vaccine uptake	Child flu vaccine uptake	Child flu vaccine uptake
Panel A						
Mandate state	-0.1	-0.07	-0.07	-0.12	-0.11	-0.18
	($p = 0.001$)	($p = 0.011$)	($p = 0.002$)	($p < 0.001$)	($p = 0.001$)	($p < 0.001$)
COVID-19	3.87	3.69	1.54	1.52	2.2	2.19
vaccination rate	($p < 0.001$)	($p < 0.001$)	($p < 0.001$)	($p < 0.001$)	($p < 0.001$)	($p < 0.001$)
Baseline mandate	0.2	0	-0.62	0	-0.74	0
attitudes	($p = 0.001$)	($p = 0.67$)	($p < 0.001$)	($p = 0.686$)	($p < 0.001$)	($p = 0.537$)
Panel B						
Mandate state	-0.07	-0.07	-0.12	-0.12	-0.19	-0.18
	($p = 0.012$)	($p = 0.02$)	($p < 0.001$)	($p < 0.001$)	($p < 0.001$)	($p < 0.001$)
COVID-19	3.67	3.67	1.48	1.55	2.18	2.18
vaccination rate	($p < 0.001$)	($p < 0.001$)	($p < 0.001$)	($p < 0.001$)	($p < 0.001$)	($p < 0.001$)
Baseline adult flu	0	0	0	0	0	0
vaccination rate	($p = 0.697$)	($p = 0.991$)	($p = 0.023$)	($p = 0.875$)	($p = 0.154$)	($p = 0.5$)
Panel C						
Mandate state	-0.05	-0.06	-0.12	-0.12	-0.19	-0.17
	($p = 0.122$)	($p = 0.038$)	($p < 0.001$)	($p < 0.001$)	($p < 0.001$)	($p < 0.001$)
COVID-19	3.68	3.69	1.51	1.52	2.18	2.19
vaccination rate	($p < 0.001$)	($p < 0.001$)	($p < 0.001$)	($p < 0.001$)	($p < 0.001$)	($p < 0.001$)
Baseline child flu	0	0	0	0	0	0
vaccination rate	($p = 0.163$)	($p = 0.437$)	($p = 0.692$)	($p = 0.387$)	($p = 0.479$)	($p = 0.627$)
Published estimates	X		X		X	
Reproduction		X		X		X

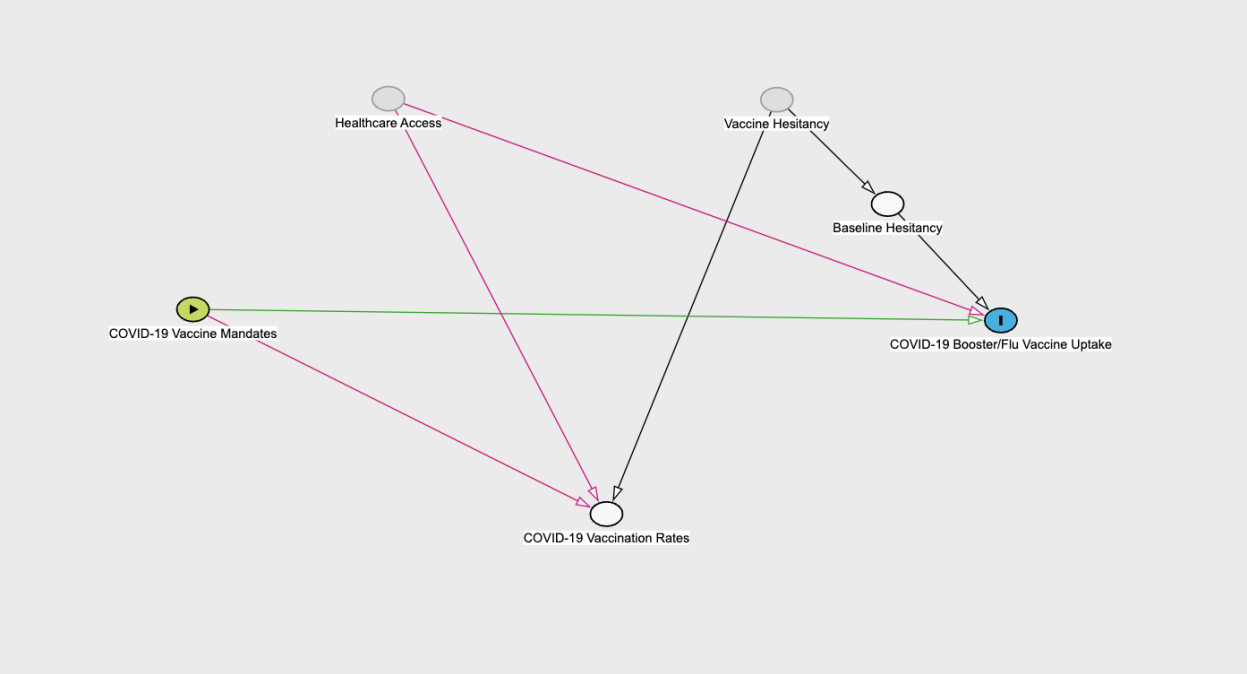
Note: Raw p-values are reported in parentheses.

Table A1: Published Estimates vs. Reproduction of Table 1 in Rains & Richards' Reply



Note: Directed acyclic graph showing relationships between variables in the models from Rains & Richards (2024a). COVID-19 vaccine mandates are the exposure of interest, and COVID-19 booster and flu vaccine uptake are the outcomes of interest. White circles denote observed variables controlled for in Rains & Richards (2024a), and grey circles denote unobserved factors not controlled for in Rains & Richards (2024a). Graph made using https://www.dagitty.net/dags.html.

Figure A1: Directed Acyclic Graph With Only Vaccine Hesitancy as a Potential Confounder



Note: Directed acyclic graph showing relationships between variables in the models from Rains & Richards (2024a). COVID-19 vaccine mandates are the exposure of interest, and COVID-19 booster and flu vaccine uptake are the outcomes of interest. White circles denote observed variables controlled for in Rains & Richards (2024a), and grey circles denote unobserved factors not controlled for in Rains & Richards (2024a). Red arrows show biasing paths. Graph made using https://www.dagitty.net/dags.html.

Figure A2: Directed Acyclic Graph With Other Potential Confounders